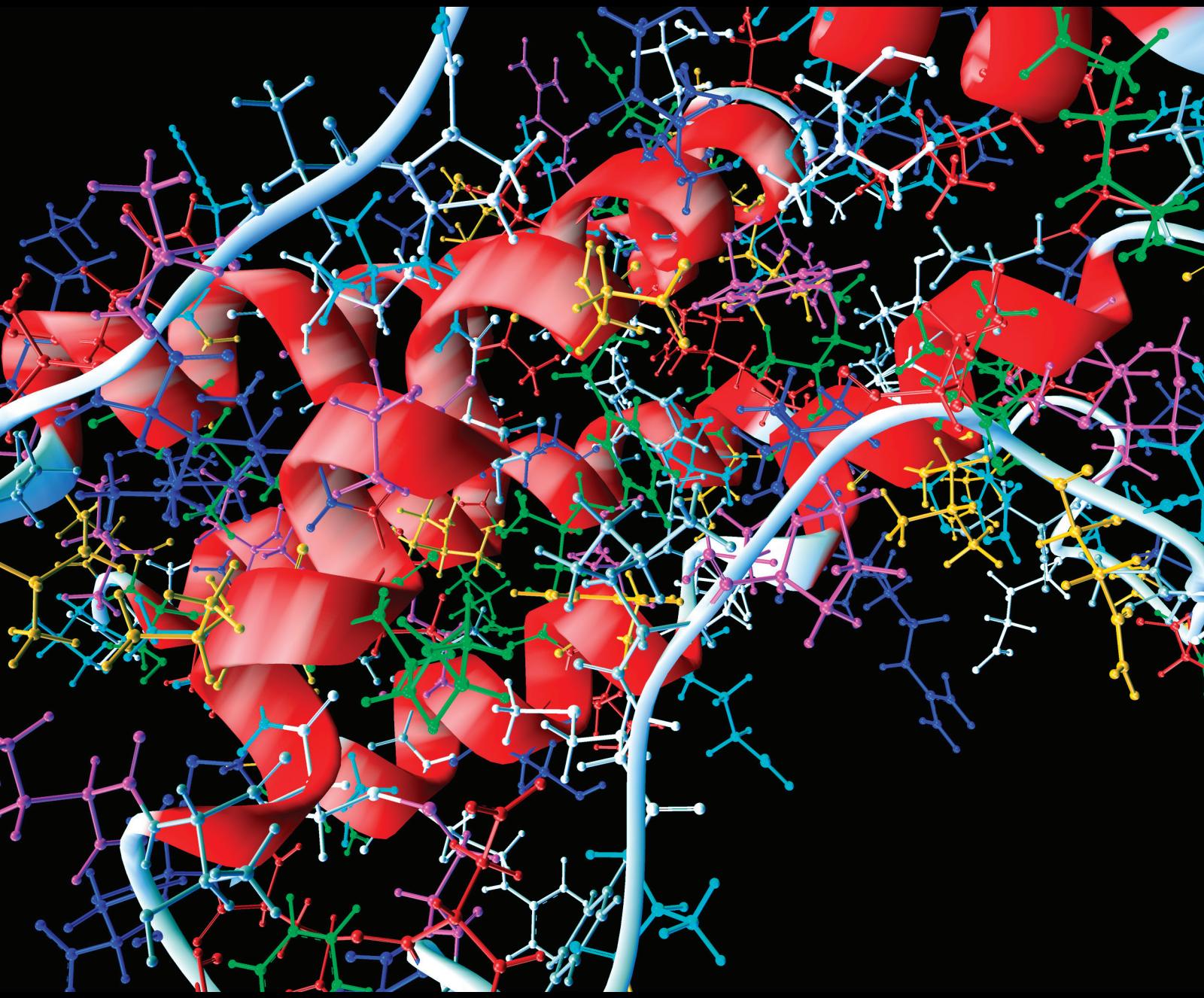
Computational and Mathematical Methods in Medicine

Mathematical Methods and Applications in Medical Imaging 2014

Guest Editors: Liang Li, Tianye Niu, and Yi Gao



Mathematical Methods and Applications in Medical Imaging 2014

Computational and Mathematical Methods in Medicine

Mathematical Methods and Applications in Medical Imaging 2014

Guest Editors: Liang Li, Tianye Niu, and Yi Gao

Editorial Board

Contents

Mathematical Methods and Applications in Medical Imaging 2014, Liang Li, Tianye Niu, and Yi Gao
Volume 2015, Article ID 685036, 2 pages

Adhesion Pulmonary Nodules Detection Based on Dot-Filter and Extracting Centerline Algorithm, Liwei Liu, Xin Wang, Yang Li, Liping Wang, and Jianghui Dong
Volume 2015, Article ID 597313, 11 pages

A Reconstruction Method of Blood Flow Velocity in Left Ventricle Using Color Flow Ultrasound, Jaeseong Jang, Chi Young Ahn, Kiwan Jeon, Jung Heo, DongHak Lee, Chulmin Joo, Jung-il Choi, and Jin Keun Seo
Volume 2015, Article ID 108274, 15 pages

Automatically Identifying Fusion Events between GLUT4 Storage Vesicles and the Plasma Membrane in TIRF Microscopy Image Sequences, Jian Wu, Yingke Xu, Zhouyan Feng, and Xiaoxiang Zheng
Volume 2015, Article ID 610482, 7 pages

An Adaptive Thresholding Method for BTV Estimation Incorporating PET Reconstruction Parameters: A Multicenter Study of the Robustness and the Reliability, M. Brambilla, R. Matheoud, C. Basile, C. Bracco, I. Castiglioni, C. Cavedon, M. Cremonesi, S. Morzenti, F. Fioroni, M. Giri, F. Botta, F. Gallivanone, E. Grassi, M. Pacilio, E. De Ponti, M. Stasi, S. Pasetto, S. Valzano, and D. Zanni
Volume 2015, Article ID 571473, 12 pages

Explicit Filtering Based Low-Dose Differential Phase Reconstruction Algorithm with the Grating Interferometry, Xiaolei Jiang, Li Zhang, Ran Zhang, Hongxia Yin, and Zhenchang Wang
Volume 2015, Article ID 623236, 7 pages

Few-View Prereconstruction Guided Tube Current Modulation Strategy Based on the Signal-to-Noise Ratio of the Sinogram, Ming Chang, Yongshun Xiao, and Zhiqiang Chen
Volume 2015, Article ID 906452, 8 pages

CT Image Reconstruction from Sparse Projections Using Adaptive TpV Regularization, Hongliang Qi, Zijia Chen, and Linghong Zhou
Volume 2015, Article ID 354869, 8 pages

Adaptive Autoregressive Model for Reduction of Noise in SPECT, Reijo Takalo, Heli Hytti, Heimo Ihalainen, and Antti Sohlberg
Volume 2015, Article ID 494691, 9 pages

The EM Method in a Probabilistic Wavelet-Based MRI Denoising, Marcos Martin-Fernandez and Sergio Villullas
Volume 2015, Article ID 182659, 21 pages

NUFFT-Based Iterative Image Reconstruction via Alternating Direction Total Variation Minimization for Sparse-View CT, Bin Yan, Zhao Jin, Hanming Zhang, Lei Li, and Ailong Cai
Volume 2015, Article ID 691021, 9 pages

Deep Adaptive Log-Demons: Diffeomorphic Image Registration with Very Large Deformations, Liya Zhao and Kebin Jia
Volume 2015, Article ID 836202, 16 pages

Optimized Parallelization for Nonlocal Means Based Low Dose CT Image Processing, Libo Zhang, Benqiang Yang, Zhikun Zhuang, Yining Hu, Yang Chen, Limin Luo, and Huazhong Shu
Volume 2015, Article ID 790313, 11 pages

Estimation of Sensitive Proportion by Randomized Response Data in Successive Sampling, Bo Yu, Zongda Jin, Jiayong Tian, and Ge Gao
Volume 2015, Article ID 172918, 6 pages

Finger Vein Segmentation from Infrared Images Based on a Modified Separable Mumford Shah Model and Local Entropy Thresholding, Marios Vlachos and Evangelos Dermatas
Volume 2015, Article ID 868493, 20 pages

Methodology for Registration of Shrinkage Tumors in Head-and-Neck CT Studies, Jianhua Wang, Jianrong Dai, Yongjie Jing, Yanan Huo, and Tianye Niu
Volume 2015, Article ID 265497, 9 pages

Recent Development of Dual-Dictionary Learning Approach in Medical Image Analysis and Reconstruction, Bigong Wang and Liang Li
Volume 2015, Article ID 152693, 9 pages

Feature Selection Based on Machine Learning in MRIs for Hippocampal Segmentation, Sabina Tangaro, Nicola Amoroso, Massimo Brescia, Stefano Cavuoti, Andrea Chincarini, Rosangela Errico, Paolo Inglese, Giuseppe Longo, Rosalia Maglietta, Andrea Tateo, Giuseppe Riccio, and Roberto Bellotti
Volume 2015, Article ID 814104, 10 pages

Hybrid Pixel-Based Method for Cardiac Ultrasound Fusion Based on Integration of PCA and DWT, Samaneh Mazaheri, Puteri Suhaiza Sulaiman, Rahmita Wirza, Mohd Zamrin Dimon, Fatimah Khalid, and Rohollah Moosavi Tayebi
Volume 2015, Article ID 486532, 16 pages

Segmentation of Regions of Interest Using Active Contours with SPF Function, Farhan Akram, Jeong Heon Kim, Chan-Gun Lee, and Kwang Nam Choi
Volume 2015, Article ID 710326, 14 pages

Identification of Mitral Annulus Hinge Point Based on Local Context Feature and Additive SVM Classifier, Jianming Zhang, Yangchun Liu, and Wei Xu
Volume 2015, Article ID 419826, 8 pages

Adaptive Ridge Point Refinement for Seeds Detection in X-Ray Coronary Angiogram, Ruoxiu Xiao, Jian Yang, Danni Ai, Jingfan Fan, Yue Liu, Guangzhi Wang, and Yongtian Wang
Volume 2015, Article ID 502573, 10 pages

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 685036, 2 pages
http://dx.doi.org/10.1155/2015/685036

Editorial

Mathematical Methods and Applications in Medical Imaging 2014

Liang Li,[1,2] Tianye Niu,[3] and Yi Gao[4]

[1]*Department of Engineering Physics, Tsinghua University, Beijing 100084, China*
[2]*Key Laboratory of Particle & Radiation Imaging (Tsinghua University), Ministry of Education, Beijing 100084, China*
[3]*Sir Run Run Shaw Hospital and Institute of Translational Medicine, Zhejiang University, Hangzhou, Zhejiang 310000, China*
[4]*Departments of Biomedical Informatics, Computer Science, and Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794-8322, USA*

Correspondence should be addressed to Liang Li; lliang@tsinghua.edu.cn

Received 21 April 2015; Accepted 21 April 2015

Medical imaging applies different techniques to acquire human images for clinical purposes, including diagnosis, monitoring, and treatment guidance. As a typical multidisciplinary field, medical imaging requires the improvements in both science and engineering to implement and maintain its noninvasive feature. Computational and mathematical methods are involved with imaging theories, models, reconstruction algorithms, image processing, quantitative imaging techniques, acceleration techniques, and multimodal imaging techniques. The main purpose of this issue is to bridge the gap between mathematical methods and their applications in medical imaging. This special issue covers most of the common medical imaging modalities, such as CT, PET, SPECT, MRI, ultrasound, phase contrast, and various image processing methods, such as segmentation, registration, fusion, identification, and denoising.

This special issue has 21 papers which were reviewed by at least two reviewers. Seven papers are involved with medical image reconstruction. Few-view CT scanning is a promising low-dose CT imaging mode in applications. B. Yan et al. propose an iterative few-view CT image reconstruction based on nonuniform fast Fourier transform and alternating direction total variation minimization. Different from Yan's algorithm, H. Qi et al. propose an adaptive TpV regularization algorithm which uses variable p value instead of the traditional constant p-norm of the image gradient magnitude. M. Brambilla et al. assess the robustness and reliability of an adaptive thresholding algorithm for the Biological Target Volume estimation incorporating PET reconstruction parameters. M. Chang et al. study the automatic exposure control strategies and propose a few-view prereconstruction guided tube current modulation method by keeping the SNR of the sinogram proximately invariable. J. Jang et al. propose a reconstruction method to quantify the distribution of blood flow velocity fields, a potentially useful index of cardiac dysfunction, inside the left ventricle from color flow ultrasound images. X-ray grating interferometry offers more information compared with traditional X-ray attenuation imaging especially for the study of weakly absorbing samples. X. Jiang et al. propose a low-dose differential phase reconstruction algorithm which adopts a differential algebraic reconstruction technique with the explicit filtering based sparse regularization rather than the common total variation. B. Wang and L. Li write a review paper on recent developments of dual-dictionary learning method in medical image analysis and reconstruction, which also discusses its role in the future studies and potential applications in medical imaging.

As a typical and important application, there are 13 papers involved with various medical image processing methods including segmentation, registration, fusion, denoising, and detection. F. Akram et al. present a region based image segmentation algorithm using active contours with signed

pressure force function. It has the potential to contemporaneously trace high intensity or dense regions in an image by evolving the contour inwards. Local feature calculation is important for delineation of hippocampus, a well-known biomarker for Alzheimer disease and other neurological and psychiatric diseases. S. Tangaro et al. compare four different techniques for feature selection from a set of 315 features extracted for each voxel. The authors obtain comparable state-of-the-art performances by using only 23 features for each voxel. M. Vlachos and E. Dermatas present a finger vein segmentation method from infrared images based on a modified separable Mumford-Shah model and local entropy thresholding method. In order to improve the performance of the registration in presence of tumor shrinkage between planning CT images and posttreatment CT images, J. Wang et al. propose a registration method by combining an image modification procedure and a fast symmetric Demons algorithm. L. Zhao and K. Jia propose a diffeomorphic image registration algorithm for capturing large and complex deformation by using a two-layer deep adaptive registration framework. S. Mazaheri et al. present an ultrasound image fusion method which weights the image information within the overlapping regions by using a combination of principal component analysis and discrete wavelet transform. It is expected to increase the segmentability of echocardiography features and decrease impact of noise and artifacts. For the purpose of quantitative analysis of the dynamic behavior about membrane-bound secretory vesicles, J. Wu et al. present a method to automatically identify the fusion events between VAMP2-pHluorin labeled GLUT4 storage vesicles and the plasma membrane in TIRF microscopy image sequences. J. Zhang et al. present a method automatically detecting the hinge point of mitral annulus in echocardiography by combining local context feature with additive support vector machines classifier. R. Xiao et al. present a seed point detection method by using adaptive ridge point extraction for coronary artery segmentation in X-ray angiogram image. L. Liu et al. present an adhesion pulmonary nodule detection method for 2D lung CT images based on the dot-filter and centerline extracting algorithm. R. Takalo et al. present an improved autoregressive model to reduce noise in SPECT images. This AR filter may be applied in both projection image and SPECT reconstruction image filtration. Nonlocal means filtering is an effective algorithm to remove the mottled noise by using large-scale similarity information in low-dose CT. However, it is very time-consuming. L. Zhang et al. present an optimized parallelization method for NLM filtering by avoiding the repeated computation with row-wise intensity calculation and the symmetry weight calculation. M. Martin-Fernandez and S. Villullas propose a MRI image denoising method by performing a shrinkage of wavelet coefficients based on the conditioned probability of noise. Instead of using an estimator of noise variance, its parameters are calculated by means of the Expectation Maximization method. B. Yu et al. present a method on estimating the binomial proportions of sensitive or stigmatizing attributes in the population of interest in successive sampling on two occasions.

We hope that this special issue may represent the state of the art and would attract wide attention of the researchers in medical imaging field.

Acknowledgment

Finally, we are grateful for the tremendous efforts by the authors and the reviewers.

Liang Li
Tianye Niu
Yi Gao

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 597313, 11 pages
http://dx.doi.org/10.1155/2015/597313

Research Article

Adhesion Pulmonary Nodules Detection Based on Dot-Filter and Extracting Centerline Algorithm

Liwei Liu,[1] Xin Wang,[1] Yang Li,[1] Liping Wang,[2] and Jianghui Dong[3]

[1]*College of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China*
[2]*Sansom Institute for Health Research and School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, SA 5001, Australia*
[3]*School of Natural and Built Environments, University of South Australia, Adelaide, SA 5095, Australia*

Correspondence should be addressed to Liwei Liu; liwei_liu2006@163.com and Liping Wang; wlp2005518@gmail.com

Received 2 August 2014; Revised 26 August 2014; Accepted 13 September 2014

Academic Editor: Yi Gao

A suspected pulmonary nodule detection method was proposed based on dot-filter and extracting centerline algorithm. In this paper, we focus on the distinguishing adhesion pulmonary nodules attached to vessels in two-dimensional (2D) lung computed tomography (CT) images. Firstly, the dot-filter based on Hessian matrix was constructed to enhance the circular area of the pulmonary CT images, which enhanced the circular suspected pulmonary nodule and suppresses the line-like areas. Secondly, to detect the nondistinguishable attached pulmonary nodules by the dot-filter, an algorithm based on extracting centerline was developed to enhance the circle area formed by the end or head of the vessels including the intersection of the lines. 20 sets of CT images were used in the experiments. In addition, 20 true/false nodules extracted were used to test the function of classifier. The experimental results show that the method based on dot-filter and extracting centerline algorithm can detect the attached pulmonary nodules accurately, which is a basis for further studies on the pulmonary nodule detection and diagnose.

1. Introduction

Pulmonary nodules are small masses of tissue in the lung, are prevalent findings on chest and abdominal CT scans, and can be cancerous, though most of them are benign [1]. Lung cancer is one of the biggest malignancy cancers among all kinds of cancers in our healthy life [2, 3] and is also the most common histological type in Aden carcinoma [4]. In recent years, the number of people suffering from lung cancer increases more and more rapidly. The early stage lung cancer is shown as lung nodules, which can be discovered and treated with the assistance of computer-aided diagnostic technique in time, which will prolong the life of lung cancer patients [5, 6]. The computer-aided diagnostic scheme can detect the nodules automatically in the pulmonary CT images and decrease the miss rate [5, 7], especially with the low-dose CT (LDCT) scanning [8].

To date, many researchers all over the world are devoted to the study of the detection of attached pulmonary nodules, for example, nodule attached to vessels and the pulmonary wall. However, limitations occur in lung cancer imaging of distinguishing nodules attached to vessels from the normal blood vessels, which infiltrate the vessels surreptitiously. Using the corrosion morphology and expansion to segment the pulmonary nodules from the vessels resulted in the corrosion of the nodule thorn, which is another important index for malignant nodule valuation [9]. A weighted fuzzy C-means clustering was developed for remotely sensed image classification but requires a given number of clustering and is easy to fall into local minimum rather than the global optimal solution [10]. A method based on EM and Mean-shift or one of the two means was proposed to detect attached nodules, but there are many conditions need to be considered, and not entirely consistent with the actual situation [11, 12]. Algorithm Based on Fuzzy Integrated Active Contour Model and Hybrid Parametric Mixture Model to detect pulmonary nodules just extracts the adhesion nodules, but it did not exclude the false positives such as ends of the vessels. For the value of the pixel

on the nodule which is close to that of pixel on the vessel, the gray threshold cannot work well and morphological operations cannot identify the adhesion nodules effectively [13–15]. Guo et al. developed a pulmonary nodule detection algorithm based on multiscale enhancement filtering of Hessian matrix and selecting of grads entropy, where Hessian matrix is relative to the gray scale of the pixel in the CT image, and grads entropy is also relative to gray scale of the pixel [16]. It worked well in the solitary pulmonary nodules detection, but it can only detect most suspect nodules and cannot exclude the false positives, especially the ends and the cross sections of the vessels or tracheas. Template matching method can be used to extract suspected nodules, but this will need more human intervention and prior information [17]. For solitary pulmonary nodules, regional growth can obtain good segmentation results [3]; for region growing segmentation results are part of vessel without separation and nodule. The method based on SVMS to detect the nodules, worked well, but it required a long processing time and lots of work [18–20].

Pulmonary nodules are similar to spherical objects, and the lung CT images are 2D. In order to enhance the dot-like regions and depress the line-like regions quickly and effectively, an algorithm named dot-filter was proposed by Li et al. [21]. However, when it was applied to detect pulmonary nodules, many false positives appeared, such as the ends and cross sections of the vessels and tracheas [16]. We found that the distances from the adhesion nodules center or false positives to the centerline of the vessel or tracheas were different. In this paper, starting from the relationship of their position, we combine dot-Filter and algorithm of extracting centerline, using which to identify which is the end or head of the vessel and which the circle formed by the intersection of the lines. In this way, we can separate the nodules from vessels and tracheas effectively with fewer steps.

2. Materials and Methods

2.1. Algorithms of Adhesion Pulmonary Nodules Detection. The process of the algorithms used in this paper was shown in Figure 1. Firstly we removed the background noise from the initial CT images and then extract the lung parenchyma. Secondly we used the Gauss function to convolute the image and a smooth image can be obtained. After that we can use dot-filter to enhance the dot-like regions to obtain suspect nodules. At last, we used the extracting centerline algorithm to analyse the relationship of the position of the suspect nodules between the vessels and tracheas, which was used to recognize the adhesion pulmonary nodules.

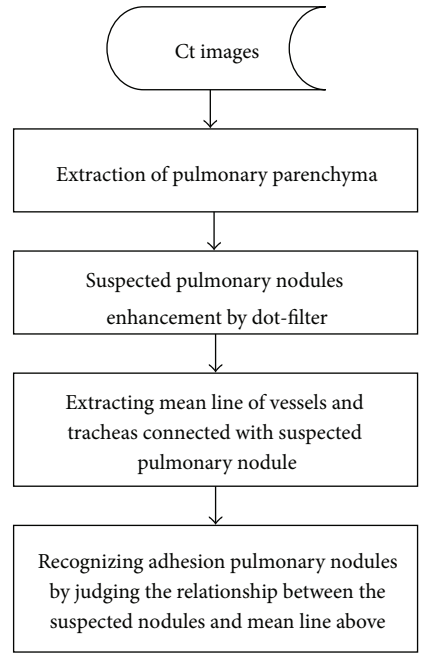


Figure 1: The process of the algorithms used to recognize adhesion pulmonary nodules.

2.2. Enhancement of Nodules by Dot-Filter

2.2.1. Dot-Filter Constructed by Hessian Matrix. To a medical CT image, the enhancement filter of local structure was used extensively which is based on the shape of organization. On a 2D image, we used the dot model conforming to Gauss distribution to represent a nodule [21, 22] as well as line model; the equation is expressed as

$$d(x,y) = \exp\left\{-\frac{x^2+y^2}{2\sigma^2}\right\},$$
$$l(x,y) = \exp\left\{-\frac{x^2}{2\sigma^2}\right\}. \quad (1)$$

Here, $d(x, y)$ denotes a dot expression expressed by a 2D Gaussian function; σ represents the dimension of the dot and the line. Because of the variety values of σ, we simulate the image of dots and lines shown in Figure 2(a).

Li et al. [21] proposed that dot-filter can be constructed by using Hessian matrix to effectively extract dot-like objects. For an original 2D image, we assume it has four second derivatives f_{xx}, f_{xy}, f_{yx}, and f_{yy}, where $f_{xy} = f_{yx}$ and its 2D Hessian matrix is

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}. \quad (2)$$

$f(x, y)$ is the value of one of the pixels in the image. Suppose the Eigenvalues of H are λ_1 and λ_2 and satisfied that abs $|\lambda_1|$ is bigger than $|\lambda_2|$. If $|\lambda_1| < |\lambda_2|$, exchange them. The $|\lambda_1|$ and $|\lambda_2|$ of the dot and line in the image satisfy the following expressions:

$$\text{dot: } \lambda_1 = \lambda_2 = -\frac{1}{\sigma^2} < 0,$$
$$\text{line: } \lambda_1 = -\frac{1}{\sigma^2} < 0, \quad \lambda_2 = 0. \quad (3)$$

(a) (b) (c)

Figure 2: (a) simulate the model of Gaussian dot; (b) enhanced with one scale; (c) enhanced with multitude scales.

The enhanced dot-filter is expressed by the following expression [21]:

$$w_d = \frac{|\lambda_2|^2}{\lambda_1} \quad \text{if } \lambda_1 < 0,\ \lambda_2 < 0,$$
$$w_d = 0 \text{ others;} \tag{4}$$

In CT images, if the semidiameter of one pulmonary nodule is σ_0, the nodule will account for 49.9% of the area of the Gauss function. If it is $2\sigma_0$, it will account for 72.0% of the area of the Gauss function. And if $3\sigma_0$, it accounts for 99.0% of the area of the Gauss function. Then, to a nodule of which the semidiameter is r, we use one Gauss function with σ_0 that equals $r/3$ to express it better [16]. For the position of the pulmonary nodules in CT images is different, the scale of the nodules is different between them. If the range of the scale of the nodules is $[r_0, r_1]$, the σ in Gaussian function is in $[r_0/3, r_1/3]$. In order to enhance all the goals in the range, we use different value of σ in Gaussian function to smooth a 2D CT image firstly; then we use the dot-filter constructed with Hessian matrix to enhance the goal area. The two steps above should be repeated N times with increasing scale of σ from σ_0 to σ_1 to obtain N enhanced CT images. If the range becomes bigger, N will become bigger [16, 21]. In lung CT images, we find that the value of N which equals 5 is better. In the range of $[r_0/3, r_1/3]$, the algorithm to obtain the σ can be shown as follows:

$$\begin{aligned} \sigma_0 &= \frac{r_0}{3}, \\ \sigma_1 &= r\sigma_0, \\ \sigma_2 &= r^2\sigma_0, \\ &\vdots \\ \sigma_N &= r^N\sigma_0 = \frac{r_1}{3}; \end{aligned} \tag{5}$$

among which the r equals $(r_1/r_0)^{1/(N-1)}$. In each N scale, we can obtain one most effective enhancement to the appointed nodule.

The steps of extracting dot with numbers of scales of dot-filter are as follows:

(1) According to the range of scale of the nodules we compute the value of σ.

(2) For every σ, repeat (3)–(8).

(3) Using Gaussian function convolve with 2D $f(x, y)$.

(4) For every pixel, repeat (5)–(7).

(5) Compute H and $|\lambda_1|$, $|\lambda_2|$.

(6) Compute w_d.

(7) Stop computing.

(8) Select the maximum of w_d.

In order to prove better the effect of using dot-filter with variety value of σ to identify the dot-like shapes, we use Figure 2(a) as input, and the output is shown as Figures 2(b) and 2(c).

Figure 2(a) is an image constructed by the expression (1) with variety scale of σ, and there are five dots and three lines. The scales of the σ among the dots are 2, 4, 6, 8, and 10 pixels. Figure 2(b) is the image, in which the better identified dot is enhanced by one dot-filter with the scale of 10 pixels. We also found that the lines is not identified and the dots smaller than 10 pixels do not have large output. Figure 2(c) is the image enhanced by four dot-filters, of which the dots equal to the scale of 2, 4, 6, 8, and 10 pixels all have large output, and the lines are depressed. According to Figure 2, we can prove that dot-filter can depress the line-like shapes and with variety value of σ it can extract all the goal areas better.

2.2.2. Application of Dot-Filter Constructed. As depicted above, we know that dot-filter can enhance the dot-like areas effectively. However, in the lung CT images, the ends and cross sections of vessels are also of dot-like shapes, which will be enhanced by using dot-filter, leading to many more false positives appearance. In order to prove that, we construct three types' vessel models, such as single line model, Y type

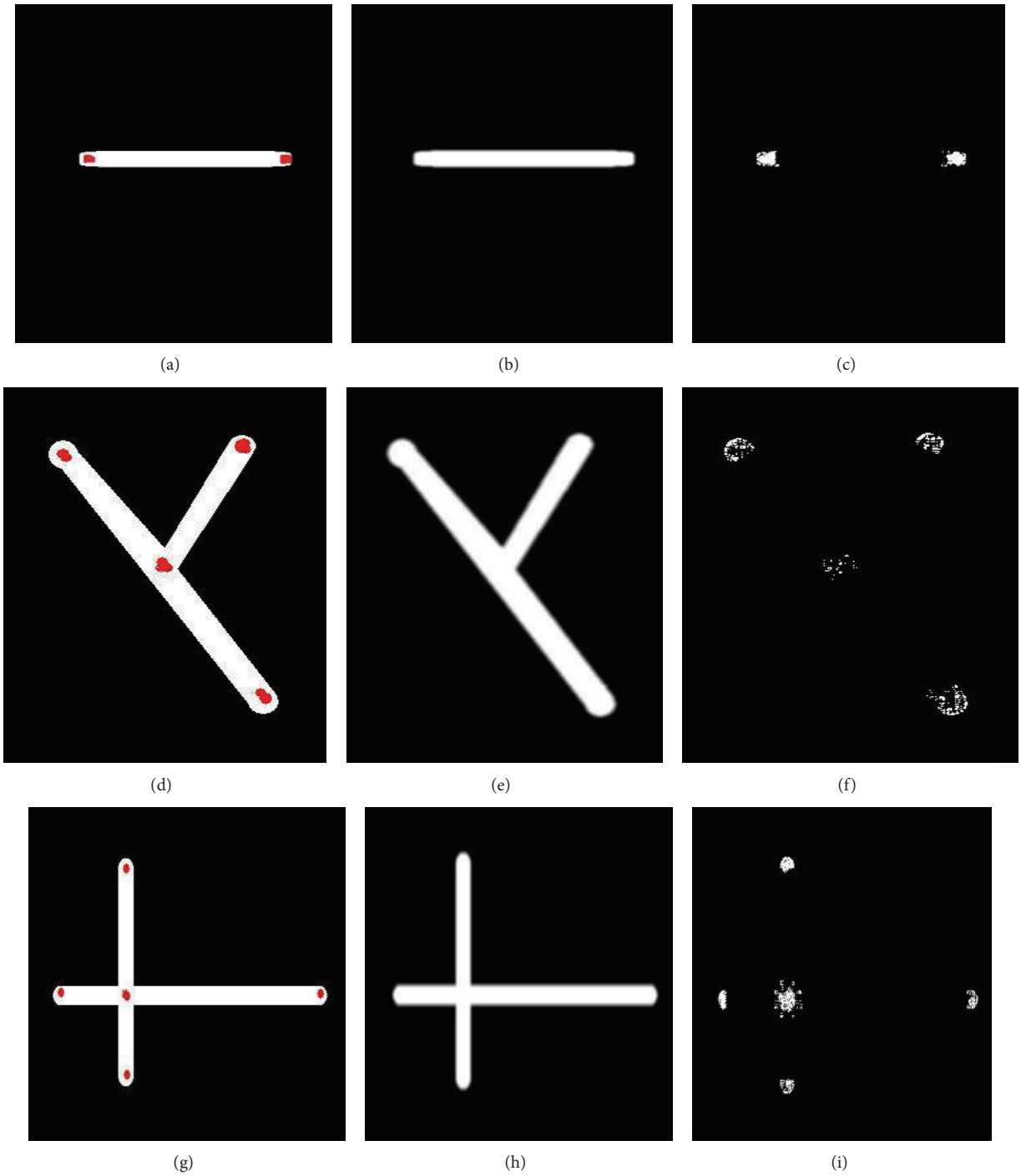


Figure 3: Models of vessels constructed. The regions marked by red in (a), (d), and (g) are dot-like regions, which will be enhanced by dot-filter. (a) the vessel of single line type; (b) smoothed by Gauss function; (c) enhanced by dot-filter; (d) the vessel of Y type; (e) smoothed by Gauss function; (f) enhanced by dot-filter; (g) the vessel of X type; (h) smoothed by Gauss function; (i) enhanced by dot-filter.

model, and X type model shown in Figures 3(a), 3(d), and 3(g). The regions marked by red in the images are dot-like regions, which will be enhanced by dot-filter. Figures 3(b), 3(e), and 3(h) were smoothed by Gauss function. Figures 3(c), 3(f), and 3(i) were the images enhanced by dot-filter. The enhanced areas that we marked in Figures 3(a), 3(d), and 3(g) are also called suspect nodules. Due to many suspect nodules that appeared after the enhanced process by dot-filter we need to eliminate these false positives which may lead to much more computation works.

Now we will use the Dot-Filter constructed above based on Hessian matrix to lung CT images, and the result is shown

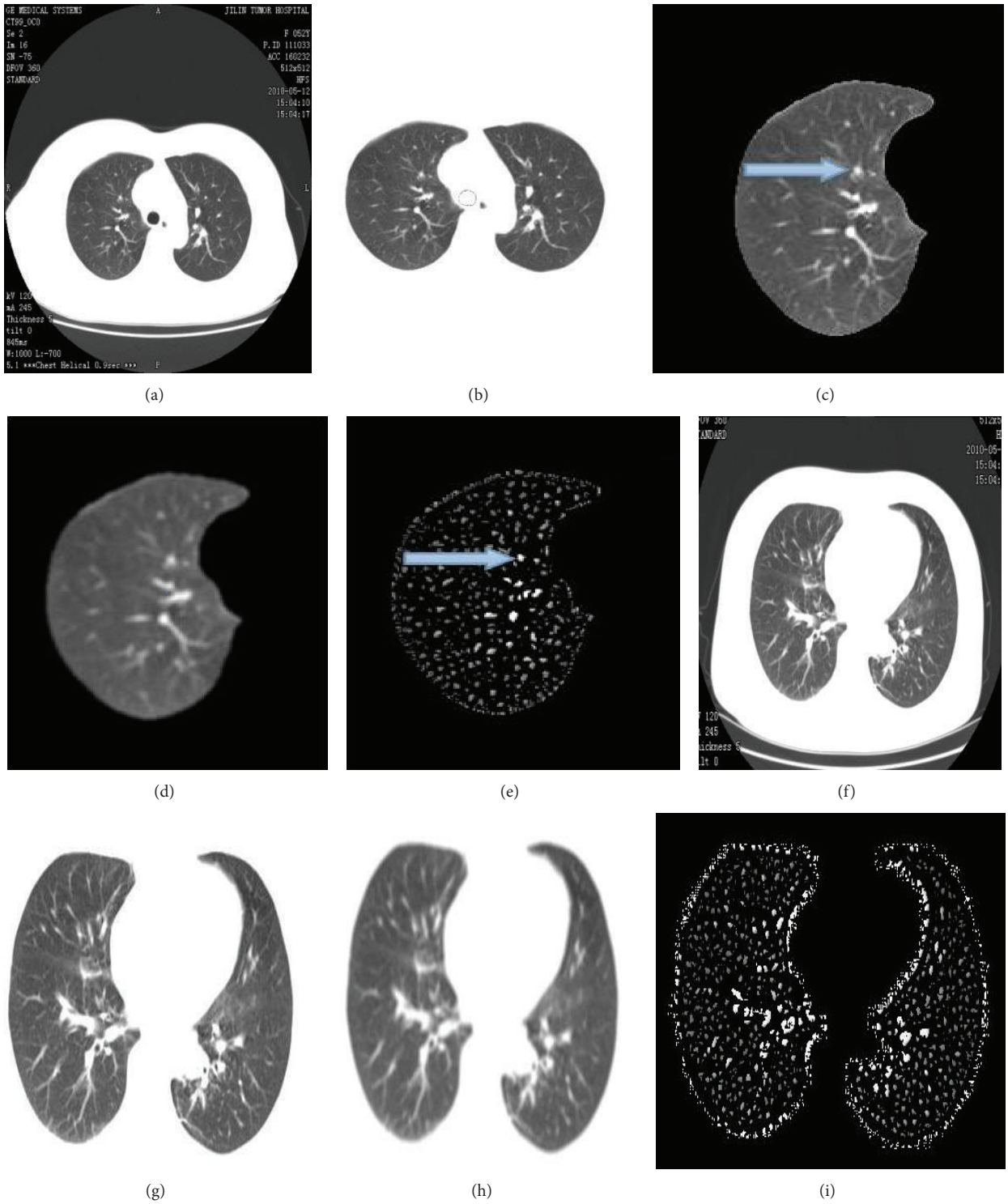


Figure 4: Detection of the solitary nodules in a 2D CT image enhanced by dot-Filter. (a) original lung image; (b) lung segment extraction; (c) the right lung segment; (d) lung CT image smoothed by Gauss function; (e) detection of the solitary nodules; (f) lung CT image without nodules; (g) pulmonary segment extracted; (h) smoothed by Gauss function; (i) enhanced by dot-filter.

in Figure 4. Lung CT images are given by one big hospital for lung nodules detection based on dot-Filter and the algorithm of extracting centerline.

Figure 4(a) is an original pulmonary CT image. Figure 4(b) is the pulmonary segment extracted through segmentation of digital image. Figure 4(c) is the right pulmonary segment. In the image the blue arrow denotes a nodule identified by the doctor. From the image we can find that there are many dot-like areas such as solitary areas and dot-like areas attached to the vessels, which we

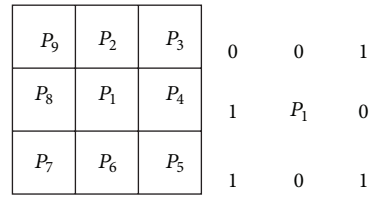


Figure 5: (a) The arrangement of the neighborhood pixels used in the thinning algorithm; (b) the explanation of expression (6) in conditions (a) and (b). At this time $N(p_1) = 4$ and $T(p_1) = 3$.

take as false positives. Figure 4(d) is smoothed by Gauss function. It is not ideal to enhance the region of interest (ROI) instantly because there is much noise in the image, so we had better firstly use Gauss function to convolute with it. Figure 4(e) is the enhancement of the solitary nodules and other dot-like area by using dot-filter. In this image we cannot identify which is real nodule without any other assistance or algorithm, because dot-filter just enhances the dot-like areas. To effectively distinguish whether the dot-like areas are nodules or just the part of the vessels we will have to use the algorithm of centerline extracted. Figure 4(f) is a lung CT image without nodules and Figure 4(g) is pulmonary segment extracted; Figures 4(h) and 4(i) are the image smoothed by Gauss function and enhanced by dot-filter, respectively. According to Figures 4(e) and 4(i), we obtain that dot-filter can enhance the dot-like areas effectively but leads to many false positives appearing.

2.3. Algorithm of Extracting Centerline. There are many algorithms used to extract the central line [23–28], such as margin of linear least square fitting legitimate, symmetric moment fitting center method, and block cancroids least squares fitting. Compared with the algorithms used in this paper, they are not stable and accurate enough and have high computational complexity. As Figure 4(e) shows, dot-filter can enhance the dot-like area significantly with more false positive increased. To overcome this shortcoming, we will combine the algorithm of extracting centerline to reduce the false positive.

2.3.1. Principle. Different from the traditional algorithms of extracting central line, area skeleton can be defined by mean axle transforming (MAT). Describe an area whose profile is b as follows: for every pixel p in the R, we search the nearest pixel in b. If p is bigger than the nearest pixel, we named p centerline (skeleton) of R, which obeys the following constraints: (1) cannot delete the endpoint; (2) cannot destroy connectivity; and (3) cannot cause excessive corrosion of the area.

We here give the mean of a refinement of two-value algorithm region: we suppose the value of the pixel in the region is 1, and the values of the pixels on the background are 0. The value of the pixels in the edge of the region is 1 and at least there is one pixel of which the value is 0. As 8 neighborhoods shown in Figure 5(a), if it meets the following conditions (a)–(d), then (step 1) we take p_1 as the pixel that will be removed as follows:

$$\begin{aligned} &\text{(a)}\ \ 2 \le N(p_1) \le 6, \\ &\text{(b)}\ \ T(p_1) = 1, \\ &\text{(c)}\ \ p_2 \times p_4 \times p_6 = 0, \\ &\text{(d)}\ \ p_4 \times p_6 \times p_8 = 0; \end{aligned} \tag{6}$$

among which $N(p_1)$ is the number of the nonzero adjacent pixels of p_1; in other words,

$$N(p_1) = p_2 + p_3 + \cdots + p_8 + p_9; \tag{7}$$

among which p_i is either 0 or 1, and $T(p_1)$ is the frequency conversion from 0 to 1 in $p_2, p_3, \ldots, p_8, p_9$. For example, in Figure 5(b), $N(p_1) = 4$ and $T(p_1) = 3$.

In Step 2, (a) and (b) remain unchanged, and (c) and (d) become

$$\begin{aligned} &\text{(c}')\ \ p_2 \times p_4 \times p_8 = 0, \\ &\text{(d}')\ \ p_2 \times p_6 \times p_8 = 0. \end{aligned} \tag{8}$$

We apply step 1 to every pixel in the edge of the two-value region. If we violate (a) or (b), the value of the pixel we talk about is unchanged. Otherwise, we take it as the pixel that will be removed after we handle all the pixels of the edge. Then, we use step 2 the same way as step 1 till there is no pixel needed to be removed any more and stop the algorithm.

Take Figure 6(a), which is processed by the mean of a refinement, for example, and the result is shown as Figures 6(b)~6(e).

Figure 6(a) is an image of a human chromosome by electron microscope magnified 30000 times and segmented using digital image processing algorithm. Figure 6(b) is the image after Gaussian smoothing. Figure 6(c) is the skeleton of the chromosome. Figure 6(d) shows skeletons after applying extinguishing the burr algorithm eight times. We found that on the skeleton there is much burr but less than that in Figure 6(c). Because this algorithm is related to the threshold of the pixel, we should increase the threshold value in the algorithm. Figure 6(e) presents seven more times for extinguishing the burr by using the algorithm.

If a line is expressed by $\alpha x + \beta y + \gamma = 0$, we will take $\Delta(-\alpha/\beta)$ and $\Delta(-\gamma/\beta)$ as the deviation [29]. If they are all very

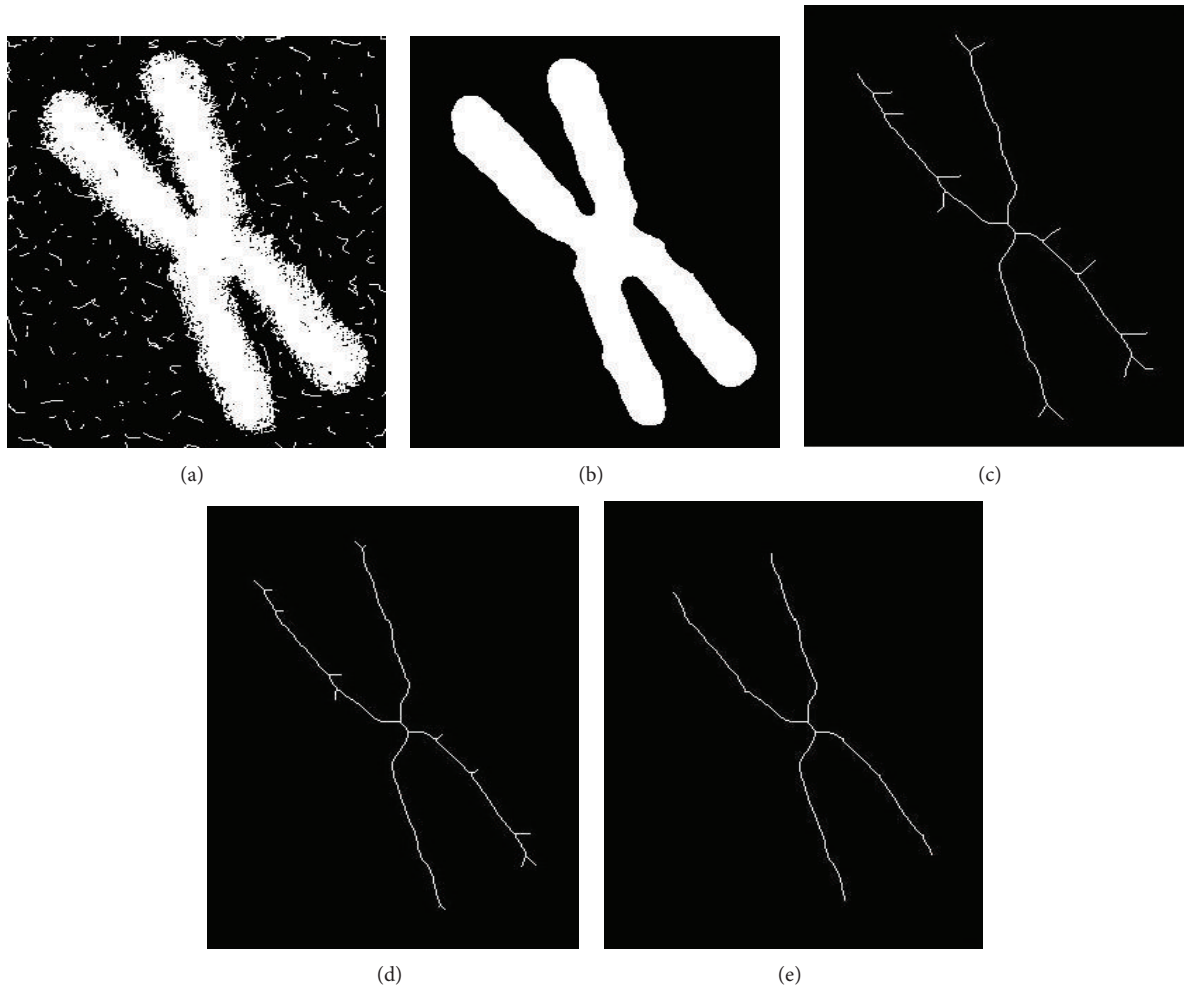


Figure 6: (a) chromosome image after segment; (b) image after Gaussian-filter; (c) the skeleton; (d) eight times for extinguishing the burr of the skeleton; (e) seven more times for extinguishing the burr.

small, we will think the algorithm works better. We use other three algorithms for extracting center line to compare with the one used in this paper; the result is shown as Table 1.

According to Table 1 we found the algorithm in this paper can work effectively, though the symmetric moment fitting center method consumes the least time, but its deviation was the bigger. The algorithm used in this paper worked steadily and time consuming is not much. Combining all the factors, the algorithm used in this paper is better.

2.3.2. Application. In this paper, to accomplish the experiment combining dot-filter with the method proposed above we use six steps shown as follows.

(1) We first selected three lung CT images after extracting the lung segment in Figures 7(a)~7(c); they were nodules attached to vessels, single vessel, and crossing vessel. There was one nodule attached to one end of the vessel noted by the doctor in Figure 7(a), shown as the arrow points to. In Figure 7(b), we can see that there is one vessel apparently and one of its ends is of dot-like shape, similar to the adhesion nodule in Figure 7(a). In Figure 7(c), the vessel is composed of two vessels and they are crossed.

(2) As the value of vessels, tracheas, and nodules are bigger than the value of lung parenchyma, in order to decrease the computation, we extracted the soft tissue of the lung based on gray threshold which we defined as 130, which is obtained after many times of drawing histogram, shown in Figures 7(d)~7(f). For the low contrast nodules, we did not consider them in this paper. According to Figures 7(d)~7(f), we found that the soft tissue of the lung parenchyma was completely extracted.

(3) To the tissues extracted in step (2), firstly we should extinguish the noise in the CT image. So we used Gauss filter to accomplish it. Then, we used the dot-filter constructed above to enhance the dot-like regions, in other words, we enhance the suspected nodules. The enhancement result was shown in Figures 7(g)~7(i).

Table 1: The difference between four algorithms for extracting center line in deviation and time consuming.

Algorithm for center-line extracted	$\Delta(-\alpha/\beta)$	$\Delta(-\gamma/\beta)$	t/s
Margin of linear least square fitting legitimate	0.142	45.437	0.042
Symmetric moment fitting center method	1.50	550.832	0.031
Block cancroids least squares fitting	0.671	332.117	0.033
Algorithm used in this paper	0.157	23.858	0.059

Table 2: Databases of testing the method used in this paper.

Name	LIDC database	Supported by Jida Hospital
Number of CT images	10	10
Pixel unit (volume)/mm^3	$0.6 \times 0.6 \times 0.6$	$0.6 \times 0.6 \times 0.6$
Average number	36	70
Number of adhesion nodules	3	13
Image size/pixels	512×512	512×512
Layer thickness	1 mm	1 mm

(4) We eliminated the tissues obtained in step (3) from the images in step (2) and then obtained the skeleton shown in Figures 7(j)~7(l). We considered the tissues obtained in step (3) as the false positives; we should remove them and use the algorithm of extracting centerline to extract the skeletons. But for Figure 7(f), the vessels became three parts, which was not beneficial for us to use the algorithm of extract the center line, so we firstly supplied the lack, making the vessels become one connected vessel and then extracted the skeleton.

(5) Firstly we worked out the center of every suspected nodule (false positive) and then computed the value of d_1, d_2. d_1 denotes the perpendicular distance from the center of mass of suspected nodules to the line of the skeleton near the suspected nodule obtained in step (4). d_2 denotes the minimum distance from the center of mass of suspected nodules to all pixels in the skeleton obtained in step (4).

(6) The diameter of the nodules is 3 mm~30 mm, so we compared the diameter with d_1 obtained in (5). If d_1 is smaller than 1.5 mm and d_2 is smaller than 1.5 mm, the suspected nodule is treated as the intersection of the vessels or the end of the vessel. If d_1 was smaller than 1.5 mm and d_2 was bigger than 1.5 mm, the suspected nodule was treated as one end of the vessel. If d_1 was bigger than 1.5 mm, the suspected nodule was treated as the attached nodule shown in Figure 7(m).

Figures 7(p)~7(r) is the three-dimension (3D) display of the nodule, the solitary vessel, and the crossing vessel. In Figure 7(p) the nodule with green edge is connected to the vessel in the yellow circle. Figures 7(q) and 7(r) are two kinds of different vessels in the green circle in order to be found easily. According to the three images we can easily find that the suspect nodules appeared in the process; they are actualized parts of vessels.

Table 3: The method used in this paper compared with current two methods.

Method	Number of false positives per set	Missing rate/(%)	Runtime/(min)
Literature [9]	8.6	33.3	2.8
Literature [12]	11.2	27.5	4.2
Method with only dot-filter	34.8	18.7	1.2
Method with dot-filter and centerline extraction	5.3	18.7	1.7

3. Results and Discussion

Table 2 depicts what are the attributes and where they come from the CT images used in this paper. 20 sets of CT images with less noise [25] were used in the experiments. They originated from LIDC database and Jida Hospital and each CT image has 512×512 pixels. Nodules in each CT image have been noted by doctors. The 20 true nodules and 20 false ones extracted were used to test the function of classifier. All experiments in this paper were based on the computer that consists of AMD CPU with a frequency of 2 GHz, 1.5 GB of RAM, and Windows XP operating system. Algorithm development code is developed on the platform of MATLAB.

For the CT images supported by the hospital it missed 3 adhesion nodules and missed none for the LIDC database. Table 3 shows the missing rate and runtime of each set of CT images. Literature [9] can better extract the solitary nodules but lacks the high capacity of extracting the adhesion lung nodules because it is based on the threshold value of the pixel and it needs much justice. Literature [12] uses the algorithm which has too much computation. If we use a method with only dot-filter, there will be more false positives appearance because dot-filter can enhance the ends and the intersections of the vessels meanwhile. By using dot-filter and the algorithm of extracting the center line, we have lower missing rate and less runtime. This is because dot-filter constructed with Hessian matrix can extract the dot-like region effectively and quickly. And after extracting the center line of the vessels, we can achieve a better understanding of the relationship between the nodules and vessels. According to the relationship, it is beneficial for us to extract the nodules. There are errors of this method because scale of some nodules is very small or the value of the pixels in nodules is very small, which can lead to the miss rate increases. The method in this paper has some limitations; for example, it cannot adapt to the lower contrast nodules and the nodules attached to the lung wall. It is just applied to the nodules attached to vessels and tracheas.

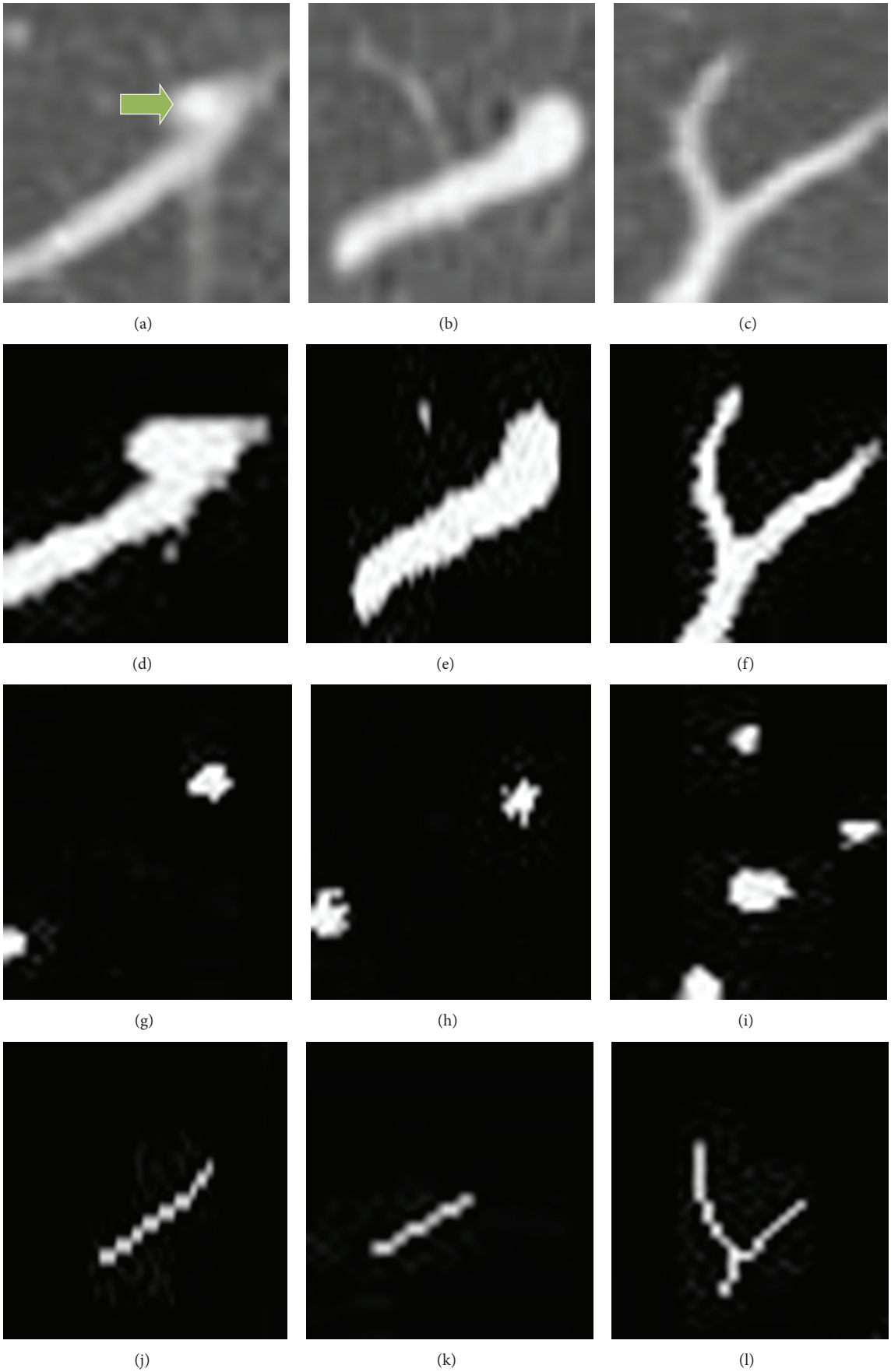


Figure 7: Continued.

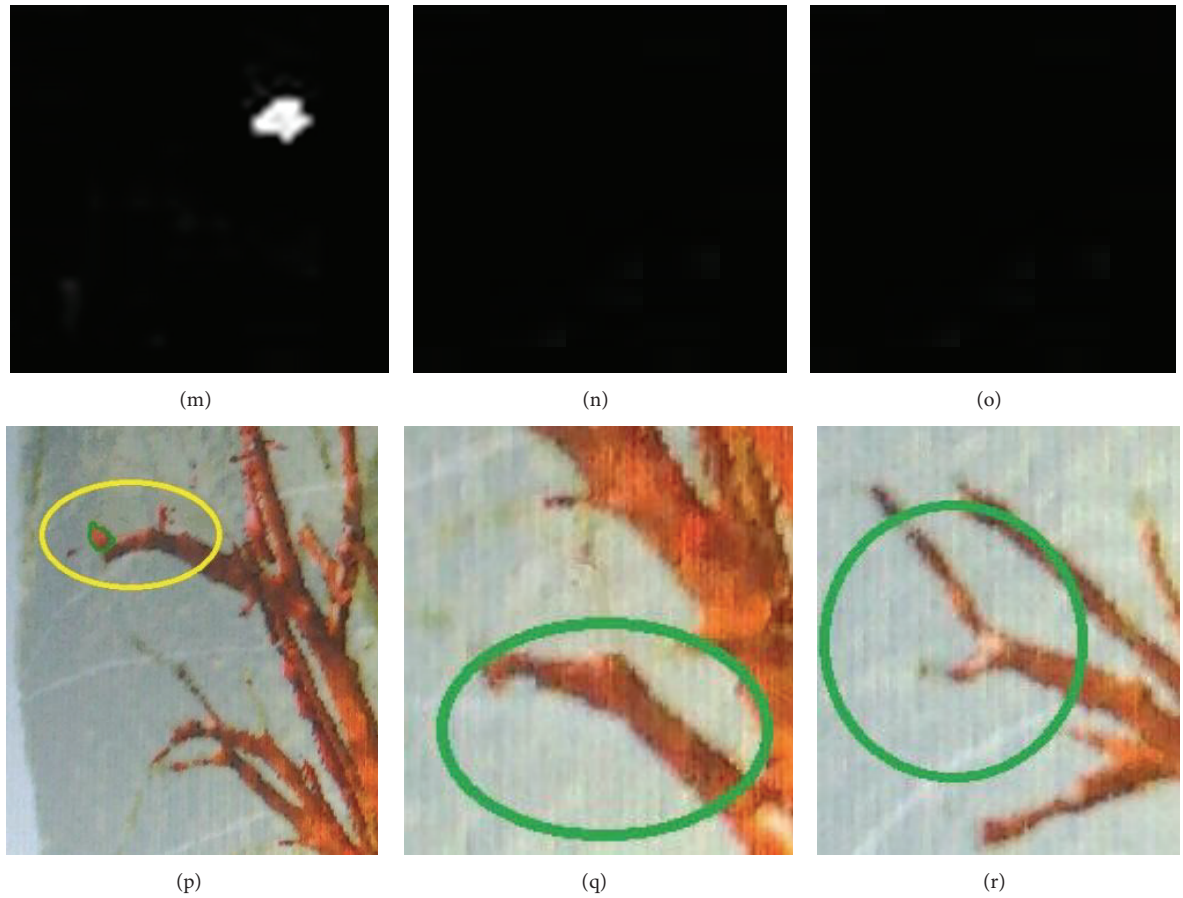


Figure 7: (a)~(c) original image; (d)~(f) soft tissue extraction of lung; (g)~(i) images after dot-filter; (j)~(l) images after seven times for extracting the skeleton; (m)~(o) attached nodule extraction; (p)~(r) is the three-dimension display of the objects.

4. Conclusion

In this paper we first use 2D Hessian matrix to construct dot-filter constructed to extract dot-like region. In order to solve the problem that the dot-filter cannot detect attached pulmonary nodules, an algorithm based on extracting centerline was used. Results of experiment indicated that the method is easy and effective while extracting attached pulmonary nodules well. In the future, we will be devoted to extracting the lung nodules contacting pulmonary wall and ground glass opacity pulmonary nodules.

Conflict of Interests

The authors declared that they have no conflict of interests regarding this work.

References

[1] K. Alexander, H. Joly, L. Blond et al., "A comparison of computed tomography, computed radiography, and film-screen radiography for the detection of canine pulmonary nodules," *Veterinary Radiology and Ultrasound*, vol. 53, no. 3, pp. 258–265, 2012.

[2] T. Hiraki, H. Gobara, T. Iguchi, H. Fujiwara, Y. Matsui, and S. Kanazawa, "Radiofrequency ablation for early-stage nonsmall cell lung cancer," *BioMed Research International*, vol. 2014, Article ID 152087, 11 pages, 2014.

[3] M. Inoue, S. Nakatsuka, and M. Jinzaki, "Cryoablation of early-stage primary lung cancer," *BioMed Research International*, vol. 2014, Article ID 521691, 8 pages, 2014.

[4] W. D. Travis, E. Brambilla, M. Noguchi et al., "International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma," *Journal of Thoracic Oncology*, vol. 6, no. 2, pp. 244–285, 2011.

[5] W.-J. Choi and T.-S. Choi, "Genetic programming-based feature transform and classification for the automatic detection of pulmonary nodules on computed tomography images," *Information Sciences*, vol. 212, pp. 57–78, 2012.

[6] C. M. Niranjana and P. Deepa, "Nodule detection in lung intervention by using VDE and morphology techniques," *International Journal of Research in Computer Applications and Robotics*, vol. 2, no. 5, pp. 114–123, 2014.

[7] S. C. Park, B. E. Chapman, and B. Zheng, "A multistage approach to improve performance of computer-aided detection of pulmonary embolisms depicted on CT Images: preliminary investigation," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1519–1527, 2011.

[8] F. Niknam, J. Chen, S. Napaki, and M. Aghmesheh, "Approach to multiple pulmonary nodules: a case report and review of literature," *The Scientific World Journal*, vol. 11, pp. 760–765, 2011.

[9] W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke, "Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images," *IEEE Transactions on Medical Imaging*, vol. 22, no. 10, pp. 1259–1274, 2003.

[10] C.-C. Hung, S. Kulkarni, and B.-C. Kuo, "A new weighted fuzzy C-means clustering algorithm for remotely sensed image classification," *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 543–553, 2011.

[11] S.-S. Sun, H. Li, X.-R. Hou, Y. Kang, and H. Zhao, "Pulmonary nodule segmentation based on EM and mean-shift," *Journal of Image and Graphics*, vol. 10, pp. 2016–2022, 2009.

[12] K. Okada, D. Comaniciu, and A. Krishnan, "Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multi-slice CT," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 281–288, 2003.

[13] V. Sudha and P. Jayashree, "Lung nodule detection in CT images by using thresholding and morphological operations," *International Journal of Emerging Science and Engineering*, vol. 1, no. 2, 2012.

[14] B. Chen, T. Kitasaka, H. Honma et al., "Automatic segmentation of pulmonary blood vessels and nodules based on local intensity structure analysis and surface propagation in 3D chest CT images," *International Journal of Computer Assisted Radiology and Surgery*, vol. 7, no. 3, pp. 465–482, 2012.

[15] A. O. de Carvalho Filho, W. B. de Sampaio, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, "Automatic detection of solitary lung nodules using quality threshold clustering, genetic algorithm and diversity index," *Artificial Intelligence in Medicine*, vol. 60, no. 3, pp. 165–177, 2014.

[16] W. Guo, Y. Wei, H. Zhou, and D. Xue, "Suspected pulmonary nodule detection algorithm based on Hessian matrix and grads entropy," *Chinese Journal of Scientific Instrument*, vol. 30, no. 8, pp. 1702–1706, 2009.

[17] Y. Lee, T. Hara, H. Fujita, S. Itoh, and T. Ishigaki, "Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 595–604, 2001.

[18] M. J. Gangeh, L. Sørensen, S. B. Shaker, M. S. Kamel, M. de Bruijne, and M. Loog, "A texton-based approach for the classification of lung parenchyma in CT images," *Medical Image Computing and Computer-Assisted Intervention*, vol. 13, part 3, pp. 595–602, 2010.

[19] S.-S. Sun, H.-Z. Ren, Y. Kang, and H. Zhao, "Lung nodule detection by GA and SVM," *Journal of System Simulation*, vol. 23, no. 3, pp. 497–501, 2011.

[20] J. S. Bie, *Segmentation and recognition of lung nodules in CT images based on SVM [M.S. thesis]*, 2012.

[21] Q. Li, S. Sone, and K. Doi, "Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans," *Medical Physics*, vol. 30, no. 8, pp. 2040–2051, 2003.

[22] Q. Li, F. Li, and K. Doi, "Computerized detection of lung nodules in thin-section CT images byuse of selective enhancement filters and an automated rule-based classifier," *Academic Radiology*, vol. 15, no. 2, pp. 165–175, 2008.

[23] S. Kurugol, E. Bas, D. Erdogmus, J. G. Dy, G. C. Sharp, and D. H. Brooks, "Centerline extraction with principal curve tracing to improve 3D level set esophagus segmentation in CT images," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2011, pp. 3403–3406, IEEE Engineering in Medicine and Biology Society, 2012.

[24] J. Hai, G. Zhang, and J. Cheng, "Estimation of mass transfer coefficient in ozone absorption by linear least square fitting and Simplex search methods," *Journal of Central South University*, vol. 19, no. 12, pp. 3396–3399, 2012.

[25] P.-H. Tseng and K.-T. Feng, "Derivation of CRLB for linear least square estimator in wireless location systems," *Wireless Networks*, vol. 18, no. 7, pp. 735–747, 2012.

[26] L. M. Alawneh, C. J. Park, M. K. Jaradat, and B. Lee, "Burnup estimation for plate type fuel assembly of research reactors through the least square fitting method," *Annals of Nuclear Energy*, vol. 71, pp. 37–45, 2014.

[27] S. Sergey, M. van der Kooij, and K. Tiampo, "A simultaneous inversion for deformation rates and topographic errors of DInSAR data utilizing linear least square inversion technique," *Computers and Geosciences*, vol. 37, no. 8, pp. 1083–1091, 2011.

[28] W.-P. Choi, K.-M. Lam, and W.-C. Siu, "Extraction of the Euclidean skeleton based on a connectivity criterion," *Pattern Recognition*, vol. 36, no. 3, pp. 721–729, 2003.

[29] K. Wang, A. You, and L. Wang, "Algorithm of centerline extracted based on eigen decomposition of Hessian matrix," *High Power Laser and Particle Beams*, vol. 25, no. 1, pp. 24–28, 2013.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 108274, 15 pages
http://dx.doi.org/10.1155/2015/108274

Research Article

A Reconstruction Method of Blood Flow Velocity in Left Ventricle Using Color Flow Ultrasound

Jaeseong Jang,[1] Chi Young Ahn,[2] Kiwan Jeon,[2] Jung Heo,[3] DongHak Lee,[3] Chulmin Joo,[3] Jung-il Choi,[1] and Jin Keun Seo[1]

[1]*Department of Computational Science and Engineering, Yonsei University, Seoul 120-749, Republic of Korea*
[2]*Division of Computational Mathematics, National Institute for Mathematical Sciences, Daejeon 305-811, Republic of Korea*
[3]*School of Mechanical Engineering, Yonsei University, Seoul 120-749, Republic of Korea*

Correspondence should be addressed to Chi Young Ahn; chiyoung@nims.re.kr

Received 29 October 2014; Accepted 25 December 2014

Academic Editor: Tianye Niu

Vortex flow imaging is a relatively new medical imaging method for the dynamic visualization of intracardiac blood flow, a potentially useful index of cardiac dysfunction. A reconstruction method is proposed here to quantify the distribution of blood flow velocity fields inside the left ventricle from color flow images compiled from ultrasound measurements. In this paper, a 2D incompressible Navier-Stokes equation with a mass source term is proposed to utilize the measurable color flow ultrasound data in a plane along with the moving boundary condition. The proposed model reflects out-of-plane blood flows on the imaging plane through the mass source term. The boundary conditions to solve the system of equations are derived from the dimensions of the ventricle extracted from 2D echocardiography data. The performance of the proposed method is evaluated numerically using synthetic flow data acquired from simulating left ventricle flows. The numerical simulations show the feasibility and potential usefulness of the proposed method of reconstructing the intracardiac flow fields. Of particular note is the finding that the mass source term in the proposed model improves the reconstruction performance.

1. Introduction

Vortex flow imaging has recently attracted much attention in the field of clinical cardiac assessment owing to reports of its feasibility for analyzing intraventricular vortex flows [1–3]. The vorticity of intraventricular blood flow describes a rotational flow pattern that offers possible clinical indices of cardiac functions such as sphericity, vortex depth, vortex length, and vortex pulsation correlation.

There are several methods to compute and visualize the velocity fields of blood flow inside the left ventricle (LV), with echo particle image velocimetry (E-PIV) being representative of the commonly used noninvasive methods [4]. It tracks the speckle patterns of blood flow to estimate blood motion within the imaging plane. Although it is generally unable to measure out-of-plane particle motion from 2D echocardiography data (called B-mode images), a recent study extending E-PIV to 3D volume data demonstrated the possibility of out-of-plane assessment [5]. However, E-PIV is not completely noninvasive because it requires the intravenous injection of a contrast agent to obtain images suitable for the speckle-tracking algorithm.

To develop less invasive techniques, methods to reconstruct blood flows from color flow images (also called C-mode images, color Doppler images, color Doppler data, or Doppler echocardiography) have been proposed. The color flow images reflect the projected velocity components in the direction of ultrasound beam propagation [6]. To compute the flow velocity from color flow images, Garcia et al. [7] assumed a 2D divergence-free condition on the velocity fields; they decomposed each 2D velocity vector into a radial component obtained from the color flow data and an unknown angular component, which was computed using their assumption of the 2D flow. Ohtsuki and Tanaka [8] also assumed 2D flows and recovered the 2D velocity fields from the color Doppler data using the concepts of stream function and streamline in 2D fluid flow. However, the assumption of

a 2D divergence-free condition is an oversimplification that ignores out-of-plane flows.

Arigovindan et al. [9] proposed a velocity reconstruction method using color Doppler data acquired from beams in two different directions. To cope with the nonuniformly sampled data of multiple imaging planes, they used 2D B-spline on each of the velocity components to be estimated, and the unknown coefficients of the 2D B-spline were calculated from the measured color Doppler data using least squares. Similar to the 2D reconstruction, Gomez et al. [10] recovered 3D velocity fields from multiple registered color Doppler images using 3D B-spline and least squares. The registration of multiple imaging planes for the above two methods remains very challenging in a practical environment.

Recently, a new imaging modality (Doppler vortography) based on 2D color Doppler data was introduced by Mehregan et al. [11], who assumed that a vortex flow pattern has axisymmetric features in the neighborhood of its center. Their method employs a simple kernel filter designed to find the positions of axisymmetry in the 2D color Doppler images. The vortex flow was recovered using a color Doppler-variable vorticity function that directly computes vorticities from color Doppler values. However, the assumption of axisymmetry does not reflect detailed flow patterns, and it may lead to inaccurate vortex positions and vorticity values in patients with severe dysfunction where axisymmetry cannot be assumed at all.

In this paper, we propose a 2D Navier-Stokes model to reconstruct intraventricular flows using color flow images and LV boundaries extracted from echocardiography data. Although the use of the full 2D Navier-Stokes equations in this setting has already been proposed and evaluated for 2D flow field regularization [12], the originality of this work is the inclusion of a source-term to deal with the out-of-plane flow component. The proposed model considers both in-plane and out-of-plane blood flows for an imaging plane in apical long-axis three-chamber (A3CH) view. Particular attention is given to the appearance and disappearance of the out-of-plane components in the imaging plane, which is modeled as a mass source term of a source-sink distribution. Blood flows in the imaging domain are reconstructed through solving a system of equations, which include a 2D incompressible Navier-Stokes equation for the mass source term and the color flow data measurement equation describing the projected velocity component for the color flow data. The boundary conditions required to solve the system of equations are given by the LV borders extracted from echocardiography data.

The performance of the proposed method is evaluated numerically using synthetic flow data with LV motion. The proposed method is shown to be feasible and potentially valuable for reconstructing intracardiac flow fields.

2. Materials and Methods

Commonly used ultrasound systems can provide not only 2D echo images but also color flow images, which represent the scanline directional components of the velocity fields using the phase-shift estimated by a standard autocorrelation algorithm [13]. Our flow reconstruction method is to reconstruct the intraventricular flows using color flow images and the LV boundaries extracted from the echo images. In this section, we describe the overall outline of our flow reconstruction method using those ultrasound measurements based on two assumptions as follows:

(i) the time difference between sequential color flow imaging frames is very small;

(ii) the echo and color flow images are acquired simultaneously and separately for the entire heart cycle.

Table 1: Comparison of L_2-norm errors for the velocity fields recovered from the synthetic and real data. The third and fourth columns represent normalized L_2-norm errors for pointwise and global energy estimates of velocity. Here, $\|\mathbf{v}\|_2 = \sqrt{u^2 + v^2}$, $\Delta\mathbf{v} = \mathbf{v}^{\text{recon}} - \mathbf{v}^{\text{exact}}$, and $\Delta k/k^{\text{exact}} = |\|\mathbf{v}^{\text{exact}}\|_2 - \|\mathbf{v}^{\text{recon}}\|_2|/\|\mathbf{v}^{\text{exact}}\|_2$.

t/T	$\|\mathbf{v}^{\text{exact}}\|_2$	$\|\mathbf{v}^{\text{recon}}\|_2$	$\|\Delta\mathbf{v}\|_2 / \|\mathbf{v}^{\text{exact}}\|_2$	$\Delta k/k^{\text{exact}}$
0.1	0.012	0.011	28.2%	11.1%
0.2	0.026	0.023	17.6%	8.9%
0.3	0.033	0.029	21.2%	9.9%
0.4	0.028	0.024	28.2%	12.0%
0.5	0.018	0.015	39.6%	17.4%
0.6	0.015	0.013	39.8%	16.8%
0.7	0.015	0.013	36.6%	13.4%
0.8	0.014	0.012	36.3%	12.0%
0.9	0.011	0.010	39.3%	13.0%
1.0	0.010	0.007	47.5%	18.6%

Table 2: Comparison of L_2-norm errors for the corresponding vortex fields. The third and fourth columns represent normalized L_2-norm errors for pointwise and global energy estimates of vorticity. Here, $\Delta\omega = \omega^{\text{recon}} - \omega^{\text{exact}}$ and $\Delta k_\omega/k_\omega^{\text{exact}} = |\|\omega^{\text{exact}}\|_2 - \|\omega^{\text{recon}}\|_2|/\|\omega^{\text{exact}}\|_2$.

t/T	$\|\omega^{\text{exact}}\|_2$	$\|\omega^{\text{recon}}\|_2$	$\|\Delta\omega\|_2 / \|\omega^{\text{exact}}\|_2$	$\Delta k_\omega/k_\omega^{\text{exact}}$
0.1	1.09	1.19	36.8%	8.7%
0.2	2.35	2.48	24.7%	5.3%
0.3	2.82	2.90	21.7%	2.8%
0.4	2.31	2.34	21.2%	1.4%
0.5	1.67	1.67	32.1%	0.4%
0.6	1.38	1.40	40.6%	1.7%
0.7	1.33	1.40	46.7%	5.2%
0.8	1.20	1.30	51.4%	7.8%
0.9	0.95	1.02	52.0%	6.7%
1.0	0.79	0.82	47.8%	4.2%

2.1. Mathematical Model on 2D Imaging Plane. We focus on the dominant vortex flow appearing in the A3CH view, which passes through the apex and the mitral and aortic valves as shown in Figure 1(b), and mathematical model for blood flows inside the LV on the imaging plane of the A3CH view.

Let D be a 2D imaging domain and $\Omega(t)$ the cross-section of the LV region in the A3CH view so that they satisfy $\Omega(t) \subseteq D \subseteq \mathbb{R}^2$. Color flow data are practically measured

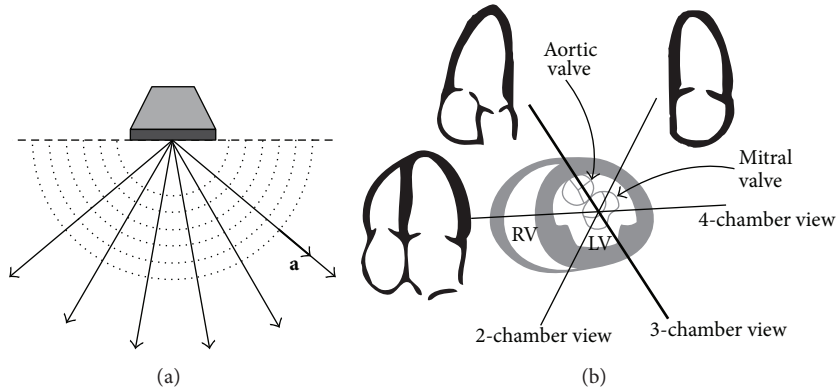


FIGURE 1: A sector scanning probe and apical long-axis views. The vector **a** in (a) means ultrasound beam propagation directions in sector-type scanning for cardiac application. (b) shows three (4-chamber, 3-chamber, and 2-chamber) apical imaging views for the cardiac scanning. They have the relationship of rotation by approximately 60 degrees toward each other. The imaging plane of the apical long-axis 3-chamber view passes through both of mitral and aortic valves.

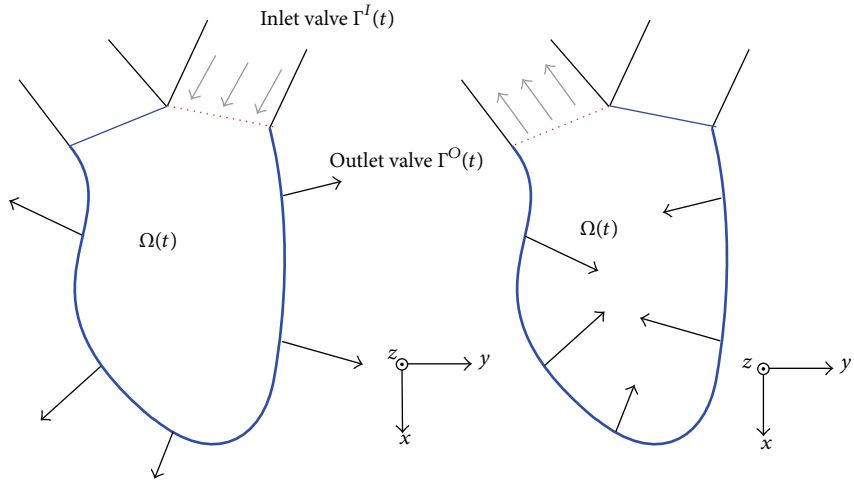


FIGURE 2: Description on the LV region $\Omega(t)$ in the A3CH view. They show the diastolic and systolic motion of $\Omega(t)$, respectively. Here, $\Gamma^I(t)$ and $\Gamma^O(t)$ refer to the mitral and aortic valves (as inlet and outlet valves), respectively.

in the imaging plane D. Figure 1(a) describes the scanline directional vectors $\mathbf{a} := (a_1, a_2)$ for the 2D imaging plane D as an example of sector scanning. As shown in Figure 2, $\Gamma^I(t)$ and $\Gamma^O(t)$ denote the mitral and aortic valves, respectively. The parameter t denotes the LV region and valves as time-varying boundaries. The superscripts I and O of $\Gamma^I(t)$ and $\Gamma^O(t)$ stand for inlet and outlet valves, respectively. The 3D coordinate system takes the xy-plane to contain $\Omega(t)$ and the z-axis to be normal to this plane. Figure 2 describes a basic LV structure and the diastolic/systolic motion of its wall observed in the A3CH view.

Let $c(\mathbf{x}, t)$ be the measured color flow data and $(u(\mathbf{x}, t), v(\mathbf{x}, t))$ the velocity fields of flow at the position $\mathbf{x} \in D$ and time t, respectively. Then the color flow data $c(\mathbf{x}, t)$ can be expressed as the inner product of the scanline vector and the velocity vector:

$$c(\mathbf{x}, t) = \left(a_1(\mathbf{x}), a_2(\mathbf{x})\right) \cdot \left(u(\mathbf{x}, t), v(\mathbf{x}, t)\right). \quad (1)$$

To recover (u, v) from knowledge of c on the imaging plane D, we propose the following 2D Navier-Stokes model:

$$\begin{aligned} \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{\mu}{\rho}\nabla^2 u + \frac{\mu}{3\rho^2}\frac{\partial s}{\partial x}, \\ \frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial y} + \frac{\mu}{\rho}\nabla^2 v + \frac{\mu}{3\rho^2}\frac{\partial s}{\partial y}, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= \frac{s}{\rho}, \end{aligned} \quad (2)$$

where $\rho = 1050\,\mathrm{kg/m^3}$ and $\mu = 0.00316\,\mathrm{Pa{\cdot}s}$ are the density and viscosity of the blood flow, respectively [14]. This model (2) is equivalent to a 2D incompressible flow having a source-sink distribution $s(\mathbf{x}, t)$ [15]. In fact, the 3D blood flows appear or disappear in the imaging plane. Therefore, we reconstruct the 2D velocity fields in the imaging plane by solving (1) and

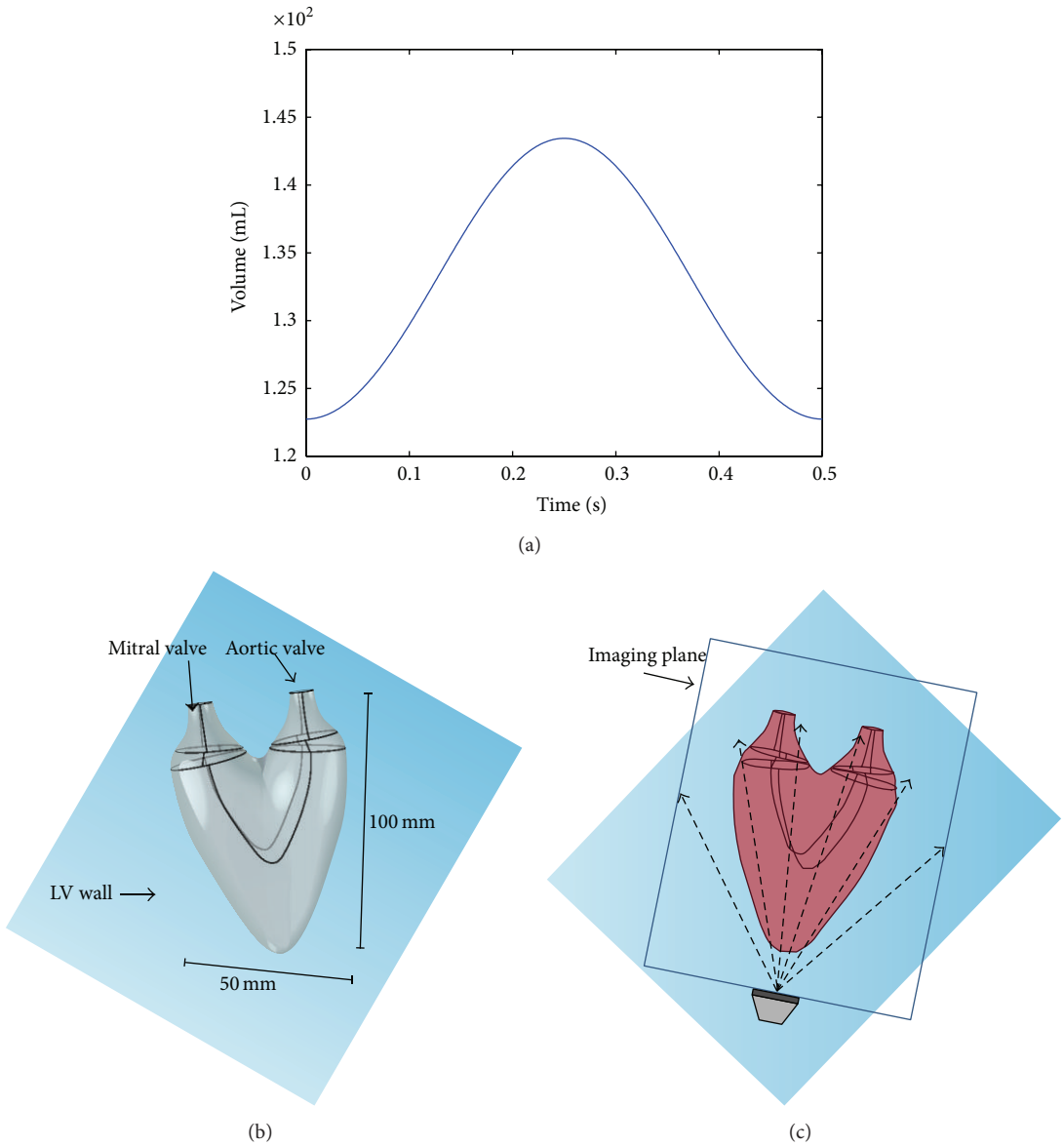


FIGURE 3: LV model for numerical experiments. (a) LV volume curve according to diastolic and systolic motions, (b) 3D LV shape model at the end systole, and (c) 2D imaging plane of A3CH view positioned in the constructed 3D LV model. The dashed arrows means the scanning lines transmitted from a linear array transducer for 2D imaging.

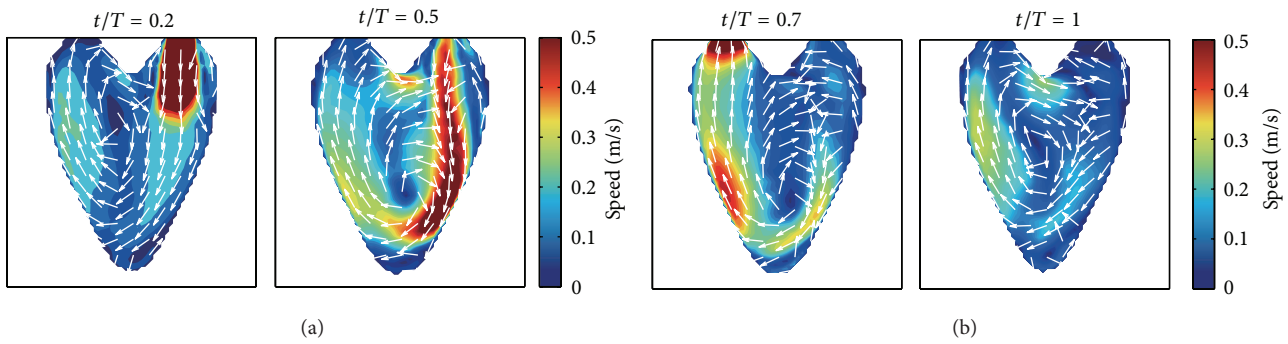


FIGURE 4: Synthetic intraventricular velocity fields projected on the imaging plane. (a) and (b) show the velocity fields for the diastole and systole process of LV, respectively.

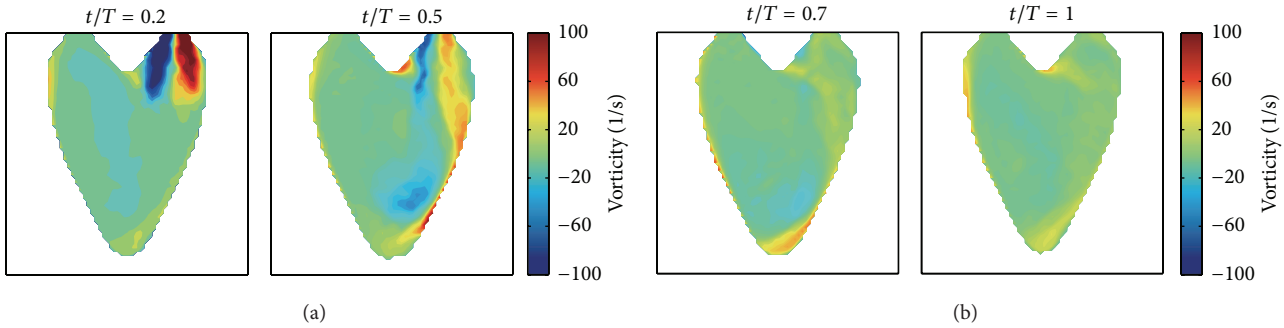


Figure 5: Vorticities corresponding to the synthetic 2D intraventricular velocity fields. (a) and (b) show the vorticity changes according to the diastole and systole process of LV, respectively.

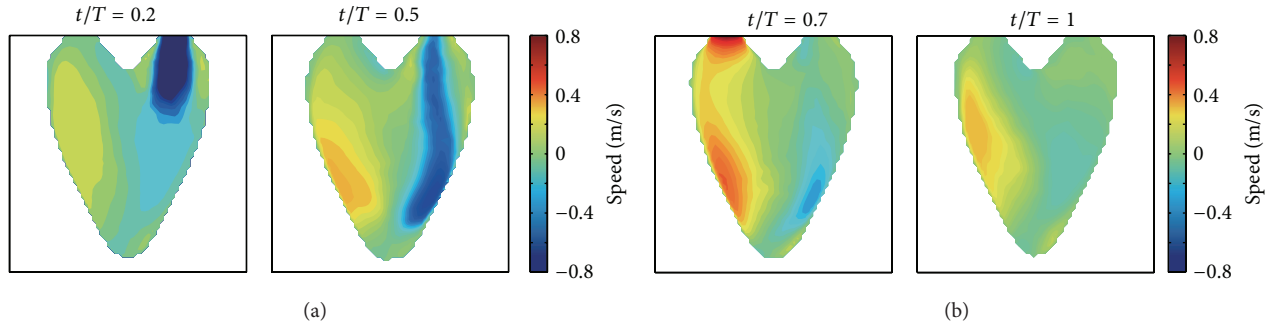


Figure 6: One-directional velocity component data. (a) and (b) show the scanline directional components of the generated velocity fields during the diastole and systole process of LV, respectively.

(2) with LV boundary conditions. We impose the following boundary conditions:

$$\begin{aligned} u &= u_{\text{wall}} \quad \text{on } \partial\Omega(t), \\ v &= v_{\text{wall}} \quad \text{on } \partial\Omega(t), \\ \frac{\partial p}{\partial \mathbf{n}} &= 0 \quad \text{on } \partial\Omega(t), \\ s &= 0 \quad \text{on } \partial\Omega(t), \end{aligned} \tag{3}$$

where u_{wall} and v_{wall} are velocity components computed by the motion of $\partial\Omega(t)$ and $\mathbf{n}$ is the unit outward normal vector to $\partial\Omega(t)$. Here, $\partial\Omega(t)$ is the boundary of $\Omega(t)$, and its motion can be extracted from echocardiography data as described in the above subsection. Further mathematical explanation of (1) and (2) is given in the Appendix A.

2.2. Reconstruction Algorithm. For numerical implementation, we write the system of (1) and (2) as the following linear second-order system:

$$\begin{bmatrix} a_1 & a_2 & 0 & 0 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & 0 & -\frac{1}{\rho} \\ \frac{\partial}{\partial t} - \frac{\mu}{\rho}\nabla^2 & 0 & \frac{1}{\rho}\frac{\partial}{\partial x} & -\frac{\mu}{3\rho^2}\frac{\partial}{\partial x} \\ 0 & \frac{\partial}{\partial t} - \frac{\mu}{\rho}\nabla^2 & \frac{1}{\rho}\frac{\partial}{\partial y} & -\frac{\mu}{3\rho^2}\frac{\partial}{\partial y} \end{bmatrix} \begin{bmatrix} u \\ v \\ p \\ s \end{bmatrix} = \begin{bmatrix} c \\ 0 \\ -u\frac{\partial u}{\partial x} - v\frac{\partial u}{\partial y} \\ -u\frac{\partial v}{\partial x} - v\frac{\partial v}{\partial y} \end{bmatrix}. \tag{4}$$

We discretize the 2D imaging region Ω into the mesh grid elements with $M \times N$ nodes and apply the standard finite

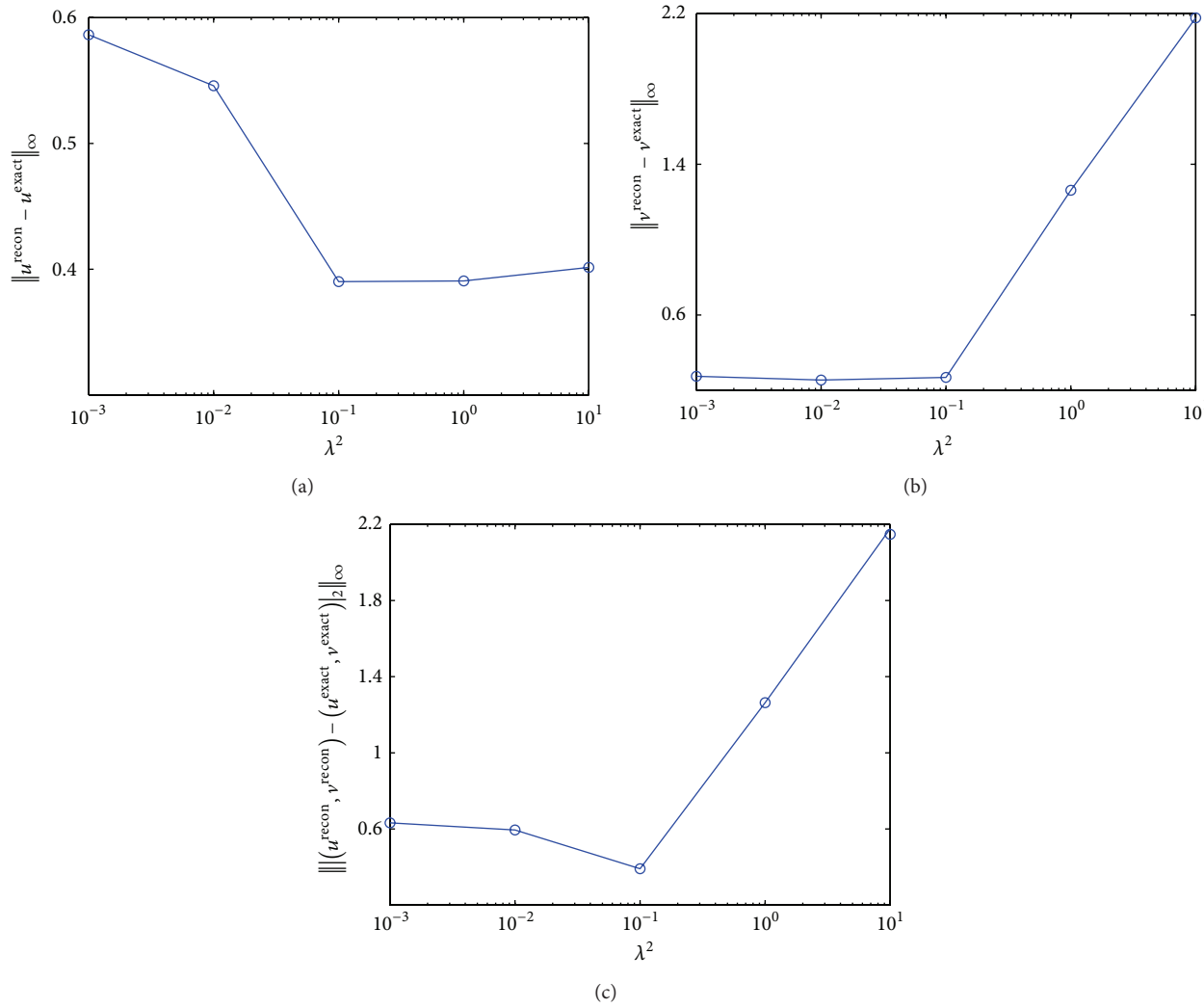


FIGURE 7: Error change depending on the parameter λ. (a) L_∞-norm errors between u^{recon} and u^{exact}, (b) between v^{recon} and v^{exact}, and (c) between $(u^{\text{recon}}, v^{\text{recon}})$ and $(u^{\text{exact}}, v^{\text{exact}})$ with respect to λ^2. Here, $(u^{\text{recon}}, v^{\text{recon}})$ and $(u^{\text{exact}}, v^{\text{exact}})$ stand for the reconstructed and the reference velocity fields, respectively.

difference method for the linear equation (4). Then we obtain the discretized linear system of the following form:

$$\begin{bmatrix} a_1 & a_2 & 0 & 0 \\ D_x & D_y & 0 & -\frac{1}{\rho} \\ 1 - \frac{\mu\Delta t}{\rho}L & 0 & \frac{\Delta t}{\rho}D_x & -\frac{\mu\Delta t}{3\rho^2}D_x \\ 0 & 1 - \frac{\mu\Delta t}{\rho}L & \frac{\Delta t}{\rho}D_y & -\frac{\mu\Delta t}{3\rho^2}D_y \end{bmatrix} \begin{bmatrix} u^{(n+1)} \\ v^{(n+1)} \\ p^{(n+1)} \\ s^{(n+1)} \end{bmatrix} \tag{5}$$

$$= \begin{bmatrix} c^{(n+1)} \\ 0 \\ u^{(n)} - \Delta t\left(u^{(n)}D_x u^{(n)} + v^{(n)}D_y u^{(n)}\right) \\ v^{(n)} - \Delta t\left(u^{(n)}D_x v^{(n)} + v^{(n)}D_y v^{(n)}\right) \end{bmatrix},$$

where D_x, D_y, and L are the x-derivative, y-derivative, and Laplace operators of finite difference, respectively; the superscript (n) denotes the nth time-step. Note that the measured data c on the right-hand side are the values at the $(n+1)$th step, not the (n)th step.

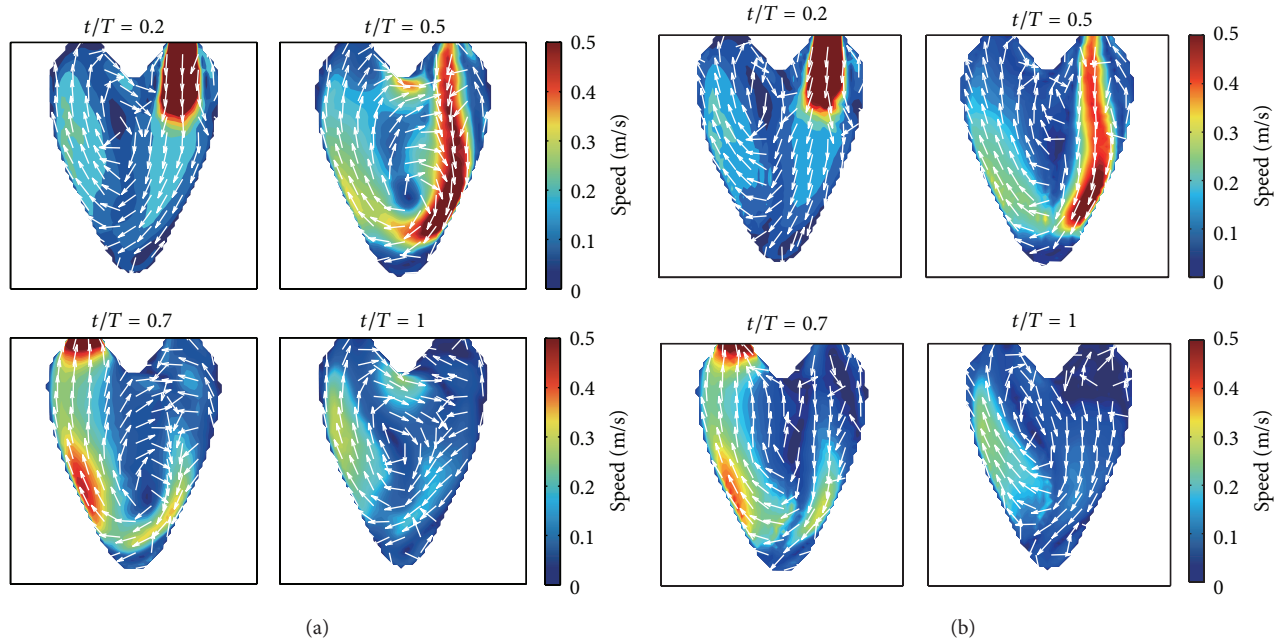


Figure 8: Reconstruction results. The reference vector field (a) and the reconstructed 2D vector fields by the proposed method (b).

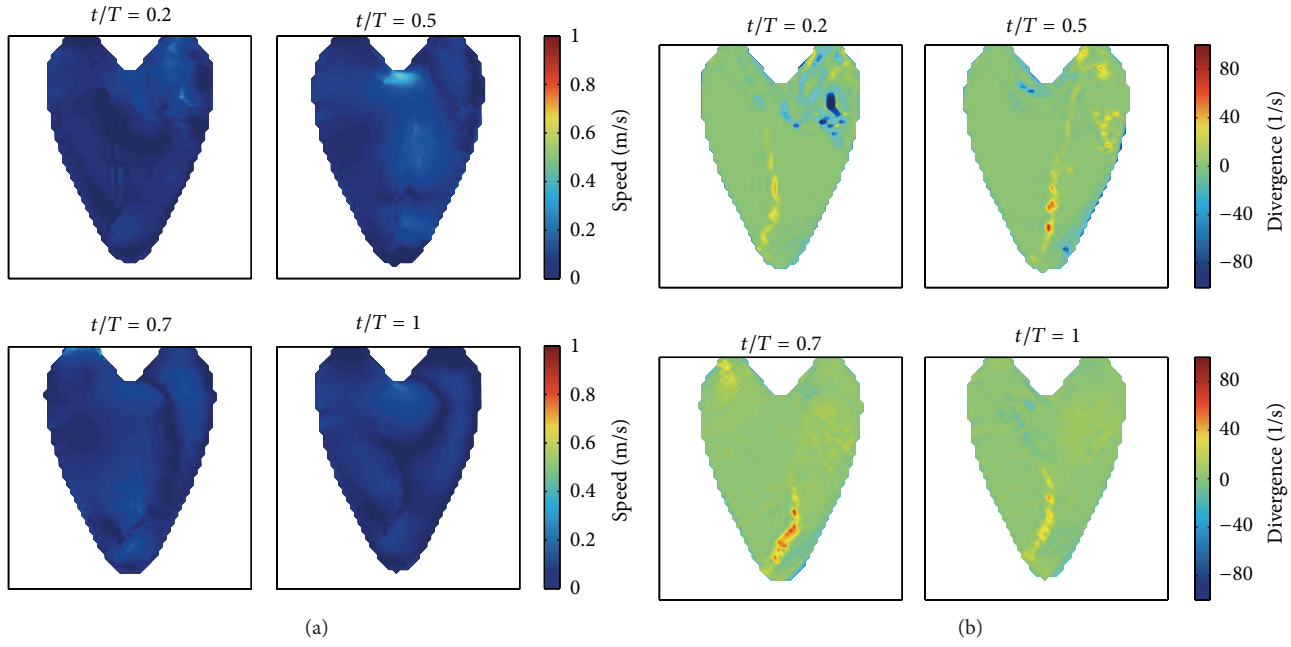


Figure 9: Errors of the reconstruction results. (a) 2D plots for error-magnitude of reconstructed (u, v) relative to the reference velocity fields and (b) the mass source term s.

Let $\mathbf{I}$ and $\mathbf{O}$ be the $N \times N$ identity and zero matrices, respectively. Also set $\mathbf{D}$ and $\mathbf{L}$ to

$$\mathbf{D} = \begin{bmatrix} 0 & \frac{1}{2h} & & \\ -\frac{1}{2h} & \ddots & \ddots & \\ & \ddots & \ddots & \frac{1}{2h} \\ & & -\frac{1}{2h} & 0 \end{bmatrix}, \tag{6}$$

$$\mathbf{L} = \frac{1}{h^2} \begin{bmatrix} -4 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -4 \end{bmatrix},$$

and let $\mathbb{A}_1$, $\mathbb{A}_2$, $\mathbb{D}_x$, $\mathbb{D}_y$, and $\mathbb{L}$ be the $(M \times N) \times (M \times N)$ matrices defined by

$$\mathbb{A}_1 = \operatorname{diag}\left(a_{1(1,1)}, \ldots, a_{1(N,M)}\right),$$

$$\mathbb{A}_2 = \operatorname{diag}\left(a_{2(1,1)}, \ldots, a_{2(N,M)}\right),$$

$$\mathbb{D}_x = \begin{bmatrix} \mathbf{D} & & \\ & \ddots & \\ & & \mathbf{D} \end{bmatrix},$$

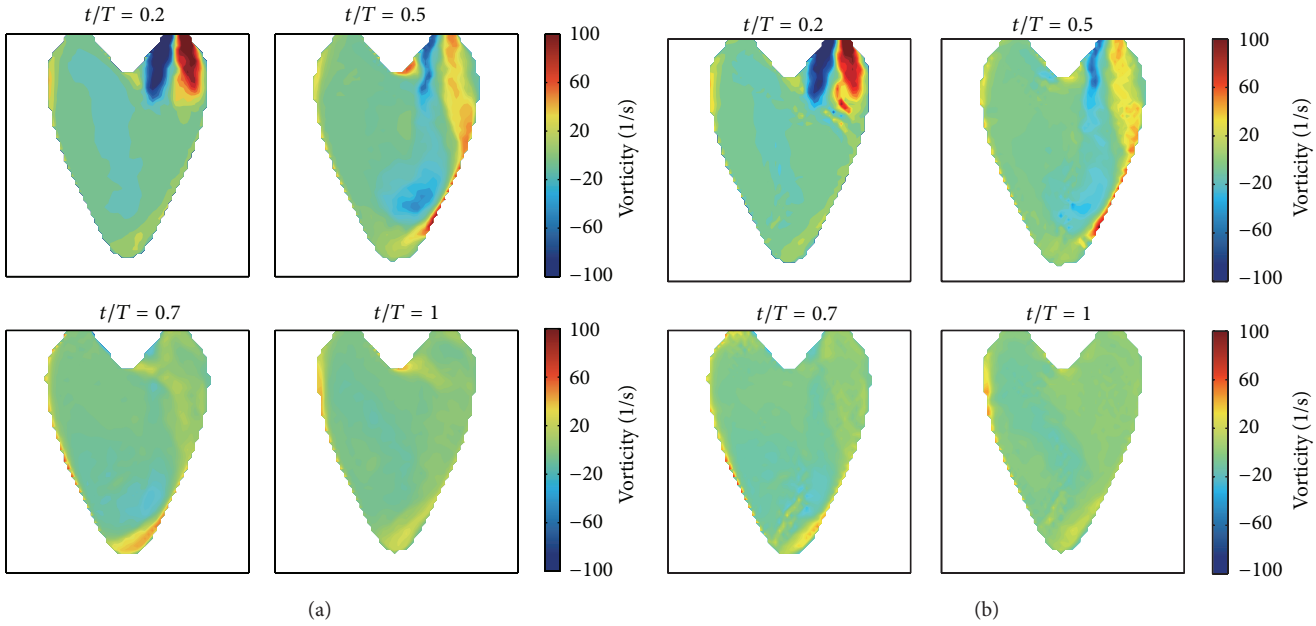


Figure 10: Reconstruction results of vorticities. These were obtained by taking the 2D curl operator to the reconstructed velocity fields.

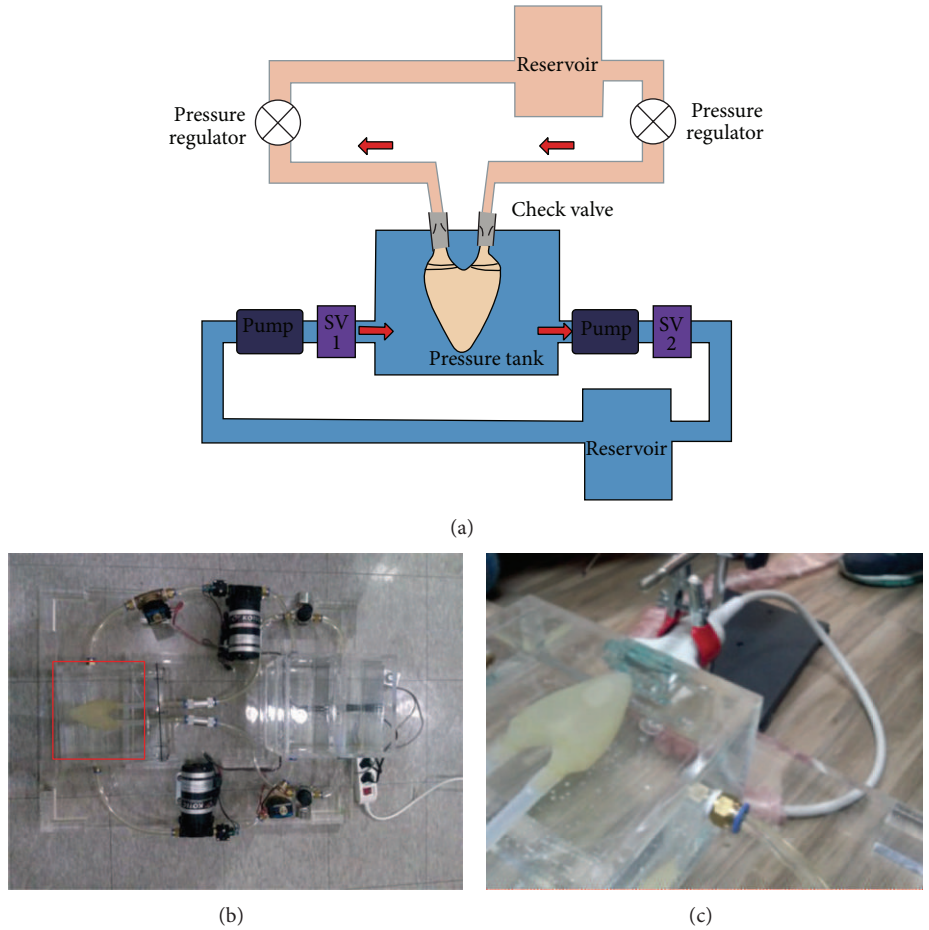


Figure 11: LV phantom operating system. The LV phantom is made of polyurethane. We control the fluid motion by the synchronous operation of two pumps.

$$\mathbb{D}_y = \begin{bmatrix} \mathbf{O} & -\frac{1}{2h}\mathbf{I} & & \\ \frac{1}{2h}\mathbf{I} & \ddots & \ddots & \\ & \ddots & \ddots & -\frac{1}{2h}\mathbf{I} \\ & & \frac{1}{2h}\mathbf{I} & \mathbf{O} \end{bmatrix},$$

$$\mathbb{L} = \begin{bmatrix} \mathbf{L} & \frac{1}{h^2}\mathbf{I} & & \\ \frac{1}{h^2}\mathbf{I} & \ddots & \ddots & \\ & \ddots & \ddots & \frac{1}{h^2}\mathbf{I} \\ & & \frac{1}{h^2}\mathbf{I} & \mathbf{L} \end{bmatrix}. \tag{7}$$

The above discretized linear system (5) can then be written in the form of $\mathbb{K}\mathbf{U} = \mathbf{F}$, where

$$\mathbb{K} = \begin{bmatrix} \mathbb{A}_1 & \mathbb{A}_2 & \mathbb{O} & \mathbb{O} \\ \mathbb{D}_x & \mathbb{D}_y & \mathbb{O} & -\frac{1}{\rho}\mathbb{I} \\ \mathbb{I} - \frac{\mu\Delta t}{\rho}\mathbb{L} & \mathbb{O} & \frac{\Delta t}{\rho}\mathbb{D}_x & -\frac{\mu\Delta t}{3\rho^2}\mathbb{D}_x \\ \mathbb{O} & \mathbb{I} - \frac{\mu\Delta t}{\rho}\mathbb{L} & \frac{\Delta t}{\rho}\mathbb{D}_y & -\frac{\mu\Delta t}{3\rho^2}\mathbb{D}_y \end{bmatrix}, \tag{8}$$

$$\mathbf{U} = \begin{bmatrix} \mathbf{u}^{(n+1)} \\ \mathbf{v}^{(n+1)} \\ \mathbf{p}^{(n+1)} \\ \mathbf{s}^{(n+1)} \end{bmatrix}, \qquad \mathbf{F} = \begin{bmatrix} \mathbf{c}^{(n+1)} \\ \mathbf{0} \\ \mathbf{k}_1^{(n)} \\ \mathbf{k}_2^{(n)} \end{bmatrix}.$$

For notational simplicity, we drop the superscripts (n) and $(n+1)$. Then, in the right-hand side, $\mathbf{u}$, $\mathbf{v}$, $\mathbf{p}$, $\mathbf{s}$, $\mathbf{c}$, $\mathbf{k}_1$, and $\mathbf{k}_2$ are the $(M \times N) \times 1$ vectors defined as follows:

$$\mathbf{u} = \begin{bmatrix} u_{(1,1)} \\ \vdots \\ u_{(N,M)} \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_{(1,1)} \\ \vdots \\ v_{(N,M)} \end{bmatrix}, \quad \mathbf{p} = \begin{bmatrix} p_{(1,1)} \\ \vdots \\ p_{(N,M)} \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} s_{(1,1)} \\ \vdots \\ s_{(N,M)} \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_{(1,1)} \\ \vdots \\ c_{(N,M)} \end{bmatrix},$$

$$\mathbf{k_1} = \begin{bmatrix} u_{(1,1)} - \Delta t\left(u_{(1,1)}\frac{u_{(2,1)} - u_{(0,1)}}{2h} + v_{(1,1)}\frac{u_{(1,2)} - u_{(1,0)}}{2h}\right) \\ \vdots \\ u_{(N,M)} - \Delta t\left(u_{(N,M)}\frac{u_{(N+1,M)} - u_{(N-1,M)}}{2h} + v_{(N,M)}\frac{u_{(N,M+1)} - u_{(N,M-1)}}{2h}\right) \end{bmatrix},$$

$$\mathbf{k_2} = \begin{bmatrix} v_{(1,1)} - \Delta t\left(u_{(1,1)}\frac{v_{(2,1)} - v_{(0,1)}}{2h} + v_{(1,1)}\frac{v_{(1,2)} - v_{(1,0)}}{2h}\right) \\ \vdots \\ v_{(N,M)} - \Delta t\left(u_{(N,M)}\frac{v_{(N+1,M)} - v_{(N-1,M)}}{2h} + v_{(N,M)}\frac{v_{(N,M+1)} - v_{(N,M-1)}}{2h}\right) \end{bmatrix}. \tag{9}$$

However, the coefficient matrix in (8) is nonsymmetric, sparse, and large scale. Moreover, it has a bad condition number and is almost singular. Hence, we use the least-squares approach with Tikhonov regularization to solve the minimization problem with a regularization term of the form:

$$\widehat{\mathbf{U}} = \underset{\mathbf{u}}{\operatorname{argmin}} \|\mathbb{K}\mathbf{U} - \mathbf{F}\|^2 + \|\Gamma\mathbf{U}\|^2. \tag{10}$$

By setting $\Gamma = \lambda\mathbb{I}_{(4\times N\times M)\times(4\times N\times M)}$, the minimization problem is written as

$$\widehat{\mathbf{U}} = \underset{\mathbf{u}}{\operatorname{argmin}} \|\mathbb{K}\mathbf{U} - \mathbf{F}\|^2 + \lambda^2 \|\mathbf{U}\|^2, \tag{11}$$

and its solution is given by

$$\widehat{\mathbf{U}} = \left(\mathbb{K}^T\mathbb{K} + \lambda^2\mathbb{I}\right)^{-1}\mathbb{K}^T\mathbf{F}. \tag{12}$$

We perform the reconstruction algorithm using the heuristically selected Tikhonov regularization parameter λ shown in the reconstruction algorithm (11).

2.3. Numerical Simulations with Moving LV Boundaries. We obtain synthetic flow data inside a virtual moving simplified LV wall. To obtain the synthetic intraventricular flows, we construct a 3D moving LV region of the stroke volume of about 20 mL. As shown in Figure 3(a), the time-dependent LV volume ranges from 123 mL to 143 mL during the entire

cycle of 0.5 s. Figure 3(b) shows a LV shape model corresponding to the volume of 133 mL at the preset pressure state. We then perform a numerical simulation of the forward problem of the Navier-Stokes equation inside the 3D moving LV for the beat cycle. For the forward simulations, the "fluid-structure interaction model" of the COMSOL software is used, and 3D intraventricular velocity fields are computed as a solution of the forward problem. A no-slip boundary condition is used for computing the velocity fields. We assume that the blood flow during the filling of the LV is ejected in the normal direction to the inlet valve surface $\Gamma^I(t)$, that the flow velocity is uniform on the entire inlet boundary, and that the LV volume change is equal to the total amount of its net inflow:

$$\frac{\partial V(\Omega(t))}{\partial t} = \int_{\Gamma^I(t)} \mathbf{n} \cdot \mathbf{u} dS, \tag{13}$$

where $\mathbf{n}$ is the normal vector to the mitral valve (inlet valve) boundary denoted by $\Gamma^I(t)$. The boundary portion containing the mitral and aortic valves was fixed not to be moved. Neumann boundary conditions of velocity fields were given by setting the zero normal derivatives at the outlet Γ^O and inlet Γ^I valves for the LV diastole and systole, respectively. For simplicity, pressure is assumed to be a constant at the inlet, while the Neumann condition for the pressure is used at the outlet. (The work of [16, 17] will also help readers create synthetic flow data.)

We project the 3D synthetic velocity fields on the imaging plane of A3CH view, which is set to be located in the constructed 3D LV model (see Figure 3(c)). The projected 2D velocity fields are used as reference data to evaluate the proposed 2D velocity reconstruction algorithm. Representative cases of these synthetic 2D flows are depicted in Figure 4, which show the time-varying dominant vortex patterns inside the virtual moving LV wall. While the LV model is relaxing, flows enter into the LV through the inlet valve. The main stream impinges against the LV wall, and a large vortex is simultaneously formed near the center. The large vortex moves upward during the contraction process and weakens at the end of the systole cycle.

Figure 5 shows the vorticity fields obtained by taking the curl operator to the synthetic flow fields. At the early stage of relaxation, small vortices are formed on the both sides of the main incoming stream. The higher vorticity near the wall is related to the fact that the high momentum incoming flow results in increasing the wall shear stress when the incoming flow impinges to the wall; we think. While the vorticity gradually grows during the expanding process, it eventually shrinks and weakens during the process of contraction.

The inner product of the synthetic 2D velocity fields and the scanline directional unit vectors gives the scanline directional velocity components, which can be regarded as the color flow data of the intraventricular flow on the A3CH view measured by a real ultrasound system. The one-directional velocity component data are used as the input data for our proposed method. Figure 6 illustrates the scanline directional velocity components. The figures in the second column corresponding to the end-diastole and end-systole reflect the overall similar flow patterns that are comparable to real color patterns (see Figures 12(c) and 12(d) in Appendix B).

2.4. Choice of the Parameter $\boldsymbol{\lambda}$. In this study, all the experiments were performed by setting the parameter $\lambda = \sqrt{0.1}$. To optimize the value of this parameter, we investigated the errors of the reconstructed velocity fields with respect to the change of λ^2. We adjusted the value of λ^2 to each power of 10 from 10^{-3} to 10^1. Figure 7 shows that the L_∞-norm errors between the reconstructed velocity fields and the reference data have formed a convex plot with the minimized L_∞-norm error attained at $\lambda^2 = 0.1$.

3. Results

We demonstrated the feasibility of the proposed method by showing the L_2-norm errors between the synthetic flow and the reconstructed flow fields. We reconstructed the velocity fields from the synthetic scanline directional velocity components by repeatedly solving the minimization problem (11) five times for the entire beat cycle. Figure 8 shows the reconstruction results of velocity fields obtained from the input data of the one-directional velocity components. As in the forward numerical simulations, the incoming stream impinged to the lower right wall. Overall, the reconstructed flows showed very similar patterns to the reference flows shown in Figure 4. However, the main vortex near the center was shifted upward relative to the reference vector field. This shifting phenomenon was also continued in the remaining steps.

Figure 9 illustrates the performance of the proposed model. Part (a) compares the pointwise error magnitude images of the reconstructed velocity fields with the reference data. While the reconstructed u-component errors were distributed in the range of −0.39~0.25 m/s, the errors of the v-components showed larger differences of −0.24~0.27 m/s for the entire cycle. Figure 9(b) shows the distribution of the mass source term s and reflects the out-of-plane components of the flows. The ratio s/ρ ranged from $-151.1\,\mathrm{s}^{-1}$ to $82.5\,\mathrm{s}^{-1}$. From these contributions of s, the averaged pointwise errors for u- and v-components were 0.06 m/s and 0.02 m/s, respectively. The averaged pointwise error-magnitude was 0.065 m/s. We also compared the reconstruction results between the 2D Navier-Stokes models with and without the mass source term s. When s was used, the largest pointwise errors for the u- and v-components were 0.39 m/s and 0.27 m/s, respectively; reconstruction results not considering s showed larger pointwise errors of 0.71 m/s and 0.78 m/s for these components, respectively, for the entire cycle. These results support that the proposed model improves reconstruction performance.

For further quantitative comparison, we computed L_2-norm errors of each velocity field on the given whole region at each 1/10 time step of the whole cycle. Table 1 shows the relatively large velocity error (≈47%) for the end of the contraction process compared with the expanding process. While the pointwise errors of velocity may be larger than

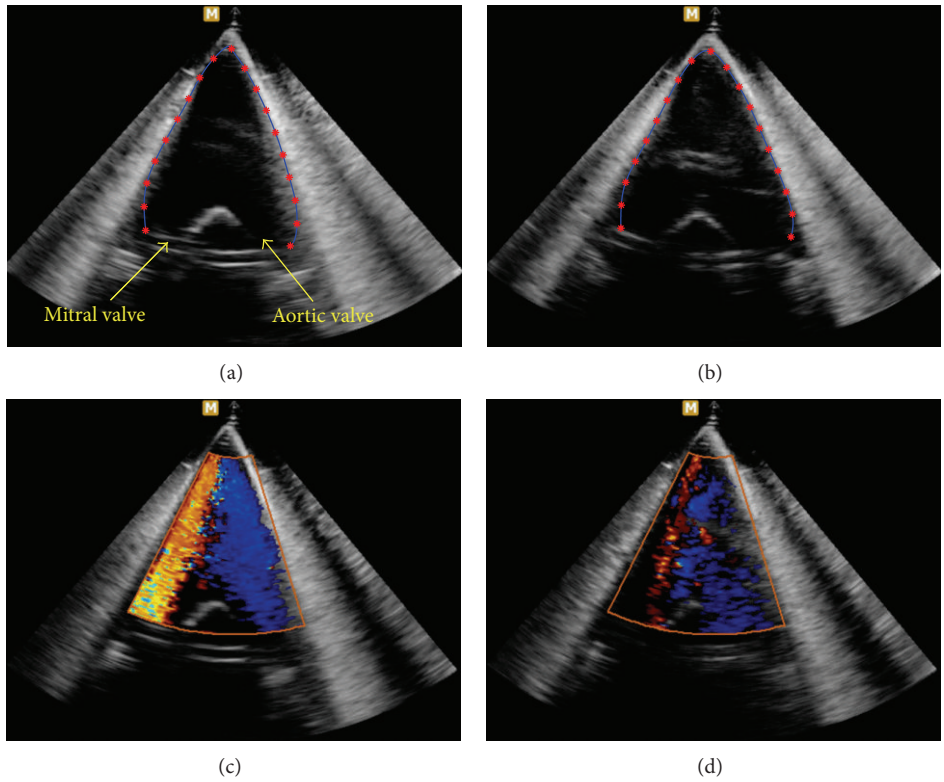


FIGURE 12: Real data acquisition using LV phantom and border extraction. (a) and (b) are echo images in ventricular shape at end-diastole and end-systole phases, respectively. (c) and (d) are color flow images corresponding to (a) and (b), respectively.

its L_2-norm errors at some local regions, those pointwise errors at the local regions did not affect the major features of the vortex flows. This implies that the proposed model is reasonably accurate for estimating the unsteady features of the blood flows when the reconstruction algorithm of flow is performed repeatedly over several cycles of diastole and systole processes.

Although the proposed reconstruction method produced nonsmooth distributions of vorticity inside the LV region (as shown in Figure 10), the evolutions of the large vortical structure inside the LV region were clearly observed, and the vorticities overall showed very similar patterns to the reference results described in Figure 5. The pointwise and global errors for the vortex fields are listed in Table 2. These nonsmooth vorticity distributions may be induced by uncertainties of the boundary geometry and the regularization term in the minimization problem (11).

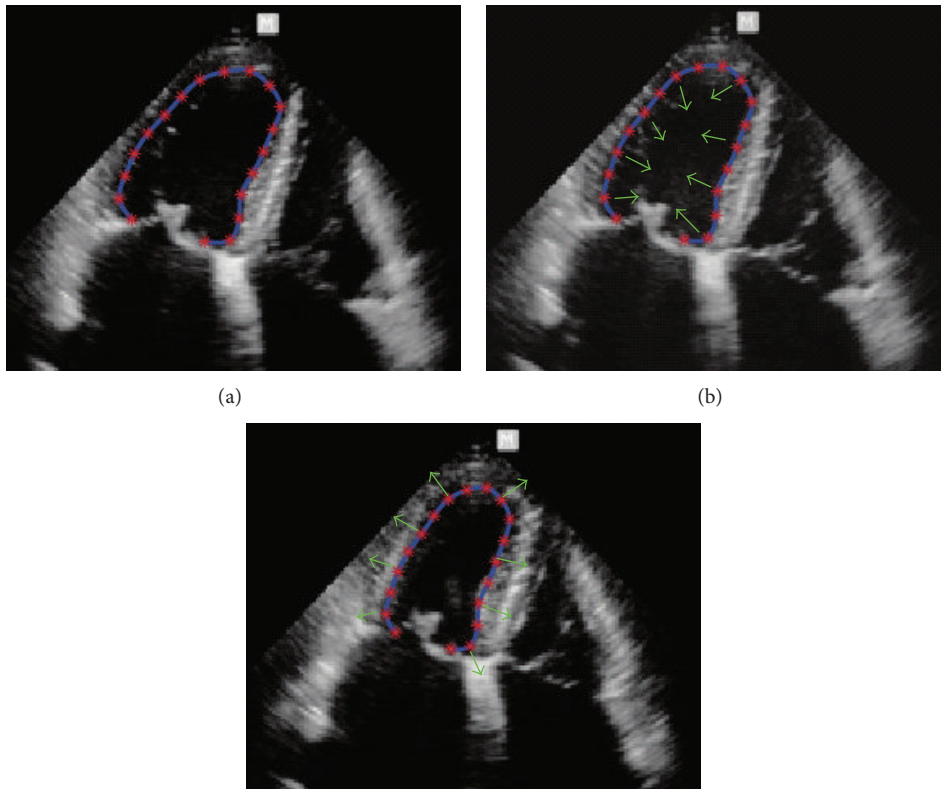
4. Discussion and Conclusions

We propose here a new 2D Navier-Stokes model that reconstructs intraventricular flows using color flow data and LV boundaries extracted from echocardiography data. The proposed model considers both in-plane and out-of-plane blood flows on the imaging plane of an apical long-axis three-chamber view. The out-of-plane components moving out of the imaging plane were modeled as the mass source term of a source-sink distribution. We reconstructed blood flows in the imaging domain by solving a system of equations, which includes the 2D incompressible Navier-Stokes equation of the mass source term and a color flow measurement equation describing the one-directional velocity component of the color flow data. The boundary conditions that are required to solve the system of equations are given by the LV borders extracted from echocardiography data. To evaluate the proposed method, numerical experiments were performed on synthetic flow data following a virtual LV motion. The results showed that the proposed method is feasible and potentially valuable for reconstructing intracardiac flow fields.

The numerical experiments used an imaging domain of size 64×64 pixels, because the computational costs of the proposed reconstruction method were large. We divided the whole cycle into 1000 time-steps for stable numerical implementation and repeatedly performed the proposed algorithm five times for the entire heartbeat cycle. We observed that the solutions obtained after two repeats were very similar to each other. A detailed mathematical description of the convergence of the solution will be given in our next work.

All the experiments were conducted using MATLAB 7.10.0, and the computational time for each step was about 12 s using an Intel i7-4702MQ Quadcore CPU running at 2.0 GHz with 8 GB of RAM. The speed of the proposed method needs to be improved for its practical application, and a study to reduce the processing time is under way.

To reconstruct 2D flow patterns using the color flow images requires an experimental ultrasound system that



(a) (b)

(c)

Figure 13: Velocities of wall motion using LV border tracking method. Initialized LV contour at end-diastole frame (a), velocity fields of contracting and expanding LV wall (b-c). Blue solid line, red dots, and green arrows are the LV contours $\mathscr{C}(t)$, contour points $\{\mathbf{x}_1(t), \ldots, \mathbf{x}_n(t)\}$, and some of velocity vectors $\{\mathbf{v}_1(t), \ldots, \mathbf{v}_n(t)\}$, respectively.

can acquire echo and color data independently; however, we had only a commercial ultrasound system that cannot support data acquisition for the experiments. The frame-rate limitation of the color flow data acquired by the ultrasound imaging system should also be overcome. The color flow images from the ultrasound imaging system are obtained by applying the autocorrelation algorithm [13] after repeatedly transmitting and receiving the ultrasound beam along the same scanline; this then causes the low frame-rate of color flow data acquisition. The proposed method assumed that the time difference between the sequential color flow images is very small and that each color value reflects the exact velocity components in the direction of ultrasound beam propagation. For the proposed method to be practically applicable, the frame-rate needs to be improved. The development of an ultrafast imaging system is ongoing. Such a system is expected to have a frame-rate higher than 1000 fps using plane wave beam-forming; it would acquire echo and color flow image data separately and support the performance of the proposed method in real application. Some recent works by other researchers have examined ultrafast flow imaging [18–20]. A review of these references will help us to perform our future work on an ultrafast imaging system. In Appendix B, we describe an *in vitro* phantom experimental setup for empirical data (color flow ultrasound and LV boundary data) to demonstrate the feasibility of our reconstruction method.

To advance the proposed algorithm, we are studying on mitral valve tracking. Based on reports [21, 22] that the motion of the mitral valve affects the vorticities of the intraventricular flows, studies on the boundary conditions containing the valve motion are under way.

This work is the first that describes overall our proposed method. More detailed mathematical analysis and further validation tests may be required to verify its scientific validity and practical applicability. We will perform several follow-up works to achieve this.

Appendices

A. Mathematical Formulation

Based on the 3D modeling of intraventricular blood flows, we derive a 2D mathematical model describing the relationship between blood flows and color flow data measured on the imaging plane.

A.1. Mathematical Model and Inverse Problem. Let D be a 3D imaging domain, $\Omega(t)$ a time-varying LV region satisfying $\Omega(t) \subseteq D \subseteq \mathbb{R}^3$, and T a beat cycle. For the beat cycle T, we consider a spatial-temporal domain Ω_T defined by $\Omega_T := \bigcup_{0<t<T} \Omega(t) \times \{t\} \subseteq D \times (0, T)$. Let $\mathbf{v}(\mathbf{r}, t)$ be a velocity field

of the blood flow within the spatial-temporal domain Ω_T, $\mathbf{a}(\mathbf{r}) = (a_1(\mathbf{r}), a_2(\mathbf{r}), a_3(\mathbf{r}))$ scanline directional unit vectors at the position $\mathbf{r} \in D$ and $c(\mathbf{r}, t)$ color flow data in Ω_T. Given $c(\mathbf{r}, t)$, we then consider an inverse problem to find a 3D vector field $\mathbf{v} = (u, v, w)$ satisfying the following condition:

$$\mathbf{a}(\mathbf{r}) \cdot \mathbf{v}(\mathbf{r}, t) = c(\mathbf{r}, t), \tag{A.1}$$

where $\cdot$ is the inner product operator.

To solve this inverse problem, we should deal with the 3D Navier-Stokes equations governing the blood flow $\mathbf{v}$:

$$\begin{aligned} \rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v}\right) &= -\nabla p + \mu \nabla^2 \mathbf{v} \quad \text{in } \Omega_T, \\ \nabla \cdot \mathbf{v} &= 0 \quad \text{in } \Omega_T. \end{aligned} \tag{A.2}$$

However, given that our aim is to reconstruct the blood flow using 2D measurements, we reduce (A.1) and (A.2) to a 2D inverse problem and the corresponding 2D Navier-Stokes equation. We compute the velocity fields of blood flows by using the scanline directional projected velocity components represented as 2D color flow data in a commonly used ultrasound 2D imaging scanner and then solving the system of equations related to them.

A.2. Mathematical Model on a 2D Imaging Plane. From now on, $\Omega(t)$ denotes the cross-section of the LV region in the A3CH view. (For simplicity, the same notation $\Omega(t)$ in 3D will be used to represent its cross-section.) The notations of $\Gamma^I(t)$ and $\Gamma^O(t)$ are same as shown in Figure 2. Let D be a 2D imaging domain satisfying $\Omega(t) \subseteq D \subseteq \mathbb{R}^2$ so that it is equal to the cross-sectioned region of the 3D imaging domain by the A3CH view. As described earlier, we try to model flows in the 2D imaging plane D, where color flow data are practically measured. In this case, the scanline direction $\mathbf{a}(\mathbf{r})$ is given to be tangent to the plane D, $a_3 = 0$ on the plane. Let $c(\mathbf{x}, t)$ be the measured color flow data at the position $\mathbf{x} \in D$ at time t. The color flow data $c(\mathbf{x}, t)$ can then be expressed as the inner product of the scanline vector and the velocity vector:

$$c(\mathbf{x}, t) = \left(a_1(\mathbf{x}), a_2(\mathbf{x})\right) \cdot \left(u(\mathbf{x}, t), v(\mathbf{x}, t)\right). \tag{A.3}$$

Hence, the corresponding inverse problem is to recover (u, v) from knowledge of c on the imaging plane D. Also, (A.2) can be rewritten as

$$\begin{aligned} \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{\mu}{\rho}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) \\ &\quad + \underbrace{\frac{\mu}{\rho}\frac{\partial^2 u}{\partial z^2} - w\frac{\partial u}{\partial z}}_{f_1}, \\ \frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial y} + \frac{\mu}{\rho}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) \\ &\quad + \underbrace{\frac{\mu}{\rho}\frac{\partial^2 v}{\partial z^2} - w\frac{\partial v}{\partial z}}_{f_2}, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= -\frac{\partial w}{\partial z}. \end{aligned} \tag{A.4}$$

Note that color flow data may not contain measurable information on the 2D plane D for $f_1 = (\mu/\rho)(\partial^2 u/\partial z^2) - w(\partial u/\partial z)$, $f_2 = (\mu/\rho)(\partial^2 v/\partial z^2) - w(\partial v/\partial z)$, and $\partial w/\partial z$. We want to keep the incompressible condition of the 3D problem. However, the third term $\partial w/\partial z$ of the divergence of 3D flow is neither of the measurable or computable quantities in the 2D plane. Here, we model it as a mass source-sink term s, which represents the linear deformation of the fluid in the out-of-plane direction. This new unknown term will be determined by solving the modified Navier-Stokes equation with the color flow measurement data. Introducing a new variable s, we obtain the following reduced 2D Navier-Stokes model:

$$\begin{aligned} \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{\mu}{\rho}\nabla^2 u + \frac{\mu}{3\rho^2}\frac{\partial s}{\partial x}, \\ \frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} &= -\frac{1}{\rho}\frac{\partial p}{\partial y} + \frac{\mu}{\rho}\nabla^2 v + \frac{\mu}{3\rho^2}\frac{\partial s}{\partial y}, \\ \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} &= \frac{s}{\rho}. \end{aligned} \tag{A.5}$$

This model is equivalent to a 2D incompressible flow with a source-sink distribution $s(\mathbf{x}, t)$ [15]. In fact, if we do not consider the external force, equation (3.3.13) in [15] implies

$$\begin{aligned} \rho\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y}\right) &= -\frac{\partial p}{\partial x} + \mu\nabla^2 u \\ &\quad + \frac{\mu}{3\rho}\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right), \\ \rho\left(\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y}\right) &= -\frac{\partial p}{\partial y} + \mu\nabla^2 v \\ &\quad + \frac{\mu}{3\rho}\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right). \end{aligned} \tag{A.6}$$

B. *In Vitro* Phantom Experiments

B.1. Experimental Setup. The experimental setup for operating the LV phantom is depicted in Figure 11. The setup was composed of main two parts: one for making the LV phantom beat periodically and the circulatory part (resp., the lower and upper parts in Figure 11(a)). The fluidic motion was controlled by the synchronous operation of two pumps, two solenoid valves (SVs), and the check valves employed to model the aortic and mitral valves. The phantom, which was constructed of polyurethane, was immersed in a water tank. The phantom was 3D printed as a half-ellipsoid [23, 24]. The experimental circulatory system including the LV phantom was completely filled with water. Figures 11(b) and 11(c) show the manufactured LV phantom with its design drawing and the experimental setting for scanning the LV phantom, respectively. For LV systolic motion, the left solenoid valve (SV1) is set to open, while the right solenoid valve (SV2) is closed. The pressure inside the tank then increases, which subsequently exerts pressure onto the LV phantom. When the pressure inside the LV phantom is greater than the preset pressure of the check valve, the fluid in the LV phantom surges into the circulatory system. The check valves restrain backward flow, making the fluid in the LV phantom pass through the aortic valve to circulate the system. Dilation of the LV phantom then follows by closing SV1 and opening SV2, allowing similar complementary processes to occur. A simple timer switch controls the opening and closing of the solenoid valves to mimic LV beating. The volume of the LV phantom under the preset pressure was measured to be 133 mL, while the stroke volume was about 20 mL during LV beating with 0.5 s period. The polyurethane is flexible and inelastic; the LV phantom therefore may retain a constant "endocardial" surface area, while the contained volume may vary owing to changes in its cross-sectional shape. Despite those limitations, vortex flow pattern inside the phantom is generated during the "diastole process." The B-mode and C-mode images were scanned using an Accuvix V10 ultrasound system (Samsung Medison, Seoul, South Korea) and a probe P2-4BA with the frequency band of 2~4 MHz. The probe was positioned just below the water surface, at a depth of approximately 10 cm, to enable the LV phantom to be positioned in the center of the ultrasound image.

Representative ultrasound echo and color flow images acquired by scanning the LV phantom are shown in the first and second rows of Figure 12, respectively. By applying an LV tracking algorithm to the acquired ultrasound echo images, we obtained LV boundaries for the whole cycle. (The LV tracking method is explained in the next subsection.) The red dotted and blue lines in Figures 12(a) and 12(b) present the LV boundaries extracted in the end-diastole and end-systole images. The LV boundaries are extracted at each frame (50 fps) during the cardiac cycle in order to impose moving boundary conditions for the numerical simulations of the next section. Figures 12(c) and 12(d) show that the color flow patterns may indicate two different vortex flows. Here, the red and blue colors represent the velocity components of the flows coming toward and receding from the ultrasound probe, respectively. The color flow pattern in the diastolic phase is clearly stronger than that in the systolic phase, as has been reported for human LV [1].

B.2. LV Wall Segmentation from B-Mode Image. We acquire ultrasound echo images for the entire cardiac cycle and from them extract the LV borders, typically by myocardial motion tracking. The myocardial motion tracking method is combined with the Lucas-Kanade method and a constraint formulated by the global deformation of nonrigid heart motion proposed in [25]. As illustrated in Figure 13, we denote the endocardial border traced at initially selected end-diastole frame by a parametric contour $\mathscr{C}^* = \{\mathbf{x}^*(s) = (x^*(s), y^*(s)) \mid 0 \le s \le 1\}$ that can be identified as its n tracking points $\mathbf{x}_1^* = \mathbf{r}^*(s_1), \ldots, \mathbf{x}_n^* = \mathbf{x}^*(s_n)$. Here, $0 = s_1 < s_2 < \cdots < s_n = 1$. Let $\mathscr{C}(t) = \{\mathbf{x}(s,t) = (x(s,t), y(s,t)) \mid 0 \le s \le 1\}$ be the contour deformed from $\mathscr{C}(0) = \mathscr{C}^*$ at time t. The velocity field $\mathbf{V}(t)$ of motion of the contour $\mathscr{C}(t)$ is determined by a time change of tracking points $\{\mathbf{x}_1(t), \ldots, \mathbf{x}_n(t)\}$:

$$\mathbf{V}(t) := \begin{bmatrix} \mathbf{v}_1(t) \\ \vdots \\ \mathbf{v}_n(t) \end{bmatrix} = \frac{d}{dt} \begin{bmatrix} \mathbf{x}_1(t) \\ \vdots \\ \mathbf{x}_n(t) \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} \mathbf{x}_1(0) \\ \vdots \\ \mathbf{x}_n(0) \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^* \\ \vdots \\ \mathbf{x}_n^* \end{bmatrix}. \tag{B.1}$$

Here, we identify the contour $\mathscr{C}(t)$ with tracking points $\{\mathbf{x}_1(t), \ldots, \mathbf{x}_n(t)\}$.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Institute for Mathematical Sciences (NIMS) and the National Research Foundation of Korea (NRF) grants funded by the Korea government (nos. NIMS-A21402, NRF-2012R1A1A2038700, and 2012R1A2A1A03670512 (2011-0028868)).

References

[1] G.-R. Hong, G. Pedrizzetti, G. Tonti et al., "Vannan, characterization and quantification of vortex flow in the human left ventricle by contrast echocardiography using vector particle image velocimetry," *JACC: Cardiovascular Imaging*, vol. 1, no. 6, pp. 705–717, 2008.

[2] S. Cimino, G. Pedrizzetti, G. Tonti et al., "In vivo analysis of intraventricular fluid dynamics in healthy hearts," *European Journal of Mechanics—B/Fluids*, vol. 35, pp. 40–46, 2012.

[3] K. Lampropoulos, W. Budts, A. van de Bruaene, E. Troost, and J. P. van Melle, "Visualization of the intracavitary blood flow in systemic ventricles of Fontan patients by contrast echocardiography using particle image velocimetry," *Cardiovascular Ultrasound*, vol. 10, no. 1, article 5, 2012.

[4] H. Gao, P. Claus, M.-S. Amzulescu, I. Stankovic, J. D'Hooge, and J.-U. Voigt, "How to optimize intracardiac blood flow tracking by echocardiographic particle image velocimetry? Exploring the influence of data acquisition using computer-generated data sets," *European Heart Journal Cardiovascular Imaging*, vol. 13, no. 6, pp. 490–499, 2012.

[5] H. Gao, B. Heyde, and J. D'Hooge, "3D Intra-cardiac flow estimation using speckle tracking: a feasibility study in synthetic ultrasound data," in *Proceedings of the IEEE International Ultrasonics Symposium (IUS '13)*, pp. 68–71, Prague, Czech Republic, July 2013.

[6] D. H. Evans and W. N. McDicken, *Doppler Ultrasound: Physics, Instrumentation, and Signal Processing*, John Wiley & Sons, Chichester, UK, 2000.

[7] D. Garcia, J. C. Del Á.lamo, D. Tanné. et al., "Two-dimensional intraventricular flow mapping by digital processing conventional color-doppler echocardiography images," *IEEE Transactions on Medical Imaging*, vol. 29, no. 10, pp. 1701–1713, 2010.

[8] S. Ohtsuki and M. Tanaka, "The flow velocity distribution from the Doppler information on a plane in three-Dimensional flow," *Journal of Visualization*, vol. 9, no. 1, pp. 69–82, 2006.

[9] M. Arigovindan, M. Sühling, C. Jansen, P. Hunziker, and M. Unser, "Full motion and flow field recovery from echo Doppler data," *IEEE Transactions on Medical Imaging*, vol. 26, no. 1, pp. 31–45, 2007.

[10] A. Gomez, K. Pushparajah, J. M. Simpson, D. Giese, T. Schaeffter, and G. Penney, "A sensitivity analysis on 3D velocity reconstruction from multiple registered echo Doppler views," *Medical Image Analysis*, vol. 17, no. 6, pp. 616–631, 2013.

[11] F. Mehregan, F. Tournoux, S. Muth et al., "Doppler vortography: a color doppler approach to quantification of intraventricular blood flow vortices," *Ultrasound in Medicine and Biology*, vol. 40, no. 1, pp. 210–221, 2014.

[12] A. Tura, H. Gao, T. Gaens, and C. Lamberti, "Recularization of blood motion fields by modified navier-stokes equations," *European Heart Journal—Cardiovascular Imaging*, vol. 21, pp. 27–36, 1999.

[13] C. Kasai, K. Namekawa, A. Koyano, and R. Omoto, "Real-time twodimensional blood flow imaging using an autocorrelation technique," *IEEE Transactions on Sonics and Ultrasonics*, vol. 32, no. 3, pp. 458–464, 1985.

[14] Y. Cheng, H. Oertel, and T. Schenkel, "Fluid-structure coupled CFD simulation of the left ventricular flow during filling phase," *Annals of Biomedical Engineering*, vol. 33, no. 5, pp. 567–576, 2005.

[15] G. K. Batchelor, *An Introduction to Fluid Dynamics*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, UK, 2000.

[16] A. Swillens, L. Løvstakken, J. Kips, H. Torp, and P. Segers, "Ultrasound simulation of complex flow velocity fields based on computational fluid dynamics," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, no. 3, pp. 546–556, 2009.

[17] S. Balocco, O. Basset, J. Azencot, P. Tortoli, and C. Cachard, "3D dynamic model of healthy and pathologic arteries for ultrasound technique evaluation," *Medical Physics*, vol. 35, no. 12, pp. 5440–5450, 2008.

[18] J. Bercoff, G. Montaldo, T. Loupas et al., "Ultrafast compound doppler imaging: providing full blood flow characterization," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 58, no. 1, pp. 134–147, 2011.

[19] B.-F. Osmanski, M. Pernot, G. Montaldo, A. Bel, E. Messas, and M. Tanter, "Ultrafast doppler imaging of blood flow dynamics in the myocardium," *IEEE Transactions on Medical Imaging*, vol. 31, no. 8, pp. 1661–1668, 2012.

[20] B. Y. S. Yiu and A. C. H. Yu, "High-frame-rate ultrasound color-encoded speckle imaging of complex flow dynamics," *Ultrasound in Medicine & Biology*, vol. 39, no. 6, pp. 1015–1025, 2013.

[21] G. Pedrizzetti, F. Domenichini, and G. Tonti, "On the left ventricular vortex reversal after mitral valve replacement," *Annals of Biomedical Engineering*, vol. 38, no. 3, pp. 769–773, 2010.

[22] T. Schenkel, M. Malve, M. Reik, M. Markl, B. Jung, and H. Oertel, "MRI-Based CFD analysis of flow in a human left ventricle: methodology and application to a healthy heart," *Annals of Biomedical Engineering*, vol. 37, no. 3, pp. 503–515, 2009.

[23] T. Plappert and M. Sutton, *The Echocardiographers' Guide*, Informa Healthcare, London, UK, 2006.

[24] A. Kheradvar and M. Gharib, "On mitral valve dynamics and its connection to early diastolic flow," *Annals of Biomedical Engineering*, vol. 37, no. 1, pp. 1–13, 2009.

[25] C. Y. Ahn, "Robust myocardial motion tracking for echocardiography: variational framework integrating local-to-global deformation," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 974027, 14 pages, 2013.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 610482, 7 pages
http://dx.doi.org/10.1155/2015/610482

Research Article

Automatically Identifying Fusion Events between GLUT4 Storage Vesicles and the Plasma Membrane in TIRF Microscopy Image Sequences

Jian Wu,[1,2,3] Yingke Xu,[1,2,3] Zhouyan Feng,[1,3] and Xiaoxiang Zheng[1,2,3,4]

[1] *Department of Biomedical Engineering, College of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou 310027, China*
[2] *Zhejiang Provincial Key Laboratory of Cardio-Cerebral Vascular Detection Technology and Medicinal Effectiveness Appraisal, Hangzhou 310027, China*
[3] *Key Laboratory for Biomedical Engineering of Ministry of Education, Hangzhou 310027, China*
[4] *Qiushi Academy for Advanced Studies, Zhejiang University, Hangzhou 310027, China*

Correspondence should be addressed to Xiaoxiang Zheng; zxx@mail.bme.zju.edu.cn

Received 31 October 2014; Revised 2 January 2015; Accepted 3 January 2015

Academic Editor: Yi Gao

Quantitative analysis of the dynamic behavior about membrane-bound secretory vesicles has proven to be important in biological research. This paper proposes a novel approach to automatically identify the elusive fusion events between VAMP2-pHluorin labeled GLUT4 storage vesicles (GSVs) and the plasma membrane. The differentiation is implemented to detect the initiation of fusion events by modified forward subtraction of consecutive frames in the TIRFM image sequence. Spatially connected pixels in difference images brighter than a specified adaptive threshold are grouped into a distinct fusion spot. The vesicles are located at the intensity-weighted centroid of their fusion spots. To reveal the true *in vivo* nature of a fusion event, 2D Gaussian fitting for the fusion spot is used to derive the intensity-weighted centroid and the spot size during the fusion process. The fusion event and its termination can be determined according to the change of spot size. The method is evaluated on real experiment data with ground truth annotated by expert cell biologists. The evaluation results show that it can achieve relatively high accuracy comparing favorably to the manual analysis, yet at a small fraction of time.

1. Introduction

Accurate regulation of insulin is essential for the maintenance of glucose homeostasis in human body. As a member of the protein family of glucose transporters (GLUTs), glucose transporter type 4 (GLUT4) proteins are preliminarily stored within intracellular membrane bound secretory vesicles inside adipose tissues and striated muscle (skeletal and cardiac), also known as GLUT4 storage vesicles (GSVs). Defects in the activity of this protein have been implicated in some forms of insulin resistance and type II diabetes mellitus. When an insulin receptor on cell surface is activated, insulin induces a rapid increase in the uptake of glucose by inducing the translocation of GSVs from intracellular compartments to the plasma membrane. It has long been essential for membrane trafficking to exactly and quantitatively decipher the dynamic behavior of membrane bound secretory vesicles. However, traditional methods from molecular biology and biochemistry are unable to resolve discrete steps of vesicle movement fundamentally [1]. Total Internal Reflection Fluorescence Microscope (TIRFM) can observe layers as thin as 100 nm of a specimen adjacent to the coverslip, making it a widely used tool for observing biological activities near the cell surface, such as endocytosis and exocytosis. Much more quantitative information can be extracted to support biological research through analyzing TIRFM image data. However, it is still a standard practice for most biologists to manually analyze high throughput images generated from

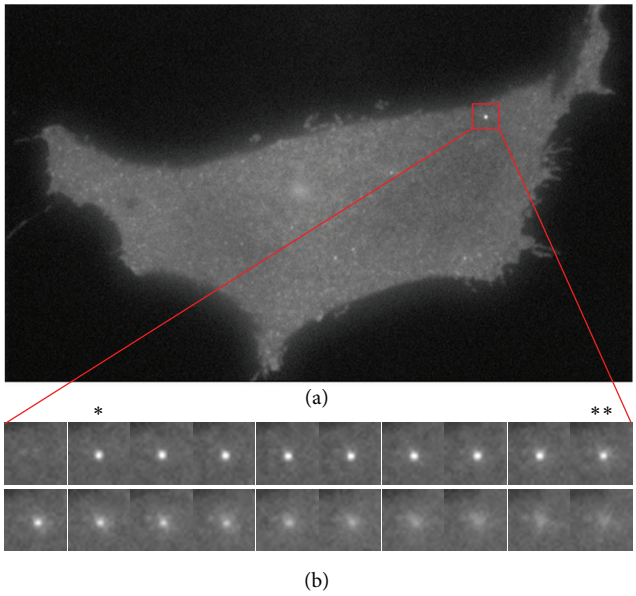


Figure 1: Consecutive time frames from 11 to 30 (b) show that a prominent fusion event corresponds to the patch of interest in a TIRFM image sequence (a). ∗ indicates the fusion pore opening, that is, the initiation of a fusion event. ∗∗ indicates the initiation of a diffusion process. Here, transition time is 1.6 s (8 frames, sampling rate is 5 frames/s).

in vivo observation and visually observe vesicle behaviors. This work is not only time consuming but is also error-prone and nonreproducible, which always induces subjective biases. It is a great need for developing an effective TIRFM image analysis system in biomedical research, which is a novel area in bioimaging, also a subsidiary branch of computing-based image processing [2].

A fusion event of GSVs comprises final steps of an exocytosis behavior, which includes the processes of fusion pore opening and vesicle diffusion. As the GSVs dock to the plasma membrane, a transient and moderate increase of fluorescence can be observed by TIRFM once the fusion pore of a GSV opens. The vesicles halt and vibrate at the same place for a period (named transition time) and then diffuse away from the fusion site visualized as a fluorescence puff to the cell surface or a small explosion at the cell membrane. GLUT4 is then inserted and becomes the integral membrane (transmembrane) protein. Glucose can be transported into the cell down its concentration gradient in a process called facilitated diffusion. The diffusion process of a fusion vesicle comprises a rapid decrease in fluorescence intensity at the fusion site, a widening of vesicle size and a spreading of signal intensity [3], which is the hallmark for identifying fusion events. A prominent fusion event which comprises fusion pore opening and diffusion process is depicted in Figure 1. While some nonfusion vesicles do not diffuse at the plasma membrane after fusion pore opening, they undock or leave the cell surface and return back into the cell at last.

Little has been done towards the identification of fusion events between GSVs and the cell membrane in TIRFM image sequences. Some of the current existing methods are not fully automated [3–5]. Image processing techniques are usually used to detect the positions of GSVs and sort them out from each frame in an image sequence. Corresponding positions for the same vesicle can be linked to a trajectory of vesicle movement. Before identifying the fusion events, the termination of GSVs trajectories (named death events) should first be located. Subsequently, each single-vesicle trajectory is screened for a possible fusion event primarily based on rules, which derived from quantitative characterization of manually identified fusion process. In Vallotton et al.'s [6], a fully automated system was designed for fusion events detection based on vesicle tracking and rigid template matching. On the basis of Vallotton's study, Mele et al. proposed an improved one, where each fusion candidate was described by a set of novel domain specific descriptors. Similarity scores between genuine fusion events (prototype events, set manually by an expert) and fusion candidates were calculated in the Principle Component Analysis (PCA) eigenspace for fusion identification [2]. The identification of fusion events, which takes the multiple vesicles tracking into account, presents a significant hurdle to surmount. A trajectory may end when a vesicle simply undocks from the plasma membrane, when two vesicles fuse together or when the trajectory is built erroneously, usually resulting in inaccurate location of death events. A strong dependence on the standard fusion template makes finding a fitted correlation kernel for different types of fusion events a nigh impossible task. Based on a usually unsatisfied assumption that a vesicle remains stationary for around N frames before it fuses or undocks, a novel approach is proposed in [7]. This method detects fusion and undocking events by first searching for docked vesicles that "appear" to and "disappear" from the field of view, then uses a diffusion model to classify them as either fusion or undocking events.

2. Method Outline

In this paper, a fully automated fusion events identification system is proposed, which comprises detecting fusion pore opening and fusion site of fusion vesicles, characterizing the fusion process by Gaussian fitting, and finally identifying the fusion events according to size change of fusion spots. To precisely detect fusion pore opening and fusion site, GSVs are labeled with VAMP2-pHluorin, which is a pH-sensitive reporter. The pH sensitivity of pHluorin has been exploited to visualize the fusion pore opening [1]. When vesicles dock to the cell surface and fusion pore opens, VAMP2-pHluorin is expressed as a transmembrane protein. The sudden rise in fluorescence can be observed due to the different pH value between the inside and outside of the cell. According to this apparent change in fluorescence, we employ the moving average differentiation instead of absolute differentiation between two consecutive frames to identify the initiation of fusion events. To derive the fusion site, an adaptive threshold more commonly known as the Mean Absolute Deviation (MAD) is applied to reduce noise saturated points caused by small variations and other artifacts in difference images. This

threshold has proven to be useful when compared to a biologist's visual identification in this experiment. The assumption given in this paper is that no other fusions occur at the same place where a fusion process already exists. Because vesicles do not exhibit much movement upon docking to the cell surface, a square patch of image sequence with each candidate vesicle in its center is cropped for further analysis. Two-dimensional (2D) Gaussian models are used to derive the size of fusion spots during the fusion process. Depending on 2D Gaussian fitting, the total intensity of a fusion spot, intensity at fusion site (intensity weighted centroid of a fusion spot) can also be calculated. The method is evaluated on real data with ground truth annotated by biologists. Evaluation results show that it can achieve relatively high accuracy at a low computation cost.

2.1. Detecting the Fusion Candidate Vesicles. In our experiment, due to the benefit of the pH-sensitive reporter (e.g., VAMP2-pHluorin), GSVs cannot be observed until fusion pore opens. The phenomenon that a transient and abrupt fluorescence increases at the cell surface is a strong indicator that vesicles are ready to fuse with the cell membrane. For these apparent changes in intensity, a forward differentiation framework can be used to detect the fusion pore opening. Considering the inherent noise existing in fluorescent imaging, a forward moving average differentiation is used instead of absolute differentiation. In this paper, each difference image ΔI_{t-1} from $t = 2$ can be achieved by subtracting an iterative background Acc from each I_t,

$$\begin{aligned} &Acc = I_1 \\ &\text{for} \quad t = 2 : N \\ &\qquad Acc = (1-\alpha)\, Acc + \alpha I_t \\ &\qquad \Delta I_{t-1} = I_t - Acc \\ &\text{end,} \end{aligned} \tag{1}$$

where N is the frame number of a TIRFM image sequence; α is a user-defined parameter according to the image qualities. In difference images, fusion candidate vesicles correspond to the region where its intensity is higher than local surroundings. To identify these vesicles, an adaptive threshold for difference images called mean absolute deviation (MAD)

$$th = \text{mean}\left(\text{abs}\left(I_t - \text{mean}\left(I_t\right)\right)\right) \tag{2}$$

is used. It has been proven that MAD can extract regions that represent real candidates, as well as eliminate interferences that come from the subtle appearance of fluorescent spots [8]. Sometimes regions corresponding to a same vesicle can be derived many times according to threshold th in the difference image sequence. To ensure that there are no other fusion events occurring at the site where a vesicle fusion process already exists, the first difference image of fusion candidate vesicles is considered.

2.2. Gaussian Fitting for Fusion Process. After detecting fusion candidate vesicles, further dynamic behaviors can be tracked and deciphered for fusion events identification. According to [2], many cues are used to identify fusion events, such as the coefficient of increased intensity, peak difference, and coefficient of maximum increase. In this paper, the vesicle's total intensity, intensity at weighted centroid, and the fusion size during fusion process are taken into consideration. A fusion event can be described as a transient behavior of intensity and size. During the transition time, there is not much change in the intensity and size of a vesicle. After a short period, the fluorescence of the vesicle rapidly diffuses into the background while the vesicle size increases in a fusion event.

For each of the previously detected fusion candidate vesicles, a patch image sequence centered on a weighted centroid with the extent of $(2w+1)\times(2w+1)$ is spatially and temporally cropped from the original TIRFM image sequence. The user defined parameter w is an integer larger than a single vesicle's radius and smaller than the nearest distance between vesicles. The extent of an image patch should include the whole vesicle spot and ensure enough space for fluorescence diffusion. Rather than an extensive analysis of the entire image sequence, it can heavily reduce the computation cost.

According to the Point Spread Function (PSF) of microscope systems, vesicles appear as symmetric and round spots in images. The intensity distribution of a vesicle is well approximated by a 2D isotropic Gaussian function using a simplex algorithm with a Least Mean Square Errors (LMSE) estimator [9]. The equation of a 2D Gaussian surface is of the general form

$$G_w\left(x, y\right) = I_c \exp\left[-\frac{\left(x - x_c\right)^2 + \left(y - y_c\right)^2}{2{\sigma_{xy}}^2}\right], \tag{3}$$

where (x_c, y_c) is the coordinate of a weighted centroid, I_c is the intensity at (x_c, y_c), and the variance σ_{xy} is dependent on the actual radius of a vesicle spot, ranging from 0 to w. The local background of each frame is subtracted and all pixel intensity values are normalized. The local background for a patch image is approximated using a boxcar average over a square region with the extent of $(4w + 1) \times (4w + 1)$ in the original image sequence:

$$B\left(x, y\right) = \frac{1}{\left(4w+1\right)^2} \sum_{i=-2w}^{2w} \sum_{j=-2w}^{2w} A\left(x + i, y + j\right). \tag{4}$$

The Full Width at Half Maximum (FWHM) is considered as the vesicle radius r and divided by σ_{xy} is approximately the constant 1.1774. The termination of vesicle's movement can be defined when the value of σ_{xy} multiplied by 1.1774 is bigger than parameter w, using the result of directly fitting the 2D Gaussian function to each patch image without local background subtraction.

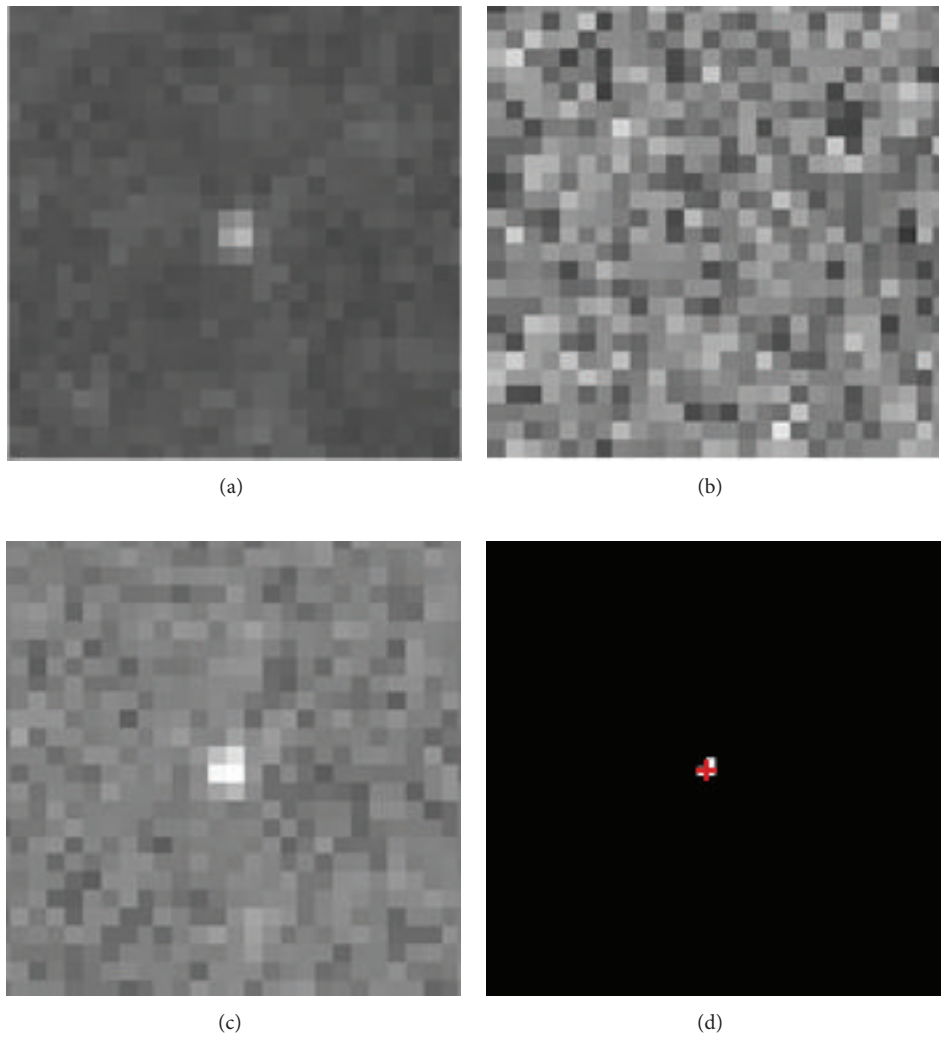


Figure 2: The process of detecting a fusion candidate vesicle. (a) A patch of image with a fusion candidate vesicle in. (b) The absolute differentiation of consecutive images results in a noisy difference image, from which the vesicle is hard to detect. (c) The forward moving average differentiation can achieve a high SNR difference image. (d) A vesicle spot mask derived from (c) with the MAD threshold, which can be used to calculate the intensity-weighted centroid (a red cross).

To eliminate the abrupt interference due to fluorescence turbulence, the integrated average of vesicle radii at each time step is used to quantify the size of the fusion spots. All fluorescence is assumed to be membrane-embedded in either vesicles, internal membrane structures, or the plasma membrane [2]. A vesicle can also be assumed as an isotropic ball with radius r when docking to the cell membrane before diffusion. The surface area of a sphere with radius r is determined using the following formula area $= 4\pi r^2$. When a vesicle fully fuses with the plasma membrane, the membrane-embedded fluorescence is released to the cell membrane. The radius of the final diffusion spot is at least two times bigger than spot radius before diffusion. According to these, during diffusion process a fusion event can be exactly deciphered if vesicle spot radius is two times bigger than its integrated average at diffusion initiation.

3. Experiment and Evaluation Results

In this experiment, VAMP2-pHluorin positive 3T3-L1 adipocytes are imaged using an IX-70 inverted TIRFM microscope (Olympus), which is equipped with both argon (488 nm) and argon/krypton (568 nm) laser lines (Melles Griot), a 60 × 1.45 N.A. oil immersion objective lens (PlanApoN; Olympus), and a TIRFM condenser. The TIRFM images were detected with a sampling frequency of 5 Hz with a back-illuminated Andor iXon887 EMCCD camera (1024 × 1024, pixel spacing 0.18 μm, 16 bits; Andor Technologies).

In this paper, three real TIRFM image sequences, each has 400 consecutive frames, are chosen to evaluate the performance of the proposed method. All image sequences are well annotated by expert cell biologists. During the detection of fusion candidate vesicles, we choose the user-defined

Table 1: Recall and precision results of detecting fusion candidate vesicles.

	Fusion candidate vesicles ground truth (TP + FN)	Fusion candidate vesicles detected (TP + FP)	Mistakenly detected (FP)	Missed (FN)	Recall	Precision
Sequence 1	131	125	2	8	93.9%	98.4%
Sequence 2	144	140	1	5	96.5%	99.3%
Sequence 3	179	173	5	11	93.9%	97.1%
		Average			94.8%	98.3%

Table 2: Results of automatically identifying fusion events.

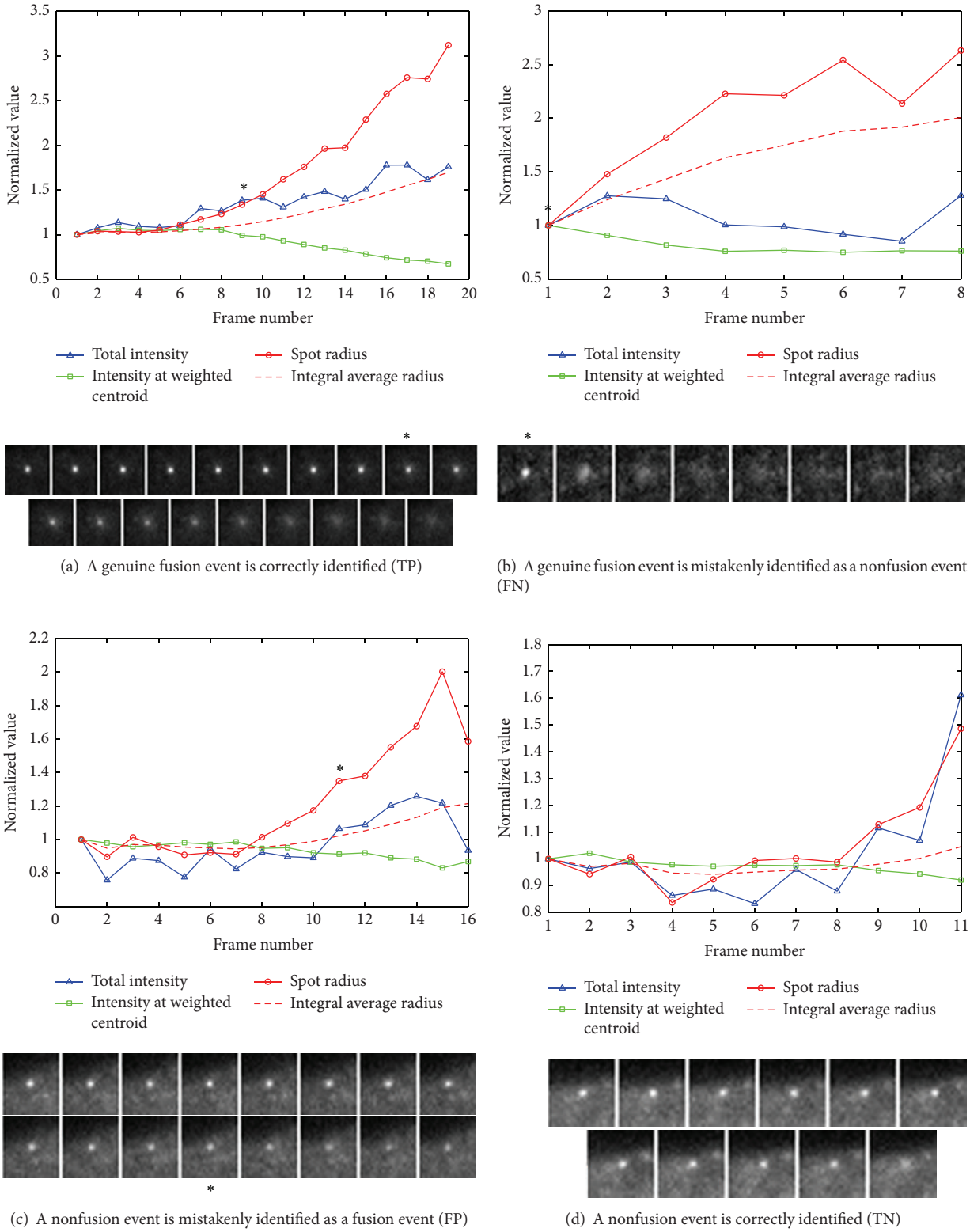
	Fusion events ground truth (TP + FN)	Fusion events detected (TP + FP)	Mistakenly identified (FP)	Missed (FN)	Recall	Precision	False +ve
Sequence 1	47	42	6	11	76.6%	85.7%	7.1%
Sequence 2	62	59	7	10	83.9%	88.1%	8.5%
Sequence 3	89	83	9	15	83.1%	89.2%	10.0%
		Average			81.2%	87.7%	8.5%

parameter $\alpha = 0.2$ for the purpose of removing irrelevant background noise and keeping the vesicle objects. As shown in Figure 2, the forward moving average differentiation can achieve a high signal-to-noise ratio (SNR) difference image comparing to the absolute differentiation method. Without initiation, the adaptive threshold MAD can heavily suppress uneven local background and extract vesicle spot masks from difference images. The intensity-weighted centroid of a fusion vesicle can be calculated using the mask imposed on the original image. The evaluation results of detecting fusion candidate vesicles in three image data sets are shown in Table 1. The detection method combined with the forward moving average differentiation and an adaptive threshold MAD can achieve a relatively high accuracy with both recall and precision up to 90%. Due to the strong noisy image background, some subtle vesicles are missed. While some artifacts suddenly appeared in cell images are mistakenly detected as vesicles through image differentiation step.

According to the image data, the parameter $w = 10$ is taken, because the largest radius of fusion spot is not bigger than 10 pixels. A sequence of patch images with extent of 21×21 for each fusion candidate vesicles is cropped. After Gaussian fitting for each frame, the vesicle radius (FWHM of a Gaussian function), intensity at weighted centroid (peak intensity), and total intensity can be derived. The relationship between them during vesicle movement is shown in Figure 3. The integral average radius is also calculated to detect the initiation of the diffusion process, marked with black asterisk. In this paper, the initiation of the diffusion process can be defined at the time when vesicle size is 1.2 times bigger than the integral average size. It means that the size change of a vesicle is not more than 20% within fusion pore duration. The candidate vesicle can be defined as a fusion one if the largest radius of vesicle spot during the diffusion process is 2 times bigger than the integral average size of the vesicle at the initiation of the diffusion. As shown in Figure 3, the assumption in [2, 6] that the integrated intensity over the entire surface is constant is not satisfied due to the consequence of the cells releasing fluorescent materials to the cell surface. The evaluation results for automatically identifying fusion events in three image data sets are shown in Table 2. Four typical identification results comparing to ground truth are shown in Figure 3. Many fusion events mistakenly identified (FP) by this method are at the edge of cell, shown in Figure 3(c). Although local background is reduced, 2D Gaussian fitting cannot output satisfied results due to uneven image background. While the missed one (FN) is due to the rigorous condition, the largest radius of fusion spot during the fusion event is 2 times bigger than the integral average size of the vesicle at the initiation of the diffusion event, shown in Figure 3(b). This method is quite dependent on precise 2D Gaussian fitting. For total image date sets, the evaluation results in Table 2 show that it can achieve relatively high accuracy compared favorably to the manual analysis, yet at a small fraction of time.

4. Conclusion

In this paper, we have proposed a fully automated and easily applied method for identifying fusion events. We have evaluated the method using real image data annotated by biologists. Evaluation results show that the forward moving average differentiation combined with an adaptive threshold MAD is more useful for detecting fusion candidate vesicles combined with a pH-sensitive reporter, while 2D Gaussian fitting of fusion process helps to annotate the fusion event in TIRFM image sequences.



(a) A genuine fusion event is correctly identified (TP)

(b) A genuine fusion event is mistakenly identified as a nonfusion event (FN)

(c) A nonfusion event is mistakenly identified as a fusion event (FP)

(d) A nonfusion event is correctly identified (TN)

Figure 3: Four typical identification results compare to ground truth. Relationship between the variety of total instensity, intnesity at weighted centroid, and spot radius during vesicle movement. The integral average radius is also calculated (red dashed line) as a threshold for identifing diffusion initiation (>1.2 times, black asterisk) and fusion event (>2 times). (c) A nonfusion at the edge of cell is mistakenly identified as a fusion event, mainly due to unsymmetrical background. All values are normalized to the initiation of fusion pore opening.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by grants from the National Basic Research Program of China (2011CB504400), National Natural Science Foundation of China (31301176, 61427818), Zhejiang Provincial Natural Science Foundation of China (LY13C050001), and Specialized Research Fund for the Doctoral Program of Higher Education (20130101120172).

References

[1] Y. Xu, B. R. Rubin, C. M. Orme et al., "Dual-mode of insulin action controls GLUT4 vesicle exocytosis," *The Journal of Cell Biology*, vol. 193, no. 4, pp. 643–653, 2011.

[2] K. Mele, A. Coster, J. G. Burchfield et al., "Automatic identification of fusion events in TIRF microscopy image sequences," in *Proceedings of the IEEE 12th International Conference on Computer Vision Workshops (ICCV '09)*, pp. 578–584, Kyoto, Japan, October 2009.

[3] N. Deng, Y. Xu, D. Sun, P. Hua, X. Zheng, and H. Duan, "Image processing for fusion identification between the GLUT4 storage vesicles and the plasma membrane," *Journal of Signal Processing Systems*, vol. 54, no. 1–3, pp. 115–125, 2009.

[4] L. Bai, Y. Wang, J. Fan et al., "Dissecting multiple steps of GLUT4 trafficking and identifying the sites of insulin action," *Cell Metabolism*, vol. 5, no. 1, pp. 47–57, 2007.

[5] S. Huang, L. M. Lifshitz, C. Jones et al., "Insulin stimulates membrane fusion and GLUT4 accumulation in clathrin coats on adipocyte plasma membranes," *Molecular and Cellular Biology*, vol. 27, no. 9, pp. 3456–3469, 2007.

[6] P. Vallotton, D. James, and W. Hughes, "Towards fully automated identification of vesicle-membrane fusion events in TIRF microscopy," in *Computational Models for Life Sciences—CMLS '07: 2007 International Symposium on Computational Models of Life Sciences*, AIP Conference Proceedings, pp. 3–10, 2007.

[7] L. Berger, M. Mirmehdi, S. Reed, and J. Tavaré, "A diffusion model for detecting and classifying vesicle fusion and undocking events," in *Medical Image Computing and Computer-Assisted Intervention*, vol. 6893 of *Lecture Notes in Computer Science*, pp. 329–336, 2012.

[8] J.-C. Olivo-Marin, "Extraction of spots in biological images using multiscale products," *Pattern Recognition*, vol. 35, no. 9, pp. 1989–1996, 2002.

[9] M. K. Cheezum, W. F. Walker, and W. H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," *Biophysical Journal*, vol. 81, no. 4, pp. 2378–2388, 2001.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 571473, 12 pages
http://dx.doi.org/10.1155/2015/571473

Research Article

An Adaptive Thresholding Method for BTV Estimation Incorporating PET Reconstruction Parameters: A Multicenter Study of the Robustness and the Reliability

M. Brambilla,[1] R. Matheoud,[1] C. Basile,[2] C. Bracco,[3] I. Castiglioni,[4] C. Cavedon,[5] M. Cremonesi,[6] S. Morzenti,[7] F. Fioroni,[8] M. Giri,[5] F. Botta,[6] F. Gallivanone,[4] E. Grassi,[8] M. Pacilio,[2] E. De Ponti,[7] M. Stasi,[3] S. Pasetto,[9] S. Valzano,[1] and D. Zanni[9]

[1]*Department of Medical Physics, University Hospital Maggiore della Carità, 28100 Novara, Italy*
[2]*Department of Medical Physics, Hospital S. Camillo Forlanini, 00152 Roma, Italy*
[3]*Department of Medical Physics, Institute for Cancer Research and Treatment (IRCC), 10060 Candiolo, Italy*
[4]*Institute of Molecular Bioimaging and Physiology, National Research Council (IBFM-CNR), 20093 Milan, Italy*
[5]*Department of Medical Physics, University Hospital, 37126 Verona, Italy*
[6]*Department of Nuclear Medicine, European Institute of Oncology, 20141 Milan, Italy*
[7]*Department of Medical Physics, Hospital San Gerardo, 20900 Monza, Italy*
[8]*Department of Medical Physics, Arcispedale S.Maria Nuova, IRCCS, 42123 Reggio Emilia, Italy*
[9]*Department of Medical Physics, Hospital Niguarda, 20162 Milan, Italy*

Correspondence should be addressed to I. Castiglioni; castiglioni.isabella@hsr.it

Received 12 September 2014; Accepted 25 December 2014

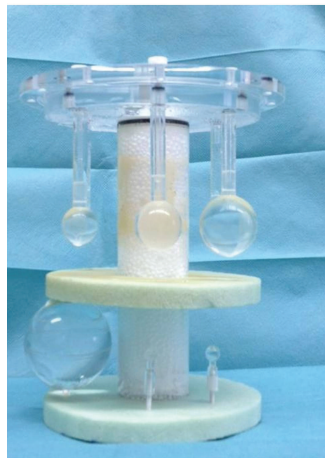
Academic Editor: Tianye Niu

Objective. The aim of this work was to assess robustness and reliability of an adaptive thresholding algorithm for the biological target volume estimation incorporating reconstruction parameters. *Method*. In a multicenter study, a phantom with spheres of different diameters (6.5–57.4 mm) was filled with ^{18}F-FDG at different target-to-background ratios (TBR: 2.5–70) and scanned for different acquisition periods (2–5 min). Image reconstruction algorithms were used varying number of iterations and postreconstruction transaxial smoothing. Optimal thresholds (TS) for volume estimation were determined as percentage of the maximum intensity in the cross section area of the spheres. Multiple regression techniques were used to identify relevant predictors of TS. *Results*. The goodness of the model fit was high (R^2: 0.74–0.92). TBR was the most significant predictor of TS. For all scanners, except the Gemini scanners, FWHM was an independent predictor of TS. Significant differences were observed between scanners of different models, but not between different scanners of the same model. The shrinkage on cross validation was small and indicative of excellent reliability of model estimation. *Conclusions*. Incorporation of postreconstruction filtering FWHM in an adaptive thresholding algorithm for the BTV estimation allows obtaining a robust and reliable method to be applied to a variety of different scanners, without scanner-specific individual calibration.

1. Introduction

In the last years the coregistration of ^{18}F-fluorodeoxyglucose positron emission tomography (^{18}F-FDG PET) images with computed tomography (CT) images has gained an increasing interest in the staging and treatment planning for radiotherapy of several tumor sites. However, a standardized way of converting PET signals into target volumes is not yet available [1].

New semiautomatic and automatic segmentation methods have been developed implying gradient, region growing, clustering, statistical methods and other approaches [2–9]. While referring to these promising methods, it should be pointed out that some of these new methods suffer from the need of extensive preprocessing of the images (e.g., edge-detection). Moreover, the majority of these new algorithms are not widely available, so their use is currently restricted to



Sphere ID (mm)	Sphere A (mm^2)	Sphere volume (mL)
6.5	33.2	0.14
8.1	51.5	0.28
10	79	0.52
13	133	1.15
17	227	2.57
22	383	5.57
28	615	11.49
37	1075	26.52
57.4	2570	99.02

Figure 1: Inserts in the IEC phantom used for the multicenter measurements comprising 9 fillable spheres of different diameters.

the developers and, as a consequence, they are not independently validated.

Apart from visual inspection of PET scans, which suffers from interobserver variability [10], thresholding methods are widely used as PET segmentation approach in clinical practice for biological target volume (BTV) delineation for radiotherapy planning. Adaptive thresholding methods based on contrast-oriented contouring algorithms have been developed independently by many groups and validated in patient data in head and neck, in lung cancer, and in lymph nodes [11–13]. These methods are based on phantom measurements to derive a relationship between the "true" volume and the threshold to be applied to PET images.

These threshold-volume curves for one PET/CT scanner have been previously obtained varying target-to-background ratio (TBR), target dimensions, and postreconstruction smoothing [14]. It has been previously demonstrated that the emission scan duration (ESD) and background activity concentration, related to the level of image noise, are not predictors of the thresholding level of PET images [15]. Moreover, adaptive-threshold segmentation algorithms are not influenced by the different conditions of attenuation and scatter which may be encountered in different anatomical districts [16] and by the degree of convergence of iterative reconstruction algorithms [14].

Although adaptive thresholding methods are applicable to every PET scanner, it is generally assumed that the values of the parameters obtained during model building are system dependent so that a specific calibration for each PET system is required. On the other hand, a less hardware-dependent solution to the problem of PET segmentation could provide a robust algorithm, easily usable with images acquired by different scanner models without needing any previous optimization of the individual image quality.

Our hypothesis is that this goal can be accomplished by incorporating in our algorithm the reconstruction parameters that impact on threshold determination. To validate this hypothesis we firstly developed an original method to adapt the thresholds (TS) used to estimate the BTV in PET images. The proposed method incorporates the PET reconstruction parameters that influence the threshold determination. Secondly, we investigated in a multicenter trial the robustness of this method with respect to various scanner models, reconstruction settings, and acquisition conditions: a multivariable approach was adopted to study the dependence of the TS that define the boundaries of ^{18}F-FDG uptake on object characteristics (contrast, size), acquisition parameters (scan duration), and reconstruction modalities (reconstruction algorithm, number of iterations, and amount of postreconstruction smoothing) in eleven state-of-the-art PET/CT scanners installed in eight different institutions. Finally, we assessed the reliability of the regression models through the use of split-sample analysis.

2. Materials and Methods

2.1. Phantoms. Measurements were performed on the NEMA IEC Body Phantom Set (Data Spectrum Corporation, Hillsborough, NC). This phantom contains 6 coplanar spheres, with internal diameters (ID) of 10, 13, 17, 22, 28, and 37 mm. A supplemental set of 2 microhollow spheres of 6.5 and 8.1 mm ID and 1 sphere of 57.4 mm ID were positioned at the bottom of the phantom. The experimental setup is depicted in Figure 1, together with sphere IDs (mm), maximum cross section areas (A) (mm^2), and volumes (mL). The same positioning of the phantom was ensured through laser localizer and a scout CT acquisition.

2.2. PET/CT Scanners. Eleven PET/CT scanners were used for the robustness study: n. 2 Discovery ST (S1, S2) [17], n. 1 Discovery STE (S3) [18], n. 2 Discovery 600 (S4, S5) [19], and n. 1 Discovery 690 (S6) [20] (GE Healthcare, Milwaukee, WI), n. 1 Biograph HI-REZ (S7) [21] and n. 1 Biograph TRUEV (S8) [22] (SIEMENS Medical Solutions, Knoxville, TN), n. 1 Gemini XL (S9) and n. 2 Gemini TF (S10, S11) [23] (Philips Medical Systems, Cleveland, OH). The technical characteristics and physical performances of the PET/CT scanners were derived from factory data and/or previous publications and are reported in Table 1.

2.3. Phantom Acquisition. The background of the IEC phantom was filled with 3 kBq/mL activity concentration of ^{18}F-FDG. A standard protocol was designed to generate the following acquisitions for each scanner model.

Table 1: PET/CT scanners: main technical characteristics and physical performances.

	Discovery ST (general electric) S1-2	Discovery STE (general electric) S3	Discovery 600 (general electric) S4-5	Discovery 690 (general electric) S6	Biograph16 Hi-REZ (Siemens) S7	Biograph 6 True V (Siemens) S8	Gemini XL (Philips) S9	Gemini TF (Philips) S10
Detector ring diameter (cm)	88.6	88.6	80.1	80.1	83.0	83.0	88.5	90.3
Detector material	BGO	BGO	BGO	LYSO	LSO	LSO	GSO	LYSO
Acquisition mode	2D/3D	2D/3D	3D	3D	3D	3D	3D	3D
Number of individual crystals	10.080	13.440	12.288	13824	24.336	32448	17.864	28.336
Number of crystal/ring	420	560	512	576	624	624	616	NA
Number of image planes	47	47	47	47	81	109	90	90
Crystal size (mm^3)	$6.3 \times 6.3 \times 30$	$4.7 \times 6.3 \times 30$	$4.7 \times 6.3 \times 30$	$4.2 \times 6.3 \times 25$	$4 \times 4 \times 20$	$4 \times 4 \times 20$	$4 \times 6 \times 30$	$4 \times 4 \times 22$
Patient port diameter (cm)	70	70	70	70	70	70	70	71.7
Axial field of view (cm)	15.7	15.7	15.7	15.7	16.2	21.8	18.0	18.0
Transaxial filed view (cm)	70	70	70	70	58.5	60.4	57.6	57.6
Axial sampling interval (mm)	3.27	3.27	3.27	3.27	2.0	2.0	2.0	2.0
Coincidence window width (ns)	11.7	9.3	9.0	4.9	4.5	4.5	7.5	6.0
Lower energy threshold (keV)	375	425	425	425	425	425	410	440
Physical performances								
Transverse resolution								
FWHM (mm) at 1 cm	6.29	5.1	4.9	4.70	4.61	4.1	5.2	4.8
FWHM (mm) at 10 cm	6.82	5.7	5.6	5.06	5.34	4.8	5.8	5.0
Axial resolution								
FWHM (mm) at 1 cm	5.68	5.2	5.6	4.74	5.10	4.7	5.8	4.8
FWHM (mm) at 10 cm	6.05	5.9	6.4	5.55	5.93	5.7	6.6	5.2
System sensitivity (cps/KBq)	8.99	8.8	9.6	7.5	4.87	8.0	8.0	6.6
Scatter fraction (%)	45	34	36.6	37	34.1	32.7	35	27

(a) Nine different TBRs (2.5 : 1, 4 : 1, 8 : 1, 16 : 1, 25 : 1, 35 : 1, 47 : 1, 55 : 1, and 70 : 1), determined by the dose calibrator and dilution, were imaged in different acquisition sessions. The measured TBRs were determined in the reconstructed image as the maximum pixel intensity in a region of interest (ROI) encircling the cross sectional area of the target, divided by the average pixel intensity of ROIs surrounding the sphere. These TBRs ranged from 70 down to 2.5 and were within the full range observed in patients.

(b) Four different ESD (2, 3, 4, and 5 min) were acquired to provide independent replicates of the experiments.

2.4. PET Image Reconstruction

2.4.1. Discovery ST, Biograph HI-REZ, and Biograph TRUEV. These systems use a 2D Fourier-rebinning (FORE) ordered subset expectation maximization (OSEM) algorithm with all corrections (scatter, random, dead time, attenuation, and normalization) incorporated into the iterative reconstruction scheme. In these systems the user can independently specify the number of iterations and subsets and the amount of the transaxial postreconstruction Gaussian smoothing, through the filter full-width-at-half-maximum (FWHM) expressed in mm.

2.4.2. Discovery STE and Discovery 600. The D-600 system uses a fully 3D-OSEM algorithm with all corrections incorporated into the iterative reconstruction scheme. The reconstruction settings are the same as above with the only difference that the axial filter is a mean filter with available kernels of 1 : 2 : 1, 1 : 4 : 1, and 1 : 6 : 1.

2.4.3. Discovery 690. The D-690 system uses a fully 3D-OSEM algorithm with all corrections incorporated into the iterative reconstruction scheme. Furthermore, new reconstruction algorithms are available on the D-690, which add to the standard configuration the time of flight information (TOF) and/or a 3D model of the D-690 PET point spread function (PSF). The activation of TOF and/or PSF does not require the setting of any new parameter compared to those used with the 3D-OSEM algorithm (number of subsets, number of iterations, reconstructed field of view (FOV), image matrix, and axial and transaxial postfilters). In this study, both TOF and PSF information were included in the reconstruction scheme.

2.4.4. Gemini XL. This system uses a fully 3D line-of-response (LOR) based iterative reconstruction algorithm named row-action maximum likelihood algorithm (RAMLA) [24]. The number of iterations is fixed (2 iterations and 33 subsets) and the reconstruction protocols contain one modifiable parameter that can be set to adjust the quality of the images as normal, smooth, or sharp.

2.4.5. Gemini TF. This system uses the TOF maximum likelihood expectation-maximization reconstruction algorithm (TF-MLEM) [24]. The reconstruction protocols contain three modifiable parameters that can be set to adjust the quality of the images: the first is the number of iterations (3 iterations and 20 subsets or 3 iterations and 33 subsets); the second is a so called relaxation parameter that can be set between 1, 0.7 and 0.5 and controls the magnitude of change that each iteration makes to the image. A third parameter, the kernel width of the TOF, can be set by the user at two levels (Gemini TF manual) [25].

The type of reconstruction algorithm, the degree of the convergence of the iterative algorithm, and the amount of the postreconstruction smoothing applied on images were varied starting from the clinical acquisition protocols used in each institution for radiotherapy planning. Overall, in each scanner, the maximum of theoretical independent combinations of acquisition parameters available for the subsequent model fitting were 9 sphere $A \times 9$ TBR $\times$ 4 ESD = 324. The 8.1 and 6.5 mm spheres was not always included in the analysis because they were not clearly visible in all the phantom acquisitions. The number of reconstruction modalities available for model fitting depends on the scanner capabilities: the details of the reconstruction parameters together with the voxel size of the reconstructed images and the number of data points that is actually available for model fitting in each scanner are shown in Table 2.

2.5. Image Analysis. TS were determined as a percentage of the maximum intensity in the cross section area of the spheres. Target cross sections of area A were selected in the middle of the spheres, which constitutes the largest cross section of the sphere. The values of TS were entirely based on the apparent activity concentration in the images and not on the known activities actually placed in the spheres. To find the TS value that yielded an area A best matching the true value, the cross sections were autocontoured in the attenuation corrected slices varying TS in step of 1%, until the area so determined differed by less than 10 mm^2 versus its known physical value.

The analysis was performed by means of an automatic routine, EyeLite RT v.1.1 (G-Squared, Vicenza, Italy), to avoid the influence of the operator in ROIs dimensioning and to minimize the influence of the operator in the ROIs positioning. The operator placed six 17 mm-diameter ROIs in the background area surrounding the spheres. The mean intensity of these 6 ROIs was used as a background value (BG). ROI analyses were performed only for visually detectable spheres: this accounted for the discrepancy between theoretical and experimental data points collected for each scanner.

2.6. Statistical Analysis. For each combination of EM-equivalent iteration number (i) and ESD (j), the following variables were evaluated: X_{1ij} defined as target cross section A, X_{2ij} defined as $1 - 1/\text{TBR}$, and X_{3ij} defined as the FWHM.

Multiple linear regression analysis was performed in order to define the relationship between the best TS (TS_{ij}) (providing the most accurate sphere cross sectional area) and X_{1ij}, X_{2ij}, and X_{3ij}. The multiple regression model used for the fit was

$$\text{TS}_{ij} = B_0 + B_1 \times X_{1ij}\left(\text{mm}^2\right) + B_2 \times X_{2ij} + B_3 \times X_{3ij} + E, \quad (1)$$

Table 2: PET/CT scanners: reconstruction parameters.

	Discovery ST S1	Discovery ST S2	Discovery STE S3	Discovery 600 S4	Discovery 600 S5	Discovery 690 S6	Biograph 16 Hi-REZ S7	Biograph 6 true V S8	Gemini XL S9	Gemini TF S10	Gemini TF S11
Reconstruction protocol	FORE-OSEM	FORE-OSEM	3D-OSEM	3D-OSEM	3D-OSEM	3D-OSEM	FORE-OSEM	FORE-OSEM	LOR-RAMLA	TF-MLEM	TF-MLEM
Number of iterations	14, 42	16, 40	14, 42	16, 32	16, 48	54, 108	16, 24	16, 42	66	60, 99	60, 99
Transaxial smoothing FWHM (mm)	6, 9, 13	5.5, 8.2, 11	6, 8, 11	6, 8, 11	5.5, 8.2, 11	4, 6, 8	4, 6, 8	4, 6, 8	—	—	—
Axial smoothing (kernel)	—	—	1 : 4 : 1	1 : 4 : 1	1 : 4 : 1	1 : 6 : 1 1 : 4.1 1 : 2 : 1	—	—	—	—	—
Kernel width (cm)	—	—	—	—	—	—	—	—	—	14.1	14.1, 18.7*
Relaxation (λ)	—	—	—	—	—	—	—	—	0.5, 0.7, 1	0.5, 0.7, 1	0.5, 0.7, 1
Voxel dimensions $L \times W \times H$ (mm)	$2.7 \times 2.7 \times 3.3$	$2.7 \times 2.7 \times 3.3$	$2.7 \times 2.7 \times 3.3$	$2.7 \times 2.7 \times 3.3$	$2.7 \times 2.7 \times 3.3$	$2.7 \times 2.7 \times 3.3$	$2.6 \times 2.6 \times 2$ $5.3 \times 5.3 \times 2$	$4.1 \times 4.1 \times 5$	$4 \times 4 \times 3$	$4 \times 4 \times 3$	$4 \times 4 \times 3$
Number of reconstructions	6	6	6	6	6	18	12	6	3	6	9
Number of data points	1641	1651	1676	1785	1727	5368	3456	1676	814	1993	2466

*Only available with 99 iterations.

where B_0, B_1, B_2, and B_3 are the regression coefficients to be estimated and E is the error term. The hypothesis of linear dependence between TS and independent variables X were already demonstrated in Brambilla et al. and in Matheoud et al. [14, 15]. However, nonlinear objective functions or even indicator functions according to the different parameters range could be investigated for fitting TS.

Additional variables, reflecting the characteristics of the reconstruction protocol of each considered scanner, were inserted in the model as independent predictors. Axial smoothing was considered for the Discovery 690 and voxel dimensions were inserted for the Biograph Hi-REZ, while relaxation parameter and TOF kernel width were accounted for in the Gemini and in the Gemini TF, respectively.

Stepwise forward selection was used as a strategy for selecting the variables and F statistic was used as a criterion for selecting a model. Goodness of fit for each regression model was expressed using the adjusted coefficient of determination (R^2). Goodness of fit was reported at each stage of model building as partial R^2. The criteria for retaining a variable in a model were $F > 4$ and an increment of at least 0.01 in the R^2 in order to be cautious in including redundant variables into the models. The weight of the different independent variables in explaining TS was quantified by means of standardized regression coefficients βi.

The reliability of the regression models was assessed through split-sample analysis [26]. Using this methodology, all observations in each scanner model were randomly assigned to one of two groups, the training group or the holdout group. The regression models were derived using the training group and the sample squared multiple correlation R^2 was obtained. Then the prediction equation for the training group was used to compute predicted values for the holdout group. Finally, the univariate correlation R^{2*} (cross validation correlation) was obtained between these predicted values and the observed responses in the holdout group. The reliability of the regression models was expressed by using the shrinkage on cross validation coefficients $R^2 - R^{2*}$. As a criterion, shrinkage values of less than 0.10 were considered as indicative of a reliable model.

In order to compare separate multiple regressions of TS as a function of the X independent variables for two different scanners, an additional dummy variable, coding for each scanner, was inserted in the model. A regression model was then built by pooling the data coming from the two scanners and inserting this dummy variable as a predictor. The criteria for retaining this variable in the model were the ones specified above.

Statistical analysis was performed using the software Statistica 6.0 (Statsoft Inc., Tulsa OK).

3. Results and Discussion

3.1. Multiple Linear Regression. Figure 2 shows the plots of averaged TS versus cross sectional areas for a coarse grouping of TBRs for each scanner model.

The TS versus predictor variables plot was fitted only for cross sectional area > 133 mm^2 that is in the range of clinically relevant volumes comprised between 1 and 100 mL. The cross sectional area of 133 mm^2 (that corresponds to a sphere ID of 13 mm and approximately the twofold FWHM of the scanners) was selected as a separator of the data due to the resolution characteristics of the scanners.

Following (1), the regression equations that best summarize the results obtained in a multiple regression model with TS as the predicted variable are reported for each scanner in Table 3 together with the values of the corresponding parameters B_0–B_3. In the third column of Table 3 the multiple-R^2 of model fitting are reported, while the last column shows the ranking of the independent predictors together with the standardized regression coefficients and the amount of TS variance explained by each predictor.

The emission scan duration and the degree of the convergence of the iterative algorithm were never significant predictors of TS. This provided a confirmation of previously reported findings. Also the axial smoothing in the Discovery 690, the voxel size in the Biograph Hi-REZ, the relaxation λ in the Gemini scanners, and the Kernel width of the time of flight correction in the Gemini TF were not significant predictors of TS.

The goodness of the model fit, assessed by the coefficient of determination R^2, was high, ranging from a minimum of 0.74 for the Discovery 690 to a maximum of 0.92 for both the Discovery 600 and the Biograph Hi-REZ. The most relevant variable for TS prediction was TBR with a partial R^2 accounting for 74% to 91% of TS variability. In the case of the Discovery 690, TBR only accounted for 40% of TS variability, although it remains the best individual predictor. Second came the amount of smoothing in the transaxial plane (FWHM) that showed an additional R^2 roughly explaining from 1 to 5% of TS variability. The only exceptions were the Gemini scanners, where this parameter cannot be varied by the user, and the Discovery 690, where its contribution is significantly increased to 29% of TS variability. Last came the lesion size (A) that played an independent role only in the Discovery 690 and in the Gemini scanners accounting for 5%–8% of TS variability.

The comparison of the regression lines obtained from two scanners of the same model (Discovery ST, Discovery 600, and Gemini TF) did not evidence any relevant difference. The test of the hypothesis of coincident regression lines for the two DST scanners provided an $F_{3,2145} = 0.73$ ($P = 0.53$). This F statistics is small (P is large), so we do not reject H_0 and therefore have no statistical basis for believing that the two lines are not coincident. The details of the test are reported in Table 4. Moreover, the additional R^2 of the "scanner" dummy variable as predictor of TS variance was below <0.001. Similar results were found for the D600 and GTF scanners (not shown). Accordingly, the results of the regression analysis obtained by pooling all the measurements from scanners of the same model are reported in Table 3.

3.2. Regression Model Reliability. The results of the reliability study on regression models are reported in Table 3. The shrinkage on cross validation was always below 0.07, which is quite small and indicative of an excellent reliability of

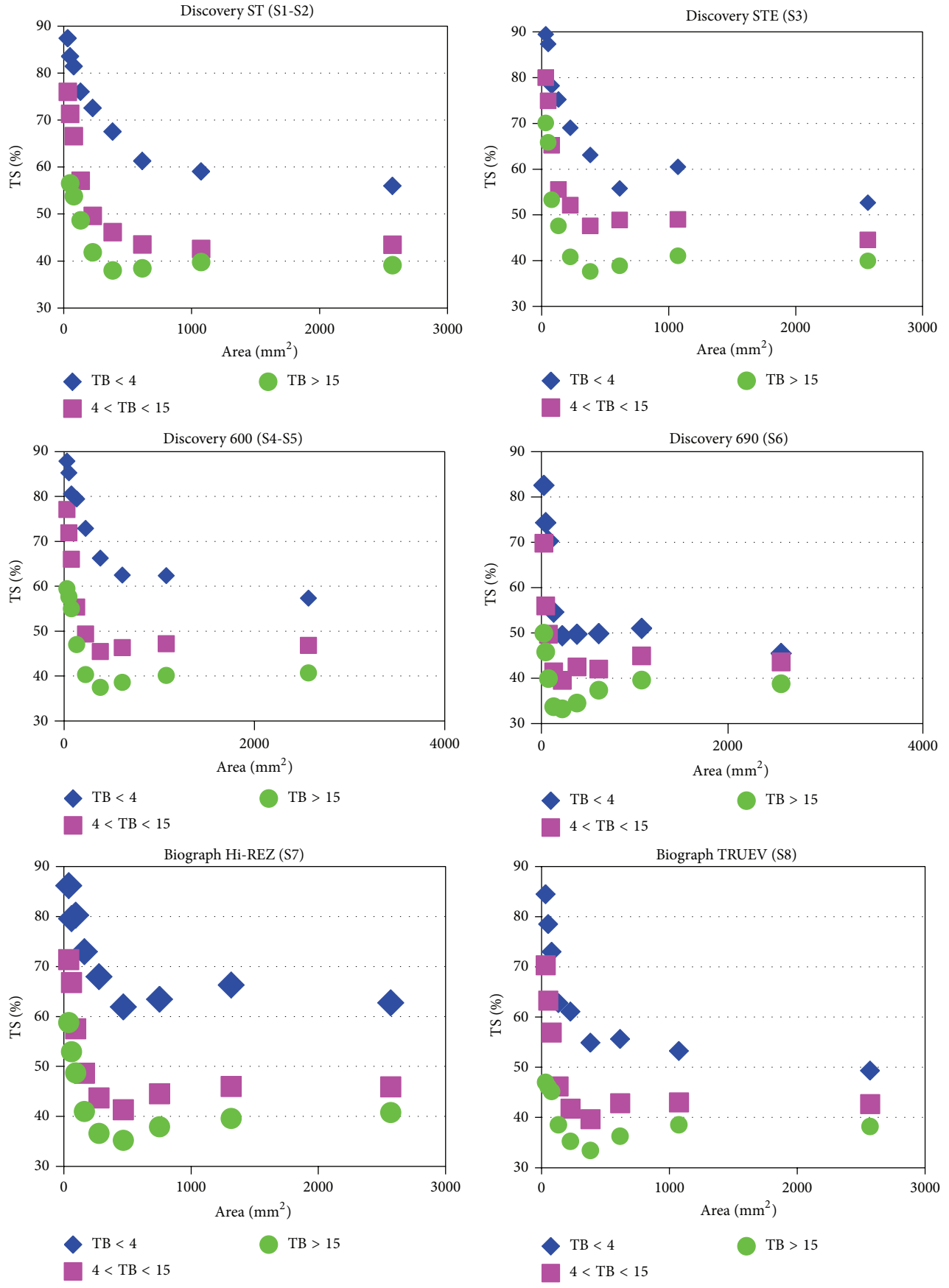


Figure 2: Continued.

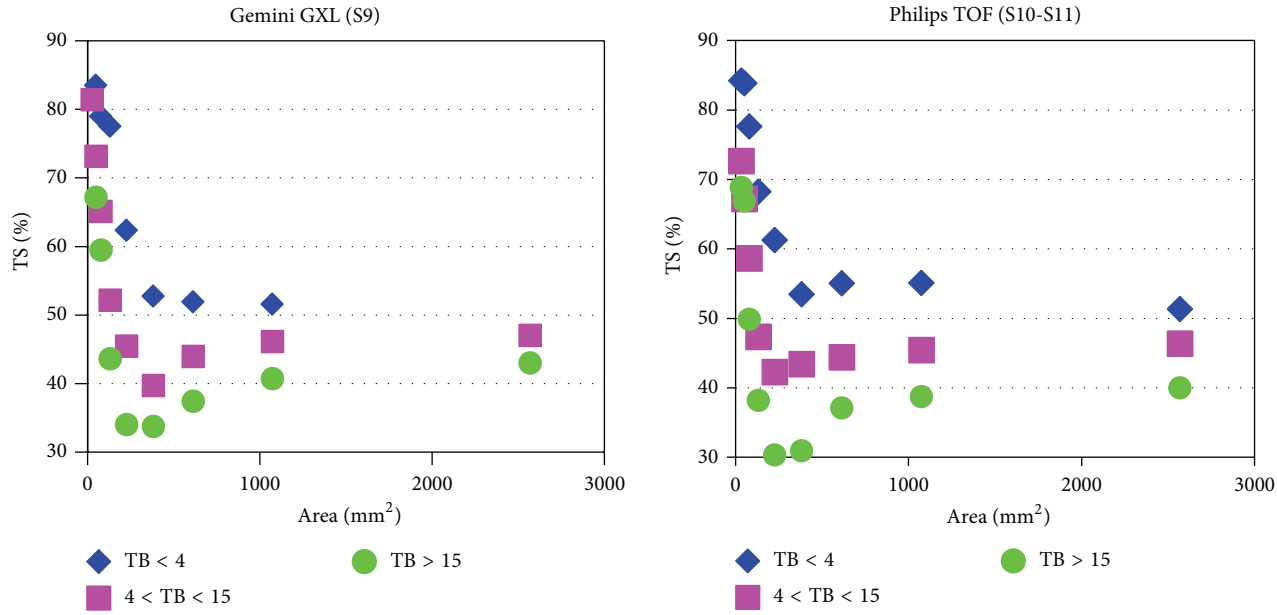


FIGURE 2: Plots of averaged TS versus cross sectional areas for a coarse grouping of TBRs for each scanner model. Measured data points have been omitted here to prevent obscuring differences in trends.

estimation. An important aspect related to assessing the reliability of a model involves considering difference score of the form $TS_{observed} - TS_{predicted}$, where only holdout cases are used and when the training sample equation is used to compute the predicted values. The "unstandardized residuals" can be subjected to various residual analyses. The most helpful entails univariate descriptive statistics such as the box and whiskers plots depicted in Figure 3. In our case a few large residuals are present, but they are neither sufficiently implausible nor influential to require further investigations.

In our investigation, we derived the calibration curve for eleven PET scanners (eight models, three manufacturers, and eight sites) to apply the adaptive-threshold algorithm for PET-based contouring. The eight scanner types investigated in this study differ in scintillation crystal, scanner electronics, and reconstruction methodologies. Methods of retrospective image resolution recovery such as PSF-reconstruction or TOF measurements were also characterized in the present study.

At present, there is considerable variability in the way standard PET/CT scans are performed in different centers [27, 28]. Thus, there was no chance for a multicenter standardization of all scanners and all imaging protocols in use. Instead, we chose to directly incorporate in our adaptive thresholding algorithm the reconstruction parameters that can be selected by the user and that are relevant for TS determination. This should increase the robustness of the proposed method by avoiding the need to perform individual calibrations in each center of the algorithm. Noteworthy, the comparison of the regression lines obtained from two scanners of the same model did not evidence any relevant difference, at least for the three scanner models tested. This brings another relevant consequence; that is, with the incorporation of the reconstruction parameters in the regression models the calibration curve in a specific scanner model need not to be obtained at each site. Instead, it can be derived once and applied irrespectively of the specific scanner being utilized provided that is of the same model.

3.3. Comparison with Previous Published Papers. To the best of our knowledge only three studies have been published so far on the integration of PET/CT scans from different hospitals into radiotherapy treatment planning. In the first study, Öllers et al. [29] used a TBR algorithm to evaluate head-and-neck tumors. To this purpose only small spheres of volumes ranging from 2 to 16 mL (i.e., sphere ID less than 3 cm) were used. TBRs, as determined by the dose calibrators, ranged from 2 to 12. The authors performed phantom measurements on three scanners of the same manufacturer (Biograph Accel, Siemens) equipped with Pico 3D (2 scanner) or standard (1 scanner) detector electronics. Identical acquisition and reconstruction protocols were used. To study the effect of different reconstruction parameters on the results PET raw data were reconstructed varying the number of iterations (IT from 2 to 64) with a fixed smoothing of FWHM = 5 mm. They found that the standardized uptake value (SUV) threshold of the scanner equipped with standard electronic differed significantly from those of the other two scanners and that at least 16 iterations are required in order to produce reliable SUV thresholds. Our own results support these findings. On the one hand, the regression lines did not differ significantly between scanners of the same type equipped with similar electronics, while the calibration curves for scanners of different type clearly differ (Figure 2). On the other hand, the number of iteration is not a significant predictor of TS, provided that this number is kept above a certain level which is both recommended by the manufactures and necessary to have good image quality.

Table 3: Scanner-model specific calibration curves, model R^2, shrinkage on cross validation and independent predictor of TS ranked in order of significance.

	Equation	R^2	$R^2(1) - R^2 * (2)$	TS predictors
Discovery ST	$TS = 90.68 - 62.44 \cdot (1 - 1/TBR) + 1.05 \cdot FWHM\ (mm)$	0.91	−0.008	$1 - 1/TBR$ ($\beta_2 = -0.91$, partial $R^2 = 0.85$) FWHM ($\beta_3 = 0.23$, additional $R^2 = 0.06$)
Discovery STE	$TS = 88.68 - 59.39 \cdot (1 - 1/TBR) + 1.03 \cdot FWHM\ (mm)$	0.87	0.004	$1 - 1/TBR$ ($\beta_2 = -0.90$, partial $R^2 = 0.82$) FWHM ($\beta_3 = 0.23$, additional $R^2 = 0.05$)
Discovery 600	$TS = 93.52 - 64.38 \cdot (1 - 1/TBR) + 0.99 \cdot FWHM\ (mm)$	0.92	0.000	$1 - 1/TBR$ ($\beta_2 = -0.93$, partial $R^2 = 0.88$) FWHM ($\beta_3 = 0.20$, additional $R^2 = 0.04$)
Discovery 690	$TS = 63.04 - 0.015 \cdot A\ (mm^2) - 40.51 \cdot (1 - 1/TBR) + 1.92 \cdot FWHM\ (mm)$	0.74	0.007	$1 - 1/TBR$ ($\beta_2 = -0.61$, partial $R^2 = 0.40$) FWHM ($\beta_3 = 0.54$, additional $R^2 = 0.29$) A ($\beta_1 = 0.22$, additional $R^2 = 0.05$)
Biograph Hi-REZ	$TS = 88.19 - 56.18 \cdot (1 - 1/TBR) + 0.67 \cdot FWHM\ (mm)$	0.92	−0.012	$1 - 1/TBR$ ($\beta_2 = -0.95$, partial $R^2 = 0.91$) FWHM ($\beta_3 = 0.10$, additional $R^2 = 0.01$)
Biograph TRUEV	$TS = 90.43 - 61.86 \cdot (1 - 1/TBR) + 0.95 \cdot FWHM\ (mm)$	0.89	−0.013	$1 - 1/TBR$ ($\beta_2 = -0.91$, partial $R^2 = 0.85$) FWHM ($\beta_3 = 0.19$, additional $R^2 = 0.04$)
Gemini XL	$TS = 92.04 + 0.0025 \cdot A\ (mm^2) - 59.15 \cdot (1 - 1/TBR)$	0.82	0.067	$1 - 1/TBR$ ($\beta_2 = -0.89$ partial $R^2 = 0.74$) A ($\beta_1 = 0.27$, additional $R^2 = 0.08$)
Gemini TF	$TS = 88.57 + 0.0027 \cdot A\ (mm^2) - 57.44 \cdot (1 - 1/TBR)$	0.84	−0.009	$1 - 1/TBR$ ($\beta_2 = -0.88$ partial $R^2 = 0.76$) A ($\beta_1 = 0.28$, additional $R^2 = 0.08$)

Table 4: Analysis of variance table. Test of H_0 = coincident regression lines for the two DST scanners.

$A > 133\,\text{mm}^2$	Sum of squares (SS)	Degrees of freedom	Mean square (MS)	F	P
Reduced model					
Regression	281844	2	140922	10445	$< 10^{-6}$
Residuals	28980	2148	13.5		
Full model					
Regression	281874	3	93958	6968	$< 10^{-6}$
Residuals	28951	2147	13.5		

$F = (28980{,}3 - 28950{,}6)/3/13.48 = 0.73$; $P = 0.53$.

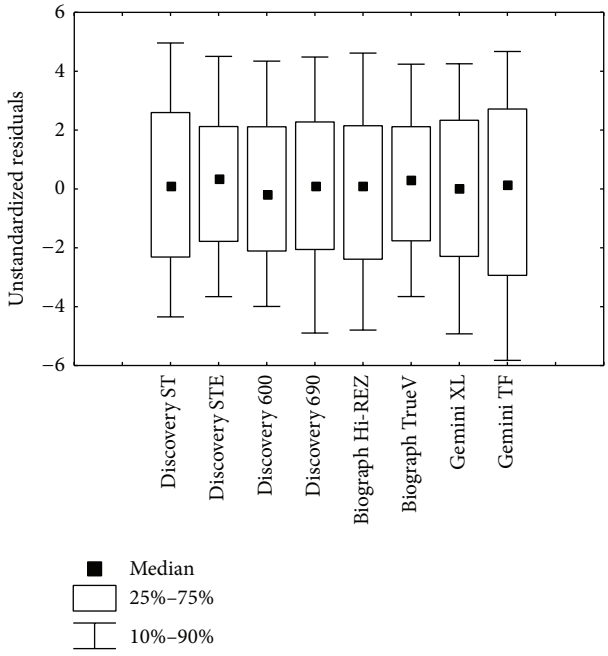


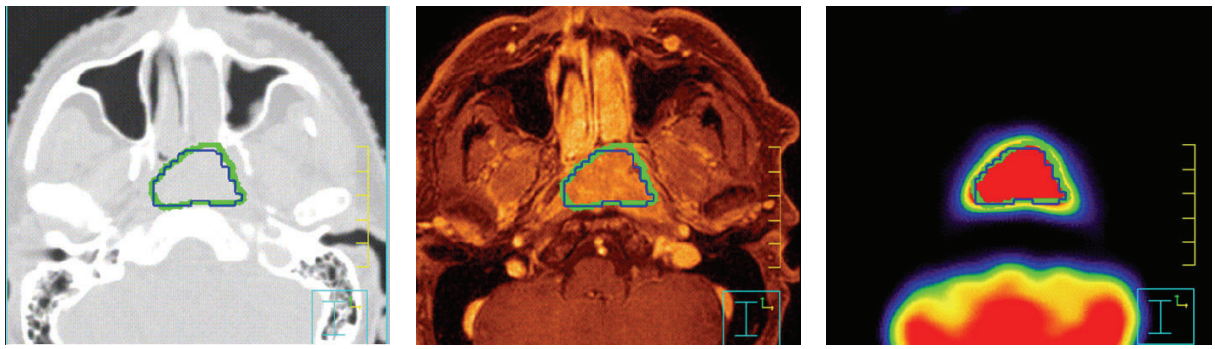
Figure 3: Box and whiskers plot of unstandardized residuals ($TS_{observed} - TS_{predicted}$) for the different scanners where only holdout cases are used and when the training sample equation is used to compute the predicted values.

In a second study, Hatt et al. [30] evaluated the robustness and repeatability of a TBR algorithm in comparison to fuzzy C-means clustering and fuzzy locally adaptive Bayesian algorithm. The authors performed phantom measurements on four different PET/CT scanners (Philips Gemini and Gemini TF, Siemens Biograph, and GE Discovery LS) using a standard acquisition protocol with two TBR (4 and 8) and Three ESD (1, 2, and 5 min). PET raw data were reconstructed using routine clinical image reconstruction and two voxel size volume for all scanners. They reported a higher robustness of the fuzzy locally adaptive Bayesian algorithm while the repeatability provided by all segmentation methods was very high with a negligible variability of <5% in comparison to that associated with manual delineation. However, as recognized by the same authors, in order to assess the robustness of the TBR approach they applied an adaptive thresholding using the parameters optimized on other scanners to the image datasets acquired with the Siemens Biograph, which is sort of misleading since the TBR approach is system dependent. Instead, by adopting scanner-model specific calibration curves, similar mean classification error (~10%) and variability (~5%) would have been obtained for the TBR algorithm and for the fuzzy locally adaptive Bayesian approach. By first principles, the inclusion of the postreconstruction smoothing (not considered in the study of Hatt) should increase both the accuracy and robustness of adaptive thresholding algorithms possibly leading to results even superior to those achieved by advanced image segmentation methods. Noteworthy, the coefficient of regression for the TBR variable reported by Hatt for the Gemini TF is very similar to the one obtained for the same variable in the present work (B_{TBR} = 61.4 versus 59.3), also considering that the two regression models are not identical. This provides, although indirectly, a further confirmation of the robustness of the scanner-model specific approach in deriving TS calibration curves.

In the last study Schaefer et al. [31] evaluated the calibration of an adaptive SUV thresholding algorithm in eleven centers equipped with 5 Siemens Biograph, 5 Philips Gemini, and one Siemens ECAT ART scanners. They reported only minor differences in calibration parameters for scanners of the same type provided that identical imaging protocols were used, whereas significant differences were found comparing scanners of different types. Moreover, they reported no statistically significant differences among SUV thresholds calculated for each site by use of the "site-specific" calibration neither among scanners of the same type at different sites nor among scanners of different types at different sites. Our own results support these findings only partially. In our study both acquisition and reconstruction parameters were varied and relevant parameters were incorporated into the "site-specific" algorithms so that there is no need to force individual centers to adopt a fixed protocol of image acquisition and image reconstruction. Bearing in mind this relevant difference, also in our study the calibration curves were not significantly different between scanners of the same type, whereas significant differences were found comparing scanners of different types. On the contrary, both the measured (Figure 2) and the calculated TS (Table 3) were significantly different among scanners of different types. For instance, the measured TS averaged over the entire spectrum of acquisition and reconstruction parameters for larger targets (sphere $A > 133\,\text{mm}^2$) for the Discovery 690 (S6) and the Discovery 600 (S4-5) were 39.0 ± 5.8% versus 45.4 ± 10.8, respectively ($P < 0.0001$). This difference largely reflects the hot-contrast recovery capabilities of the different scanners which, in the case of the Discovery 690, are emphasized by the introduction of PSF techniques in the reconstruction process.

3.4. Study Advantages and Limitations. The proposed method for the definition of BTV has several advantages, even if the results of this study must be interpreted in the context of some limitations.

Our method is feasible in a clinical context in those lesions presenting a uniform radiotracer uptake, as for



(a) (b) (c)

Figure 4: Application of the proposed method to a head and neck tumor: CT (a), MR (b), and PET (c) images (acquired on the biograph Hi-REZ PET scanner). The green ROI corresponds to the GTV delineation manually performed by the radiation oncologist, while the blue one is the result of the application of the 39% TS derived from the thresholding algorithm (TBR = 12, FWHM = 4 mm). The differences between manual GTV and GTV obtained from the thresholding algorithm is 4.3%.

different oncological lesions. In this case, the proposed method can be effective for extracting functional biomarkers and for using PET imaging in image-guided radiotherapy treatments. As a representative example, Figure 4 shows the application of the proposed method to a head and neck oncological patient, a candidate for image-guided radiotherapy. In these patients, BTV can be used to optimize radiotherapy treatment taking advantages from the information of functional imaging.

The effects of lesion movement in lung tumors have been recently incorporated in an adaptive thresholding algorithm using multiple regression techniques similar to those in the present study [32]. Though the effects of lesion movement were not included in this study, we believe that the conclusions regarding the effect of smoothing and TBR on thresholds still apply in the case of moving targets.

Threshold techniques do not take into account variations in tumor heterogeneity. This has motivated the investigation of advanced segmentation techniques not based on thresholding. While referring to these important methods for segmentation of nonuniform tracer concentration it should be pointed out that until they are further developed and validated, adaptive threshold segmentation methods are and will be used in most clinics and therefore need to be accurately characterized.

Furthermore, it has to be pointed out that in our work BTVs were fitted only for cross sections larger than 133 mm^2. This choice is justified by the fact that several studies found severe errors in the volume estimation for tumor volume < 2 mL corresponding to cross sections < 192 mm^2 (in terms of sphere-equivalent cross section) [33–35].

4. Conclusion

This study demonstrated that the calibration curves for the proposed adaptive thresholding method were not significantly different between scanners of the same type at different sites. The incorporation of the postreconstruction Gaussian smoothing in the algorithms avoids the need of system-dependent optimization procedures. This, together with the demonstrated high level of reliability of this approach, may provide robust and reliable tools to aid physicians as an initial guess in segmenting biological volumes on FDG-PET images.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors wish to thank the Tecnologie Avanzate TA Srl for the constant support provided for the realization of the study.

References

[1] U. Nestle, W. Weber, M. Hentschel, and A.-L. Grosu, "Biological imaging in radiation therapy: role of positron emission tomography," *Physics in Medicine and Biology*, vol. 54, no. 1, pp. R1–R25, 2009.

[2] X. Geets, J. A. Lee, A. Bol, M. Lonneux, and V. Grégoire, "A gradient-based method for segmenting FDG-PET images: methodology and validation," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 34, no. 9, pp. 1427–1438, 2007.

[3] H. Li, W. L. Thorstad, K. J. Biehl et al., "A novel PET tumor delineation method based on adaptive region- growing and dual-front active contours," *Medical Physics*, vol. 35, pp. 3711–3721, 2008.

[4] A. J. Green, R. J. Francis, S. Baig, and R. H. J. Begent, "Semiautomatic volume of interest drawing for 18F-FDG image analysis—method and preliminary results," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 35, no. 2, pp. 393–406, 2008.

[5] E. Day, J. Betler, D. Parda et al., "A region growing method for tumor volume segmentation on PET images for rectal and anal cancer patients," *Medical Physics*, vol. 36, no. 10, pp. 4349–4358, 2009.

[6] D. W. G. Montgomery, A. Amira, and H. Zaidi, "Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model," *Medical Physics*, vol. 34, no. 2, pp. 722–736, 2007.

[7] M. Halt, C. C. Le Rest, A. Turzo, C. Roux, and D. Visvikis, "A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET," *IEEE Transactions on Medical Imaging*, vol. 28, no. 6, pp. 881–893, 2009.

[8] I. El Naqa, D. Yang, A. Apte et al., "Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning," *Medical Physics*, vol. 34, no. 12, pp. 4738–4749, 2007.

[9] S. Belhassen and H. Zaidi, "A novel fuzzy C-means algorithm for unsupervised heterogeneous tumor quantification in PET," *Medical Physics*, vol. 37, no. 3, pp. 1309–1324, 2010.

[10] A. C. Riegel, A. M. Berson, S. Destian et al., "Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion," *International Journal of Radiation Oncology Biology Physics*, vol. 65, no. 3, pp. 726–732, 2006.

[11] J.-F. Daisne, T. Duprez, B. Weynand et al., "Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen," *Radiology*, vol. 233, no. 1, pp. 93–100, 2004.

[12] A. Schaefer, S. Kremp, D. Hellwig, C. Rübe, C. M. Kirsch, and U. Nestle, "A contrast-oriented algorithm for FDG-PET-based delineation of tumour volumes for the radiotherapy of lung cancer: derivation from phantom measurements and validation in patient data," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 35, no. 11, pp. 1989–1999, 2008.

[13] U. Nestle, A. Schaefer-Schuler, S. Kremp et al., "Target volume definition for ^{18}F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 34, no. 4, pp. 453–462, 2007.

[14] R. Matheoud, P. D. Monica, G. Loi et al., "Influence of reconstruction settings on the performance of adaptive thresholding algorithms for FDG-PET image segmentation in radiotherapy planning," *Journal of Applied Clinical Medical Physics*, vol. 12, no. 2, pp. 3363–3381, 2011.

[15] M. Brambilla, R. Matheoud, C. Secco, G. Loi, M. Krengli, and E. Inglese, "Threshold segmentation for PET target volume delineation in radiation treatment planning: the role of target-to-background ratio and target size," *Medical Physics*, vol. 35, no. 4, pp. 1207–1213, 2008.

[16] R. Matheoud, P. D. Monica, C. Secco et al., "Influence of different contributions of scatter and attenuation on the threshold values in contrast-based algorithms for volume segmentation," *Physica Medica*, vol. 27, no. 1, pp. 44–51, 2011.

[17] V. Bettinardi, M. Danna, A. Savi et al., "Performance evaluation of the new whole-body PET/CT scanner: discovery ST," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 31, no. 6, pp. 867–881, 2004.

[18] M. Teräs, T. Tolvanen, J. J. Johansson, J. J. Williams, and J. Knuuti, "Performance of the new generation of whole-body PET/CT scanners: discovery STE and Discovery VCT," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 34, no. 10, pp. 1683–1692, 2007.

[19] E. de Ponti, S. Morzenti, L. Guerra et al., "Performance measurements for the PET/CT Discovery-600 using NEMA NU 2-2007 standards," *Medical Physics*, vol. 38, no. 2, pp. 968–974, 2011.

[20] V. Bettinardi, L. Presotto, E. Rapisarda, M. Picchio, L. Gianolli, and M. C. Gilardi, "Physical performance of the new hybrid PETCT Discovery-690," *Medical Physics*, vol. 38, no. 10, pp. 5394–5411, 2011.

[21] M. Brambilla, C. Secco, M. Dominietto, R. Matheoud, G. Sacchetti, and E. Inglese, "Performance characteristics obtained for a new 3-dimensional lutetium oxyorthosilicate-based whole-body PET/CT scanner with the National Electrical Manufacturers Association NU 2-2001 standard," *Journal of Nuclear Medicine*, vol. 46, no. 12, pp. 2083–2091, 2005.

[22] B. W. Jakoby, Y. Bercier, M. Conti, M. E. Casey, B. Bendriem, and D. W. Townsend, "Physical and clinical performance of the mCT time-of-flight PET/CT scanner," *Physics in Medicine and Biology*, vol. 56, no. 8, pp. 2375–2389, 2011.

[23] S. Surti, A. Kuhn, M. E. Werner, A. E. Perkins, J. Kolthammer, and J. S. Karp, "Performance of Philips Gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities," *Journal of Nuclear Medicine*, vol. 48, no. 3, pp. 471–480, 2007.

[24] J. Browne and A. R. De Pierru, "A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography," *IEEE Transactions on Medical Imaging*, vol. 15, no. 5, pp. 687–699, 1996.

[25] T. F. Gemini, *Key to Success, Philips Medical Systems*, Koninklijke Philips Electronics N.V., Cleveland, Ohio, USA, 2007.

[26] D. G. Kleinbaum, L. L. Kupper, and K. E. Muler, *Applied Regression Analysis and Other Multivariable Methods*, PWS-Kent, Boston, Mass, USA, 1988.

[27] M. M. Graham, R. D. Badawi, and R. L. Wahl, "Variations in PET/CT methodology for oncologic imaging at U.S. academic medical centers: an imaging response assessment team survey," *The Journal of Nuclear Medicine*, vol. 52, no. 2, pp. 311–317, 2011.

[28] T. Beyer, J. Czernin, and L. S. Freudenberg, "Variations in clinical PET/CT operations: results of an international survey of active PET/CT users," *Journal of Nuclear Medicine*, vol. 52, no. 2, pp. 303–310, 2011.

[29] M. Öllers, G. Bosmans, A. van Baardwijk et al., "The integration of PET-CT scans from different hospitals into radiotherapy treatment planning," *Radiotherapy and Oncology*, vol. 87, no. 1, pp. 142–146, 2008.

[30] M. Hatt, C. Cheze le Rest, N. Albarghach, O. Pradier, and D. Visvikis, "PET functional volume delineation: a robustness and repeatability study," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 38, no. 4, pp. 663–672, 2011.

[31] A. Schaefer, U. Nestle, S. Kremp et al., "Multi-centre calibration of an adaptive thresholding method for PET-based delineation of tumour volumes in radiotherapy planning of lung cancer," *NuklearMedizin*, vol. 51, no. 3, pp. 101–110, 2012.

[32] A. C. Riegel, M. K. Bucci, O. R. Mawlawi et al., "Target definition of moving lung tumors in positron emission tomography: Correlation of optimal activity concentration thresholds with object size, motion extent, and source-to-background ratio," *Medical Physics*, vol. 37, no. 4, pp. 1742–1752, 2010.

[33] J. F. Daisne, M. Sibomana, A. Bol, T. Doumont, M. Lonneux, and V. Grégoire, "Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: Influence of reconstruction algorithms," *Radiotherapy and Oncology*, vol. 69, no. 3, pp. 247–250, 2003.

[34] P. Tylski, S. Stute, N. Grotus et al., "Comparative assessment of methods for estimating tumor volume and standardized uptake value in ^{18}F-FDG PET," *The Journal of Nuclear Medicine*, vol. 51, no. 2, pp. 268–276, 2010.

[35] F. Gallivanone, A. Stefano, E. Grosso et al., "PVE correction in PET-CT whole-body oncological studies from PVE-affected images images," *IEEE Transactions on Nuclear Science*, vol. 58, no. 3, pp. 736–747, 2011.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 623236, 7 pages
http://dx.doi.org/10.1155/2015/623236

Research Article

Explicit Filtering Based Low-Dose Differential Phase Reconstruction Algorithm with the Grating Interferometry

Xiaolei Jiang,[1,2] Li Zhang,[1,2] Ran Zhang,[1,2] Hongxia Yin,[3] and Zhenchang Wang[4]

[1] *Key Laboratory of Particle & Radiation Imaging (Tsinghua University), Ministry of Education, Beijing 100084, China*
[2] *Department of Engineering Physics, Tsinghua University, Beijing 100084, China*
[3] *Department of Radiology, Beijing Tongren Hospital, Capital Medical University, Beijing 100730, China*
[4] *Department of Radiology, Beijing Friendship Hospital, Capital Medical University, Beijing 100050, China*

Correspondence should be addressed to Li Zhang; zli@mail.tsinghua.edu.cn

Received 15 September 2014; Revised 8 December 2014; Accepted 20 December 2014

Academic Editor: Yi Gao

X-ray grating interferometry offers a novel framework for the study of weakly absorbing samples. Three kinds of information, that is, the attenuation, differential phase contrast (DPC), and dark-field images, can be obtained after a single scanning, providing additional and complementary information to the conventional attenuation image. Phase shifts of X-rays are measured by the DPC method; hence, DPC-CT reconstructs refraction indexes rather than attenuation coefficients. In this work, we propose an explicit filtering based low-dose differential phase reconstruction algorithm, which enables reconstruction from reduced scanning without artifacts. The algorithm adopts a differential algebraic reconstruction technique (DART) with the explicit filtering based sparse regularization rather than the commonly used total variation (TV) method. Both the numerical simulation and the biological sample experiment demonstrate the feasibility of the proposed algorithm.

1. Introduction

X-ray grating interferometry [1–4] with conventional X-ray tubes develops rapidly in recent years and is becoming the most promising technology among various phase contrast imaging methods for clinical applications. Three kinds of information, that is, the attenuation, differential phase contrast (DPC), and dark-field images, can be obtained through one single scanning, and the latter two images provide additional and complementary information to the conventional attenuation image. The DPC method measures phase shifts of X-rays by obtaining the line integral of the directional derivatives of refractive index decrements (δ) [5], that is, the refraction angle of the beam. The refraction index reconstructed afterwards is 1000 times larger than the absorption index.

The phase-stepping approach of the grating interferometry, which requires a number of images to retrieve information, significantly increases the examine time and the dose delivered to the patient [6]. The problem becomes even more severe for DPC-CT because of the requirement of multiangle scanning. Therefore, reducing the number of projections, the exposure time, and the delivered dose is of great value. And that is why the low dose DPC reconstruction algorithm is proposed.

As mentioned above, the reconstruction problem for DPC-CT is to obtain the refraction index from the refraction angle data. The analytical method, the filtered backprojection (FBP) algorithm with the Hilbert transform, was first applied [7, 8]. Afterwards, several iterative algorithms, such as the maximum likelihood (ML) algorithm [9] and the differential algebraic reconstruction technique (DART) [10], were proposed. However, these algorithms rely on the completeness of data and the large number of projections.

The recently proposed compressed sensing (CS) theory [11] makes image reconstruction from incomplete data possible. Essentially, it illustrates that if the image is sparse in a domain which has small coherence with the sampling domain, according to the Shannon/Nyquist sampling theorem, fewer projections can almost accurately recover

the images. A typical image reconstruction method exploits TV as the sparse regularization [12] (from CS measurements). Applications in both the absorption imaging [12] and the DPC imaging [10] have been implemented. Instead of the implicit regularization coming from the penalty, another sparse regularization method based on explicit filtering is proposed, which exploits spatially adaptive filters sensitive to image features and details [13]. However, no similar algorithms for DPC imaging have been suggested so far, which is just the problem to be solved in this paper.

In this work, we propose an explicit filtering based low-dose differential phase reconstruction algorithm. The algorithm combines the DART iterative algorithm and the explicit filtering based CS method. It has the potential to exactly reconstruct the refractive index distribution using few-view projections, thus reducing the exposure time and the delivered dose, making DPC-CT closer to clinical applications. The feasibility of the low dose reconstruction algorithm is verified by both the numerical simulation and the biological sample experiments.

2. Methods

2.1. Grating-Based Imaging. Figure 1 illustrates the schematic diagram of a typical grating interferometry. Two kinds of apparatuses are shown, the Talbot effect based interferometry with coherent source, that is, Figure 1(a), and the Talbot-Lau effect based interferometry with incoherent source, that is, Figure 1(b). The first grating G1 creates its self-image through Talbot-Lau effect or classical optics in the position of G2 where Moire fringes occur. The source grating G0 splits the source into an array of line sources, enabling the use of the large-focal-spot X-ray tube, that is, the incoherent source. The phase-stepping approach is adopted for image acquisition, capturing a series of raw images at every step of one of the gratings along the transverse direction, obtaining the intensity oscillation curve, Figure 1(c). The changes of the background oscillation curves determine three kinds of information, namely, the attenuation image, the DPC image, and the dark-field image.

To analyze the changes quantitatively, the oscillation curve for each pixel is expressed by the Fourier expansion series:

$$I\left(m,n,x_g\right) = \sum_i a_i\left(m,n\right)\cos\left(\frac{2\pi x_g}{p_2} + \phi_i\left(m,n\right)\right) \approx a_0\left(m,n\right) + a_1\left(m,n\right)\cos\left(\frac{2\pi x_g}{p_2} + \phi_1\left(m,n\right)\right), \tag{1}$$

where a_i is the amplitude coefficient, ϕ_i is the corresponding phase coefficient, and p_2 is the period of G2. Then, the attenuation, dark-field, and DPC images are given by $T = -\log(a_0^s/a_0^r)$, $S = -\log(V_s/V_r)$, and $\mathrm{DP} = \phi_1^s - \phi_1^r$, respectively, where the superscripts (s) and (r) denote the values with the sample in place and as a reference without, respectively, and $V = a_1/a_0$ is the visibility of the oscillation curve.

2.2. Differential Phase-Contrast Reconstruction Algorithm. The DPC image measured by the grating interferometer is the refraction angle, which is related to the phase shift $\Phi(r)$ of the sample:

$$\mathrm{DP} = \Delta\theta \approx \frac{\lambda}{2\pi}\frac{\partial\Phi\left(r\right)}{\partial r}. \tag{2}$$

In other words, the grating interferometer obtains the line integral of the directional derivatives of refraction index decrements. Therefore, the reconstruction problem of DPC-CT can be expressed by

$$y_\theta = \int \frac{\partial\delta\left(i,j\right)}{\partial l'}dl, \tag{3}$$

where y_θ is the refraction angle projection, δ is the refraction index decrement of the samples, l is the path of X-ray beam in the medium and l' is the perpendicular direction to l.

By contrast, the projection of the conventional X-ray transmission comes from the linear integration of attenuation coefficient; hence, the reconstruction problem can be expressed by

$$y_I = \int \mu\left(x,y\right)dl, \tag{4}$$

where y_I is the intensity projection and μ is the linear attenuation coefficient.

Such difference in mathematical expressions of the projections requires different reconstruction algorithms.

Inspired by the widely used algebraic reconstruction technique (ART) [14], Wang et al. proposed a differential algebraic reconstruction technique (DART) by discretizing the projection process of the differential phase contrast imaging into a linear partial derivative matrix [10]:

$$y_\theta = \int \frac{\partial\delta\left(i,j\right)}{\partial l'}dl \approx \sum \frac{\partial\delta\left(i,j\right)}{\partial l'} = \left\langle b_i, x\right\rangle, \tag{5}$$

where b_i denotes the net interpolation coefficient corresponding to each pixel.

Therefore, the forward projection process of differential phase contrast imaging can be expressed by

$$\Delta\theta = Bx, \tag{6}$$

where B is named as the linearly partial-derivative matrix. Equation (3) can be used to reconstruct the refractive index by algorithms similar to ART directly shown as (4):

$$x^{k+1} = x^k + \frac{y_\theta - Bx^k}{\|B\|}B^T, \tag{7}$$

where x^k is the image vector x in the kth iteration; Bx^k presents the forward projection process.

However, the DART algorithm relies on the completeness of data [10]. It is incompetent at the ill-posed reconstruction problems, such as the cases of few-view or limited-angle projections.

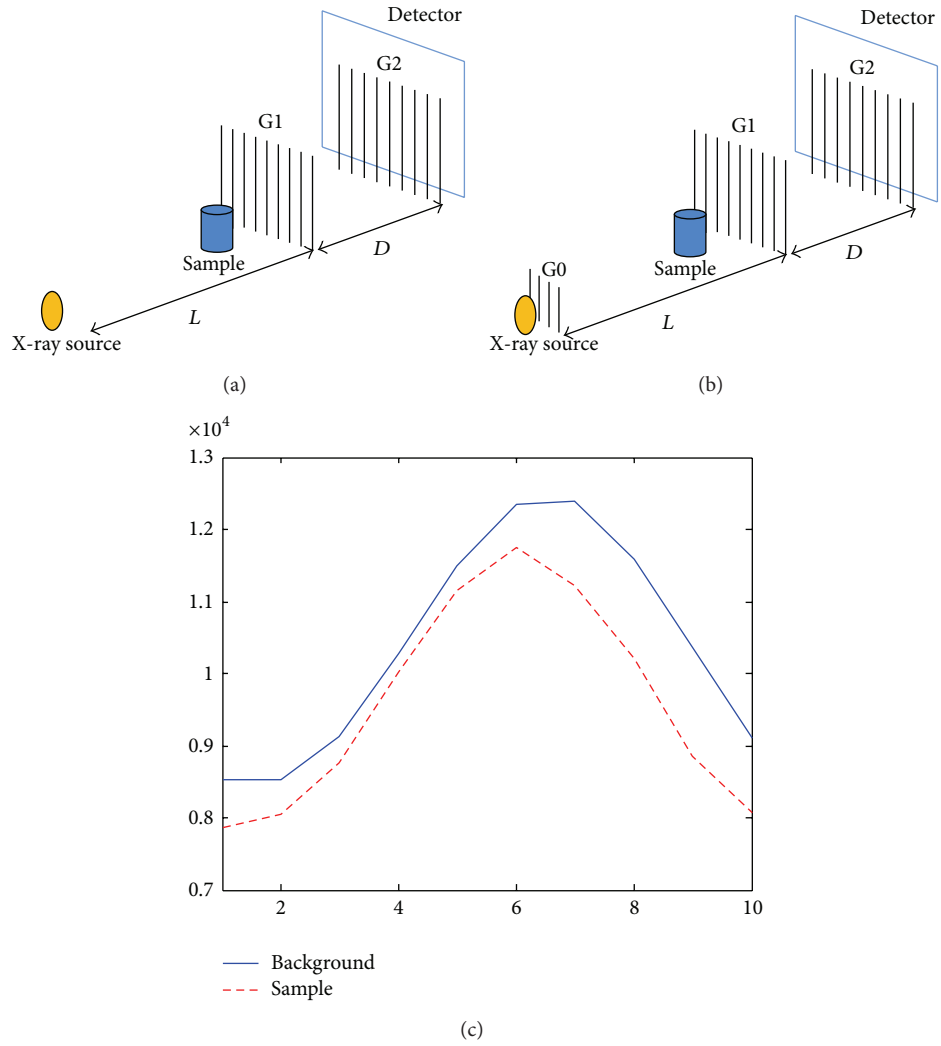


FIGURE 1: The grating interferometry. (a) Talbot effect based interferometry with coherent source; (b) Talbot-Lau effect based interferometry with incoherent source; the sample intensity oscillation curve (the dash line curve) and the background intensity oscillation curve (the solid line curve) measured with 10 steps during the phase stepping approach with the grating interferometer in Tsinghua University, China.

2.3. Explicit Filtering Based Low-Dose Differential Phase-Contrast Reconstruction Algorithm. In this part, we propose an explicit filtering based low-dose differential phase-contrast reconstruction algorithm based on the above DART algorithm for DPC reconstruction.

Generally, the compressed sensing DPC-CT reconstruction method can be summarized as

$$\widehat{x} = \arg\min_{x} \left\{ \left\| y_\theta - Bx \right\|^2 + \lambda\varphi(x) \right\}. \tag{8}$$

Here, the fitness function $\|y_\theta - Bx\|^2$, which tries to match the estimation to the data, is accomplished by the DART algorithm above, where the operation $\|\cdot\|$ represents the l-2 norm. The regularization function $\varphi(x)$ expresses some priori known property of the unknown object, which usually defines a sparse representation of x after a specific transformation. The regularization parameter λ balances the fitness function with the regularization function. Different regularization methods are used to find the solution of the mathematic model under different constraints of minimizing the complexity of reconstructed signals in different representations. The typical TV-type regularization method uses the total variations as the sparse transformation, which is a parametric regression technique. Another solution is to replace the parametric regression by spatially adaptive filters, which is sensitive to image features and details. In this paper, the BM3D filter is adopted, that is, DART_BM3D algorithm.

The BM3D algorithm is based on an enhanced sparse representation in transform domain [15]. The enhancement of the sparsity is achieved by grouping similar 2D fragments of the image into 3D data arrays. And then the collaborative filtering with the hard thresholding is carried out, which enhances the similarity between the blocks while at the same time preserves even the finest details with their essential unique features shared by the jointly filtered 2D fragments. The implementation of BM3D filter is shown as Figure 2.

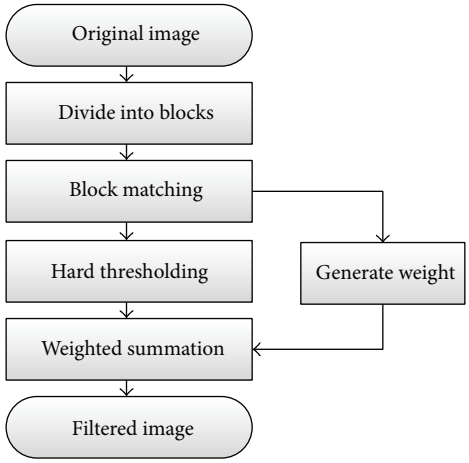


Figure 2: The flow chart of BM3D filter.

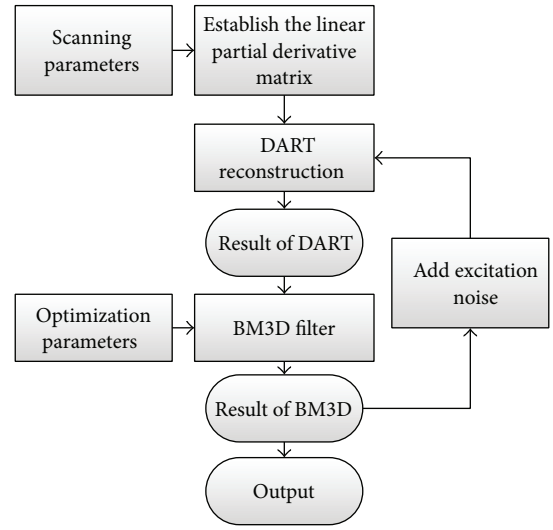


Figure 3: The flow chart of the explicit filtering based low-dose differential phase-contrast reconstruction algorithm.

BM3D filter matches the similar block regions in the image (block matching, BM) using the similarity defined as

$$d\left(Z_1 - Z_2\right) = \frac{\left\|Z_1 - Z_2\right\|_2^2}{N_{\text{mat}}^2}, \tag{9}$$

where Z is the numerical metric of the image, N_{mat} is the side length of block, and $d(Z_1 - Z_2)$ is the difference between two blocks.

Blocks with small differences can be grouped as a set S. For blocks in the same group, a hard threshold is used to assign zeros for small value pixels to generate a weight for blocks:

$$w_Z^S = \begin{cases} \dfrac{1}{\sigma^2 N_{\text{har}}^2} & N_{\text{har}} \neq 0 \\ 1 & N_{\text{har}} = 0, \end{cases} \tag{10}$$

where N_{har} is the number of nonzero pixels. Then, the filtered image can be obtained as a similarity-weighted average:

$$\lambda(x) = \frac{\sum_{x \in Z} \sum_{Z \in S} w_Z^S Z(x)}{\sum_{x \in Z} \sum_{Z \in S} \chi(x)}, \quad \chi(x) = \begin{cases} 1 & Z(x) \neq 0 \\ 0 & Z(x) = 0. \end{cases} \tag{11}$$

The process of our reconstruction algorithm is shown in Figure 3 and the steps are described as follows.

(A) Initialization. Creating the linearly partial-derivative matrix B for few-view scanning based on the geometry parameters. An initial guess of the reconstruction image $\widehat{x}_0 = 0$ is given and $n = 1$.

(B) DART reconstruction:

$$\widehat{x}^{k+1} = \text{DART}\left(\widehat{x}_k\right), \tag{12}$$

where DART represents the reconstruction method in (4).

(C) BM3D filter:

$$\widehat{x}_{k+1} = \text{BM3D}\left(\widehat{x}^{k+1}\right), \tag{13}$$

where BM3D represents the method shown as in Figure 2.

(D) Add excitation noise into $\widehat{x}_{k+1}$ and go back to step (B).

Gaussian noise is added in the unobserved portions, that is, the missing angles caused by the few-view or limited-angle scanning, in frequency domain, which works as a random generator of the missing components in the spectrum:

$$\widehat{x}_{k+1} = \text{FFT}^{-1}\left(\text{FFT}\left(\widehat{x}_{k+1}\right) + \text{Gaussian}\left(0, \sigma\right)\right), \tag{14}$$

where FFT represents the Fourier transform operation and FFT^{-1} is the inverse Fourier transform operation, σ is the standard deviation of the Gaussian noise which is determined empirically by the noise of the projections caused by the quantum noise and electronics noise.

3. Experiments and Results

The proposed explicit filtering based low-dose differential phase-contrast reconstruction algorithm was validated by both the numerical simulation and biological experiments.

3.1. Numerical Simulation. The Shepp-Logan phantom with the resolution of 256 × 256 was used in the study. The phantom is shown as in Figure 4(a) and the reconstructed results are shown as in Figures 4(b)–4(g). All images are shown in the same display window [0, 1].

Figures 4(b) and 4(c) are the reconstructed results with the FBP and DART (500 iterations) methods, respectively, with 180 views within 180° with 1° angular interval, which verified the effectiveness of the FBP and DART methods in case of complete data. Figures 4(d) and 4(e) are the reconstructed results with the FBP and DART, respectively, methods with 10 views within 180° with 18° angular interval, with an angular downsampling factor of 18. The results

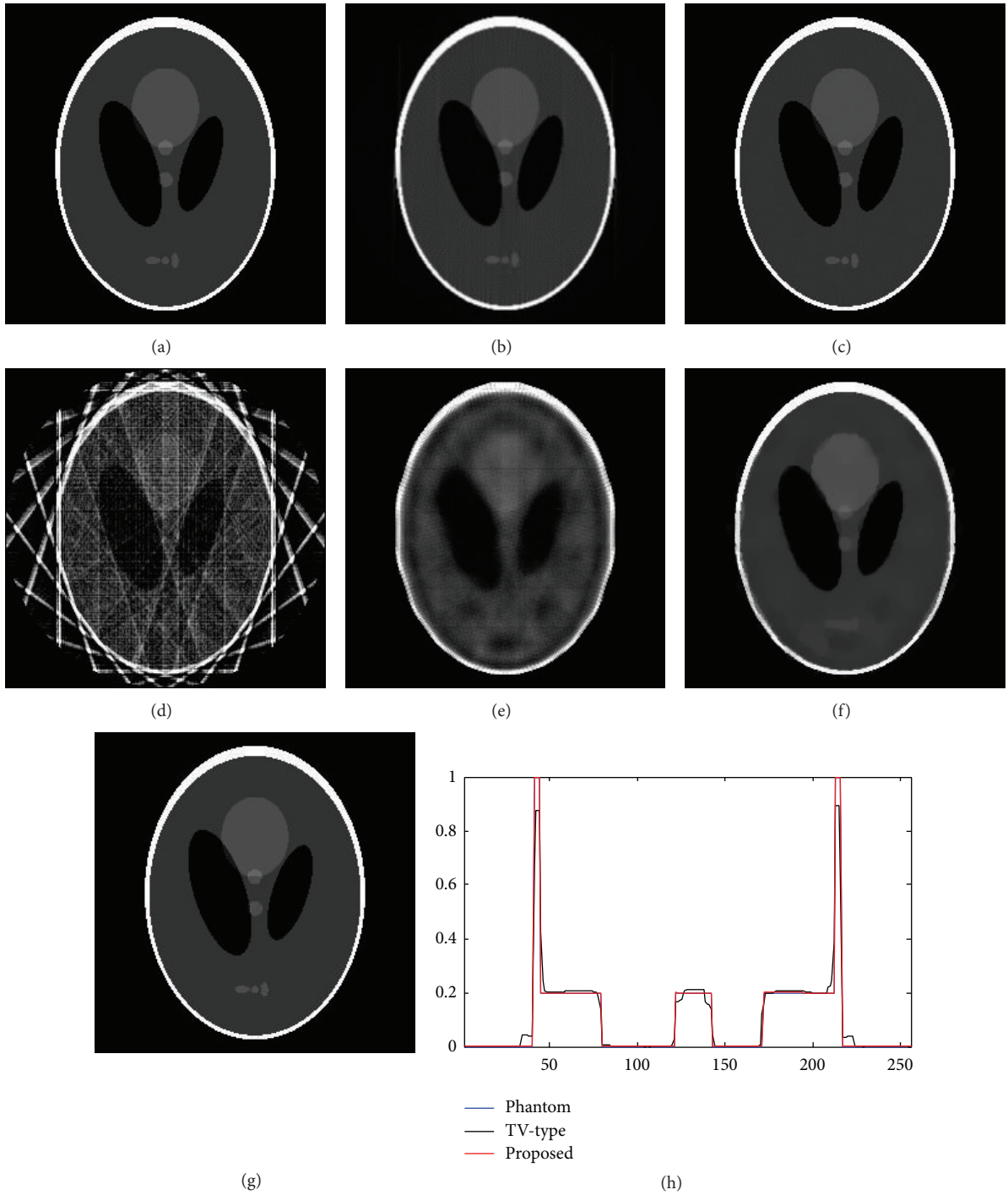


Figure 4: The reconstruction results. (a) The phantom; (b) the result of FBP with 180 views, MSE = $4.51e-3$; (c) the result of DART with 180 views, 500 iterations, MSE = $2.84e-5$; (d) the result of FBP with 10 views, MSE = $5.15e-2$; (e) the result of DART with 10 views, 2000 iterations, MSE = $5.80e-3$; (f) the result of DART_TV with 10 views, 2000 iterations, MSE = $1.63e-3$; (g) the result of the proposed algorithm with 10 views, 2000 iterations, MSE = $4.62e-6$; (h) the profiles of the 128th line of (a), (f), and (g).

demonstrate the dependency on the completeness of the data of the FBP and DART (2000 iterations) methods.

Figure 4(f) is the reconstructed result with the typical TV-type compressed sensing method (2000 iterations). The method is an effective reconstruction method dealing with incomplete data. Figure 4(g) is the reconstructed results with proposed explicit filtering based compressed sensing methods (2000 iterations), which is also skilled in dealing with incomplete data. The reconstructed results with the compressed sensing methods are almost artifact-free and

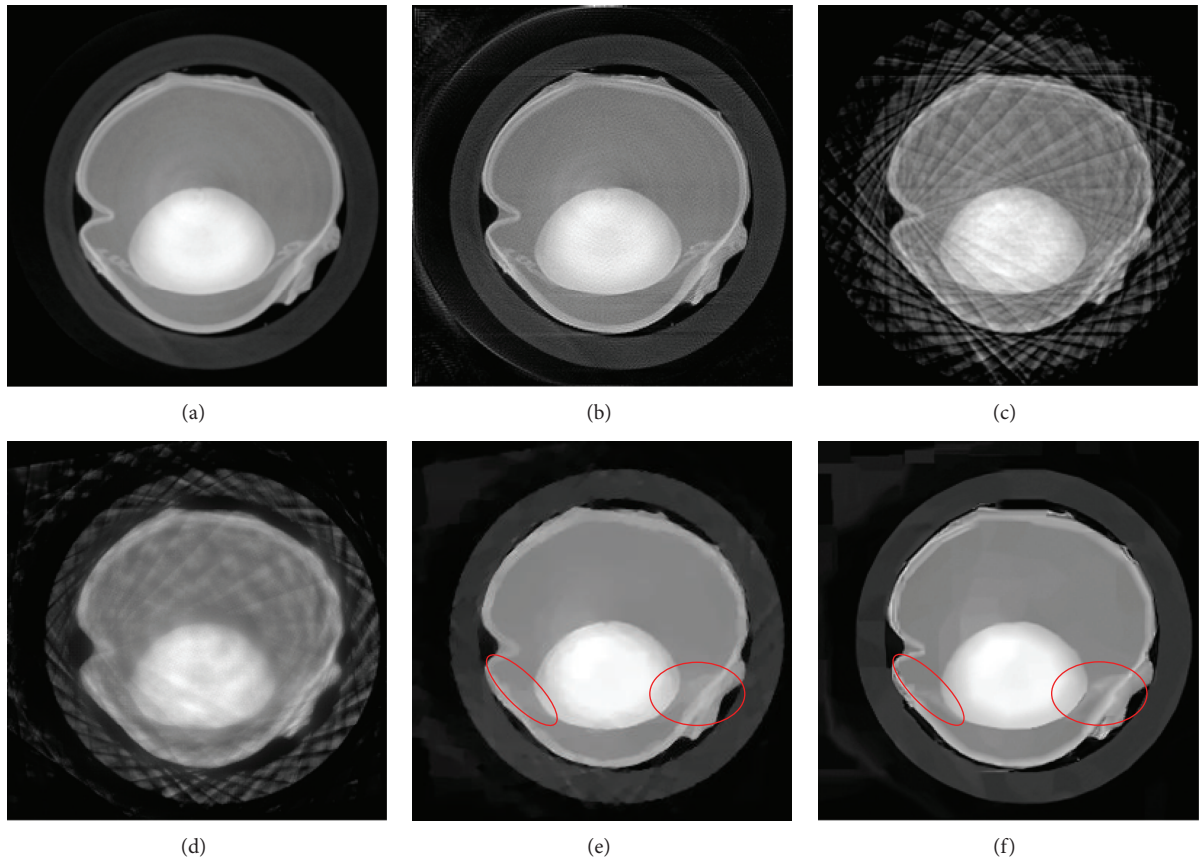


Figure 5: The reconstruction results. (a) the result of FBP with 180 views; (b) the result of DART with 180 views; (c) the result of FBP with 10 views; (d) the result of DART with 20 views; (e) the result of DART_TV with 20 views; (f) the result of the proposed algorithm with 20 views.

they are in high accordance with the phantom in both appearance and values, while the results of FBP and DART algorithms show severe streaking artifacts due to the downsampling. Furthermore, profiles as shown in Figure 4(h) show that the proposed method is better than the TV-type method in preserving details.

3.2. Biological Sample Experiment. The biological sample results have been used to test the proposed reconstruction method. The experiments were performed at the TOMCAT beamline using a two-grating interferometer operated at 25 KeV and in the 3rd Talbot distance at the Swiss Light Source of the Paul Scherrer Institute in Switzerland. The pitch of the phase grating p_1 was 3.981 μm with a height of h_1 = 31.7 μm. The corresponding values for the second grating (gold absorber grating) were p_2 = 2.00 μm and h_2 = 24 μm. The sample was a guinea pig eyeball packed in a plastic pipe and was scanned within 180° with equivalent angular interval of 1°. For each view, an eight-step phase stepping process was adopted and the refraction angular projections were retrieved by information retrieving algorithm.

The reconstruction results are shown in Figure 5. Figures 5(a) and 5(b) are the results of FBP algorithm and DART algorithm with 180 views, respectively. Since the reconstruction image is more complex than the phantom in the previous section, an angular down-sampling factor of 9 was adopted. Figures 5(c) and 5(d) is the results of FBP algorithm and DART algorithm with 20 views, respectively. The results show severe streaking artifacts, especially the FBP result. As shown in Figures 5(e) and 5(f), the results of the two compressed sensing methods are nearly artifact-free. Both the results of the two compressed sensing methods show great image quality improvement compared with DART and FBP in few-view reconstruction. And the red ellipses in Figures 5(e) and 5(f) illustrate the better capabilities of preserving details of the proposed explicit filtering based method than the TV-type method, which can also be seen from Figure 6, which shows the profiles of the 188th line in the red circles on the right.

4. Conclusions

In this paper, an explicit filtering based low-dose differential phase-contrast reconstruction algorithm is proposed. The algorithm is an application of the compressed sensing theory and has the potential to accurately reconstruct the distribution of the refractive index with few-view projections.

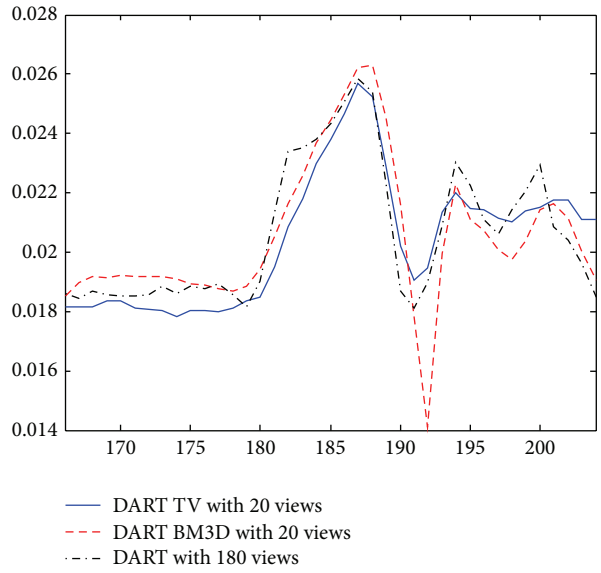


FIGURE 6: The profiles of the 188th line in the red circles on the right.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank Marco Stampanoni for the help with biological sample experiments. The work was supported by a grant from the National Natural Science Foundation of China (NNSFC), Grants 11235007, and a Tsinghua University Independent Research Project Grant, "Research on Key Technologies and CT Reconstruction Methods of Multienergy X-Ray imaging".

References

[1] C. David, B. Nöhammer, H. H. Solak, and E. Ziegler, "Differential x-ray phase contrast imaging using a shearing interferometer," *Applied Physics Letters*, vol. 81, no. 17, pp. 3287–3289, 2002.

[2] A. Momose, S. Kawamoto, I. Koyama, Y. Hamaishi, K. Takai, and Y. Suzuki, "Demonstration of X-ray Talbot interferometry," *Japanese Journal of Applied Physics*, vol. 42, no. 7, pp. L866–L868, 2003.

[3] F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David, "Phase retrieval and differential phase-contrast imaging with low-brilliance X-ray sources," *Nature Physics*, vol. 2, no. 4, pp. 258–261, 2006.

[4] Z. Huang, K. Kang, L. Zhang et al., "Alternative method for differential phase-contrast imaging with weakly coherent hard x rays," *Physical Review A*, vol. 79, no. 1, Article ID 013815, 2009.

[5] A. G. Michette, *X-Ray Science and Technology*, CRC Press, 1993.

[6] Z. Wang, Z. Huang, Z. Chen et al., "Low-dose multiple-information retrieval algorithm for X-ray grating-based imaging," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 635, pp. 103–107, 2011.

[7] Z. Huang, K. Kang, Z. Li et al., "Direct computed tomographic reconstruction for directional-derivative projections of computed tomography of diffraction enhanced imaging," *Applied Physics Letters*, vol. 89, no. 4, Article ID 041124, 2006.

[8] G.-H. Chen and Z. Qi, "Image reconstruction for fan-beam differential phase contrast computed tomography," *Physics in Medicine and Biology*, vol. 53, no. 4, pp. 1015–1025, 2008.

[9] Y. J. Liu, P. P. Zhu, B. Chen et al., "A new alterative algorithm to reconstruct the refractive index," *Physics in Medicine and Biology*, vol. 52, no. 12, pp. L5–L13, 2007.

[10] Z. Wang, Z. Huang, L. Zhang et al., "Low dose reconstruction algorithm for differential phase contrast imaging," *Journal of X-Ray Science and Technology*, vol. 19, no. 3, pp. 403–415, 2011.

[11] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[12] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-Ray Science and Technology*, vol. 14, no. 2, pp. 119–139, 2006.

[13] K. Egiazarian, A. Foi, and V. Katkovnik, "Compressed sensing image reconstruction via recursive spatially adaptive filtering," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 1, pp. I549–I552, September 2007.

[14] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York, NY, USA, 1999.

[15] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 906452, 8 pages
http://dx.doi.org/10.1155/2015/906452

Research Article

Few-View Prereconstruction Guided Tube Current Modulation Strategy Based on the Signal-to-Noise Ratio of the Sinogram

Ming Chang,[1,2] Yongshun Xiao,[1,2] and Zhiqiang Chen[1,2]

[1] *Department of Engineering Physics, Tsinghua University, Beijing 100084, China*
[2] *Key Laboratory of Particle & Radiation Imaging, Tsinghua University, Ministry of Education, Beijing 100084, China*

Correspondence should be addressed to Zhiqiang Chen; czq@mail.tsinghua.edu.cn

Received 23 September 2014; Revised 27 December 2014; Accepted 28 December 2014

Academic Editor: Yi Gao

The radiation dose reduction without sacrificing the image quality as an important issue has raised the attention of CT manufacturers and different automatic exposure control (AEC) strategies have been adopted in their products. In this paper, we focus on the strategy of tube current modulation. It is deduced based on the signal-to-noise (SNR) of the sinogram. The main idea behind the proposed modulation strategy is to keep the SNR of the sinogram proximately invariable using the few-view reconstruction as a good reference because it directly affects the noise level of the reconstructions. The numerical experiment results demonstrate that, compared with constant tube current, the noise distribution is more uniform and the SNR and CNR of the reconstruction are better when the proposed strategy is applied. Furthermore it has the potential to distinguish the low-contrast target and to reduce the radiation dose.

1. Introduction

X-ray CT has played an important role in primary diagnostic imaging and radiotherapy since its introduction in 1973. It is estimated that 67 million CT examinations were performed in 2006 in the USA while the number was about 3 million in 1980 [1]. With the increasing utilization of CT, the radiation dose and corresponding potential risks associated with CT scanning raise the ongoing concern to both the patients and CT manufactures. According to the as low as reasonably achievable principle (ALARA), the radiation dose reduction is an important issue in clinical routine, methodology research, and system development.

Various factors, including system scanning parameters, the difference of the patients, and the requirement for the following diagnoses, have influences on the CT imaging dose. The CT dose index (CTDI) is a measure of the absorbed dose to a standard plastic phantom, which is commonly used as the CT dose metric [2, 3]. It is affected by the scanning mode, exposure time, tube current, tube potential, field of measurement, and beam shape filtering [2, 4–6]. Generally it is directly proportional to the tube current and the exposure time. It is approximately proportional to the square of the percentage change in tube potential. These parameters can be adjusted flexibly according to the specific imaging task. Therefore the adjustment strategy of these parameters according to the specific applications is the key to achieve a lower CTDI in practical CT examinations.

The automatic exposure control (AEC) technology is to automatically adapt the tube current or tube potential according to patient's attenuation to achieve a specified image quality. In the currently used commercial CT system, the AEC technologies based on different strategies are adopted [7]. Automatic tube potential selection is to choose the tube potential according to the patient size or image task in order to achieve the desired image quality with a lower CTDI. Angular and longitudinal current modulation are to adjust the tube current according to the patient size, shape, and attenuation changes at different projection views [4]. Different vendors provide strategies based on slightly different principles [8]. GE adopts AutomA and SmartmA theory based on the noise index (NI) of the reference image, which is used to control the average image noise level [9, 10]. It is possible to achieve clinically acceptable images at the

lowest radiation dose to patients using the optimal NI selection and model-based iterative reconstruction (MBIR) method [11]. Philips offers the DoseRight technology and uses a reference image concept. Quality reference mAs and reference standard deviation are selected by Siemens and Toshiba, respectively [7]. However the adjustment based on the online feedback can be only used for helical scanning [12] and the strategy based on the predictive calculation or sinusoidal-type function may be far apart from the real situation. To our knowledge, most currently used strategies are designed based on the guideline of image quality and adaptive statistical iterative reconstruction (ASIR) or MBIR methods are necessary [11, 13]. As we know, the noise level of the sinogram directly affects the CT image quality no matter what reconstruction method is used. Therefore the SNR of the sinogram is considered when the proposed method is designed.

In this study, we focus on the tube current modulation strategy in the CT scanning. In order to achieve a better image quality of the reconstruction, we propose the few-view prereconstruction guided tube current modulation strategy. It is established based on the analysis of the noise in the sinogram, which directly affects the final image quality. The main idea behind the proposed strategy is to keep the SNR of the sinogram proximately invariable at different angle views using the few-view prereconstruction as a good reference for the adjustment of the tube current. The rest of the paper is organized as follows. In the next section, the proposed strategy derived from the noise analysis is introduced. A specific workflow using the proposed strategy to enhance the image quality of the reconstructed images is also presented in this section. In the third section, numerical experiments are carried out and qualitative and quantitative results are shown correspondingly. In the end, the conclusions are made for this work.

2. Materials and Methods

2.1. Tube Current Modulation Strategy Based on the SNR of the Sinogram. According to the Lambert-Beer law, the ideal attenuated X-ray photon I is expressed as

$$I = I_0 \exp(-p), \tag{1}$$

where I_0 is the initial photon and p is the integration of the linear attenuation coefficients along the X-ray path.

Generally, the quantum noise and the system electronic noise exist in the practical CT measurements. The system electronic noise should be taken into account in the case of low-dose CT [14]. However, in this work, only the Poisson distributed quantum noise is considered in the following analysis. With the Poisson statistics, the practical measurement I_m is presented in

$$I_m = \text{Poisson}(I) \approx I + \sqrt{I}x(0,1), \tag{2}$$

where $x(0,1)$ represent the random value, which is satisfied with the standard normal distribution.

Before CT reconstruction, the measurements are converted into the sinogram p_m by the negative logarithm operation:

$$p_m = -\log\left(\frac{I_m}{I_0}\right) = p - \log\left[1 + \frac{1}{\sqrt{I}}x(0,1)\right]. \tag{3}$$

Then the noise of the sinogram Δp is presented as follows:

$$\Delta p = |p_m - p| \approx \frac{1}{\sqrt{I}}x(0,1). \tag{4}$$

The SNR of the sinogram is defined as the ratio of the mean value to the standard variance, which should be expressed as

$$\text{SNR}_p = \frac{p}{\sigma_{\Delta p}} = \frac{p\sqrt{I_0}}{\exp(p/2)} = k. \tag{5}$$

If the p is a variable in (5), the SNR_p achieves a global minimization by making $p = 2.0$. It can be used for the tube potential modulation. However, we only focus on the current modulation strategy in this work. The p is no longer a variable after the tube potential is set for a specific imaging task. To achieve a reconstructed image with uniformly distributed noise, the SNR_p should be a constant k or at least should not be changed significantly. This can be realized by adjusting the initial intensity I_0 according to the attenuation coefficient p. Based on such principle, the tube current modulation strategy can be deduced from (5):

$$I_0 = \frac{k^2 \exp(p)}{p^2}. \tag{6}$$

Equation (6) provides the basis of tube current modulation. For a fixed total photon count I_{total}, the allocation of the photon count for each view should be done based on

$$I_{0i} = \frac{w_i}{\sum_i w_i} I_{\text{total}}, \quad \text{where } w_i = \frac{\exp(p_i)}{{p_i}^2}. \tag{7}$$

The subscript i indicates the index of the view angles. By (7), on one hand, the number of required initial photons increases with the growth of the p_i when the p_i is greater than 2.0. On the other hand, it also increases with the decrease of p_i when p_i is less than 2.0. The least amount of initial photos is needed by $p_i = 2.0$.

2.2. Strategy Implementation. However, there are still two problems to be addressed when the strategy is implemented. The first one is how to determine the weight factor w_i for each view in (7). In the practical CT scanning, the attenuations of different detector bins are generally not the same at a certain view angle and it is difficult to adjust the tube current for each detector bin. As aforementioned, the SNR of the sinogram is expected to be constant or invariable approximately. Therefore the difference of the sinogram's SNR

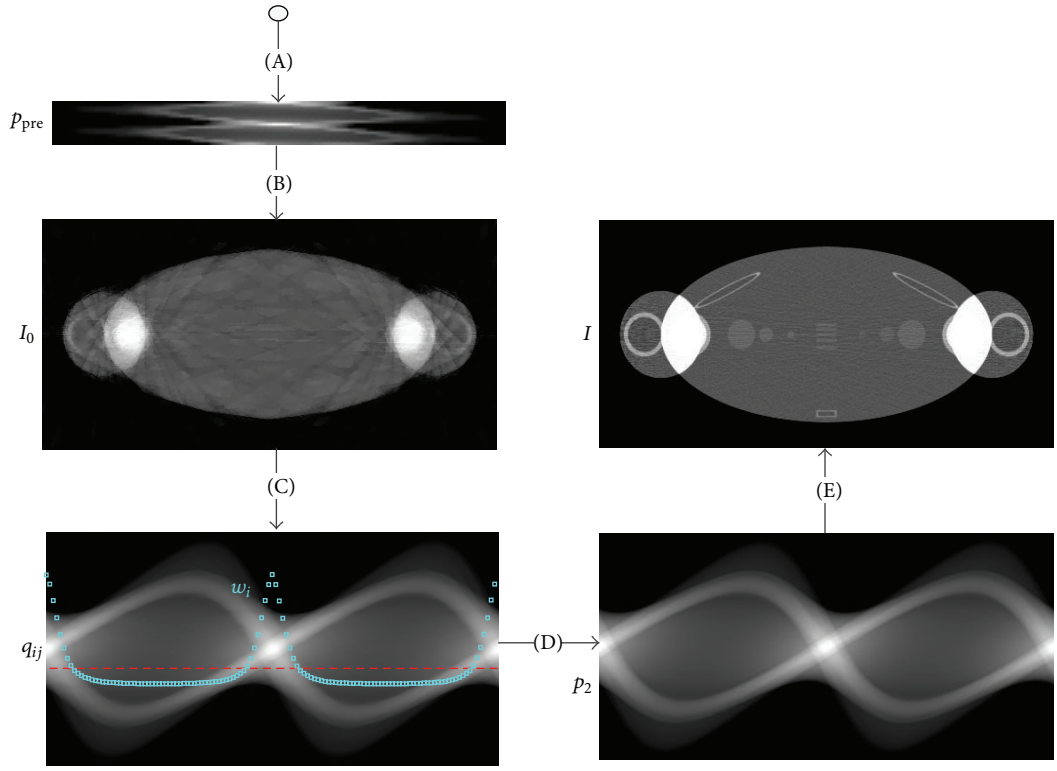


Figure 1: The complete work flow of the proposed strategy.

from the desired one within a certain projection view should be minimized:

$$p_i = \arg\min_{p_i} \left\{\Delta\text{SNR}_i\left(p_i\right)\right\},$$
$$\text{where } \Delta\text{SNR}_i\left(p_i\right) = \sum_j \left(\text{SNR}_{q_{ij}} - \text{SNR}_{p_i}\right)^2. \tag{8}$$

The q_{ij} is the attenuation coefficient obtained by jth detector bin at ith projection view. The p_i determines the SNR_{p_i} at the ith projection view. Then the weight factor w_i in (7) can be expressed as

$$w_i = \left[\frac{M}{\sum_{j=1}^{M} q_{ij} \exp\left(-q_{ij}/2\right)}\right]^2. \tag{9}$$

The M indicates the total number of the detector bins. It should be noted that the weight factor in (9) is a compromise of the high attenuation projection and the low attenuation projections in terms of the SNR. In fact, the tube current is too low for the projections of high attenuation and relatively high for the projections of low attenuation. In order to suppress the strips artifacts caused by the poor SNR of high attenuation projection, the projections q_{ij} used to calculate the weight factor w_i in (9) should be above a certain threshold. The threshold T_i is projection view related and it is determined by the median of q_{ij} $(j = 1, 2, \ldots, M)$ within the corresponding view in the following experiments. It should be noted that it is an empirical parameter and it may not be the optimal choice. But it works well in our study.

The second problem is to get the sinogram q_{ij} as a reference for the tube current modulation. However it is impossible to achieve such a sinogram before completing the scanning. In practical CT scanning, patient sizes, shapes, and compositions may differ from the assumption, which has the negative influence on the tube current modulation. In this study, a few-view prereconstruction image is used to acquire the sinogram for the determination of the weight factors in the proposed strategy. Recently the compressed sensing based reconstruction methods make it possible to achieve the reconstructed image of the acceptable image quality using few-view projection data [15–18]. The few-view reweighted sparsity hunting (FRESH) method, which is demonstrated to have good performance in the case of few-view tomography, is adopted to complete the task. You can refer to [17] for the details about the method. Then the forward projection operation is done with the few-view prereconstruction and the simulating sinograms provide a good guide for the specific modulation plan in the following routine scanning.

The complete work flow of the proposed strategy is shown in Figure 1. (A) few-view scanning to get the sinogram p_{pre}; (B) carrying out prereconstruction I_0 by FRESH method using p_{pre}; (C1) estimating the complete sinogram q_{ij} by forward projection of I_0; (C2) calculating the weight factor w_i, according to (9) using q_{ij}; (D) routine scanning with the

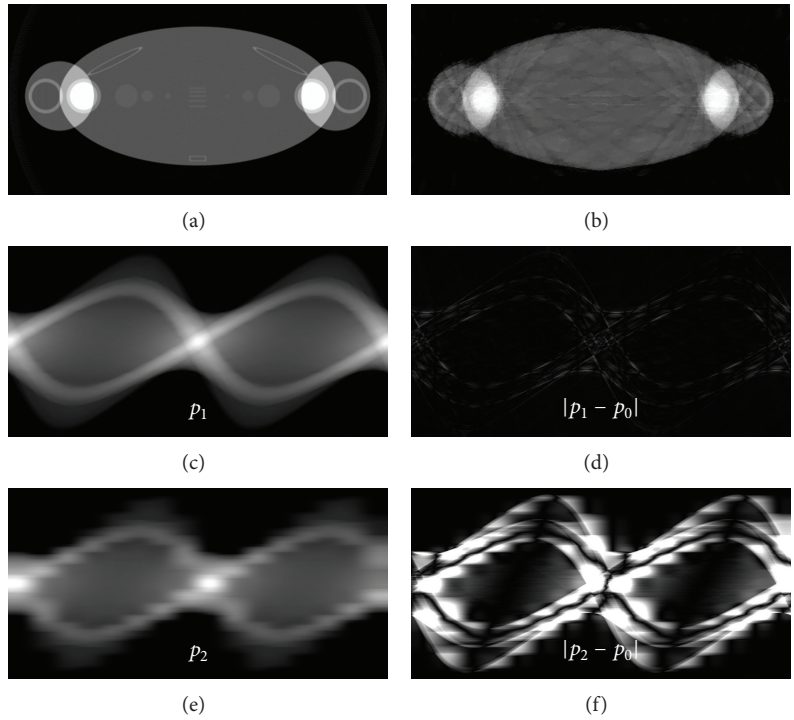


Figure 2: The thorax phantom and the sinogram used for the determination of the weight factors. (a) The thorax phantom; (b) the prereconstruction using 16-view projections; (c) the estimated sinogram using the prereconstruction p_1; (d) the difference image $|p_1 - p_0|$; (e) the estimated sinogram by interpolation p_2; (f) the difference image $|p_2 - p_0|$. The display windows are set to $[0, 0.3]$ for (a) and (b), $[0, 7]$ for (c) and (e), and $[0, 1]$ for (d) and (f).

strategy based on the weight factor w_i to get the sinogram p_2; (E) carrying out final reconstruction I with p_2.

3. Results and Discussion

In the first experiment, as shown in Figure 2(a), the thorax phantom was used to demonstrate the feasibility and effectiveness of the proposed strategy. In the phantom, there were arms, clavicle, humerus, and shoulder blades of high attenuations. Several low-contrast disks and the line-pair were placed in the center of the phantom. They were used to test the distinguish ability of the soft tissue. The initial photon intensity was set to 1.0×10^5 for each view. The geometrical configuration parameters in the simulations are listed in Table 1.

To have a clear analysis of the proposed strategy, the intermediate result is presented in the first numerical simulations. Firstly, a prereconstruction by the FRESH method was done using the uniformly distributed 16-view projections and the result is shown in Figure 2(b). It was used to estimate the sinogram for the current modulation strategy in the following routine scanning. As the results shown in Figures 2(c) and 2(d), the estimated sinogram p_1 using the prereconstruction image as a good reference was very close to the ideal one p_0. In contrast, there were significant errors in the direct interpolations' result p_2 using the few-view projections, which eventually leads to an inappropriate current modulation. To give a clear illustration of their influence on the proposed strategy, the comparison of the determined weight factors based on different references was plotted in Figure 3(a). It can be found that the prereconstruction based result was consistent with the ideal one. More photons were distributed at the projection views where the attenuation is relatively high and slightly less photons were used in the other view angles. It makes it possible to achieve a more uniformly distributed SNR of the sinogram. However the interpolation based result did not match well at some view angles due to the errors of the interpolation based sinogram. Figures 3(b)–3(d) show the reciprocal of the sinogram's SNR using different weight factors for the tube current modulations. As the white arrows indicated, compared to the other strategy, the SNR was enhanced at the projection views where the attenuation is

Table 1: The geometrical configurations in the numerical simulations.

Scanning configuration parameters	Values
Trajectory radius (cm)	40.0
Object radius (cm)	22.5
Source-to-detector distance (cm)	70.0
Projection number per circle	1024
Linear array detector size (cm)	76.8
Detector unit number	512
Reconstructed image dimensions	512×512

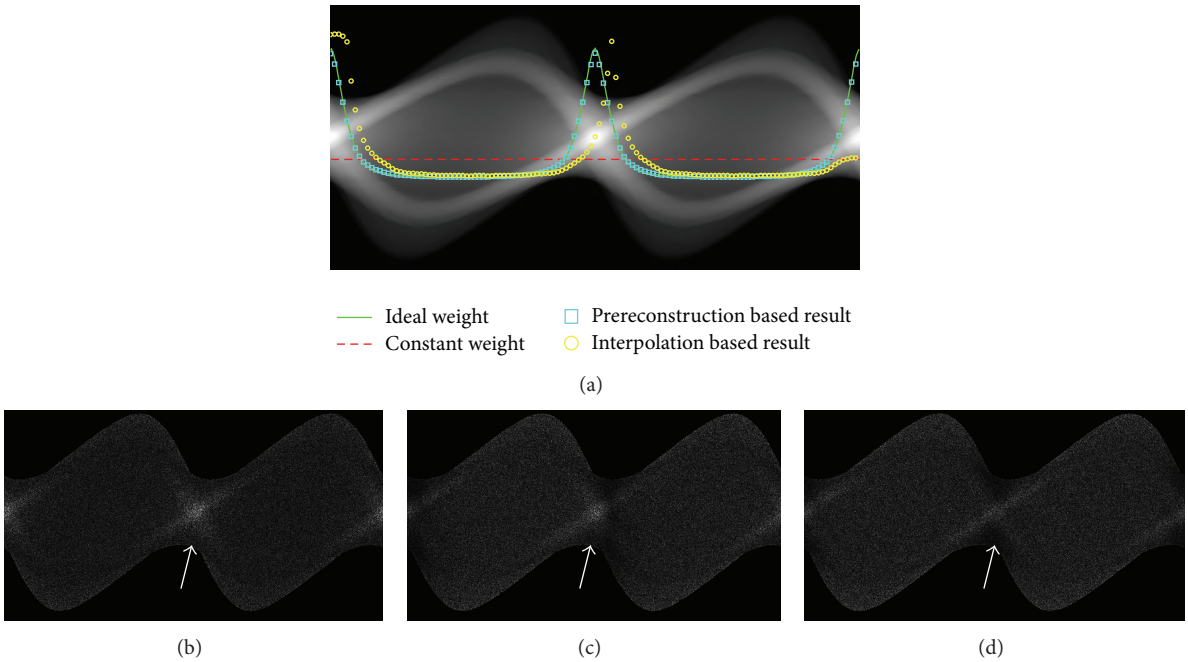


Figure 3: The weight factors based on different references for tube current modulation and the corresponding sinogram acquired based on such strategy. (a) The weight factor based on different strategies; (b)–(d) the reciprocal of the sinogram's SNR using constant weight factor and interpolation based and prereconstruction based weight factors, respectively. The display windows are set to [0, 7] for (a) and [0.00, 0.035] for (b)–(d).

relatively high by the proposed strategy, which yields a more uniformly distributed SNR of the sinogram.

Then the routine scanning with the proposed strategy was done. It was compared with the scanning using a constant tube current. However, it should be noted that the total photon of each whole scanning is the same. The comparisons in the following experiments were all based on such premise. The results in Figure 4 show that the noise distribution in the final reconstruction was more uniform using the proposed strategy than using a constant tube current or the interpolation based method. As Figure 4(a) has shown, the result of constant weight is corrupted with severe strips artifacts. Although some improvements on image quality have been achieved in the interpolation based result Figure 4(b), there were some obvious strips artifacts due to the imperfect current modulation. By contrast, the low-contrast disk and the line-pair in Figure 4(c) could be easily distinguished. To make a quantitative analysis of the image quality, the SNR and CNR of the $10 \times 10\,\text{pixel}^2$ region of interest (ROI) were calculated. The pixel size is 0.0879 cm and the dimensions of the ROI are $0.879 * 0.879\,\text{cm}^2$. The definitions are given in (10) and (11). The ROIs were labeled as R1–R4 in Figure 4(a) and they were used to calculate the SNR. The CNR was estimated based on R2 and R3. As listed in Table 2, the SNR and CNR along the lateral direction have been greatly improved due to enhancements of the sinogram's SNR in these directions:

$$\text{SNR} = \frac{\mu}{\delta}, \tag{10}$$

Table 2: The SNR and CNR of the reconstructed images in Figure 4.

Reconstructed images	Figure 4(a)	Figure 4(b)	Figure 4(c)
SNR			
R1	10.93	11.55	15.71
R2	17.26	20.01	24.63
R3	15.93	21.73	30.94
R4	23.14	28.00	27.29
CNR			
R2 and R3	1.58	1.81	2.58

where the μ and δ are the mean and the standard deviation of the ROI. Consider

$$\text{CNR} = \frac{|\mu_1 - \mu_2|}{\delta_0}, \tag{11}$$

where μ_1 and μ_2 are the mean of R1 and R2 and δ_0 is the standard deviation of the pure image noise.

In the second experiment, a more complicated dental phantom, as shown in Figure 5(a), was used to further demonstrate the effectiveness of the proposed strategy for the practical applications. It was done on the CT simulation platform developed by our group [19]. The phantom was designed based on a real CT image using B-spline curves to approximate the edges of different compositions. The X-ray spectrum of 160 kV used in this experiment was simulated by Monte Carlo method. The corresponding attenuation coefficients of various tissues, including the adipose, dentin,

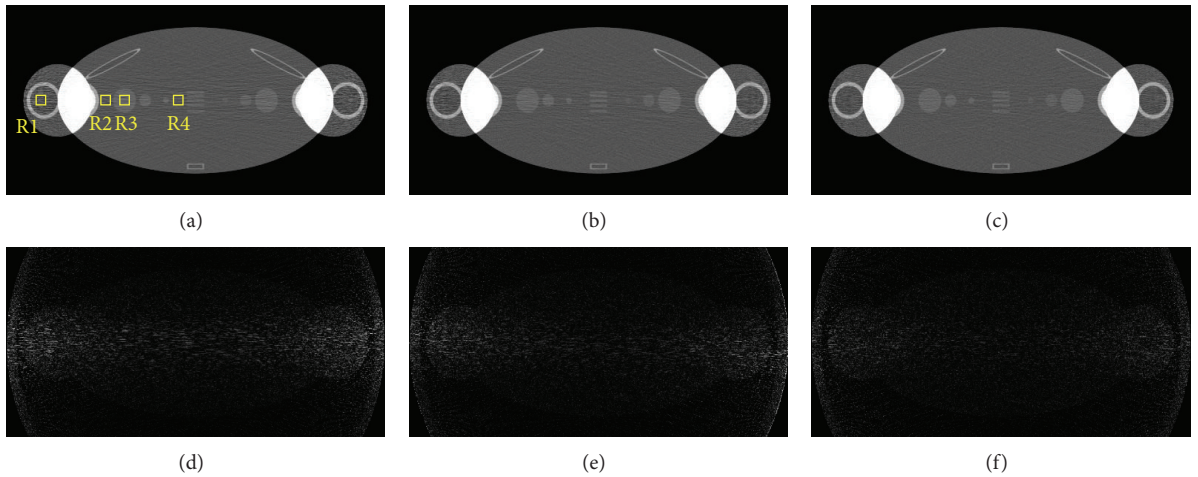


FIGURE 4: The comparisons of the reconstructions using different strategies with the thorax phantom. Reconstructions using a constant current and interpolation based strategy and the proposed strategy are shown in (a)–(c). The corresponding difference images from the phantom are presented in (d)–(f), respectively. The display windows are set to [0.05, 0.20] for (a)–(c) and [0.00, 0.05] for (d)–(f).

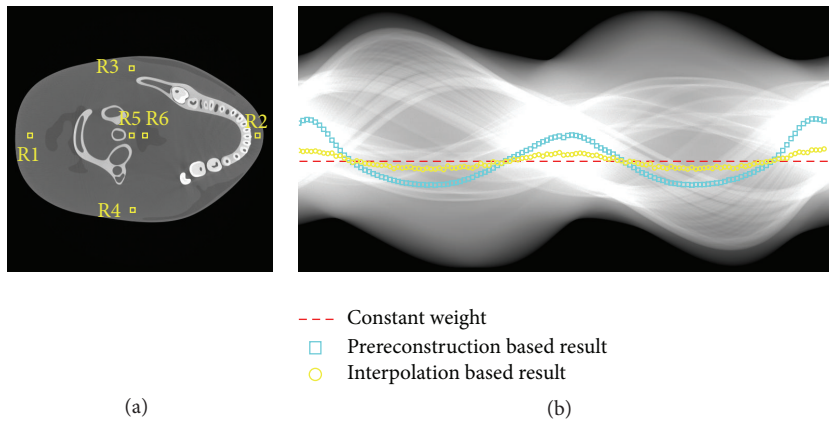


FIGURE 5: (a) The dental phantom and (b) the sinogram obtained using the proposed strategy with the weight factor based on the prereconstruction plotted on it. The display window is set to [0.0, 0.5] and [0, 7], respectively.

brain, and enamel, at different X-ray energy were obtained from the web of the National Institute of Standards and Technology (NIST). The material of the detectors was CsI and the response to different X-ray energy was also considered in our simulations. The geometrical configuration parameters in the simulations were similar to the previous set. Figure 5(b) shows the sinogram obtained by the proposed strategy. The weight factor in such strategy is determined by the prereconstruction. Similar to the results in the first experiment, the distribution of the initial photon intensity is more reasonable in terms of the sinogram's SNR. Compared with the result in the first experiment, the amplitude of the weight curve becomes small because the differences of the sinogram at different views are not as significant as they were in the first experiment.

The results using the proposed method as well as other strategies are shown in Figure 6. The distribution of the image noise in our result was more uniform, especially along the direction where the X-ray was attenuated seriously. As the arrow indicated, the edge of the low-contrast structure was easily distinguished in Figure 6(c) while it was corrupted by the noises and artifacts in Figures 6(a) and 6(b). The SNR and CNR were also calculated for the quantitative analysis in this experiment. The ROIs were labeled in Figure 5(a). As the results listed in Table 3, the SNR and the CNR increase by about 10%. The improvements in image quality make it possible to distinguish tiny low-contrast abnormal tissues with the same imaging doses. On the other hand, it has the potential to reduce the radiation doses under the same image quality.

4. Conclusions

As a conclusion, the few-view prereconstruction based tube current modulation strategy is proposed in this work. It is derived from the SNR analysis in the sinogram domain. The main idea behind the strategy is to make a more uniform distribution of the sinogram to enhance the CT

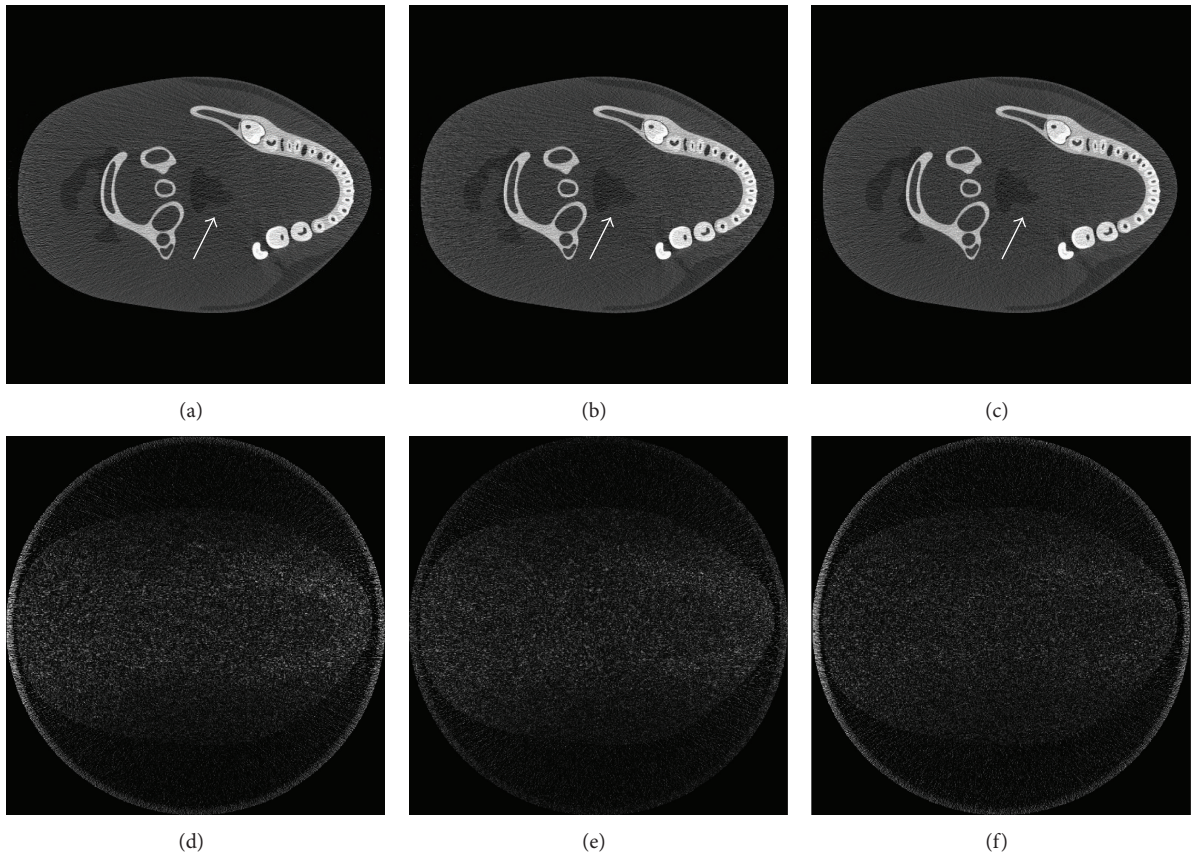


Figure 6: The comparisons of the reconstructions using different strategies with the dental phantom. Reconstructions using a constant current and interpolation based strategy and the proposed strategy were shown in (a)–(c). The corresponding difference images from the phantom were presented in (d)–(f), respectively. The display windows are set to [0.1, 0.5] for (a)–(c) and [0.00, 0.15] for (d)–(f).

Table 3: The SNR and CNR of the reconstructed images in Figure 6.

Reconstructed images	Figure 6(a)	Figure 6(b)	Figure 6(c)
SNR			
R1	6.54	7.28	8.59
R2	6.34	6.00	7.11
R3	12.62	10.88	13.43
R4	13.18	14.29	14.41
CNR			
R5 and R6	1.01	1.02	1.12

image quality. In the strategy, exact allocation of the X-ray photon at various projection views is made with the reference sinogram provided by the prereconstruction of FRESH method. Its feasibility and effectiveness have been demonstrated by the experiment results. The SNR and CNR of the final reconstruction are enhanced by more than 10% using the proposed method. In general, a lowered imaging dose results in a high noise level of CT image. But the noise can be reduced if the proposed strategy is adopted. Therefore the proposed strategy has the potential to achieve the same noise levels of CT image while reducing the overall radiation dose to patients. We will apply it to the practical CT system and make more experiments for the quantitative analysis in the future.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was partly supported by the grants from the Transportation Construction Science and Technology Project 2013 328 49A090.

References

[1] D. J. Brenner and E. J. Hall, "Computed tomography—an increasing source of radiation exposure," *The New England Journal of Medicine*, vol. 357, no. 22, pp. 2277–2284, 2007.

[2] C. Zacharias, A. M. Alessio, R. K. Otto et al., "Pediatric CT: strategies to lower radiation dose," *The American Journal of Roentgenology*, vol. 200, no. 5, pp. 950–956, 2013.

[3] L. Yu, X. Liu, S. Leng et al., "Radiation dose reduction in computed tomography: techniques and future perspective," *Imaging in Medicine*, vol. 1, no. 1, pp. 65–84, 2009.

[4] M. K. Kalra, M. M. Maher, T. L. Toth et al., "Techniques and applications of automatic tube current modulation for CT," *Radiology*, vol. 233, no. 3, pp. 649–657, 2004.

[5] C. H. McCollough, M. R. Bruesewitz, and J. M. Kofler Jr., "CT dose reduction and dose management tools: overview of available options," *Radiographics*, vol. 26, no. 2, pp. 503–512, 2006.

[6] A. Gervaise, P. Naulet, F. Beuret et al., "Low-dose CT with automatic tube current modulation, adaptive statistical iterative reconstruction, and low tube voltage for the diagnosis of renal colic: impact of body mass index," *American Journal of Roentgenology*, vol. 202, no. 3, pp. 553–560, 2014.

[7] M. Sderberg and M. Gunnarsson, "Automatic exposure control in computed tomography an evaluation of systems from different manufacturers," *Acta Radiologica*, vol. 51, no. 6, pp. 625–634, 2010.

[8] J. N. Althén, "Automatic tube-current modulation in CT—a comparison between different solutions," *Radiation Protection Dosimetry*, vol. 114, no. 1–3, pp. 308–312, 2005.

[9] K. M. Kanal, B. K. Stewart, O. Kolokythas, and W. P. Shuman, "Impact of operator-selected image noise index and reconstruction slice thickness on patient radiation dose in 64-MDCT," *American Journal of Roentgenology*, vol. 189, no. 1, pp. 219–225, 2007.

[10] S. Namasivayam, M. K. Kalra, K. M. Pottala, S. M. Waldrop, and P. A. Hudgins, "Optimization of Z-axis automatic exposure control for multidetector row CT evaluation of neck and comparison with fixed tube current technique for image quality and radiation dose," *American Journal of Neuroradiology*, vol. 27, no. 10, pp. 2221–2225, 2006.

[11] J. H. Chen, E. H. Jin, W. He, and L. Q. Zhao, "Combining automatic tube current modulation with adaptive statistical iterative reconstruction for low-dose chest CT screening," *PLoS ONE*, vol. 9, no. 4, Article ID e92414, 2014.

[12] H. Greess, A. Nömayr, H. Wolf et al., "Dose reduction in CT examination of children by an attenuation-based on-line modulation of tube current (CARE Dose)," *European Radiology*, vol. 12, no. 6, pp. 1571–1576, 2002.

[13] M. Ibrahim, H. Parmar, E. Christodoulou, and S. Mukherji, "Raise the bar and lower the dose: current and future strategies for radiation dose reduction in head and neck imaging," *American Journal of Neuroradiology*, vol. 35, no. 4, pp. 619–624, 2014.

[14] J. Ma, Z. Liang, Y. Fan et al., "Variance analysis of x-ray CT sinograms in the presence of electronic noise background," *Medical Physics*, vol. 39, no. 7, pp. 4051–4065, 2012.

[15] E. Y. Sidky, C.-M. Kao, and X. H. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-Ray Science and Technology*, vol. 14, no. 2, pp. 119–139, 2006.

[16] L. Li, Z. Chen, G. Wang, J. Chu, and H. Gao, "A tensor PRISM algorithm for multi-energy CT reconstruction and comparative studies," *Journal of X-Ray Science and Technology*, vol. 22, no. 2, pp. 147–163, 2014.

[17] M. Chang, L. Li, Z. Chen, Y. Xiao, L. Zhang, and G. Wang, "A few-view reweighted sparsity hunting (FRESH) method for CT image reconstruction," *Journal of X-Ray Science and Technology*, vol. 21, no. 2, pp. 161–176, 2013.

[18] T. Niu and L. Zhu, "Accelerated barrier optimization compressed sensing (ABOCS) reconstruction for cone-beam CT: phantom studies," *Medical Physics*, vol. 39, no. 7, pp. 4588–4598, 2012.

[19] L. Shen and Y. Xing, "An edge-preserving total variation denoising method for DECT image," in *Proceedings of the Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC '13)*, pp. 1–4, Seoul, South Korea, October–Novembe 2013.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 354869, 8 pages
http://dx.doi.org/10.1155/2015/354869

Research Article

CT Image Reconstruction from Sparse Projections Using Adaptive TpV Regularization

Hongliang Qi, Zijia Chen, and Linghong Zhou

School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

Correspondence should be addressed to Linghong Zhou; smart@smu.edu.cn

Received 4 August 2014; Revised 12 September 2014; Accepted 18 September 2014

Academic Editor: Liang Li

Radiation dose reduction without losing CT image quality has been an increasing concern. Reducing the number of X-ray projections to reconstruct CT images, which is also called sparse-projection reconstruction, can potentially avoid excessive dose delivered to patients in CT examination. To overcome the disadvantages of total variation (TV) minimization method, in this work we introduce a novel adaptive TpV regularization into sparse-projection image reconstruction and use FISTA technique to accelerate iterative convergence. The numerical experiments demonstrate that the proposed method suppresses noise and artifacts more efficiently, and preserves structure information better than other existing reconstruction methods.

1. Introduction

X-ray computed tomography (CT), as an important medical imaging protocol, has been widely used in clinical applications. However, the involved X-ray radiation dose delivered to patients may potentially increase the probability of causing cancer [1–4]. In this sense, reducing radiation dose without significantly losing image quality is highly required.

Radiation dose in CT examination can be reduced by decreasing the number of projections. However, conventional filtered back-projection (FBP) reconstruction algorithm suffers from systematic geometric distortion and streak artifacts when the measured projection data is not sufficient [5–7]. Iterative methods have been proposed to overcome this problem. Recently, compressed sensing (CS) theory [8] has been applied in CT image reconstruction. It is possible to reconstruct high-quality images from sparse-projection data under the frame of CS. Many optimization methods have been studied following such concepts. Among these optimization methods, total variation (TV) minimization has been widely used. The most famous reconstruction model with TV is ART-TV, firstly proposed by Sidky et al. [9, 10]. This method consists of two steps: ART reconstruction and TV minimization. However, TV is based on an assumption that the signal is piecewise smooth, so this makes TV algorithm suffer from oversmoothing in image edges. To solve this problem, many improved TV methods have been proposed. Tian et al. proposed a TV-based edge preserving (EPTV) model [11]. This model can preserve edges by bringing in different weights in the TV term from edges and constant areas of the to-be-estimated image. Different from the EPTV model, Liu et al. considered the anisotropic edge property of an image and proposed a novel adaptive-weighted TV (AwTV) model [12] for low-dose CT image reconstruction from sparse-sampled projection data. Zhang et al. used a high-order norm coupled within TV to overcome the disadvantages of traditional TV minimization [13]. Chang et al. proposed a few-view reweighed sparsity hunting (FRESH) method for CT image reconstruction [14]. Sidky et al. replaced L_1 norm with L_p $(0 < p < 1)$ norm in the minimization function and investigated image reconstruction by minimizing the L_p norm of the image gradient magnitude or the so-called total p-variation (TpV) [15]. Chen et al. proposed a CT reconstruction algorithm based on L_p $(p = 1/2)$ regularization, where $L_{1/2}$ norm is used as the regularization norm and gradient as the sparse conversion [16]. However, the TpV and L_p $(p = 1/2)$ regularization methods choose p value as a constant in the whole image without identifying edges and constant areas. The disadvantage is that larger p value can oversmooth edges and sometimes produce blocky artifacts, while smaller p value can preserve edges well but enhance blocky artifacts in

constant areas when the projection data is noisy (as shown in Figure 3(f) in [16]). The blocky artifacts are introduced by the noise in the projections whenever p is less than 1 or $p = 1$. Although L_0 regularization is the sparest and most ideal regularization norm, L_0-norm minimization problem is known to be NP-hard, and it is difficult to solve equations. Theoretically, a regularization, which is closer to L_0 norm, could obtain higher-quality CT images in CT reconstruction. It should be noted that TV is the L_1 norm of gradient image. Traditional TpV ($0 < p < 1$) is sparser than TV, and the success of traditional TpV is sharpening image edges, but blocky artifacts still exist in homogeneous regions due to the noisy projection data. The same disadvantage of TV and TpV is their tendency to uniformly penalize the image gradient irrespective of the underlying image structures.

In this study, to deal with the trade-off between smoothing nonedge part and preserving edge part of the image, we propose a CT reconstruction algorithm using adaptive TpV regularization wherein each pixel in reconstructed image corresponds to one p value determined by the pixel's gradient magnitude. From our experiments, one can see that the low-contrast features can be reconstructed better than other methods and blocky artifacts are reduced much to a certain extent. The rest of the paper is organized as follows. In Section 2, ART-TV, traditional TpV, adaptive TpV regularization, and the proposed CT reconstruction algorithm are introduced, respectively. In Section 3, quantitative and qualitative experimental results are shown. Section 4 concludes the paper.

2. Materials and Methods

2.1. ART-TV Reconstruction. CT reconstruction problem can be converted to a constrained optimization problem

$$f = \arg\min_f R(f) \quad \text{subject to } Af = y, \tag{1}$$

where $A = \{a_{ij}\}$ denotes the system matrix, y represents the projection data, and f is the reconstructed image. $R(f)$ is the regularization function.

To solve (1), Sidky et al. proposed famous total variation (TV) based reconstruction method (ART-TV). In their method, $R(f)$ in (1) was considered as a L_1 norm of the first-order gradient image or the so-called TV norm. In a 2D image f with the size $L \times M$, whose pixel values are labeled by $f_{l,m}$, its gradient magnitude with respect to $f_{l,m}$ can be expressed as

$$\nabla f_{l,m} = \sqrt{(f_{l,m} - f_{l-1,m})^2 + (f_{l,m} - f_{l,m-1})^2}, \quad 2 \le l \le L, \quad 2 \le m \le M. \tag{2}$$

The TV of image f is defined as

$$\|f\|_{\mathrm{TV}} = \sum_{l,m}^{L,M} |\nabla f_{l,m}| = \sum_{l,m}^{L,M} \sqrt{(f_{l,m} - f_{l-1,m})^2 + (f_{l,m} - f_{l,m-1})^2 + \varepsilon}, \tag{3}$$

where the parameter ε is a small positive constant to avoid discontinuities.

The ART-TV method is implemented by performing ART algorithm as the first step and TV minimization using gradient descent method as the second step. One can see [9] for more implementation details.

2.2. Traditional TpV and Adaptive TpV Regularization. For traditional TpV algorithm, the quantity $\|\nabla f\|_p$ is the p-norm of the image gradient magnitude. For a 2D image it can be defined by

$$\|f\|_{\mathrm{TpV}} = \sum_{l,m}^{L,M} \|\nabla f_{l,m}\|_p = \sum_{l,m}^{L,M} \left(\sqrt{(f_{l,m} - f_{l-1,m})^2 + (f_{l,m} - f_{l,m-1})^2 + \varepsilon} \right)^p, \quad 0 < p \le 1. \tag{4}$$

When p is set to be 1, the $\|f\|_{\mathrm{TpV}}$ reduces to conventional $\|f\|_{\mathrm{TV}}$ and reconstructed image will suffer from oversmoothing image feature details. When $p < 1$ (e.g., $p = 1/2$ in [16]), the structural information can be efficiently preserved, but at the same time the blocky artifacts in nonedge regions will be enhanced when the projection data is noisy.

To overcome this limitation, in this work we propose an adaptive TpV (ATpV) regularization defined by

$$\|f\|_{\mathrm{ATpV}} = \sum_{l,m}^{L,M} \left(\sqrt{(f_{l,m} - f_{l-1,m})^2 + (f_{l,m} - f_{l,m-1})^2 + \varepsilon} \right)^{p_{l,m}}, \quad 0 < p_{l,m} \le 1, \tag{5}$$

where $p_{l,m}$ is determined by pixel $f_{l,m}$ in a 2D image. On one hand, if a pixel's gradient magnitude is large, this pixel is on the edge and corresponds to a small p value to avoid oversmoothing edges. On the other hand, if the gradient magnitude of one pixel is small, this pixel is in the nonedge area and corresponds to a large p value to suppress noise and artifacts. In this study, we define $p_{l,m}$ as

$$p_{l,m} = \frac{1}{1 + \nabla(B * f)_{l,m}}, \tag{6}$$

where $B * f$ denotes image f is filtered by the well-known bilateral filter which does well in denoising and preserving edge information. ∇ is gradient operator. Adding bilateral filter on image f is to avoid treating noise point as edge point when calculating gradient magnitude. Obviously in (6), a pixel with large gradient magnitude corresponds to a small p value and a pixel with small gradient magnitude corresponds to a large p value. The "1" in the denominator in (6) may not be the best option and may be correlated with the contrast value of the reconstruction, but in our experiments,

the defined (6) can produce good result. How to replace "1" with an optimal value will be an interesting topic in our future study.

In summary, the benefit of the proposed ATpV is that the parameter p is dynamically adopted by identifying edges and nonedges, and larger p value is chosen to smooth constant areas while smaller p value is chosen to preserve edge part, which will improve the reconstruction quality for sparse-view reconstruction.

2.3. CT Reconstruction Algorithm Based on Adaptive TpV Regularization. According to aforementioned methods, in this paper we propose CT image iterative reconstruction using ATpV regularization. The reconstruction is implemented by solving the following constrained minimization problem:

$$\begin{aligned} & f \\ & = \arg\min_{f} \|f\|_{\mathrm{ATpV}} \\ & = \arg\min_{f} \left\{ \sum_{l,m}^{L,M} \left(\sqrt{(f_{l,m} - f_{l-1,m})^2 + (f_{l,m} - f_{l,m-1})^2 + \varepsilon} \right)^{p_{l,m}} \right\} \\ & \quad \text{subject to } Af = y. \end{aligned} \tag{7}$$

The algorithm implementation can follow ART-TV in [9].

Besides, we apply fast iterative shrinkage/thresholding algorithm (FISTA) [17] to accelerate iterative convergence. In FISTA, the key idea is that initial value of the next iteration is determined by a linear combination of the two previous iterate results. For simplicity, in this study the proposed reconstruction method is termed ART-ATpV without using FISTA or ART-ATpV-FISTA using FISTA.

In summary, the main steps of ART-ATpV-FISTA are as follows.

(A) Initialization: $f^0 = 0$, iteration index $k = 1, 2 \ldots K$, $t_0 = 0$.

(B) ART reconstruction:

$$f^{k+1} = f^k + \lambda A_i \frac{y_i - A_i f^k}{A_i A_i^T}. \tag{8}$$

(C) Positivity constraint: $f^{k+1} = \max(f^{k+1}, 0)$.

(D) ATpV minimization:

calculate $p_j = 1/(1 + \nabla(B * f^{k+1})_j)$ for each pixel f_j^{k+1},

minimize $\|f\|_{\mathrm{ATpV}}$ using gradient descent algorithm to get updated f_j^{k+1}.

Figure 1: Shepp-Logan phantom.

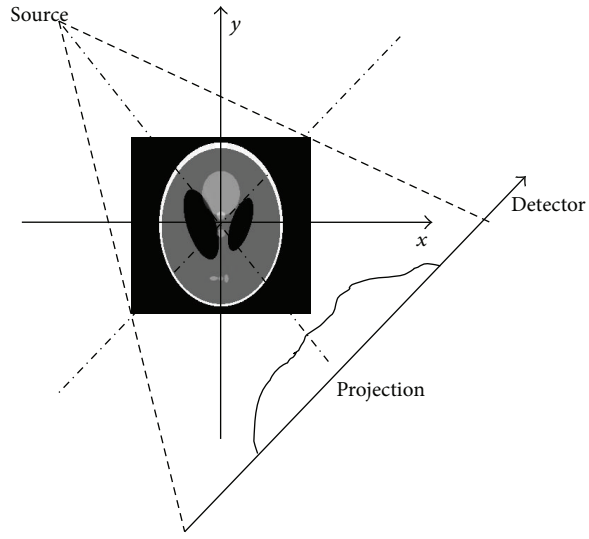


Figure 2: Fan beam CT geometry configuration.

(E) FISTA acceleration:

$$\begin{aligned} t &= \frac{1 + \sqrt{1 + 4t_0^2}}{2}, \\ f_j^{k+1} &= f_j^{k+1} + \frac{t_0 - 1}{t} \left(f_j^{k+1} - f_j^k \right), \\ t_0 &= t, \quad k = k + 1. \end{aligned} \tag{9}$$

(F) Return to (B) until the stopping criterion is satisfied.

In our experimental implementation, the initial to-be-reconstructed image was set to be uniform with pixel values of 0. The relaxation parameter λ for the ART was fixed at 1.0, and the step-size used in ATpV minimization using the gradient descent was set to be constant 0.2. The parameter ε in (7) was fixed as 10^{-5}.

3. Experimental Study

3.1. Numerical Simulation. In this section, we study the ART, ART-TV, ART-TpV ($p = 0.5$), and our proposed

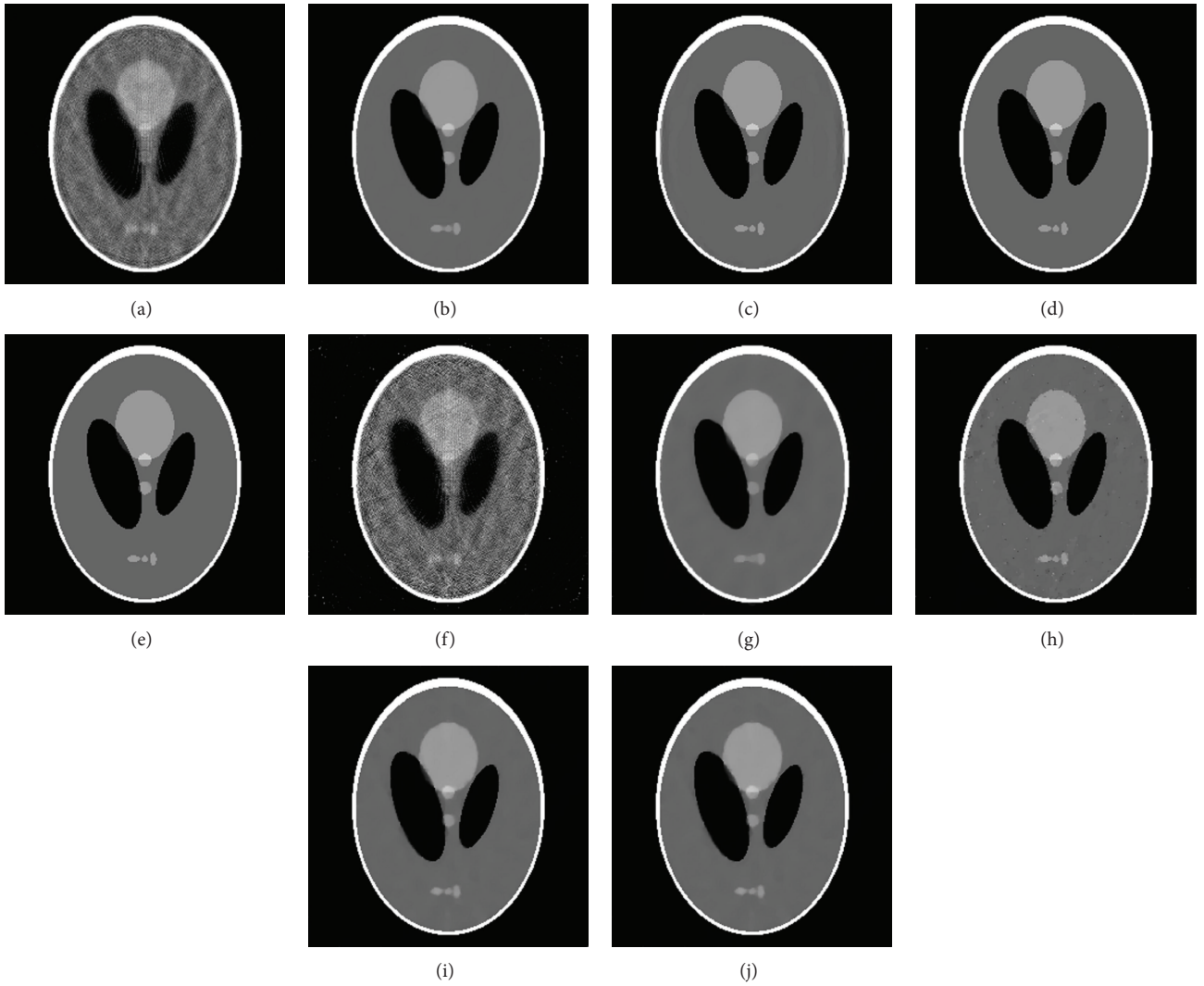


Figure 3: The images reconstructed by different reconstruction algorithms from the noise-free and noisy data. ((a)–(e)) Reconstructed images from noise-free data: (a) image reconstructed by ART, (b) image reconstructed by ART-TV, (c) image reconstructed by ART-TpV, (d) image reconstructed by ART-ATpV, and (e) image reconstructed by ART-ATpV-FISTA; ((f)–(j)) reconstructed images from noisy data: (f) image reconstructed by ART, (g) image reconstructed by ART-TV, (h) image reconstructed by ART-TpV, (i) image reconstructed by ART-ATpV, and (j) image reconstructed by ART-ATpV-FISTA.

algorithm. Numerical experiment results are given. Shepp-Logan phantom is tested in this paper as shown in Figure 1, and the size of phantom image is 256×256. Without losing generality, we choose a fan beam imaging geometry to capture the projection data as illustrated in Figure 2. The source to rotation center distance is 40 cm and the detector to rotation center is 40 cm. The image array is $25.6 \times 25.6\,\text{cm}^2$. The detector whose length is 61.44 cm is modeled as a straight-line array of 512 detector bins. All the tests are performed by MATLAB on a PC with Intel (R) Core (TM) 2 Quad CPU 2.50 GHz and 3.25 GB RAM.

The generation of projection data is using Siddon's ray-driven algorithm [18, 19]. Projection data is sampled evenly over 360° and the number of projections in this experiment is set 20 over 360°. In the simulation, we add 0.2% Gaussian noise to noise-free projection data. The iteration number for all iterative methods in this experiment is 200, which makes sure each method reaches convergence.

The reconstruction results are shown in Figure 3. Figures 3(a)–3(e) show the reconstructed images using different methods from noise-free projection data. In Figure 3(a), it is obvious that, due to insufficiency of projection data, ART reconstructs low-quality image containing severe artifacts and noises. In Figure 3(b), artifacts and noises are effectively suppressed by using ART-TV, but three oval organs in the bottom of image are blurred to some extent. The ART-TpV can overcome the limitation of TV and the result can be seen from Figure 3(c). Similarly, the image edges reconstructed by ART-ATpV are well kept and three oval organs in the bottom of image are more distinguishable in Figure 3(d). Figure 3(e) shows the image reconstructed by ART-ATpV-FISTA after only 50 iterations which has almost the same image quality

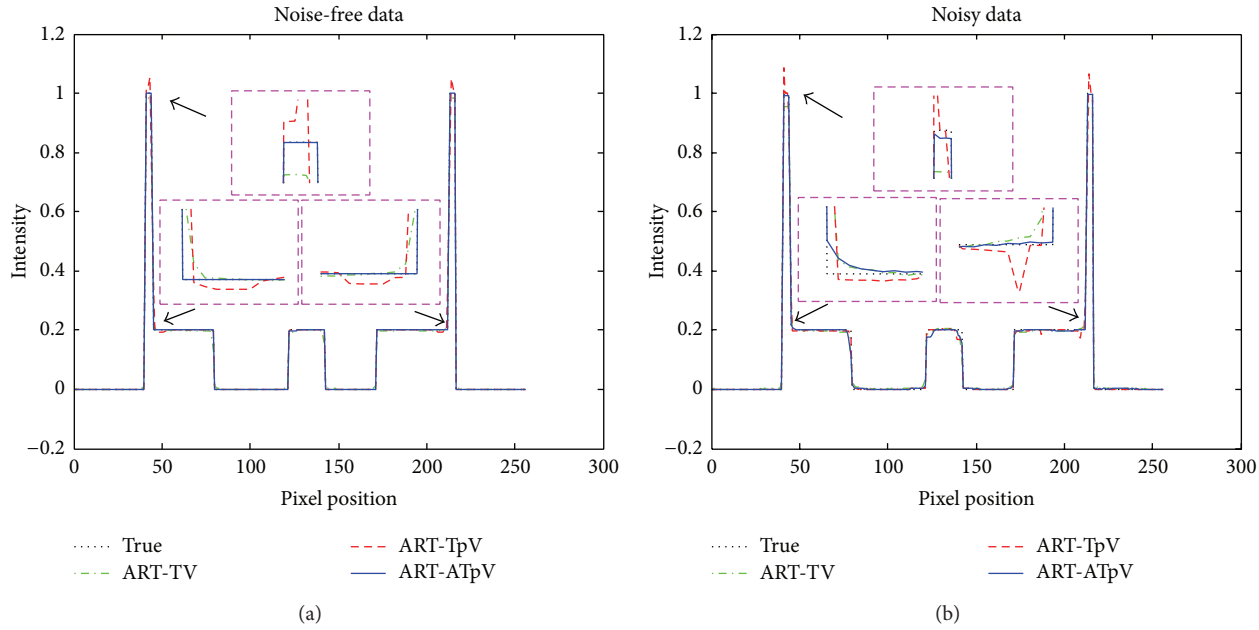


FIGURE 4: The comparison between reconstructed images using three different reconstruction algorithms and original Shepp-Logan phantom. (a) The horizontal profiles in reconstructed images using ART-TV, ART-TpV, and ART-ATpV methods from noise-free data and original Shepp-Logan phantom and (b) the horizontal profiles in reconstructed images using ART-TV, ART-TpV, and ART-ATpV methods from noisy data and original Shepp-Logan image.

as Figure 3(d). One can see that ART-ATpV-FISTA can reconstruct high-quality image using less iterations. The same results can be seen in Figures 3(f)–3(j) with noisy projection data. It is noticed that image (Figure 3(h)) reconstructed by ART-TpV has some small enhanced blocky artifacts due to noise existing in projection data.

Figure 4 depicts the horizontal profiles through the center pixels of the ART-TV, ART-TpV, and ART-ATpV reconstructed images corresponding to Figure 3. It can be clearly seen that the ART-ATpV algorithm produces a closer profile to the true image compared to ART-TV and ART-TpV.

To assess the accuracy of the reconstructed image, the mean absolute error (MAE) is used and defined by

$$\mathrm{MAE} = \frac{\left\| x - x^{\mathrm{true}} \right\|_1}{N}, \tag{10}$$

where N is the total number of pixels of the reconstructed image x and x^{true} is the original image.

As shown in Figure 5, the ART-ATpV with/without FISTA algorithms can reconstruct high-quality images at less iteration numbers, and application of FISTA to ART-ATpV (denoted ART-ATpV-FISTA) can further remarkably accelerate iterative convergence.

To challenge our ART-ATpV/ART-ATpV-FISTA method further, we use a complicated low-contrast FORBILD phantom to reconstruct image and compare it to other methods. The corresponding images are in Figure 6. One can see that high-quality images can be obtained by our proposed method.

3.2. Real Data Experiment. In this section, we use a real CT image (head phantom) obtained from a commercial medical CT scanner to test the effectiveness of our ART-ATpV and ART-ATpV-FISTA algorithms and compare them with other algorithms. 40 projections are simulated in this case, with the aforementioned geometrical parameters unchanged. For all iterative methods except ART-ATpV-FISTA, the number of iterations is 50. For ART-ATpV-FISTA, the number of iterations is 20.

As shown in Figure 7, due to the incompleteness of projection data, the reconstructed image using ART method has more artifacts and noises than the reconstructed images by using ART-TV, ART-TpV, ART-ATpV, and ART-ATpV-FISTA method. ART-TV suppresses most of the artifacts and noises, but the oversmoothing effect and blocky artifacts exist in the images, as indicated by the white arrows in Figure 7(c). Although ART-TpV method suppresses most of the artifacts and noises, the blocky artifacts are still visible and enhanced in the images, as indicated by the white arrows in Figure 7(d). ART-ATpV overcomes the flaws and preserves the structure information well, without obvious blocky artifacts. Figure 7(f) shows the image reconstructed by ART-ATpV-FISTA after only 20 iterations. It is seen that the reconstructed image has almost the same image quality as Figure 7(e) (ART-ATpV), which means that ART-ATpV-FISTA can reconstruct high-quality image using less iterations. The FISTA technique is a useful tool to accelerate convergence rate and can be applied to other iterative reconstruction approaches for speedup.

4. Discussions and Conclusion

In this paper, we present ART-ATpV-FISTA method for X-ray CT reconstruction from few-view or sparse projections.

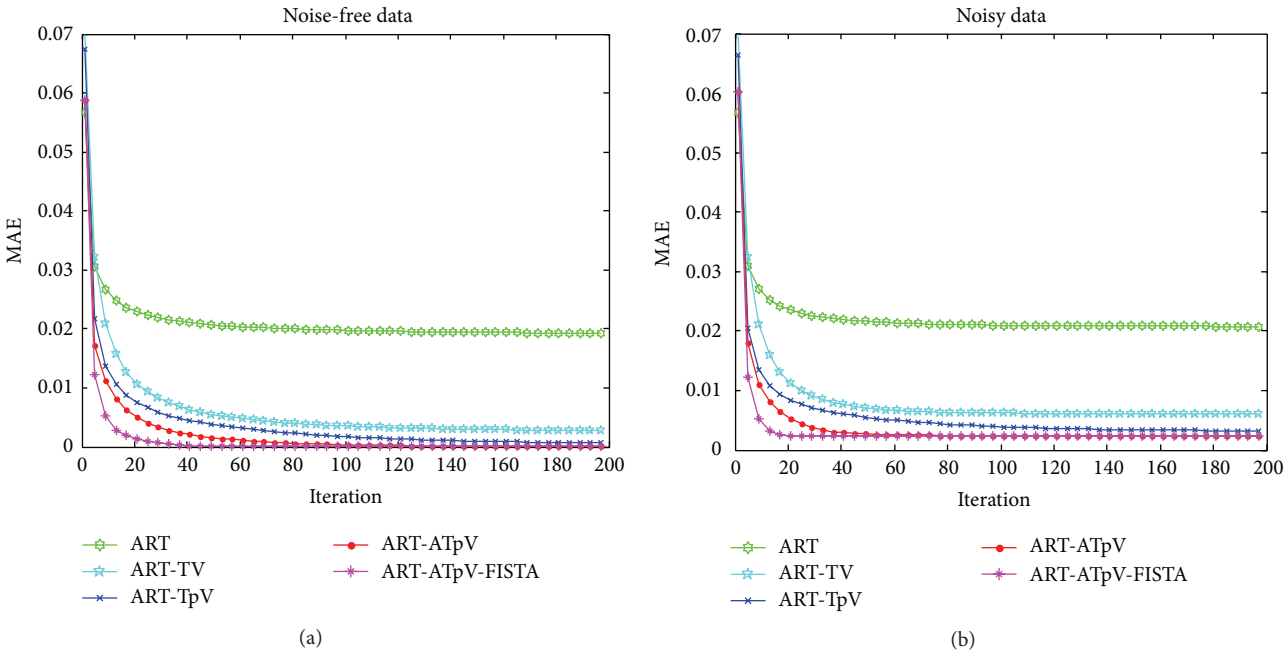


Figure 5: The MAE curves of reconstructed images with different reconstruction algorithms at 20 projection angles and different iteration numbers, and the iteration numbers range from 1 to 200. (a) The MAE curves of reconstructed images from noise-free projection data and (b) the curves of reconstructed images from noisy projection data.

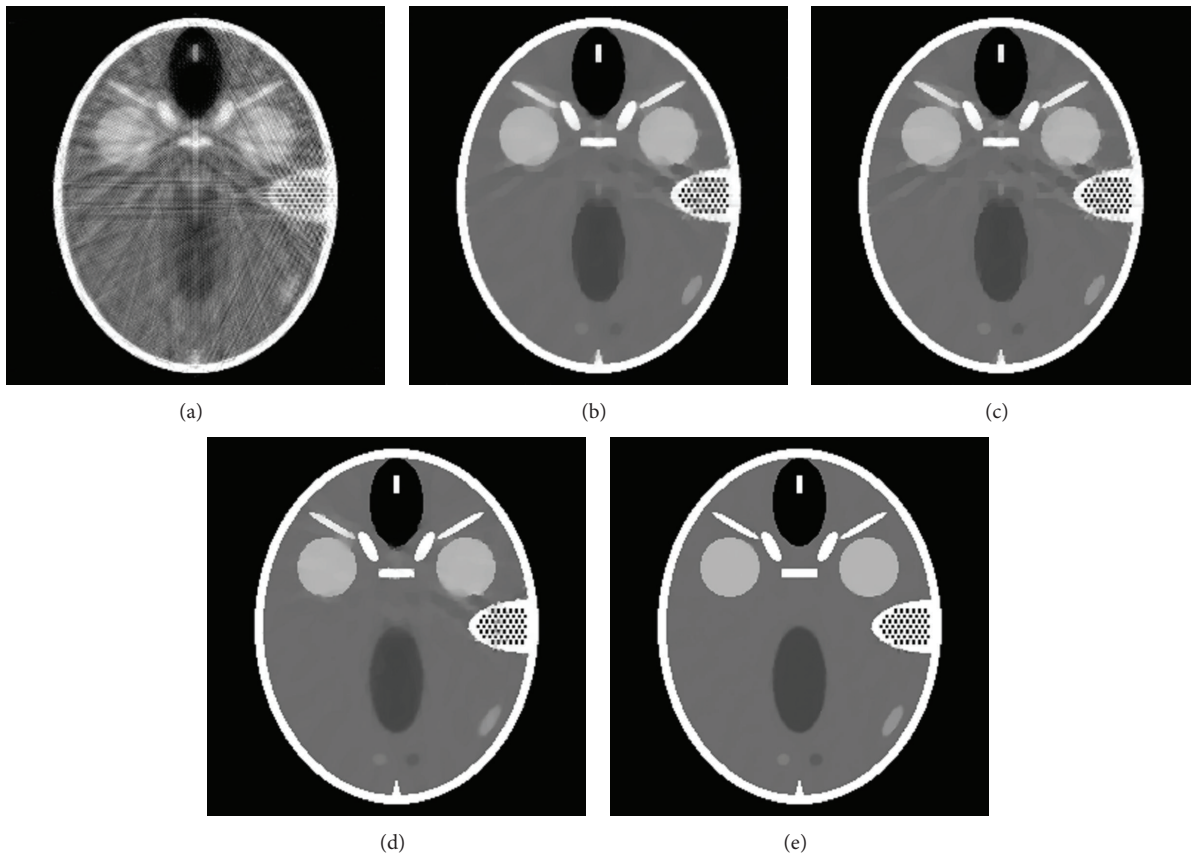


Figure 6: The images reconstructed by different reconstruction algorithms from noisy data: (a) image reconstructed by ART, (b) image reconstructed by ART-TV, (c) image reconstructed by ART-TpV, (d) image reconstructed by ART-ATpV, and (e) image reconstructed by ART-ATpV-FISTA.

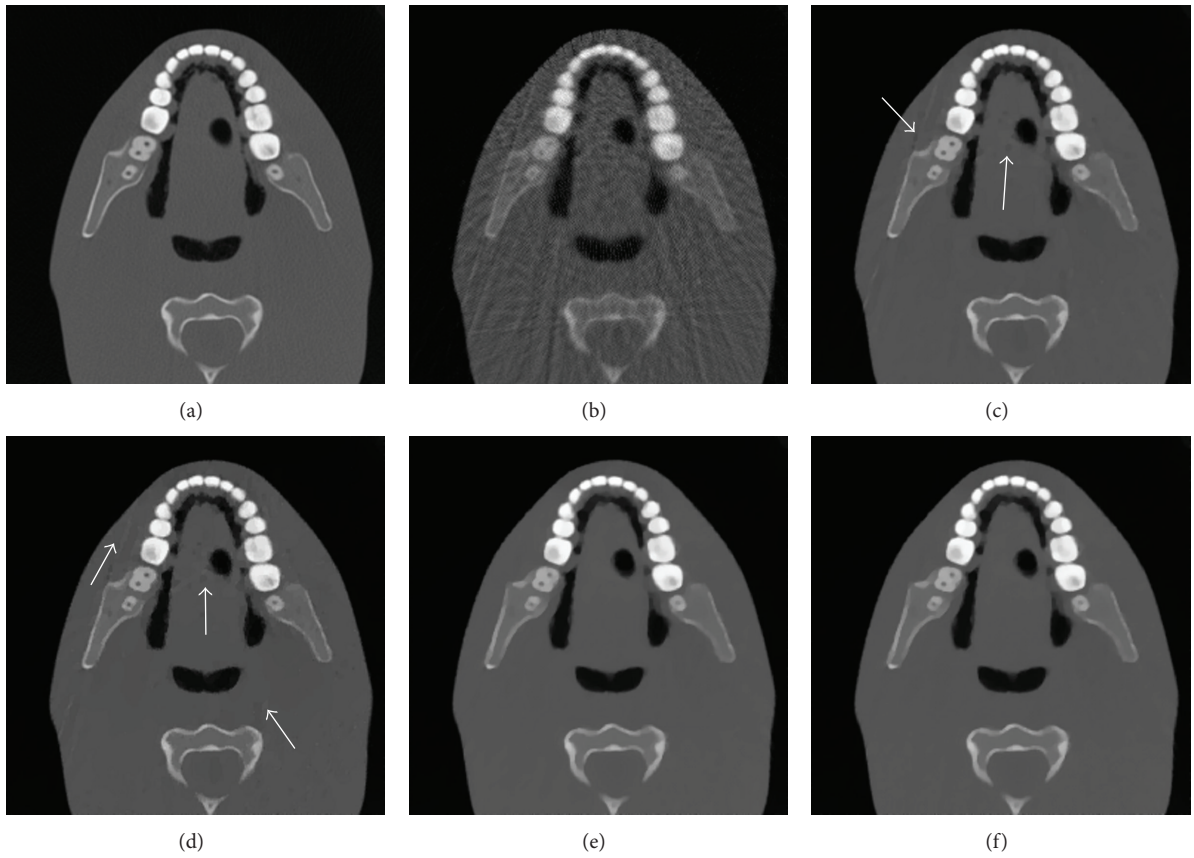


Figure 7: The reconstructed images by different algorithms from real head phantom projection data. (a) The original image, (b) the reconstructed image using ART, (c) the reconstructed image using ART-TV, (d) the reconstructed image using ART-TpV, (e) the reconstructed image using ART-ATpV, and (f) the reconstructed image using ART-ATpV-FISTA.

The main contribution of this work is to minimize adaptive TpV norm of reconstructed image instead of traditional TpV norm and TV norm. FISTA technique is employed to speed up iterative convergence rate.

The advantage of adaptive TpV is that if a pixel's gradient magnitude is large, this pixel is on the edge and corresponds to a small p value to avoid oversmoothing. If gradient magnitude of one pixel is small, this pixel is in the constant area and corresponds to a large p value to smooth noise and artifacts. In our primary experiment, the Gaussian filter which was chosen in (6) may not lead to some losses of details, because the Shepp-Logan phantom is simple. When we use low-contrast FORBILD phantom to reconstruct images, as we expect, the Gaussian filter causes the losses of low-contrast details in FORBILD phantom. Therefore, to effectively remove noise and preserve low-contrast structures, we use bilateral filter instead of Gaussian filter when we calculate p values, and the results are better.

The performance of the propose method is compared to ART, ART-TV, and ART-TpV methods on Shepp-Logan phantom, low-contrast FORBILD phantom, and a real head phantom. Both qualitative and quantitative comparisons are performed to show the proposed method provides more superior results than other existing methods. Since the main goal of this work is to demonstrate the effectiveness of the proposed ATpV-based regularization, the parameters were empirically set through extensive experiments by visual inspection and quantitative measures in this study.

Although the presented ART-ATpV-FISTA algorithm in this paper is used in fan beam CT geometry, it is also easily extended to cone beam CT (CBCT) geometry due to its iterative-correction property. Furthermore, the ART-ATpV-FISTA algorithm may also be useful for other tomographic imaging modalities. In the further research, we will apply the developed algorithm in CBCT system and study the few-view CBCT reconstruction, which will reduce radiation dose as much as possible.

Similar to FBP, serious streaking artifacts in the reconstructed CT images using FDK type algorithms [20, 21] exist when number of X-ray projections is not sufficient. Besides, using GPU to speed up the computationally intensive tasks of CBCT reconstruction has drawn a lot of attention recently [22–26] and we would like to include it in our future study.

In conclusion, the proposed algorithm using adaptive TpV regularization in this work can reconstruct high-quality images from few-view projections and will have great potential clinical applications.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by Guangdong strategic emerging industry core technology research (Grant no. 2011A081402003) and Guangzhou municipal special major science and technology programs (Grant no. 2011Y1-00019).

References

[1] D. J. Brenner and E. J. Hall, "Computed tomography—an increasing source of radiation exposure," *The New England Journal of Medicine*, vol. 357, no. 22, pp. 2277–2284, 2007.

[2] M. K. Islam, T. G. Purdie, B. D. Norrlinger et al., "Patient dose from kilovoltage cone beam computed tomography imaging in radiation therapy," *Medical Physics*, vol. 33, no. 6, pp. 1573–1582, 2006.

[3] N. Wen, H. Guan, R. Hammoud et al., "Dose delivered from Varian's CBCT to patients receiving IMRT for prostate cancer," *Physics in Medicine and Biology*, vol. 52, no. 8, pp. 2267–2276, 2007.

[4] M. J. Murphy, J. Balter, S. Balter et al., "The management of imaging dose during image-guided radiotherapy: report of the AAPM Task Group 75," *Medical Physics*, vol. 34, no. 10, pp. 4041–4063, 2007.

[5] L. A. Shepp and B. F. Logan Jr., "The Fourier reconstruction of a head section," *IEEE Transactions on Nuclear Science*, vol. 21, no. 3, pp. 21–43, 1974.

[6] Y.-M. Zhu, "Extrapolation method to obtain a full projection from a limited one for limited-view tomography problems," *Medical and Biological Engineering and Computing*, vol. 31, no. 3, pp. 323–324, 1993.

[7] P. Oskoui and H. Stark, "Comparative study of three reconstruction methods for a limited-view computer tomography problem," *IEEE Transactions on Medical Imaging*, vol. 8, no. 1, pp. 43–49, 1989.

[8] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[9] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-Ray Science and Technology*, vol. 14, no. 2, pp. 119–139, 2006.

[10] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in Medicine and Biology*, vol. 53, no. 17, pp. 4777–4807, 2008.

[11] Z. Tian, X. Jia, K. Yuan, T. Pan, and S. B. Jiang, "Low-dose CT reconstruction via edge-preserving total variation regularization," *Physics in Medicine and Biology*, vol. 56, no. 18, pp. 5949–5967, 2011.

[12] Y. Liu, J. Ma, Y. Fan, and Z. Liang, "Adaptive-weighted total variation minimization for sparse data toward low-dose x-ray computed tomography image reconstruction," *Physics in Medicine and Biology*, vol. 57, no. 23, pp. 7923–7956, 2012.

[13] Y. Zhang, W.-H. Zhang, H. Chen, M.-L. Yang, T.-Y. Li, and J.-L. Zhou, "Few-view image reconstruction combining total variation and a high-order norm," *International Journal of Imaging Systems and Technology*, vol. 23, no. 3, pp. 249–255, 2013.

[14] M. Chang, L. Li, Z. Chen, Y. Xiao, L. Zhang, and G. Wang, "A few-view reweighted sparsity hunting (FRESH) method for CT image reconstruction," *Journal of X-Ray Science and Technology*, vol. 21, no. 2, pp. 161–176, 2013.

[15] E. Y. Sidky, R. Chartrand, and X. Pan, "Image reconstruction from few views by non-convex optimization," in *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS '07)*, pp. 3526–3530, Honolulu, Hawaii, USA, November 2007.

[16] M. Chen, D. Mi, P. He, L. Deng, and B. Wei, "A CT reconstruction algorithm based on $L_{1/2}$ regularization," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 862910, 8 pages, 2014.

[17] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.

[18] R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," *Medical Physics*, vol. 12, no. 2, pp. 252–255, 1985.

[19] F. Jacobs, E. Sundermann, B. de Sutter, M. Christiaens, and I. Lemahieu, "A fast algorithm to calculate the exact radiological path through a pixel or voxel space," *Journal of Computing and Information Technology*, vol. 6, no. 1, pp. 89–94, 1998.

[20] I. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, vol. 1, no. 6, pp. 612–619, 1984.

[21] L. Li, Y. Xing, Z. Chen, L. Zhang, and K. Kang, "A curve-filtered FDK (C-FDK) reconstruction algorithm for circular cone-beam CT," *Journal of X-Ray Science and Technology*, vol. 19, no. 3, pp. 355–371, 2011.

[22] X. Jia, Y. Lou, J. Lewis et al., "GPU-based fast low-dose cone beam CT reconstruction via total variation," *Journal of X-Ray Science and Technology*, vol. 19, no. 2, pp. 139–154, 2011.

[23] X. Jia, B. Dong, Y. Lou, and S. B. Jiang, "GPU-based iterative cone-beam CT reconstruction using tight frame regularization," *Physics in Medicine and Biology*, vol. 56, no. 13, pp. 3787–3807, 2011.

[24] Z. Tian, X. Jia, K. Yuan, T. Pan, and S. B. Jiang, "GPU-based low dose CT reconstruction via edge-preserving total variation regularization," *Physics in Medicine and Biology*, vol. 56, no. 18, pp. 1359–1365, 2011.

[25] Z. Tian, X. Jia, B. Dong, Y. Lou, and S. B. Jiang, "Low-dose 4DCT reconstruction via temporal nonlocal means," *Medical Physics*, vol. 38, no. 3, pp. 1359–1365, 2011.

[26] X. Jia, H. Yan, L. Cerviño, M. Folkerts, and S. B. Jiang, "A GPU tool for efficient, accurate, and realistic simulation of cone beam CT projections," *Medical Physics*, vol. 39, no. 12, pp. 7368–7378, 2012.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 494691, 9 pages
http://dx.doi.org/10.1155/2015/494691

Research Article

Adaptive Autoregressive Model for Reduction of Noise in SPECT

Reijo Takalo,[1] Heli Hytti,[2] Heimo Ihalainen,[2] and Antti Sohlberg[3]

[1] *Division of Nuclear Medicine, Department of Diagnostic Radiology, Oulu University Hospital (OYS), P.O. Box 500, 90029 Oulu, Finland*
[2] *Department of Automation Science and Engineering, Tampere University of Technology, P.O. Box 692, 33101 Tampere, Finland*
[3] *Joint Authority for Päijät-Häme Social and Health Care, Department of Clinical Physiology and Nuclear Medicine, Keskussairaalankatu 7, 15850 Lahti, Finland*

Correspondence should be addressed to Reijo Takalo; reijo.takalo@ppshp.fi

Received 2 September 2014; Revised 5 November 2014; Accepted 2 December 2014

Academic Editor: Liang Li

This paper presents improved autoregressive modelling (AR) to reduce noise in SPECT images. An AR filter was applied to prefilter projection images and postfilter ordered subset expectation maximisation (OSEM) reconstruction images (AR-OSEM-AR method). The performance of this method was compared with filtered back projection (FBP) preceded by Butterworth filtering (BW-FBP method) and the OSEM reconstruction method followed by Butterworth filtering (OSEM-BW method). A mathematical cylinder phantom was used for the study. It consisted of hot and cold objects. The tests were performed using three simulated SPECT datasets. Image quality was assessed by means of the percentage contrast resolution (CR%) and the full width at half maximum (FWHM) of the line spread functions of the cylinders. The BW-FBP method showed the highest CR% values and the AR-OSEM-AR method gave the lowest CR% values for cold stacks. In the analysis of hot stacks, the BW-FBP method had higher CR% values than the OSEM-BW method. The BW-FBP method exhibited the lowest FWHM values for cold stacks and the AR-OSEM-AR method for hot stacks. In conclusion, the AR-OSEM-AR method is a feasible way to remove noise from SPECT images. It has good spatial resolution for hot objects.

1. Introduction

Numerous methods for removing noise from SPECT images have been proposed [1, 2]. This indicates the difficulty of the task. Noise removal can be performed before reconstruction (prefiltering), during reconstructions or after reconstruction (postfiltering). In modern iterative methods, collimator correction denoises images during reconstruction, but the reconstructed images may still require postfiltering [3]. Earlier we introduced an adaptive autoregressive (AR) filter to reduce noise in scintigraphic planar images or projection images of a SPECT study [4]. In the present work, the AR filter was further improved to reduce noise from the projection images and also from three-dimensionally reconstructed data. It is important to apply the best AR filter to the projection data of SPECT, because a small change in the projection data may cause a large change in the estimated transaxial image [5]. Our method was compared with two established methods for improving image quality in SPECT. The methodical comparison was carried out using a three-dimensional mathematical cylinder phantom (3D-MAC) [6], and it was illustrated with patient data.

2. Methods

2.1. AR Model. In two-dimensional AR modelling, each value of an image is regressed on its neighbourhood pixel values, called the prediction region. An AR model can be regarded as a low-pass filter that divides the image into two additive components, a predictable image and a prediction error image. An AR process $X(n_1, n_2)$ is defined by

$$X_{\text{pred}}(n_1, n_2) = -\sum_{k_1}\sum_{k_2} a(k_1, k_2)\, X_{\text{orig}}(n_1 - k_1, n_2 - k_2) + w(n_1, n_2), \quad (1)$$

where $a(k_1, k_2)$ are the predictor (weighting) coefficients, indices k_1 and k_2 define the type of prediction region in a two-dimensional array (n_1, n_2 matrix), and $w(n_1, n_2)$ represents prediction error, that is, the difference between the predicted value and the current value in this pixel. The predictable image X_{pred} is the image obtained by applying the AR model to the original image X_{orig}. The prediction error image $X_{\text{err}} = X_{\text{orig}} - X_{\text{pred}}$.

In a typical scintigraphic image, there are large local spatial variations in the count number of the image. Therefore, the same model cannot be applied to the entire image, but the model must be adapted to the variations. In this adaptive method, the image area is divided into smaller blocks and the AR model is then fitted into each block separately by using MATLAB subroutines. Recently, a block-wise denoising method has been introduced also for three-dimensional ultrasound images [7]. In the AR model, a prediction region of four orthogonal neighbours of the predicted pixel with a block size of 5 × 5 pixels was used [4]. Seventy-five percent overlap of the image blocks in combination with one iteration of the filtering procedure was used. The two error term images were summed up and subjected to AR filtering, and the resulting image was then added to the iteratively filtered image (Figure 1). In the present study, we tested the effect of using another AR model for the summed error term images than for the original image. We used the same transaxial slice of the Zubal phantom [8] and the same simulation conditions as in our previous work [4], and image quality was assessed by means of the mean squared error (MSE) of the image. It is of note that the Poisson-noise-corrupted slice of the phantom actually represented an artificial scintigraphic planar image or a projection image of a SPECT study. The AR model with the lowest MSE was then used to prefilter the SPECT projection images and also to postfilter iteratively reconstructed data. The filter was applied to each set of orthogonal plane images separately. The software was based on MATLAB subroutines (The MathWorks, Inc.).

2.2. Phantom. Data were simulated using a 3D-MAC phantom [6]. The phantom measured 200 mm in both diameter and length. It comprised three imbedded objects: two hot objects and a cold one. Each object consisted of five stacked cylinders. The cylinders had diameters of 4, 10, 20, 40, and 60 mm and a length of 30 mm. The smallest cylinder was not utilised in the present study because its dimensions were beyond the resolution of the simulated SPECT system used. Relative activities were 1, 0, 2, and 4 for the background, a cold stack, and two hot stacks, respectively. The tests were performed using three SPECT datasets with different image statistics. Total counts of the projection images were approximately 50000 (low level), 100000 (intermediate level), and 150000 (high level) per projection. A built-in MATLAB function was used to add Poisson noise to ideal projection images of the 3D-MAC. The mean counts of a pixel in the projection images were 12, 24, and 37, respectively, and the range of pixel values was 0–52, 0–104, and 0–156, respectively. The matrix size was 64 × 64 pixels, pixel size was 4 mm, and the number of projections was 120. There was no scatter or attenuation component and perfect depth-independent resolution was assumed in the simulated data. Thus, the only factor degrading image quality in the projection images was the Poisson noise.

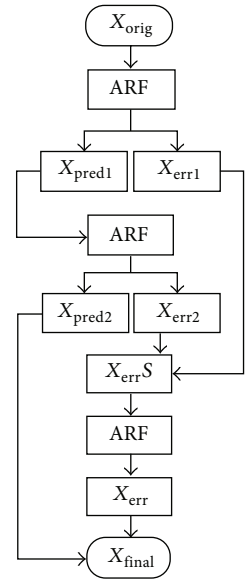
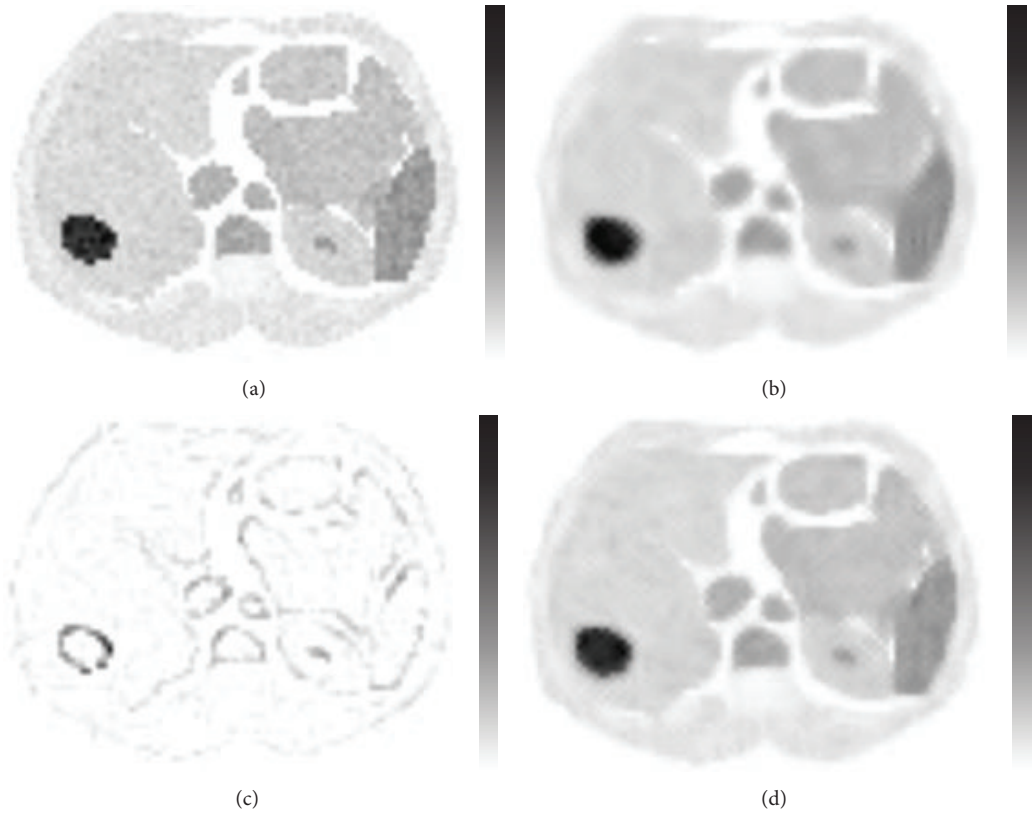


Figure 1: Flowchart of the autoregressive denoising process. ARF: autoregressive filtering; X_{orig}: original noise-corrupted image; X_{pred1} and X_{pred2}: predictable images; $X_{\text{err}\,1}$, $X_{\text{err}\,2}$, X_{err}: prediction error images; $X_{\text{err}}S$: sum of two prediction error images; X_{final}: final image.

Table 1: Effect of changing the block size of the summed error term image with a prediction region of 3 × 3 pixels. For the predictable image a prediction region of four orthogonal neighbours with a block size of 5 × 5 pixels was used.

Total counts	Block size	Mean squared error
28705	5 × 5	0.87
28705	6 × 6	0.86
28705	7 × 7	0.86
54469	5 × 5	2.12
54469	6 × 6	2.10
54469	7 × 7	2.14
108938	5 × 5	6.61
108938	6 × 6	6.56
108938	7 × 7	7.04

2.3. Reconstruction Methods. Transaxial slices were reconstructed using either the filtered back projection method (FBP) [9] or an iterative ordered subset expectation maximisation (OSEM) algorithm [10]. The reconstruction methods were implemented on the Hermes SPECT (G) reconstruction software (version 3.8) and reconstruction engine of Hermes HybridRecon (Hermes Medical Solutions, Stockholm, Sweden), respectively. Three methods were compared: AR filtering before and after ordered OSEM reconstruction (AR-OSEM-AR), two-dimensional Butterworth filtering before FBP reconstruction in combination with a ramp filter during



(a) (b) (c) (d)

Figure 2: Transaxial slice of the Zubal phantom. (a) Poisson-noise-corrupted transaxial slice. (b) Iteratively filtered predictable image. (c) Filtered summed error term image. (d) The final image. The images are individually scaled to their own maximum. Inverse linear grey scale is used for comparison with original phantom. The total count level is 108791 in the Poisson-noise-corrupted image, 102237 in the iteratively filtered predictable image, and 6205 in the filtered summed error term image.

reconstruction (BW-FBP), and OSEM reconstruction followed by three-dimensional Butterworth filtering (OSEM-BW). The Butterworth filter was originally designed for one-dimensional data [11]. In the OSEM method, the number of subsets was set to 8 and the number of iterations to 10. Postfiltering was performed using Multimodality software (Hermes Medical Solutions, Stockholm, Sweden). Noise-free projection images were also reconstructed using the OSEM (Ideal-OSEM) method.

2.4. Assessment of Image Quality. To obtain a fair comparison of the methods, the same amount of filtering was applied in each method. This was done by drawing a circular region-of-interest (ROI) 150 mm in diameter in the uniform part of the phantom and calculating the percentage coefficient of variation (CoV%) in the ROI, that is, the ratio of the standard deviation to the mean multiplied by 100. This kind of presentation ensures that filtering between each method is equal.

Percentage contrast resolution (CR%) values for the activity in each cylinder and uniform activity were calculated. CR% can be expressed by the following formula [12]:

$$\text{CR\%} = \frac{|M - m|}{M} \times 100\%, \quad (2)$$

where M is the count value of uniform activity and m is the count value in each cylinder. Activity in each cylinder was analysed using a circular ROI with the same diameter as the cylinder. The ROIs were drawn on noise-free transaxial slices and were copied to each set of reconstructed data, so their

Table 2: Percentage coefficient of variation for different reconstruction techniques.

Method	Count level	RelAct	CF (cycles/cm)	CoV%
AR-OSEM-AR	50000	0	—	6.35
BW-FBP	50000	2	0.83	6.37
OSEM-BW	50000	4	0.84	6.36
AR-OSEM-AR	100000	0	—	4.60
BW-FBP	100000	2	0.80	4.56
OSEM-BW	100000	4	0.84	4.65
AR-OSEM-AR	150000	0	—	4.43
BW-FBP	150000	2	0.89	4.50
OSEM-BW	150000	4	0.86	4.43

AR-OSEM-AR: autoregressive filtering before and after ordered subset expectation maximisation algorithm; BW-FBP: Butterworth prefiltering and filtered back projection; OSEM-BW: ordered subset expectation maximisation algorithm and Butterworth postfiltering; RelAct: activity relative to background activity of 1; CF: cut-off frequency. The order of the filter was 2; —: not definable.

Table 3: Percentage contrast resolution values for the different methods.

Method	RelAct	10 Ø	20 Ø	40 Ø	60 Ø
		A			
AR-OSEM-AR	0	18.4	56.3	74.6	84.2
BW-FBP	0	28.7	67.2	83.1	91.6
OSEM-BW	0	24.9	59.5	75.7	85.6
AR-OSEM-AR	2	9.0	86.1	84.3	103.5
BW-FBP	2	16.7	75.5	86.6	97.2
OSEM-BW	2	9.9	65.3	76.9	98.3
AR-OSEM-AR	4	136.5	253.0	270.4	296.5
BW-FBP	4	84.9	229.2	259.2	289.1
OSEM-BW	4	70.2	204.1	247.1	300.0
		B			
AR-OSEM-AR	0	21.5	58.0	76.1	84.8
BW-FBP	0	20.4	61.6	83.2	91.8
OSEM-BW	0	32.9	61.9	78.0	86.6
AR-OSEM-AR	2	32.8	81.0	81.9	95.7
BW-FBP	2	35.1	80.2	87.4	102.7
OSEM-BW	2	29.7	75.4	78.8	98.3
AR-OSEM-AR	4	132.8	230.2	269.8	295.7
BW-FBP	4	94.6	236.9	270.3	316.2
OSEM-BW	4	85.6	224.6	255.1	308.5
		C			
AR-OSEM-AR	0	24.0	55.9	75.0	85.5
BW-FBP	0	32.9	68.7	83.0	93.0
OSEM-BW	0	31.2	60.7	76.8	87.8
AR-OSEM-AR	2	41.1	76.0	86.9	94.9
BW-FBP	2	38.8	82.4	93.9	103.6
OSEM-BW	2	34.3	83.1	85.5	100.6
AR-OSEM-AR	4	119.4	237.1	266.3	297.7
BW-FBP	4	112.7	248.5	274.5	318.2
OSEM-BW	4	108.1	260.5	261.0	316.9

A: low count level; B: intermediate count level; C: high count level; AR-OSEM-AR: autoregressive filtering before and after ordered subset expectation maximisation algorithm; BW-FBP: Butterworth prefiltering and filtered back projection; OSEM-BW: ordered subset expectation maximisation algorithm and Butterworth postfiltering; RelAct: activity relative to background activity of 1; Ø: diameter.

Table 4: Full width at half maximum values for the different methods.

Method	RelAct	10 Ø	20 Ø	40 Ø	60 Ø
		A			
AR-OSEM-AR	0	19.7	25.0	45.8	64.8
BW-FBP	0	17.9	23.2	43.5	63.8
OSEM-BW	0	19.4	23.2	43.6	62.8
AR-OSEM-AR	2	—	21.5	35.3	59.1
BW-FBP	2	—	23.2	36.6	59.6
OSEM-BW	2	—	22.2	35.8	59.1
AR-OSEM-AR	4	13.4	18.4	37.8	57.8
BW-FBP	4	15.8	20.4	37.9	57.9
OSEM-BW	4	15.2	19.9	37.7	57.9
		B			
AR-OSEM-AR	0	23.7	23.7	43.0	62.3
BW-FBP	0	18.6	22.4	41.2	61.4
OSEM-BW	0	18.8	23.2	41.9	61.7
AR-OSEM-AR	2	19.4	19.8	38.9	59.0
BW-FBP	2	19.0	21.0	38.5	58.9
OSEM-BW	2	17.7	19.9	38.5	59.0
AR-OSEM-AR	4	13.0	19.7	37.8	58.2
BW-FBP	4	15.4	21.0	37.7	57.9
OSEM-BW	4	14.2	20.2	37.9	58.1
		C			
AR-OSEM-AR	0	20.2	23.9	40.9	63.2
BW-FBP	0	18.5	21.1	40.6	61.7
OSEM-BW	0	16.8	23.2	40.3	62.3
AR-OSEM-AR	2	15.8	19.5	39.0	58.6
BW-FBP	2	16.9	20.8	38.6	58.7
OSEM-BW	2	16.5	20.3	39.1	58.6
AR-OSEM-AR	4	12.5	19.0	37.6	58.2
BW-FBP	4	14.0	19.6	36.8	58.0
OSEM-BW	4	13.8	19.3	37.4	57.9

A: low count level; B: intermediate count level; C: high count level; AR-OSEM-AR: autoregressive filtering before and after ordered subset expectation maximisation algorithm; BW-FBP: Butterworth prefiltering and filtered back projection; OSEM-BW: ordered subset expectation maximisation algorithm and Butterworth postfiltering; RelAct: activity relative to background activity of 1; Ø: diameter.

position and area were equal in every image. The ROIs were drawn using Multimodality software. The CR% values were obtained using the average counts in the ROIs.

Spatial resolution was estimated by the full width at half maximum (FWHM) of the line spread functions of the cylinders. One-, two-, four- and six-pixel-thick profiles were drawn through the 10-, 20-, 40- and 60-mm-wide cylinders, respectively. The FWMH values were calculated using Hermes quality control software (version 2.0).

2.5. Patient Study. Skeletal SPECT was performed three hours after an intravenous injection of 925 MBq of ^{99m}Technetium-labelled methylene diphosphonate. The images were obtained over a 360° arc, using 64 projections at 20 sec per projection. The images were acquired into a 128 × 128 matrix with a pixel size of 4.8 mm. Total counts of the projection images were 41006–66830 counts per projection.

2.6. Statistical Methods. The data were analysed using WinSTAT for Excel (version 2007.1; R. Fitch Software, Staufen, Germany). Pair-wise comparisons were performed with the nonparametric Wilcoxon's rank-sum test. Comparisons were made between the AR-OSEM-AR and BW-FBP methods, the AR-OSEM-AR and OSEM-BW methods, and the BW-FBP and OSEM-BW methods. Data from the cold stacks and the pooled hot stacks were analysed separately. For each cylinder, the paired difference between the values of a variable was computed. The values of the differences were sorted to get a rank order. Finally, the mean rank of negative differences was compared with that of positive differences. Wilcoxon's rank-sum test determines to what extent the difference in mean rank is significant. A P value of less than 0.05 was considered significant.

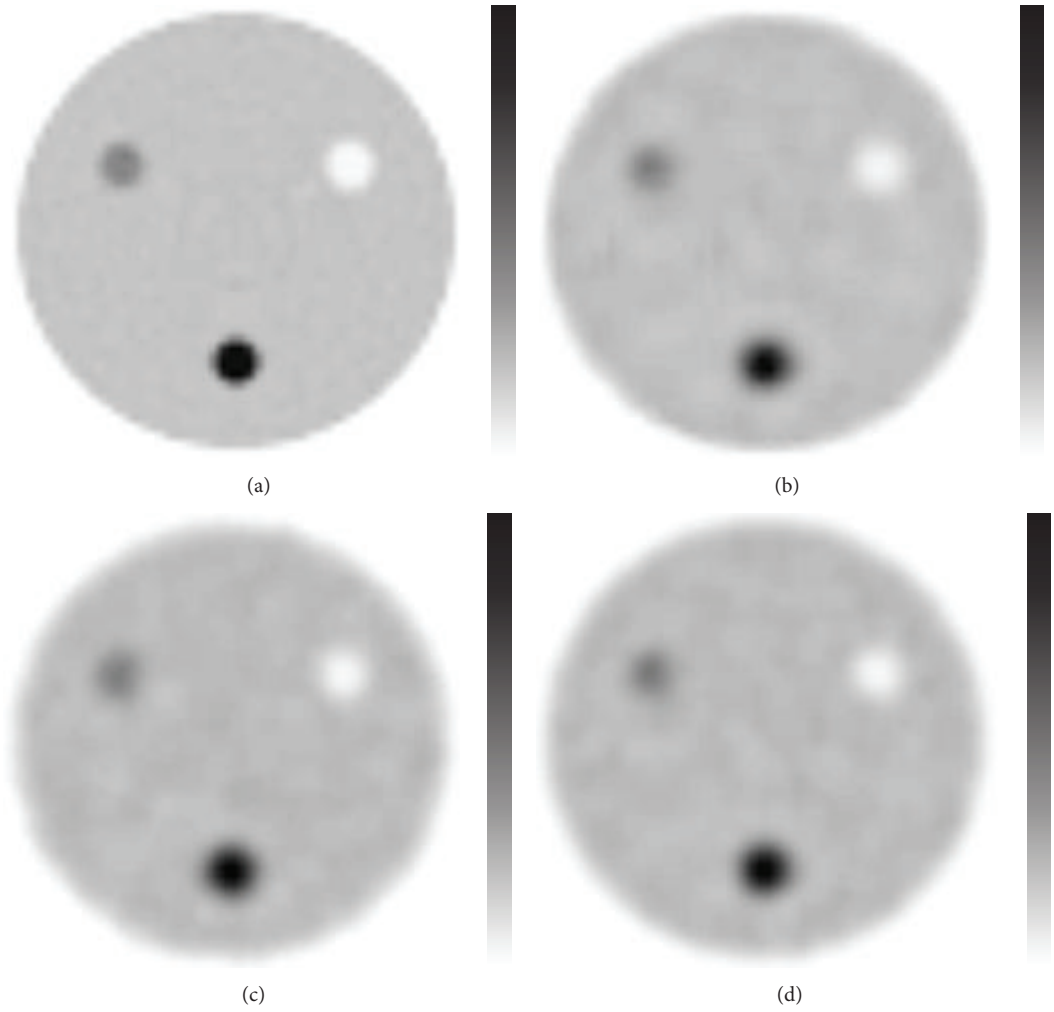


Figure 3: Transaxial slices of the phantom at the level of cylinders with a diameter of 20 mm. (a) Images reconstructed from noise-free projection images using ordered subset expectation maximisation reconstruction. (b) Autoregressive filtering before and after ordered subset expectation maximisation reconstruction. (c) Filtered back projection reconstruction method preceded by Butterworth filtering. (d) Ordered subset expectation maximisation reconstruction followed by Butterworth filtering. Intermediate count level. The images are individually scaled to their own maximum.

3. Results

The MSE of the images improved when a different AR model was used for the summed error term image rather than for the original image. A prediction region of four orthogonal neighbours with a block size of 5×5 pixels for the original image and a prediction region of 3×3 and a block size of 6×6 for the summed error term image produced the lowest MSE, although the differences were small (Table 1). Part of the counts at the edges of the image could be returned to the filtered image to reduce blurring of the image (Figure 2).

Butterworth filtering was chosen so that the methods had the same amount of statistical fluctuation in the uniform part of the phantom, as confirmed by the CoV% values (Table 2). For the cold stacks, the BW-FBP method showed higher CR% values than the AR-OSEM-AR and OSEM-BW methods (Table 3). The P values were 0.003 and 0.04, respectively. The BW-FBP method had the highest CR% values for all other cold cylinders except the two smallest cylinders at the intermediate count level. The OSEM-BW method, in turn, displayed better performance than the AR-OSEM-AR method ($P = 0.002$). When the hot stacks were assessed, there were no statistically significant differences between the AR-OSEM-AR and BW-FBP methods nor between the AR-OSEM-AR and OSEM-BW methods, but the BW-FBP method showed higher CR% values than the OSEM-BW method ($P = 0.001$).

In the analysis of spatial resolution, without exception, the BW-FBP and OSEM-BW methods exhibited lower FWHM values for the cold stacks than the AR-OSEM-AR method ($P = 0.002$ for both comparisons), but there was no statistical difference between the BW-FBP and OSEM-BW methods (Table 4). For the hot stacks, the AR-OSEM-AR method showed lower FWHM values than the BW-FBP and OSEM-BW methods. The P values were 0.01 and 0.04, respectively.

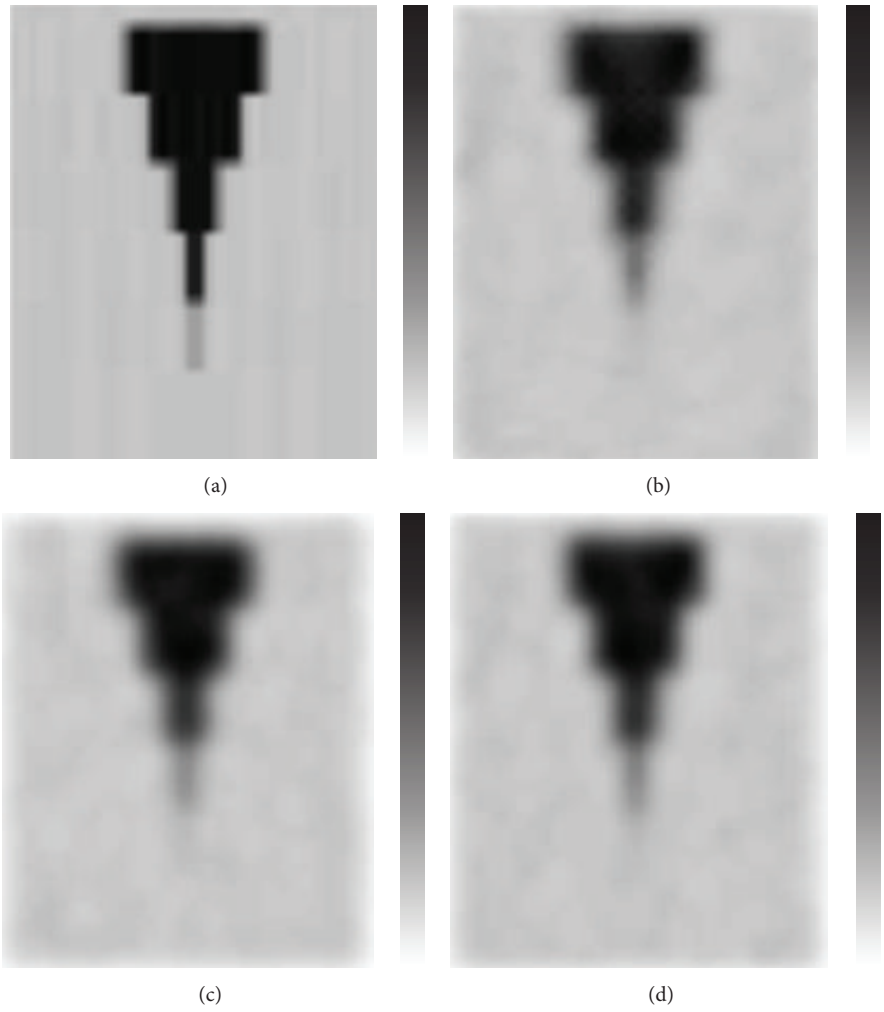


Figure 4: Reformatted coronal slices of the phantom. (a) Images reconstructed from noise-free projection images using ordered subset expectation maximisation reconstruction. (b) Autoregressive filtering before and after ordered subset expectation maximisation reconstruction. (c) Filtered back projection reconstruction method preceded by Butterworth. (d) Ordered subset expectation maximisation reconstruction followed by Butterworth filtering. Stacks with the highest activity. Intermediate count level. The images are individually scaled to their own maximum.

Furthermore, the OSEM-BW method had lower FWHM values than the BW-FBP method ($P = 0.008$). It is of note that the AR-OSEM-AR method showed better resolution than the other two methods in the analysis of the two smallest hot stacks, with two exceptions in the analysis of cylinders with two times the background activity and a diameter of 10 mm (Table 4).

Visually, the differences between the images produced by the three methods were small (Figures 3, 4, and 5). When comparing the skeletal SPECT data, the BW-FBP method showed lower image quality than the two other methods because of streak artefacts (Figure 6).

4. Discussion

This paper presented an improved two-dimensional adaptive AR filter and introduced a three-dimensional adaptive AR model for reduction of noise in SPECT images. We demonstrated that the quality of scintigraphic images can be improved when the same AR procedure is not applied to the original image and the summed error term image. We have previously shown that if a prediction region of four orthogonal neighbours of the predicted pixel with a block size of 5×5 pixels is used for both the original image and the summed error term image in the same simulation conditions, then the mean squared errors for the three different images with Poisson statistics are 0.85, 2.23, and 7.12 [4]; that is, this combination exhibits lower performance than any of those presented in Table 1.

The goal of filtering in SPECT is to suppress statistical noise and simultaneously preserve contrast and spatial resolution [1]. In the present study we showed that the AR-OSEM-AR method simultaneously provides both efficient noise rejection and good spatial resolution for hot objects. The methodological comparison was done using the well-known *de facto* reconstruction standards, FBP and OSEM,

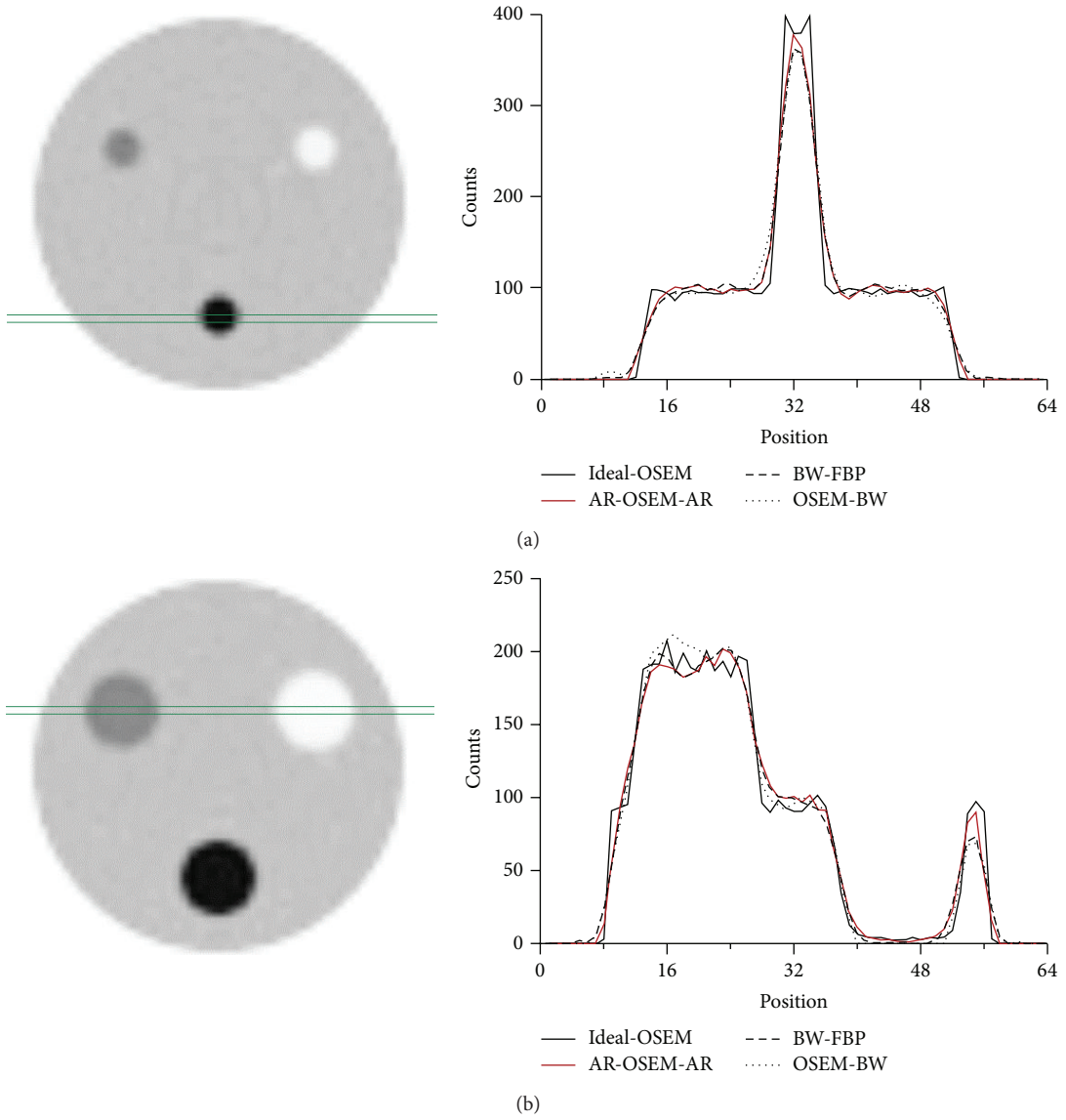
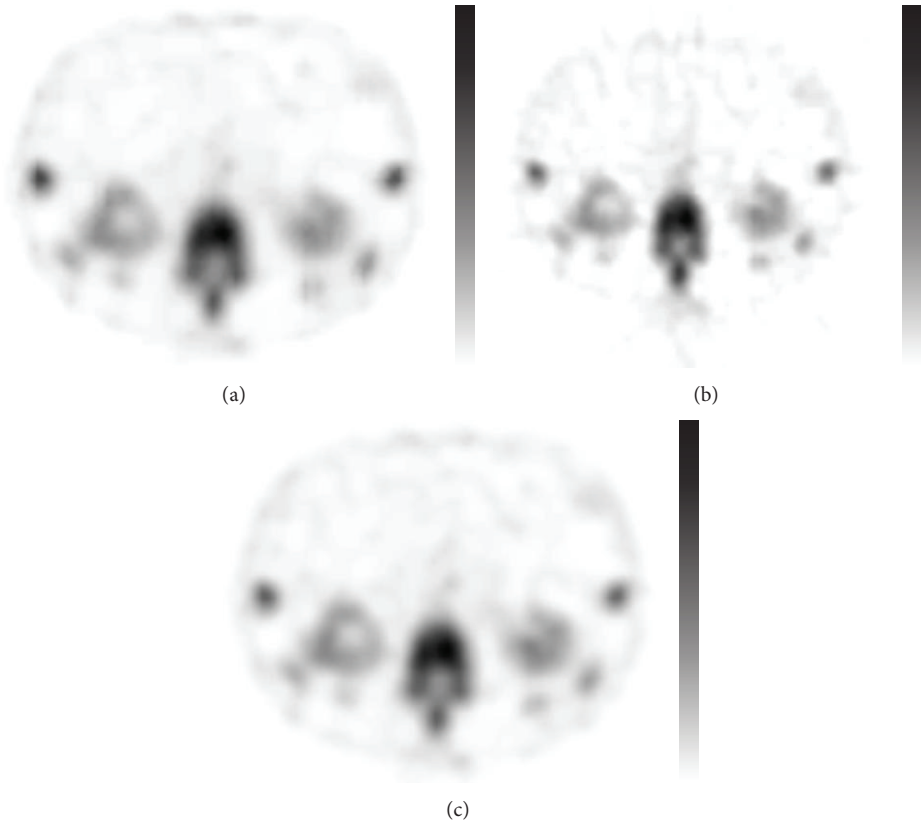


Figure 5: One-pixel-thick profiles drawn through the phantom. (a) Profiles at the level of a cylinder with a diameter of 20 mm. (b) Profiles at the level of cylinders with a diameter of 40 mm. Ideal OSEM: images reconstructed from noise-free projection images using ordered subset expectation maximisation reconstruction; AR-OSEM-AR: autoregressive filtering before and after ordered subset expectation maximisation reconstruction; BW-FBP: filtered back projection reconstruction method preceded by Butterworth filtering; OSEM-BW: ordered subset expectation maximisation reconstruction followed by Butterworth filtering. Intermediate count level. Profiles were rescaled so that they had the same amount of counts as the profile of the image reconstructed from noise-free projection images using ordered subset expectation maximisation reconstruction.

and by using one of the most commonly used filters in nuclear medicine, the Butterworth filter.

The BW-FBP method produced better performance than the two other methods in the analysis of the cold stacks. FBP's good performance with cold features has been noticed before [5, 13]. OSEM's built-in nonnegativity constraint explains its poor contrast in cold regions. In the analysis of the hot stacks, the magnitude of the differences between the three methods proved to be small, but the AR-OSEM-AR method had statistically the best performance. This is obviously due to the fact that part of the counts at the edges of the error term images could be returned back to the filtered image. Adding postfiltering to the method produced efficient noise reduction without compromising contrast or spatial resolution significantly.

The FBP method consists of filtering of the projection data and back projection of the filtered data [5, 10]. Prefiltering is generally not applied in the OSEM method because it lowers spatial resolution. Secondly, OSEM assumes that projection pixel values are independent and the number of counts is Poisson-distributed. Filtering might hamper



(a) (b) (c)

Figure 6: Transaxial slice of skeletal SPECT. (a) Autoregressive filtering before and after ordered subset expectation maximisation reconstruction. (b) Filtered back projection reconstruction method preceded by Butterworth filtering. (c) Ordered subset expectation maximisation reconstruction followed by Butterworth filtering.

these assumptions. OSEM reconstructed images are usually postfiltered because the images become noisier as the iterations proceed.

The disadvantage of FBP is that it can produce radial streak artefacts because filtered noisy projection profiles do not cancel each other out in back projection. In the present study, the phenomenon was seen in the clinical data. Iterative reconstruction algorithms also provide some other advantages over FBP. They permit the use of several important corrections, such as scatter, attenuation, and collimator response corrections, which can be included in the image reconstruction procedure. Incorporation of anatomical information derived from magnetic resonance imaging or computerized tomography is possible as well [14]. For the abovementioned reasons, FBP has in recent years been progressively replaced with iterative reconstruction algorithms.

The Butterworth filter is defined by two parameters: cut-off frequency and order [2]. In the present study, order was set to 2 because a ringing artefact is imperceptible to Butterworth filters of order 2 but can become a significant factor in filters of a higher order [15]. Filter orders much higher than 2 are often seen in clinical praxis. Edge sharpness in images produced by the BW-FBP and OSEM-BW methods can be improved by increasing the cut-off frequency, but the improvement occurs at the expense of increased noise.

In our opinion, a strength of the AR-OSEM-AR method is its simplicity, but the lack of user-controlled variables can also be regarded as a limitation. Sometimes, adjustable parameters are needed. No particular filter can emerge as the best filter for any organ system. However, filtering should be performed locally in the spatial domain, not globally in the frequency domain, because the correct trade-off between resolution and smoothing will vary at different points within the image.

Because the AR-OSEM-AR method was only marginally better than the OSEM-BW method, an additional study is needed to find out whether image quality will be even better if the AR method is applied to the intermediate results in between the iterations. Secondly, the AR-OSEM-AR method has not yet been tested with positron emission tomography (PET) data, but the method should also be suitable for PET data. The signal-to-noise ratio is considerably higher in PET than in SPECT. Therefore, our model will probably provide a good fit for PET data.

5. Conclusions

The AR-OSEM-AR method is a feasible denoising method in SPECT. It has good spatial resolution for hot features and it is simple to use. It does not have any adjustable parameters.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M. Lyra and A. Ploussi, "Filtering in SPECT image reconstruction," *International Journal of Biomedical Imaging*, vol. 2011, Article ID 693795, 14 pages, 2011.

[2] P. Zanzonico, "Technical requirements for SPECT: instrumentation, data acquisition and processing, and quality control," in *Clinical SPECT Imaging*, E. L. Kramer and J. J. Sanger, Eds., pp. 7–41, Raven Press, New York, NY, USA, 1995.

[3] T. S. Kangasmaa, J. T. Kuikka, E. J. Vanninen, H. M. Mussalo, T. P. Laitinen, and A. O. Sohlberg, "Half-time myocardial perfusion SPECT imaging with attenuation and Monte Carlo-based scatter correction," *Nuclear Medicine Communications*, vol. 32, no. 11, pp. 1040–1045, 2011.

[4] R. Takalo, H. Hytti, and H. Ihalainen, "Adaptive autoregressive model for reduction of poisson noise in scintigraphic images," *Journal of Nuclear Medicine Technology*, vol. 39, no. 1, pp. 19–26, 2011.

[5] J. Biemond, R. L. Lagendijk, and R. M. Mersereau, "Iterative methods for image deblurring," *Proceedings of the IEEE*, vol. 78, no. 5, pp. 856–883, 1990.

[6] H. Onishi, N. Motomura, M. Takahashi, M. Yanagisawa, and K. Ogawa, "A 3-dimensional mathematic cylinder phantom for the evaluation of the fundamental performance of SPECT," *Journal of Nuclear Medicine Technology*, vol. 38, no. 1, pp. 42–48, 2010.

[7] L. Li, W. Hou, X. Zhang, and M. Ding, "GPU-based block-wise nonlocal means denoising for 3D ultrasound images," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 921303, 10 pages, 2013.

[8] I. G. Zubal, C. R. Harrell, E. O. Smith, Z. Rattner, G. Gindi, and P. B. Hoffer, "Computerized three-dimensional segmented human anatomy," *Medical Physics*, vol. 21, no. 2, pp. 299–302, 1994.

[9] R. A. Brooks and D. G. Chiro, "Principles of computer assisted tomography (CAT) in radiographic and radioisotopic imaging," *Physics in Medicine and Biology*, vol. 21, no. 5, pp. 689–732, 1976.

[10] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[11] S. Butterworth, "On the theory of filter amplifiers," *Wireless Engineering*, vol. 7, pp. 536–541, 1930.

[12] W. A. Edelstein, P. A. Bottomley, H. R. Hart, and L. S. Smith, "Signal, noise, and contrast in nuclear magnetic resonance (NMR) imaging," *Journal of Computer Assisted Tomography*, vol. 7, no. 3, pp. 391–401, 1983.

[13] S. D. Wollenweber and K. L. Gould, "Investigation of cold contrast recovery as a function of acquisition and reconstruction parameters for 2D cardiac PET," in *Proceedings of the IEEE Nuclear Science Symposium Conference Record (NSS/MIC '05)*, vol. 5, pp. 2552–2556, October 2005.

[14] G. Gindi, M. Lee, A. Rangarajan, and I. G. Zubal, "Bayesian reconstruction of functional images using anatomical information as priors," *IEEE Transactions on Medical Imaging*, vol. 12, no. 4, pp. 670–680, 1993.

[15] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Prentice Hall, Upper Saddle River, NJ, USA, 2002.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 182659, 21 pages
http://dx.doi.org/10.1155/2015/182659

Research Article

The EM Method in a Probabilistic Wavelet-Based MRI Denoising

Marcos Martin-Fernandez[1] and Sergio Villullas[2]

[1]*Laboratorio de Procesado de Imagen, Escuela Técnica Superior de Ingenieros de Telecomunicación, Campus Miguel Delibes s.n., 47011 Valladolid, Spain*
[2]*Departamento de Álgebra, Análisis Matemático, Geometría y Topología, Facultad de Ciencias, Campus Miguel Delibes s.n., 47011 Valladolid, Spain*

Correspondence should be addressed to Sergio Villullas; svillu@am.uva.es

Received 1 July 2014; Revised 23 September 2014; Accepted 30 September 2014

Academic Editor: Yi Gao

Human body heat emission and others external causes can interfere in magnetic resonance image acquisition and produce noise. In this kind of images, the noise, when no signal is present, is Rayleigh distributed and its wavelet coefficients can be approximately modeled by a Gaussian distribution. Noiseless magnetic resonance images can be modeled by a Laplacian distribution in the wavelet domain. This paper proposes a new magnetic resonance image denoising method to solve this fact. This method performs shrinkage of wavelet coefficients based on the conditioned probability of being noise or detail. The parameters involved in this filtering approach are calculated by means of the expectation maximization (EM) method, which avoids the need to use an estimator of noise variance. The efficiency of the proposed filter is studied and compared with other important filtering techniques, such as Nowak's, Donoho-Johnstone's, Awate-Whitaker's, and nonlocal means filters, in different 2D and 3D images.

1. Introduction

Magnetic resonance imaging (MRI) is one of the most important imaging acquisition techniques [1], which allows studying the structural features of the internal body parts noninvasively. This procedure is based on the principle of nuclear magnetic resonance (NMR) [2], and the power of this technique over other noninvasive techniques, such as ultrasound, is the high quality of its images, despite the inconvenience of being a larger and expensive equipment. However, because the ultrasonic signal is not transmitted through bones, in cases such as imaging the brain which is surrounded by the skull, ultrasound modalities are not viable and resorting to magnetic resonance (MR) imaging is needed.

For a given acquisition time, in MRI there is a fundamental agreement between resolution and signal to noise ratio (SNR) [3]. MRI is affected by noise mainly produced by interference due to human body heat emission (Gaussian at frequency space and Rayleigh at envelope of its inverse Fourier transform [4]), which prevents correct identification of shapes and details. Moreover, there is a relationship between noise level and image resolution in MRI acquisition, that is, the larger the resolution of the acquired image, the lower the SNR [5]. The simplest method to reduce the noise level is to increase the acquisition time of the machine (thus increasing the number of images averaged, that is, increasing the machine number of experiments (NEX)), which would cause a large increase in spending and long waiting lists, but long acquisition times could be problematic for patients who are not able to remain in a resting state (due to stress, pain, and so on). To avoid these problems, a filter, which acts on the acquired image, can be applied. This filter must eliminate noise trying to preserve details. The main problem of removing noise by means of softening an image is the resulting loss of information on the edges and contours (image blur), which is typical in Gaussian filter

convolution. Perona and Malik [6] proposed a new type of filtering, based on partial differential equations and the heat diffusion equation, which is the origin of a family of filters that allow homogenizing regions while maintaining or enhancing borders between them. Gerig et al. [7] propose the use of the so-called nonlinear anisotropic filter, which gives very good results in the context of MRI and is an important practical application of the ideas proposed by Perona and Malik. From the same diffusion equation, [7] proposes an alternative discretization with simpler formulation, whose stability is subject to certain restrictions of the parameters. Linearly optimal methods in the sense of minimum mean square error (Wiener filter) have also been adapted to the case of MRI [8, 9]. Another filter with good MR image denoising results is the so-called nonlocal means (NLM) filter [10, 11], which averages similar image pixels as a function of their intensity distance (some filters, like the bilateral filter [4], are based on the same proposition, but the advantage of the NLM over other methods is that the similarity measure used is more robust to noise due to region comparisons rather than pixel comparisons). Principles of nonparametric statistical methods are also the base of the iterated conditional entropy reduction (ICER) proposed by Awate and Whitaker [12], a Bayesian-inference algorithm based on Markov random field which estimates the uncorrupted-image statistics by optimizing an information-theoretic metric using the expectation-maximization algorithm. This filter method incorporates a Rician noise model, unlike NLM method, which is more general. He and Greenshields [13] designed another filter method that improves NLM filter by adding Rician noise information. Dabov et al. [14] also proposed a filter similar to NLM. This method creates 3D arrays formed by stacking together similar image 2D neighborhoods. The importance of grouping is to enable the use of a higher-dimensional filtering of each group, which exploits the potential similarity (correlation, affinity, etc.) between grouped fragments. More generally, Sivaramakrishnan and Weissman [15] designed a universal filter that does not need *a priori* noise information which is asymptotically optimal. Furthermore, Awate and Whitaker also proposed a patch-based method [16, 17] that tries to optimize the entropy of the noisy image to reduce noise.

A large interesting characteristic of the wavelet transform is its capacity to preserve detail at different scales, because of its ability to model the information locally present in the image due to the multiresolution decomposition [18]. In the literature, other authors have performed MR image filtering using wavelet techniques. Donoho and Johnstone [19] proved that a simple thresholding algorithm with an appropriated base may be a (nonlinear) filter which is almost optimal. Nowak [20] and Pizurica et al. [21] propose to perform the filtering using the discrete wavelet transform. In particular, the significance of Nowak's work is that it uses the fact that MR magnitude image obeys a Rician distribution and its square image noise obeys a noncentral Chi-square distribution. In addition, other researches, like Sijbers et al. [22], use this fact. A diffusion method such as Perona-Malik's, but adapted to the Rician distribution case, has been proposed in [23]. In [20], the wavelet transform is performed on the square of the amplitude image where the noise and the bias of approximation coefficients are reduced. Another interesting work is that of Anand and Sahambi [4]. In this case, the square of the amplitude image is also used, correcting the bias and applying a bilateral filter (Gaussian filtering in the spatial and amplitude domains) over approximation coefficients (it is also based on the Rician image distribution). On the other hand, Yang and Fei [24] combined the 1D wavelet transform with the Radon transform to denoise Rician noise in MR images. Finally, we can also mention the method proposed by Wirestam et al. [25], where a technique of shrinkage coefficients of the filter is based on a Wiener filter. The novelty of this method is that the filtering is performed in the complex Fourier image, where noise data are complex Gaussian. This method has the problem that complex data are not always available in MRI acquisition (usually, MR machines provide only data in the image plane after envelope calculation in DICOM format. Complex raw data in the k-space are normally stored in a proprietary format which is not open and is brand-dependent. The complex data in the image plane are not even stored in the machine).

Based on wavelets state of art and given locality property of wavelet transform, the alternative that we propose in this paper is to perform filtering in the domain of the transform coefficients, adapting to MRI the method proposed in [26], originally designed for mammographic images. We also propose to estimate the parameters of the model by means of the expectation maximization (EM) method, proposed by Dempster et al. [27] and Moon [28]. Moreover, this fact makes the filter independent of noise variance estimators, in contrast to other filter methods.

This paper is structured as follows. In Section 2, some different MR image denoising methods are detailed. First, in Section 2.1, the methods proposed by Donoho and Johnstone [19] (Section 2.1.1) and Nowak [20] (Section 2.1.2) are described. Section 2.1.3 shows shift variance of wavelets coefficients and how to correct it. This section also contains Section 2.2 (with Sections 2.2.1 and 2.2.2), where Awate-Whitaker's algorithm [16, 17] and nonlocal means filter [10, 11] are presented. Finally, Section 2.3 describes the used technique to estimate noise variance. In Section 3, the new wavelet denoising method, based on [26], is presented. Section 4 presents the algorithms corresponding to the methods showed in Sections 2 and 3 and some practical experiments. First, Section 4.1 shows a step-by-step explanation of the different denoising algorithms presented before. Second, in Section 4.2, the images used to test the efficiency of the proposed method are detailed. The measurements used to compare the different filter methods are described in Section 4.3. Fourth, in Section 4.4, the numerical experiments and some remarks, obtained after performing these experiments, are presented. Finally, Section 5 contains a brief summary of the obtained results and some possible future research lines. Appendix A contains the details about the optimization method proposed to estimate the parameters of the filter presented in Section 3 and Appendix B contains some useful auxiliary functions.

2. State of Art

This section reviews some filters used in the practical experiments of Section 4. The new filter proposed in this paper is compared with other two wavelet filters, Donoho-Johnstone's hard thresholding filter [19] and Nowak's [20] filter, the Awate and Whitaker's [16, 17] patch-based (Gaussian and Rician) filters and the nonlocal means [10, 11] filter. Moreover, wavelet-based filtering is shift-variant, as Section 2.1.3 shows. This subsection also presents how to avoid this problem. Finally, an estimator of noise variance, required in both wavelet-based filters, is given in Section 2.3.

2.1. Wavelet-Domain Filters. An image/volume can be interpreted as a 2-dimensional/3-dimensional function with compact support. The values of this function, represented in a matrix/3D array I, are a good approximation to scale coefficients, s_0, in discrete wavelet transform. The fast wavelet transform algorithm let us calculate the scale, s_j, and detail, d_j, coefficients in the following levels $0 \leq j \leq J$, that is, if $s_0 = I$, s_{j+1} and d_{j+1} can be calculated in function of s_j. Noise interferences modify the details of the MR image/volume; as noise grows more levels are affected. The wavelet coefficients are filtered to denoise the image/volume. The wavelet coefficients $d^\alpha[k]$ (the index α corresponds to level and orientation: horizontal, vertical, diagonal, and so on, and index k corresponds to scale and position) can be determined by means of

$$d^\alpha[k] := \sum_m \psi^\alpha[k](m) I(m), \tag{1}$$

where $\psi^\alpha[k]$ is the discrete wavelet function at α level and orientation and k scale and position, and index m represents pixel/voxel position. Similarly the scaling coefficient $s^j[k]$ (the index j corresponds to level and index k corresponds to scale and position) can be determined by means of

$$s^j[k] := \sum_m \varphi^j[k](m) I(m), \tag{2}$$

where $\varphi^j[k]$ is the scaling function for level j and k scale and position. Given a sequence of wavelet coefficients $d^\alpha = (d^\alpha[k])_{k=1}^{N^\alpha}$ and scaling coefficients $s^j = (s^j[k])_{k=1}^{N^j}$ for the image/volume I (where N^α represent the number of scales and positions for each α and N^j is the number scales and positions for each j) the filters can be defined in the wavelet domain.

The following two filters, proposed by Donoho and Johnstone and Nowak, are used to analyze the efficiency of the new filter presented in next section.

2.1.1. Donoho-Johnstone's Filter. The classic hard thresholding filter described by Donoho and Johnstone [19] is given by

$$F_{\mathrm{DJ}}(d^\alpha[k]) := \begin{cases} d^\alpha[k], & \text{if } |d^\alpha[k]| > T^\alpha, \\ 0, & \text{if } |d^\alpha[k]| \leq T^\alpha, \end{cases} \tag{3}$$

where $|\cdot|$ is the module operator and $T^\alpha := \sigma_{\mathrm{noise}}\sqrt{2\log_e(N^\alpha)}$ with σ_{noise} standard deviation of the noise in the image/volume I.

2.1.2. Nowak's Filter. Another filter is proposed by Nowak [20] by

$$F_N(d^\alpha[k]) := C^\alpha[k] d^\alpha[k], \tag{4}$$

where

$$C^\alpha[k] := \left(\frac{(d^\alpha[k])^2 - 3(\sigma^\alpha)^2[k]}{(d^\alpha[k])^2}\right)_+, \qquad (x)_+ := \begin{cases} x, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0, \end{cases} \tag{5}$$

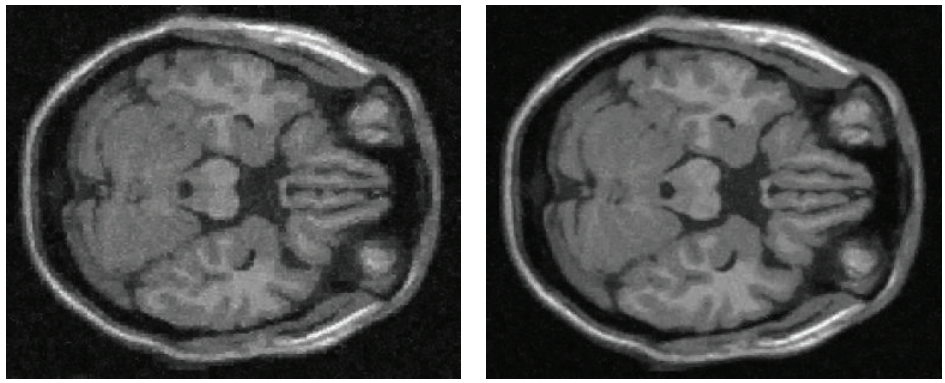
and $(\sigma^2)^\alpha[k]$ is the variance of the wavelet coefficient $d^\alpha[k]$. An estimation for that, $\widehat{(\sigma^\alpha)^2[k]}$, is proposed in [20]

$$\widehat{(\sigma^\alpha)^2[k]} := 4\sigma_{\mathrm{noise}}^4 \max\left[\frac{\sum_{m\in I}(\psi^\alpha)^2[k](m) I^2(m)}{\sigma_{\mathrm{noise}}^2} - 1, 1\right], \tag{6}$$

where σ_{noise} is the standard deviation of the noise in the image/volume I.

2.1.3. Shift-Invariant Filtering. As [20] shows, wavelet coefficients filtering (based on discrete wavelet transform) is shift-variant, which can produce the appearance of artifacts, because the wavelet coefficient values depend on the alignment between the data and the wavelet basis functions. Shift-invariant (translation-invariant, undecimated) methods [29–33] can provide better performance, avoiding the appearance of artifacts (see an example in Figure 1). Shift-invariant wavelet transforms provide a higher degree of regularity [29, 33, 34] than standard wavelet analysis approaches, so shift-invariant estimation algorithms usually outperform standard methods. A shift-invariant filtering can be obtained by applying the filter for every possible shift of the image, unshifting each filtered result, and averaging all the results obtained. As performing all shifts of the image would be computationally too expensive, to reduce computational burden of filtering, an approximately shift-invariant scheme is proposed; that is, all shifts are replaced by a small range of shifts. More specifically, the image is shifted in both horizontal and vertical directions (left, right, up, and down), at most K pixels/voxels in each direction in steps of one pixel/voxel. While this does not guarantee shift-invariance, it does reduce the dependence of the filter output on the alignment between the data and the wavelet basis functions. In our experiments, we used $K = 2$, that is, movements of $\{-2, -1, 0, 1, 2\}$ in each direction giving rise to a total of 25 shiftings for the 2D images (the values $K > 2$ enlarge computational time excessively with minor improvements), and $K = 1$, that is, movements of $\{-1, 0, 1\}$ in each direction giving rise to a total of 27 shiftings for the 3D volumes (the same remark as in 2D also applies in 3D for $K > 1$). The final image/volume is constructed after unshifting each image/volume and averaging the 25/27 resulting images/volumes.

2.2. Awate-Whitaker's Patch-Based Filter and Nonlocal Means Filter. For quantitative evaluation, these two wavelet methods (and the new proposed filter) are compared (in 2D case)



(a) (b)

Figure 1: Comparative images: (a) without shift-invariant filtering; (b) with shift-invariant filtering ($K = 2$). Saw tooth (artifacts) can be seen in the case (a).

with a patch-based method proposed by Awate and Whitaker [16, 17] and the nonlocal means (NLM) filter [10, 11]. This comparison will be very significative as this method was recently shown to be competitive with respect to wavelet methods.

2.2.1. Awate-Whitaker's Filter. In this approach, given an image I, a random vector $Z(m) = [X(m), Y(m)]$, where m represents the pixel position on the image I, is generated, where $X(m)$ is the intensity of I at the pixel m and $Y(m)$ is the intensity of I at the pixels in a neighborhood ($Y(m)$ is a vector) of pixel m (we use $\widetilde{X}(m)$, $\widetilde{Y}(m)$, and $\widetilde{Z}(m)$ for degraded image random variables). The target of the method is to minimize the entropy of the conditional PDF, $h(\widetilde{X} \mid \widetilde{Y})$, for which a descending gradient method, given by

$$\widehat{x} = x - \delta \frac{\partial h}{\partial x}, \tag{7}$$

where δ is a parameter, is used. In this paper, we used two versions of this filter method using Gaussian and Rician models for the PDF.

2.2.2. Nonlocal Means Filter. For a given image I, the NLM filtered image at pixel position m (F_{NLM} is the NLM filter operator) is given by the weighted average of all the pixels in a searched area Ω_m of pixel position m in the image I,

$$F_{\text{NLM}}\left(I\left(m\right)\right) = \sum_{n \in \Omega_m} w\left(m,n\right) I\left(n\right), \tag{8}$$

where $0 \leq w(m,n) \leq 1$, $\sum_{n \in \Omega} w(m,n) = 1$. The weights $w(m,n)$ are based on the similarity between the neighbourhoods of pixels $I(m)$ and $I(n)$ and are defined as

$$w\left(m,n\right) = \frac{e^{-(d(N_m,N_n)/h^2)}}{\sum_{n \in I} e^{-(d(N_m,N_n)/h^2)}}, \tag{9}$$

where N_m and N_n are the neighborhoods of the pixel positions m and n, respectively, d is a Gaussian weighted squared Euclidean distance, and h is the exponential decay control smoothing parameter. Region Ω can be the whole image, but, because of computational reasons, Ω uses to be a smaller region in the local neighborhood.

2.3. Estimation of σ^2_{noise}. Donoho and Johnstone's [19] and Nowak's [20] methods are highly dependent on the noise estimate. We assume that noise distribution, in complex domain, is zero mean Gaussian. Then, noise estimate means noise standard deviation/variance ($\sigma_{\text{noise}}/\sigma^2_{\text{noise}}$) estimate. In MR images, σ_{noise} is usually unknown *a priori* and it must be estimated from the data. A good estimator of σ^2_{noise} of I is given in [35, 36] by

$$\widehat{\sigma}^2_{\text{noise}} := \frac{2}{4-\pi} \text{Mode}\left\{\sigma^2_{\text{local}}\right\}, \tag{10}$$

where σ^2_{local} is the local variance of I, defined as

$$\sigma^2_{\text{local}}\left(m\right) := \text{LV}\left(I,m\right), \tag{11}$$

where the operator LV is defined in Appendix B.

3. A New Probabilistic Wavelet Filter

The wavelet filter proposed in this paper, which will be named hereafter as Villullas-Martin's filter, is defined by the shrinkage

$$F_{\text{VM}}\left(d^{\alpha}\left[k\right]\right) := S^{\alpha}_{\lambda}\left[k\right] d^{\alpha}\left[k\right], \tag{12}$$

where d^{α} are the wavelet coefficients at level/orientation α of an image/volume I and

$$S^{\alpha}_{\lambda}\left[k\right] := \frac{(1-\lambda)\, P\left(d^{\alpha}\left[k\right] \mid \text{detail}\right)}{(1-\lambda)\, P\left(d^{\alpha}\left[k\right] \mid \text{detail}\right) + \lambda P\left(d^{\alpha}\left[k\right] \mid \text{noise}\right)} \tag{13}$$

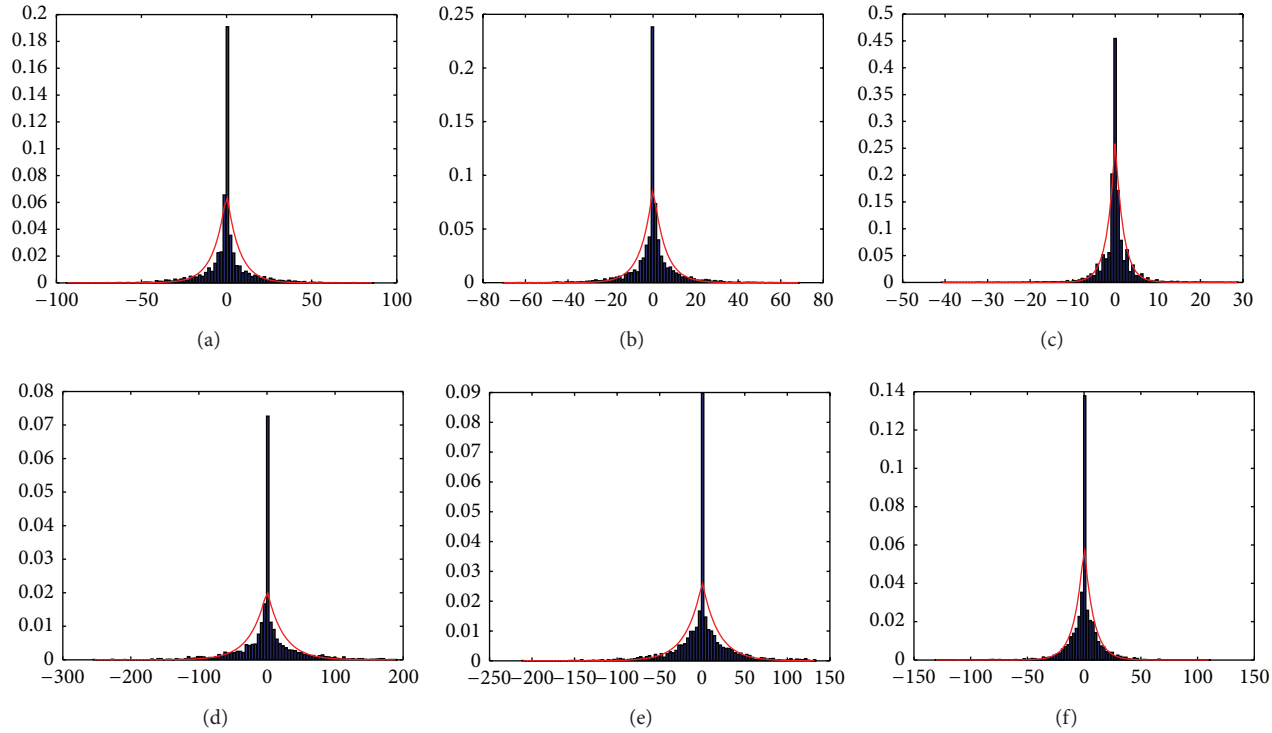


Figure 2: Noiseless image/volume wavelet coefficients (blue bar chart) against approximated Laplacian distribution (red graph).

with parameter λ ($S_\lambda^\alpha[k]$ represents the posterior probability of being detail given $d^\alpha[k]$ with λ the prior probability of a coefficient being noise). This filter is based on the filter proposed by Gorgel et al. [26]. The novelty of our filter is the proposed distribution models for details and noise coefficients in the wavelet domain to MRI and the use of the EM method for estimating the parameters. This avoids the problem setting any free parameter such as the noise variance which is usually problematic.

Noiseless MR images/volumes have a distribution in the wavelet domain with a pronounced maximum at the origin (due to smooth regions) and long tails produced by edges and different structures contained in the image/volume (as shown in Figure 2), so this distribution can be approximately modeled by a Laplacian function, given by

$$P\left(d^\alpha[k] \mid \text{detail}\right) := \frac{1}{2b}\, e^{-(|d^\alpha[k]-\mu|/b)}, \tag{14}$$

with parameters μ and b. This distribution was first proposed to model wavelet coefficients distribution of mammographic images by Gorgel et al. [26].

MR magnitude image/volume distribution is Rician [5, 37]. In high signal to noise ratio (high intensity, bright) regions, Rician distribution tends to a Gaussian distribution and in low signal to noise ratio (low intensity, dark) regions, Rician distribution tends to a Rayleigh distribution [4]; that is, MR noise (signal free) can be modeled by a Rayleigh distribution. In the wavelet domain, the distribution of this noise has a maximum at the origin with short tails (see Figure 3). In this case, we approximate this distribution by a Gaussian distribution, given by

$$P\left(d^\alpha[k] \mid \text{noise}\right) := \frac{1}{\sigma_{\text{noise}}\sqrt{2\pi}}\, e^{-((d^\alpha[k])^2/2\sigma_{\text{noise}}^2)}. \tag{15}$$

The definition of S_λ^α as a conditioned probability lets us modify the wavelet coefficients taking into account noise intensity. Besides, the larger image/volume modification by wavelet coefficients shrinkage, the bigger module of the shrunk coefficients; that is, a change in low module wavelet coefficient does not change significantly the image/volume. So, the importance of S_λ^α is focused on the tails of the distributions. Figure 4 shows the mixture model of detail and noise, with the two approximated distribution models presented before, superimposed on real image/volume wavelet histograms at different levels and orientations.

Parameters estimation in the original paper [26] consider both distributions independently choosing the parameter λ depending on the ratio noise/detail. In this paper we consider the joint distribution resulting from considering both distributions allowing us to find the parameters that maximize the similarity between the real and the theoretical distributions. For this purpose, parameters μ, b, σ_{noise}, and λ are to be calculated by the EM method as Appendix A shows in detail. Here a brief summary is given as follows; given the (independent) known data $\mathbf{X} = \{X_i\}_{i=1}^N$ (vector $\mathbf{X}$ represents each set of wavelet coefficients d^α for each scale and orientation α of an image/volume. We also drop the upper index α for the sake of simplicity along this

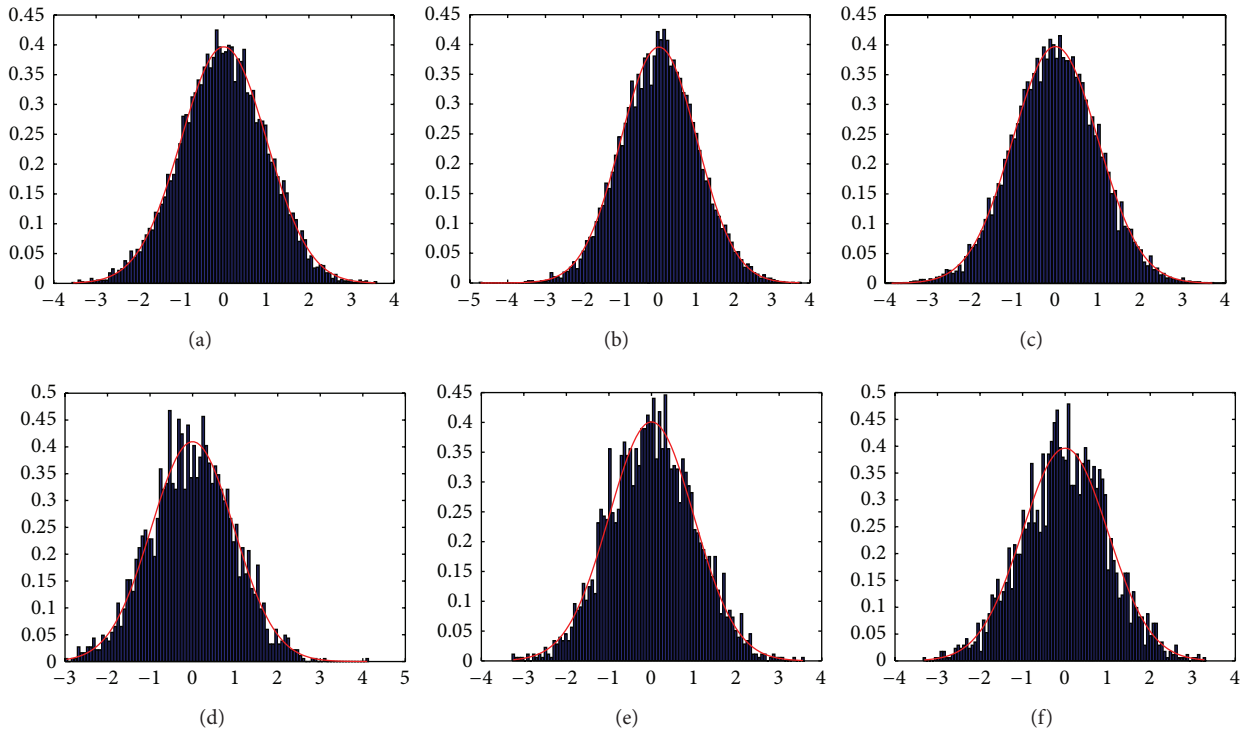


Figure 3: Rayleigh noise image/volume wavelet coefficients (blue bar chart) against approximated Gaussian distribution (red graph).

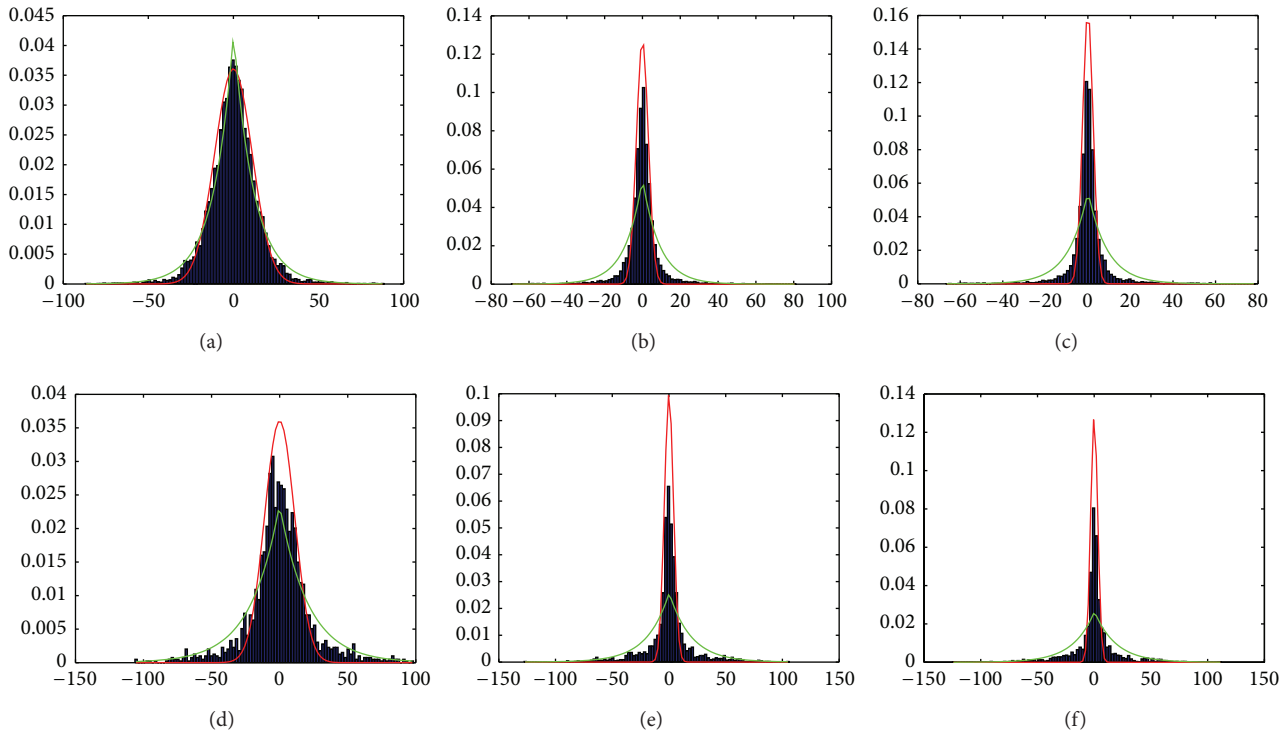


Figure 4: Mixture approximated model distributions (Detail/Laplace = green graph, Noise/Gauss = red graph) superimposed on real image wavelet coefficients histogram (blue bar chart) at level/orientation | NEX: (a) 1/horizontal | 1; (b) 1/horizontal | 10; (c) 1/horizontal | 20; (d) 2/vertical | 1; (e) 2/vertical | 10; (f) 2/vertical | 20.

section.) and the hidden (auxiliary) variables $\mathbf{Z} = \{Z_i\}_{i=1}^N$, defined by

$$Z_i := \begin{cases} 1, & \text{if } X_i \text{ is detail,} \\ 0, & \text{if } X_i \text{ is noise,} \end{cases} \quad i = 1, \ldots, N, \tag{16}$$

the likelihood function to be maximized in this method is

$$L(\boldsymbol{\Theta} \mid \mathbf{X}, \mathbf{Z}) := \sum_{i=1}^{N} \ln\left(P(X_i \mid Z_i, \boldsymbol{\Theta})\right) + \sum_{i=1}^{N} \ln\left(P(Z_i \mid \boldsymbol{\Theta})\right), \tag{17}$$

where $\boldsymbol{\Theta} = [\mu, b, \sigma_{\text{noise}}, \lambda]$ and $P(X_i \mid \boldsymbol{\Theta}) = \lambda P_{\text{Gauss}}(X_i \mid \sigma_{\text{noise}}) + (1-\lambda) P_{\text{Laplace}}(X_i \mid \mu, b)$, and its expected value is

$$\begin{aligned} &E\left(L(\boldsymbol{\Theta} \mid \mathbf{X}, \mathbf{Z})|_{\boldsymbol{\Theta}, \mathbf{X}}\right) \\ &= -\sum_{i=1}^{N} \left[\gamma_i \left(\ln(2) + \ln(b) + \frac{|X_i - \mu|}{b} - \ln(1-\lambda) \right) \right. \\ &\quad + (1-\gamma_i) \left(\frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln\left(\sigma_{\text{noise}}^2\right) \right. \\ &\quad \left. \left. + \frac{X_i}{2\sigma_{\text{noise}}^2} - \ln(\lambda) \right) \right], \end{aligned} \tag{18}$$

with

$$\gamma_i = \frac{(1-\lambda) P_{\text{Laplace}}(X_i \mid \mu, b)}{(1-\lambda) P_{\text{Laplace}}(X_i \mid \mu, b) + \lambda P_{\text{Gauss}}(X_i \mid \sigma_{\text{noise}})}. \tag{19}$$

Maximizing this expression, we obtain

$$\begin{aligned} \widehat{\lambda} &:= 1 - \frac{1}{N}\sum_{i=1}^{N} \gamma_i; \\ \widehat{\sigma}_{\text{noise}}^2 &:= \frac{\sum_{i=1}^{N} (1-\gamma_i) X_i^2}{\sum_{i=1}^{N} (1-\gamma_i)}; \\ \widehat{\mu} &:= \arg\min_{\mu = X_m} \sum_{i=1}^{N} \gamma_i |\mu - X_i|; \\ \widehat{b} &:= \frac{\sum_{i=1}^{N} \gamma_i |X_i - \widehat{\mu}|}{\sum_{i=1}^{N} \gamma_i}. \end{aligned} \tag{20}$$

It is an implicit system, so estimators γ_i, $\widehat{\lambda}$, $\widehat{\sigma}_{\text{noise}}^2$, $\widehat{\mu}$, and $\widehat{b}$ are calculated by fixed-point iteration with initial conditions

$$\begin{aligned} \widehat{\lambda}_{\text{Ini}} &:= \frac{1}{2}, \\ \widehat{\mu}_{\text{Ini}} &:= \text{median}(\mathbf{X}), \\ \widehat{b}_{\text{Ini}} &:= \frac{1}{N}\sum_{i=1}^{N} |X_i - \widehat{\mu}_{Ini}|, \\ \left(\widehat{\sigma}_{\text{noise}}^2\right)_{\text{Ini}} &:= \text{median}\left(\sigma_{\text{local}}^2\right), \end{aligned} \tag{21}$$

where $\sigma_{\text{local}}^2 := \text{LV}(\widetilde{\mathbf{X}})$ is defined in (B.2) (where $\widetilde{\mathbf{X}}$ is the vector $\mathbf{X}$ as a matrix [of wavelet coefficients d^α at scale and orientation α]), and

$$\begin{aligned} (\gamma_i)_{\text{Ini}} &:= \left(\left(1 - \widehat{\lambda}_{\text{Ini}}\right) P_{\text{Laplace}}\left(X_i \mid \widehat{\mu}_{\text{Ini}}, \widehat{b}_{\text{Ini}}\right) \right) \\ &\quad \times \left(\left(1 - \widehat{\lambda}_{\text{Ini}}\right) P_{\text{Laplace}}\left(X_i \mid \widehat{\mu}_{\text{Ini}}, \widehat{b}_{\text{Ini}}\right) \right. \\ &\quad \left. + \widehat{\lambda}_{\text{Ini}} P_{\text{Gauss}}\left(X_i \mid \left(\widehat{\sigma}_{\text{noise}}\right)_{\text{Ini}}\right) \right)^{-1}. \end{aligned} \tag{22}$$

4. Experiments

In this section, some noisy MR 2D and 3D volumes are filtered to compare the algorithm proposed in this paper with other wavelet filters as well as patch-based and NLM filters found in the reviewed literature and described in Section 2. Section 4.1 contains the description, step by step, of the different filter algorithms applied on the images defined in Section 4.2. The three wavelet methods, proposed by Donoho and Johnstone, Nowak, and Villullas and Martin, apply the approximately shift-invariant scheme, with $K = 2$ in the 2D images and $K = 1$ in the 3D volumes, as proposed in Section 2.1.3, and Donoho-Johnstone's and Nowak's filters use the noise estimator defined in Section 2.3. Patch-based filter, proposed by Awate-Whitaker, uses Gaussian and Rician models (this method can be obtained at the web http://www.itk.org/Doxygen42/html/group_ITKDenoising.html with the predefined parameters values except m_NoiseModel = Gaussian or Rician and m_NoiseModelFidelityWeight = 0.5.), and nonlocal means filter uses predeterminated parameters (this method can be obtained at the web http://www.mathworks.com/matlabcentral/fileexchange/40162-james-stein-type-center-pixel-weights-for-non-local-means) with the same noise estimator as Nowak's and Donoho-Johnstone's filters. The three wavelet filters and NLM filter have been programming in MATLAB and Awate-Whitaker's filter has been programmed in C.

4.1. Filter Algorithms. The different filtering methods used in experiments shown in Section 4.4 are next described step by step for a given image/volume I. The wavelet used in these experiments is the Haar wavelet. As the Haar wavelet support is minimal and a wavelet coefficient image/volume is a windowed weighted average of an image/volume, this wavelet seems the most logical choice when we want to remove noise and preserve details in an image/volume filtering, against other wavelet families with larger supports.

4.1.1. Donoho-Johnstone's Algorithm

(i) Calculate noise variance estimator $\widehat{\sigma}_{\text{noise}}^2$ as Section 2.3 shows.

(ii) Compute the ($J = 2$)-scale DWT of the magnitude image/volume I.

(iii) Filter the wavelet coefficients d^α through the Donoho-Johnstone's filter, F_{DJ}, defined in Section 2.1.

(iv) Compute the inverse DWT of the filtered wavelet and scaling coefficients to obtain an estimate of denoised image/volume.

(v) Repeat Steps (ii)–(iv), with the image/volume shifted $\{-K, \ldots, -1, 0, 1, \ldots, K\}$ pixels/voxels in each direction, as Section 2.1.3 explains, and average all shifted denoised images/volumes (original image/volume I corresponds to null shift).

4.1.2. Nowak's Algorithm

(i) Calculate noise variance estimator $\widehat{\sigma}^2_{\text{noise}}$ as Section 2.3 shows.

(ii) Compute the $(J = 2)$-scale DWT of the squared magnitude image/volume I^2.

(iii) Remove the bias from the scaling coefficients s^J by subtracting $C = 2^{J+1}\sigma_{\text{noise}}$ from each (see [20]).

(iv) Filter the wavelet coefficients d^α through the Nowak's filter, F_{N}, defined in Section 2.1.

(v) Compute the inverse DWT of the filtered wavelet and unbiased scaling coefficients to obtain an estimate of squared denoised image/volume.

(vi) Take the pixel-by-pixel/voxel-by-voxel square-root of the result to obtain an estimate of the unbiased image/volume.

(vii) Repeat Steps (ii)–(vi), with the image/volume shifted $\{-K, \ldots, -1, 0, 1, \ldots, K\}$ pixels/voxels in each direction, as Section 2.1.3 explains, and average all shifted denoised images/volumes (original image/volume I corresponds to null shift).

4.1.3. Awate-Whitaker's Algorithm

(i) Let $I_0 = I$ the noisy image/volume and $k > 0$.

(ii) For each region $\widehat{z}^k = [\widehat{x}^k, \widehat{y}^k]$ (i.e., for each pixel m) of the image/volume I_k, compute $\partial h(\widehat{X} \mid \widehat{Y} = \widehat{y}^k)/\partial \widehat{x}^k$.

(iii) Use the descending gradient differences to calculate $\widehat{x}^{k+1}$ by $\widehat{x}^{k+1} = \widehat{x}^k - \delta(\partial h/\partial \widehat{x}^k)$.

(iv) The image/volume I_{k+1} is given by the intensity $\widehat{x}^{k+1}$ at pixel t. If $k + 1$ is less than the maximum number of iterations, go to Step (ii) with $k = k + 1$. Otherwise, the filtered image/volume is I_{k+1}.

4.1.4. Nonlocal Means Algorithm

(i) Given the parameters values, calculate the weights function w for each pixel of the image/volume I.

(ii) For each pixel, compute the NLM filtered image/volume with the weight function.

4.1.5. Villullas-Martin's Algorithm

(i) Compute the $(J = 2)$-scale DWT of the magnitude image/volume I.

(ii) Filter the wavelet coefficients d^α through the Villullas-Martin's filter, F_{VM}, defined in Section 3.

(iii) Compute the inverse DWT of the filtered wavelet and scaling coefficients to obtain an estimate of the denoised image/volume.

(iv) Repeat Steps (i)–(iii), with the image/volume shifted $\{-K, \ldots, -1, 0, 1, \ldots, K\}$ pixels/voxels in each direction, as Section 2.1.3 explains, and average all shifted denoised images/volumes (original image/volume I corresponds to null shift).

4.2. MR Data Sets. The experiments were conducted on fifth MRI data sets. The first and second data sets consist of simulated MR volumes and images obtained from the Brainweb database [38]. The third and fourth data sets were collected from Centro de Diagnóstico Valladolid (CDV) QDIAGNOSTICA in Valladolid (Spain). The last data set was obtained from a database of the Laboratory of Image Processing (LPI) of the University of Valladolid. Details about these data sets are described in following subsections.

4.2.1. Simulated MR Images. Simulated MR images/volumes are a useful data set which allows a first evaluation of different analysis methods. In the 3D case, the noiseless image volume consists of a 3-dimensional volume of resolution $180 \times 216 \times 4$ extracted from a volume generated in the Brainweb database of resolution $181 \times 217 \times 181$, T1-weighted, 1 mm slice thickness, 0% of noise, and RF = 0%. Image intensity values vary in $\{0, 1, \ldots, 255\}$. In the 2D case, the noiseless image consists of a 2-dimensional axial section of the 3-dimensional volume of resolution 180×216. Rician noise was generated by the equation

$$\widehat{I} := \sqrt{(I + N_1)^2 + N_2^2}, \tag{23}$$

where I is the noiseless MR image/volume and $N_k/k \in \{1, 2\}$ are $N(0, \sigma^2)$ independent identically distributed Gaussian random variables, with $\sigma \in \{5, 6, \ldots, 20\}$.

4.2.2. Real MR Images. Real data set consists of two data sets. Real data images are 2D axial section of a brain, acquired using a General Electric Signa 1.5 T scanner, T1-weighted, and $0.9375 \times 0.9375\,\text{mm}^2$ pixel size for both data sets. The first data set is composed of 4 images of dimension 256×256, varying NEX $\in \{4, 8, 16, 64\}$, TR = 6 ms, TE = 1.588 ms, and flip angle = 15°. The second data set consists of 32 images with dimension 192×160, NEX = 1, TR = 40 ms, TE = 9 ms, and flip angle = 90°. As noiseless image cannot be obtained for real data, noiseless image approximations are given by the following: in the first data set, the noiseless image approximation corresponds to NEX = 64; in the second data set, the noiseless image approximation is the average in the complex domain of the 32 images that this set consists of.

4.2.3. Clinical MR Images. Real clinical data set consists of 20 MR T1-weighted acquisitions from different subjects. The dataset was acquired in a General Electric Signa 1.5 T scanner

with NEX = 1, 2.1875 × 2.1875 mm^2 pixel size, and 256 × 256 image dimension. Other parameters values are TR = 5.8020 ms, TE = 1.7280 ms, and flip angle = 10°. This set of images try to evaluate the different filter methods with images taken from different subjects (anatomy variations). In this case, as no ground truth is available, the evaluation will be performed based on experts' rankings.

4.3. Quality Measures. After filtering these images/volumes (I is the noiseless image/volume, I_N is the noisy image/volume, and I_F is the filtered noisy image/volume), the results obtained are compared using the following efficiency measurements (see Appendix B for details).

4.3.1. Averaged Error Local Variance (AELV). AELV is an objective quality measure [39] that quantifies the deviation of estimated values from the true value. Specifically, the AELV of I_F with respect to I is measured as

$$\mathrm{AELV}\left(I_F, I\right) := \frac{1}{M}\sum_m \mathrm{LV}\left(I - I_F, m\right), \tag{24}$$

where M is the total number of pixels/voxels of I and the operator LV is defined in Appendix B.

4.3.2. Normalized Averaged Error Local Variance (NAELV). Variations in AELV can be difficult to understand. To test at what rate reduces its value after filtering, AELV normalization is used as

$$\mathrm{NAELV}\left(I_F, I_N, I\right) := \frac{\mathrm{AELV}\left(I_F, I\right)}{\mathrm{AELV}\left(I_N, I\right)}. \tag{25}$$

4.3.3. Structural Similarity (SSIM). Although AELV is a useful measure of similarity, it is not suitable to obtain a comparison similar to that performed by the human eye [40, 41]. Most common alternative is SSIM which is consistent with the visual perception. The SSIM index is estimated as

$$\begin{aligned}\mathrm{SSIM}\left(I_F, I\right) := {} & \frac{2\mathrm{GM}\left(I\right)\mathrm{GM}\left(I_F\right) + C_1}{\mathrm{GM}^2\left(I\right) + \mathrm{GM}^2\left(I_F\right) + C_1} \\ & \cdot \frac{2\mathrm{GCV}\left(I, I_F\right) + C_2}{\mathrm{GV}\left(I\right) + \mathrm{GV}\left(I_F\right) + C_2},\end{aligned} \tag{26}$$

where the operators GM(I), GV(I), and GCV(I, I_F) are defined in Appendix B, $C_1 := 6.5025$ and $C_2 := 58.5225$.

4.3.4. Averaged Local Signal to Noise Ratio (ALSNR). Another simple method of checking the noise level in an image is the averaged local SNR, measured as

$$\mathrm{ALSNR}\left(I_F, I\right) := \frac{1}{M}\sum_m \frac{\mathrm{LV}\left(I, m\right)}{\mathrm{LV}\left(I - I_F, m\right)}, \tag{27}$$

where M is the total number of pixels/voxels of I and the operator LV is defined in Appendix B.

4.3.5. Contrast. The measure normally used to calculate images contrast is given by

$$\mathrm{Cont}\left(I\right) := \frac{S_{\max} - S_{\min}}{S_{\max} + S_{\min}}, \tag{28}$$

where $S_{\max}$ and $S_{\min}$ are the maximum and minimum value in a specific ROI of the image/volume I. This quality measurement is not effective enough because some filters (e.g., some wavelet filters) generate a bias that modifies the values of this quality measurement but does not affect the contrast of the image/volume. Besides, there are many tools, used by radiologists, which let them modify the window/level of the image/volume. We can observe the variation of contrast taking a section of the corresponding images/volumes to compare the different changes of intensity. Figure 5 shows that wavelet filters preserve edge changes of intensity very well due to its property of locality.

4.4. Numerical Experiments. In these experiments, the size of a neighborhood at pixel/voxel m (on the image/volume I), size (N_m), used is size (N_m) = 5 × 5 ($|N_m|$ = 25) in 2-dimensional case and size (N_m) = 5 × 5 × 5 ($|N_m|$ = 125) in 3-dimensional case in (B.2) and (B.1) in Appendix B. Others values give us different measurement values but the same visual distribution.

4.4.1. Experiment 1: Filtering Simulated Data. In the first experiment, the 16 2D simulated data sets described in Section 4.2.1 are filtered. The parameter values used to generate noisy images are $\sigma \in \{5, 6, \ldots, 20\}$. Figure 6 shows the comparison of the different averaged quality measurements as function of the parameter σ (each parameter σ has associated 20 simulated images experiments and the corresponding quality measurement value is the average of the 20 quality measurement values of the corresponding images. We use this 20 data sets to reduce the variability of the quality measurements). For low values of the parameter σ (almost noiseless images), Awate-Whitaker's filter (Rician over Gaussian in SSIM measurement and Gaussian over Rician in AELV measurement) have a good noise removal and it beats the other methods in this case. Nowak's filter gets results close to Villullas-Martin's filter (in fact, Nowak's filter improves Villullas-Martin filter in SSIM for tiny values of σ) but as parameter value increases, difference between Villullas-Martin's filter and Nowak's filter grows. Villullas-Martin's filter is clearly the best method for medium and high values of σ (hard noisy images. This value corresponds to low NEX in image acquisition). NAELV graph shows us that Awate-Whitaker's filter has constant improvement proportion in contrast to wavelet filters, which amplify its improvement proportion with σ. Then, Awate-Whitaker's filter is a good denoising method for very low values of σ but as σ increases, the filter's strength decreases. The NLM filter has a behavior near to Nowak's filter, better than Nowak's filter in high values of σ and worst in low and medium values of σ. Figure 7 shows an example (σ = 15) where visual differences between the filtering methods can be seen. In this case, Villullas-Martin's filtering and Nowak's filtering are similar

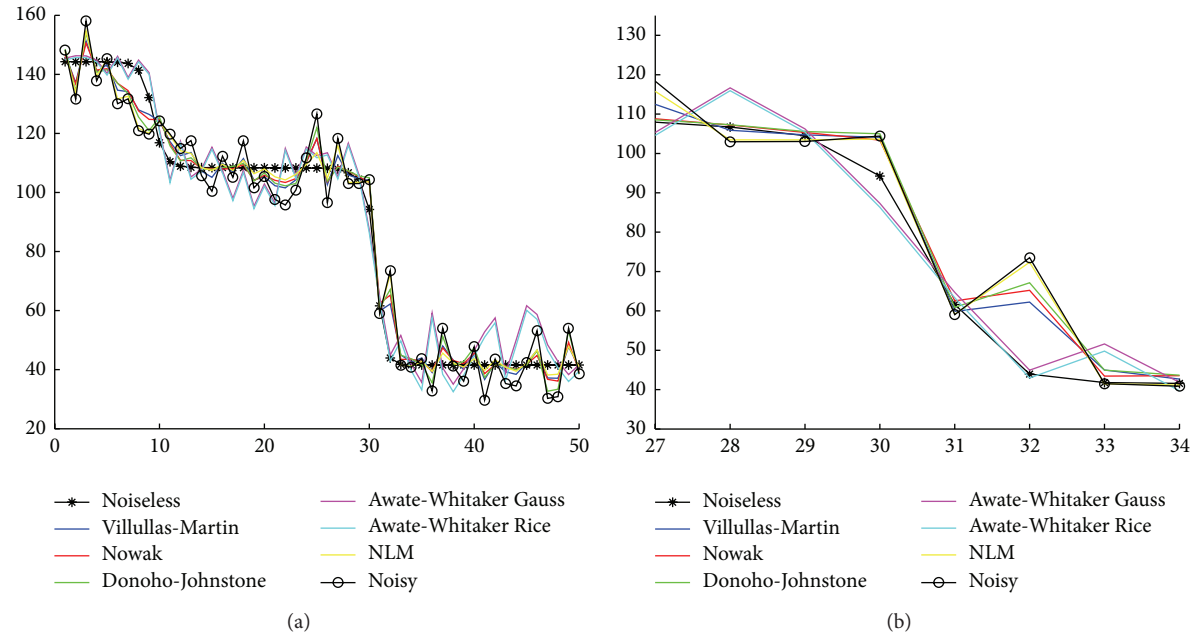


FIGURE 5: Comparative profiles of 1D section of noiseless, noisy, and the different filtered 2D images. (a) 1D section of Section 4.2.1 data set image with parameter $\sigma = 10$ from pixel $(50, 126)$ to pixel $(99, 126)$; (b) detail of profile (a) in coordinates $[27, 28, \ldots, 34]$.

to each other and better than Donoho-Johnstone's filtering which has lower noise removal. NLM filter is the worst choice because the noise has not been properly eliminated inside brain structure.

4.4.2. Experiment 2: Filtering Real Data. In the second experiment, we have filtered the first real data set described in Section 4.2.2. Noisy images set is the images belonging to NEX $\in \{4, 8, 16\}$ and noiseless image is generated with NEX = 64. Figure 8 shows the comparisons of the different quality measurements as function of NEX. It can be seen that Villullas-Martin's filter improves Nowak's and Donoho-Johnstone's filters, with a slight increase as NEX increase. Awate-Whitaker's filters cannot improve wavelet filters results (moreover, as AELV graph shows, Awate-Whitaker's filter ruins low NEX images). As Experiment 1, NLM filter works as good as Nowak's filter. Figure 9 shows the example of NEX = 8. In this example we can see that Villullas-Martin's filter provide a better filtering, where more noise is removed than Nowak's, Donoho-Johnstone's, and NLM filters. Awate-Whitaker's filter with Gaussian model obtains a bit worse denoising than Nowak's, Donoho-Johnstone's, and NLM filters and Awate-Whitaker's filter with Rice model does not remove enough noise inside brain structures and generates artifacts.

4.4.3. Experiment 3: 3D Wavelet Filtering. This experiment shows the strength of the wavelet filters in the 3-dimensional case. In this case, the 16 3D simulated data volumes described in Section 4.2.1 for $\sigma \in \{5, 6, \ldots, 20\}$ are filtered. Figure 10 shows the comparison of the different averaged (20 volumes for each σ level whose quality measurements are averaged for each σ) quality measurements as function of σ. This experiment shows the same behavior as Experiment 1 in wavelet filters. All of these methods improve the filtering in 3D case preserving the relationship between them.

TABLE 1: Measurement results for NEX = 2, 3, 4 and Villullas-Martin's method for NEX = 2.

Image\measurement	AELV	SSIM	ALSNR
NEX = 2	60.5142	0.6355	2.1124
NEX = 3	39.2910	0.7162	3.2623
NEX = 4	28.6274	0.7698	4.5433
NEX = 2 filtered	30.5097	0.7331	3.4963

4.4.4. Experiment 4: Filtered Real Image against Higher NEX. The forth experiment evaluates the filtering power of Villullas-Martin's filter. This filter is compared with NEX rising. For this comparison, the second real data set described in Section 4.2.2 is used. The "same" 32 images with NEX = 1 let us control noise level, which can be reduced by averaging N images in the complex domain, where higher values of N imply lower noise level. The selected image to be filtered belongs to $N = 2$. Table 1 shows that the first value of averaged images (N) which improves the quality measurements of filtered image with $N = 2$ is $N = 4$; that is, Villullas-Martin's filter lets us obtain images with quality as good as with double acquisition time. Figure 11 shows the images involved in this experiment. It can be seen that filtered image with $N = 2$ has less noise at smooth regions inside the structure than the image with no filtering with $N = 3$.

4.4.5. Experiment 5: Clinical Experts Comparison. In the last experiment, data set images of Section 4.2.3 are filtered by

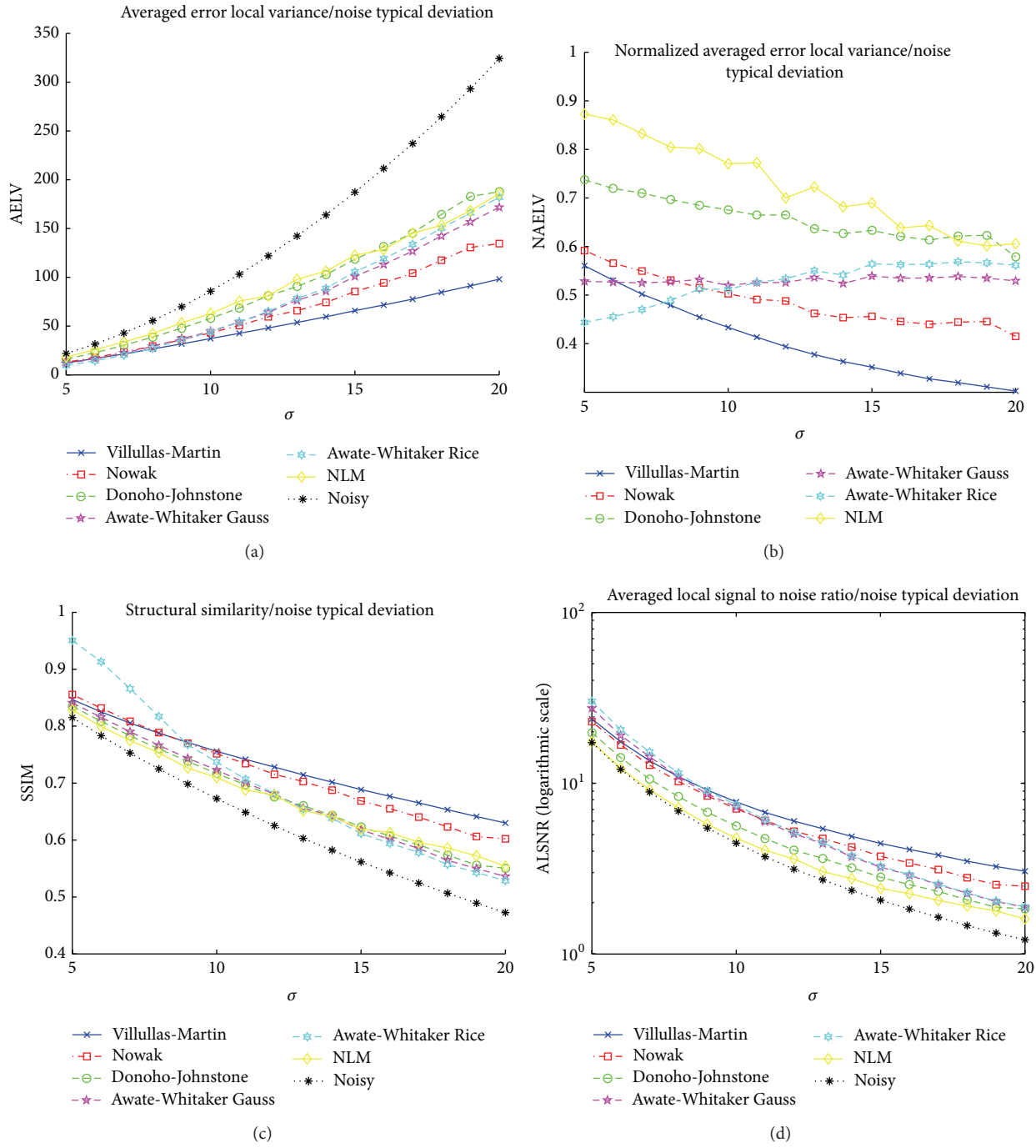


Figure 6: Comparative results for the different filtering methods in experiment 1 with parameter $\sigma \in \{5, 6, \ldots, 20\}$.

the wavelet filter methods described in Section 2.1. The goal of this experiment is to evaluate the clinical relevance of the proposed filter over other wavelet filters with the help of 3 radiology experts at high SNR images (where the three wavelet filters are more similar). The images are filtered with Villullas-Martin's, Nowak's, and Donoho-Johnstone's filter methods, which have proven quite efficient in previous experiments. The experts rank the filtered images by their visual quality based on their experience (1 = best, 2 = medium, and 3 = worst). The criteria to evaluate this quality are based on the ease of recognizing different structures in the image (each expert uses different regions that he judges to be key). The ranking is blind in the sense that the filtering method used to filter each image is hidden to the expert. To evaluate the intraexpert variability the experiment was repeated during two different days. Tables 2 and 3 show results achieved for the intraexpert and interexpert variability, respectively. Although the Villullas-Martin's filter does not remove as

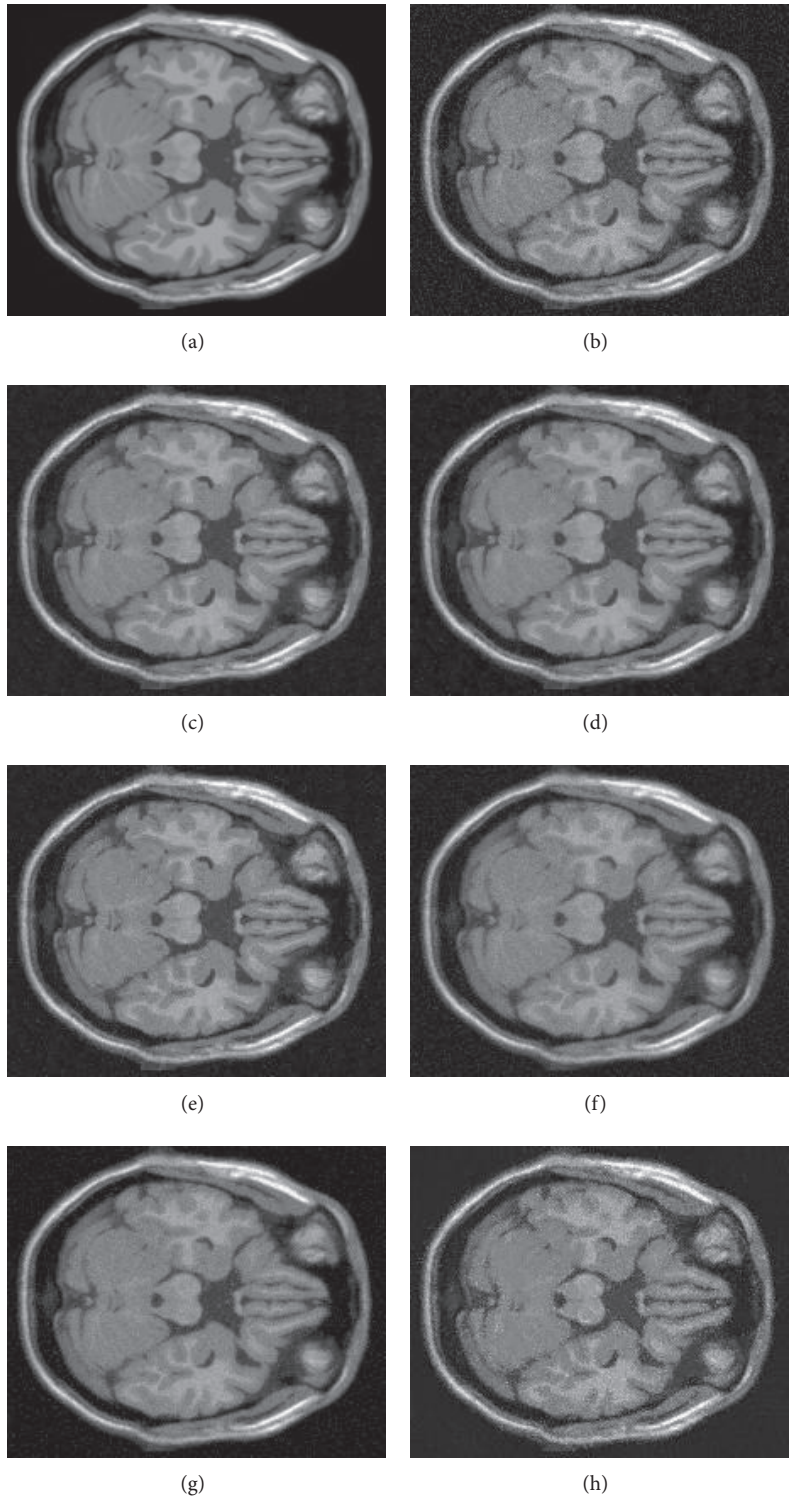


FIGURE 7: Example of experiment 1 with parameter $\sigma = 15$. (a) Noiseless image; (b) noisy image; (c) noisy image filtered by Villullas-Martin's method; (d) noisy image filtered by Nowak's method; (e) noisy image filtered by Donoho-Johnstone's method; (f) noisy image filtered by Awate-Whitaker's method with Gaussian model; (g) noisy image filtered by Awate-Whitaker's method with Rician model; (h) noisy image filtered by nonlocal means method.

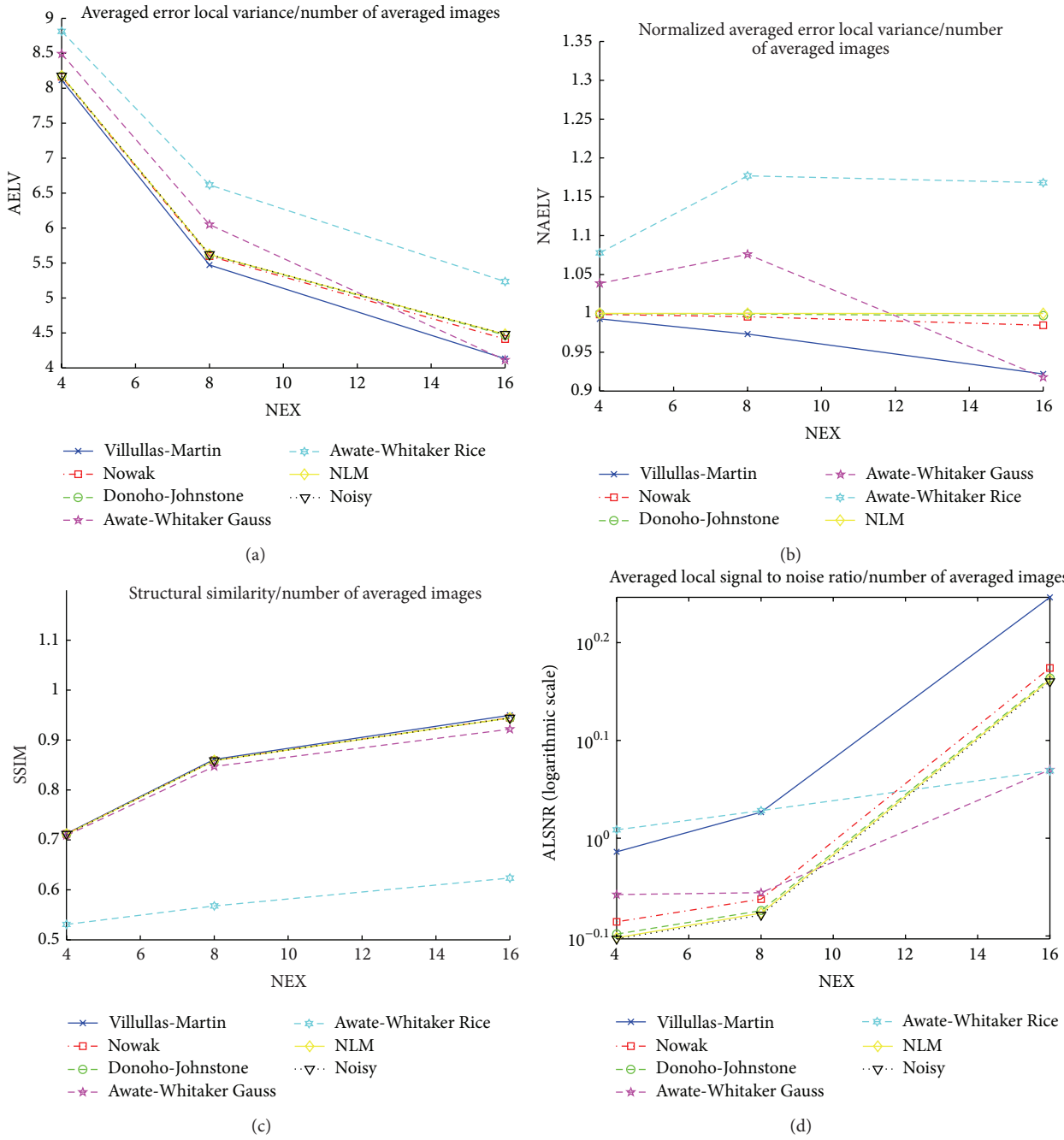


Figure 8: Comparative results for the different filtering methods in experiment 2 with number of averaged images NEX $\in \{4, 8, 16\}$.

much noise as other filters (apparently), which obtains similar filtering results in these cases, the result image given by our method was considered by the experts visually nicer and with higher contrast in the boundaries of the image, which eases radiologist work in identifying the different brain structures; that is, Villullas-Martin's filter preserves structures better than the other wavelet filters after denoising. The experiment results can be summarized as follows. In 84.2% of cases, Villullas-Martin's filter was chosen as the best filter method (rank = 1), whereas Nowak's filter was chosen as the worst filter in 88.3% of cases (rank = 3). Donoho-Johnstone's filter was chosen as the medium filter (rank = 2) in 75% of cases. Table 4 shows all percentages. Figure 12 shows some example images in this experiment.

5. Conclusion

A new wavelet-domain filtering has been proposed in this paper. Assuming that wavelet coefficients of a noiseless MR image/volume can be modeled by a Laplacian distribution (as we know in brain, the filter can be adapted to other human parts easily with minor or no changes), and the wavelet

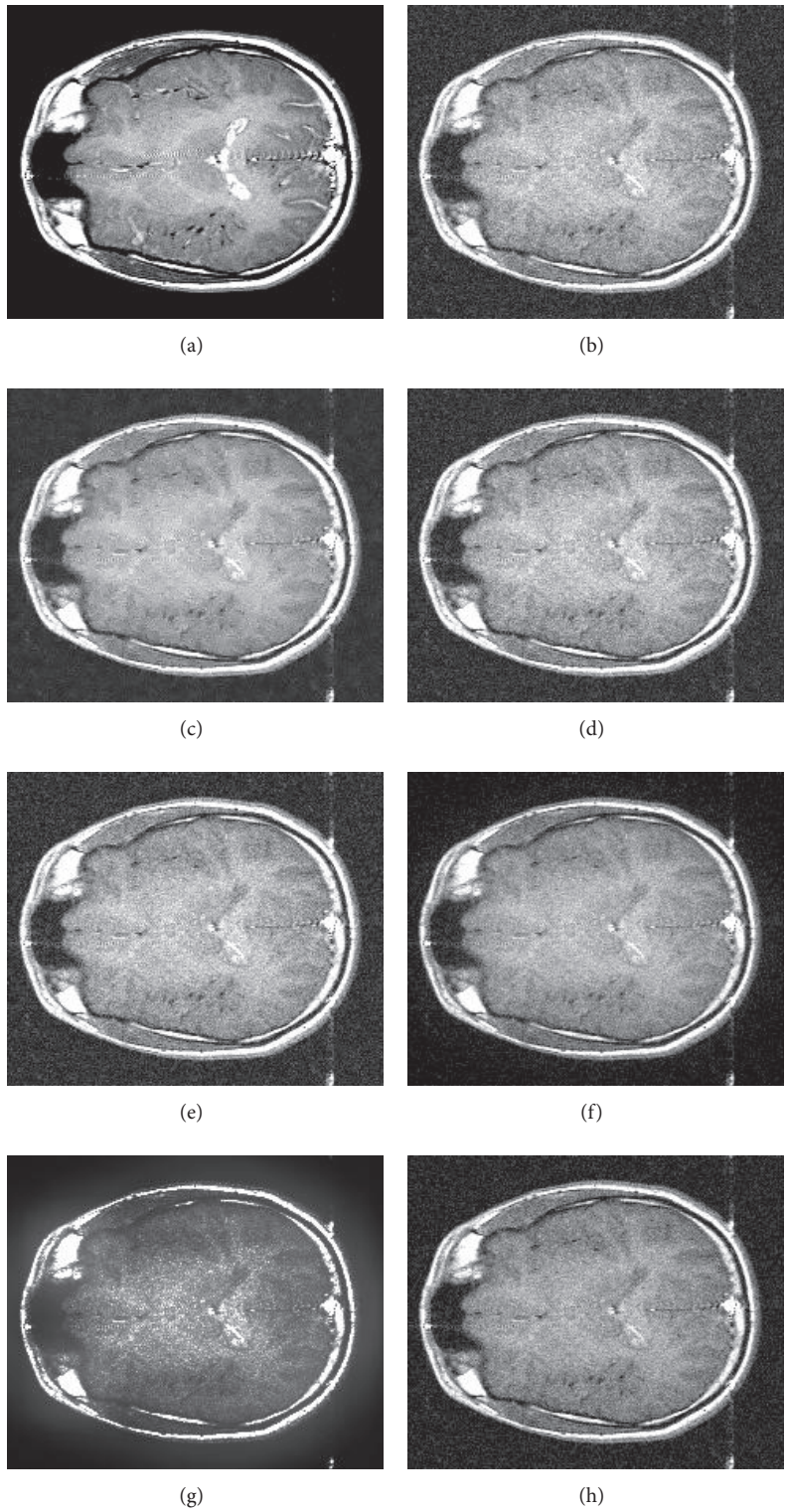


Figure 9: Example of experiment 2 with number of averaged images NEX = 8. (a) Noiseless image; (b) noisy image; (c) noisy image filtered by Villullas-Martin's method; (d) noisy image filtered by Nowak's method; (e) noisy image filtered by Donoho-Johnstone's method; (f) noisy image filtered by Awate-Whitaker's method with Gaussian model; (g) noisy image filtered by Awate-Whitaker's method with Rician model; (h) noisy image filtered by nonlocal means method.

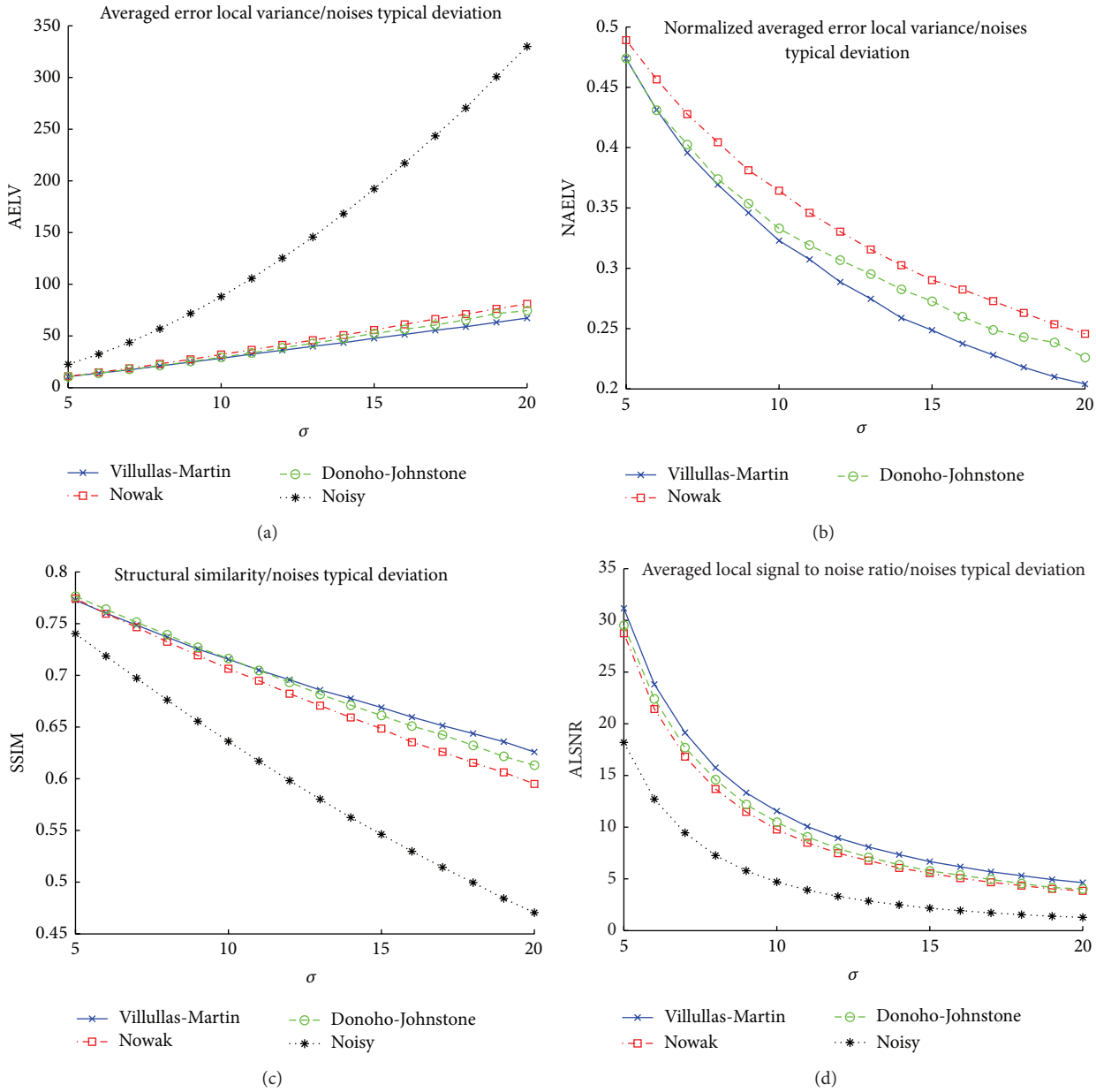


Figure 10: Comparative results for the different filtering methods in experiment 3 with parameter $\sigma \in \{5, 6, \ldots, 20\}$.

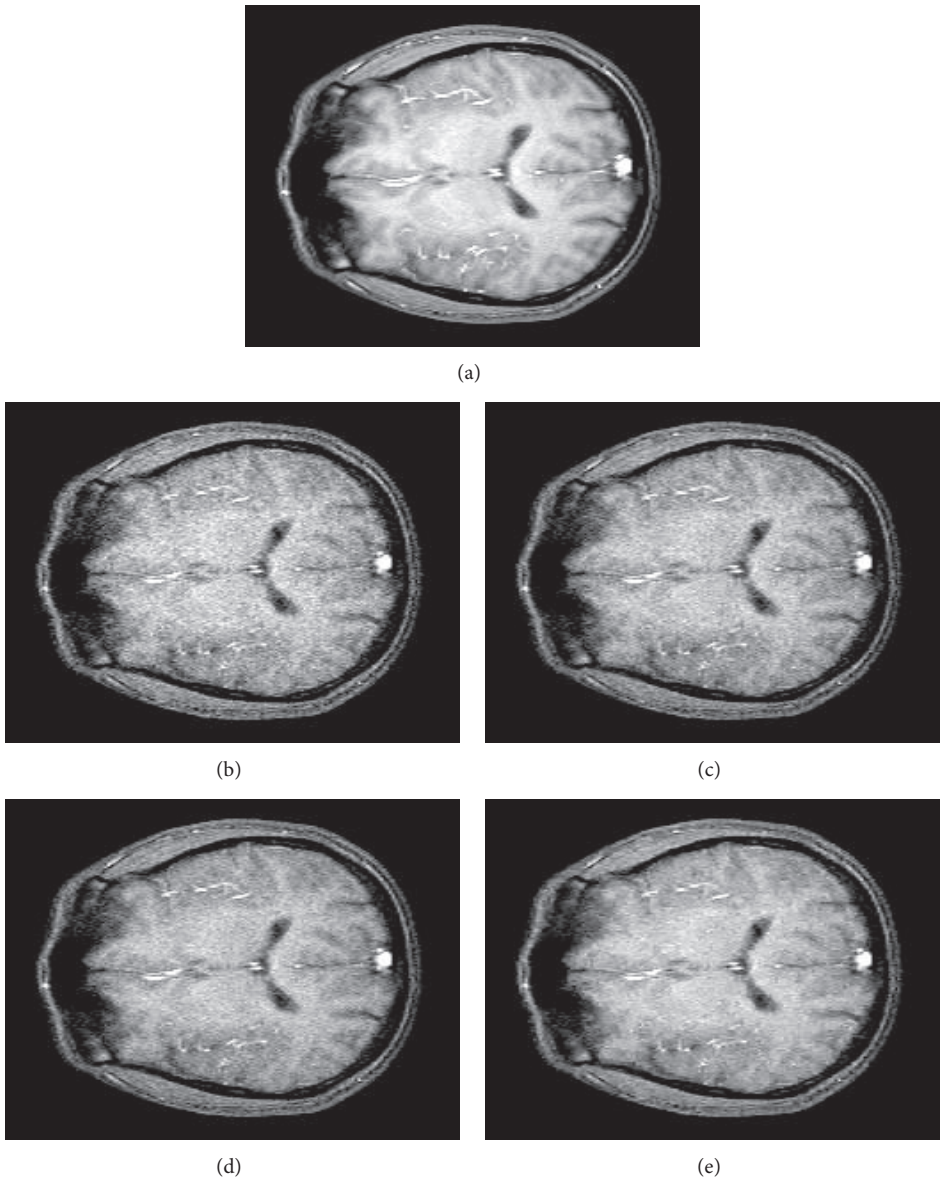
Table 2: Intraexpert evaluation percentages. Success implies that the expert chooses the same quality rank for the same subject for the two different instants of time; fail implies that the expert chooses different quality rank for the same subject.

	Expert 1	Expert 2	Expert 3
Success	95%	60%	70%
Fail	5%	40%	30%

Table 3: Interexpert evaluation percentages. Success implies that the experts choose the same quality rank for the same subject for the two different instants of time; fail implies that the experts choose different quality rank for the same subject. In these results only intraexpert coincidences are taken into account.

	Experts 1 and 2	Experts 1 and 3	Experts 2 and 3
Success	75%	85%	65%
Fail	25%	15%	35%

coefficients distribution of Rayleigh noise can be approximated by a Gaussian distribution, a probabilistic method has been proposed; that is, wavelet coefficients are shrunk depending on its conditioned probability of being noise or detail (posterior probability). To calculate the parameters involved in the expression of F_{VM}, EM method has been used. This fact makes a filter independent of σ_{noise} estimators,



(a) (b) (c) (d) (e)

Figure 11: Experiment 4. (a) Noiseless image; (b) noisy image for NEX = 2; (c) noisy image for NEX = 3; (d) noisy image for NEX = 4; (e) noisy image for NEX = 2 filtered by Villullas-Martin method.

Table 4: Total percentages in all of the experiment images used in experiment 5. The different wavelet filters were ranked by 3 radiology experts according to their visual quality in two different instants of time.

	Villullas-Martin	Nowak	Donoho-Johnstone
Best	84.2%	0.8%	15%
Middle	14.1%	10.9%	75%
Worst	1.7%	88.3%	10%

in contrast to other methods like Donoho-Johnstone's and Nowak's. Section 4 shows that Villullas-Martin's filter can provide better noise removal than Nowak's, Donoho-Johnstone's, Awate-Whitaker's, and nonlocal means filters for several signal to noise ratios and let us obtain high quality images/volumes with lower acquisition time, as much in 2D images as 3D volumes. In addition, a clinical evaluation performed by 3 radiology experts shows that our filter clearly outperforms the others. This experiment allows us to state that apparently to eliminate more noise not always means the visual quality expected by the experts is better, because this noise removal could imply low contour conservation and less details. Our filter result seems to the experts more similar to what they expect to see to better identify anatomical structures based on their experience. In addition, F_{VM} can be faster than F_N due to the expensive computation of $\sum_m (\psi^\alpha)^2[k](m)I^2(m)$ at $\widehat{(\sigma^\alpha)^2[k]}$ estimator for the case of a general wavelet (see Section 2.1 and [20]).

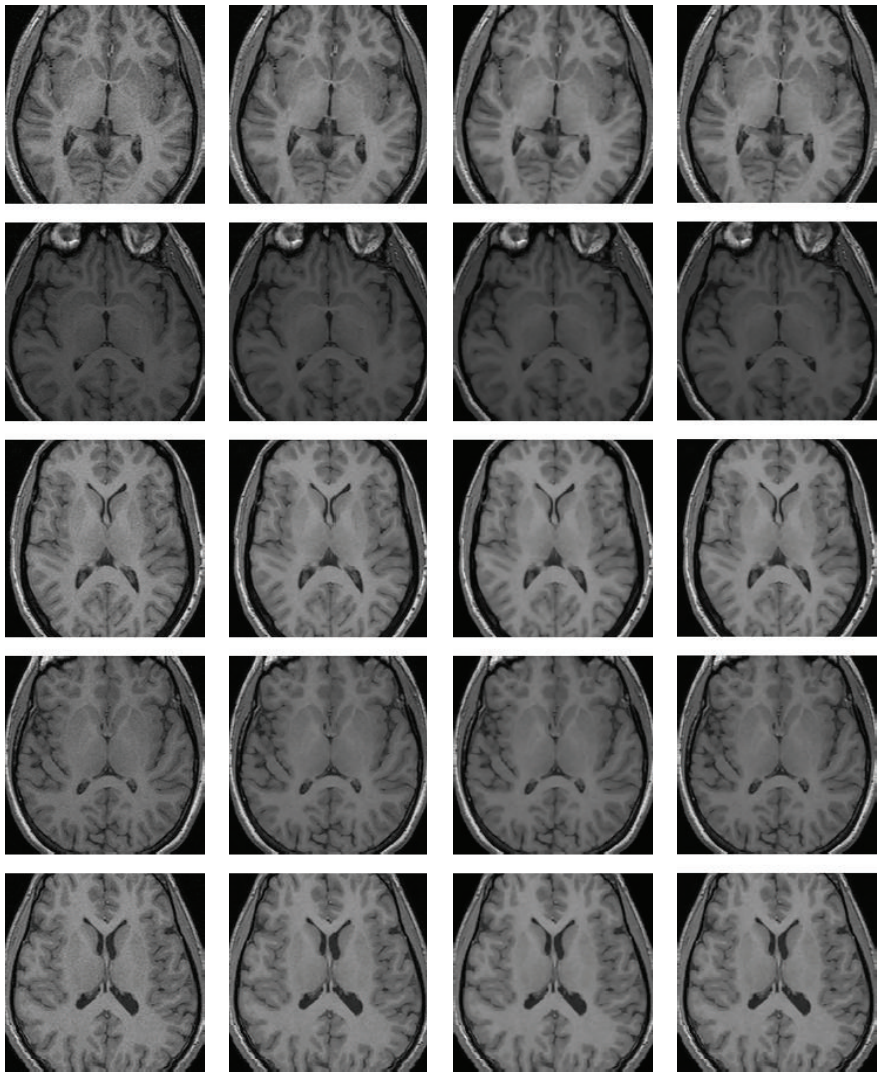


Figure 12: Examples subset of experiment 5. (a) Noisy images; (b) noisy images filtered by Villullas-Martin's method; (c) noisy images filtered by Nowak's method; (d) noisy images filtered by Donoho-Johnstone's method.

As future research lines, we plan to generalize the distribution models of the wavelet coefficients both for the details and for the noise. For the details a good candidate will be the generalized Gaussian distribution which seems very promising. In addition to that, we plan to check whether the proposed method can be also applied to other body parts as well as to other MR modern modalities such as fast, ultrafast, and low-field MRI for which the noise level is known to be very high.

Appendices

A. E.M. Parameter Estimation

We estimate the filter parameters by the EM method. This appendix shows the detailed derivation of the estimation equations.

Let $\boldsymbol{\Theta} = [\mu, b, \sigma_{\text{noise}}, \lambda]$ be the unknown parameter vector, $\mathbf{X} = \{X_i\}_{i=1}^{N}$ the known data, and $\mathbf{Z} = \{Z_i\}_{i=1}^{N}$ hidden auxiliary variables defined by

$$Z_i := \begin{cases} 1, & \text{if } X_i \text{ is detail,} \\ 0, & \text{if } X_i \text{ is noise,} \end{cases} \qquad i = 1, \ldots, N. \tag{A.1}$$

We assume that $\mathbf{X} = \{X_i\}_{i=1}^{N}$ are independent. So

$$\begin{aligned} P\left(X_i \mid \boldsymbol{\Theta}\right) &= \lambda P_{\text{Gauss}}\left(X_i \mid \sigma_{\text{noise}}\right) \\ &\quad + (1-\lambda)\, P_{\text{Laplace}}\left(X_i \mid \mu, b\right), \\ P\left(\mathbf{X} \mid \boldsymbol{\Theta}\right) &= \prod_{i=1}^{N} P\left(X_i \mid \boldsymbol{\Theta}\right). \end{aligned} \tag{A.2}$$

Including hidden variables,

$$P(\mathbf{X}, \mathbf{Z} \mid \mathbf{\Theta}) = \prod_{i=1}^{N} P(X_i, Z_i \mid \mathbf{\Theta}). \tag{A.3}$$

Moreover,

$$P(X_i, Z_i \mid \mathbf{\Theta}) = P(X_i \mid Z_i, \mathbf{\Theta}) P(Z_i \mid \mathbf{\Theta}), \tag{A.4}$$

so

$$\begin{aligned} P(X_i \mid \mathbf{\Theta}) &= P(X_i \mid Z_i = 1, \mathbf{\Theta}) P(Z_i = 1 \mid \mathbf{\Theta}) \\ &\quad + P(X_i \mid Z_i = 0, \mathbf{\Theta}) P(Z_i = 0 \mid \mathbf{\Theta}) \\ &= P_{\text{Laplace}}(X_i \mid \mu, b)(1 - \lambda) \\ &\quad + P_{\text{Gauss}}(X_i \mid \sigma_{\text{noise}})\lambda. \end{aligned} \tag{A.5}$$

We define the likelihood function that we optimize as

$$\begin{aligned} L(\mathbf{\Theta} \mid \mathbf{X}, \mathbf{Z}) &:= \ln(P(\mathbf{X}, \mathbf{Z} \mid \mathbf{\Theta})) = \sum_{i=1}^{N} \ln(P(X_i, Z_i \mid \mathbf{\Theta})) \\ &= \sum_{i=1}^{N} \ln(P(X_i \mid Z_i, \mathbf{\Theta})) + \sum_{i=1}^{N} \ln(P(Z_i \mid \mathbf{\Theta})). \end{aligned} \tag{A.6}$$

The expected value of this function is

$$\begin{aligned} &E(L(\mathbf{\Theta} \mid \mathbf{X}, \mathbf{Z})|_{\mathbf{\Theta},\mathbf{X}}) \\ &= E\left(\sum_{i=1}^{N} \ln(P(X_i \mid Z_i, \mathbf{\Theta})) + \sum_{i=1}^{N} \ln(P(Z_i \mid \mathbf{\Theta}))\Big|_{\mathbf{\Theta},\mathbf{X}}\right) \\ &= \sum_{i=1}^{N} E\left(\ln(P(X_i \mid Z_i, \mathbf{\Theta}))\big|_{\mathbf{\Theta},\mathbf{X}}\right) \\ &\quad + \sum_{i=1}^{N} E\left(\ln(P(Z_i \mid \mathbf{\Theta}))\big|_{\mathbf{\Theta},\mathbf{X}}\right) \\ &= \sum_{i=1}^{N} E\left(\ln(P(X_i \mid Z_i, \mathbf{\Theta}))\big|_{\mathbf{\Theta},X_i}\right) \\ &\quad + \sum_{i=1}^{N} E\left(\ln(P(Z_i \mid \mathbf{\Theta}))\big|_{\mathbf{\Theta},X_i}\right) \\ &= \sum_{i=1}^{N} [P(Z_i = 1 \mid \mathbf{\Theta}, X_i) \ln(P(X_i \mid Z_i = 1, \mathbf{\Theta})) \\ &\qquad + P(Z_i = 0 \mid \mathbf{\Theta}, X_i) \ln(P(X_i \mid Z_i = 0, \mathbf{\Theta})) \\ &\qquad + P(Z_i = 1 \mid \mathbf{\Theta}, X_i) \ln(P(Z_i = 1 \mid \mathbf{\Theta})) \\ &\qquad + P(Z_i = 0 \mid \mathbf{\Theta}, X_i) \ln(P(Z_i = 0 \mid \mathbf{\Theta}))]. \end{aligned} \tag{A.7}$$

Let $\gamma_i := P(Z_i = 1 \mid \mathbf{\Theta}, X_i)$ (then $P(Z_i = 0 \mid \mathbf{\Theta}, X_i) = 1 - \gamma_i$), so

$$\begin{aligned} &E(L(\mathbf{\Theta} \mid \mathbf{X}, \mathbf{Z})|_{\mathbf{\Theta},\mathbf{X}}) \\ &= \sum_{i=1}^{N} \left[\gamma_i \ln\left(P_{\text{Laplace}}(X_i \mid \mu, b)\right)\right. \\ &\qquad + (1 - \gamma_i) \ln\left(P_{\text{Gauss}}(X_i \mid \sigma_{\text{noise}})\right) \\ &\qquad \left. + \gamma_i \ln(1 - \lambda) + (1 - \gamma_i) \ln(\lambda)\right] \\ &= \sum_{i=1}^{N} \left[\gamma_i \left(\ln\left(P_{\text{Laplace}}(X_i \mid \mu, b)\right) + \ln(1 - \lambda)\right)\right. \\ &\qquad \left. + (1 - \gamma_i)\left(\ln\left(P_{\text{Gauss}}(X_i \mid \sigma_{\text{noise}})\right) + \ln(\lambda)\right)\right]. \end{aligned} \tag{A.8}$$

As

$$\begin{aligned} \gamma_i &= P(Z_i = 1 \mid \mathbf{\Theta}, X_i) \\ &= \frac{P(X_i \mid Z_i = 1, \mathbf{\Theta}) P(Z_i = 1 \mid \mathbf{\Theta})}{P(X_i \mid \mathbf{\Theta})} \\ &= \left(P(X_i \mid Z_i = 1, \mathbf{\Theta}) P(Z_i = 1 \mid \mathbf{\Theta})\right) \\ &\quad \times \left(P(X_i \mid Z_i = 1, \mathbf{\Theta}) P(Z_i = 1 \mid \mathbf{\Theta})\right. \\ &\qquad \left. + P(X_i \mid Z_i = 0, \mathbf{\Theta}) P(Z_i = 0 \mid \mathbf{\Theta})\right)^{-1}, \end{aligned} \tag{A.9}$$

then

$$\gamma_i = \frac{(1 - \lambda) P_{\text{Laplace}}(X_i \mid \mu, b)}{(1 - \lambda) P_{\text{Laplace}}(X_i \mid \mu, b) + \lambda P_{\text{Gauss}}(X_i \mid \sigma_{\text{noise}})}. \tag{A.10}$$

Replacing P_{Laplace} y P_{Gauss} at expected value expression

$$\begin{aligned} &E(L(\mathbf{\Theta} \mid \mathbf{X}, \mathbf{Z})|_{\mathbf{\Theta},\mathbf{X}}) \\ &= -\sum_{i=1}^{N} \left[\gamma_i \left(\ln(2) + \ln(b) + \frac{|X_i - \mu|}{b} - \ln(1 - \lambda)\right)\right. \\ &\qquad + (1 - \gamma_i)\left(\frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln\left(\sigma_{\text{noise}}^2\right)\right. \\ &\qquad \left.\left. + \frac{X_i}{2\sigma_{\text{noise}}^2} - \ln(\lambda)\right)\right]. \end{aligned} \tag{A.11}$$

To maximize this function, we calculate the gradient and make it equal to 0:

$$\begin{aligned} 0 &= \frac{\partial}{\partial \lambda} E(L(\mathbf{\Theta} \mid \mathbf{X}, \mathbf{Z})|_{\mathbf{\Theta},\mathbf{X}}) \\ &= \sum_{i=1}^{N} \left[\gamma_i \frac{-1}{1 - \lambda} + (1 - \gamma_i) \frac{1}{\lambda}\right] = \sum_{i=1}^{N} \frac{(1 - \lambda)(1 - \gamma_i) - \lambda\gamma_i}{(1 - \lambda)\lambda}. \end{aligned} \tag{A.12}$$

Therefore

$$\widehat{\lambda} := 1 - \frac{1}{N}\sum_{i=1}^{N}\gamma_i,$$

$$0 = \frac{\partial}{\partial\sigma^2_{\text{noise}}} E\left(L\left(\boldsymbol{\Theta} \mid \mathbf{X}, \mathbf{Z}\right)|_{\boldsymbol{\Theta},\mathbf{X}}\right) \tag{A.13}$$

$$= \sum_{i=1}^{N}\left(1-\gamma_i\right)\left(\frac{1}{2\sigma^2_{\text{noise}}} - \frac{X_i^2}{2\sigma^4_{\text{noise}}}\right).$$

Therefore

$$\widehat{\sigma}^2_{\text{noise}} := \frac{\sum_{i=1}^{N}\left(1-\gamma_i\right)X_i^2}{\sum_{i=1}^{N}\left(1-\gamma_i\right)}. \tag{A.14}$$

The estimator of μ, $\widehat{\mu}$ cannot be determined using this technique as the expression $\sum_{i=1}^{N}\gamma_i|\cdot - X_i|$ is piecewise-linear and its minimum is always in a vertex. So in this case a direct method is proposed:

$$\widehat{\mu} := \arg\min_{\mu=X_m}\sum_{i=1}^{N}\gamma_i\left|\mu - X_i\right|. \tag{A.15}$$

Finally the estimation of the shape parameter b is given by

$$0 = \frac{\partial}{\partial b}E\left(L\left(\boldsymbol{\Theta} \mid \mathbf{X}, \mathbf{Z}\right)|_{\boldsymbol{\Theta},\mathbf{X}}\right) = \sum_{i=1}^{N}\gamma_i\left(\frac{1}{b} - \frac{\left|X_i-\mu\right|}{b^2}\right). \tag{A.16}$$

Therefore

$$\widehat{b} := \frac{\sum_{i=1}^{N}\gamma_i\left|X_i-\widehat{\mu}\right|}{\sum_{i=1}^{N}\gamma_i}. \tag{A.17}$$

B. Measurements

This appendix contains different functions used along the paper. The function $|\cdot|$ is the cardinal of the corresponding set ($|A|$ = Cardinal of the set A).

(i) The local mean of an image/volume I at pixel/voxel position m is defined as

$$\text{LM}\left(I, m\right) := \frac{1}{\left|N_m\right|}\sum_{i\in N_m} I\left(i\right), \tag{B.1}$$

where N_m is a neighborhood of pixel/voxel m.

(ii) The local variance of an image/volume I at pixel/voxel position m is defined as

$$\text{LV}\left(I, m\right) := \frac{1}{\left|N_m\right|}\sum_{i\in N_m} I^2\left(i\right) - \text{LM}(I,m)^2, \tag{B.2}$$

where N_m is a neighborhood of pixel/voxel m.

(iii) The global mean of an image/volume I is defined as

$$\text{GM}\left(I\right) := \frac{1}{M}\sum_{m} I\left(m\right), \tag{B.3}$$

where M is the total number of pixels/voxels and m represents the different pixel/voxel positions on the image/volume I.

(iv) The global variance of an image/volume I is defined as

$$\text{GV}\left(I\right) := \frac{1}{M}\sum_{m} I^2\left(m\right) - \text{GM}(I)^2, \tag{B.4}$$

where M is the total number of pixels/voxels and m represents the different pixel/voxel positions on the image/volume I.

(v) The global covariance of images/volumes I and $\widetilde{I}$ is defined as

$$\text{GCV}\left(I, \widetilde{I}\right) := \frac{1}{M}\sum_{m} I\left(m\right)\widetilde{I}\left(m\right) - \text{GM}\left(I\right)\text{GM}\left(\widetilde{I}\right), \tag{B.5}$$

where M is the total number of pixels/voxels and m represents the different pixel/voxel positions on the image/volume I.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Sergio Villullas is supported by the University of Valladolid under the Grant FPI 2010 (rectoral resolution of 16/07/2010, Royal Decree 63/2006). This work was supported by the Ministry of Education and Science under the Grants MTM2009-12561 and MTM2012-31439, by the Institute of Health Carlos III under the Grant PI11-01492, by the Health Regional Council under the Grant GRS 783/A/13, and by the Regional Council of Education and Science under the Grant VA136U13. The authors would like to thank Dr. Fernando Gómez Cubillo, Dr. Miguel Ángel Martín Fernández, Susana Merino Caviedes, and Daniel Argibay Quiñones for their contribution in this paper. Additionally, the authors also thank Centro de Diagnóstico Valladolid (CDV) QDIAGNOSTICA for their collaboration in MRI data sets acquisition and Dr. José María Sierra Román, Dr. Nuria Fabiola Marcos del Río, and Dr. Jaime de la Peña Cadenato for their collaboration in the clinical expert evaluation.

References

[1] G. A. Wright, "Magnetic resonance imaging," *IEEE Signal Processing Magazine*, vol. 14, no. 1, pp. 56–66, 1997.

[2] Z.-P. Liang and P. C. Lauterbur, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*, IEEE Press, 2000.

[3] M. Fuderer, "The information content of MR images," *IEEE Transactions on Medical Imaging*, vol. 7, no. 4, pp. 368–380, 1988.

[4] C. S. Anand and J. S. Sahambi, "Wavelet domain non-linear filtering for MRI denoising," *Magnetic Resonance Imaging*, vol. 28, no. 6, pp. 842–861, 2010.

[5] A. Macovski, "Noise in MRI," *Magnetic Resonance in Medicine*, vol. 36, no. 3, pp. 494–497, 1996.

[6] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[7] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of MRI data," *IEEE Transactions on Medical Imaging*, vol. 11, no. 2, pp. 221–232, 1992.

[8] S. Aja-Fernandez, C. Alberola-Lopez, and C.-F. Westin, "Noise and signal estimation in magnitude MRI and Rician distributed images: a LMMSE approach," *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1383–1398, 2008.

[9] M. Martin-Fernandez, C. Alberola-Lopez, J. Ruiz-Alzola, and C.-F. Westin, "Sequential anisotropic Wiener filtering applied to 3D MRI data," *Magnetic Resonance Imaging*, vol. 25, no. 2, pp. 278–292, 2007.

[10] P. Coupe, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, "An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images," *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 425–441, 2008.

[11] J. V. Manjon, J. Carbonell-Caballero, J. J. Lull, G. García-Marti, L. Martí-Bonmati, and M. Robles, "MRI denoising using Non-Local Means," *Medical Image Analysis*, vol. 12, no. 4, pp. 514–523, 2008.

[12] S. P. Awate and R. T. Whitaker, "Feature-preserving MRI denoising: a nonparametric empirical bayes approach," *IEEE Transactions on Medical Imaging*, vol. 26, no. 9, pp. 1242–1255, 2007.

[13] L. He and I. R. Greenshields, "A nonlocal maximum likelihood estimation method for Rician noise reduction in MR images," *IEEE Transactions on Medical Imaging*, vol. 28, no. 2, pp. 165–172, 2009.

[14] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[15] K. Sivaramakrishnan and T. Weissman, "A context quantization approach to universal denoising," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2110–2129, 2009.

[16] S. P. Awate and R. T. Whitaker, "Higher-order image statistics for unsupervised, information-theoretic, adaptive, image filtering," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 2, pp. 44–51, San Diego, Calif, USA, June 2005.

[17] S. P. Awate and R. T. Whitaker, "Unsupervised, information-theoretic, adaptive image filtering for image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 3, pp. 364–376, 2006.

[18] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 2nd edition, 1998.

[19] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[20] R. D. Nowak, "Wavelet-based Rician noise removal for magnetic resonance imaging," *IEEE Transactions on Image Processing*, vol. 8, no. 10, pp. 1408–1419, 1999.

[21] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A versatile wavelet domain noise filtration technique for medical imaging," *IEEE Transactions on Medical Imaging*, vol. 22, no. 3, pp. 323–331, 2003.

[22] J. Sijbers, A. J. den Dekker, P. Scheunders, and D. van Dyck, "Maximum-likelihood estimation of rician distribution parameters," *IEEE Transactions on Medical Imaging*, vol. 17, no. 3, pp. 357–361, 1998.

[23] S. Basu, T. Fletcher, and R. Whitaker, "Rician noise removal in diffusion tensor MRI," in *Proceedings of the 9th International Conference on Medical Image Computing and Compute r-Assisted Intervention*, vol. 1, pp. 117–125, Copenhagen, Denmark, 2006.

[24] X. Yang and B. Fei, "A wavelet multiscale denoising algorithm for magnetic resonance (MR) images," *Measurement Science and Technology*, vol. 22, no. 2, Article ID 025803, pp. 1–26, 2011.

[25] R. Wirestam, A. Bibic, J. Lätt, S. Brockstedt, and F. Ståhlberg, "Denoising of complex MRI data by wavelet-domain filtering: application to high-b-value diffusion-weighted imaging," *Magnetic Resonance in Medicine*, vol. 56, no. 5, pp. 1114–1120, 2006.

[26] P. Gorgel, A. Sertbas, and O. N. Ucan, "A wavelet-based mammographic image denoising and enhancement with homomorphic filtering," *Journal of Medical Systems*, vol. 34, no. 6, pp. 993–1002, 2010.

[27] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society B: Methodological*, vol. 39, no. 1, pp. 1–38, 1977.

[28] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, 1996.

[29] R. Coifman and D. Donoho, "Translation invariant de-noising," in *Wavelets and Statistics*, pp. 125–150, Springer, New York, NY, USA, 1995.

[30] M. Figueiredo and R. Nowak, "Bayesian wavelet-based signal estimation using non-informative priors," in *Proceedings of the 32nd Asilomar Conference on Signals, Systems & Computers*, pp. 1368–1373, Pacific Grove, Calif, USA, November 1998.

[31] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells Jr., "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 10–12, 1996.

[32] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 617–643, 1992.

[33] R. Nowak, "Shift invariant wavelet-based statistical models and 1/f processes," in *Proceedings of the IEEE Digital Signal Processing Workshop*, vol. 83, pp. 1428–1450, Bryce Canyon, Utah, USA, 1998.

[34] K. Berkner and J. Wells, "Smoothness estimates for soft-threshold denoising via translation-invariant wavelet transforms," *Applied and Computational Harmonic Analysis*, vol. 12, no. 1, pp. 1–24, 2002.

[35] S. Aja-Fernandez, A. Tristán-Vega, and C. Alberola-Lopez, "Noise estimation in single- and multiple-coil magnetic resonance data based on statistical models," *Magnetic Resonance Imaging*, vol. 27, no. 10, pp. 1397–1409, 2009.

[36] S. Aja-Fernandez and A. Tristan-Vega, "A review on statistical noise models for magnetic resonance imaging," Technical Report of the LPI, University of Valladolid, Valladolid, Spain, 2013.

[37] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magnetic Resonance in Medicine*, vol. 34, no. 6, pp. 910–914, 1995.

[38] D. L. Collins, A. P. Zijdenbos, V. Kollokian et al., "Design and construction of a realistic digital brain phantom," *IEEE*

Transactions on Medical Imaging, vol. 17, no. 3, pp. 463–468, 1998.

[39] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1988.

[40] S. C. Kale, X. J. Chen, and R. M. Henkelman, "Trading off SNR and resolution in MR images," *Nuclear Magnetic Resonance in Biomedicine*, vol. 22, no. 5, pp. 488–494, 2009.

[41] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 691021, 9 pages
http://dx.doi.org/10.1155/2015/691021

Research Article

NUFFT-Based Iterative Image Reconstruction via Alternating Direction Total Variation Minimization for Sparse-View CT

Bin Yan, Zhao Jin, Hanming Zhang, Lei Li, and Ailong Cai

China National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450002, China

Correspondence should be addressed to Lei Li; leehotlin@aliyun.com

Received 18 September 2014; Accepted 11 January 2015

Academic Editor: Yi Gao

Sparse-view imaging is a promising scanning method which can reduce the radiation dose in X-ray computed tomography (CT). Reconstruction algorithm for sparse-view imaging system is of significant importance. The adoption of the spatial iterative algorithm for CT image reconstruction has a low operation efficiency and high computation requirement. A novel Fourier-based iterative reconstruction technique that utilizes nonuniform fast Fourier transform is presented in this study along with the advanced total variation (TV) regularization for sparse-view CT. Combined with the alternating direction method, the proposed approach shows excellent efficiency and rapid convergence property. Numerical simulations and real data experiments are performed on a parallel beam CT. Experimental results validate that the proposed method has higher computational efficiency and better reconstruction quality than the conventional algorithms, such as simultaneous algebraic reconstruction technique using TV method and the alternating direction total variation minimization approach, with the same time duration. The proposed method appears to have extensive applications in X-ray CT imaging.

1. Introduction

X-ray computed tomography (CT) has been widely used for imaging applications in various fields, such as industrial nondestructive testing [1] and medical diagnosis [2], for its advantages of noninvasive and high spatial resolution. However, in many practical applications of X-ray computed tomography, complete projection data set cannot be obtained because of the limitation of scanning time, space, dose, and so on. Therefore, sparse angle scanning scheme is adopted to tackle these problems. On one hand, this scheme can speed up the scanning rate and decrease the X-ray radiation dose, such as breast and vascular imaging [3–6]. On the other hand, sparse data sampling can save much scanning time and it is of practice value when high reconstruction precision is not that urgent. To solve the sparse-view reconstruction problem, the classical methods should be upgraded, and a new algorithm needs to be developed.

Given the unsatisfactory Tuy-Smith condition [7, 8] in sparse-view, a CT image cannot be accurately reconstructed via analytic method. To solve the ill-posed problem [9, 10], numerous iterative algorithms [11–13] have been proposed based on spatial domain. However, these iterative algorithms are time-consuming and have a great demand for hardware resources. Despite applying hardware speedup technology, such as an ordinary graphics processing unit [14], these algorithms still consume a considerable amount of time.

Compressive sensing theory by Candés et al. [15–17] provided a new idea for the exact recovery of an image from the sparse samples of its discrete Fourier transform. The exact reconstruction relies on the assumption that there exists sparse representation for an image. A number of cases are known to have sparse gradient-magnitude images. In some cases, minimizing the total variation (TV) can generate accurate images from sparse samples [18–20]. Therefore, combining TV regularization with the iterations of the simultaneous algebraic reconstruction technique (SART), hereinafter called SART–TV [21], can improve reconstruction image quality while decreasing mean-squared error. Based on the projection onto the convex sets (POCS) algorithm,

the adaptive steepest descent-POCS (ASD-POCS) algorithm [20] can effectively handle incomplete datasets and demonstrates excellent performance in sparse-view CT applications.

The rapid increase in the size of scanning data has highlighted the importance of reducing reconstruction time and improving reconstruction quality. It is known that processing the same signal in the frequency space is faster than that in the spatial domain by fast Fourier transform (FFT). Several algorithms for image reconstruction in the frequency space can also be developed on the basis of fast Fourier transform. Several studies have been conducted to achieve this goal. In 1981, Stark et al. [22] developed direct Fourier methods (DFM) using central slice theorem and obtained favorable results. In 2003, Seger and Danielsson [23] analyzed the missing projection data in the frequency domain and proposed a reconstruction method for the scanned timber data according to Fourier transform. In 2013, Fahimian et al. [24] presented a Fourier-based iterative reconstruction in medical X-ray CT, and numerical experiment results showed that this method required less computation time than other iterative algorithms. These achievements facilitated the development of an improved algorithm for solving the sparse-view reconstruction problem in the frequency domain.

Because of the limitation of FFT, that is, its unsuitability for application to nonuniform samples, this technique requires further enhancement to improve its universality. To this end, nonuniform FFT (NUFFT) [25] has been recently developed to overcome this limitation without increasing the computation complexity of FFT. NUFFT is also basis of the proposed reconstruction algorithm in this study. Motivated by the feature of NUFFT for data distribution, some approaches have been proposed to reconstruct a CT image to deal with frequency data. In 2004, Matej et al. [26] proposed an iterative tomographic image reconstruction method using NUFFT and obtained better results with this technique than with the filtered back-projection (FBP) algorithm. In 2006, Zhang-O'Connor and Fessler [27] proposed Fourier-based forward- and back-projectors for fan-beam tomographic image reconstruction. However, these proposed NUFFT cannot effectively solve the problems in sparse-view image reconstruction. The NUFFT just was especially applied as a transition during iteration in spatial domain, which in turn burdened computation consumption.

In this paper, our study aims to present a promising contribution to the task of image reconstruction from sparse-view by combining the alternating direction total variation minimization (ADTVM) technique with NUFFT to establish a new method which is suitable for large-scale reconstruction because of its low computational requirement. The algorithm is developed under the framework of alternating direction method (ADM) which shows high efficiency and stability. The advantages of the proposed algorithm are verified by the results of several groups of experiments.

The organization of this paper is organized as follows. Section 1 concisely reviews the basic CT reconstruction and the state of the art of sparse-view image reconstruction. Section 2 describes the basic principles of the proposed method, including the reconstruction model and the corresponding algorithm based on NUFFT and ADTVM.

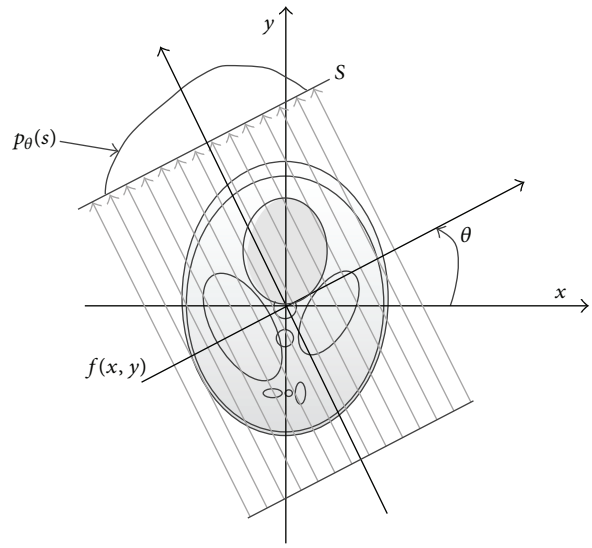


Figure 1: An object $f(x, y)$ and its parallel projection $p_\theta(s)$.

Section 3 demonstrates groups of typical experiments results, including numerical simulation and real data ones. Section 4 discusses the findings of the experiments and concludes the paper.

2. Method and Material

2.1. Image Reconstruction Model. In this work, we consider temporarily parallel geometry. In parallel geometry, 2D function f (objection function) is defined in a compact support of spatial domain, which means that it vanishes outside a finite region of the plane. In the (x, y) plane, the general formation for the line integral, known as the Radon transform of $f(x, y)$, is

$$p_\theta(s) = \iint_{-\infty}^{\infty} f(x, y)\, \delta(x\cos\theta + y\sin\theta - s)\, dx\, dy. \quad (1)$$

In Figure 1, projection $p_\theta(s)$ consists of a collection of line integrals (1) taken along straight parallel lines in the plane that means a collection of $p_\theta(s)$ with constant $\theta \in [0, \pi/2]$ and $s \in [-S/2, S/2]$.

CT image reconstruction is an inverse problem, and the observed projections should be converted into images which reflect the distribution of the attenuation coefficient of the interested physical object. A conventional reconstruction method, that is, the direct Fourier methods (DFM), is established based on the Fourier slice theorem. Basically, the steps of DFM can be summarized as follows:

(a) 1D discrete Fourier transform of the parallel projections taken at different angles;

(b) polar to Cartesian grid interpolations;

(c) 2D inverse Fourier transform.

From the perspective of DFM, the observation equation in the Fourier domain can be expressed as follows:

$$P_\theta(\rho) = \int_{-\infty}^{\infty} p_\theta(s)\, e^{-j2\pi\rho s} ds, \tag{2}$$

where $P_\theta(\rho)$ is observed by the Fourier transform of the measured projection data $p_\theta(s)$ and ρ is the frequency variable of Fourier transform. The process proceeds with the use of FFT and is characterized by high accuracy. According to Fourier central slice theory, the 1D FFT of the projection $P_\theta(\rho)$ is equal to $\widehat{f}(u, v)$, which is derived from the 2D FFT of the reconstructed image in a certain angle. Therefore, the relationship is described as follows:

$$\begin{aligned} P_\theta(\rho) &= \int_{-\infty}^{+\infty} p_\theta(s)\, e^{-j2\pi\rho s} ds \\ &= \widehat{f}(u, v)\Big|_{\substack{u=\rho\cos\theta \\ v=\rho\sin\theta}}, \end{aligned} \tag{3}$$

where $\widehat{f}(u, v) = \iint_{-\infty}^{+\infty} f(x, y) \mathrm{e}^{-j2\pi(xu+yv)} dx\, dy$. The equation shows an obvious equivalence corresponding to frequency projection $P_\theta(\rho)$ with $\widehat{f}(u, v)$ in polar coordinates.

To avoid interpolation errors, such as DFM in the image and frequency domains, this study introduces NUFFT, which can translate polar coordinates into the image space without interpolation. This technique can significantly improve the accuracy of reconstruction. Let F_N represent the NUFFT operator, such that the following equation can be derived:

$$\begin{aligned} \mathsf{F}_N(f) &= \iint_{-\infty}^{+\infty} f(x, y)\, e^{-j2\pi(xu+yv)} dx\, dy\Big|_{\substack{u=\rho\cos\theta \\ v=\rho\sin\theta}} \\ &= \int_{-\infty}^{\infty} p_\theta(s)\, e^{-j2\pi\rho s} ds \\ &= P_\theta(\rho). \end{aligned} \tag{4}$$

The reconstruction module can be discretely shown as follows:

$$\mathbf{P} = \mathsf{F}_N \mathbf{f}, \tag{5}$$

where the (observed) constant F_N and the variant $\mathbf{f}$ are the vector forms of Fourier sampling and objection function, respectively. Matrix F_N stands for the NUFFT of $\mathbf{f}$. The Fourier transform $\mathsf{F}_N\mathbf{f}$ in (5) and its adjoint $\mathsf{F}_N^H\mathbf{f}$ can be implemented by using FFT to generate a fast and accurate evaluation.

In sparse-view reconstruction, (5) is ill-posed, and the projection data are insufficient for exact reconstruction. Mathematically, the problem that we consider here involves insufficient data, such that (5) is underdetermined. To solve this linear and underdetermined equation, we specify a TV minimization algorithm that considers the reconstruction to be a task of finding the best solution to the following optimization problem:

$$\begin{aligned} &\mathbf{f}^* = \arg\min \|\mathbf{f}\|_{\mathrm{TV}}, \\ &\text{subject to (s.t.)} \quad \mathbf{P} = \mathsf{F}_N\mathbf{f}, \quad \mathbf{f} \geq 0, \end{aligned} \tag{6}$$

where $\|\mathbf{f}\|_{\mathrm{TV}}$ denotes the discretization of the TV term and $\|\mathbf{f}\|_{\mathrm{TV}} = \sum_i \|D_i\mathbf{f}\|_1$. By applying the directional gradients operators D_i [20, 28], model (6) can also be written as follows:

$$\begin{aligned} \min \quad & \sum_i \|\mathbf{w}_i\|_1 + \frac{\lambda}{2}\|\mathsf{F}_N\mathbf{f} - \mathbf{P}\|_2^2, \\ \text{s.t.} \quad & D_i\mathbf{f} = \mathbf{w}_i, \quad \mathbf{f} \geq 0, \end{aligned} \tag{7}$$

where λ is the fidelity parameter to control the data consistency in the object function.

Therefore, the overall reconstruction flowchart can be summarized as Figure 2.

2.2. NUFFT with ADM for the Model. The above constrained optimization is addressed by converting the equation into its unconstrained form by applying the augmented Lagrange function:

$$\begin{aligned} \min L(\mathbf{f}, \mathbf{w}_i) = \sum_i \Big(& \|\mathbf{w}_i\|_1 + u_i^T (D_i\mathbf{f} - \mathbf{w}_i) \\ & + \frac{\rho_i}{2}\|D_i\mathbf{f} - \mathbf{w}_i\|_2^2 \Big) \\ & + \frac{\lambda}{2}\|\mathsf{F}_N\mathbf{f} - \mathbf{P}\|_2^2, \end{aligned} \tag{8}$$

where ρ_i is a scalar that denotes the penalty coefficient and u_i denotes the multipliers. The minimization processes with respect to variables $\mathbf{f}$ and $\mathbf{w}_i$ cannot be easily realized simultaneously by directly performing the optimization. Moreover, decomposing the variables by using ADM has a low computation cost. The ADM approach decouples the augmented Lagrange function into two subproblems, namely, the $\mathbf{w}$-subproblem and the $\mathbf{f}$-subproblem [29].

The $\mathbf{w}$-subproblem can be written as follows:

$$\min L_{\mathbf{f}}(\mathbf{w}_i) = u_i^T (D_i\mathbf{f} - \mathbf{w}_i) + \frac{\rho_i}{2}\|D_i\mathbf{f} - \mathbf{w}_i\|_2^2 + \|\mathbf{w}_i\|_1. \tag{9}$$

The $\mathbf{w}$-subproblem is separable with respect to $\mathbf{w}$, and problem (9) can be efficiently solved by using the shrinkage operator [30], which is expressed as follows:

$$\mathbf{w}_i^* = \max\left\{\left|D_i\mathbf{f} + \frac{u_i}{\rho_i}\right| - \frac{1}{\rho_i}, 0\right\} \cdot \operatorname{sgn}\left(D_i\mathbf{f} + \frac{u_i}{\rho_i}\right). \tag{10}$$

In addition, with the aid of $\mathbf{w}_i$, the optimization of $\mathbf{f}$-subproblem can be achieved by solving the following:

$$\begin{aligned} \min L_{\mathbf{w}_i}(\mathbf{f}) = & \frac{\lambda}{2}\|\mathsf{F}_N\mathbf{f} - \mathbf{P}\|_2^2 \\ & + \sum_i \left(u_i^T (D_i\mathbf{f} - \mathbf{w}_i) + \frac{\rho_i}{2}\|D_i\mathbf{f} - \mathbf{w}_i\|_2^2 \right). \end{aligned} \tag{11}$$

$L_{\mathbf{w}_i}(\mathbf{f})$ is clearly a quadratic function, the gradient of which is expressed as follows:

$$l(\mathbf{f}) = \lambda\mathbf{F}_N^H(\mathsf{F}_N\mathbf{f} - \mathbf{P}) + \sum_i \left(D_i^T u_i + \rho_i D_i^T (D_i\mathbf{f} - \mathbf{w}_i)\right). \tag{12}$$

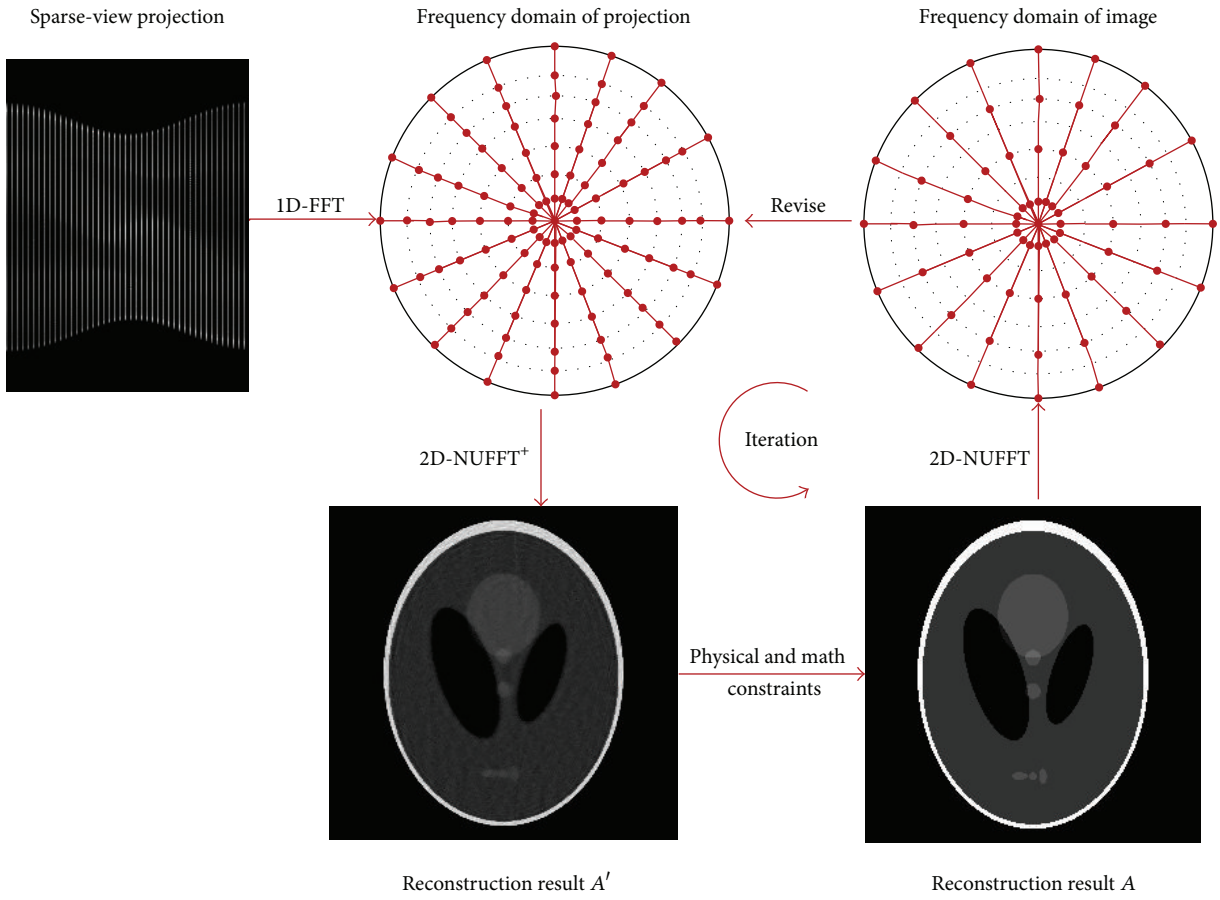


Figure 2: The flowchart of sparse-view image reconstruction for the model.

Force $l(\mathbf{f}) = 0$ and the exact solution for $L(\mathbf{f})$ is presented as follows:

$$\mathbf{f}^* = \left(\lambda \mathsf{F}_N^T \mathsf{F}_N + \sum_i \rho_i D_i^T D_i\right)^+ \cdot \left(\lambda \mathsf{F}_N^T P - \sum_i D_i^T \left(u_i^T - \rho_i \mathbf{w}_i\right)\right), \tag{13}$$

where M^+ denotes the Moore-Penrose inverse of matrix M. Theoretically, the exact minimizer can be used to solve the $\mathbf{f}$-subproblem. However, the inverse or pseudo-inverse is too costly to compute numerically at each iteration. The augmented Lagrangian function (8) is expected to be minimized by solving the $\mathbf{w}$-subproblem and the $\mathbf{f}$-subproblem alternately. Therefore, solving the $\mathbf{f}$-subproblem accurately at each sweep may be unnecessary. A robust and efficient nonmonotone alternating direction algorithm [31] is used to solve problem (13).

By using the solutions of $\mathbf{w}_i^*$ and $\mathbf{f}^*$, the multipliers are updated as follows:

$$u_i = u_i + \rho_i \left(D_i \mathbf{f}^* - \mathbf{w}_i^*\right). \tag{14}$$

The optimized solution for (8) is attained by circularly applying (10) and (13) until $L(\mathbf{f}, \mathbf{w}_i)$ is minimized jointly with respect to $(\mathbf{f}, \mathbf{w}_i)$.

2.3. Algorithm of the Overall Framework. All issues in handling the subproblems have been addressed in Section 2.2. In light of all derivations presented above, the new algorithm for solving the reconstruction problem can be stated as follows.

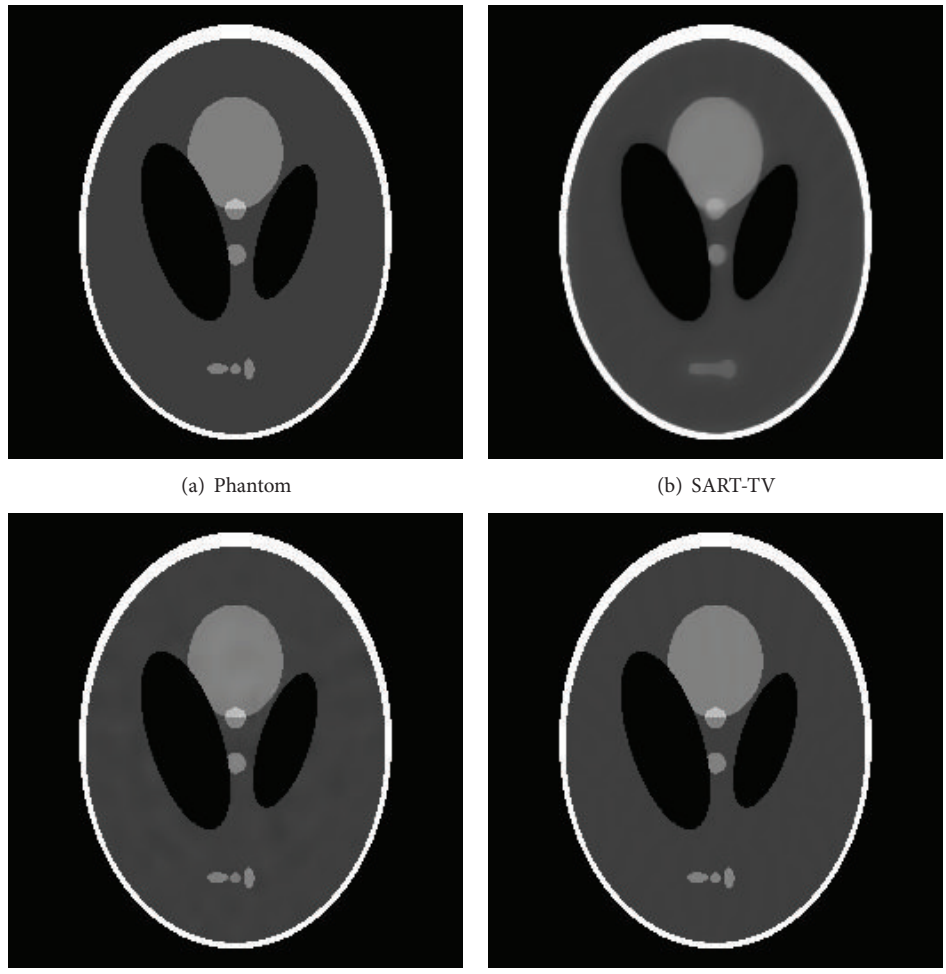
Algorithm 1. Input projection data $p, \lambda, \rho_i > 0$. Initialize $u_i = u_i^{(0)}$ and starting points w_i^0, u_i^0 for all i. Set $k = 0$.

(1) make 1D FFT of $p_\theta(s)$ with respect to s

$$P_\theta(\rho) \longleftarrow \int_{-\infty}^{+\infty} p_\theta(s)\, e^{-j2\pi\rho s} ds; \tag{15}$$

while "not achieved maximum iteration loops," **Do**

(2) compute frequency domain $P_\theta(\rho)$ via NUFFT;



(a) Phantom (b) SART-TV

(c) ADTVM (d) New method

Figure 3: Image reconstructions of the Shepp-Logan phantom in 60-view scan. Display Window (0.1 0.5). (a) shows the original image. (b) shows the image after applying the SART-TV algorithm at 200 iterations. (c) gives the image after vua the ADTVM algorithm at 200 iterations. (d) presents the reconstruction image after using the NUFFT-ADM algorithm at 200 iterations.

(3) compute $\mathbf{f}$ by

$$\mathbf{f}_i^{k+1} \longleftarrow \left(\lambda \mathbf{F}_N^T \mathbf{F}_N + \sum_i \rho_i D_i^T D_i\right)^+ \cdot \left(\lambda \mathbf{F}_N^T P - \sum_i D_i^T \left(u_i^T - \rho_i \mathbf{w}_i\right)\right); \quad (16)$$

(4) compute $\mathbf{w}$ by

$$\mathbf{w}_i^{k+1} \longleftarrow \max\left\{\left|D_i \mathbf{f} + \frac{u_i}{\rho_i}\right| - \frac{1}{\rho_i}, 0\right\} \cdot \operatorname{sgn}\left(D_i \mathbf{f} + \frac{u_i}{\rho_i}\right); \quad (17)$$

(5) update u_i by

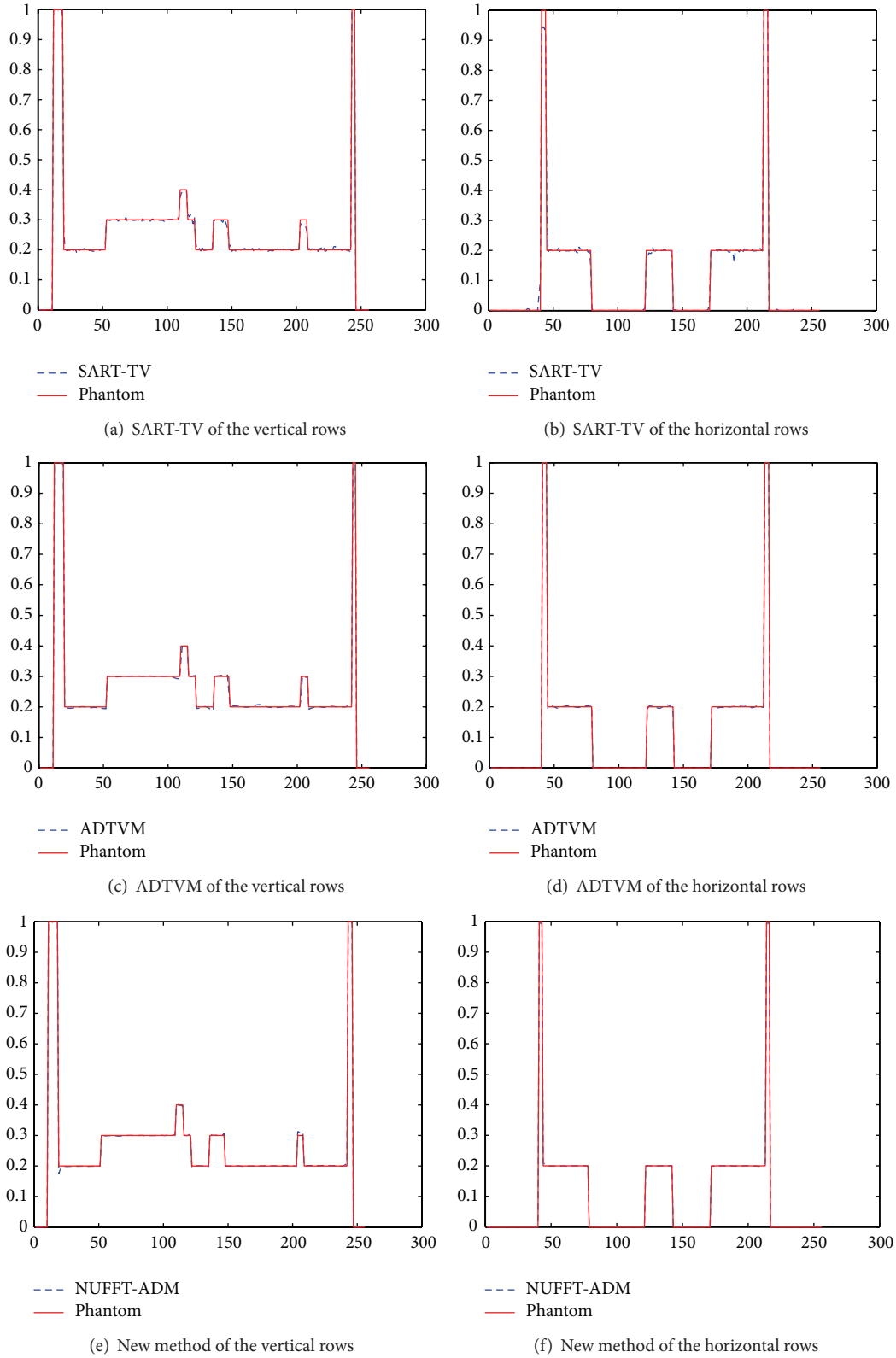
$$u_i^{k+1} \longleftarrow u_i + \rho_i \left(D_i \mathbf{f}^* - \mathbf{w}_i^*\right); \quad (18)$$

(6) $k \leftarrow k + 1$
End Do

In this study, the proposed NUFFT reconstruction technique is developed on the basis of ADTVM. This technique is called NUFFT-ADM. According to the above algorithm, the proposed method demonstrates fast convergence and effective iteration through ADM. This method can be effectively implemented in large-scale reconstruction in sparse-view because of its low computational cost, thus making this technique promising in practical applications.

3. Experiments

To verify the performance of the proposed algorithm, both numerical simulations and real CT scan data experiments are conducted. All experiments are performed on an AMAX Tesla workstation with Intel Xeon E5520 dual-core CPU 2.27 GHz and 24 GB memory. All programs are performed using MATLAB 2011a. In all experiments, the parameter of TV is that primary penalty parameter μ and secondary penalty parameter β are 1024 and 32, respectively.



(a) SART-TV of the vertical rows

(b) SART-TV of the horizontal rows

(c) ADTVM of the vertical rows

(d) ADTVM of the horizontal rows

(e) New method of the vertical rows

(f) New method of the horizontal rows

Figure 4: The image profile in Figure 1 shows (a) vertical profiles along the center of the SART-TV result, (b) horizontal profiles along the center of the SART-TV result, (c) vertical profiles along the center of the ADTVM result, and (d) horizontal profiles along the center of the ADTVM result. (e) Vertical profiles along the center of the NUFFT-ADM result and (f) horizontal profiles along the center of the NUFFT-ADM result.

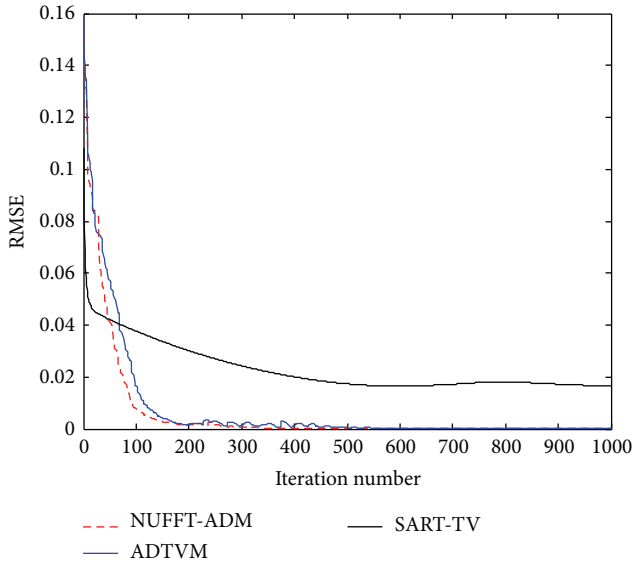


Figure 5: RMSEs as functions of iterations for three different algorithms.

Table 1: Parameters in the simulation of a sparse-view scan.

Parameters	Configuration
Detection elements	512
Source to axis distance	300 mm
Source to detection distance	600 mm
Views of projection data	18
Projection data	512 × 18
Reconstruction size	256 × 256 pixels
Pixel size	0.127 mm × 0.127 mm

3.1. Numerical Phantom Simulation. The above algorithm is applied to validate its high efficiency. A group of 2D image reconstruction experiments are performed using a 2D Shepp-Logan phantom with a size of 256 × 256. This phantom is generated according to the definition of the ellipse phantom image. The scanning and reconstruction parameters in the experiment are listed in Table 1. The detector elements are equidistantly spaced 0.127 mm from one another.

To demonstrate the reconstruction accuracy quantitatively, the root-mean-squared error (RMSE) is used as a measurement of the reconstruction error. RMSE is defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_i \sum_j \left| f(i,j) - g(i,j) \right|^2}{N}}, \tag{19}$$

where f and g denote the ideal phantom and the reconstruction image, respectively; and N denotes the total number of pixels of the image. The image was reconstructed using SART-TV, ADTVM, the proposed method, respectively, and their results are presented in Figure 3. Two hundred iterations are performed for each algorithm. The profiles of these images along the central horizontal and vertical rows are shown in Figure 4 for the different methods.

RMSE is used as an evaluation criterion for different iteration times. The result is shown in Figure 5. The accuracy and running time of each reconstruction method at different iterations are presented in Table 2 for comparison.

The RMSEs, as well as the accuracy and running time of different methods, show that NUFFT-ADM significantly outperforms SART-TV and ADTVM. On one hand, the convergence of the new method is faster than that of SART-TV because of the use of the optimal solution with ADM. On the other hand, by taking advantage of the frequency NUFFT operator instead of the projection and back-projection in the spatial domain, which consumes the greatest amount of time among all components, NUFFT-ADM has a higher computation capability than the general algorithm in the spatial domain.

3.2. Reconstruction Using Real Data. To further verify the performance of the proposed algorithm, several experiments are performed to reconstruct a head model from real data using the new method. We compare the proposed algorithm with SART-TV and ADTVM. Table 3 lists the scanning and reconstruction parameters. The detector elements are equidistantly spaced 0.635 mm from one another. The number of iterations for NUFFT-ADM, ADTVM, and SART-TV is 200.

The reconstruction results are presented in Figure 6.

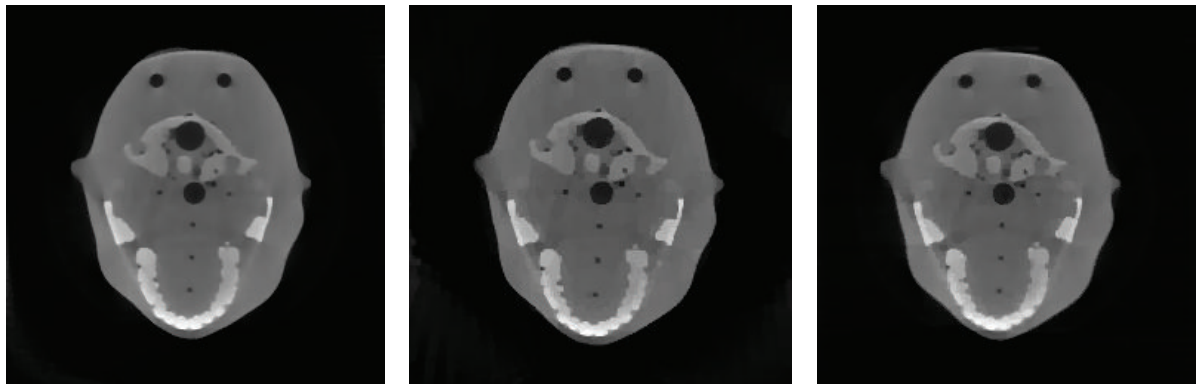
The reconstruction acquired results using real data clearly show that the quality of the reconstructed image is improved as the number of iterations is increased. Under the same number of iterations, the reconstruction results of NUFFT-ADM are superior to those of SART-TV, especially in terms of the high-frequency information that shows the image detail or volatile part. Compared to ADTVM, the detail of the image by the new method is nearly the same.

4. Conclusions

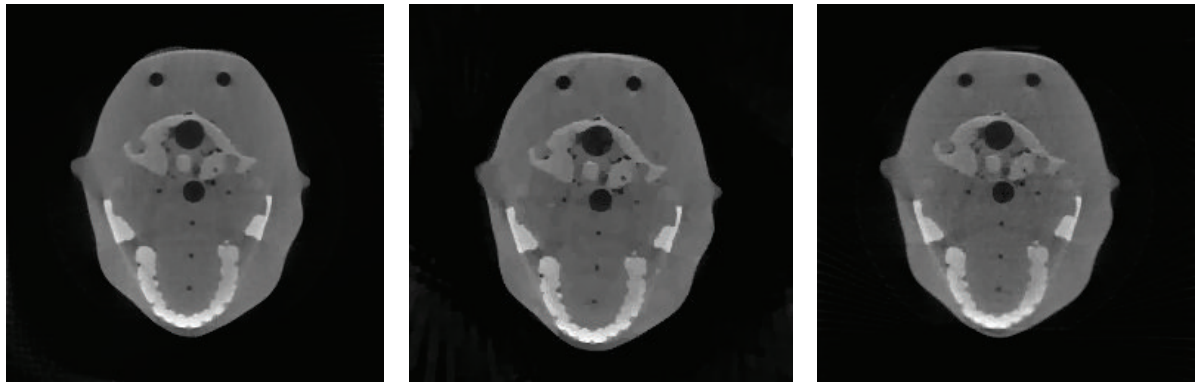
An optimal algorithm based on NUFFT for CT image reconstruction is presented in this work. The validity of the NUFFT-ADM algorithm is verified by conducting numerical simulations and real data experiments. The reconstruction results show that the proposed reconstruction algorithm improves reconstruction quality, accelerates convergence speed, and demonstrates lower computation complexity than other iterative algorithms. That is, the NUFFT-ADM algorithm can practically deal with fast image reconstruction from sparse projection measurements to reduce the radiation dose in X-ray CT. In principle, the proposed method can be extended to fan-beam geometry via the rebinning method to broad its application.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.



(a) SART-TV (at 100 iterations) (b) ADTVM (at 100 iterations) (c) New method (at 100 iterations)



(d) SART-TV (at 300 iterations) (e) ADTVM (at 300 iterations) (f) New method (at 300 iterations)

Figure 6: Reconstruction results of SART-TV, ADTVM, and NUFFT-ADM.

Table 2: Accuracy and the running time of different method.

Iteration number	SART-TV		ADTVM		NUFFT-ADM	
	RMSE	Running time	RMSE	Running time	RMSE	Running time
100	0.0377	177.786 s	0.0165	107.439 s	0.0079	8.986 s
200	0.0302	349.544 s	0.0015	199.133 s	0.0012	16.752 s
500	0.0174	856.158 s	$4.8927e-4$	508.376 s	$1.6378e-4$	51.863 s

Table 3: Parameters in the simulation of a sparse-view scan.

Parameters	Configuration
Detection elements	640
Source to axis distance	678 mm
Source to detection distance	1610 mm
Views of projection data	60
Projection data	60×72
Reconstruction size	256×256 pixels
Pixel size	$0.582 \times 0.582\ \text{mm}^2$

Acknowledgments

This work was supported in part by the National High Technology Research and Development Program of China under the Grant no. 2012AA011603 and the National Natural Science Foundation of China under Grant no. 61372172.

References

[1] H. E. Martz, S. G. Azevedo, J. M. Brase, K. E. Waltjen, and D. J. Schneberk, "Computed tomography systems and their industrial applications," *International Journal of Radiation Applications and Instrumentation, Part A: Applied Radiation and Isotopes*, vol. 41, no. 10-11, pp. 943–961, 1990.

[2] S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 6, pp. 490–498, 2001.

[3] X. Gong, S. J. Glick, B. Liu, A. A. Vedula, and S. Thacker, "A computer simulation study comparing lesion detection accuracy with digital mammography, breast tomosynthesis, and cone-beam CT breast imaging," *Medical Physics*, vol. 33, no. 4, pp. 1041–1052, 2006.

[4] S. J. Glick, S. Thacker, X. Gong, and B. Liu, "Evaluating the impact of X-ray spectral shape on image quality in flat-panel CT breast imaging," *Medical Physics*, vol. 34, no. 1, pp. 5–24, 2007.

[5] C. J. Henri and T. M. Peters, "Three-dimensional reconstruction of vascular trees. Theory and methodology," *Medical Physics*, vol. 23, no. 2, pp. 197–204, 1996.

[6] S. M. Jorgensen, O. Demirkaya, and E. L. Ritman, "Three-dimensional imaging of vasculature and parenchyma in intact rodent organs with X-ray micro-CT," *American Journal of Physiology—Heart and Circulatory Physiology*, vol. 275, no. 3, pp. H1103–H1114, 1998.

[7] H. K. Tuy, "An inversion formula for cone-beam reconstruction," *SIAM Journal on Applied Mathematics*, vol. 43, no. 3, pp. 546–552, 1983.

[8] B. D. Smith, "Image reconstruction from cone-beam projections: necessary and sufficient conditions and reconstruction methods," *IEEE Transactions on Medical Imaging*, vol. 4, no. 1, pp. 14–25, 1985.

[9] F. Natterer, *The Mathematics of Computerized Tomography*, B. G. Teubner, Stuttgart, Germany, 1986.

[10] F. Natterer, *Mathematical Methods in Image Reconstruction*, SIAM, 2001.

[11] L.-Y. Wang, L. Li, B. Yan, C.-S. Jiang, H.-Y. Wang, and S.-L. Bao, "An algorithm for computed tomography image reconstruction from limited-view projections," *Chinese Physics B*, vol. 19, no. 8, Article ID 088106, 2010.

[12] M. Beister, D. Kolditz, and W. A. Kalender, "Iterative reconstruction methods in X-ray CT," *Physica Medica*, vol. 28, no. 2, pp. 94–108, 2012.

[13] H.-M. Zhang, L.-Y. Wang, B. Yan, L. Li, X.-Q. Xi, and L.-Z. Lu, "Image reconstruction based on total-variation minimization and alternating direction method in linear scan computed tomography," *Chinese Physics B*, vol. 22, no. 7, Article ID 078701, 2013.

[14] X. Jia, Y.-F. Lou, R.-J. Li, W. Y. Song, and S. B. Jiang, "GPU-based fast cone beam CT reconstruction from undersampled and noisy projection data via total variation," *Medical Physics*, vol. 37, no. 4, pp. 1757–1760, 2010.

[15] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[16] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[17] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[18] M. Chang, L. Li, Z. Chen, Y. Xiao, L. Zhang, and G. Wang, "A few-view reweighted sparsity hunting (FRESH) method for CT image reconstruction," *Journal of X-Ray Science and Technology*, vol. 21, no. 2, pp. 161–176, 2013.

[19] A.-L. Cai, L.-Y. Wang, H.-M. Zhang et al., "Edge guided image reconstruction in linear scan CT by weighted alternating direction TV minimization," *Journal of X-Ray Science and Technology*, vol. 22, no. 3, pp. 335–349, 2014.

[20] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in Medicine and Biology*, vol. 53, no. 17, pp. 4777–4807, 2008.

[21] H. Y. Yu and G. Wang, "Compressed sensing based interior tomography," *Physics in Medicine and Biology*, vol. 54, no. 9, pp. 2791–2805, 2009.

[22] H. Stark, J. W. Woods, I. Paul, and R. Hingorani, "Direct Fourier reconstruction in computer tomography," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 2, pp. 237–245, 1981.

[23] M. M. Seger and P.-E. Danielsson, "Scanning of logs with linear cone-beam tomography," *Computers and Electronics in Agriculture*, vol. 41, no. 1-3, pp. 45–62, 2003.

[24] B. P. Fahimian, Y. Zhao, Z. Huang et al., "Radiation dose reduction in medical X-ray CT via Fourier-based iterative reconstruction," *Medical Physics*, vol. 40, no. 3, Article ID 031914, 2013.

[25] A. Dutt and V. Rokhlin, "Fast Fourier transforms for nonequispaced data," *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1368–1393, 1993.

[26] S. Matej, J. A. Fessler, and I. G. Kazantsev, "Iterative tomographic image reconstruction using Fourier-based forward and back-projectors," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 401–412, 2004.

[27] Y. Zhang-O'Connor and J. A. Fessler, "Fourier-based forward and back-projectors in iterative fan-beam tomographic image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 25, no. 5, pp. 582–589, 2006.

[28] C. B. Li, *An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing [M.S. thesis]*, Rice University, 2009.

[29] Y. Xiao, J. Yang, and X. Yuan, "Alternating algorithms for total variation image reconstruction from random projections," *Inverse Problems and Imaging*, vol. 6, no. 3, pp. 547–563, 2012.

[30] Y. Wang, J. Yang, W. Yin, and Y. Zhang, "A new alternating minimizationalgorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.

[31] C. Li, W. Yin, H. Jiang, and Y. Zhang, "An efficient augmented Lagrangian method with applications to total variation minimization," *Computational Optimization and Applications*, vol. 56, no. 3, pp. 507–530, 2013.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 836202, 16 pages
http://dx.doi.org/10.1155/2015/836202

Research Article

Deep Adaptive Log-Demons: Diffeomorphic Image Registration with Very Large Deformations

Liya Zhao and Kebin Jia

College of Electronic Information & Control Engineering, Beijing University of Technology, Beijing 100124, China

Correspondence should be addressed to Kebin Jia; kebinj@bjut.edu.cn

Received 31 October 2014; Revised 28 December 2014; Accepted 28 December 2014

Academic Editor: Yi Gao

This paper proposes a new framework for capturing large and complex deformation in image registration. Traditionally, this challenging problem relies firstly on a preregistration, usually an affine matrix containing rotation, scale, and translation and afterwards on a nonrigid transformation. According to preregistration, the directly calculated affine matrix, which is obtained by limited pixel information, may misregistrate when large biases exist, thus misleading following registration subversively. To address this problem, for two-dimensional (2D) images, the two-layer deep adaptive registration framework proposed in this paper firstly accurately classifies the rotation parameter through multilayer convolutional neural networks (CNNs) and then identifies scale and translation parameters separately. For three-dimensional (3D) images, affine matrix is located through feature correspondences by a triplanar 2D CNNs. Then deformation removal is done iteratively through preregistration and demons registration. By comparison with the state-of-the-art registration framework, our method gains more accurate registration results on both synthetic and real datasets. Besides, principal component analysis (PCA) is combined with correlation like Pearson and Spearman to form new similarity standards in 2D and 3D registration. Experiment results also show faster convergence speed.

1. Introduction

The aim of image registration is to establish spatial correspondences between two or more images of the same/or different scene acquired at different times, from different viewpoints, and/or by different sensors. Usually the ability to capture complex and large image deformations is vital to many computer vision applications including image registration and atlas construction. The problem becomes more challenging when the object in the image or edge of the image undergoes severe deformation [1].

Take medical image registration for example, tissues and organs or body itself are prone to deform, move, and rotate under most circumstances. Most methods iteratively reach a satisfying overlap under specific mathematical criterions, maximizing or minimizing deformation energy as described in (1). Fixed image is defined as F, while moving image as M. Registration aims to find the optimal model T that best satisfies energy S. As a result, model T, objective function S (similarity metric), and optimization method constitutes the three main components of image registration. Consider

$$S(F, M \circ T) + R(W). \tag{1}$$

According to a state-of-the-art survey [2], registration can be classified into rigid and nonrigid registration. Rigid models restrain the optimum to a few parameters to achieve global registration, while nonrigid models recover local deformation through physical model like elastic or viscous, or statistical model or support vector regression framework, and so forth. In order to fully overlap two images, researchers commonly adopt the two-step strategy, which contains initial registration and following iterative registration [3].

In the two-step strategy, registration firstly begins with a global affine transformation for initial global alignment, take state-of-the-art method FLIRT [4] and ELASTIX [5, 6] for example. Or fiducial markers are firstly detected through feature descriptors, for example, the SIFT method

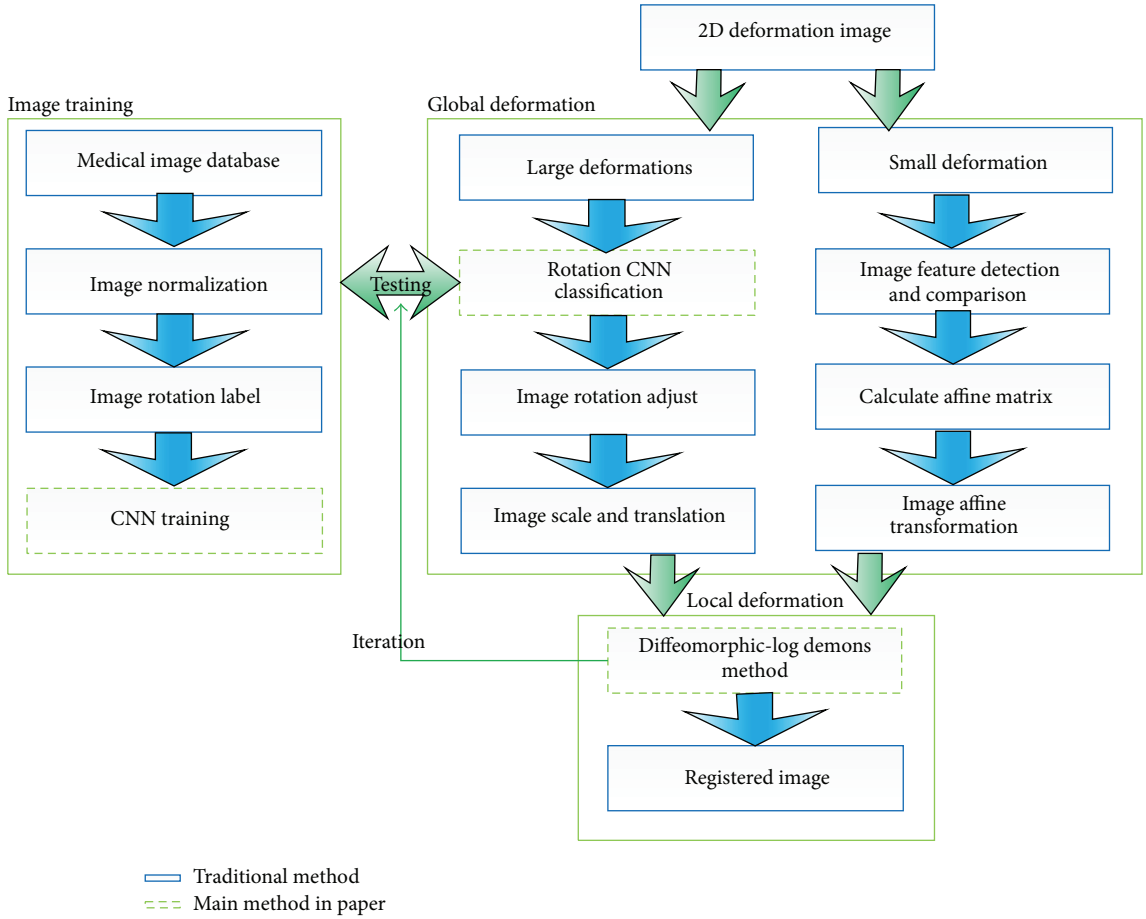


Figure 1: The work flow of 2D image registration.

[7], so the initial registration is carried out to establish correspondences between these point sets. In preregistration procedure, rotation, scale and translation of the moving image are modified by the calculated affine matrix. After that, nonrigid registration iteratively goes on. One severe problem of preregistration in affine matrix is that when large distortion and rotation both exists, accuracy is limited by correspondences between those region-based descriptors. If descriptor itself is not accurate, problem becomes more severe. Once descriptors fail to discover point correspondences, accuracy of following registration would be badly influenced. As a result, imprecision may also be introduced misleading the following procedure. Besides, traditional FLIRT and ELASTIX method declares that images for registration must be with the best quality, otherwise poor registration may occur.

In order to address these above limitations and capture very complex and large deformations, we proposed a new approach for image registration based on a two-layer deep adaptive registration framework. Firstly, in the preregistration procedure, rotation, scale and translation extent between two images are obtained separately to achieve initial registration. This is quite different from traditional "one time calculated" affine matrix. For rotation parameter, a CNN classifier is trained offline in order to identify the level of current image rotation under sever distortion. Then scale and translation parameters are obtained. An optimum preregistration is calculated relating to above gained parameters. As for 3D images, a triplanar 2D CNNs [8] around each voxel is utilized for calculation of final affine matrix. Until now, preregistration is done. Secondly, the rectified images are further recovered through the following nonrigid demons registration procedure. In the next circle, the former registration further facilitates results of the later registration of last iteration. This iterative procedure is carried out until an optimum overlap between the two images is achieved. Besides, PCA is introduced to extract the most valuable features, and detected features are put into SSD, Kendall, Pearson and Spearman, and so forth to form new similarity metric. Also, a triplanar 2D PCA is proposed to process 3D registration problem and Figure 1 gives details of the algorithm. As a result, convergence speed is accelerated while maintaining the same registration accuracy. Figures 1, 2 and Algorithm 3 illustrate work flow of our framework of processing 2D and 3D image registration.

The work introduced in this paper contributes in the following aspects:

(i) Preregistration is improved through estimation of rotation, scale, and translation separately. A multi-source CNNs is developed to precisely classify various levels of rotation under sever distortion and help identify rotation extent with high accuracy. For 3D

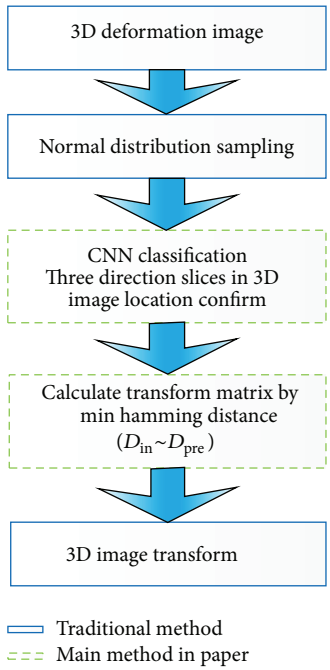


Figure 2: 3D image preregistration structure.

images, triplanar 2D CNNs is constructed to estimate parameters of affine matrix. This new preregistration performs better than state-of-the-art ELASTIX and SURF-based methods.

(ii) A two-layer adaptive registration framework is constructed and it performers better than other so-called two-step strategies.

(iii) PCA is used to extract valuable features and introduced into traditional similarity metric as SSD, Pearson, and so forth. For 3D images, triplanar 2D PCA is proposed to process 3D registration problem. Experiment results show that convergence speed is accelerated with the new similarity standard.

(iv) The proposed framework is tested under both synthetic and nature 2D and 3D images under various extent deformation. Experiment results show that our two-layer deep adaptive registration framework is able to identify the extent of rotation under sever deformation more precisely and correct large and complex distortions with high dice ratio than the comparative methods as it adaptively modify differences between images while others does not have any deep insight of deformation between images.

The rest of the paper is organized as follows. The whole architecture of the proposed two-layer adaptive registration framework for 2D and 3D images is illustrated in Section 2; Section 3 explains methodology of our CNNs classifier preregistration; Section 4 introduces our preregistration in combination with demons nonrigid registration and our new PCA related similarity metric; the proposed methods are evaluated in Section 5 under different datasets and evaluation principles; finally, the conclusion of this work is given in Section 6.

2. Architecture

2.1. CNNs for 2D Images. The whole workflow of our 2D image preregistration compared with traditional method is illustrated in Figure 1. In traditional algorithms, an affine matrix is calculated through correspondences between detected features, containing information of rotation, scale, and translation. This procedure is significantly influenced by accuracy of detected feature points. And under sever deformation images, traditional feature methods usually corrupt. Our algorithm processes each of the three above elements separately. By refining each procedure, accurate correspondences between fixed and moving image is obtained. It works as follows.

(i) For rotation, firstly the CNNs classifier is trained offline in order to rectify rotation extent of image under sever deformation. The trained CNNs classifier can identify as much as 360 classes of rotation.

(ii) For scale, image size information is utilized to achieve consistency between fixed and moving image.

(iii) For translation, centroid of each image is calculated through statistical algorithm and translation is achieved by utilizing position information of centroids.

2.2. Triplanar CNNs for 3D Images. Different from the 2D image preregistration, CNNs classifier here is used for the slice location of one voxel (xy, yz, xz three directions) instead of the rotation identifier. The work flow of 3D image preregistration is showed in Figure 2. The main procedure includes sampling, slices classification, transform matrix calculation, and image transformation by the matrix. Using CNNs on 3D image registration is a new attempt to resolve image registration for high deformation. Detailed method is introduced in Section 5.2.

3. Preregistration

Our strategy consists of firstly preregistration through CNNs classifier on both 2D and 3D images and then utilizing CNNs and demons algorithm adaptively in the following nonrigid registration and finally improving similarity metric for acceleration of registration convergence speed. In this section, we show our preregistration methodology by introducing our CNNs rotation classifier.

3.1. Why We Use CNNs

(1) The Robustness of Classification. CNNs are a kind of data based classification method which undertakes training by appropriate amount of data. CNN is suitable for nearly any types of data and can make classification with high accuracy, especially for the low quality of fMRI, CT images or images under high deformation (Experiment in Section 5 shows

these two kinds of real data are suitable for CNN processing method). Detailed CNNs structure and back propagation training method will be described in Section 3.5.

(2) Automatic Image Feature Perception. Nearly all kinds of preregistration method are based on precise feature perception so that the different feature perception methods are playing the key role in this procedure. Traditional image feature perception method is usually based on expert designated data feature. Usually experts give some fixed method to detect specific features of limited kinds of images. For example FLIRT method using inter-model voxel similarity measures where correlation ratio and mutual information are used to detect voxel relationships of different parts. This method has high limitation to the image sources, quality and variable settings. When exceptional case happens, some large deformation images are input for example, it will not work well. While the features from CNNs method are learned by network itself from training data such as edge, brightness, high or low frequency feature, distribution features and so on. Once the training data is updated, the network will get fit for more features automatically at the same time. Although long training time and complicated network variable learning makes CNN method not so easy to use, because of its high accuracy, it is still the image processing trend and future.

(3) High Efficiency Classification. Although the data training time of CNNs is long (depending on the detailed training method, network layer structure and hardware equipment like GPU), the total time spent on testing or classification is very short. Once the network is trained well, the only time consuming for processing is as short as linear operation.

Above all, even though there are some good affine transformation methods based on expert knowledge, we still need a smarter one to adapt to more complex image processing tasks in the future.

3.2. Theory of CNNs. The concept of deep learning was raised by Hinton and Salakhutdinov [9] in 2006, and it has brought great advances to machine learning since then. Deep learning aims to construct/use brain simulations to recognize data such as image/video, audio and text in an unsupervised way. Deep learning framework uses a multilayer "encoder" network to transform the high-dimensional data into a low-dimensional code and a similar "decoder" network to recover the data from the code. Outputs of low layer network acts as inputs of higher layer network. The whole network aims to equal inputs and outputs without loss of information. By using lower layer features to represent higher layer feature/classification, distributed feature representation of data is found. Auto encoder, Sparse coding, Restricted Boltzmann Machine (RBM), Deep Belief Networks (DBNs) and CNNs are five kinds of deep learning framework. Convolutional neural networks are excellent deep learning architectures, which were firstly introduced by Fukushima [10] and applied for handwritten digit recognition. Image recognition and segmentation tasks have also successfully used CNNs since then, with an error rate as low as 0.23 percent on the MINST database [11]. Besides, it is of high speed and accuracy for image classification in [12]. In facial recognition [13, 14] and video quality analysis [15], CNNs also gained large decrease in error rate and root mean square error.

A CNN is a multilayer perceptron consisting of multilayers, each layer with a convolutional layer followed by a subsampling layer. Through locally connected networks, stationary features of natural images are exploited by the network topology. Firstly, images are sampled into small patches. In the convolutional layers, small feature detectors are learned based on these extracted samples. Then, a feature is calculated by convolution of the feature detector and the image at that point. In the sampling layer, the number of features is reduced to reduce computational complexity and introduce invariance properties. One significant property of features learnt by CNNs is invariance to translation, rotation, scale and other deformations. This twice feature extraction structure enables CNNs with high distortion tolerance when identifying input samples.

3.3. CNNs Methodology. The goal of CNNs has no difference with other classification methods. They both focus on minimal total square error. Here we use c to denote the class number, and N to denote the training dataset, the total square error function can be shown:

$$E^N = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{c}\left(t_k^n - y_k^n\right). \tag{2}$$

Here t_k^n is the k-dimension of the n dataset, the y_k^n stands for the k output from the network, activation function in CNNs is sigmoid function for faster convergence rate. For each single dataset n, (2) can be describes as (3). The final aim of CNNs is to achieve smallest total square error between t_k^n and y^n. Consider

$$E^n = \frac{1}{2}\sum_{k=1}^{c}\left(t_k^n - y_k^n\right)^2 = \frac{1}{2}\left\|t^n - y^n\right\|_2^2. \tag{3}$$

For traditional full connection neural network, BP (Back propagation method) is used to calculate partial derivative to get the minimum square error, usually I the current layer, the output of I can be shows as (4), where f is sigmoid function. Consider

$$x^l = f\left(u^l\right), \quad \text{with } u^l = W^l x^{l-1} + b^l, \tag{4}$$

$$x_j^l = f\left(\sum_{i\in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right). \tag{5}$$

Unlike (4), as (5) shows, for the convolutional layer I, the image features (x) from prior layer is convoluted by kernel which is different in different layers, b_j^l is the offset of sigmoid function f. Consider

$$x_j^l = f\left(\beta_j^l \text{down}\left(x_j^{l-1}\right) + b_j^l\right). \tag{6}$$

For the sample layer, the image feature numbers and styles are the same with prior layer except the feature size is scaled

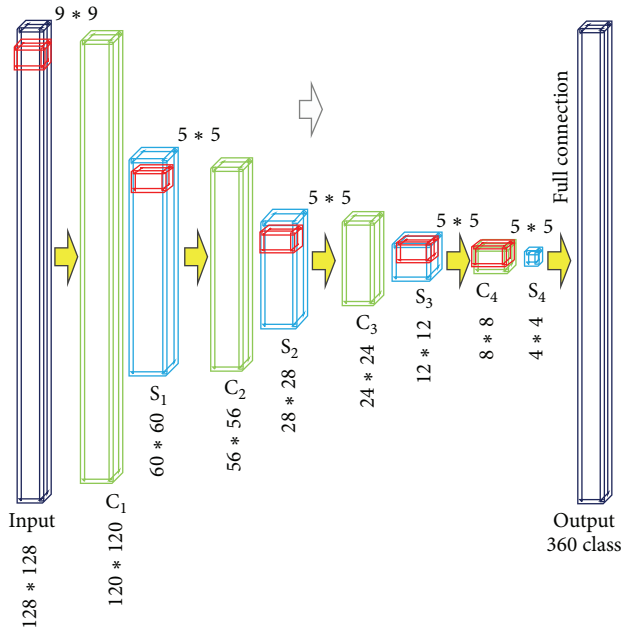


Figure 3: An illustration of CNNs.

down. Each feature contains a multi and addition kind offset. The down sample size in this paper is 2 which means the next layer image size is shrink two times by both weight and height. So through combination of (4) and (5), we can get sample equation (7) in which α_{ij} stands for the value of no. j output with no. i input features. By calculating α_{ij} and training kernels by back propagation method we can finally get the best features from different layers with high classification accuracy. Consider

$$x_j^l = f\left(\sum_{i=1}^{N_{in}} \alpha_{ij}\left(x_i^{l-1} * k_i^l\right) + b_j^l\right). \tag{7}$$

Constraint condition $\sum_i \alpha_{ij} = 1$, and $0 \le \alpha_{ij} \le 1$.

As shown in Figure 3, input images are defined as input layer; detailed introduction can be found in Sections 3.4 and 3.5. Hidden layer is the four pairs of convolutional and subsampling layer, which are denoted as $S_{l(l=1,2,3,4)}$, $C_{l(l=1,2,3,4)}$ and called local connection layer. The output layer is a combination of full connection layer and softmax classifier for classification. Each layer of $S_{l(l=1,2,3,4)}$ and $C_{l(l=1,2,3,4)}$ is constructed with multi-maps and each map is consisted of multi independent neural cells. Let $d^{(l-1)}$ and $d^{(l)}$ be the input and output for the lth layer, $S_I^{(l)} \times S_I^{(l)}$ and $S_O^{(l)} \times S_O^{(l)}$ be the size of the input and output map, $N_I^{(l)}$ and $N_O^{(l)}$ be the number of input and output maps respectively of that layer. According to CNNs, $N_I^{(l)} = N_O^{(l-1)}$ $S_I^{(l)} = S_O^{(l-1)}$.

3.4. CNNs Structure Design. We adopt a ten-layer CNNs perceptron network (input and output layers are included; convolutional and sample layers are separately calculated). Key variables setting including kernel size and sample rate of different layers in proposed CNN is showed in Table 1 and Figure 3. Learning rate alpha = 1, variable update batch size = 10, iteration times = 1000, any training and test images are normalized to 128 * 128 size gray images with [0, 1] pixel size.

Table 1: Key setting variables in CNN network.

Layer	Name	Kernel size or sample rate
1	Input layer	None
2	1st convolutional layer	9 * 9
3	1st sample layer	2
4	2nd convolution layer	5 * 5
5	2nd sample layer	2
6	3rd convolutional layer	5 * 5
7	3rd sample layer	2
8	4th convolution layer	5 * 5
9	4th sample layer	2
10	Output layer	None

3.5. Training Image Rotation Classifier through CNNs. Our input images for training are difference images between fixed and moving image: $F - M$. M is under deformation with different extent of rotation. Each rotation angle of 360° is defined as one class, producing as much as 360 classes. Two distinguishing characters of CNNs are perception field and shared weights. Perception field means each neural cell in each layer is not connected wtih all neural cells in adjacent layers, but limited to a local area of neural cells (9 * 9 as in Figure 3). Shared weights means the connection weight parameters (9 * 9) of every neural cell to the local area cell are the same. As shown in Figure 3, suppose size of input image T is 128 * 128. After convolution with filters, the kernel size of which is 9 * 9, image changes into Ts1 of 120 * 120 size. Image then scales into Tc1 60 * 60 in layer S1. After four pairs of S and C, the original image is represented as Ts4 of only 4 * 4 matrix. In this hidden layer, all neural cells on feature maps are not all connected but with same weights. As a result, only 9*9 weight parameters need to be calculated, greatly reducing computation complexity. An all connection exists between Ts4 matrix and output layer, eliminating disparity caused by partial connection in the hidden layer. Then softmax classifier identifies the matrix and outputs the detected results. After that, the parameters are fine-tuned through back propagation of 1000 times until convergence. After all these steps, a finite classifier is obtained.

4. Two-Layer Deep Iterative Registration Framework

4.1. Diffeomorphic Log Demons Registration. In the 19th century, Maxwell firstly introduced the concept of demons to illustrate a paradox of thermodynamics. In 1998, Thirion [16] proposed a registration algorithm under demons model, which had a high registration precision and efficiency through pixel velocities caused by edge based forces.

(i) Theory and Improvements of Demons Registration. Demons registration utilizes optical flow equation as basis forces for finding tiny deformations in temporal image sequences. For

point p in space, let f and m be intensity values in fixed image F and moving image M respectively. According to Thirion's theory, (8) shows calculation of velocity u allowing point p to match the corresponding point in M. Here, ∇f called internal edge force is the gradient image of fixed image and $(m - f)$ called the external force. In order to make the equation more stable and appropriate for image registration, Thirion added term $(m - f)^2$. Later on, He Wang et al. added image forces of the moving image in the equation to improve convergence speed and stability of the registration as shown in (9). Parameter α was proposed by Cachier et al. to adjust force strength. Consider

$$u = \frac{(m - f)\,\Delta f}{|\Delta f|^2 + (m - f)^2}, \tag{8}$$

$$u = \frac{(m - f)\,\Delta f}{|\Delta f|^2 + \alpha^2 (m - f)^2} + \frac{(m - f)\,\Delta m}{|\Delta m|^2 + \alpha^2 (m - f)^2}. \tag{9}$$

Vercauteren et al. [17] proposed nonparametric diffeomorphic demons algorithm. It considers the demons algorithm as a procedure of optimization on the whole space of velocity fields and adapts that procedure in a space of diffeomorphic transformations. The transformation result is smoother and more accurate. Then Vercauteren et al. [18] brings the process into log-domain, that is, he uses a stationary velocity field. Besides, the algorithm is symmetric with respect to the order of the input images. Lorenzi et al. [19] implements a symmetric local correlation coefficient to log-demons diffeomorphic algorithm. Lombaert et al. [1] proposed spectral log-demons to capture large deformations. Peyrat et al. [20] implements multichannel demons to register 4D time-series cardiac images.

(ii) Diffeomorphic Log Demons Algorithm. Here, diffeomorphic log demons algorithm is briefly reminded. A diffeomorphic transformation ϕ is related to the exponential map of the velocity field $v : \phi = \exp(v)$ (Algorithm 1) [1]. The log-demons framework alternates between optimization of a similarity metric updated by Euler-Lagrangian function in (10). In general, procedure of diffeomorphic log demons framework is described in Algorithm 2. Consider

$$\mathrm{Sim}\left(F, M \circ \exp(v)\right) = \left\| I_F - I_{M \circ \exp(v)} \right\|^2. \tag{10}$$

4.2. New Similarity Metric by Combination of PCA. Mathematically, PCA is defined as an orthogonal linear transformation that transforms the data to a new coordinate system to extract the greatest variance in the data set. As a result, it is able to avoid influences caused by image biases. Traditionally, PCA is used for dimensionality reduction to facilitate classification, visualization, communication, and storage of high-dimensional data. Here, PCA is applied in both 2D and 3D medical and usual images, and the detected feature representations are used as inputs of similarity metric to achieve anatomical correspondence and assist optimization procedure in registration.

There are many classical metric measures, such as SSD, mutual information (MI), cross correlation (CC), pattern intensity and also their corresponding improved edition. In this paper, Pearson, Spearman, kendall, SSD together with extracted features by PCA are utilized as the new similarity metric. Pearson, Spearman and Kendall are concepts in statistics and are frequently used in data mining. Pearson is short for Pearson product-moment correlation coefficient (PPMCC), which was developed to measure the linear correlation between two variables. Spearman's rank correlation coefficient is a nonparametric measure of statistical dependence between two variables. Both of their value is between +1 and −1. Spearman has no requirement on variables, while pearson insists variables meets normal distribution. Our utilization of log demons registration avoids the influence brought by this.

Input: Velocity field v.
Output: Diffeomorphic map $\phi = \exp(v)$.
(i) Choose N such that $2^{-N}v$ is close to 0
e.g., such that $\max \|2^{-N}v\| \leq 0.5$ pixels
(ii) Scale velocity field $\phi \leftarrow 2^{-N}v$.
for N times **do**
(iii) Square $\phi \leftarrow \phi \circ \phi$.
end for

Algorithm 1: Exponential $\phi = \exp(v)$ [1].

(i) For 2D images of size $m \times n$, firstly, PCA is applied to both fixed image F and registered moving image M, gaining pca_F and pca_M. Thus, most important information of image can be fully utilized by combination of pca_F and pca_M as inputs of pearson, spearman, and so forth, forming new similarity metric.

(ii) For 3D images of size $m \times n \times k$, firstly, PCA is applied to every slice of x axis and gains a series of pca_{x_i} $(i = 1, 2, \ldots, m)$. By summarizing each of pca_{x_i} $(i = 1, 2, \ldots, m)$, pca_x is obtained. The same operation is carried out on y and z axis data, obtaining pca_y, pca_z. Then, PCA of both fixed (f_x, f_y, f_z) and registered moving (m_x, m_y, m_z) image is calculated. Thus, information of image can be fully utilized by combination of (f_x, f_y, f_z) and (m_x, m_y, m_z) as inputs of PPMCC, Spearman, and so forth. Workflow of this part is shown in Figure 4.

4.3. Two-Layer Iterative Registration Framework. Traditionally, the two step registration means an initial affine registration in the very beginning to coarsely rectify deformation and a following iterative registration to optimize a similarity metric achieving fine registration. We also adopt the two step strategy. But before the two step registration, we build up a classifier offline under CNNs training to identify rotation between fixed F and moving M image under very large distortions, then scale and translation. Also in each iteration, the initial and following registration are carried out iteratively. This feed-back procedure assists achieving higher

Registration:
Input: Images F, M and initial velocity field v
Output: Transformation $\phi = \exp(v)$ from F to M
Pre-registration: through SURF related affine transformation or ELASTIX to globally register F and M
Repeat
Demons registration:
(i) Find updates $u_{F \to M}$ mapping F to $M \circ \exp(v)$
(ii) Smooth updates: $u \leftarrow K_{\text{fluid}} * u$
(iii) Update velocity field: $v \leftarrow \log(\exp(v) \circ \exp(u))$
(approximated with $v \leftarrow v + u$)
(iv) Smooth velocity field: $v \leftarrow K_{\text{fluid}} * v$
(v) **Calculate traditional similarity metric** SSD.
until convergence

Algorithm 2: SURF/ELSTIX related registration framework.

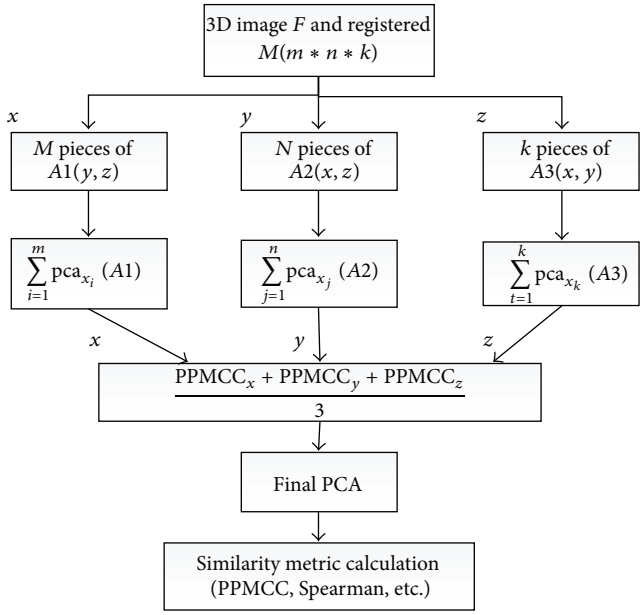


Figure 4: Calculation procedure of 3D PCA-related similarity metric.

registration accuracy comparing with traditional SURF and affine method.

Besides, at the end of each iteration, we utilized a new similarity metric by combining PCA with traditional SSD, pearson, and so forth, fully containing most important features of image. As a result, convergence speed is highly accelerated than traditional SSD without PCA while maintaining the same registration accuracy. Algorithm 3 shows the over flow of the framework.

5. Experiment Results

In this section, the performance of the whole two-layer registration method is evaluated on both 2D and 3D images, synthetic and nature datasets. For comparison, traditional two step methods, ELASTIX and SURF related algorithm are used to preregistrate moving and fixed image. Then demons nonrigid registration is conducted. These methods are set as the baseline methods, which are denoted as ELASTIX+demons and SURF+demons. They all firstly use detected features initially to register images through affine transformation and original SSD as similarity metric under the diffeomorphic log demons framework. Our method is different from their framework both in preregistration and following nonrigid registration framework. For 2D images, firstly train a rotation classifier through CNNs and preregister moving image under large distortion and rotation, then together with scale and translation transformation, preregistration is done. For 3D images, a pretrained triplanar 2D CNNs is utilized to locate voxels, establishing feature correspondences. Finally, PCA related similarity metric iteratively registering images under diffeomorphic log demons framework.

The improvement of our two-layer method in registration accuracy, robustness to large deformation and rotation, and convergence speed are all assessed with ground truth data. Our matlab code is under Lombaert's work [21] and Toolbox [22].

Specifically, we downloaded brain and lung dataset from BrainWeb MRI Simulated Normal Brain Database [23, 24] and Empire 10 challenge [25] for training and testing. Besides, training and tests were also carried out on classical Lena image, which is mostly used for image processing. T1, T2 image from [26] and ITK image [27] is also utilized in our experiment. BrainWeb MRI dataset contains 20 both 2D and 3D new normal anatomical models. Empire 10 challenge lung dataset includes 20 3D scans, each containing fixed and moving image pair. Description of lung dataset is shown in Table 2 [28].

5.1. 2D Image Registration

5.1.1. Synthetic Deformation Tests. A lot of registration has been evaluated on synthetic deformation images for algorithm test according to previous work [1, 26].

(1) Training CNNs Deformation Classifier. In this work, ten-layer CNNs are constructed to train various sources of

Classifier Training: Eight-layer CNNs trains different levels of rotation and deformation images
Registration:
Input: Images F, M and initial velocity field v
Output: Transformation $\phi = \exp(v)$ from F to M
Repeat
Pre-registration: For 2D image, through CNNs classifier to rectify rotation then scale and transformation;
For 3D image, triplanar 2D CNNs to pre-registrate
Demons registration:
(i) Find updates $u_{F \to M}$ mapping F to $M \circ \exp(v)$
(ii) Smooth updates: $u \leftarrow K_{\text{fluid}} * u$
(iii) Update velocity field: $v \leftarrow \log(\exp(v) \circ \exp(u))$
(approximated with $v \leftarrow v + u$)
(iv) Smooth velocity field: $v \leftarrow K_{\text{fluid}} * v$
(v) **Use PCA to extract features and calculate new similarity metric** as PCA-SSD, PCA-pearson, PCA-spearman, and kendall.
until convergence

Algorithm 3: Two-layer unsupervised deep adaptive registration framework.

Table 2: Listing of data used in lung registration [28].

Pair	Data category	Pair	Data category	Pair	Data category
1	Insp-Exp	11	Insp-Insp	21	Insp-Exp
2	Insp-Insp	12	Warped	22	Insp-Insp
3	Insp-Insp	13	4D	23	4D
4	Ovine	14	Insp-Exp	24	Ovine
5	Warped	15	Insp-Insp	25	Warped
6	Contrast	16	4D	26	Contrast
7	Insp-Exp	17	4D	27	Insp-Insp
8	Insp-Exp	18	Insp-Exp	28	Insp-Exp
9	Insp-Insp	19	Insp-Insp	29	Ovine
10	Ovine	20	Insp-Exp	30	Warped

images. We also tested other number layer CNNs, results showed that ten-layer CNNs achieved highest score when classifying rotation of deformed images. Four kinds of 2D source images [24–27] served as samples. An example of sample image is shown in Figure 5. Take image T_1 for example, linear transformation like rotation or translation is added to image T_1 by multipling rotation matrix r coded through matlab; then four kinds of large and complex nonlinear transformation is added to T_1 through special processing by photoshop. T_1 image with only rotation is denoted as $T_1 \circ r$, with only deformation noted as T_1', and with both rotation and deformation denoted as $T_1' \circ r$. The same notation is with F_1 and Lena image. Figure 6 is an illustration after all those processing. In order for accurate identification of rotation, here for training, difference image of F and M (with only rotation) $D_{T_1} = \|T_1 - T_1 \circ r\|$ is input of CNNs. After training, each angle of 360° is defined as one class, obtaining 360 class of distortion. For other CNNs, number of classes is 180, 90, 36.

Our test is carried out on computer of windows 7 system, with 8 GB RAM, i7-4770 CPU @ 3.4 GHz. Take BrainWeb data [23, 24], for example, Table 3 shows test results of the classifier according to these data.

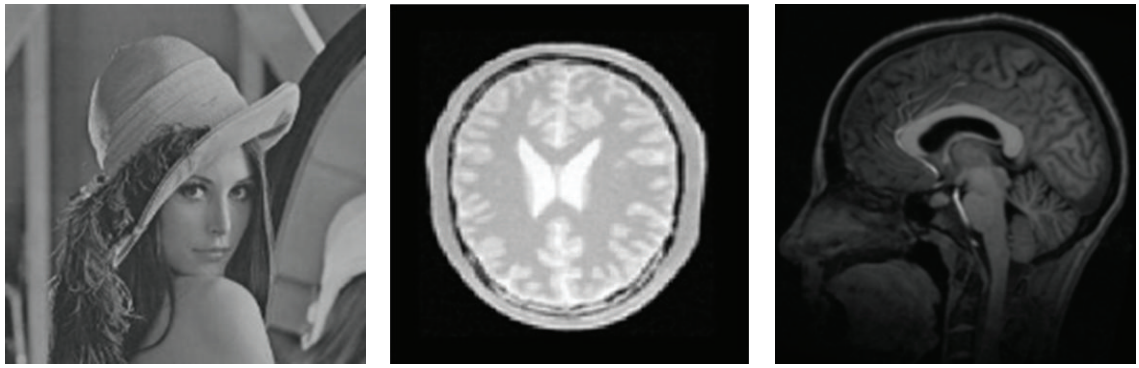
As we can see from Table 3, when input images are resized into 64×64 pixels, identification of rotation can reach as much as 99.86% for classifier 90; while images are resized into 28×28 pixels, the identification accuracy for classifier 36 is 99.97%. All these are done under condition that training data is also for testing. When the testing data of BrainWeb is put into the trained classifier, accuracy reaches 99.56%, even lower than the training data itself, but still very high according to many usual classifiers. For Lena, ITK and T1 training data, classifier 36 gains an accuracy rate of 99.94%. Number of iteration is set to 1000 for every training.

(2) CNNs Preregistration Test. SURF related method, ELASTIX and our CNNs method are tested. Here, SURF related method means using firstly SURF algorithm to detect features and then affine transformation to initially register images.

(i) When only rotation exists as Lena $\circ$ r in Figure 7, ELASTIX method failed; SURF method is able to identify rotation invariant features and establish accurate correspondences between Lena and Lena $\circ$ r. Established correspondences are shown as $\text{Lena}_{\text{corr}}$ in Figure 7. Vertices of the yellow lines stand for feature

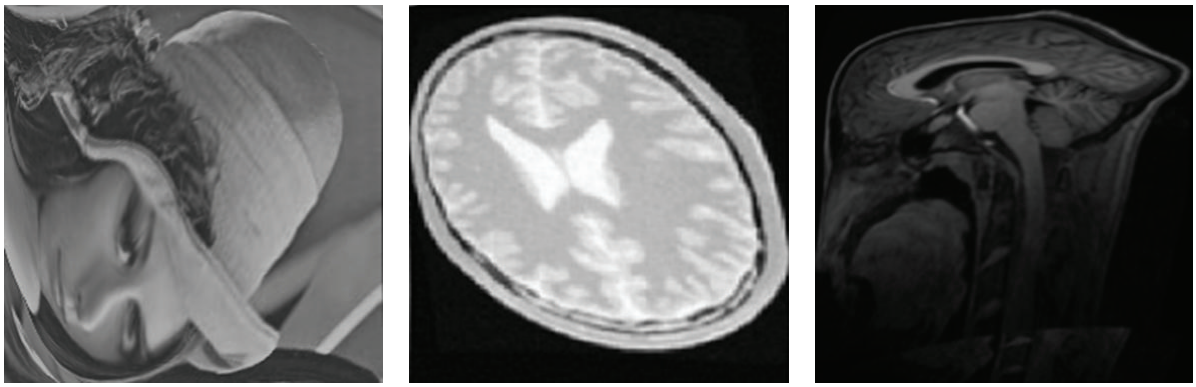
Table 3: Performance of classifier.

	Image size	Accuracy of classifier 36	Accuracy of classifier 90	Time (s each iteration)
BrainWeb Training data	64 × 64	※	99.86%	41.2
	28 × 28	99.97%	※	6.17
BrainWeb Testing	28 × 28	99.56%	※	※
Lena, ITK, T1	28 × 28	99.94%	※	2.4



Lena F_1 T_1

Figure 5: An example of original sample image.



$\text{Lena}' \circ r'$ $F_1' \circ r$ $T_1' \circ r$

Figure 6: An illustration of sample image after sever distortion and large rotation.

Lena $\text{Lena} \circ r$ $\text{Lena}_{\text{corr}}$ $\text{Lena}_{\text{SURF}}$ $\text{Lena}_{\text{ELASTIX}}$

Figure 7: Preregistration result of ELASTIX and SURF method with only rotation on image.

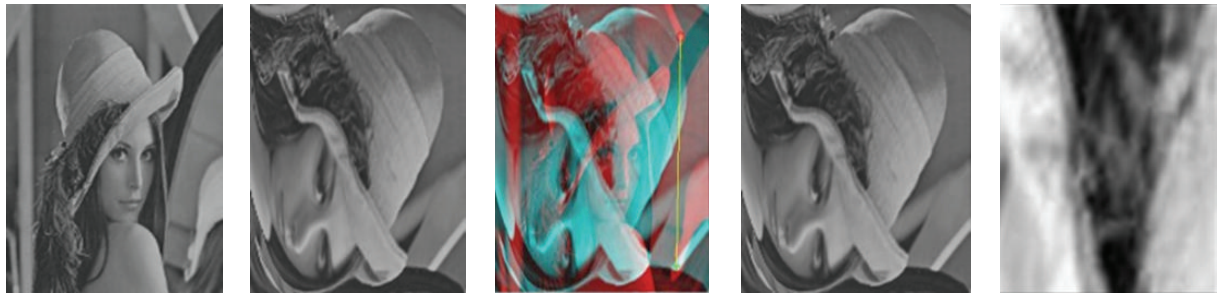


Figure 8: Preregistration result of ELASTIX and SURF method with both rotation and large deformation on image.

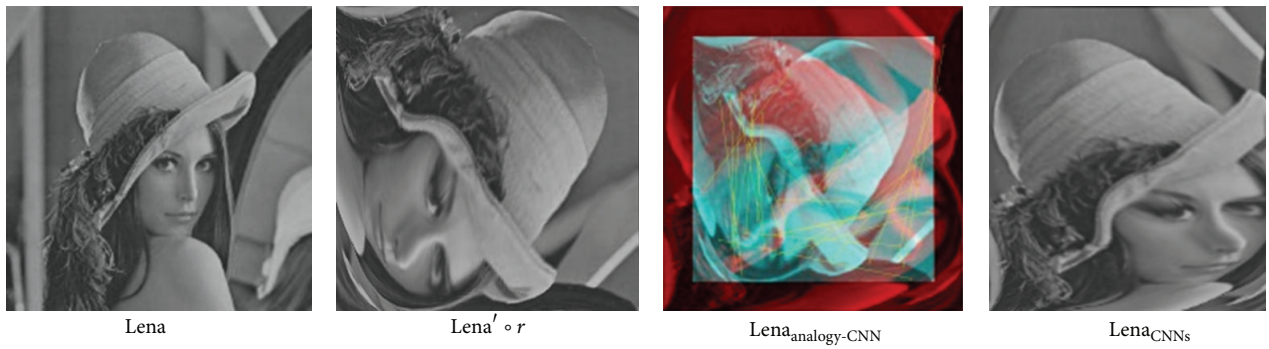


Figure 9: Preregistration result of CNNs method with both rotation and large deformation on image.

correspondences. Red circle vertex stands for feature in original moving image Lena $\circ r$ while green cross stands for corresponding feature point in registered moving image $\text{Lena}_{\text{SURF}}$. Registered image is denoted as $\text{Lena}_{\text{ELASTIX}}$ and $\text{Lena}_{\text{SURF}}$.

(ii) However, when rotation and large deformation simultaneously appears in moving image as Lena$' \circ r$ in Figure 8, both ELASTIX and SURF method crushed. Under such circumstances, in our tests SURF only found one pair of correspondence points. As there are not enough feature correspondences, initial registration failed.

(iii) On the contrary, our trained CNNs classifier and following scale and translation operation directly identified Lena image's rotation angle accurately (90° rotation), and turned it back to Lena$'$ as in Figure 9. For better comparison, we used software to show ways of rotation processing in CNNs as SURF's manner, feature detecting and matching in $\text{Lena}_{\text{analogy-CNN}}$. As enough number of so-called features are detected, CNNs is able to recover rotation added to Lena$'$.

(3) Accuracy Evaluation of Registration. Mathematically, dice ratio is used to evaluate overlap between two datasets. It is defined in (11). In this section, both dice ratio and subjective human evaluation method is used to assess accuracy of ELASTIX and SURF related registration and our method result

$$O_{\text{overlap}} = \frac{2F \cap M}{F + M}. \tag{11}$$

After preregistration in Section 5.1.1, ELASTIX and SURF related method performs diffeomorphic log demons algorithm iteratively to achieve for best registration; while our method iteratively carries out CNNs classifier and diffeomorphic log demons algorithm to optimize registration. This new two-layer registration framework makes full use of both preregistration and following demons method and registration results show that it indeed improves accuracy.

Figure 10 shows registration procedure and result of ELASTIX (Figure 10(c)) and SURF (Figure 10(b)) related method, while Figure 11 shows that of our method. When both rotation and deformation exists in image F_1, our registration result $F_1 - F_{1\,\text{C+demons}}$ is much better than $F_{1\,\text{ELASTIX+demons}}$ and $F_{1\,\text{SURF+demons}}$ apparently. Besides, to test dice ratio of registration, original fixed image F_1 and registered moving image of two methods are put into function (11) separately. Dice ratio of ELASTIX and SURF-demons method is 0.889 and 0.88, while our CNNs-demons-iterative method achieves as much as 0.8964.

5.1.2. Lung Atlases. Description of lung dataset can be found in Table 2 [28]. Empire 10 lung datasets are firstly used for the MICCAI conference 2010. It contains 20 intra-patient thoracitic CT image pairs. Figures 12 and 13 shows our

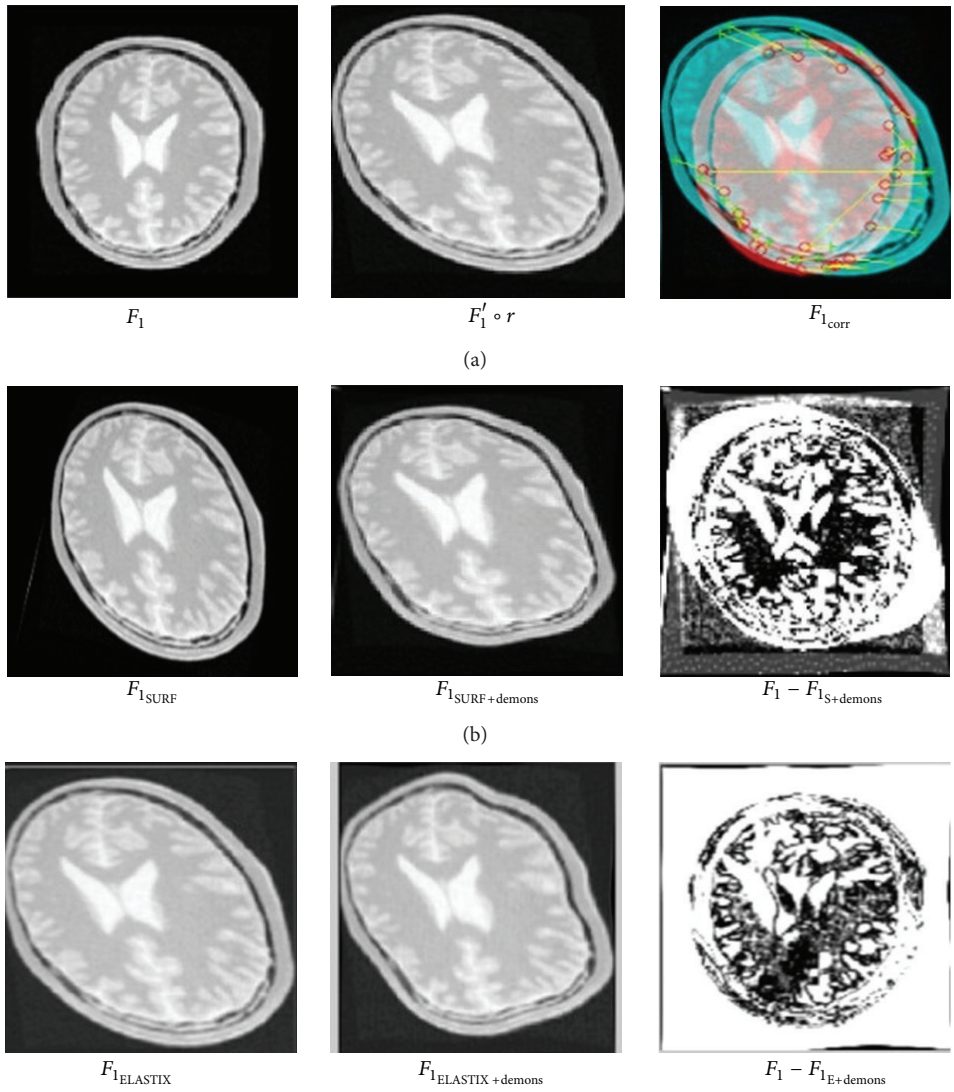


Figure 10: Preregistration result of SURF related method (b) and ELASTIX method (c) with both rotation and large deformation on image.

preregistration results of slice image 8 and 6 of 4D image pairs compared with that of Elastix tool. All images are shown with the help of tool vv [29]. From left to right in the figures are fixed image, moving image, preregistration result of ELASTIX and our proposed CNNS method, final registration result of above two methods with demons. It is obvious that our method can accurately rectify rotation, scale and translation deformation added to moving image. While ELASTIX preregistration failed to rectify rotation differences between fixed and moving images. Registered images are denoted as Elastix-demons and CNNs-demons. Diff images between fixed and registered image are denoted as Elastix-diff and CNNs-diff.

ELASTIX preregistration consuming time of each slice is shown in Figure 14. The shortest time of one slice is more than 1000 ms (1 s) and time for slice 8 is 3500 ms. Although training of our CNNs classifier costs long time, it is offline. And our CNNs rotation, scale and translation operation costs a total of only 39 ms. As a result, it is quite attractive for real-time clinical applications.

5.1.3. Brain Atlases. We select the cross section 2D image of the BrainWeb MRI 20 object, 10 for training and the other 10 for testing. From Figure 15, we can see that our proposed preregistration can rectify both rotation and translation more successfully than traditional Elastix affine registration.

5.2. An Attempt on 3D Image Registration by Using CNNs. For the 3D image registration part we focus on the brain atlases registration and give a CNNs 3D image registration method. We train brain atlas from 18 people's 3D image data in BrainWeb Brain database by four steps: (1) Randomly select 10 label points by Normal distribution in 3D image. (2) Adjust 3D brain image and separate it to 2D image on three directions (xy, yz, xz). (3) Test each 2D slice position by triplanar after-trained CNNs classifier (each dimension enjoys one CNN

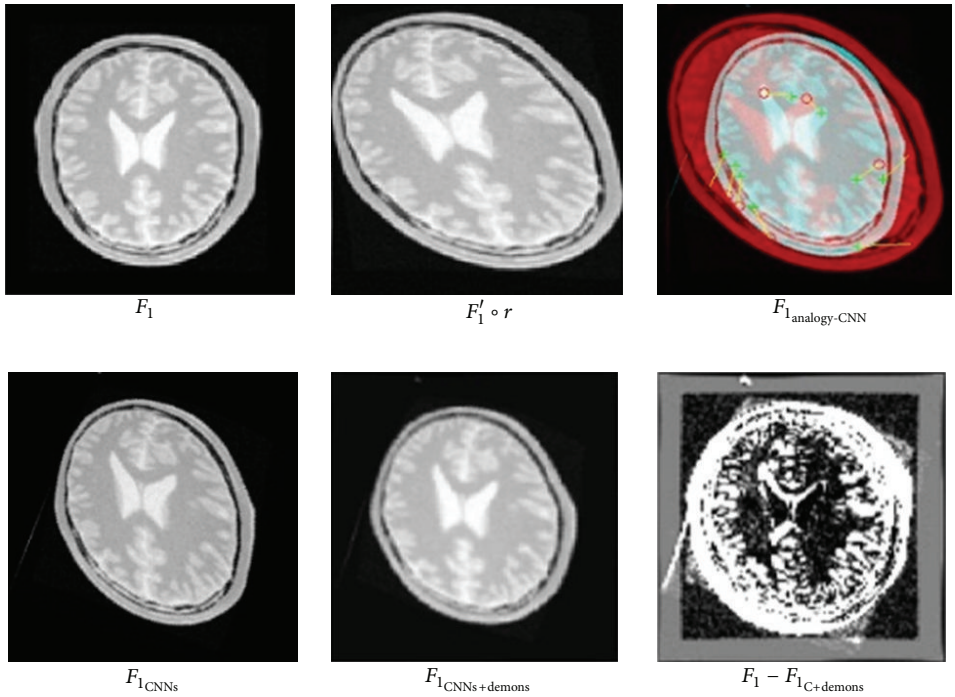


Figure 11: Preregistration result of CNN method with both rotation and large deformation on image.

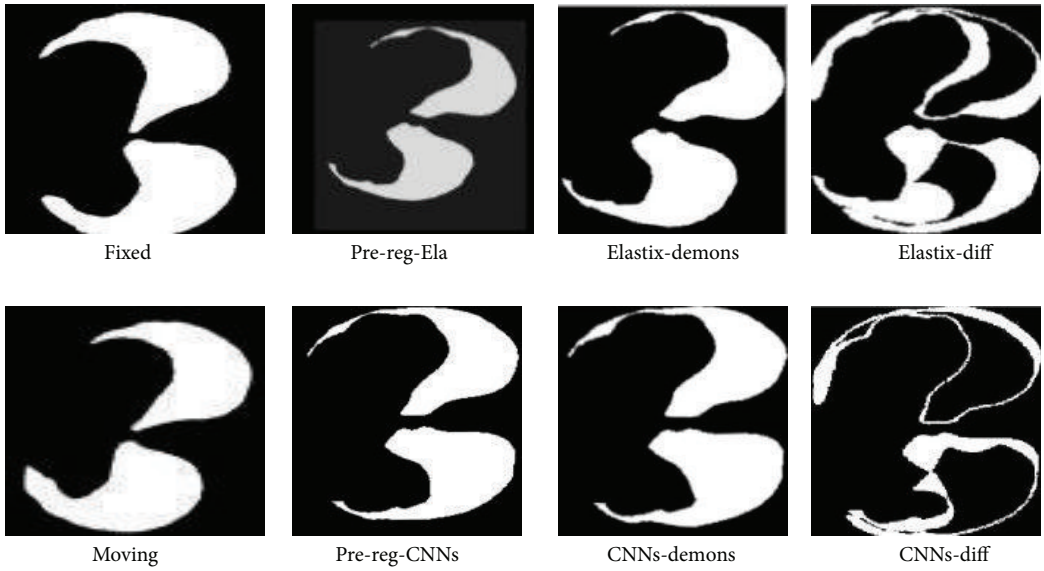


Figure 12: Lung slice 8 registration.

network) and get the right slice position (predicted voxels). (4) Adjusting the 3D image to make label voxels and predict voxels that enjoy smallest hamming distance. Experiments shows the high accuracy CNNs classify results will greatly improve moving 3D image's similarity to the fixed 3D image. The detailed procedure is shown in Figure 16.

5.3. Convergence Speed Evaluation of Registration. Sections 5.1 and 5.2 improve both registration accuracy and speed. In this section, we test registration accuracy on T2 brain medical data and focus on accelerating convergence speed of registration. PCA is introduced to extract valuable features and by combining features with SSD, Pearson, Spearman, Kendall, we get new similarity PCA-SSD, PCA-pearson, PCA-spearson and Kendall. Original SSD is denoted as energy in Figure 17(e). The course-to-fine (in here, three level is recommended) registration strategy is adopted in here. In Figure 17, horizontal axis stands for iteration times and vertical axis stands for the values of metric. Firstly, mean convergence extent of the three-level registration is

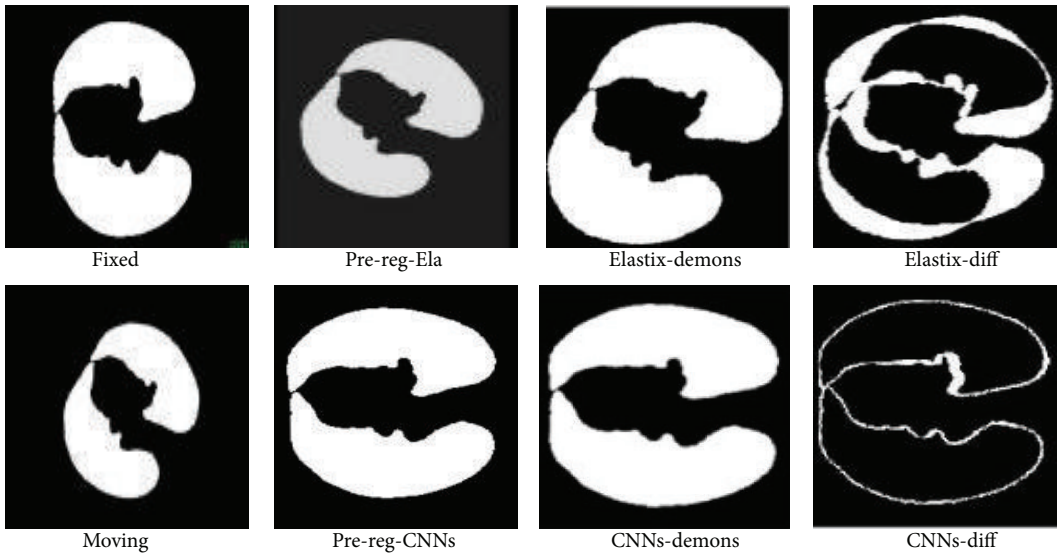


Figure 13: Lung slice 6 registration.

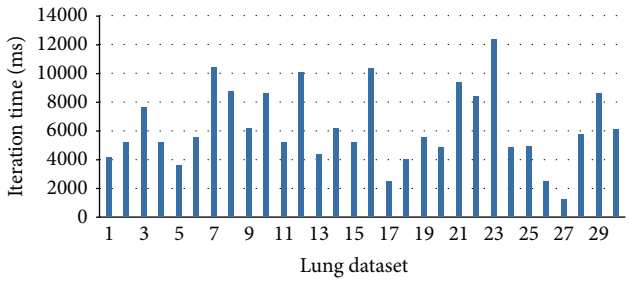


Figure 14: ELASTIX preregistration consuming time of the 30 slices of 4D lung dataset.

calculated. Then normalization is carried out on the mean value. Several conclusions can be gained:

(1) both PCA related and original SSD methods converge regularly,

(2) as a whole, PCA-SSD and PCA-Pearson methods perform best and converge faster than original SSD metric;

(3) PCA-spearman metric firstly converges fastest, but latterly it slows down;

(4) Kendall metric performs worst compared with other metrics.

6. Conclusion

In this paper, a comprehensive method of constructing rotation classifier for images under severe deformation and rotation was proposed through CNNs. The classifier is able to identify distortion as much as 360 classes according to analysis of rotation angles. The classifier is utilized to assist our proposed two-layer deep adaptive registration framework. In each registration iteration, preregistration with identification of the trained classifier, scale, and translation operator and following diffeomorphic log demons registration facilitates each other one after another. Besides, proposed PCA related similarity metric helps achieve faster convergence speed. The new two-layer registration framework is compared with traditional diffeomorphic log demons registration in combination with state-of-the-art ELASTIX and SURF preregistration. As baseline method carries out preregistration only once, large deformations cannot be fully modified. From tests on different image resources containing various kinds of both 2D and 3D, MRI, and CT datasets, our framework indeed outperforms the baseline method on both registration quality and convergence speed.

In the following work, we would combine other kinds of deep learning framework as independent subspace analysis (ISA) [30], sparse coding [31], and so forth to improve current registration. Also, more performance tests of the proposed two-layer registration framework should be carried out on more data resources. Besides, the proposed method performance should be compared with other deep learning models.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This paper is supported by The Project for the National Key Technology R&D Program under Grant no. 2011BAC12B03, The Innovation Team of Beijing, The National Natural Science Foundation of China under Grant nos. 81370038 and 61100131, The Beijing Natural Science Foundation under Grant no. 7142012, The Beijing Nova Program under Grant no. Z141101001814107, The Science and Technology Project of

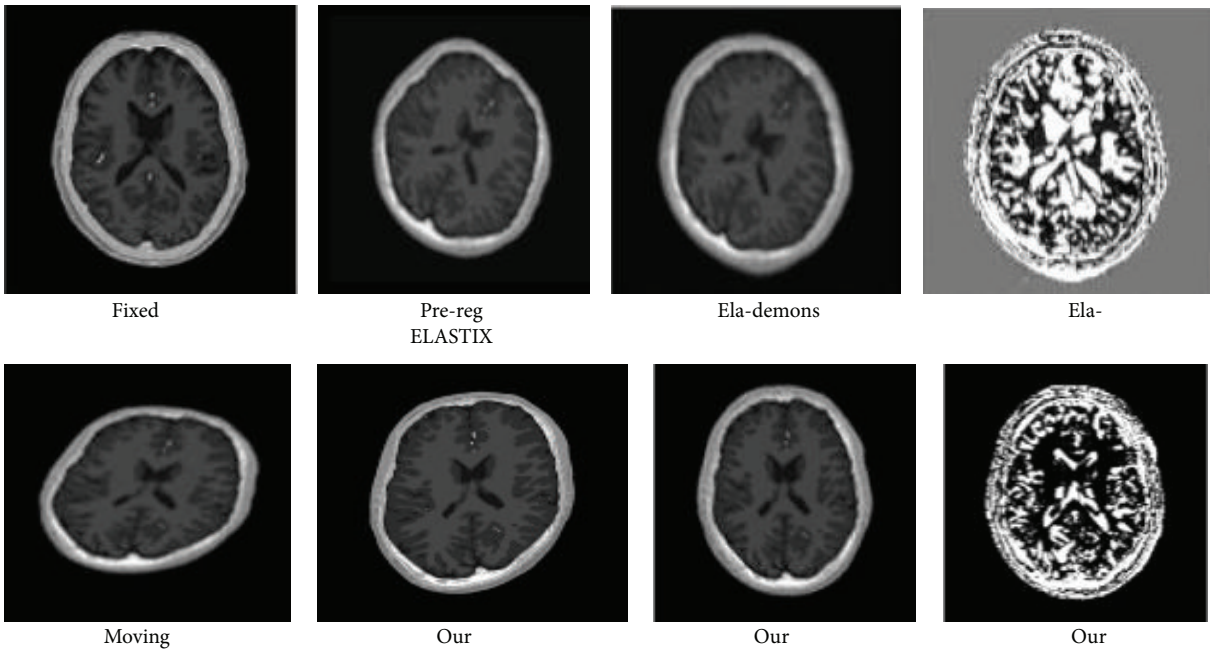


Figure 15: Brain slice registration.

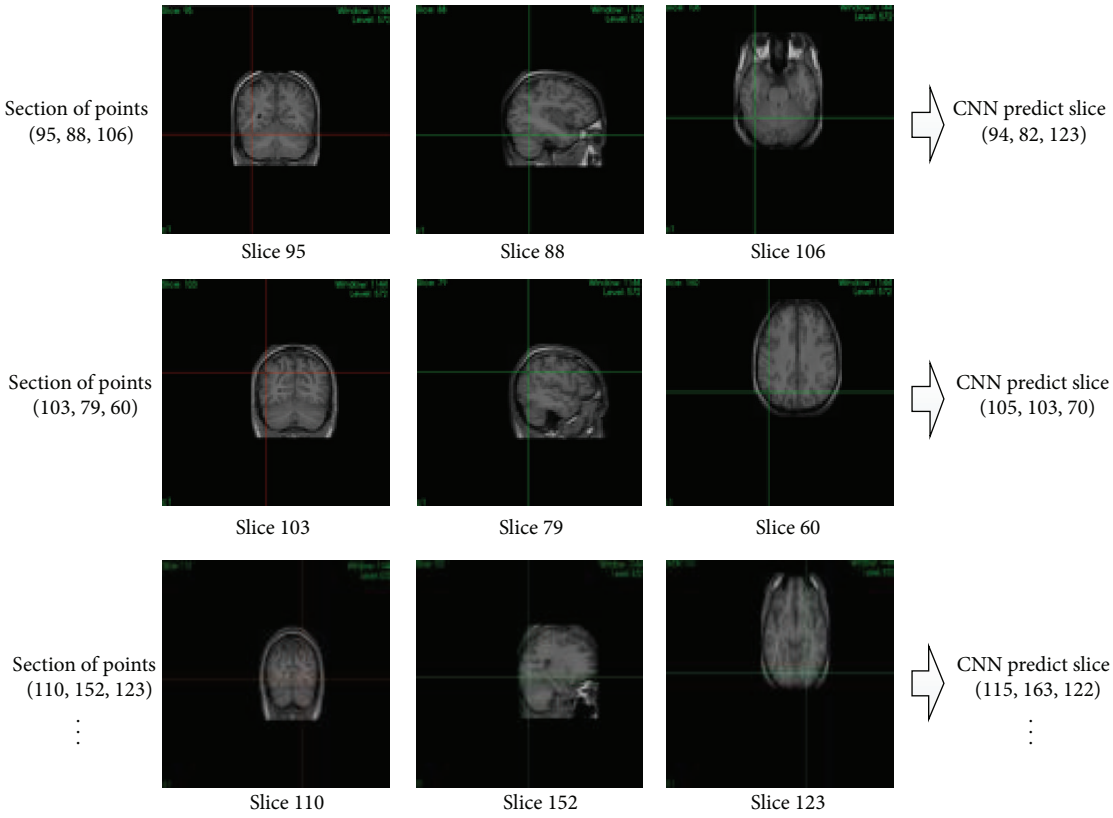
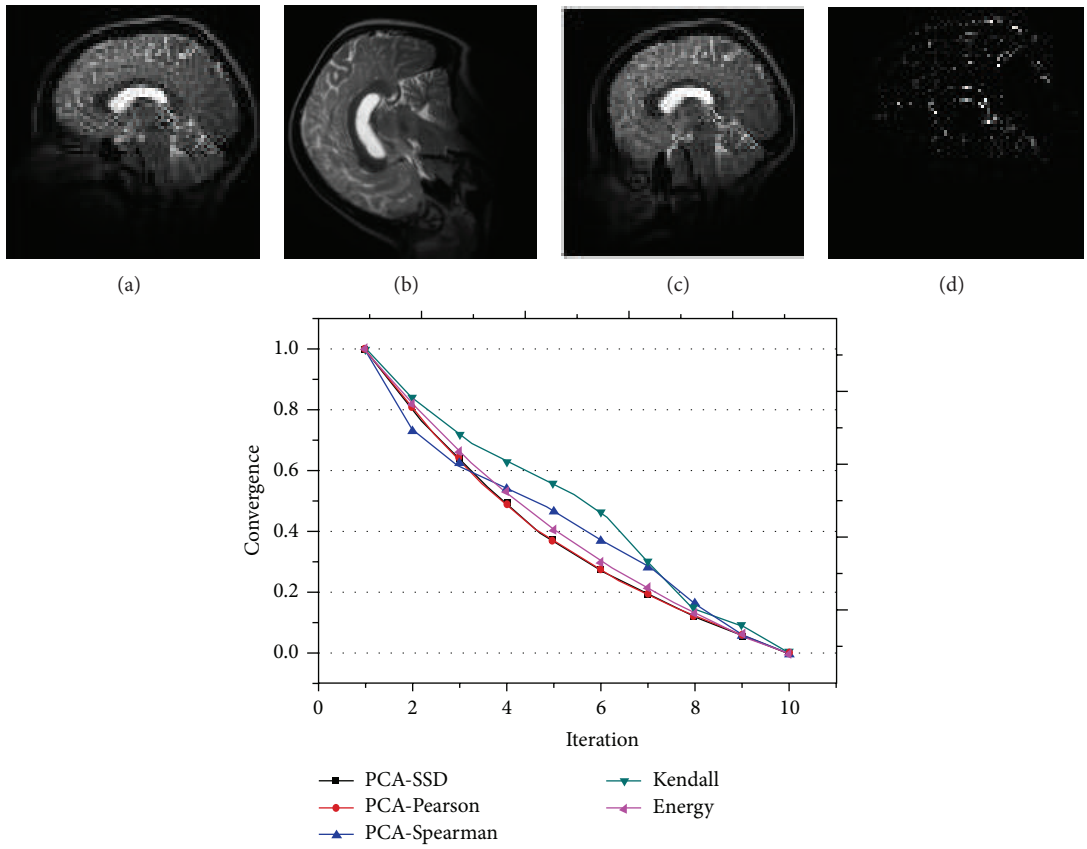


Figure 16: 3D Sample voxel slice images (three slices, xy, yz, xz) classification.



(e) Convergence extent

Figure 17: T2 Data: (a) fixed image, (b) moving image, (c) registered moving image, (d) difference between (a) and (c), (e) convergence extent of the first ten iteration.

Beijing Municipal Education Commission under Grant nos. KZ201310005004 and km201410005003, The Rixin Fund of Beijing University of Technology under Grant nos. 2013-RX-L04 and 2012-RX-03, and The Basic Research Fund of Beijing University of Technology under Grant no. 002000514312015.

References

[1] H. Lombaert, L. Grady, X. Pennec, N. Ayache, and F. Cheriet, "Spectral log-demons: diffeomorphic image registration with very large deformations," *International Journal of Computer Vision*, vol. 107, no. 3, pp. 254–271, 2014.

[2] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, "A review of 3D/2D registration methods for image-guided interventions," *Medical Image Analysis*, vol. 16, no. 3, pp. 642–661, 2012.

[3] G. Song, N. Tustison, B. Avants, and J. Gee, "Lung CT image registration using diffeomorphic transformation models," in *Medical Image Analysis for the Clinic—A Grand Challenge*, pp. 23–32, 2010.

[4] http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT.

[5] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "Elastix: a toolbox for intensity-based medical image registration," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 196–205, 2010.

[6] D. P. Shamonin, E. E. Bron, B. P. F. Lelieveldt, M. Smits, S. Klein, and M. Staring, "Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease," *Frontiers in Neuroinformatics*, vol. 7, article 50, 2014.

[7] C. Liu, J. Yuen, and A. Torralba, "SIFT flow: dense correspondence across scenes and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 978–994, 2011.

[8] A. Prasoon, K. Peters, C. Lgel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '13)*, vol. 8150, pp. 246–253, 2013.

[9] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[10] K. Fukushima, "Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.

[11] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 3642–3649, June 2012.

[12] D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification," in *Proceedings of*

the 22nd International Joint Conference on Artificial Intelligence (IJCAI '11), vol. 22, pp. 1237–1242, July 2011.

[13] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98–113, 1997.

[14] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network," *Neural Networks*, vol. 16, no. 5-6, pp. 555–559, 2003.

[15] P. Le Callet, C. Viard-Gaudin, and D. Barba, "A convolutional neural network approach for objective video quality assessment," *IEEE Transactions on Neural Networks*, vol. 17, no. 5, pp. 1316–1327, 2006.

[16] J.-P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[17] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Non-parametric diffeomorphic image registration with the demons algorithm," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007*, vol. 4792 of *Lecture Notes in Computer Science*, pp. 319–326, Springer, Berlin, Germany, 2007.

[18] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Symmetric log-demons diffeomorphic registration: a demons-based approach," in *Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '08)*, vol. 5241 of *Lecture Notes in Computer Science*, pp. 754–761, Springer, New York, NY, USA, 2008.

[19] M. Lorenzi, N. Ayache, G. B. Frisoni, and X. Pennec, "LCC-Demons: a robust and accurate symmetric diffeomorphic registration algorithm," *NeuroImage*, vol. 81, pp. 470–483, 2013.

[20] J.-M. Peyrat, H. Delingette, M. Sermesant, C. Xu, and N. Ayache, "Registration of 4D cardiac CT sequences under trajectory constraints with multichannel diffeomorphic demons," *IEEE Transactions on Medical Imaging*, vol. 29, no. 7, pp. 1351–1368, 2010.

[21] http://www.mathworks.cn/matlabcentral/fileexchange/39194-diffeomorphic-log-demons-image-registration.

[22] https://github.com/rasmusbergpalm/DeepLearnToolbox.

[23] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "BrainWeb: online interface to a 3D MRI simulated brain database," *NeuroImage*, vol. 5, p. 425, 1997.

[24] B. Aubert-Broche, A. C. Evans, and L. Collins, "A new improved version of the realistic digital brain phantom," *NeuroImage*, vol. 32, no. 1, pp. 138–145, 2006.

[25] http://empire10.isi.uu.nl/.

[26] http://www.mathworks.cn/matlabcentral/fileexchange/21451-multimodality-non-rigid-demon-algorithm-image-registration.

[27] http://www.itk.org/ITK/resources/software.html.

[28] K. Murphy, B. van Ginneken, J. M. Reinhardt et al., "Evaluation of registration methods on thoracic CT: the empire 10 challabnge," *IEEE Transactions on Medical Imaging*, vol. 30, pp. 1901–1920, 2011.

[29] http://elastix.isi.uu.nl/index.php.

[30] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen, "Unsupervised deep feature learning for deformable registration of MR brain images," in *Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '13)*, vol. 8150 of *Lecture Notes in Computer Science*, pp. 649–656, Springer, Nagoya, Japan, 2013.

[31] M. Liu, D. Zhang, P. T. Yap, and D. Shen, "Tree-guided sparse coding for brain disease classification," *Medical Image Computing and Computer-Assisted Intervention*, vol. 7512, no. 15, part 3, pp. 239–247, 2012.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 790313, 11 pages
http://dx.doi.org/10.1155/2015/790313

Research Article

Optimized Parallelization for Nonlocal Means Based Low Dose CT Image Processing

Libo Zhang,[1] Benqiang Yang,[1] Zhikun Zhuang,[2,3,4] Yining Hu,[2,3,4] Yang Chen,[2,3,4] Limin Luo,[2,3,4] and Huazhong Shu[2,3,4]

[1]*Department of Radiology, General Hospital of Shenyang Military Area Command, Shenhe District, Shenyang 110840, China*
[2]*Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, China*
[3]*The Key Laboratory of Computer Network and Information Integration, Southeast University and Ministry of Education, Nanjing 210096, China*
[4]*Centre de Recherche en Information Biomedicale Sino-Francais (LIA CRIBs), 35000 Rennes, France*

Correspondence should be addressed to Yang Chen; chenyang20071979@hotmail.com

Received 19 July 2014; Revised 19 September 2014; Accepted 3 October 2014

Academic Editor: Yi Gao

Low dose CT (LDCT) images are often significantly degraded by severely increased mottled noise/artifacts, which can lead to lowered diagnostic accuracy in clinic. The nonlocal means (NLM) filtering can effectively remove mottled noise/artifacts by utilizing large-scale patch similarity information in LDCT images. But the NLM filtering application in LDCT imaging also requires high computation cost because intensive patch similarity calculation within a large searching window is often required to be used to include enough structure-similarity information for noise/artifact suppression. To improve its clinical feasibility, in this study we further optimize the parallelization of NLM filtering by avoiding the repeated computation with the row-wise intensity calculation and the symmetry weight calculation. The shared memory with fast I/O speed is also used in row-wise intensity calculation for the proposed method. Quantitative experiment demonstrates that significant acceleration can be achieved with respect to the traditional straight pixel-wise parallelization.

1. Introduction

X-ray Computed Tomography (CT) can reflect human attenuation map in millimeter level, in which rich 3D information of tissues, organs or lesions can be provided for clinical diagnosis. Though its wide application in clinics, the radiation delivered to patients during CT examinations is always a wide-spread concern. It was reported in [1] that CT radiation may increase the risk of developing metabolic abnormalities even cancer. The most practical means to lower radiation dose is to decrease tube current (milliampere (mA)) or tube current time products (milliampere second (mAs)). However, lowering mA/mAs settings often leads to degraded CT images with increased mottled noise and streak artifacts [2, 3], which will influence the diagnosis accuracy [4–7]. Researchers can suppress noise and artifacts in low dose CT (LDCT) images by developing new reconstruction or postprocessing algorithms. Current solutions to improve the quality of LDCT images can be roughly divided into three categories: preprocessing approaches, iterative reconstruction approaches, and postprocessing approaches. The first one refers to those techniques that improve CT imaging by suppressing the noise in projected raw data before the routine FBP reconstructions. The key of these techniques is to find the accurate statistical distribution of projected data and design effective restoration algorithms [5, 6]. The second one refers to iterative reconstruction approaches which treat the LDCT imaging as an ill-posed inverse problem and solve the problem as a prior-regularized cost function via some iterative optimization solutions [7, 8]. Though effective in

obtaining favorable reconstructed image quality, the most well-known limit for iterative reconstructions is the required intensive computation in iterative optimization. Additionally, for patent protection consideration, current mainstream CT device suppliers normally do not provide well-formatted projected data, which severely restricts the research and the possible clinical application of these two study directions.

The third one refers to postprocessing methods, which can be directly applied to improve LDCT images. Distribution and scale features of noise, artifacts, and normal tissues in CT images need to be jointly considered in designing effective postprocessing algorithms [9, 10]. It was pointed in [9–14] that the nonlocal means (NLM) filtering, which utilizes the information redundancy property, can effectively suppress noise and artifacts without obviously blurring image details. We would also note that the patch similarity metric in NLM has also been used to build regularization term for tomographic reconstruction [15, 16].

However, since noise and artifacts often distribute with prominent amplitudes in LDCT images, a large searching window is practically required to include more structure information in noise/artifact suppression, which implies a large computation cost. This will strongly limit its clinical application considering the large workload in current radiology departments. To overcome this, this paper presents an improved GPU-based parallelization approach to accelerate the NLM filtering. The proposed approach optimizes the computation in NLM filtering by avoiding repeated computation with row-wise intensity calculation and weight calculation. The fast I/O data access speed for shared memory in GPU is also well exploited to reduce data I/O operation cost. Experiment results on 2D LDCT images demonstrate that the improved parallelization can significantly shorten computation time and, thus, making itself a potentially applicable processing procedure in LDCT imaging.

2. Nonlocal Means Based Low Dose CT Image Processing

Compared to restoration algorithms based on intensity gradient information, NLM filtering can provide edge-preserving noise/artifact suppression without blurring image structures. In NLM filtering, one image patch is matched with a group of similar patches in a large neighboring area, and in this way more structure similarity information in large neighboring scale can be used to suppress noise and artifacts in LDCT images. The NLM algorithm replaces pixel intensities by the weighted average of intensities within a searching window N. Each weight expresses the similarity between the central pixel and the neighboring pixels in the searching window and is calculated by the Euclidian distance between patches surrounding these two pixels. Let p ($p = (p_x, p_y)$) denote the pixel to be processed, let q denote a pixel in the search neighborhood window, let X denote the processed image, and let Y denote the image to be processed; the 2D NLM filtering algorithm can be formulated as follows [17]:

$$\widehat{X}(p) = \frac{\sum_{q\in N_p} \omega(p,q)\, Y(q)}{\sum_{q\in N_p} \omega(p,q)}, \tag{1}$$

$$w(p,q) = \exp\left(-\frac{\sum_{(\Delta x,\Delta y)\in[-B,\ldots,B]^2} \left|d_{p,q}^{(\Delta x,\Delta y)}\right| G(\Delta x,\Delta y)}{h(2B+1)(2B+1)}\right), \tag{2}$$

$$d_{p,q}^{(\Delta x,\Delta y)} = Y\left(p_x+\Delta x, p_y+\Delta y\right) - Y\left(q_x+\Delta x, q_y+\Delta y\right), \tag{3}$$

$$G(\Delta x,\Delta y) = \begin{cases} \sqrt{2} & (\Delta x,\Delta y) = (0,0) \\ \dfrac{1}{\sqrt{\Delta x^2+\Delta y^2}} & (\Delta x,\Delta y) \neq (0,0), \end{cases} \tag{4}$$

where N_p denotes the searching window centered at p; $w(p,q)$ denotes the similarity of the two patches centered at p and q, respectively, with the radius B; $G(\Delta x, \Delta y)$ is a distance dependent Gaussian kernel function; the number of pixels in a patch is $(2B+1)(2B+1)$. We routinely use the parameter h in (2) to control the smoothing effect.

In Figures 1–4, we give the NLM filtering processed results of four 2D LDCT images in Figures 1(a), 2(a), 3(a), and 4(a) and the two corresponding standard dose CT (SDCT) images are given in Figures 1(b), 2(b), 3(b), and 4(b) as references. The LDCT and SDCT images were collected using the reduced tube current 80 mA and the routine tube current 240 mA, respectively. We can see that CT images are mainly composed of pixels with limited intensity range, and the intensities representing different tissues spread over the whole image domain. Other scanning parameters were set by default. Compared to the reference SDCT images, we can see that tube current reduction will lead to severely increased noise and artifacts in LDCT images. Figures 1(c), 2(c), 3(c), and 4(c) illustrate the processing results of NLM filtering, in which the size of searching window is 81×81, patch size is 9×9 ($B = 4$), and parameter h is set to 10. Figure 5 shows the results of 3D volumes by processing a set of 2D thoracic LDCT images. The illustrations in Figures 1–4 are presented in suitable windows. All the parameters were set under the guide of an experienced doctor in radiology department. We can see that the NLM filtering can effectively suppress both mottled noise and artifacts in LDCT images without leading to significantly blurred structures.

To highlight the importance of a large searching window, we also list in Figure 1(d) the NLM processed result with 21×21 searching window (other parameters are set to be same as the result in Figure 1(c)). We can see that processing with this smaller 21×21 searching window fails to give satisfying artifact suppression (see the arrows). Therefore, a large searching window (up to 81×81) needs to be used to in NLM filtering to include enough large-scale similarity information

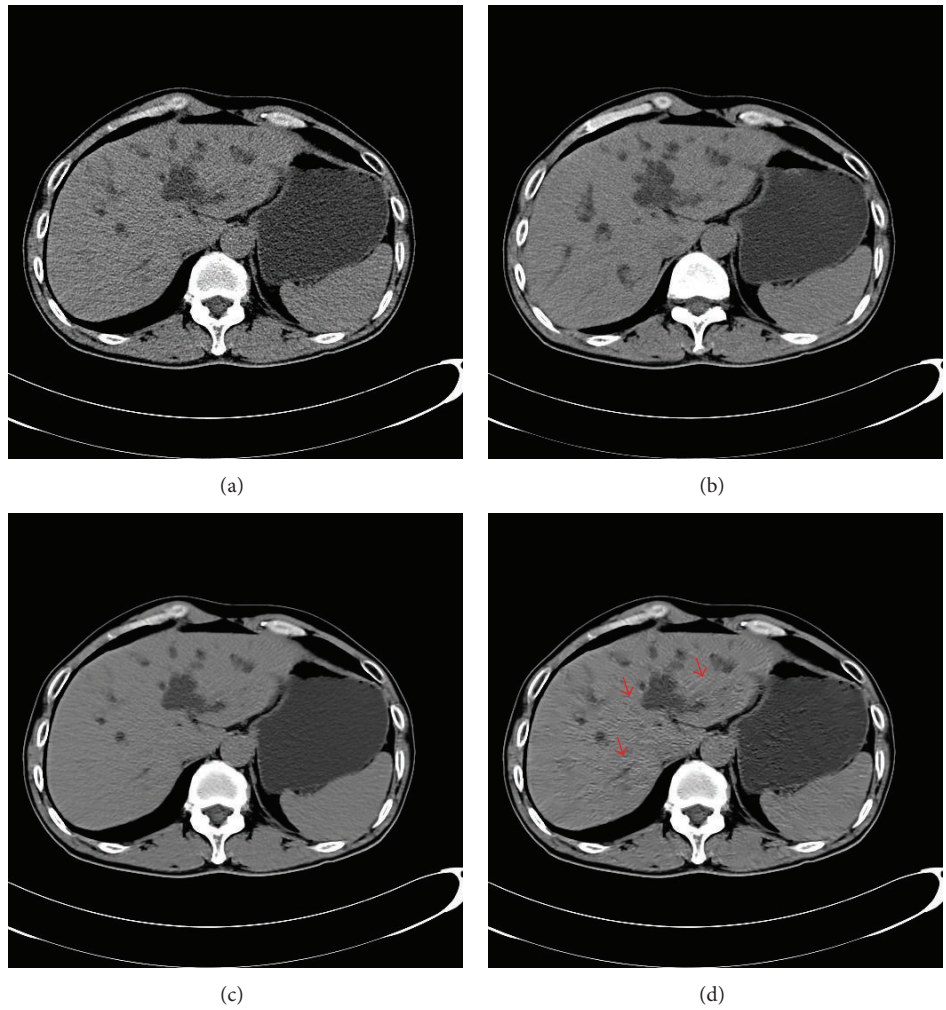


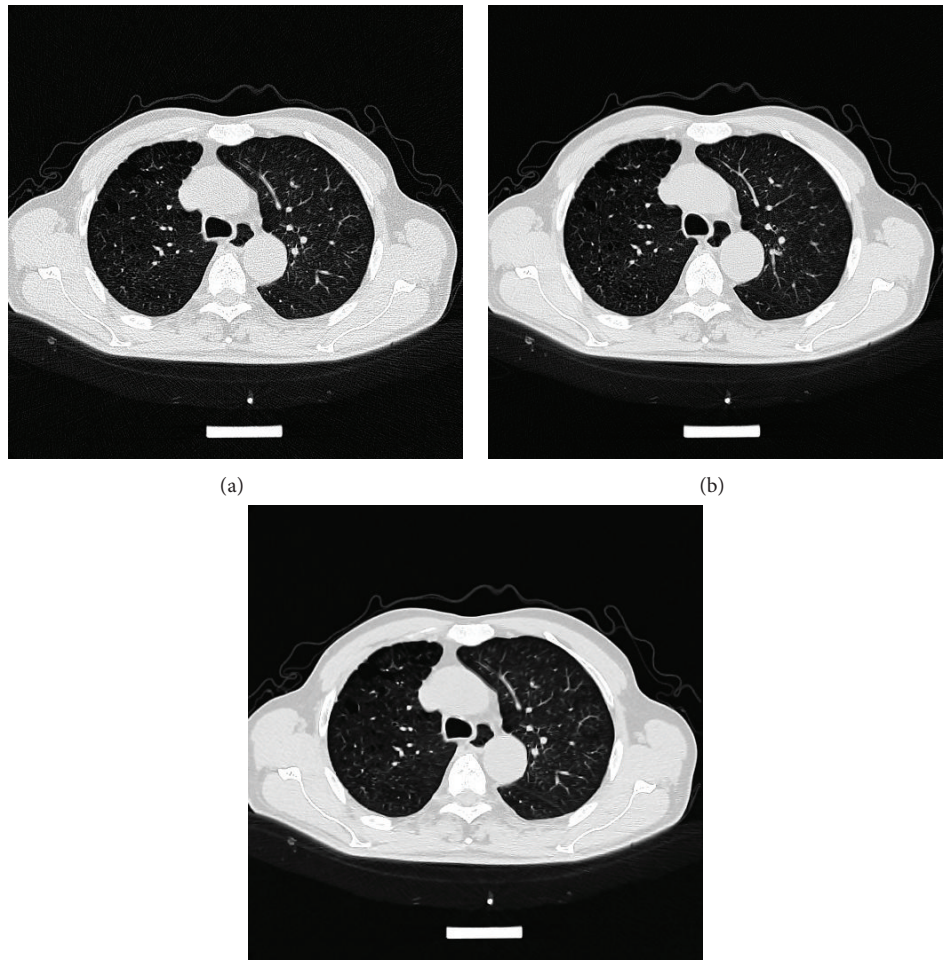
Figure 1: Processed result of one 512 × 512 abdomen LDCT image. (a) and (b) are the original LDCT image and the corresponding SDCT image. (c) and (d) are the LDCT images processed by NLM filtering with 81 × 81 and 81 × 81 21 × 21 searching windows, respectively.

to suppress noise and artifacts in LDCT images, as can be seen in Figure 1(c) [10]. However, the patch similarity calculation within a large searching window is always accompanied with large computation load. For a $m \times n$ sized image, with the pixel number of searching window being $|N|$ and patch radius being B, we get the computational complexity $O(mn|N|(2B+1)(2B+1))$ for the original CPU based serial processing, and the total computational complexity amounts to $O(512\times512\times81\times81\times9\times9) = O(1.3931\times10^{11})$ for 512 × 512 sized images. This computation cost is too high to provide real-time CT imaging for radiology department routine; so we need to accelerate the NLM filtering in order to give fast clinical application.

3. CUDA-Based GPU Acceleration for NLM Algorithm

3.1. Introduction to CUDA-Based GPU Acceleration. Utilizing GPU based techniques to parallelize algorithm has already become a notable trend in the field of parallel computing. The GPU based parallelization is achieved by jointly parallelizing coarse-scale patches and fine-scale threads in the original grid computation task which is parallelizable [18–20]. The CUDA (Compute Unified Device Architecture) technology provides a software platform for developers to design parallelized tasks with C-style code, with direct access to the virtual instruction set and GPU memories. Each parallelization function running on CUDA is called a kernel, and we use $f^{(I)} \circ f^{(I-1)} \circ \cdots \circ f^{(1)}$ to represent the connected parallelized cascade functions based on [20]. The output of kernel function $f^{(i)}$ is the input of $f^{(i+1)}$. We use $(U_1^{(i)}, \ldots, U_k^{(i)})$ to represent the input data of kernel function $f^{(i)}$ in processing and $(U_1^{(i+1)}, \ldots, U_k^{(i+1)})$ to represent the output data of function $f^{(i)}$. We can use (5) to represent the kernel function as follows:

$$\left[U_1^{(i+1)}, \ldots, U_k^{(i+1)}\right](p) = f^{(i)}_{U_1^{(i)},\ldots,U_k^{(i)}}(p), \tag{5}$$



(a) (b) (c)

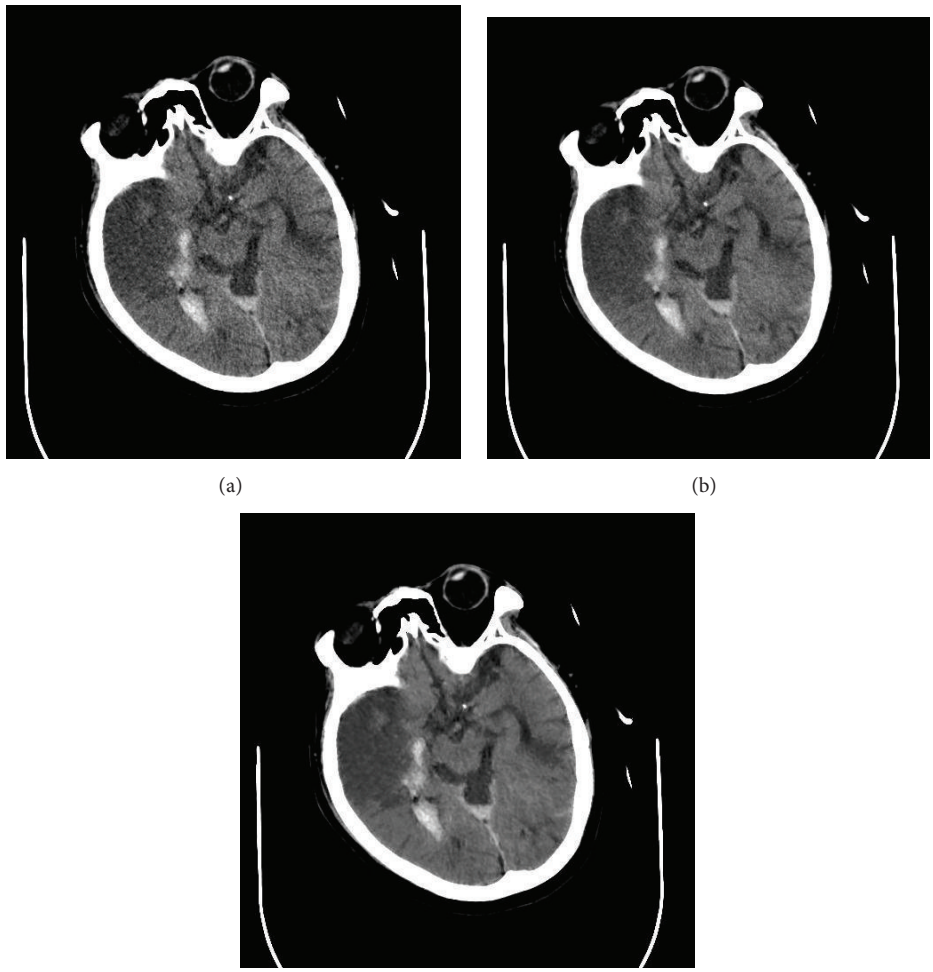
Figure 2: Processed result of one 512×512 thoracic LDCT image. (a), (b), and (c) correspond to the original LDCT image, the corresponding SDCT image, and the LDCT image processed by NLM filtering, respectively.

where p represents the pixel position in image. It should be noted that in some cases not all the input data need to be updated (e.g., $U_j^{(i+1)} = U_j^{(i)}$).

3.2. Conventional GPU Based Parallelization for NLM Filtering Algorithm. The conventional GPU based parallelization accelerates the NLM filtering algorithm by direct pixel-wise parallelization. Based on above (1)–(4), we routinely break the algorithm into four parts specified by the following (6)–(9), which are computed in loops. The number of loops is set as the searching window size $|N|$ to traverse all the neighboring points in window N. The first kernel function (6) computes intensity differences in parallel via GPU and has computational complexity $O(1)$. Here we quantify the computational complexity using the operation times in parallel. In (6), $(p_x + i_x, p_y + i_y)$ denotes the spatial position of the neighboring pixel in the searching window centered at p, which can also be represented by spatial position (p_x, p_y). We set $U_3^{(1)} = U_4^{(1)} = 0$ as initialization. Consider

$$\begin{aligned}
&\left[U_1^{(3i-1)}, \ldots, U_4^{(3i-1)}\right](p) \\
&\quad = f^{(3i-2)}_{U_1^{(3i-2)},\ldots,U_4^{(3i-2)}}(p) \\
&\quad = \begin{pmatrix} \left|Y(p_x, p_y) - Y(p_x + i_x, p_y + i_y)\right| \\ U_2^{(3i-2)}(p) \\ U_3^{(3i-2)}(p) \\ U_4^{(3i-2)}(p) \end{pmatrix}.
\end{aligned} \tag{6}$$

For data U_2, the second kernel function computes the patch similarity using (7) based on the patch difference computed by (6) in U_1. We can see that the computational



(a) (b) (c)

Figure 3: Processed result of one 512 × 512 brain LDCT image. (a), (b), and (c) correspond to the original LDCT image, the corresponding SDCT image, and the LDCT image processed by NLM filtering, respectively.

complexity of the second kernel function is around $O((2B + 1)(2B + 1))$ because there are $(2B + 1)(2B + 1)$ weighted summation operations for each pixel pair in the two comparing patches. Consider

$$\begin{aligned}&\left[U_1^{(3i)},\ldots,U_4^{(3i)}\right](p)\\&= f^{(3i-1)}_{U_1^{(3i-1)},\ldots,U_4^{(3i-1)}}(p)\\&= \begin{pmatrix}(3i-1)\\ \exp\left(-\frac{\sum_{(\Delta x,\Delta y)\in[-B,\ldots,B]^2} U_1^{(3i-1)}\left(p_x+\Delta x,p_y+\Delta y\right)G(\Delta x,\Delta y)}{h(2B+1)(2B+1)}\right)\\ U_3^{(3i-1)}\\ U_4^{(3i-1)}\end{pmatrix}.\end{aligned} \tag{7}$$

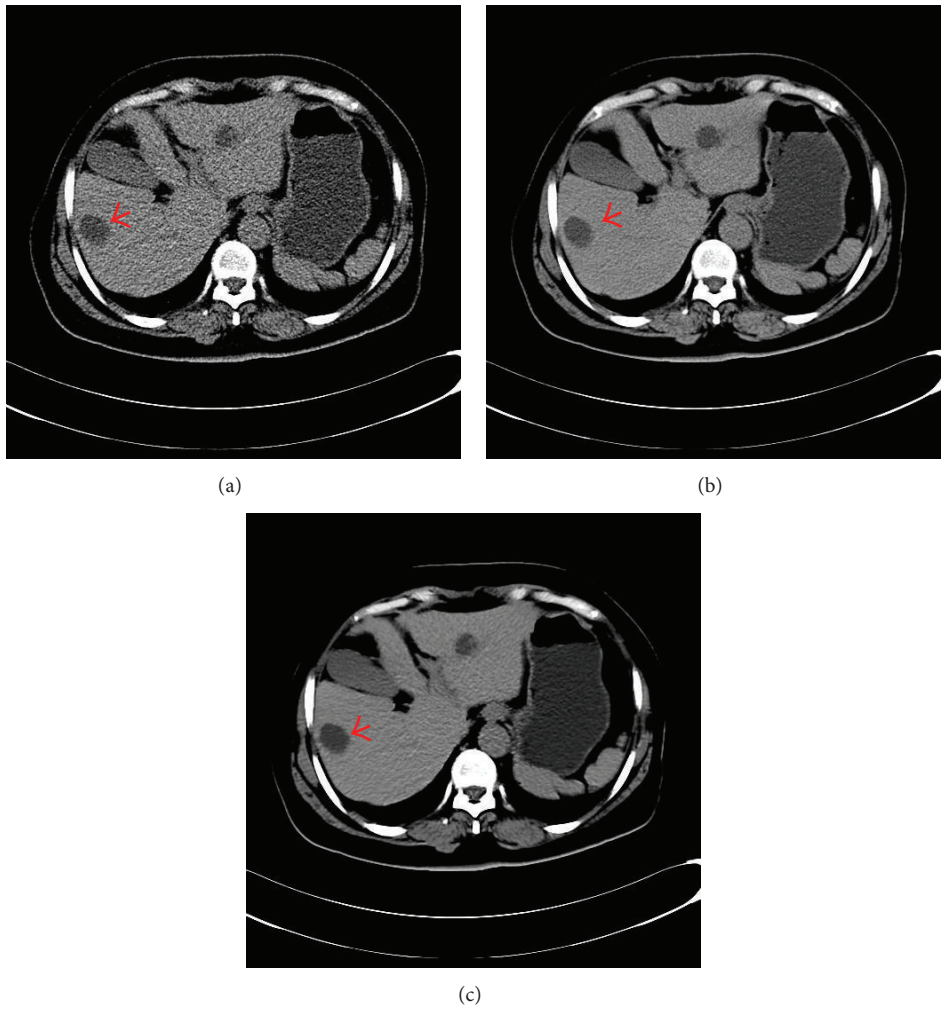
The third kernel function (8) computes the summation of weights and intensities in U_3 and U_4, and the computational complexity is $O(1)$ for this operation. Consider

$$\begin{aligned}&\left[U_1^{(3(i+1)-2)},\ldots,U_4^{(3(i+1)-2)}\right](p)\\&\quad= f^{(3i)}_{U_1^{(3i)},\ldots,U_4^{(3i)}}(p)\\&\quad= \begin{pmatrix}U_1^{(3i)}(p)\\ U_2^{(3i)}(p)\\ U_3^{(3i)}(p)+U_2^{(3i)}(p)\\ U_4^{(3i)}(p)+U_2^{(3i)}(p)\,Y\left(p_x+i_x,p_y+i_y\right)\end{pmatrix}.\end{aligned} \tag{8}$$

In the final loop $I = |N| + 1$, a last kernel function in (9) is applied to compute the final output image $\widehat{X}(p)$. Consider

$$f^{(I)}_{U_1^{(I)},\ldots,U_4^{(I)}}(p) = \begin{pmatrix}0\\0\\0\\ \dfrac{U_4^I(p)}{U_3^I(p)}\end{pmatrix}. \tag{9}$$

Here I represents the last loop number. The computational complexity for the operation (9) is also $O(1)$. The final image is outputted as $\widehat{X}(p) = U_4^{(I)}(p)/U_3^{(I)}(p)$. Combing all the operations from (6)–(9), we can see that the whole



(a) (b) (c)

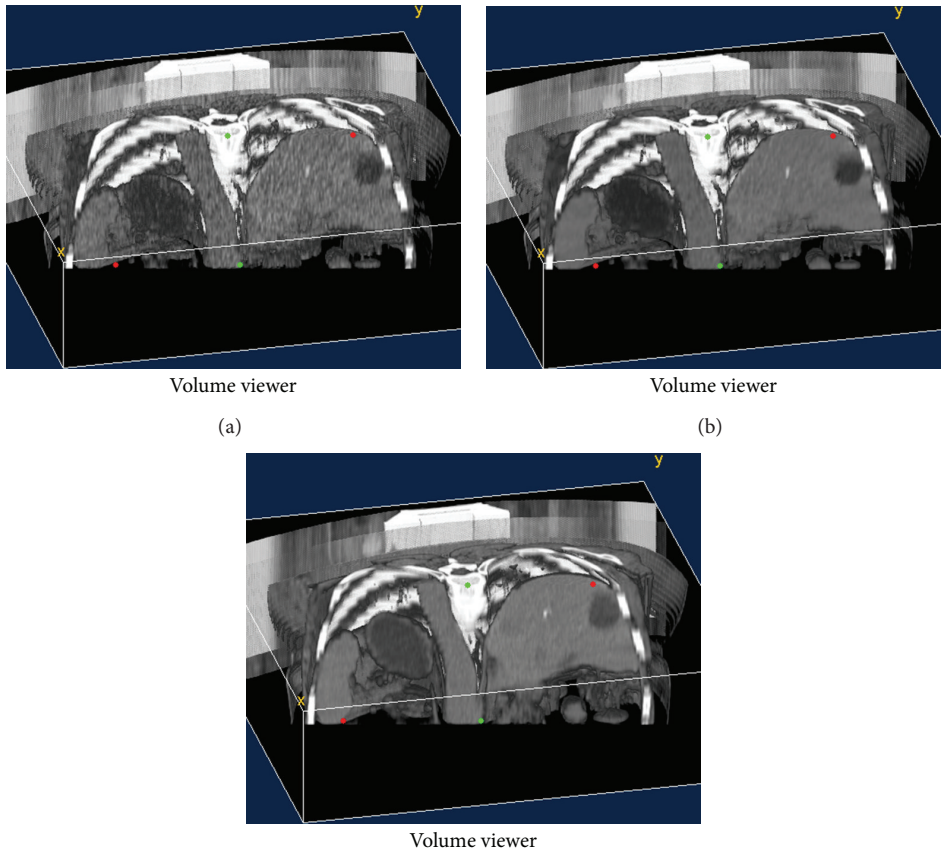
Figure 4: Processed result of one 512×512 abdomen LDCT image with tumor (pointed by arrows). (a), (b), and (c) correspond to the original LDCT image, the corresponding SDCT image, and the LDCT image processed by NLM filtering, respectively.

computational complexity of the conventional parallelization algorithm amounts to $O(|N|((2B+1)(2B+1)+2)+1)$.

3.3. Improved GPU Acceleration for NLM Filtering Algorithm. In the above conventional parallelization approach, the second kernel function in (7) is serially applied to compute the patch similarity, which leads to large computation cost when large searching window is used. Our first improvement is, thus, devoted to reduce the computational complexity in this part. Figure 6 illustrates that a patch is of size 5×5 ($B = 2$) with the red point indicating the center point. Equations (1)–(4) show that the patch similarity in NLM filtering can be quantified by the weighted sum of intensity differences of the corresponding pixels in the two patches. In Figure 6, we can see that, for the center points located at the green points in the two patches, the summed intensity difference of the blue points in the same rows is in fact the same value as the case when the center points moves down to the red points. This implies that the intensity differences of rows are repeatedly computed when the center points move within $(B+1)$ pixel distances. Therefore, we can efficiently calculate patch differences via the following row-wise calculation:

$$\bigcup_{\Delta y\in[0,\ldots,B]} \sum_{\Delta x\in[-B,\ldots,B]} \left|Y\left(p_x+\Delta x, p_y\right) - Y\left(q_x+\Delta x, q_y\right)\right| \times G\left(\Delta x, \Delta y\right), \tag{10}$$

where the intensity difference between two individual pixels is $|Y(p_x+\Delta x, p_y)-Y(q_x+\Delta x, q_y)|$ and $q=(q_x,q_y)$ represents the neighboring point positions in the searching window. $\bigcup_{\Delta y\in[0,\ldots,B]}$ denotes the dataset that includes the $(B+1)$ different points in the vertical direction. Thus, with patches of size $(2B+1)\times(2B+1)$, we know from (10) that $(B+1)$ values can be obtained for $(B+1)$ different row pairs. *The row difference needs to be calculated only once before being stored in the shared memory, and the other B operations in (10) can be easily obtained by loading data from the shared memory and then performing Gaussian weighting.* For GPU with fast single-precision floating processing, the main computation



(a) (b) (c)

FIGURE 5: 3D illustration of a set of thoracic LDCT images. (a), (b), and (c) correspond to the original LDCT volume, the corresponding SDCT volume, and the LDCT volume processed by NLM filtering, respectively.

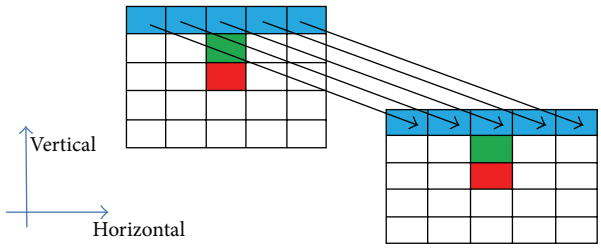


FIGURE 6: Row-wise calculation in patch difference calculation.

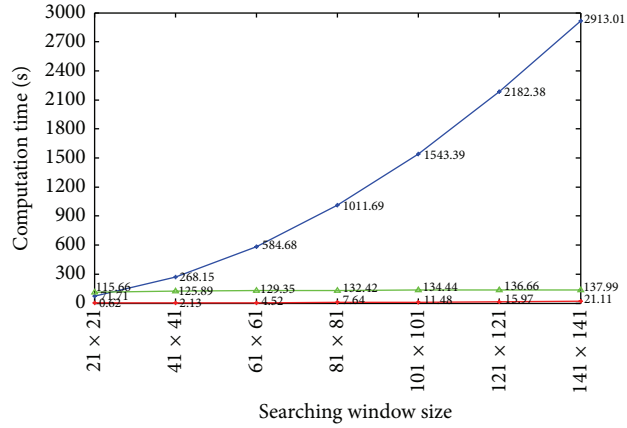


FIGURE 7: Comparison of the computation time of the serial algorithm and the conventional parallelization algorithm for NLM filtering.

cost of (10) lies in the data accessing operation of the global memory because the time cost in shared memory accessing is trivial when compared to global memory accessing. The computational complexity of (11) can be roughly estimated to be $O(2B + 1)$ [21].

Similar to the conventional GPU parallelization, we also divide the algorithm into the following four parts (11)–(14) and compute in loops. Suppose that the input image is of size $m \times n$, we set the size of $U_1^{(i)}$ to be $m \times n \times (B+1)$. The data $U_2^{(i)}$, $U_3^{(i)}$, $U_4^{(i)}$, $U_5^{(i)}$ are of size $m \times n$, and $(p_x + i_x, p_y + i_y)$ denotes the neighboring points in the searching window centered at p. The initialization is also set with data $U_3^{(1)} = U_4^{(1)} = 0$, $U_5^{(i)} = Y$.

The first kernel function computes the sum of the intensity differences for each row pair, which is multiplied by the Gaussian weight calculated based on the perpendicular distance from the row to the center point. The computational complexity of this kernel function is $O(2B + 1)$. Consider

$$\begin{aligned}\left[U_1^{(3i-1)},\ldots,U_5^{(3i-1)}\right](p) &= f^{(3i-2)}_{U_1^{(3i-2)},\ldots,U_5^{(3i-2)}}(p)\\ &= \begin{pmatrix} \bigcup_{\Delta y\in[0,\ldots,B]}\sum_{\Delta x\in[-B,\ldots,B]}\left|Y\left(p_x+\Delta x,p_y\right)-Y\left(p_x+i_x+\Delta x,p_y+i_y\right)\right|G(\Delta x,\Delta y)\\ U_2^{(3i-2)}(p)\\ U_3^{(3i-2)}(p)\\ U_4^{(3i-2)}(p)\\ U_5^{(3i-2)}(p)\end{pmatrix}.\end{aligned} \tag{11}$$

The second kernel function calculates the similarity of patches based on (12). In (12), we compute the patch similarity by accumulating the absolute values of the weighted sum of the intensity differences calculated via the first kernel function (11). The computational complexity of the kernel function (12) is $O(2B + 1)$. Consider

$$\begin{aligned}&\left[U_1^{(3i)},\ldots,U_5^{(3i)}\right](p)\\ &\quad= f^{(3i-1)}_{U_1^{(3i-1)},\ldots,U_5^{(3i-1)}}(p)\\ &\quad= \begin{pmatrix} U_1^{(3i-1)}(p)\\ \exp\left(-\dfrac{\sum_{\Delta y\in[-B,\ldots,B]}U_1^{(3i-1)}\left(p_x,p_y,\Delta y\right)}{h(2B+1)(2B+1)}\right)\\ U_3^{(3i-1)}(p)\\ U_4^{(3i-1)}(p)\\ U_5^{(3i-1)}(p)\end{pmatrix}.\end{aligned} \tag{12}$$

The second improvement is saving one half computation cost by exploiting the symmetry property of weights calculated in (2). Apparently, we have $w(p, p + \Delta p) = w(p + \Delta q, p)$ (Δq represents the location offset of pixel p in the searching window). Based on this symmetry property $w(p - \Delta q, p)Y(p-\Delta q) = w(p, p-\Delta q)Y(p-\Delta q)$, we also accumulate $w(p-\Delta q, p)Y(p-\Delta q)$ when accumulating weighted intensity $w(p, p+\Delta q)Y(p+\Delta q)$ for location p. In this way, we only need to go through half of the pixels in the searching window. The third kernel function is given by (13), whose computational complexity is $O(1)$. Consider

$$\begin{aligned}&\left[U_1^{(3(i+1)-2)},\ldots,U_5^{(3(i+1)-2)}\right](p)\\ &= f^{(3i)}_{U_1^{(3i)},\ldots,U_5^{(3i)}}(p)\\ &= \begin{pmatrix} U_1^{(3i)}(p)\\ U_2^{(3i)}(p)\\ U_3^{(3i)}(p)+U_2^{(3i)}(p+\Delta q)+U_2^{(3i)}(p-\Delta q)\\ U_4^{(3i)}(p)+U_2^{(3i)}(p+\Delta q)U_5^{(3i)}(p+\Delta q)+U_2^{(3i)}(p-\Delta q)U_5^{(3i)}(p-\Delta q)\\ U_5^{(3i)}(p)\end{pmatrix}.\end{aligned} \tag{13}$$

Then, a final kernel function (14) can be applied to obtain the finally processed image:

$$f^{(I)}_{U_1^{(I)},\ldots,U_5^{(I)}}(p) = \begin{pmatrix} 0\\ 0\\ 0\\ 0\\ \dfrac{U_4^{(I)}(p)}{U_3^{(I)}(p)}\end{pmatrix}. \tag{14}$$

The final loop number with respect to the required operation number as to the searching window now becomes $(2T + 1) \times (T + 1) + 1$ (T denotes the radius of the searching window), which is approximately $0.5|N|$. The final output image is $\widehat{X}(p) = U_4^{(I)}(p)/U_3^{(I)}(p)$. To conclude, the total computational complexity of the improved algorithm is around $O(0.5|N|(2(2B + 1) + 1) + 1)$, which approximately equals to $O(|N|(2B+1)+1)$. We can see that the computational complexity has been reduced to $1/(2B+1)$ with respect to the conventional parallelization.

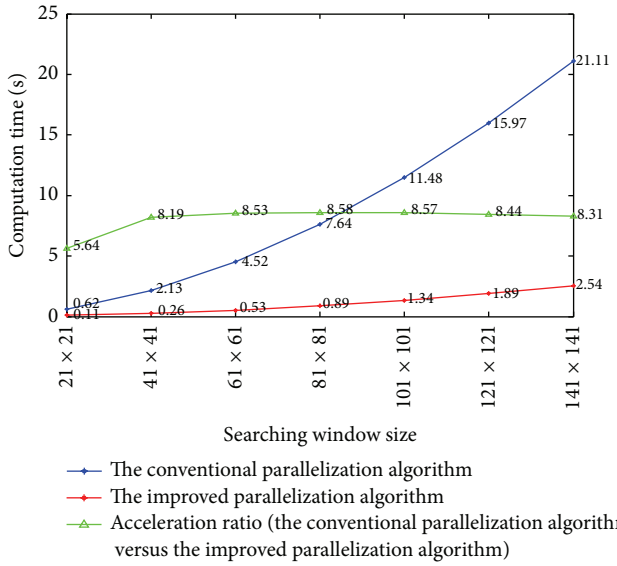


Figure 8: Comparison of the computation time with respect to searching window size for the conventional parallelization algorithm and the improved parallelization algorithm.

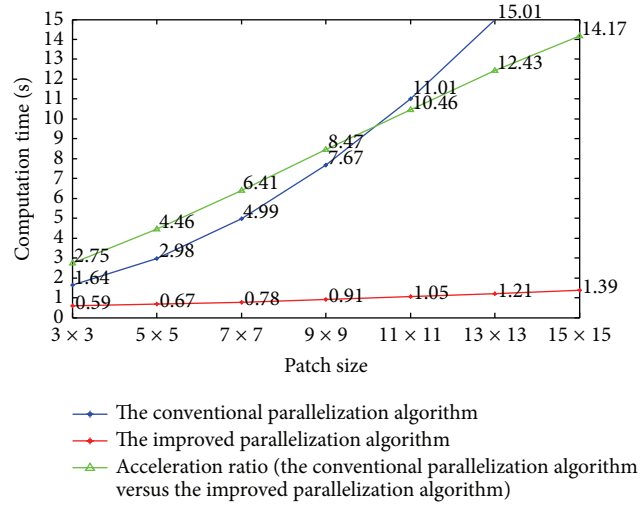


Figure 9: Comparison of the computation time with respect to patch size for the conventional parallelization algorithm and the improved parallelization algorithm.

4. Experimental Results and Analyses

In this section we compare the computation cost of different methods. To verify the improvement brought by the proposed acceleration method for the NLM filtering, we process the same 512 × 512 LDCT image in Figure 1(a) using the serial algorithm (CPU based), the conventional parallelization algorithm (GPU based), and the improved parallelization algorithm (GPU based). In this section, we do not illustrate the processed images because the same images as Figure 1(a) were obtained. We set the patch size to be 9 × 9 and record the computation time with respect to the size of searching window. Figure 7 illustrates the computation time of the serial algorithm and the conventional parallelization algorithm. We can observe that the conventional parallelization significantly reduces the computation cost through straight pixel-wise parallelization and achieves an acceleration ratio of more than 100 times of the original serial algorithm. The system configuration for our experiments is given as follows.

4.1. Hardware Environment. CPU: Inter(R) Core(TM) i7-3770 CPU @ 3.40 GHz; Memory: 8 GB; Graphics Card: NVIDIA GeForce GTX 680 with 1536 CUDA cores; Effective memory clock: 6008 MHz; Memory bandwidth: 192 GB/s; Memory size: 2 GB; Memory bus type: 256 bit.

4.2. Software Environment. Operating System: Win7 64 bit; Matlab: R2011a; CUDA: 4.0.

Then, we compare the computation time of the conventional parallelization algorithm and the improved parallelization algorithm with respect to the size of searching window. The patch size is fixed to 9 × 9. As can be seen in Figure 8, when the searching window size becomes larger than 41 × 41, the acceleration ratio approximately equals to $2B + 1 = 2 \times 4 + 1 = 9$, and this observation is consistent with the above deduced acceleration ratio $2B + 1$. In addition, we compare the computation time of the conventional parallelization algorithm and the improved parallelization algorithm with respect to patch size. The searching window size is fixed to 81 × 81. Since a large patch size often leads to blurred images, we hereby set the maximal patch size to be 15 × 15. We can observe in Figure 9 an obvious increment of acceleration ratio when the patch size increases, and this again verifies the above deduced acceleration ratio $2B + 1$.

5. Discussion and Conclusion

In this paper we further optimize the parallelization for NLM filtering in CT image processing. The proposed approach optimizes the parallelized computation in NLM filtering by avoiding repeated computation with row-wise intensity calculation and weight calculation. The fast I/O data access speed for shared memory in GPU is also well exploited. We applied our improved algorithm to LDCT image processing and find that the improved algorithm can achieve a significant acceleration ratio with respect to the conventional parallelization algorithm. For now, it takes about 0.8 second to process one 512×512 CT image with 81×81 searching window and 9×9 patch, and this parameter setting in NLM filtering is found to be able to provide effective processing. This paper only provides the results on 2D NLM filtering, and we would stress that the same parallelization strategy can be easily extended to accelerate the more computationally intensive 3D NLM filtering, and the same acceleration ratio as 2D case can be expected because they have the same calculation structures. To be specific, this extension can be realized by replacing the row-wise optimization in (11) by a plane-wise optimization. Nevertheless, we would also point it out that 3D NLM filtering should not be suggested for the processing of CT slices with

large slice thickness (>2 mm) because of the poor interslice continuity in this case.

Currently, the structure similarity idea in NLM has got wide applications in the other field of image processing (e.g., image segmentation and image reconstruction) [15, 16, 22, 23]. The proposed parallelization optimization can be directly applied to accelerate the patch similarity calculation in these applications. In current parallelization approach, the weights reflecting patch similarity are calculated via a serial loop in (7), which can be further parallelized via interkernel operations to realize a further acceleration. Accelerating the computation speed by combining multiple-core CPU strategy with GPU parallelization technique will also be explored. This optimization strategy can be easily used to accelerate other reconstruction or restoration tasks using the patch similarity type metrics [24–26]. We can also consider improving the performance of the NLM filtering by incorporating the fractional metric into the calculation of patch similarity [27]. Evaluation of the potential accuracy enhancement in segmentation/registration (related with CT images) that can be brought by the proposed processing also needs to be performed [28–30]. All these issues will be addressed in the future work.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study is supported by the Science and Technology Research Project of Liaoning Province (2013225089), the National Natural Science Foundation under Grants (81370040, 31100713), and the Qing Lan Project in Jiangsu Province. This research was also partly supported by National Basic Research Program of China under Grant (2010CB732503).

References

[1] G. Zhang, D. Sun, P. Yan, H. Zhao, and Z. Li, "A LDCT image contrast enhancement algorithm based on single-scale retinex theory," in *Proceedings of the International Conference on Computational Intelligence for Modelling Control & Automation*, pp. 1282–1287, Vienna, Austria, December 2008.

[2] M. K. Kalra, M. M. Maher, T. L. Toth et al., "Strategies for CT radiation dose optimization," *Radiology*, vol. 230, no. 3, pp. 619–628, 2004.

[3] E. Angel, N. Yaghmai, C. M. Jude et al., "Monte Carlo simulations to assess the effects of tube current modulation on breast dose for multidetector CT," *Physics in Medicine and Biology*, vol. 54, no. 3, pp. 497–511, 2009.

[4] R. Nelson, "Low-dose CT screening for lung cancer produces high rate of false positives," in *Proceedings of the 45th Annual American Society of Clinical Oncology Meeting (ASCO '09)*, 2009.

[5] J. Wang, H. Lu, J. Wen, and Z. Liang, "Multiscale penalized weighted least-squares sinogram restoration for low-dose X-ray computed tomography," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 3, pp. 1022–1031, 2008.

[6] T. Kubo, Y. Ohno, S. Gautam et al., "Use of 3D adaptive raw-data filter in CT of the lung: effect on radiation dose reduction," *The American Journal of Roentgenology*, vol. 191, no. 4, pp. W167–W174, 2008.

[7] Y. Chen, D. Gao, C. Nie et al., "Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior," *Computerized Medical Imaging and Graphics*, vol. 33, no. 7, pp. 495–500, 2009.

[8] I. A. Elbakri and J. A. Fessier, "Efficient and accurate likelihood for iterative image reconstruction in X-ray computed tomography," in *Medical Imaging 2003: Image Processing*, vol. 5032 of *Proceedings of the SPIE*, pp. 1839–1850, San Diego, Calif, USA, February 2003.

[9] Y. Chen, Z. Yang, Y. Hu et al., "Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means," *Physics in Medicine and Biology*, vol. 57, no. 9, pp. 2667–2688, 2012.

[10] Y. Chen, W. Chen, X. Yin et al., "Improving low-dose abdominal CT images by weighted intensity averaging over large-scale neighborhoods," *European Journal of Radiology*, vol. 80, no. 2, pp. e42–e49, 2011.

[11] F. P. X. de Fontes, G. A. Barroso, P. Coupé, and P. Hellier, "Real time ultrasound image denoising," *Journal of Real-Time Image Processing*, vol. 6, no. 1, pp. 15–22, 2011.

[12] W. Xu and K. Mueller, "A reference image database approach for NLM filter-regularized CT reconstruction," in *Proceedings of the Fully3D*, pp. 116–119, 2011.

[13] X. Jia, Z. Tian, Y. Lou, J.-J. Sonke, and S. B. Jiang, "Four-dimensional cone beam CT reconstruction and enhancement using a temporal nonlocal means method," *Medical Physics*, vol. 39, no. 9, pp. 5592–5602, 2012.

[14] Z. Li, L. Yu, J. D. Trzasko et al., "Adaptive nonlocal means filtering based on local noise level for CT denoising," *Medical Physics*, vol. 41, no. 1, Article ID 011908, 2014.

[15] Y. Chen, D. Gao, C. Nie et al., "Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior," *Computerized Medical Imaging and Graphics*, vol. 33, no. 7, pp. 495–500, 2009.

[16] Y. Chen, J. Ma, Q. Feng, L. Luo, P. Shi, and W. Chen, "Nonlocal prior Bayesian tomographic reconstruction," *Journal of Mathematical Imaging and Vision*, vol. 30, no. 2, pp. 133–146, 2008.

[17] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 60–65, June 2005.

[18] Y. Chen, L. Luo, W. Chen et al., "Joint-MAP tomographic reconstruction with patch similarity based mixture prior model," *Multiscale Modeling and Simulation*, vol. 9, no. 4, pp. 1399–1419, 2011.

[19] D. Gembris, M. Neeb, M. Gipp, A. Kugel, and R. Männer, "Correlation analysis on GPU systems using NVIDIA's CUDA," *Journal of Real-Time Image Processing*, vol. 6, no. 4, pp. 275–280, 2011.

[20] B. Goossens, H. Luong, J. Aelterman, A. Pižurica, and W. Philips, "A GPU-accelerated real-time NLMeans algorithm for denoising color video sequences," in *Advanced Concepts for Intelligent Vision Systems*, vol. 6475 of *Lecture Notes in Computer Science*, pp. 46–57, Springer, Berlin, Germany, 2010.

[21] Z. Zhuang, Y. Chen, H. Shu, L. Luo, C. Toumoulin, and J.-L. Coatrieux, "Fast low-dose CT image processing using improved parallelized nonlocal means filtering," in *Proceedings of the International Conference on Medical Biometrics*, pp. 147–150, Shenzhen, China, June 2014.

[22] B. Caldairou, N. Passat, P. A. Habas, C. Studholme, and F. Rousseau, "A non-local fuzzy segmentation method: application to brain MRI," *Pattern Recognition*, vol. 44, no. 9, pp. 1916–1927, 2011.

[23] M. Protter, M. Elad, H. Takeda, and P. Milanfar, "Generalizing the nonlocal-means to super-resolution reconstruction," *IEEE Transactions on Image Processing*, vol. 18, no. 1, pp. 36–51, 2009.

[24] X. Qu, Y. Hou, F. Lam, D. Guo, J. Zhong, and Z. Chen, "Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator," *Medical Image Analysis*, vol. 18, no. 6, pp. 843–856, 2014.

[25] J. Ma, J. Huang, Q. Feng et al., "Low-dose computed tomography image restoration using previous normal-dose scan," *Medical Physics*, vol. 38, no. 10, pp. 5713–5731, 2011.

[26] Y. Chen, L. Shi, Q. Feng et al., "Artifact suppressed dictionary learning for low-dose CT image processing," *IEEE Transactions on Medical Imaging*, vol. 33, no. 12, pp. 2271–2292, 2014.

[27] Z. Liao, "A new definition of fractional derivatives based on truncated left-handed Grünwald-Letnikov formula with $0 < \alpha < 1$ and median correction," *Abstract and Applied Analysis*, vol. 2014, Article ID 914386, 9 pages, 2014.

[28] J. Yang, Y. Wang, Y. Liu, S. Tang, and W. Chen, "Novel approach for 3-D reconstruction of coronary arteries from two uncalibrated angiographic images," *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1563–1572, 2009.

[29] Q. Feng, M. Foskey, W. Chen, and D. Shen, "Segmenting CT prostate images using population and patient-specific statistics for radiotherapy," *Medical Physics*, vol. 37, no. 8, pp. 4121–4132, 2010.

[30] J. Yang, Y. Wang, S. Tang, S. Zhou, Y. Liu, and W. Chen, "Multiresolution elastic registration of X-ray angiography images using thin-plate spline," *IEEE Transactions on Nuclear Science*, vol. 54, no. 1, pp. 152–166, 2007.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 172918, 6 pages
http://dx.doi.org/10.1155/2015/172918

Research Article

Estimation of Sensitive Proportion by Randomized Response Data in Successive Sampling

Bo Yu,[1,2] Zongda Jin,[3] Jiayong Tian,[4] and Ge Gao[1]

[1] *School of Public Health, Medical College, Soochow University, Suzhou 215123, China*
[2] *School of Mathematical Sciences, Dezhou University, Dezhou 253023, China*
[3] *Department of Public Health, Zhejiang Medical College, Hangzhou 310053, China*
[4] *Critical Care Medicine, People's Hospital of Linshu County, Linyi, Shandong 276700, China*

Correspondence should be addressed to Ge Gao; gaoge01@163.com

Received 31 October 2014; Revised 22 November 2014; Accepted 5 December 2014

Academic Editor: Yi Gao

This paper considers the problem of estimation for binomial proportions of sensitive or stigmatizing attributes in the population of interest. Randomized response techniques are suggested for protecting the privacy of respondents and reducing the response bias while eliciting information on sensitive attributes. In many sensitive question surveys, the same population is often sampled repeatedly on each occasion. In this paper, we apply successive sampling scheme to improve the estimation of the sensitive proportion on current occasion.

1. Introduction

Social survey sometimes includes stigmatizing or sensitive issues of enquiry, such as habitual tax evasion, sexual behaviour, substance abuse, and excessive gambling that it is difficult to obtain valid and trustworthy information. If the respondents are asked directly about controversial matters, it often results in refusal or untruthful answers, especially when they have committed stigmatizing behaviour. To overcome this difficulty, Warner [1] introduced randomized response techniques to estimate the proportion of people bearing such a stigmatizing or sensitive characteristic in a given community. This technique allows the respondent to answer sensitive questions truthfully without revealing embarrassing behaviour. Following the pioneering work of Warner [1], some researchers have made important contributions in this area, such as Christofides [2, 3], Singh [4], Kim and Elam [5], Huang [6, 7], Singh and Sedory [8], Chang and Kuo [9], Arnab et al. [10]. All these results are based on a sample on one occasion, which is not the case in the present study.

In many sensitive question surveys, the same population is often sampled repeatedly on each occasion, so that the development over time can be followed. In such situations, the use of successive sampling scheme can be attractive alternative to improve the estimators of level at a point in time or to measure the change between two time points. In successive sampling on two occasions, previous theory [11, 12] aimed at providing the optimum estimator of mean on the current (second) occasion. Successive sampling has also been discussed in some detail by Narain [13], Raj [14], Singh [15], Ghangurde and Rao [16], Okafor [17], Arnab and Okafor [18], Biradar and Singh [19], G. N. Singh and V. K. Singh [20], Artes et al. [21], and so forth, and Singh et al. [22]. However no effort has been made to estimate the proportions of sensitive infinite population on the current occasion. This motivation led the authors to consider the problem of estimating the binomial proportions of sensitive or stigmatizing attributes in the population of interest in successive sampling on two occasions. In addition, cluster sampling is usually preferred when the target population is geographically diverse. In this paper, we utilize the rotation cluster sample design to construct a class of estimators for the case of randomized response survey. The rest of the paper is organized as follows. In Section 2, we proposed a new scientific survey method using the Simmons model with cluster rotation sampling. In Section 3, corresponding formulas for the mentioned survey method are

found followed by the aforementioned method and corresponding formulas were successfully designed and applied in a survey of premarital sexual behaviour among students at Soochow University in Section 4. Section 5 contains the conclusion.

2. The Proposed Survey Methods

2.1. Simmons Model. Simmons model which is based on Warner's randomized response technique was put forward by Horvitz et al. [23]. The basic thought is to develop a random rapport between the individuals and two unrelated questions. Simmons design consists of two unrelated questions, A and B, to be answered on probability basis, where A is "do you possess the sensitive characteristic" and B is a nonsensitive question such as "is your birthday number odd or not." The two questions A and B are presented to respondents with preset probabilities P and $1 - P$, respectively. The simple random sampling with replacement (SRSWR) is assumed. The selected respondent is asked to select a question A or B and report "yes" if his/her actual status matches with the selected question and "no" otherwise.

2.2. Simmons Model in Cluster Rotation Sampling. In the following sampling on two occasions is considered to estimate population proportion with a sensitive characteristic on second occasion when the rotation sampling units are clusters. The sampling steps for Simmons model under partial clusters rotation are as follows.

Firstly, the population is divided into primary sampling units (or cluster) and the units within the clusters are the secondary sampling units (persons).

Secondly, in the first occasion a random sample of n clusters with replacement is drawn from the population. The people within the drawn clusters are asked to select a question A or B and report "yes" if his/her actual status matches with the selected question and "no" otherwise, using the Simmons model.

Thirdly, in the second occasion m of the n clusters selected on the first occasion are retained at random and the remaining $u = n - m$ of the clusters are replaced by a fresh selection. All the people within the total clusters in the second occasion are investigated using the Simmons model.

3. Formulas Deduction

3.1. The Estimator of the Population Proportion on the Second Occasion and Its Variance. Consider a random sample of n clusters with replacement drawn from the population which consists of N clusters and the ith cluster of M_i units ($i = 1, 2, \ldots, N$).

In the second (current) occasion m of the n clusters selected on the first occasion are retained at random and the remaining $u = n - m$ of the clusters are replaced by a fresh selection. Let $a_{1,l}$ be the number of the lth retained cluster (including M_l units) with the sensitive characteristic under study on the first occasion ($l = 1, 2, \ldots, m$) and let $a_{1,m+k}$ be the number of the kth rotated cluster (including M_{m+k} units) with the sensitive characteristic under study on the first occasion, respectively ($k = 1, 2, \ldots, u$). $a_{2,l}$ is the number of the lth retained cluster (including M_l units) with the sensitive characteristic under study on the second (current) occasion ($l = 1, 2, \ldots, m$) and $a_{2,m+j}$ is the number of the jth fresh cluster (including M'_{m+j} units) with the sensitive characteristic under study on the second (current) occasion ($j = 1, 2, \ldots, u$). Similarly, let $\pi_{1,l}$ be the proportion of the lth retained cluster with the sensitive characteristic under study on the first occasion ($l = 1, 2, \ldots, m$) and let $\pi_{1,m+k}$ be the proportion of the kth rotated cluster with the sensitive characteristic under study on the first occasion ($k = 1, 2, \ldots, u$), respectively. $\pi_{2,l}$ is the proportion of the lth retained cluster with the sensitive characteristic under study on the second (current) occasion ($l = 1, 2, \ldots, m$) and $\pi_{2,m+j}$ is the proportion of the jth fresh cluster with the sensitive characteristic under study on the second (current) occasion ($j = 1, 2, \ldots, u$). Assume that the variance and the correlation coefficient between the first occasion and second occasion are constant s and the overall correction coefficient is ignored.

Define the following:

π_1: the population proportion of the sensitive characteristic on the first occasion;

π_2: the population proportion of the sensitive characteristic on the second occasion;

π_{1m}: the proportion of m retained clusters with the sensitive characteristic on the first occasion;

π_{2m}: the proportion of m retained clusters with the sensitive characteristic on the second occasion;

π_{1u}: the proportion of u rotated clusters with the sensitive characteristic on the first occasion;

π_{2u}: the proportion of u fresh clusters with the sensitive characteristic on the second occasion.

The following is according to the formula and results given by Cochran [24].

The estimator of π_1 is

$$\hat{\pi}_1 = \frac{\sum_{i=1}^{n} a_{1,i}}{\sum_{i=1}^{n} M_i}. \tag{1}$$

The estimator of π_{1m} is

$$\hat{\pi}_{1m} = \frac{\sum_{i=1}^{m} a_{1,i}}{\sum_{i=1}^{m} M_i}. \tag{2}$$

The estimator of π_{1u} is

$$\hat{\pi}_{1u} = \frac{\sum_{j=1}^{u} a_{1,m+j}}{\sum_{j=1}^{u} M_{m+j}}. \tag{3}$$

The estimator of π_{2m} is

$$\hat{\pi}_{2m} = \frac{\sum_{j=1}^{m} a_{2,j}}{\sum_{j=1}^{m} M_j}. \tag{4}$$

The estimator of π_{2u} is

$$\hat{\pi}_{2u} = \frac{\sum_{j=1}^{u} a_{2,m+j}}{\sum_{j=1}^{u} M'_{m+j}}. \tag{5}$$

Consider a generalized estimator $\hat{\pi}_2$ of the population proportion of the sensitive characteristic on the second occasion or current occasion as

$$\hat{\pi}_2 = a\hat{\pi}_{1u} + b\hat{\pi}_{1m} + c\hat{\pi}_{2u} + d\hat{\pi}_{2m}, \tag{6}$$

where a, b, c, and d are suitable constants.

We have

$$E\left(\hat{\pi}_2\right) = (a+b)\,\pi_1 + (c+d)\,\pi_2. \tag{7}$$

Because the estimator $\hat{\pi}_2$ of π_2 is an unbiased estimator of π_2, we have $a + b = 0$ and $c + d = 0$.

Hence, the estimator (6) takes the form

$$\hat{\pi}_2 = a\left(\hat{\pi}_{1u} - \hat{\pi}_{1m}\right) + c\hat{\pi}_{2u} + (1-c)\,\hat{\pi}_{2m}. \tag{8}$$

The variance of estimator $\hat{\pi}_2$ is

$$\begin{aligned}
\operatorname{Var}\left(\hat{\pi}_2\right) &= a^2 \operatorname{Var}\left(\hat{\pi}_{1u}\right) + a^2 \operatorname{Var}\left(\hat{\pi}_{1m}\right) + c^2 \operatorname{Var}\left(\hat{\pi}_{2u}\right) \\
&\quad + (1-c)^2 \operatorname{Var}\left(\hat{\pi}_{2m}\right) - 2a(1-c)\operatorname{cov}\left(\hat{\pi}_{1m}, \hat{\pi}_{2m}\right) \\
&= a^2\left(\frac{S_1^2}{u}\right) + a^2\left(\frac{S_1^2}{m}\right) + c^2\left(\frac{S_2^2}{u}\right) \\
&\quad + (1-c)^2\left(\frac{S_2^2}{m}\right) - 2a(1-c)\left(\frac{\rho S_1 S_2}{m}\right) \\
&= a^2\left(\frac{S_1^2}{u} + \frac{S_1^2}{m}\right) + c^2\left(\frac{S_2^2}{u}\right) + (1-c)^2\left(\frac{S_2^2}{m}\right) \\
&\quad - 2a(1-c)\left(\frac{\rho_b S_1 S_2}{m}\right).
\end{aligned} \tag{9}$$

Other covariance terms are zero.

Minimizing the variance of estimator $\hat{\pi}_2$ with respect to a and c when N is sufficiently large,

$$\begin{aligned}
\frac{\partial \operatorname{Var}\left(\hat{\pi}_2\right)}{\partial a} &= 2a\left(\frac{1}{u} + \frac{1}{m}\right)S_1^2 - 2(1-c)\,\frac{\rho_b S_1 S_2}{m} = 0, \\
\frac{\partial \operatorname{Var}\left(\hat{\pi}_2\right)}{\partial c} &= 2c\left(\frac{S_2^2}{u}\right) - 2(1-c)\left(\frac{S_2^2}{m}\right) + 2a\left(\frac{\rho_b S_1 S_2}{m}\right) = 0.
\end{aligned} \tag{10}$$

Then we get

$$\begin{aligned}
a\left(\frac{1}{u} + \frac{1}{m}\right)S_1^2 &= (1-c)\,\frac{\rho_b S_1 S_2}{m}, \\
a\left(\frac{\rho_b S_1 S_2}{m}\right) &= c\left(\frac{S_2^2}{u}\right) - (1-c)\left(\frac{S_2^2}{m}\right).
\end{aligned} \tag{11}$$

We derive

$$\frac{1/u + 1/m}{\rho_b/m} = \frac{(1-c)\left(\rho_b/m\right)}{c/u - (1-c)/m}. \tag{12}$$

One has

$$\frac{(u+m)/um}{\rho_b/m} = \frac{(1-c)\left(\rho_b/m\right)}{\left(c(u+m) - u\right)/um}, \tag{13}$$

for $u + m = n$.

We have

$$\frac{n/u}{\rho_b} = \frac{(1-c)\,\rho_b}{(cn-u)/u}. \tag{14}$$

Hence,

$$c = \frac{u/n - \left(u^2/n^2\right)\rho_b^2}{1 - \left(u^2/n^2\right)\rho_b}. \tag{15}$$

Define $\mu = u/n$.

We get

$$c = \frac{\mu - \mu^2\rho_b^2}{1 - \mu^2\rho_b^2}. \tag{16}$$

By (16), we derive

$$a = \frac{(1-c)\left(\rho_b S_2\right)}{(m/u+1)\,S_1}, \tag{17}$$

for $u + m = n$ and $\mu = u/n$.

One has

$$a = \frac{(1-c)\,\mu\left(\rho_b S_2\right)}{S_1}. \tag{18}$$

By (16) and (18), we get

$$a = \frac{S_2}{S_1} \cdot \frac{\mu(1-\mu)\,\rho_b}{1 - \mu^2{\rho_b}^2}, \tag{19}$$

where

$$\begin{aligned}
S_h^2 &= \frac{1}{N-1}\sum_{i=1}^{N}\left(\pi_{h,i} - \pi_h\right)^2, \quad h = 1, 2; \qquad \mu = \frac{u}{n}, \\
\rho_b &= \frac{\sum_{i=1}^{N}\left(\pi_{2,i} - \pi_2\right)\left(\pi_{1,i} - \pi_1\right)}{\sqrt{\sum_{i=1}^{N}\left(\pi_{2,i} - \pi_2\right)^2 \sum_{i=1}^{N}\left(\pi_{1,i} - \pi_1\right)^2}}.
\end{aligned} \tag{20}$$

Theorem 1. *Under the Simmons model in partial clusters rotation, one has*

$$\begin{aligned}
\hat{\pi}_2 &= \frac{S_2}{S_1} \cdot \frac{\mu(1-\mu)\,\rho_b}{1 - \mu^2\rho_b^2}\left(\hat{\pi}_{1u} - \hat{\pi}_{1m}\right) + \left(\frac{\mu - \mu^2\rho_b^2}{1 - \mu^2\rho_b^2}\right)\hat{\pi}_{2u} \\
&\quad + \left(\frac{1-\mu}{1 - \mu^2\rho_b^2}\right)\hat{\pi}_{2m},
\end{aligned} \tag{21}$$

and the variance of estimator $\hat{\pi}_2$ is

$$\operatorname{Var}_{\min}\left(\hat{P}_2\right) = \left[\frac{1 - \mu\rho_b^2}{1 - \mu^2\rho_b^2}\right]\frac{S_2^2}{n}. \tag{22}$$

Remark 2. In practice, the ρ_b and S_h^2 are unknown. The estimator of ρ_b is

$$\widehat{\rho}_b = \frac{\sum_{k=1}^{m} (\pi_{2,k} - \widehat{\pi}_{2m})(\pi_{1,k} - \widehat{\pi}_{1m})}{\sqrt{\sum_{k=1}^{m} (\pi_{2,k} - \widehat{\pi}_{2m})^2 \sum_{k=1}^{m} (\pi_{1,k} - \widehat{\pi}_{1m})^2}}. \quad (23)$$

And the estimator of S_h^2 is

$$s_h^2 = \frac{1}{m-1} \sum_{i=1}^{m} (\pi_{h,i} - \widehat{\pi}_{hm})^2, \quad (h = 1, 2). \quad (24)$$

Theorem 3. *Under the Simmons model in partial clusters rotation, one has the optimum rotation rate as*

$$\mu = \left[1 + \sqrt{1 - \rho_b{}^2}\right]^{-1}. \quad (25)$$

And the optimum variance of estimator $\widehat{\pi}_2$ is

$$\mathrm{Var}_{\min(\mathrm{opt})}(\widehat{\pi}_2) = \left[1 + \sqrt{1 - \rho_b^2}\right] \frac{S_2^2}{2n}. \quad (26)$$

Practically, the costs of sample survey usually represent the following simple function, according to Cochran [24]:

$$C_T = c_0 + c_1 m + c_2 u, \quad (27)$$

where C_T is the total cost of sampling, c_0 is the fundamental cost of the survey, c_1 is the average fundamental cost of investigating one retained cluster on the second occasion, and c_2 is the average fundamental cost of investigating one fresh cluster on the second occasion.

Theorem 4. *Under the given cost of sample survey C_T, one has*

$$\mathrm{Var}_{\mathrm{opt}}(\widehat{\pi}_2) = \frac{\left(c_1 \sqrt{1 - \rho_b^2} + c_2\right) S_2^2}{2(C_T - c_0)}. \quad (28)$$

And the estimation of sample size in partial clusters rotation is

$$n' = \frac{M(C_T - c_0)\left(1 + \sqrt{1 - \rho_b^2}\right)}{c_1 \sqrt{1 - \rho_b^2} + c_2}, \quad (29)$$

where

$$S_2^2 = \frac{1}{N-1} \sum_{i=1}^{N} (\pi_{2,i} - \pi_2)^2. \quad (30)$$

3.2. The Estimator of $\pi_{h,i}$. Let $R_{h,i}$ be the proportion of the selected ith cluster (including M_i units) with the nonsensitive characteristic under study on the hth occasion; $\lambda_{h,i}$ and $m_{h,i}$ denote the number and the proportion of "yes" answers in the ith cluster, respectively, where $\widehat{\lambda}_{h,i} = m_{h,i}/M_i$, $h = 1, 2$, $(i = 1, 2, \ldots, n)$.

From the total probability formulas (see [25]), we can get

$$\pi_{h,i} = \frac{\lambda_{h,i} - (1 - P) R_{h,i}}{P}, \quad h = 1, 2, \quad i = 1, 2, \ldots, n. \quad (31)$$

Hence

$$a_{h,i} = M_i \pi_{h,i}, \quad h = 1, 2, \ i = 1, 2, \ldots, n. \quad (32)$$

4. Applications

4.1. Survey Design. The survey is about premarital sexual behavior among students in Dushu Lake Campus of Soochow University. We regard every class as a cluster of 45 persons per class on average. In the first occasion (2011), 12 classes were drawn from all the classes randomly. All the persons in the selected 12 classes are surveyed by Simmons model for sensitive questions. In the second occasion (2013), 8 of the 12 classes selected on the first occasion are retained at random and the remaining 4 classes are replaced by a fresh selection. Then all the persons in the selected classes that consist of 8 retained classes and 4 fresh classes are surveyed by Simmons model for sensitive questions.

In our design, each person was asked to draw a ball at random with replacement from a bag containing 6 red balls and 4 white balls with known probability (the proportion of red balls was 0.6). If a red ball was selected by the respondent, then he or she would be asked the sensitive question A, where A is "are you a member of the group having premarital sexual behavior." If a white ball was selected, he or she would answer the nonsensitive question B, where B is "is your student number odd or not." The respondent reports "yes" if his/her actual status matches with the selected question and "no" otherwise.

All the questionnaires of two occasions had been checked to ensure that they are completed independently and no questions were omitted. The recovery rate of the survey was 100% with no failure questionnaire. All data was processed and analyzed by Excel 2003 and SAS 9.13.

4.2. Results

4.2.1. Result of the Survey. In our design, each person was asked to draw a ball at random with replacement from a bag containing 6 red balls and 4 white balls with known probability (the proportion of red balls was 0.6). If a red ball was selected by the respondent, then he or she would be asked the sensitive question A, where A is "are you a member of the group having premarital sexual behavior." If a white ball was selected, he or she would answer the nonsensitive question B, where B is "is your student number odd or not." The respondent reports "*yes*" if his/her actual status matches with the selected question and "*no*" otherwise. According to (31), we get the sample proportion of the undergraduate students who have premarital sexual behavior $\pi_{h,i}$, $h = 1, 2$, as is shown in Table 1.

4.2.2. The Estimator of the Population Proportion on the Second Occasion and Its Variance. By (1), the estimator of the population proportion with premarital sexual behavior on the first occasion is as follows: $\widehat{\pi}_1 = 0.142933$.

According to (24), (2), and (3), we have

$$s_1^2 = 0.005853, \quad \widehat{\pi}_{1m} = 0.1458, \quad \widehat{\pi}_{1u} = 0.1372, \quad (33)$$

respectively.

Table 1: The proportion of every class with premarital sexual behavior in first and second occasions.

Class number	$\pi_{1,i}$	$\pi_{2,i}$
1	0.2624	0.2348
2	0.1631	0.1945
3	0.2101	0.2264
4	0.2063	
5	0.1556	0.1986
6	0.2390	
7	0.1783	
8	0.1970	0.1550
9	0.0123	0.0114
10	0.0476	0.0738
11	0.0455	
12	0.1185	0.1187
13		0.2035
14		0.1587
15		0.1926
16		0.1583

According to the results of investigation premarital sexual behavior among students in Dushu Lake Campus of Soochow University on the second occasion, from formulae (4) and (5),

$$\widehat{\pi}_{2m} = 0.1562, \qquad \widehat{\pi}_{2u} = 0.1783. \tag{34}$$

By (23) and (24), we obtain $S_2^2 = 0.004262$ and $\widehat{\rho}_b = 0.936$, respectively.

From formula (25), $\mu = 0.7985$; then according to formula (21), we get

$$\begin{aligned}\widehat{\pi}_2 &= 0.2912\,(0.1372 - 0.1458) + 0.5453 \times 0.1783 \\ &\quad + (1 - 0.5453) \times 0.1562 = 0.1657.\end{aligned} \tag{35}$$

Using (22), we get $\mathrm{Var}_{\min(\mathrm{opt})}(\widehat{\pi}_2) = 0.00024$. Hence, standard deviation is as follows:

$$\sqrt{\mathrm{Var}_{\min(\mathrm{opt})}\left(\widehat{\pi}_2\right)} = 0.01549. \tag{36}$$

So, 95% confidence interval of the population proportion with the premarital sexual is [0.1353, 0.1961].

5. Discussion and Conclusion

To sum up, in this study, we proposed a new sampling method to solve the question of sensitive questions surveys repeated over time, which is the first attempt made by the authors in this direction. Then the corresponding formulas for the estimator of the population proportion with sensitive characteristic and its variance for the proposed sampling method are provided. In addition, formulas for the optimal rotation rate and sample size under the given cost of sample survey are given.

The aforementioned method and corresponding formulas were successfully designed and applied in the premarital sex survey in Dushu Lake Campus of Soochow University. In a word, the designed sampling method and corresponding formulas have important theory and application value to achieve the sensitive questions continuous survey.

6. Proofs of Theorems

Proof of Theorem 1. Using the optimum values of a and c given by (16) and (19), estimator $\widehat{\pi}_2$ reduces to (21).

By (9), (16), and (19), we have

$$\mathrm{Var}_{\min}\left(\widehat{P}_2\right) = \left[\frac{1 - \mu{\rho_b}^2}{1 - \mu^2{\rho_b}^2}\right]\frac{S_2^2}{n}. \tag{37}$$

□

Proof of Theorem 3. The optimum value of μ is given by further minimizing (22) with respect to μ,

$$\frac{\partial \mathrm{Var}_{\min}\left(\widehat{\pi}_2\right)}{\partial \mu} = 0. \tag{38}$$

So

$$\mu = \left[1 + \sqrt{1 - \rho_b^2}\right]^{-1}. \tag{39}$$

Substituting (39) in (22), we have the optimum variance of estimator $\widehat{\pi}_2$ as

$$\mathrm{Var}_{\min(\mathrm{opt})}\left(\widehat{\pi}_2\right) = \left[1 + \sqrt{1 - \rho_b^2}\right]\frac{S_2^2}{2n}. \tag{40}$$

□

Proof of Theorem 4. ByTheorem 3,

$$\mu = \left[1 + \sqrt{1 - \rho_b^2}\right]^{-1}, \quad \mu = \frac{u}{n},\ u = n - m. \tag{41}$$

Substituting (41) in (27), we obtain

$$n = \frac{\left(C_T - c_0\right)\left(1 + \sqrt{1 - \rho_b^2}\right)}{c_1\sqrt{1 - \rho_b^2} + c_2}. \tag{42}$$

Suppose the average cluster consists of $\overline{M}$ units; then

$$n' = \frac{\overline{M}\left(C_T - c_0\right)\left(1 + \sqrt{1 - \rho_b^2}\right)}{c_1\sqrt{1 - \rho_b^2} + c_2}. \tag{43}$$

Substituting (42) in (26), we have

$$\mathrm{Var}_{\mathrm{opt}}\left(\widehat{\pi}_2\right) = \frac{\left(c_1\sqrt{1 - \rho_b^2} + c_2\right)S_2^2}{2\left(C_T - c_0\right)}. \tag{44}$$

□

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to express their deep thanks to the related referees for carefully reading the paper and for comments which greatly improved the paper. This paper is supported by a grant from the National Natural Science Foundation of China (no. 81273188 to G. Gao). The authors are grateful to G. Gao (corresponding author) for his invaluable help.

References

[1] S. L. Warner, "Randomized response: a survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69, 1965.

[2] T. C. Christofides, "A generalized randomized response technique," *Metrika*, vol. 57, no. 2, pp. 195–200, 2003.

[3] T. C. Christofides, "Randomized response in stratified sampling," *Journal of Statistical Planning and Inference*, vol. 128, no. 1, pp. 303–310, 2005.

[4] G. N. Singh, "On the use of chain-type ratio to difference estimator in successive sampling," *International Journal of Applied Mathematics and Statistics*, vol. 6, pp. 41–49, 2006.

[5] J.-M. Kim and M. E. Elam, "A stratified unrelated question randomized response model," *Statistical Papers*, vol. 48, no. 2, pp. 215–233, 2007.

[6] K.-C. Huang, "Estimation for sensitive characteristics using optional randomized response technique," *Quality and Quantity*, vol. 42, no. 5, pp. 679–686, 2008.

[7] K.-C. Huang, "Unbiased estimators of mean, variance and sensitivity level for quantitative characteristics in finite population sampling," *Metrika*, vol. 71, no. 3, pp. 341–352, 2010.

[8] S. Singh and S. A. Sedory, "A true simulation study of three estimators at equal protection of respondents in randomized response sampling," *Statistica Neerlandica*, vol. 66, no. 4, pp. 442–451, 2012.

[9] H.-J. Chang and M.-P. Kuo, "Estimation of population proportion in randomized response sampling using weighted confidence interval construction," *Metrika*, vol. 75, no. 5, pp. 655–672, 2012.

[10] R. Arnab, S. Singh, and D. North, "Use of two decks of cards in randomized response techniques for complex survey designs," *Communications in Statistics. Theory and Methods*, vol. 41, no. 16-17, pp. 3198–3210, 2012.

[11] R. J. Jessen, "Statistical investigation of a survey for obtaining farm facts," *Iowa Agricultural Experiment Station Research Bulletin*, vol. 304, pp. 1–104, 1942.

[12] F. Yates, *Sampling Methods for Censuses and Surveys*, Charles Griffin, London, UK, 1949.

[13] R. D. Narain, "On the recurrence formula in sampling on successive occasions," *Journal of the Indian Society of Agricultural Statistics*, vol. 5, pp. 96–99, 1953.

[14] D. Raj, "On sampling over two occasions with probability proportionate to size," *Annals of Mathematical Statistics*, vol. 36, pp. 327–330, 1965.

[15] D. Singh, "Estimates in successive sampling using a multi-stage design," *Journal of the American Statistical Association*, vol. 63, pp. 99–112, 1968.

[16] P. D. Ghangurde and J. N. Rao, "Some results on sampling over two occasions," *Sankhya*, vol. 31, pp. 463–472, 1969.

[17] F. C. Okafor, "The theory and application of sampling over two occasions for the estimation of current population ratio," *Statistica*, vol. 42, pp. 137–147, 1992.

[18] R. Arnab and F. C. Okafor, "A note on double sampling over two occasions," *Pakistan Journal of Statistics*, vol. 8, no. 3, pp. 9–18, 1992.

[19] R. S. Biradar and H. P. Singh, "Successive sampling using auxiliary information on both the occasions," *Calcutta Statistical Association Bulletin*, vol. 51, no. 203-204, pp. 243–251, 2001.

[20] G. N. Singh and V. K. Singh, "On the use of auxiliary information in successive sampling," *Journal of the Indian Society of Agricultural Statistics*, vol. 54, no. 1, pp. 1–12, 2001.

[21] R. Artes, M. Eva, L. Garcia, and V. Amelia, "Estimation of current population ratio in successive sampling," *Journal of the Indian Society of Agricultural Statistics*, vol. 54, no. 3, pp. 342–354, 2001.

[22] H. P. Singh, R. Tailor, S. Singh, and J.-M. Kim, "Estimation of population variance in successive sampling," *Quality & Quantity*, vol. 45, no. 3, pp. 477–494, 2011.

[23] D. G. Horvitz, B. V. Shah, and W. R. Simmons, "The unrelated question randomized response model," *Proceedings of the Social Statistics Section: American Statistical Association*, vol. 326, pp. 65–72, 1967.

[24] W. G. Cochran, *Sampling Techniques*, John Wiley & Sons, New York, NY, USA, 3rd edition, 1977.

[25] Z. Du, *Sampling Techniques and Its Application*, Tsinghua University Press, Beijing, China, 1st edition, 2005.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 868493, 20 pages
http://dx.doi.org/10.1155/2015/868493

Research Article

Finger Vein Segmentation from Infrared Images Based on a Modified Separable Mumford Shah Model and Local Entropy Thresholding

Marios Vlachos and Evangelos Dermatas

Department of Electrical & Computer Engineering, Polytechnic Faculty, University of Patras, Rio Campus, 26504 Patras, Greece

Correspondence should be addressed to Marios Vlachos; mvlachos@george.wcl2.ee.upatras.gr

Received 3 September 2014; Accepted 26 January 2015

Academic Editor: Yi Gao

A novel method for finger vein pattern extraction from infrared images is presented. This method involves four steps: preprocessing which performs local normalization of the image intensity, image enhancement, image segmentation, and finally postprocessing for image cleaning. In the image enhancement step, an image which will be both smooth and similar to the original is sought. The enhanced image is obtained by minimizing the objective function of a modified separable Mumford Shah Model. Since, this minimization procedure is computationally intensive for large images, a local application of the Mumford Shah Model in small window neighborhoods is proposed. The finger veins are located in concave nonsmooth regions and, so, in order to distinct them from the other tissue parts, all the differences between the smooth neighborhoods, obtained by the local application of the model, and the corresponding windows of the original image are added. After that, veins in the enhanced image have been sufficiently emphasized. Thus, after image enhancement, an accurate segmentation can be obtained readily by a local entropy thresholding method. Finally, the resulted binary image may suffer from some misclassifications and, so, a postprocessing step is performed in order to extract a robust finger vein pattern.

1. Introduction

The problem of finger vein extraction from infrared images arises mainly for biometrics purposes but it is also very important for the biomedical research community.

The general structure of a biometric system based on finger veins consists of five main stages: (1) *acquisition* of the infrared images exploiting the absorption of light in near infrared and infrared wavelengths by the different human tissues, (2) *preprocessing* of the acquired images which includes *ROI* (region of interest) extraction, image intensity normalization (in this type of images intensity is usually uneven due to the image acquisition system and may suffer from shading artefacts), and noise reduction, (3) *segmentation or classification* stage in which the preprocessed image divided into two (or more depending on the application) regions associated with veins and surrounding tissues, (4) *postprocessing* of the binary images which delivers the final segmentation result, free of outliers and misclassifications, and finally (5) *matching* of the extracted veins in order to perform the desired person identification/verification procedure. Matching procedure can be applied either directly in the extracted finger vein patterns or in their skeletons, depending on the matching algorithm that has to be used. This general structure described so far involves all the stages that may have such a system but it is worth mentioning that these stages are independent and some of them can be skipped in some applications depending on its specific requirements.

Related Work. Several methods which adopt this general architecture have already been presented starting from the pioneering work of Park et al. [1]. In this important research work, an application specific processor for vein pattern extraction and its application to a biometric identification system is proposed. The conventional vein pattern recognition algorithm [1–3] consists of a preprocessing part, applying

sequentially an iterative Gaussian low pass, a high pass, and a modified median filter, a recognition part which includes the extraction of the binary veins via local thresholding, and finally the matching between the individual patterns.

An improved vein pattern extraction algorithm is proposed in [4], which compensates the loss of vein patterns in the edge area, gives more enhanced and stabilized vein pattern information, and shows better performance than the existing algorithms. Also, the problem arising from the iterative nature of filtering preprocess is solved by designing a filter that is processed only once, increasing significantly the recognition speed and reducing the hardware complexity. The proposed algorithm is implemented with an *FPGA* device and the false acceptance rate is five times better than the existing algorithm and the recognition speed is measured to be 100 (ms/person).

The problem with conventional hand vascular technology mentioned above is that the vascular pattern is extracted without taking into account its direction. So, there is a loss of vascular connectivity which leads to a degradation of the performance of the verification procedure. An attempt to improve this problem can be found in [5], where a direction based vascular pattern extraction algorithm based on the directional information of vascular patterns is presented for biometric applications. It applies two different filters: row vascular pattern extraction filter for abscissa vascular pattern extraction and column vascular pattern extraction filter for effective extraction of the ordinate vascular patterns. The combined output of both filters produces the final hand vascular patterns. Unlike the conventional hand vascular pattern extraction algorithm, the directional extraction approach prevents loss of the vascular pattern connectivity.

Although the above algorithm considers the directionality of veins, it also assumes that the veins oriented in only two principal directions. In [6, 7] a method for personal identification based on finger vein patterns is presented and evaluated using line tracking starting from various positions. This method allows vein patterns to have an arbitrary direction. Local dark lines are identified and line tracking is executed by moving along the lines pixel by pixel. When a dark line is not detectable, a new tracking operation starts at another position. This procedure executes repeatedly, so the dark lines that tracked multiple times are classified as veins.

Typically, the infrared images of finger veins are low contrast images, due to the light scattering effect. An algorithm for finger vein pattern extraction in infrared images is proposed in [8]. This algorithm embeds all the above issues and proposes novel preprocessing and postprocessing algorithms. Initially, the image is enhanced and the fingerprint lines are removed using *2D* discrete wavelet filtering. Kernel filtering produces multiple images by rotating the kernel in six different directions, focus on the expected directions of the vein patterns. The maximum of all images is transformed into a binary image. Further improvement is achieved by a two-level morphological process; that is, a majority filter smoothes the contours and removes some of the misclassified isolated pixels, and a reconstruction procedure removes the remaining misclassified regions. The final image is segmented into two regions, the vein and the tissue.

In [9] new issues are considered and a certification system that compares vein images for low cost, high speed, and high precision certification is proposed. The equipment for authentication consists of a near infrared light source and a monochrome *CCD* to produce contrast enhanced images of the subcutaneous veins. The phase correlation and template matching methods are used for classification. Several noise reduction filters, sharpness filters, and histogram manipulations tested for the best effort. As a result, a high certification ratio in this system obtained.

In [10], the theoretical foundation and difficulties of hand vein recognition are introduced at first. Then, the optimum threshold of the segmentation process and the vein lines thinning problem of infrared hand images are deeply studied, followed by the presentation of a novel estimator for the segmentation threshold and an improved conditional thinning method. The method of hand vein image feature extraction based on end points and crossing points is studied initially, and the matching method based on a distance measure is used to match vein images. The matching experiments indicated that this method is efficient in terms of biometric verification.

An efficient automatic method for robust segmentation of finger vessel network and vein pattern extraction from infrared images acquired by a low-cost monochrome or multichannel camera is proposed in [11]. After brightness normalization, the fingerprint lines are eliminated using the 2D dimensional discrete wavelet transformation. A set of twelve directional kernels is constructed, based on a dyadic wavelet transform, for each scale, and is used to enhance the directional properties of veins. From maximum filters' response along scale and direction, a neighbourhood thresholding derives a binary segmented image to produce reliable patterns of finger veins. A postprocessing module is used in case where low quality images are to be segmented. Preliminary evaluation experiments of the proposed method demonstrate a number of advantages, compared to recently published methods.

In the narrow bandwidth of the near infrared spectrum, the light is propagated through human tissue with low absorption rates, but strong scattering effects produce extremely low contrast images. In [12], an algorithm for finger vein segmentation and centerlines extraction in infrared images is presented. Finger veins are detected in pixels with positive divergence of the gradient vector field, while centerlines are extracted in pixels with positive divergence of the normalized gradient vector field estimated at various orientations. The segmentation algorithm has been evaluated on both artificial and real finger infrared images and high segmentation rates are achieved in terms of sensitivity, specificity, and accuracy using manual annotation data obtained by human observers.

A new algorithm for vein matching based on log-polar transform to address problems that occur with the changing of finger position and from differences between imaging devices for current vein matching algorithms is discussed in [13]. The new algorithm first extracts the feature area, which contains enough characteristics for image matching, depending on the structure of the finger vein ridge alignment. It then calculates the degree of similarity between the log-polar transform results of the model image feature areas and

the sample image and finally analyzes the result by the degree of similarity and the relationship of relative positions between feature areas. Experiments show that the algorithm is robust for rotating and zooming images of the finger vein.

In [14], four principles (caliber uniformity, node replication, loop splitting, and virtual connection) are proposed, first to simplify the finger vein structure as a binary tree structure. Then a modified binary tree model is proposed based on the binary tree structure. The new model uses the binary tree to describe the relationships between different vein branches and uses a B-spline function to describe the spatial structure of vein branches. Experiments show that this model can quantitatively describe the relationships between, and the spatial structure of, vein branches with little representation error and low storage space requirements.

The method proposed in [15] is rooted in a local binary pattern based method and then inclined to use the best bits only for matching. After presenting the concept of PBBM and the generating algorithm, authors propose the finger vein recognition framework, which consists of preprocessing, feature extraction, and matching. Experimental results show that PBBM achieves not only better performance but also high robustness and reliability. In addition, PBBM can be used as a general framework for binary pattern based recognition.

In [16], a new and robust approach for finger vein ROI localization is introduced, and then a new scheme for effectively improving the visibility of finger vein imageries is proposed. Extensive experiments are conducted to validate this method.

A new method of personal identification based on finger vein recognition is presented in [17]. First, a stable region representing finger vein network is cropped from the image plane of an imaging sensor. A bank of Gabor filters is then used to exploit the finger vein characteristics at different orientations and scales. Based on the filtered image, both local and global finger vein features are extracted to construct a finger vein code (FVCode). Finally, finger vein recognition is implemented using the cosine similarity measure classifier, and a fusion scheme in decision level is adopted to improve the reliability of identification. Experimental results show that this method exhibit an exciting performance in personal identification.

Finally in [18], a finger vein system using the mean curvature that can be used for personal verification is proposed. As a robust extraction method, authors propose the mean curvature method, which views the vein image as a geometric shape and finds the valley-like structures with negative mean curvatures. When the matched pixel ratio is used in matching vein patterns, experimental results show that, while maintaining low complexity, the proposed method achieves 0.25% equal error rate, which is significantly lower than what existing methods can achieve.

However, the finger vein technology, as mentioned above, has also important applications in biomedical field. An initial work for localizing surface veins via near infrared (*NIR*) imaging and structured light ranging is presented in [19]. The eventual goal of the system is to serve as the guidance for a fully automatic (i.e., robotic) catheterization device. The proposed system is based upon near infrared (*NIR*) imaging, which has previously been shown effective in enhancing the visibility of surface veins. The vein regions in the *2D NIR* images located using standard image processing techniques. An *NIR* line generating *LED* module is used for to implement structured light ranging and construct a *3D* topographic map of the arm surface. The located veins are mapped to the arm surface to provide a camera registered representation of the arm and veins.

Also in [20, 21], a vein contrast enhancer (*VCE*) has been constructed to make vein access easier by capturing an infrared image of veins, enhancing the contrast using software, and projecting the vein image back onto the skin. The *VCE* also uses software to align the projected image with the original vein and with accuracy of 0.06 mm. Clinical evaluation of earlier monitor based vein enhancement test systems has demonstrated the clinical utility of the infrared imaging technology used in the *VCE*.

Although these methods achieve segmenting the infrared images, the finger vein pattern extraction task is still challenging mainly due to the fact that infrared images suffer from strong noise presence, uneven illumination, and shading, factors that complicate the application of automatic image segmentation techniques. Thus, another way to segment this kind of images is to assume that veins located in thin and concave regions (a reasonable assumption obtained by a careful inspection of the image intensity across the image) of infrared images based on this concept to extract them by optimizing a mathematical model. This can be done by using the Mumford Shah Model which has well-known capabilities in the image processing applications such as image segmentation, restoration, and image inpainting [22, 23]. Thus, in this paper, an analytical solution to a modified Mumford Shah Model minimization problem is derived and a local application of its results, in order to perform fast and accurate finger vein extraction, is proposed.

The remainder of this paper is organized as follows. In Section 2 the experimental device and the image acquisition procedure is presented. In Section 3, a detailed presentation of the finger vein pattern extraction method is given. The experimental results and discussion are included in Section 4. Finally, the most significant conclusions and some directions for future work are presented in the last section of this paper.

2. Image Acquisition

A typical hardware used to acquire infrared images consists of a finger probe, an array of infrared leds with adjustable illumination, and a video camera focus on frame, as shown in Figure 1. The finger *ROI* was placed inside the probe, between the open frame and the array of infrared leds light source which consists of a number of leds with adjustable illumination. The finger probe eliminates the influence of the external light sources.

The acquired image is produced as a result of several physical phenomena that happen during light propagation through human tissue, that is, absorption, diffusion, and scattering [24]. The great number of substances contained in the human body, the blood dynamics, and the mass transfer

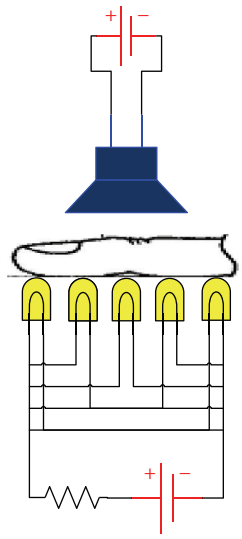


Figure 1: A typical low-cost device used for digital image acquisition of finger infrared images.

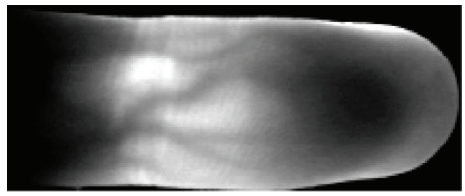


Figure 2: Original image.

phenomena complicates significantly the light transformation effects. Therefore, the solution of the inverse problem, that is, the derivation of the arterial network from the image data, becomes unrealistic. A totally different and popular approach, adopted also in the proposed method, uses several image enhancements, feature extraction, and path reconstruction methods to derive the vein network, based on the fact that substance haemoglobin presents strong absorption in the infrared wavelengths, and therefore the veins appear in the image darker than the other human tissues (Figure 2).

3. Detection of Vein Network

In Figure 3, a flowchart of the proposed vein extraction method is given.

3.1. Image Preprocessing

3.1.1. ROI Extraction. From the original image (Figure 2) a region of interest (*ROI*) is defined based on several statistical properties of the histogram; that is, very low or very high brightness areas are excluded from the *ROI*. In the designed hardware each infrared led has adjustable intensity, giving excellent image quality, minimizing also the variance of the automatic exposure times of the image acquisition system.

The acquired image suffers from shading and nonuniform illumination both in left side and in right side of the image. This effect usually influences the performance of automatic image processing methods applied in order to extract the finger vein pattern. Thus, *ROI* extraction is used in order to localize the finger region and to isolate the shading artifacts. In this paper, the method proposed in [25] is adopted for *ROI* extraction. This method is based on the cutoff of regions with shading taking care about the different dimensions of the fingers among each person. Two masks, one for x and one for y direction, are used to isolate the boundary and localize the effective finger region. A typical *ROI* extracted by the method proposed in [25] is shown in Figure 4.

3.1.2. Brightness Normalization Based on Local Statistical Measures. In general, even in case where led's intensity is adjusted to satisfy several statistical properties, in few areas of the acquired image unsatisfactory illumination or strong noise distortions are met. Therefore an image normalization procedure is applied to restore partially the desirable characteristics.

The proposed local normalization procedure unifies the local mean and variance of the *ROI*, especially useful technique for correcting nonuniform illumination or shading artifacts, using a linear transformation scheme applied on pixels' brightness,

$$u_0(x,y) = \frac{r(x,y) - m_r(x,y)}{\sigma_r(x,y)}, \tag{1}$$

where $r(x, y)$ is the brightness of the original *ROI* image at pixel (x, y), $m_r(x, y)$ is the brightness local mean, $\sigma_r(x, y)$ is the corresponding local standard deviation, and $u_0(x, y)$ is the normalized image. The estimation of local mean and standard deviation is performed inside small window neighborhoods by averaging pixel intensities, a process also known as spatial smoothing.

3.2. Minimization of the Mumford Shah Model. The human veins in finger are significantly thinner than the darker structures observed in typical infrared images, as shown in Figure 2. Multiple scattering of the propagated photons reduces significantly the contrast, eliminates the tiny veins, and increases the transition regions between the vein and the surrounding tissue. The "fog" effect hides the vein structures in concave regions of the *ROI*. This assumption could be verified by observing the cross section profile of the veins which is Gaussian-like, as claimed in [26]. The aim of the proposed system is to focus on the concave regions enhancement, based on several connectivity properties in order to simply separate them from the rest tissue by a local entropy thresholding technique.

The concave regions are regions in the image domain where the second order derivative is positive. Direct estimation of the derivatives in digital images is an ill-posed problem due to noise presence and the variations in illumination. Instead of seeking regions which have positive second order derivatives, the minimization of an objective function similar to the objective function of the Mumford Shah Model [27] is proposed. The objective of the Mumford Shah Model is to estimate a smooth function u such that it is similar to

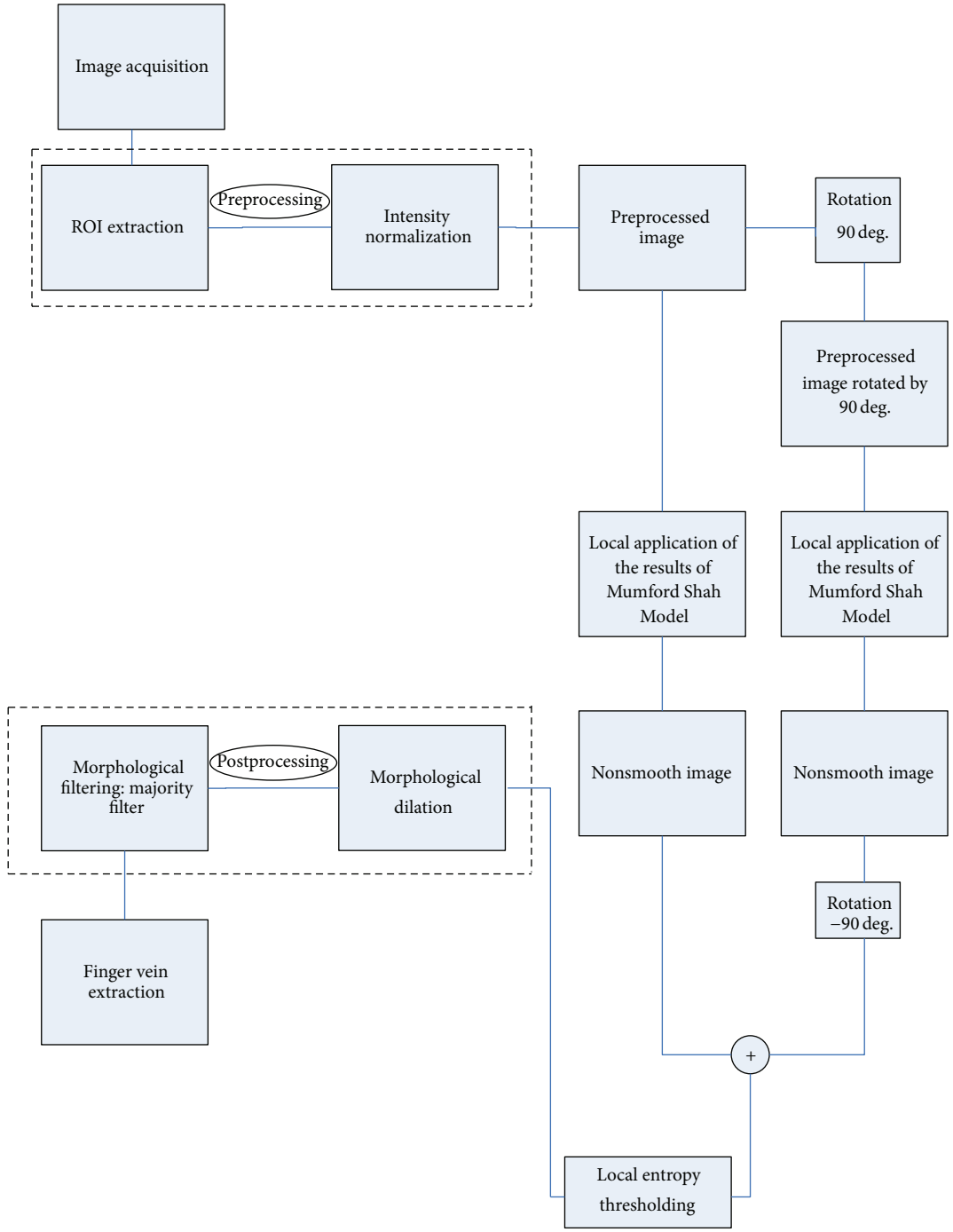


Figure 3: Flowchart of the proposed vein detection method.

the original image u_0. The equivalent mathematical expression leads to the minimization problem of the following objective function:

$$J(u) = \int_{\Omega} |\nabla u|^2 \, dx + \lambda \int_{\Omega} |u - u_0|^2 \, dx, \tag{2}$$

where Ω is the image domain and λ is a user defined parameter. The minimization of this function is computationally intensive and can be performed by the method proposed by Chan and Shen [22]. This method belongs to the category of segmentation methods which use partial differential equations (*PDE*) and it is iterative. Instead of deriving the global minimum of (2), in this paper a close form solution of a discrete objective function (3) similar to the objective function of the Mumford Shah Model is proposed by processing small rectangular areas. The proposed solution accelerates

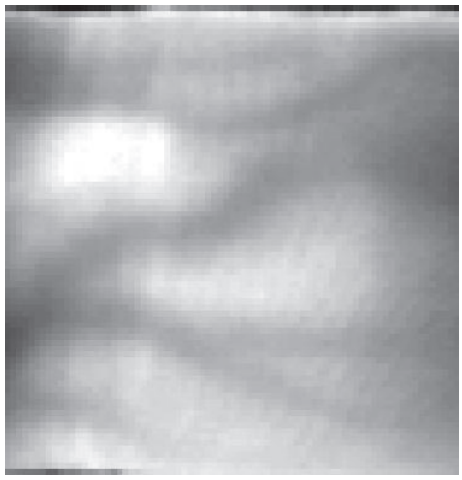


Figure 4: Region of interest extraction (*ROI*).

significantly the processing time and outperforms the classical approach [22]. As a result, fast and accurate extraction of the finger veins is obtained.

Assuming, without loss of generality, that the original image has N rows and N columns (however the image can have different dimensions along the two axes), the function $J(u)$ in the discrete space is defined as follows:

$$J(u) = \frac{1}{2} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} |\nabla u|^2 + \frac{\lambda}{2} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} \left|u - u_0\right|^2, \quad (3)$$

where ∇u is the gradient of the image $u(\cdot, \cdot)$ and can be written as

$$|\nabla u|^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2. \quad (4)$$

If the partial derivatives in (4) are approximated using local differences,

$$\begin{aligned} \frac{\partial u}{\partial x} &= u(x+1, y) - u(x, y), \\ \frac{\partial u}{\partial y} &= u(x, y+1) - u(x, y), \end{aligned} \quad (5)$$

and (4) is substituted in (3) the following formula is obtained:

$$\begin{aligned} J(u) = \frac{1}{2} \cdot \sum_{x=1}^{N-1} \sum_{y=1}^{N-1} &\left[\left(u(x+1, y) - u(x, y)\right)^2 \right. \\ &\left. + \left(u(x, y+1) - u(x, y)\right)^2\right] \\ + \frac{\lambda}{2} \cdot \sum_{x=1}^{N-1} \sum_{y=1}^{N-1} &\left[u(x, y) - u_0(x, y)\right]^2. \end{aligned} \quad (6)$$

The minimum of this objective function (6) regarding $u(\cdot, \cdot)$ can be derived in a close form by differentiating the second order, positively defined function:

$$\begin{aligned} \frac{\partial J(u)}{\partial u} = \frac{1}{2} \cdot \sum_{x=1}^{N-1} \sum_{x=1}^{N-1} \frac{\partial}{\partial u} &\left[\left(u(x+1, y) - u(x, y)\right)^2 \right. \\ &\left. + \left(u(x, y+1) - u(x, y)\right)^2\right] \\ + \frac{\lambda}{2} \cdot \sum_{x=1}^{N-1} \sum_{x=1}^{N-1} \frac{\partial}{\partial u} &\left[u(x, y) - u_0(x, y)\right]^2 = 0 \end{aligned} \quad (7)$$

$$\begin{aligned} \Longrightarrow\ & (\lambda + 4) \cdot u(x, y) - u(x+1, y) \\ & - u(x, y+1) - u(x, y-1) - u(x-1, y) \\ & = \lambda \cdot u_0(x, y), \quad \forall x, y \in [1, N-1]. \end{aligned}$$

As it can easily be observed, (7) is applied for pixels $\forall(x, y) \in [1, N-1]$ in order to avoid the boundary treatment problem for the finite difference scheme. Thus, the necessary image points are available for this type of approximation scheme. Obviously, this approach has a truncation error but, as proved experimentally, it has negligible impact in the performance of the proposed method.

Equation (7) can be rewritten in matrix form as

$$\begin{bmatrix} \lambda+4 & -1 & 0 & 0 & \cdot & -1 & \cdot & 0 \\ -1 & \lambda+4 & -1 & 0 & \cdot & \cdot & -1 & 0 \\ 0 & -1 & \lambda+4 & -1 & \cdot & \cdot & \cdot & -1 \\ 0 & 0 & -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & 0 \\ \cdot & -1 & \cdot & \cdot & \cdot & -1 & \lambda+4 & -1 \\ 0 & 0 & -1 & 0 & \cdot & 0 & -1 & \lambda+4 \end{bmatrix}$$

$$\cdot \begin{bmatrix} u(1,1) \\ u(1,2) \\ \cdot \\ u(1,N) \\ u(2,1) \\ u(2,2) \\ \cdot \\ u(2,N) \\ \cdot \\ u(N,1) \\ \cdot \\ u(N,N) \end{bmatrix} = \lambda \cdot \begin{bmatrix} u_0(1,1) \\ u_0(1,2) \\ \cdot \\ u_0(1,N) \\ u_0(2,1) \\ u_0(2,2) \\ \cdot \\ u_0(2,N) \\ \cdot \\ u_0(N,1) \\ \cdot \\ u_0(N,N) \end{bmatrix}, \quad (8)$$

which is in the form $A \cdot x = \lambda \cdot b$. A is a sparse Hermitian matrix that depends only on parameter λ coefficients and size $(N^2 \times N^2)$, x is the vector of unknown image $u(\cdot, \cdot)$ $(N^2 \times 1)$,

and b is the vector of the original image $u_0(\cdot,\cdot)$ $(N^2\times 1)$. If the matrix A is invertible, the brightness of the unknown image is derived from the solution of the system of linear equations:

$$x = A^{-1}\cdot(\lambda\cdot b). \tag{9}$$

The matrix is invertible if the determinant is nonzero (see Appendix—Lemmas A.1 and A.2 and Theorem A.3).

3.3. Modified Mumford Shah Model. From the above analysis, a sparse Hermitian matrix A has been arisen. This matrix, as it is evident from (8), has the value $\lambda+4$ in its central diagonal, the value −1 in the next up and down diagonal, and the value −1 N positions before and after the central diagonal. Thus, it is a block tridiagonal matrix which can be inverted using an iterative algorithm such as one presented in [28]. Instead of using this computationally exhaustive algorithm, the fact that the above form of matrix A can be obtained from two independent minimizations is exploited: one for the second order partial derivative in the x-axis and one for the second directional derivative in the y-axis. Thus, the following two objective functions have to be minimized with respect to $u(\cdot,\cdot)$:

$$\begin{aligned} J(u) &= \frac{1}{2}\cdot\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\partial u}{\partial x}\right)^2 + \frac{\lambda}{2}\cdot\sum_{i=1}^{N}\sum_{j=1}^{N}\left|u-u_0\right|^2,\\ J(u) &= \frac{1}{2}\cdot\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\partial u}{\partial y}\right)^2 + \frac{\lambda}{2}\cdot\sum_{i=1}^{N}\sum_{j=1}^{N}\left|u-u_0\right|^2. \end{aligned} \tag{10}$$

These minimization problems (10) led to the following set of equations:

$$\begin{aligned} J(u) &= \frac{1}{2}\cdot\sum_{x=1}^{N-1}\sum_{y=1}^{N-1}\left[\left(u(x+1,y)-u(x,y)\right)^2\right]\\ &\quad+\frac{\lambda}{2}\cdot\sum_{x=1}^{N-1}\sum_{y=1}^{N-1}\left[u(x,y)-u_0(x,y)\right]^2,\\ J(u) &= \frac{1}{2}\cdot\sum_{x=1}^{N-1}\sum_{y=1}^{N-1}\left[\left(u(x,y+1)-u(x,y)\right)^2\right]\\ &\quad+\frac{\lambda}{2}\cdot\sum_{x=1}^{N-1}\sum_{y=1}^{N-1}\left[u(x,y)-u_0(x,y)\right]^2. \end{aligned} \tag{11}$$

By differentiating above two equations the following formulas are obtained:

$$\begin{aligned} \frac{\partial J(u)}{\partial u} &= \frac{1}{2}\cdot\sum_{x=1}^{N-1}\sum_{x=1}^{N-1}\frac{\partial}{\partial u}\left[\left(u(x+1,y)-u(x,y)\right)^2\right]\\ &\quad+\frac{\lambda}{2}\cdot\sum_{x=1}^{N-1}\sum_{x=1}^{N-1}\frac{\partial}{\partial u}\left[u(x,y)-u_0(x,y)\right]^2=0\\ &\Longrightarrow(\lambda+2)\cdot u(x,y)-u(x+1,y)-u(x-1,y)\\ &\quad=\lambda\cdot u_0(x,y),\quad \forall x,y\in[1,N-1], \end{aligned} \tag{12}$$

$$\begin{aligned} \frac{\partial J(u)}{\partial u} &= \frac{1}{2}\cdot\sum_{x=1}^{N-1}\sum_{x=1}^{N-1}\frac{\partial}{\partial u}\left[\left(u(x,y+1)-u(x,y)\right)^2\right]\\ &\quad+\frac{\lambda}{2}\cdot\sum_{x=1}^{N-1}\sum_{x=1}^{N-1}\frac{\partial}{\partial u}\left[u(x,y)-u_0(x,y)\right]^2=0\\ &\Longrightarrow(\lambda+2)\cdot u(x,y)-u(x,y+1)-u(x,y-1)\\ &\quad=\lambda\cdot u_0(x,y),\quad \forall x,y\in[1,N-1]. \end{aligned} \tag{13}$$

Equations (12), (13) can be written in matrix form as

$$\begin{bmatrix} \lambda+2 & 0 & 0 & 0 & \cdot & -1 & \cdot & 0\\ 0 & \lambda+2 & 0 & 0 & \cdot & \cdot & -1 & 0\\ 0 & 0 & \lambda+2 & 0 & \cdot & \cdot & \cdot & -1\\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot\\ -1 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0\\ \cdot & -1 & \cdot & \cdot & \cdot & 0 & \lambda+2 & 0\\ 0 & 0 & -1 & 0 & \cdot & 0 & 0 & \lambda+2 \end{bmatrix}$$

$$\cdot\begin{bmatrix} u(1,1)\\ u(1,2)\\ \cdot\\ u(1,N)\\ u(2,1)\\ u(2,2)\\ \cdot\\ u(2,N)\\ \cdot\\ u(N,1)\\ \cdot\\ u(N,N) \end{bmatrix} = \lambda\cdot\begin{bmatrix} u_0(1,1)\\ u_0(1,2)\\ \cdot\\ u_0(1,N)\\ u_0(2,1)\\ u_0(2,2)\\ \cdot\\ u_0(2,N)\\ \cdot\\ u_0(N,1)\\ \cdot\\ u_0(N,N) \end{bmatrix}, \tag{14}$$

$$\begin{bmatrix} \lambda+2 & -1 & 0 & 0 & \cdot & \cdot & \cdot & 0\\ -1 & \lambda+2 & -1 & 0 & \cdot & \cdot & \cdot & 0\\ 0 & -1 & \lambda+2 & -1 & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & -1 & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \lambda+2 & -1\\ 0 & 0 & \cdot & 0 & \cdot & 0 & -1 & \lambda+2 \end{bmatrix}$$

$$\cdot \begin{bmatrix} u(1,1) \\ u(1,2) \\ \cdot \\ u(1,N) \\ u(2,1) \\ u(2,2) \\ \cdot \\ u(2,N) \\ \cdot \\ u(N,1) \\ \cdot \\ u(N,N) \end{bmatrix} = \lambda \cdot \begin{bmatrix} u_0(1,1) \\ u_0(1,2) \\ \cdot \\ u_0(1,N) \\ u_0(2,1) \\ u_0(2,2) \\ \cdot \\ u_0(2,N) \\ \cdot \\ u_0(N,1) \\ \cdot \\ u_0(N,N) \end{bmatrix}. \tag{15}$$

From (14) and (15) it can be observed that in (14) the matrix A has the value $\lambda + 2$ in its central diagonal and the value -1 N positions before and after the central diagonal while in (15) the matrix A has the value $\lambda + 2$ in its central diagonal and the value -1 in the next up and down diagonal. Thus, the combination of the two matrices can give us the same results (see Appendix—Statement) as the matrix A presented in (8). By exploiting this fact we can use only the matrix A obtained by (15) instead of using the matrix A obtained from (8). The advantage of this approach is that in this case the inversion of a symmetric matrix (15) is required which can be obtained analytically (see Appendix) and without the computationally exhaustive approaches adopted in [28–34]. In the sequel, (9) is applied one time for the original image $u_0(\cdot,\cdot)$ and one for its rotated version by 90 degrees. This is done because two objective functions (one in the x direction and one in the y direction) have to be minimized. The enhanced image is derived by the addition (after normalization) of the two independent results (after the rotation of the second image $u(\cdot,\cdot)$ by -90 degrees in order to have a meaningful result):

$$x_{(0\text{degrees})} = A^{-1} \cdot \left(\lambda \cdot b_{(0\text{degrees})}\right), \tag{16}$$

$$x_{(90\text{degrees})} = A^{-1} \cdot \left(\lambda \cdot b_{(90\text{degrees})}\right), \tag{17}$$

$$u = x_{(0\text{degrees})} + x_{(90\text{degrees})}, \tag{18}$$

where $b_{(0\text{degrees})}$, $b_{(90\text{degrees})}$ are the original image (rearranged as vector) and its rotated version (also rearranged as vector) by 90 degrees, respectively, A^{-1} is the inverse of the matrix presented in (15), and $x_{(0\text{degrees})}$, $x_{(90\text{degrees})}$ are the resulting smooth image and its rotation by 90 degrees version. Finally, u is their sum after the rotation of $x_{(90\text{degrees})}$ by -90 degrees in order to have a meaningful result.

The determinant of the matrix A must be nonzero in order to be invertible. This is proved in Lemmas A.1 and A.2 (see Appendix).

In practice, even in the case of small images, that is, 100×100 pixels, the inversion of a sparse matrix A of 10000×10000 is required. This is obvious of too high computational cost. An effective reduction of the matrix A dimensionality can be achieved using subimages, that is, multiple solutions of (13), using only a small number of neighbor pixels. In this case the number of linear equations depends on the size of the chosen window.

From the optimization criterion, the new image u is a smooth image similar to the original u_0. As a result of the image transformation method the veins are located in concave nonsmooth regions, so the veins network is enhanced in the nonsmooth image $u - u_0$.

In order to obtain the image $u - u_0$ the following process is applied. Initially a window of size MxM is selected and the corresponding matrix A $(M^2 \times M^2)$ and its inverse inv$A$ $(M^2 \times M^2)$ are estimated. Then, the sliding window is moved along each pixel of the original image and the difference between the pixel values (of the image u inside the window) obtained by (16) or (17) and the pixel values of the original image u_0 inside the window is computed. In the next step all these window differences are added and the image $u - u_0$ is obtained by keeping the central $N \times N$ part of the result (where $N \times N$ is the size of the original image).

The result of this process is the nonsmooth image $u - u_0$. In this image the veins located in concave regions and for this reason a local entropy thresholding technique is applied in order to segment the nonsmooth image in concave (veins) and nonconcave (other tissue parts) regions.

This procedure is repeated for two times: one time for the original image and one for its rotation version by 90 degrees. The two nonsmooth images are added in order to obtain the final enhanced image u (18).

To conclude, the concept of the proposed method originates from the continuous Mumford Shah Model. However, in this paper, a modified separable discrete model (11) is proposed and the transition between continuous and discrete space is not straightforward. Equation (3) is used only for indicating the conceptual similarity between the proposed discrete model and the continuous Mumford Shah Model because it is deemed that it is not fair to present the proposed model as an entirely novel model. Thus, the actual minimization is performed in models presented in (6) and (11) for the modified discrete and the modified separable discrete model, respectively.

3.4. Local Entropy Thresholding. Among the various existing methods used to automatically define the threshold for segmentation, the local entropy thresholding is selected, which has been successfully used in [35], because pixel intensities are not independent and this efficient entropy based thresholding takes into account the spatial distribution of intensities. It is based on the estimation of the cooccurrence matrix of the image $u - u_0$ which is a measure of the transition of intensities between adjacent pixels. Specifically, a local entropy thresholding technique, described in [36], is implemented which can preserve the structure details of an image. Two images with identical histograms but different spatial distribution will result in different entropy (also different threshold values).

In a correction note N. R. Pal and S. K. Pal [37] propose two modifications to improve the results of blood vessel

extraction that are essential to the performance of image registration. These modifications were adopted also in our study because they experimentally proved superior to [36] (see Appendix—Local Entropy Thresholding).

3.5. Postprocessing

3.5.1. Morphological Dilation. The resulting binary image tends to suffer from some misclassifications (outliers). In order to have a robust segmentation in this postprocessing substep a morphological dilation [38] with a line structuring element oriented along the x-axis and elongated Y pixels are performed. The output of this step is an image with less outliers but with still evident misclassifications. The final morphological filtering substep that follows gives us the desired robust finger vein pattern.

3.5.2. Morphological Filtering. In this substep the image is postprocessed by applying iteratively a morphological filter called majority [38]. This filter sets a pixel to 1 if five or more pixels in its 3-by-3 neighbourhood has the value 1; otherwise, it sets the pixel to 0. This filter is applied iteratively until the output image remains unchanged. This application clears the image from small misclassified regions which appears due to the presence of noise and smoothes the contours.

4. Experimental Results

4.1. Real Image Database. The original image was acquired under infrared light using an inexpensive *CCD* camera. The finger was placed between the camera and the light source which consists of a row of infrared leds (five elements) with adjustable illumination. The intensity of the leds adjusted as far as the illumination of the image was good enough. However, the problem of acquisition of infrared images is not a trivial task. The phenomena which were involved in the transmission of light inside the human tissue are very complicated. This fact does not permit us to acquire a sufficient number of images in order to construct an infrared finger image database and to release it in the research community for evaluation and comparison purposes. Our future work is mainly focused on the direction of improvement of the image acquisition system.

An excellent image illumination is not a strict requirement because the good performance of the proposed method remains also under adverse illumination conditions. Due to the fact that haemoglobin has strong absorption in the infrared light the veins are shown in the image darker than the other human tissues. So, the goal of our study is to extract these dark regions, corresponding to veins, from the background, corresponding to the other human parts (tissue). The original image which was acquired as described above is shown in Figure 2.

In this section the results of the application of our method and the methods proposed in [6, 7, 11, 12] are presented. The qualitative evaluation of methods performed in the *ROI* image is shown in Figure 4 and in the original image shown in Figure 2.

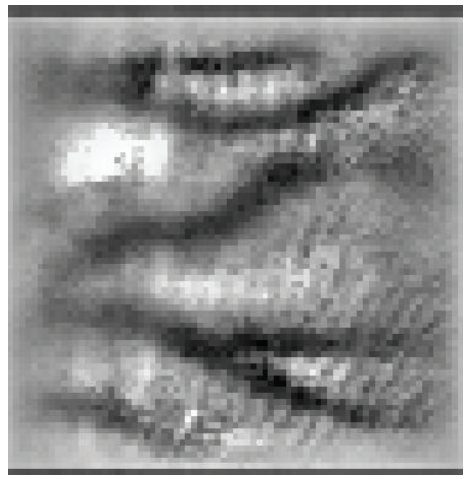


Figure 5: Nonsmooth image.

4.2. Proposed Method. For both images a window neighbourhood of size 9×9 is used which results in an 81×81 matrix A which can be inverted very quickly. The selection of the parameter λ does not affect the performance of our method and thus it is arbitrarily selected as $\lambda = 1$. Parameter λ is a weighting factor between the two terms of (2). It accounts for the degree of similarity between the original image and the estimated image. Extensive experiments are conducted in order to justify the selection of parameter λ. During the experiments the evaluation rates did not vary very much by changing the parameter λ from 0.1 to 1. The average divergence on the results was almost 2%. For the threshold computation the modified local entropy thresholding technique is used as described in the correction note [37]. In postprocessing step, a line structuring element with $Y = 5$ pixels in length and oriented in the x-axis is employed.

Figures 5–8 present the results of the application of the proposed method in the *ROI* image of Figure 4. Figure 5 shows the nonsmooth image obtained after the application of the modified separable Mumford Shah Model in both 0 and 90 degrees direction and the addition of the results. Figure 6 shows the detection, with the help of the modified local entropy thresholding, of the concave regions of the image, where the veins tend to locate. In this binary image, concave regions (candidate pixels to be detected as veins) are shown in black while other tissue parts shown in white. Figure 7 shows the binary image after the application of the morphological dilation substep and, finally, the image in Figure 8(a) shows the extracted finger vein pattern obtained after the final morphological filtering substep.

Moreover, the results of the application of our method in the original image of Figure 2 (before *ROI* extraction) are presented. Figure 9 shows the nonsmooth image obtained after the application of the modified separable Mumford Shah Model in both 0 and 90 degrees directions and the addition of the results. Figure 10 shows the detection, with the help of the modified local entropy thresholding, of the concave regions of image. In this binary image, concave regions (candidate pixels to be detected as veins) are shown in black while other

Figure 6: Concave (black) and nonconcave regions (white).

Figure 7: Morphological dilation.

tissue parts shown in white. Figure 11 shows the binary image after the application of the morphological dilation substep and finally Figure 12(a) shows the extracted finger vein pattern obtained after the final morphological filtering substep.

4.3. Method [6, 7]. As mentioned, the proposed method is compared with our implementation of the method presented in [6, 7]. The repeated line tracking method requires some parameter tuning in order to run and the robustness of the extracted finger vein pattern is strongly affected by the number of repetitions. The parameters used in the experiments are $N = 10000$ for the number of iterations, $p_{lr} = 50$ and $p_{ud} = 25$ for the probabilities of selecting the three neighboring pixels in the horizontal or vertical direction, respectively, $W = 9$ for the width of the profiles, and $r = 1$ for the distance between the testing pixel and the cross section. In order to perform a fair comparison between method [6, 7] and the proposed method the same morphological postprocessing step is used in all experiments.

4.4. Other Methods. Moreover, the proposed method is also compared with more recent methods such as those presented in [11, 12], briefly described in the Related Work section, due to the fact that these methods have been applied in the same real and artificial image databases and thus the comparison is reasonable.

For purposes of comparison the results of the application of methods [6, 7, 11, 12] in original and ROI image are presented in the same figures as the results of the proposed method. Thus, Figure 8(b) shows the extracted finger vein pattern produced after the application of the line tracking method in the *ROI* image of Figures 4 and 8(c) shows the extracted finger vein pattern produced after the application of the method [11] in the *ROI* image of Figure 4 while Figure 12(b) shows the extracted finger vein pattern produced after the application of the line tracking method in the image of Figures 2 and 12(c) shows the extracted finger vein pattern produced after the application of the method [12] in the image of Figure 2.

Observing Figures 8 and 12, it is obvious that no safe conclusions regarding the performance of all methods can be conducted by visual inspection of the images. Instead, a comparison based on widely known measures should be done. Unfortunately, there is not a publicly available database existing that can be used for evaluation and comparison purposes between various methods. However, an artificial finger image database is constructed in this study in order to evaluate the proposed method and to compare it with the methods presented in [6, 7, 11, 12]. It is worth mentioning that the results presented for the method [6, 7] are produced by the application of our implementation to images since the code of the method is not publicly available by the authors.

4.5. Artificial Image Database. A quantitative evaluation of the proposed method in real infrared images is difficult due to the absence of manual segmentation data. The extremely low contrast images increase the disagreement of human annotation. Therefore, the proposed method is evaluated using a small set of images, each one created by the weighed sum of two artificial images. The first image is constructed using an artificial vein-like network. This network consists of connected lines of different widths with junctions and bifurcations and multiple low pass filtering to simulate the blurriness of the edges which is apparent to the real images due to the blood flow and scattering effects. The second artificial image is used to simulate the nonuniform image background of real infrared images created by applying an iterative spatial low pass Gaussian filter with a large window size to the original infrared image.

4.6. Evaluation Rates. In the finger vein segmentation process, each pixel is classified as tissue (nonvein) or vein. Consequently, there are four events, true positive (TP) and true negative (TN) when a pixel is correctly segmented as vein or nonvein and two misclassifications: a false negative (FN) appears when a pixel in a vein is segmented in the nonvein area, and a false positive (FP) when a nonvein pixel is segmented as a vein pixel.

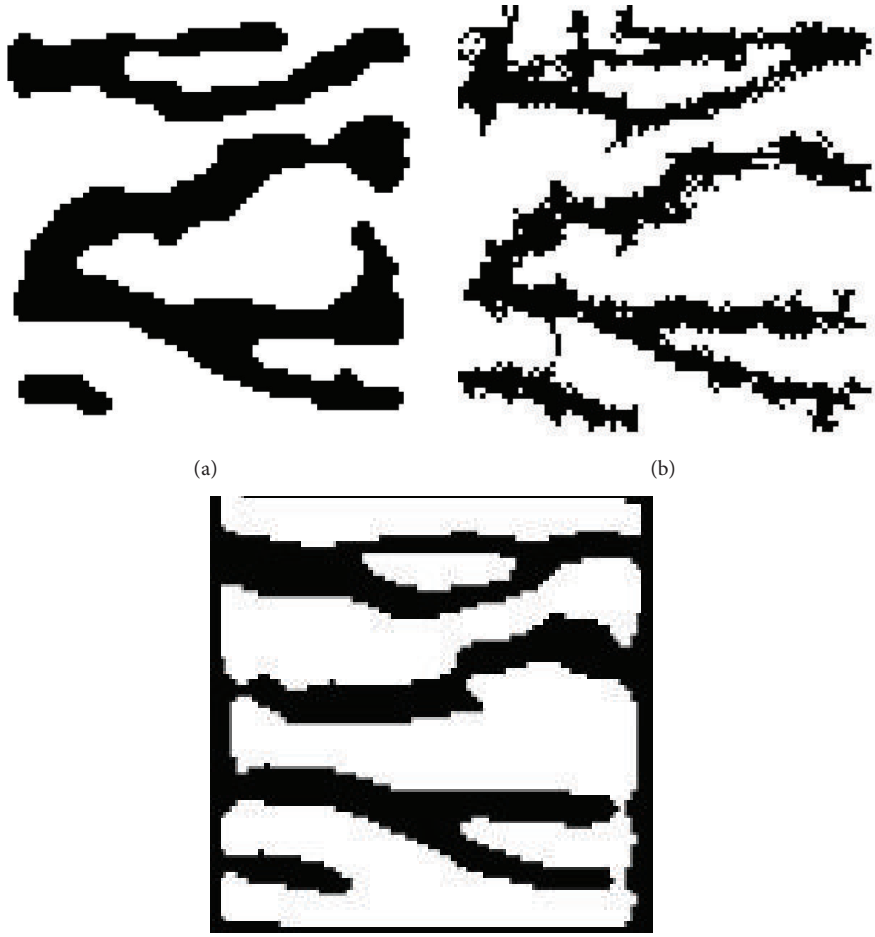


Figure 8: (a) Morphological (majority) filtering: extracted finger vein pattern using the proposed method, (b) extracted finger vein pattern using method [6, 7], and (c) extracted finger vein pattern using method [11].

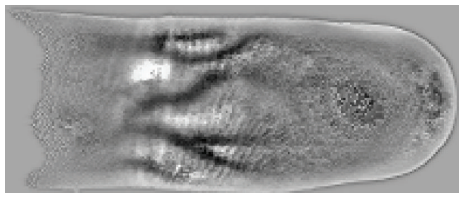


Figure 9: Nonsmooth image.

Figure 10: Concave (black) and nonconcave regions (white).

Two widely known statistical measures are used for method evaluation: sensitivity and specificity, which are used to evaluate the performance of the binary segmentation outcome. The sensitivity is a normalized measure of true positives, while specificity measures the proportion of true negatives:

$$\begin{aligned} \text{sensitivity} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{specificity} &= \frac{\text{TN}}{\text{TN} + \text{FP}}. \end{aligned} \tag{19}$$

Usually, there is a tradeoff between two measures. Finally, the accuracy of the binary classification is defined by

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{P + N}, \tag{20}$$

where P and N represent the total number of positive (vein) and negative (nonvein) pixels in the segmentation process and are the degree of conformity between the estimated binary classification and the ground truth obtained through a manual segmentation. Thus, the accuracy is strongly related

Figure 11: Morphological dilation.

to the segmentation quality and for this reason it is used to evaluate and compare different methods.

4.7. Evaluation Results. The proposed method and the methods [6, 7, 11, 12] are evaluated quantitatively on the artificial image database. Each image of the set is constructed according to the above procedure. The evaluation is performed using the widely known statistical measures of sensitivity, specificity, and accuracy.

Table 1 shows the mean sensitivity, specificity, and accuracy of the proposed method and the methods [6, 7, 11] (without preprocessing/postprocessing) and [12] on the artificial finger image database, while Figure 13(a) shows the *ROC* curve of the proposed method produced by varying the local segmentation threshold and estimating the corresponding measures, Figure 13(b) shows the *ROC* curve of the method [6, 7] produced by varying the local segmentation threshold and estimating the corresponding measures, Figure 13(c) shows the *ROC* curve of the method [11] produced by varying the local segmentation threshold and estimating the corresponding measures, and Figure 13(d) shows the *ROC* curve of the method [12] produced by varying the value of the threshold parameter φ_3 and estimating the corresponding measures.

Figure 14(a) shows the first image of the artificial image database used for the evaluation of the proposed method and the corresponding results of the segmentation for the proposed method (Figure 14(b)) and for the methods [6, 7] (Figure 14(c)), [11] (Figure 14(d)), and [12] (Figure 14(e)), respectively.

By observing the results presented in Figure 14 and in Table 1, it seems that the proposed method performs better than the methods presented in [6, 7, 11] in artificial finger image database in terms of sensitivity, specificity, and accuracy and has comparable performance with the method presented in [12] using the same evaluation criteria. In addition, the visual inspection of images shown in Figure 12 leads to the conclusion that the proposed method performs extremely better than the method [12] in real finger images. Thus, the proposed method, unlike the others compared, robustly extracts the finger vein network and preserves its connectivity against various conditions (shading, intensity variations, and noise distortion) both to real and to artificial finger images.

Apart from evaluating and comparing the proposed method to real and artificial finger vein images authors present also the results of the application of the proposed method and method [6, 7] in artificially distorted images using different types of noise and different levels of distortion. The study of the effect of noise led to an adaptation and a modification of the proposed method in order to robustly perform under extreme conditions. These conditions, although they seem to be unrealistic, may be produced by a low quality acquisition system and its noncareful setup.

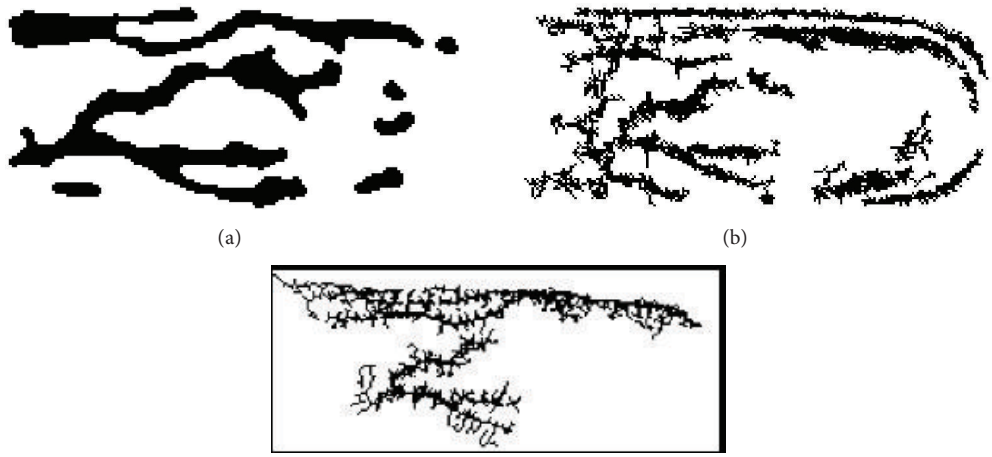
Table 2 shows mean sensitivity, specificity, and accuracy of the proposed method and the method [6, 7] for artificial infrared images with different level of distortion. These results are presented in order to indicate the robustness of the proposed method especially in low quality images. Figure 15 shows the distorted artificial images and the corresponding finger vein patterns produced after the application of the proposed method while Figure 16 shows the distorted artificial images and the corresponding finger vein patterns produced after the application of the method presented in [6, 7].

Comparing the results presented in Tables 1 and 2, it seems that the proposed method is superior to method [6, 7] in terms of sensitivity, specificity, and accuracy for the artificial finger image database. Moreover, it is worth noting that the results of the method [6, 7] are slightly different among different executions with the same parameters and images due to the randomness introduced in the selection of the current tracking pixel and the parameters p_{lr} and p_{ud}. Finally, regarding the computational complexity, the proposed method also outperforms the method [6, 7] as the number of iterations of line tracking increases.

By carefully observing the results in Table 2 and the images in Figure 16 it is obvious that the performance of the method presented in [6, 7] in low quality images is degraded and becomes unacceptable for images with high level of distortion. Thus, the proposed method outperforms line tracking method when applied to low quality image which may be produced as a result of a noncareful image acquisition setup.

As the above results show, the proposed method performs well in the majority of cases and achieves efficiently segmenting the finger vein images. However, no scientific method is perfect. Every method has also drawbacks. Thus, the main drawback of the proposed method is that its performance both in terms of segmentation accuracy and in terms of computational complexity is strongly related to the size of window neighbourhood. An appropriate selection of window neighbourhood size must be done in order to achieve meaningful results. On the other side, in case of an accurate image acquisition setup which will acquire images of specific resolution, the window size could be derived experimentally and could be then set at once.

4.8. Matching. In general, two methods are commonly used for matching of line-shaped patterns: structural matching [39] and template matching [40, 41]. As stated in [6, 7], structural matching requires additional extraction of feature points such as line endings and bifurcations. Since a finger vein pattern has few of these points, template matching based on comparison of pixel values is more appropriate for finger vein pattern matching. Thus, in this paper the robustness of the proposed method and the method [6, 7] for finger vein identification is evaluated by estimating the mismatch ratio



(a) (b) (c)

Figure 12: (a) Morphological (majority) filtering: extracted finger vein pattern using the proposed method, (b) extracted finger vein pattern using method [6, 7], and (c) extracted finger vein pattern using method [12].

Table 1: Mean sensitivity, specificity, and accuracy of the proposed method and the methods [6, 7], [11] (without preprocessing/postprocessing) and [12].

	Proposed method			Method [6, 7]		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
Mean	0.907	0.910	0.909	0.782	0.928	0.896
Standard deviation	0.073	0.018	0.027	0.075	0.037	0.041
	Method [11]			Method [12]		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
Mean	0.869	0.898	0.892	0.943	0.941	0.942
Standard deviation	0.081	0.025	0.036	0.112	0.032	0.048

between the registered and the input data for the artificial finger image database. In the matching process, the extracted finger vein pattern is converted into matching data, and these data are compared with the recorded raw data. For the case of artificial image database the ground truth data was used as recorded raw data for each image.

Mismatch ratio R_m is calculated to examine whether or not two sets of data have a correlation with each other. The ratio R_m is defined as the difference between two sets of data to be matched. $R(x, y)$ and $I(x, y)$ are the values at position (x, y) of the registered and input matching data, w and h are the width and height of both sets of data, c_w and c_h are the distances in which motion in the vertical and horizontal directions, respectively, is required to adjust the displacement between the two sets of data, and the template data are defined as the rectangular region within $R(x, y)$ whose upper left position is $R(c_w, c_h)$ and lower right position is $R(w-c_w, h-c_h)$.

The value of mismatch $N_m(s, t)$, which is the difference between the registered and input data at the positions where $R(c_w, c_h)$ overlaps with $I(s, t)$, is defined as follows:

$$N_m(s,t) = \sum_{y=0}^{h-2c_h-1} \sum_{x=0}^{w-2c_w-1} \left\{\phi\left(I\left(s+x, t+y\right), R\left(c_w+x, c_h+y\right)\right)\right\}, \tag{21}$$

where $w = 212$ and $h = 87$ in consideration of the finger size in the captured image, c_w and c_h are set at $c_w = 57$ and $c_h = 38$ in order to adjust the finger position in the captured image by up to about 1 cm, and φ in (21) is a parameter that indicates whether a pixel labeled as part of the background region and a pixel labeled as part of a vein region overlapped with each other. When P_1 is defined as the pixel value of one pixel and P_2 is defined as the pixel value of the other pixel, φ can be described as follows:

$$\phi\left(P_1, P_2\right) = \begin{cases} 1, & \text{if } \left|P_1 - P_2\right|, \\ 0, & \text{otherwise.} \end{cases} \tag{22}$$

The minimum value of mismatch N_m, which is the smallest $N_m(s, t)$ calculated under the condition that the template overlaps with the input matching data $I(x, y)$ at all positions, can be defined as follows:

$$N_m = \min_{0 \le s < 2c_w, 0 \le t < 2c_h} N_m(s,t). \tag{23}$$

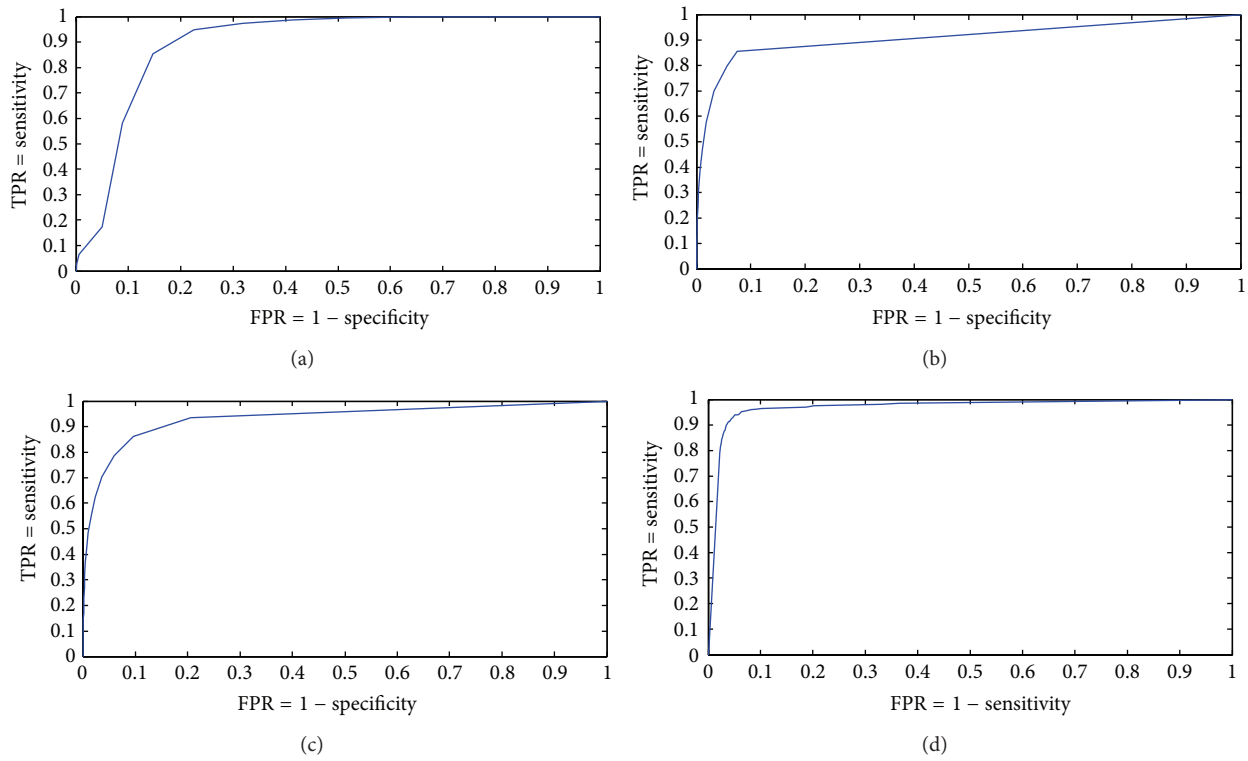


Figure 13: (a) *ROC* curve of the proposed method produced by varying the local segmentation threshold, (b) *ROC* curve of the method proposed in [6, 7] produced by varying the local segmentation threshold, (c) *ROC* curve of the method proposed in [11] produced by varying the local segmentation threshold, and (d) *ROC* curve of the method proposed in [12] produced by varying the value of the threshold parameter φ_3.

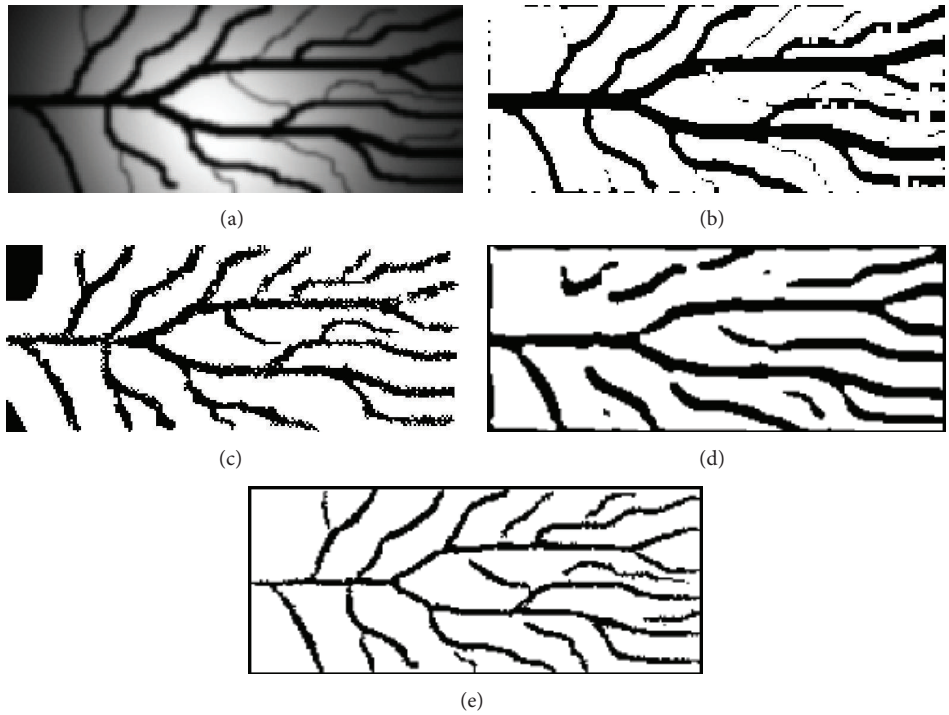


Figure 14: (a) Artificial infrared finger image, (b) extracted vein pattern using the proposed method, (c) extracted vein pattern using the method [6, 7], (d) extracted vein pattern using the method [11], and (e) extracted vein pattern using the method [12].

Table 2: Mean sensitivity, specificity, and accuracy of the proposed method and the method [6, 7].

Image name	Proposed method			Method [6, 7]		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
Artificial_finger_vein.bmp	0.8779	0.9266	0.9149	0.7876	0.9047	0.8765
Artificial_finger_vein2.bmp	0.5234	0.9462	0.8444	0.5221	0.9151	0.8205
Artificial_finger_vein3.bmp	0.6144	0.9716	0.8856	0.6185	0.8395	0.7863
P1.bmp	0.8845	0.8425	0.8526	0.5477	0.9176	0.8286
P2.bmp	0.8097	0.8295	0.8247	0.4338	0.7743	0.6923
P3.bmp	0.7748	0.8139	0.8045	0.4680	0.5865	0.5580
P4.bmp	0.7349	0.8091	0.7913	0.5561	0.4013	0.4386

Using the definitions given above, the mismatch ratio R_m is defined as follows:

$$R_m = N_m \cdot \left(\left\{ \sum_{j=t_0}^{t_0+h+2c_h-1} \sum_{i=s_0}^{s_0+w-2c_w-1} \phi\left(I\left(i,j\right),0\right) + \sum_{j=c_h}^{h-c_h-1} \sum_{i=c_w}^{w-c_w-1} \phi\left(0,R\left(i,j\right)\right) \right\} \right)^{-1}, \tag{24}$$

where s_0 and t_0 are s and t such that (23) is minimized. As is shown by (24), R_m is described as the ratio between N_m and the total number of pixels that are classified as belonging to the vein region in the two data sets.

As shown in Table 3, the average mismatch ratio for the artificial finger image database, estimated using (24), is 24.65% for the proposed method while for the method [6, 7] it is 43.64%. Although database contains twenty images and the results cannot be generalized, the proposed method seems to be more appropriate for finger vein identification purposes. The extracted patterns by the proposed method have significantly lower value of mismatch ratio than those extracted by the method [6, 7].

Table 3: Average mismatch ratio of the proposed method and the method [6, 7].

Parameters $w = 212$, $h = 87$ $c_w = 57$, $c_h = 38$	Proposed method	Method [6, 7]
Average mismatch ratio	24.65%	43.64%

5. Conclusions

In this paper an efficient finger vein pattern extraction method is presented. The proposed method is based on the minimization of the objective function of a modified Mumford Shah Model and the local application of its results. This application produces two nonsmooth images where veins located in concave regions. The two images are then combined simply by addition. Detection of concave regions is achieved via a modified local entropy thresholding technique. The preliminary segmentation result was unsatisfactory due to the presence of some outliers (misclassifications).

Thus, a final morphological postprocessing step followed in order to clean the image from the misclassifications and to produce a robust finger vein pattern. Future work includes the improvement of our imaging device in order to acquire images with less shading and noise artefacts, something that will guarantee the successful application of our method in the majority of cases. In case of images with high quality the preprocessing and/or postprocessing step can be skipped.

The experimental evaluation of the proposed method shows that it can robustly segment the finger vessel network and that the extracted finger vein pattern is appropriate for finger vein identification purposes. Finally, the proposed method is robust against strong distortions met in the acquired images.

Appendix

Lemma A.1. *The determinant of the symmetric matrix A in the form*

$$A_N(x) = \begin{bmatrix} x & -1 & 0 & 0 & \cdot & 0 \\ -1 & x & -1 & 0 & \cdot & 0 \\ 0 & -1 & x & -1 & \cdot & 0 \\ 0 & 0 & -1 & x & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & \cdot & x \end{bmatrix} = \left\{ a_{i,j} \mid a_{i,j} = \begin{cases} x, & i = j \\ -1, & i = j+1 \vee i = j-1, \ i,j = 1,N \\ 0, & elsewhere \end{cases} \right\} \tag{A.1}$$

can be estimated using the following recursive formula:

$$\left|A_n(x)\right| = x\left|A_{n-1}(x)\right| - \left|A_{n-2}(x)\right|, \quad n = 3,N, \qquad \left|A_2(x)\right| = x^2 - 1, \qquad \left|A_1(x)\right| = x. \tag{A.2}$$

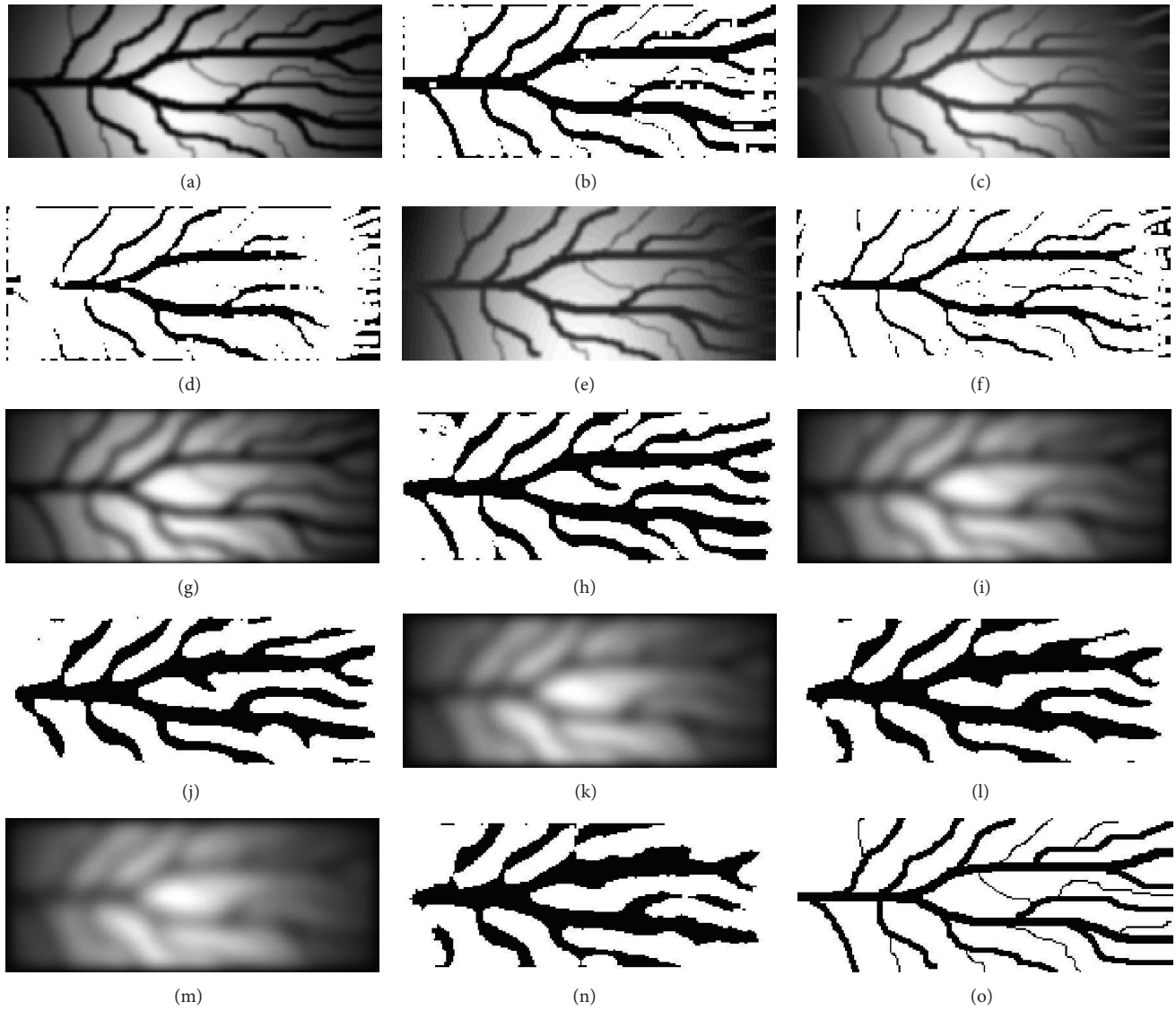


FIGURE 15: (a, c, e, g, i, k, and m) Artificial infrared images with different level of distortion, (b, d, f, h, j, l, and n) corresponding finger vein patterns extracted using the proposed method, and (o) manual segmentation (ground truth).

Proof. From the definition of matrix A it can be derived that

$$\begin{aligned}\left|A_n(x)\right| &= \begin{vmatrix} x & -1 & 0 \\ -1 & x & -1 \\ 0 & -1 & A_{n-2}(x) \end{vmatrix} \\ &= x\left|A_{n-1}(x)\right| + \begin{vmatrix} -1 & -1 \\ 0 & A_{n-2}(x) \end{vmatrix} \\ &= x\left|A_{n-1}(x)\right| - \left|A_{n-2}(x)\right|. \end{aligned} \tag{A.3}$$

□

Lemma A.2. *The matrix A determinant is an increased sequence of positive real numbers for any $x > 2$.*

Proof. If $x > 2$, then

$$\begin{aligned}\left|A_2(x)\right| &= x^2 - 1 > 3 \wedge \left|A_1(x)\right| \\ &= x > 2 \Longrightarrow \left|A_2(x)\right| > \left|A_1(x)\right|. \end{aligned} \tag{A.4}$$

If $|A_n(x)| > |A_{n-1}(x)|$, then

$$\begin{aligned}\left|A_{n+1}(x)\right| &= x\left|A_n(x)\right| - \left|A_{n-1}(x)\right| > x\left|A_n(x)\right| - \left|A_n(x)\right| \\ &= (x-1)\left|A_n(x)\right| > (2-1)\left|A_n(x)\right| \\ &= \left|A_n(x)\right| \\ &\Longrightarrow \left|A_{n+1}(x)\right| > \left|A_n(x)\right|. \end{aligned} \tag{A.5}$$

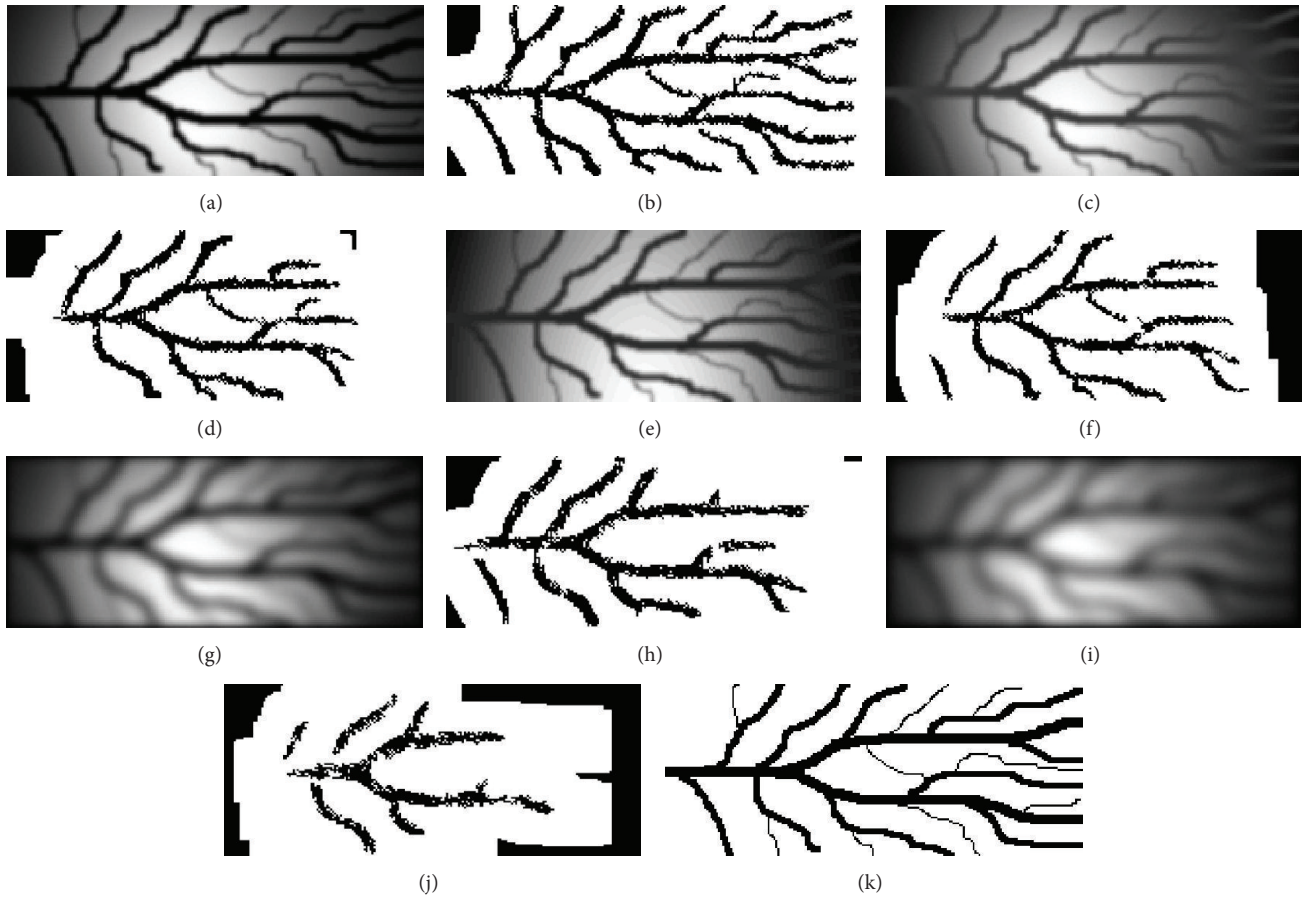


FIGURE 16: (a, c, e, g, and i) Artificial infrared images with different level of distortion, (b, d, f, h, and j) corresponding finger vein patterns extracted using the method [6, 7], and (k) manual segmentation (ground truth).

The induction rule leads to a sequence of positive numbers for the determinant of A:

$$\left|A_{n+1}(x)\right| > \left|A_n(x)\right| > \cdots > \left|A_2(x)\right| > \left|A_1(x)\right| > 0. \quad \text{(A.6)}$$

□

Theorem A.3. *The system of linear equations is solvable for any positive value of* λ.

Proof. Based on Lemma A.2, $\lambda > 0 \Rightarrow \lambda + 4 > 2 \Rightarrow |A_n(\lambda + 4)| > 0$. □

Statement. Modified model produces the same results as the original model.

Proof. The minimization of the objective function regarding $u(\cdot,\cdot)$ leads to the following equation:

$$\begin{aligned}&(\lambda + 4)\cdot u(x,y) - u(x+1,y)\\ &\quad - u(x,y+1) - u(x,y-1) - u(x-1,y) \qquad \text{(A.7)}\\ &= \lambda \cdot u_0(x,y), \quad \forall x, y \in [1, N-1],\end{aligned}$$

while the minimization of the two independent objective functions (one for the second order partial derivative in the x-axis and one for the second directional derivative in the y-axis) leads to the following equations:

$$\begin{aligned}&(\lambda + 2)\cdot u(x,y) - u(x+1,y) - u(x-1,y)\\ &\quad = \lambda \cdot u_0(x,y), \quad \forall x, y \in [1, N-1],\end{aligned} \quad \text{(A.8)}$$

$$\begin{aligned}&(\lambda + 2)\cdot u(x,y) - u(x,y+1) - u(x,y-1)\\ &\quad = \lambda \cdot u_0(x,y), \quad \forall x, y \in [1, N-1].\end{aligned} \quad \text{(A.9)}$$

Adding (A.8) and (A.9) results in

$$\begin{aligned}&(2\cdot\lambda + 4)\cdot u(x,y) - u(x+1,y)\\ &\quad - u(x,y+1) - u(x,y-1) - u(x-1,y) \qquad \text{(A.10)}\\ &= 2\cdot\lambda \cdot u_0(x,y), \quad \forall x, y \in [1, N-1].\end{aligned}$$

The only difference between (A.7) and (A.10) is that in (A.10) parameter λ is multiplied by a factor of 2. This makes no difference because the parameter λ can be arbitrarily set and has no strong impact on the final results. In other words, (A.7) and (A.10) are equivalent and produce exactly the same

results if the parameter λ set in (A.8) and (A.9) to have the half value from the value has the corresponding parameter in (A.7). □

Local Entropy Thresholding. The cooccurrence matrix of the image $F = u - u_0$ is a $P \times Q$ dimensional matrix $T = [t_{ij}]_{P\times Q}$ that gives an idea about the transition of intensities between adjacent pixels, indicating spatial structural information of an image. Depending upon the ways in which the gray level i follows gray level j different definitions of cooccurrence matrix are possible. Here, we made the cooccurrence matrix asymmetric by considering the horizontally right and vertically lower transitions. This choice has been made in order to decrease the computational cost and it is reasonable since the horizontally left and vertically upper transitions do not add more information to cooccurrence matrix. Thus, t_{ij} is defined as follows:

$$t_{ij} = \sum_{l=1}^{P}\sum_{k=1}^{Q} u\left(\delta\left(f\left(l,k\right)-i\right)\delta\left(f\left(l,k+1\right)-j\right)\right.$$
$$\left.+\delta\left(f\left(l,k\right)-i\right)\delta\left(f\left(l+1,k\right)-j\right)-1\right). \tag{A.11}$$

The probability of cooccurrence p_{ij} of gray levels i and j can therefore be written as

$$p_{ij} = \frac{t_{ij}}{\sum_i \sum_j t_{ij}}. \tag{A.12}$$

If s, $0 \le s \le L-1$, is a threshold, then s can partition the cooccurrence matrix into 4 quadrants, namely, A, B, C, and D (Figure 17).

Let us define the following quantities:

$$P_A = \sum_{i=0}^{s}\sum_{j=0}^{s} p_{ij},$$
$$P_C = \sum_{i=s+1}^{L-1}\sum_{j=s+1}^{L-1} p_{ij}. \tag{A.13}$$

From the occurrence matrix, the corresponding probabilities within each individual quadrant must sum to one. Thus, we get the following cell probabilities for different quadrants:

$$P_{ij}^{A} = \frac{p_{ij}}{P_A} = \frac{t_{ij}}{\sum_{i=0}^{s}\sum_{j=0}^{s} t_{ij}}, \quad \text{for } 0 \le i \le s,\ 0 \le j \le s,$$
$$P_{ij}^{C} = \frac{p_{ij}}{P_C} = \frac{t_{ij}}{\sum_{i=s+1}^{L-1}\sum_{j=s+1}^{L-1} t_{ij}}, \text{ for}$$
$$s+1 \le i \le L-1, \quad s+1 \le j \le L-1. \tag{A.14}$$

The second order entropy of the object can be defined as

$$H_A^{(2)}(s) = -\frac{1}{2}\cdot\sum_{i=0}^{s}\sum_{j=0}^{s} P_{ij}^{A}\cdot\log_2 P_{ij}^{A}. \tag{A.15}$$

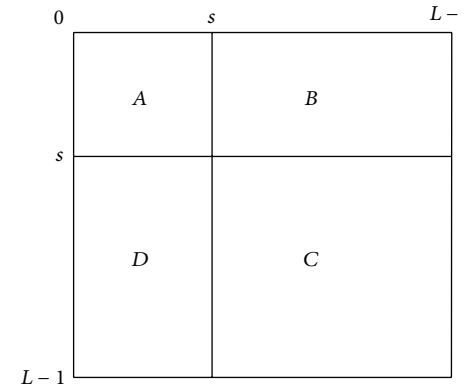


Figure 17: Quadrants of cooccurrence matrix.

Similarly, the second order entropy of the background can be written as

$$H_C^{(2)}(s) = -\frac{1}{2}\cdot\sum_{i=s+1}^{L-1}\sum_{j=s+1}^{L-1} P_{ij}^{C}\cdot\log_2 P_{ij}^{C}. \tag{A.16}$$

Hence, the total second order local entropy of the object and the background can be written as

$$H_T^{(2)}(s) = H_A^{(2)}(s) + H_C^{(2)}(s). \tag{A.17}$$

The gray level corresponding to the maximum of $H_T^{(2)}(s)$ gives the optimal threshold for object background classification.

In the first modification a different definition of the cooccurrence matrix is adopted increasing the local entropy values in vein structures. As mentioned, the cooccurrence matrix of an image shows the intensity transitions between adjacent pixels. The original cooccurrence matrix is asymmetric by considering the horizontally right and vertically lower transitions. They added some jittering effect to the cooccurrence matrix that tends to keep the similar spatial structure but with much less variations; that is, $T = [t_{ij}]_{P\times Q}$ is computed as follows:

For every pixel (l,k) in an image F

$$i = F(l,k),$$
$$j = F(l,k+1),$$
$$d = F(l+1,k+1),$$
$$t_{ij} = t_{id} + 1. \tag{A.18}$$

One may wonder whether the modified cooccurrence matrix still will represent the original spatial structure. Actually, considering a smooth area in an image where j and d should be very close or identical, the above computation in (A.18) implicitly introduces certain smoothing effect and adds some structured noise to the cooccurrence matrix. The two matrices still share a similar structure that is important for the valid thresholding result. Also, the latter one has larger entropy with a much smaller standard deviation, which is more desirable for local entropy thresholding.

Secondly, they considered the sparse foreground in selecting the optimal threshold. The original threshold selection criterion aims to maximize the local entropy of foreground and background in a gray scale image without considering the small proportion of foreground. Therefore, they proposed selecting the optimal threshold that maximizes the local entropy of the binarized image that indicates the foreground/background ratio. The larger the local entropy the more balanced the ratio between foreground and background in the binary image is.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] G. T. Park, S. K. Im, and H. S. Choi, "A person identification algorithm utilizing hand vein pattern," in *Proceedings of the Korea Signal Processing Conference*, vol. 10, pp. 1107–1110, 1997.

[2] D. U. Hong, S. K. Im, and H. S. Choi, "Implementation of real time system for personal identification algorithm utilizing hand vein pattern," *Proceedings of the IEEK Fall Conference*, vol. 22, no. 2, pp. 560–563, 1999.

[3] S.-K. Im, H.-M. Park, Y.-W. Kim et al., "An biometric identification system by extracting hand vein patterns," *Journal of the Korean Physical Society*, vol. 38, no. 3, pp. 268–272, 2001.

[4] S. K. Im, H. M. Park, S. W. Kim, C. K. Chung, and H. S. Choi, "Improved vein pattern extracting algorithm and its implementation," in *Proceedings of the International Conference on Consumer Electronics, Digest of Technical Papers (ICCE '00)*, pp. 2–3, Los Angles, Calif, USA, June 2000.

[5] S. K. Im, H. S. Choi, and S.-W. Kim, "A direction-based vascular pattern extraction algorithm for hand vascular pattern verification," *ETRI Journal*, vol. 25, no. 2, pp. 101–108, 2003.

[6] N. Miura, A. Nagasaka, and T. Miyatake, "Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification," *Machine Vision and Applications*, vol. 15, no. 4, pp. 194–203, 2004.

[7] N. Miura, A. Nagasaka, and T. Miyatake, "Feature extraction of finger vein patterns based on iterative line tracking and its application to personal identification," *Systems and Computers in Japan*, vol. 35, no. 7, pp. 61–71, 2004.

[8] M. Vlachos and E. Dermatas, "A finger vein pattern extraction algorithm based on filtering in multiple directions," in *Proceedings of the 5th European Symposium on Biomedical Engineering*, Patras, Greece, July 2006.

[9] T. Tanaka and N. Kubo, "Biometric authentication by hand vein patterns," in *Proceedings of the SICE Annual Conference*, pp. 249–253, Sapporo, Japan, August 2004.

[10] Y. Ding, D. Zhuang, and K. Wang, "A study of hand vein recognition method," in *Proceedings of the IEEE International Conference on Mechatronics and Automation (ICMA '05)*, vol. 4, pp. 2106–2110, Niagara Falls, Canada, August 2005.

[11] M. Vlachos and E. Dermatas, "A segmentation method using multiscale and multidirectional matched filters based on dyadic wavelet transform for finger vein pattern extraction," *Electrical Engineering Research*, vol. 1, no. 1, 2013.

[12] M. Vlachos and E. Dermatas, "Finger vein segmentation and centerlines extraction in infrared images," in *Advanced Topics in Scattering Theory and Biomedical Engineering: Proceedings of the 9th International Workshop on Mathematical Methods on Scattering Theory and Biomedical Engineering, Patras, Greece, 9–11 October 2009*, pp. 345–352, 2009.

[13] W. Yan, J. Xie, P. Li, T. Liu, and X. Guo, "An algorithm for vein matching based on log-polar transform," *Smart Computing Review*, vol. 3, no. 5, pp. 336–345, 2013.

[14] T. Liu, J. Xie, H. Lu, W. Yan, and P. Li, "Finger vein representation by modified binary tree model," *The Smart Computing Review*, vol. 3, no. 2, pp. 54–61, 2013.

[15] G. P. Yang, R. Y. Xiao, Y. L. Yin, and L. Yang, "Finger vein recognition based on personalized weight maps," *Sensors*, vol. 13, no. 9, pp. 12093–12112, 2013.

[16] J. F. Yang and Y. H. Shi, "Finger-vein ROI localization and vein ridge enhancement," *Pattern Recognition Letters*, vol. 33, no. 12, pp. 1569–1579, 2012.

[17] J. Yang, Y. Shi, and J. Yang, "Personal identification based on finger-vein features," *Computers in Human Behavior*, vol. 27, no. 5, pp. 1565–1570, 2011.

[18] W. Song, T. Kim, H. C. Kim, J. H. Choi, H.-J. Kong, and S.-R. Lee, "A finger-vein verification system using mean curvature," *Pattern Recognition Letters*, vol. 32, no. 11, pp. 1541–1547, 2011.

[19] V. Paquit, J. R. Price, R. Seulin et al., "Near-infrared imaging and structured light ranging for automatic catheter insertion," in *Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display*, vol. 6141 of *Proceedings of SPIE*, San Diego, Calif, USA, March 2006.

[20] H. D. Zeman, G. Lovhoiden, and C. Vrancken, "The clinical evaluation of vein contrast enhancement," in *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEMBS '04)*, vol. 1, pp. 1203–1206, San Francisco, Calif, USA, 2004.

[21] H. D. Zeman, G. Lovhoiden, and C. Vrancken, "Prototype vein contrast enhancer," in *Advanced Biomedical and Clinical Diagnostic Systems II*, vol. 5318 of *Proceedings of SPIE*, pp. 39–49, January 2004.

[22] T. F. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM, Philadelphia, Pa, USA, 2005.

[23] S. Esedoglu and J. Shen, "Digital inpainting based on the Mumford-Shah-Euler image model," *European Journal of Applied Mathematics*, vol. 13, no. 4, pp. 353–370, 2002.

[24] T. Vo-Dinh, *Biomedical Photonics Handbook*, CRC Press, New York, NY, USA, 2003.

[25] E. C. Lee, H. C. Lee, and K. R. Park, "Finger vein recognition using minutia-based alignment and local binary pattern-based feature extraction," *International Journal of Imaging Systems and Technology*, vol. 19, no. 3, pp. 175–178, 2009.

[26] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Transactions on Medical Imaging*, vol. 8, no. 3, pp. 263–269, 1989.

[27] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[28] R.-S. Ran and T.-Z. Huang, "The inverses of block tridiagonal matrices," *Applied Mathematics and Computation*, vol. 179, no. 1, pp. 243–247, 2006.

[29] M. E. A. El-Mikkawy, "On the inverse of a general tridiagonal matrix," *Applied Mathematics and Computation*, vol. 150, no. 3, pp. 669–679, 2004.

[30] C. M. da Fonseca and J. Petronilho, "Explicit inverses of some tridiagonal matrices," *Linear Algebra and Its Applications*, vol. 325, no. 1–3, pp. 7–21, 2001.

[31] G. Meurant, "A review on the inverse of symmetric tridiagonal and block tridiagonal matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 3, pp. 707–728, 1992.

[32] M. E. A. El-Mikkawy, "A note on a three-term recurrence for a tridiagonal matrix," *Applied Mathematics and Computation*, vol. 139, no. 2-3, pp. 503–511, 2003.

[33] G. H. Golub and C. F. van Loan, *Matrix Computation*, The Science Press, Beijing, China, 3rd edition, 2001.

[34] L. J. Ding, *Numerical Computing Method*, Science and Engineering University Press, Beijing, China, 1997.

[35] T. Chanwimaluang and G. Fan, "An efficient algorithm for extraction of anatomical structures in retinal images," in *Proceedings of the International Conference on Image Processing (ICIP '03)*, vol. 1, pp. 93–96, September 2003.

[36] N. R. Pal and S. K. Pal, "Entropic thresholding," *Signal Processing*, vol. 16, no. 2, pp. 97–108, 1989.

[37] N. R. Pal and S. K. Pal, Correction Note on "An efficient algorithm for extraction of anatomical structures inretinal images", http://www.vcipl.okstate.edu/Publications/TITBcorrection.pdf.

[38] R. Gongalez, R. Woods, and S. Eddins, *Digital Image Processing Using Matlab*, Prentice Hall, New York, NY, USA, 2003.

[39] Z. Weiwei and W. Yangsheng, "Core-based structure matching algorithm of fingerprint verification," in *Proceedings of the IEEE International Conference on Pattern Recognition,*, vol. 1, 2002, pp. 70–74.

[40] M. Nagao, *Methods of Image Pattern Recognition*, Corona, San Antonio, Tex, USA, 1983.

[41] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, 2000.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 265497, 9 pages
http://dx.doi.org/10.1155/2015/265497

Research Article

Methodology for Registration of Shrinkage Tumors in Head-and-Neck CT Studies

Jianhua Wang,[1] Jianrong Dai,[2] Yongjie Jing,[3] Yanan Huo,[4] and Tianye Niu[5]

[1]*Department of Radiation Oncology, Ningbo Treatment Center, Ningbo Lihuili Hospital, Ningbo, Zhejiang 315000, China*
[2]*Department of Radiation Oncology, Cancer Institute (Hospital), Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100021, China*
[3]*Key Laboratory of Particle and Radiation Imaging of Chinese Ministry of Education, Department of Engineering Physics, Tsinghua University, Beijing 100084, China*
[4]*The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang 310009, China*
[5]*Sir Run Run Shaw Hospital, Institute of Translational Medicine, Zhejiang University, Hangzhou, Zhejiang 310016, China*

Correspondence should be addressed to Tianye Niu; tyniu@zju.edu.cn

Received 31 October 2014; Revised 15 January 2015; Accepted 28 January 2015

Academic Editor: Panagiotis Mavroidis

Tumor shrinkage occurs in many patients undergoing radiotherapy for head-and-neck (H&N) cancer. However, one-to-one correspondence is not always available between voxels of two image sets. This makes intensity-based deformable registration difficult and inaccurate. In this paper, we describe a novel method to increase the performance of the registration in presence of tumor shrinkage. The method combines an image modification procedure and a fast symmetric Demons algorithm to register CT images acquired at planning and posttreatment fractions. The image modification procedure modifies the image intensities of the primary tumor by calculating tumor cell survival rate using the linear quadratic (LQ) model according to the dose delivered to the tumor. A scale operation is used to deal with uncertainties in biological parameters. The method was tested in 10 patients with nasopharyngeal cancer (NPC). Registration accuracy was improved compared with that achieved using the symmetric Demons algorithm. The average Dice similarity coefficient (DSC) increased by 21%. This novel method is suitable for H&N adaptive radiation therapy.

1. Introduction

Image registration is becoming a key tool in modern radiation therapy. In image-guided radiation therapy (IGRT), CT images acquired before each treatment are registered with planning CT images to verify patient position. In adaptive radiation therapy, similar registrations are needed to segment structures automatically and to evaluate the dose received in each fraction [1–3]. Rigid registrations are sufficient in situations where only rigid movements occur. However, deformable registrations are required in cases where structure deformation and/or tumor shrinkage occur.

The majority of patients with head-and-neck (H&N) cancer who undergo fractionated radiation therapy experience significant anatomic changes, such as tumors shrinkage, changes in overall body habitus, and weight loss [4–6]. Tumor shrinkage is mainly due to the death and destruction of cells killed by radiation. The loss of these cells causes tumor density and tumor volume to decrease. This means that physical one-to-one correspondence may no longer exist between voxels of the two image sets. This problem severely impairs the performance of an intensity-based deformable registration algorithm. The "Demons" family of algorithms is such a registration form, which can accurately account for anatomical changes at relatively low computational expense [7–10], and is a suitable choice for registration of H&N images [11].

The gas pocket mismatching problem in abdominal images also provides challenges for the intensity-based deformable image registration. Several methods have been used to overcome the problems associated with lack of correspondence and mismatched objects that occur with deformable

registration of abdominal CT images [12–16]. In images of the pelvic regions, the issue of no correspondence is associated with the presence of bowel gas. Because the gas is not clinically important (the contents of the bowel need not be registered), the artificial gas [12, 13] or constant intensity masks [14, 15] were introduced to create "virtual" correspondence. This improved the registration accuracy for the surrounding tissues, such as the rectal wall or prostate, which were clinically important organs. However, in the case of tumor shrinkage, the tumor itself is the treatment target. We could not use the constant intensity masks (or artificial intensity pattern) in the tumor, where the noncorrespondence problem occurred. For this reason, the approach mentioned above is infeasible. To our knowledge, the issue of lack of correspondence induced by the radiation therapy has not been adequately described in the literature.

In this paper, we describe a novel method for dealing with the effects of radiation on tumor tissues in the H&N during deformable registration. The method involved modification of image intensities of the primary tumor by calculating tumor cell survival rate using the linear quadratic (LQ) model according to radiation dose delivered to the tumor [17].

2. Materials and Methods

2.1. Data Extraction. The study included 10 patients with nasopharyngeal cancer (NPC) who underwent intensity-modulated radiation therapy (IMRT) at Cancer Institute (Hospital) of Chinese Academy of Medical Sciences, China. Patients were immobilized by custom-made thermoplastic masks. Two CT image sets were acquired for each patient (before the start of treatment and at the completion of the treatment course) using a Philips Brilliance Big Bore 16-slice CT scanner. The image sets were composed of 3 mm thick slices; the matrix size of each slice was 512 × 512 and the pixel size was approximately 1 mm. The initial image sets were loaded into a Pinnacle version 8.2g system (Philips Medical Systems, Cleveland, OH) for treatment planning. To improve the registration speed, the number of slices for each image pair was reduced so that it just covered the whole primary tumor. According to our standard treatment protocol, the primary tumor was prescribed 70–72.6 Gy. The tumor volume was contoured by a single physician for all planning and posttreatment CTs. Contouring was undertaken to evaluate the registration results, rather than to help the deformation process.

The CT images, contours, and treatment plan doses for each patient were exported from Pinnacle using the DICOM RT protocol. The exported dose array was a 3D dose map. The dose voxel size was $4 \times 4 \times 4\,\text{mm}^3$.

The study was approved by the local ethical committee, and informed consent was obtained from all patients.

2.2. Image Preparation. The posttreatment CT was used as the static image and the planning CT as the moving image for each image pair. This represented a worst-case scenario for tumor shrinkage to test our proposed algorithm. The couch region was manually selected in the transverse plane and the voxels CT numbers were set to that of air (i.e., −1000 HU). Voxels CT numbers below the empirically determined −500 HU were also set to −1000 HU. This was done to reduce the interference from nonuniformities outside the patient.

Both the planning CT images and the posttreatment CT images were cropped on the transverse plane, to restrict the regions of interest (ROI) to the head. The images were resampled and determined to have the same voxel dimensions of $2 \times 2 \times 3\,\text{mm}^3$ using the nearest interpolation.

2.3. Rigid Registration. Rigid registration was performed in advance to improve the speed and accuracy of the deformable registration. A minimum threshold of around 500 HU was applied to the two images so only the bony structures of the skeleton remained. The rigid registration determined a translation that minimized the correlation coefficient between voxels in the two images. To evaluate the accuracy of the rigid registration, we created a set of simulated translations of patient CT images. Compared to the introduced (known) shifts, the mean residual shifts (error) were 1.1 mm, within the voxel size of the patient images. This rigid alignment provided the basic initialization for subsequent deformable registration.

2.4. Deformable Registration. The "Demons" algorithm is an image intensity-based, deformable registration method that is widely used in medical practice as it demands relatively low computational expense. A variant of this algorithm, proposed by Wang, was used in the present study [18].

In the original Demons algorithm, the displacement for each voxel is obtained using the spatial gradient of static target image intensities. The displacement field is then regularized using a Gaussian smoothing filter to suppress noise and preserve the geometric continuity of the moving image. This iterative process alternates between calculation of the displacement field and regularization. By introducing an "active force," Wang et al. modified the standard Demons algorithm to obtain faster convergence and improve registration performance. The updated deformation field D_n for the current iteration was written as follows:

$$D_n = G_\sigma * \left(D_{n-1} + \left(I' - I\right) \cdot \left(\frac{\vec{\nabla}_S}{\left|\vec{\nabla}_I\right|^2 + k^2\left(I - I'\right)^2} + \frac{\vec{\nabla}_m}{\left|\vec{\nabla}_{I'}\right|^2 + k^2\left(I - I'\right)^2} \right) \right), \tag{1}$$

where G_σ is the Gaussian kernel, $*$ denotes the convolution operator, and the width of the Gaussian kernel σ was fixed to 1. D_{n-1} is the deformation field at iteration $n-1$. I' is the intensity of the moving image and I is the intensity of the

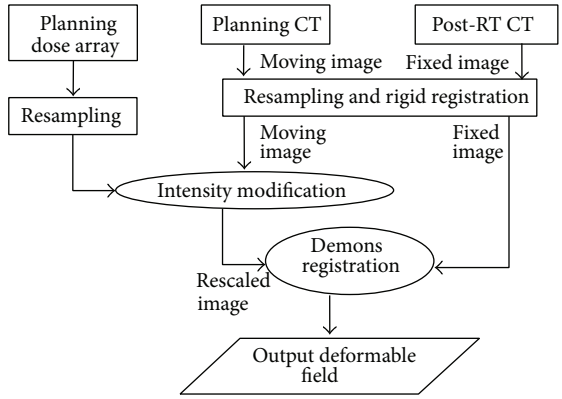


Figure 1: Method workflow consisting of the following procedures: rigid registration, intensity modification, intensity scale, and deformable registration.

static image; the respective gradient images are denoted by $\overrightarrow{\nabla}_{I'}$ and $\overrightarrow{\nabla}_{I}$. k is a normalization factor and we used $k = 0.4$.

2.5. Intensity-Modification Procedure (IMP). In patients undergoing radiation therapy, including H&N cancer, it is important that target volumes are accurately registered. Due to the cells killed by radiation, the density and even the volume of primary tumor may decrease. Unless this one-to-one correspondence is addressed, a large registration error will occur. We thus proposed to change the image intensities for the target volume in the planning CT image (moving image) based on the linear quadratic (LQ) model. The intensity modification was only done for the planning tumor and all other voxels of planning CT remained unchanged.

The workflow of the proposed method is shown in Figure 1. We implemented all procedures described here using in-house program written in Matlab (version 7.1, MathWorks) software. Based on the planning dose array, we modified the image intensities within the primary tumor using the LQ model. The details of this procedure are described below.

2.5.1. Data Import. The primary tumor volume contours as manually delineated during the routine treatment process were loaded into the planning CT images. The dose array was resampled to have the same spacing as the planning CT. The dose array and the planning CT were aligned according to their positions in the DICOM patient coordinate system. Thus, each element in the dose array represented the dose to be delivered to the corresponding voxel in the planning CT images.

2.5.2. Image Intensities Modification. We calculated voxel intensity within the primary tumor using the LQ model. For every slice in the planning CT images, I_0 represented the intensity of a voxel within the primary tumor. By definition, the relationship between I_0 (CT numbers, expressed in Hounsfield units) and the corresponding linear coefficient μ_0 was calculated as

$$I_0 = \frac{1000\,(\mu_0 - \mu_w)}{\mu_w}. \tag{2}$$

Rearranging this gives

$$\mu_0 = \frac{I_0 \mu_W}{1000} + \mu_W, \tag{3}$$

where μ_W is the attenuation coefficient of water (approximately 0.1928^{-1}).

After the primary tumor cells in a given voxel were irradiated at dose D, the intensity value would decease to I_S (HU) with the corresponding attenuation coefficient μ_S. I_S was converted to μ_S by the equation

$$\mu_S = \frac{I_S \mu_W}{1000} + \mu_W. \tag{4}$$

As shown in (5), μ_S is proportional to the number of primary tumor cells N_S in a given voxel that survives irradiation, and μ_0 is proportional to the number of cells in that voxel prior to irradiation N_0:

$$\mu_S = \frac{\mu_0 N_S}{N_0} = \mu_0 \mathrm{SF}, \tag{5}$$

where, according to the LQ model, the survival fraction SF is given by

$$\mathrm{SF} = \exp\left(-\alpha D\left(1 + \frac{d}{\alpha/\beta}\right)\right). \tag{6}$$

In this situation, cell proliferation is negligible. d and D are the dose per fraction and the total dose to the voxel, respectively. α and β are the LQ parameters. For the purposes of illustration, we assumed that α was $0.33\,\mathrm{Gy}^{-1}$ and α/β was 10 Gy [19–22]. Note that the result is not sensitive to the variation of α and α/β values as demonstrated in the result section. Using (3), (4), (5), and (6), we obtained

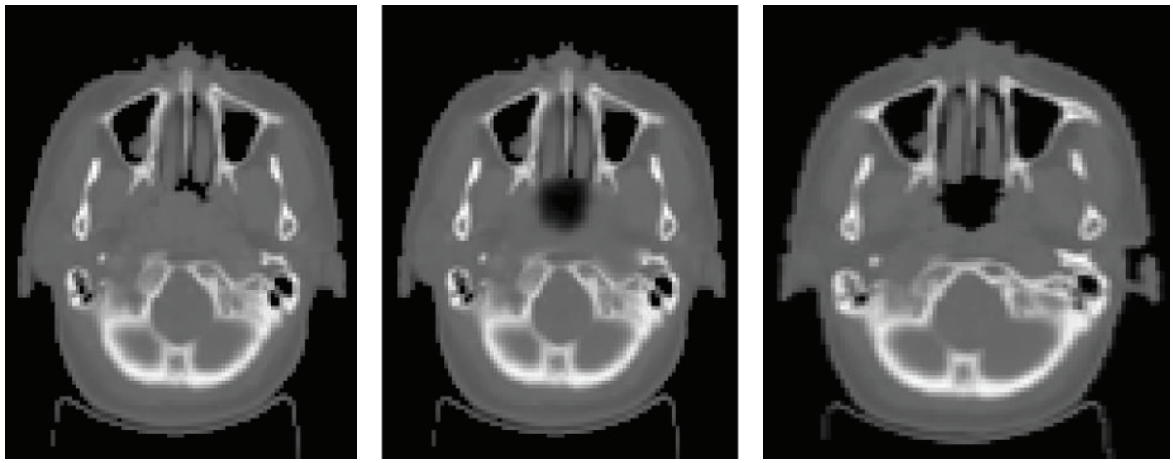
$$I_S = \left[\left(\frac{I_0 \mu_W}{1000} + \mu_W\right)\exp\left(-\alpha D\left(1 + \frac{d}{\alpha/\beta}\right)\right) - \mu_W\right] \times \frac{1000}{\mu_W}. \tag{7}$$

2.5.3. Voxel Intensity Scale. In the planning CT image, S_0 was defined as the sum of voxel values within the primary tumor, and S denoted the sum of the voxel values within the corresponding region in the posttreatment CT image. The intensity value (I_S) of the primary tumor voxel was scaled as follows:

$$I' = \frac{I_S S}{S_0}. \tag{8}$$

Equations (7) and (8) were used to modify the intensity of every voxel of primary tumor in the planning CT images. Figure 2 shows an axial slice of a planning image before and after image modification. This procedure was coupled with Wang's "Demons" algorithm to calculate the displacement field.

2.6. Evaluation. To quantitatively evaluate the performance of the proposed method, the Dice similarity coefficient (DSC) was calculated for the tumor. For two segmentations on the



(a) (b) (c)

Figure 2: Example of intensity-modification procedure. (a) An axial slice of a planning image. (b) The same image after intensity modification. This modified planning image can be accurately registered using deformable image registration. (c) The corresponding slice from the posttreatment CT image.

target images, given by the deformed contours using the computed motion fields and manual contours, respectively, their corresponding volumes were denoted by V_d and V_m. The DSC was then defined as [23]

$$\text{DSC} = \frac{2V_d \cap V_m}{V_d + V_m} \times 100\%. \quad (9)$$

DSC ranged between 0 and 100%. A DSC value of 0 indicated two completely uncorrelated images and a DSC value of 100% indicated a perfect match.

To assess the effectiveness of the proposed intensity modified procedure (IMP), we compared results of deformable image registration with and without intensity modification when all other parameters were set to the same. We also calculated the DSC for the tumor performed only with the rigid registration between these 10 pairs of CT images. The Wilcoxon signed-rank test was used to compare each method.

Table 1: Statistics for Dice similarity coefficient (DSC) for the tumor using rigid registration and deformable image registration with and without the intensity-modification procedure (IMP).

Patient	DSC (rigid)	DSC (IMP)	DSC (no IMP)
1	86.7	91.6	88.0
2	78.9	95.1	82.3
3	73.5	90.9	60.7
4	86.9	96.8	86.2
5	52.5	78.2	63.4
6	71.5	89.4	53.9
7	76.8	94.0	82.5
8	79.3	95.8	75.1
9	80.4	95.2	87.5
10	76.0	93.3	80.7
Mean	76.3	92.0	76.0

3. Results

3.1. Registration Example. To compare the effectiveness of the intensity modification, we performed deformable image registration in 10 pairs of head CT images with and without the intensity-modification procedure. An example is shown in Figure 3. The target volume change for this case was found to be 33% (from 15.59 cc to 10.47 cc). The left row of Figure 3 shows the planning CT in axial and sagittal views together with manually delineated tumor volume (green contours). The right row shows the posttreatment CT. Both sets of contours were overlaid in these images: the deformed contours without IMP (shown in blue) and the deformed contours with IMP (shown in red). As it can be seen in the figure, the deformed contours without using IMP clearly did not match with the reduced tumor target, due to the lack of one-to-one correspondence.

Figure 4 shows the dense displacement fields with and without IMP. The arrows indicate the displacement in 3D but are projected as 2D images for display purposes. A vector field was used to assess the result and detect errors. It is evident from these displays that the displacement vectors (within the smaller region around the primary tumor where noticeable shrinkage occurs) are more chaotic, discontinuous, and abrupt, when IMP is not applied (Figure 4(a)). The voxels in this region exhibit unrealistic displacement due to the lack of correspondence. This is improved when the deformable registration is embedded with the intensity-modification procedure (Figure 4(b)).

3.2. Registration Statistics. The DSC values calculated for the tumor using rigid and deformable registration with and without IMP are summarized in Table 1. The Wilcoxon signed-rank test showed that there was little difference between the rigid registration and the deformable registration method without using the IMP method ($P = 0.721$). Nevertheless, there was a significant improvement when the IMP method

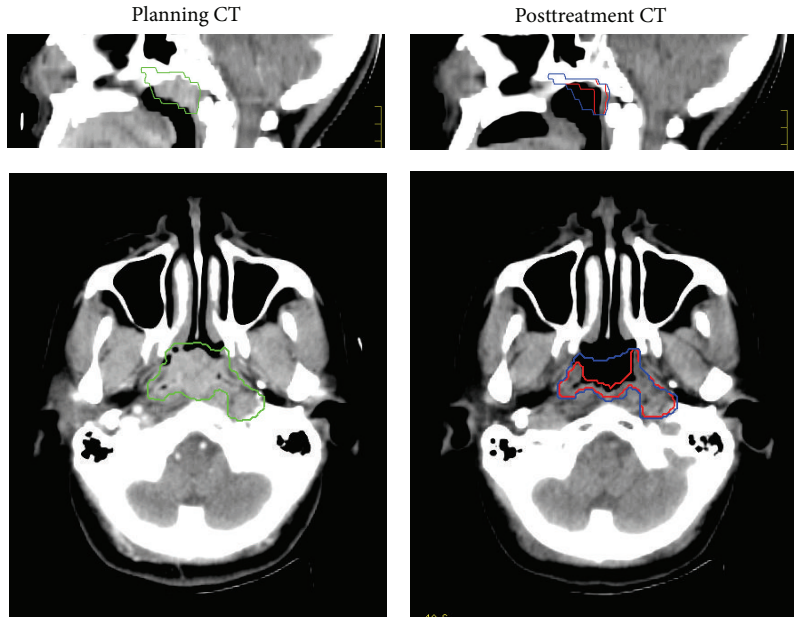


Figure 3: The CT images of one patient with the registered contours overlaid. (a) Planning CT in sagittal and axial slices and the manual contours overlaid. (b) The corresponding slices of posttreatment CT with the deformed contours overlaid. The red and blue contours represent the segmentation results with (red) and without (blue) the IMP.

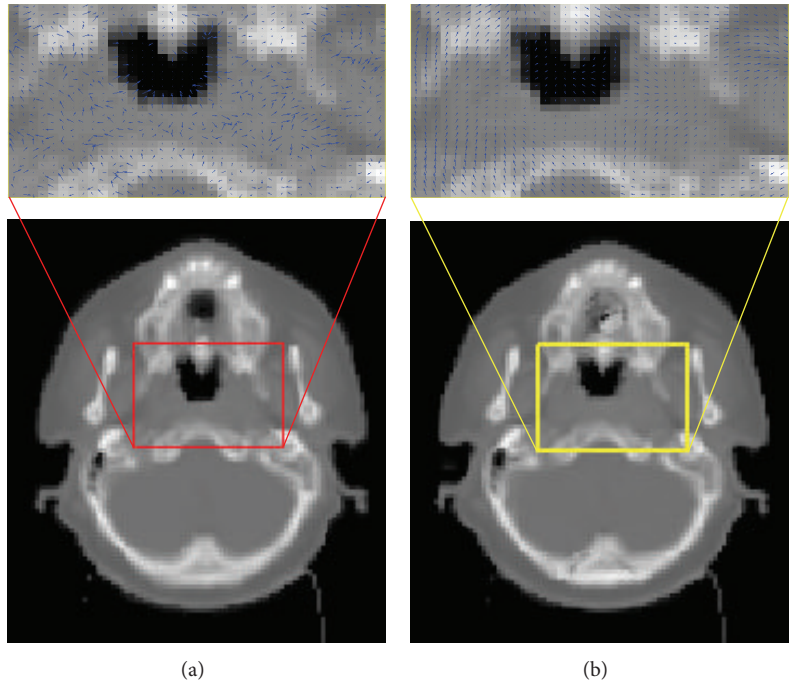


Figure 4: Displacement field calculated using the deformable registration algorithm overlaid on an axial slice. Displacement field: (a) computed without using the intensity-modification procedure (IMP) and (b) computed using the IMP.

was used over the rigid registration ($P < 0.005$) or the deformable registration method without using the IMP method ($P < 0.005$). The images with rigid registration had a mean overlap value of DSC = 76.3%. The mean value of the calculated DSC was 76.0% for deformable registration without IMP. The application of the IMP leads to a mean value of DSC = 92.0%. On average, the improvement in DSC was 21% with IMP in these 10 cases.

In this study, hypotheses were made on the values of α and α/β ratio. The uncertainty in radiosensitivity parameters may

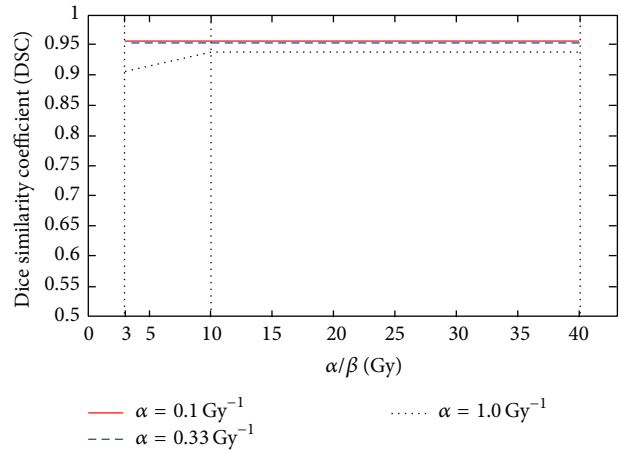


Figure 5: Sensitivity of the Dice similarity coefficient (DSC) with respect to α and α/β for patient 2.

affect the reliability of the outcome. Therefore, a sensitivity analysis was performed to quantify the uncertainty of DSC when different parameter combinations were applied in the study. A typical example is shown in Figure 5. We adjusted the key parameters (α and α/β) in a large range of possible values and provided the DSC results. It indicated that our method is not sensitive to the radiosensitivity parameters. For example, DSC slightly increased from 93.6% to 95.3%, even with more than one order of magnitude change on α and α/β values (α from 0.1 Gy^{-1} to 1 Gy^{-1} and α/β from 3 Gy to 40 Gy).

The computing time to perform the deformable registration varied depending on the number of registered image slices. For a typical planning and posttreatment CT pair, the run time for the full registration was approximately 15 min. These results were obtained on a Dell desktop PC with dual 2.13 GHz Intel Xeon processors and 2.25 GB of RAM.

As our focus was the accuracy of registration, we made no special effort to improve the speed of the registration process. However, time saving could be achieved by improving computer hardware such as GPU-based implementation of the Demons algorithm [24, 25].

4. Discussion

Previous studies have shown that nonrigid anatomical changes can occur in H&N during the course of fractionated radiation therapy [26]. The Demons algorithm has been shown to be a good choice for registration of H&N images [11]. However, the issue of primary tumor density change/volume shrinkage due to radiation treatment has not been addressed by the authors. In our experience, registration accuracy using this algorithm is reduced under these circumstances.

In the current study, we used a modification to the Demons algorithm which adjusted the intensities of the primary tumor voxels according to the LQ model. Qualitative and quantitative results demonstrated that the proposed method increased the performance of the registration in presence of tumor shrinkage.

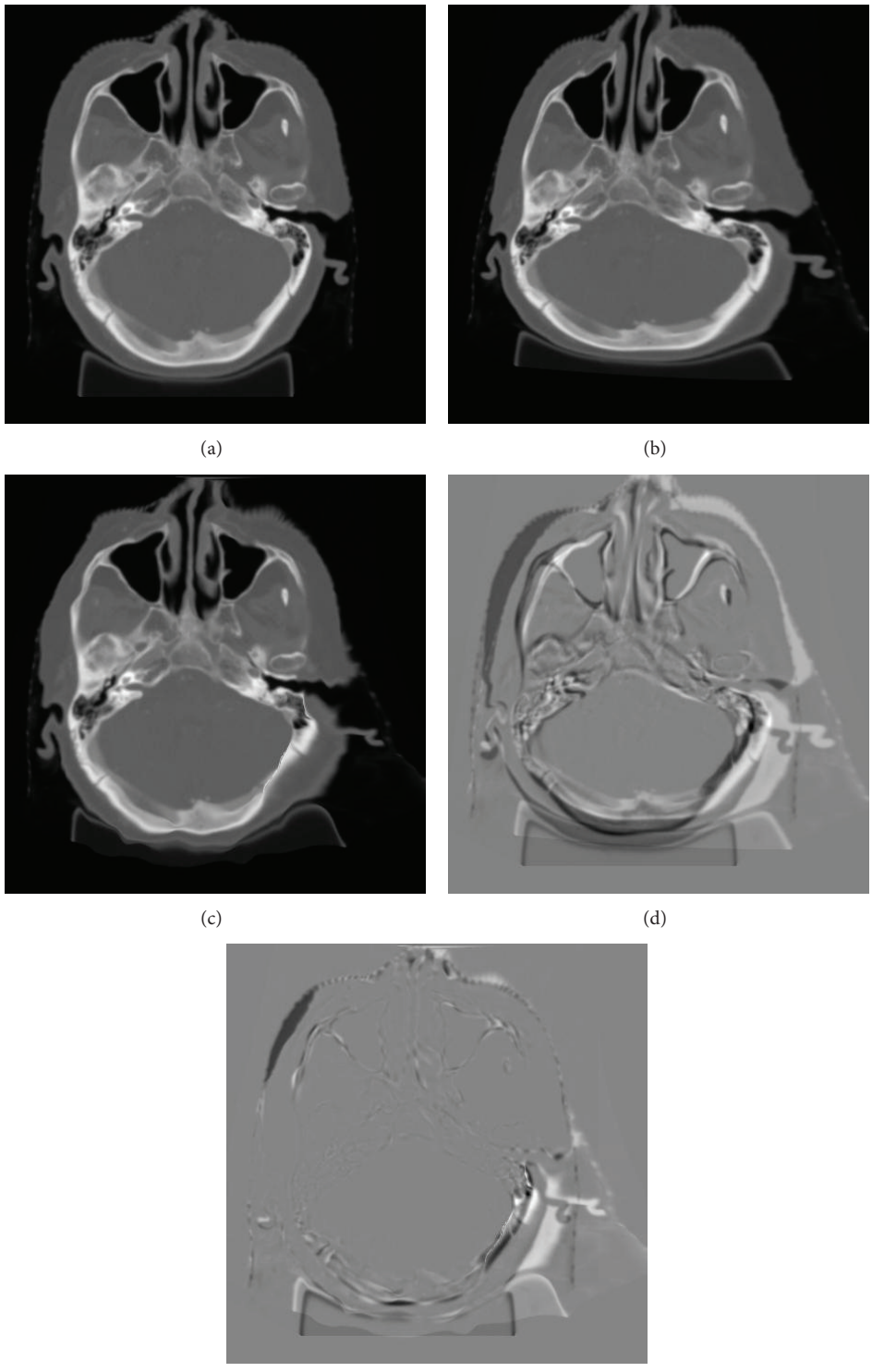
The intensity-modification procedure used in this study is essentially a preprocessing method that can be used in conjunction with other intensity-based deformable registration algorithms. While the focus of the current work was to evaluate deformable registration in H&N region, the method may also be adapted to other anatomical regions where tumor shrinking occurs (e.g., in lung cancer). In lung cancer, noncoplanar beams may be used and more slices should be included along both the superior and inferior directions to the tumor.

The LQ model is generally effective in describing tumor response to radiation and is widely used in experimental and clinical radiobiology. So we chose this model to calculate the NPC cell killing by radiation. The cell killing process is a complicated biological process. However, this preliminary model could serve as a basis for more complex models to deal with the problem of tumor shrinkage. Considering the nonhomogeneous dose distributions in IMRT, we assumed that the primary tumor volume was composed of a series of subvolumes (voxels), each one receiving a homogeneous dose. In this setting, the basic LQ model may be used to estimate the NPC cell killing by radiation. We also recognize that further studies would be desirable to validate this assumption.

The radiation therapy parameters values used in our study were derived from the literature [19–22]. The ratio α/β was assumed to be 10 Gy, which is a nominal value for most tumors, and α was set at 0.33 Gy^{-1}. Radiobiology parameters have a high degree of uncertainty caused by interpatient variations, tumor heterogeneity, and effects of hypoxia and chemotherapy. To deal with this uncertainty, we used a scale procedure within the LQ model to fix the potential problems caused by the patient and tumor specific biological parameters. The scale adjustments were based on information from pretreatment CT images. The sensitivity analysis showed that our method was not sensitive to the radiosensitivity parameters. It may also be possible to obtain biological information for individual patients noninvasively using functional imaging [27, 28].

In the formulation for the deformation field, we have chosen σ to be 1 and k to be 0.4 based on previous studies done by other investigators [7, 8, 11, 14, 18] and our own initial testing experience. The impacts of the σ and k parameters were not investigated thoroughly in this preliminary study. However, these specific values work well in most of our cases. Adjusting parameters by a trial-and-error method was also tested. No significant improvement was observed compared to the fixed values. In addition, our emphasis is on the comparison of the various schemes and not the final performance. Therefore, we used the same set of parameters for all experiments, without multiresolution adjustments.

Validation of deformable image registration remains a difficult task due to the lack of the ground truth. Our deformable registration algorithm was validated by simulating the deformation on patient CT image. We applied a 2nd-order polynomial transformation to the original H&N image (Figure 6(a)) and deformed its shape by more than 5 mm on average (Figure 6(b)). Our algorithm automatically generated the deformation field and deformed the original image to



(a) (b) (c) (d) (e)

Figure 6: (a) Original image; (b) original image transformed by a "polynomial" transformation; (c) deformed image derived from (a) using the Demons algorithm; (d) difference image of (b) and (a); (e) difference image of (b) and (c).

the mathematically transformed one (Figure 6(c)). The difference images (Figures 6(d) and 6(e)) were, respectively, Figures 6(b)-6(a) and Figures 6(b)-6(c). Figure 6(e) showed that the original CT image registered with the mathematically deformed CT image with little difference. Quantitative validation results showed that more than 90% of the voxels were within 2 mm of their intended shifts. Future work includes a further improvement to the performance of IMP method by implanting fiducial markers into the patient for further validation.

5. Conclusion

We developed and tested a novel method for performing deformable registration between planning and posttreatment CT images in the H&N region. The technique was able to account for tumor responses to radiation therapy by modifying image intensities of the primary tumor voxels according to the LQ model and to deal with the interpatient heterogeneity of radiobiology parameters by a scaling factor. The preliminary tests resulted in higher registration accuracy than current methods, indicating a role in H&N adaptive radiation therapy.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (NSFC Grants nos. 81201091 and 81200662), the Fundamental Research Funds for the Central Universities in China, the Fund Project for Excellent Abroad Scholar Personnel in Science and Technology, and the Natural Science Foundation of Zhejiang Province (Grant no. LY12H12010).

References

[1] D. Sarrut, "Deformable registration for image-guided radiation therapy," *Zeitschrift fur Medizinische Physik*, vol. 16, no. 4, pp. 285–297, 2006.

[2] J. Orban de Xivry, P. Castadot, G. Janssens et al., "Evaluation of the radiobiological impact of anatomic modifications during radiation therapy for head and neck cancer: can we simply summate the dose?" *Radiotherapy and Oncology*, vol. 96, no. 1, pp. 131–138, 2010.

[3] P. Castadot, X. Geets, J. A. Lee, and V. Grégoire, "Adaptive functional image-guided IMRT in pharyngo-laryngeal squamous cell carcinoma: is the gain in dose distribution worth the effort?" *Radiotherapy and Oncology*, vol. 101, no. 3, pp. 343–350, 2011.

[4] J. L. Barker Jr., A. S. Garden, K. K. Ang et al., "Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system," *International Journal of Radiation Oncology Biology Physics*, vol. 59, no. 4, pp. 960–970, 2004.

[5] E. K. Hansen, M. K. Bucci, J. M. Quivey, V. Weinberg, and P. Xia, "Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer," *International Journal of Radiation Oncology Biology Physics*, vol. 64, no. 2, pp. 355–362, 2006.

[6] Q. Wu, Y. Chi, P. Y. Chen, D. J. Krauss, D. Yan, and A. Martinez, "Adaptive replanning strategies accounting for shrinkage in head and neck IMRT," *International Journal of Radiation Oncology Biology Physics*, vol. 75, no. 3, pp. 924–932, 2009.

[7] J. P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[8] H. Wang, L. Dong, J. O'Daniel et al., "Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy," *Physics in Medicine and Biology*, vol. 50, no. 12, pp. 2887–2905, 2005.

[9] R. Varadhan, G. Karangelis, K. Krishnan, and S. Hui, "A framework for deformable image registration validation in radiotherapy clinical applications," *Journal of Applied Clinical Medical Physics*, vol. 14, no. 1, article 4066, 2013.

[10] N. Stanley, C. Glide-Hurst, J. Kim et al., "Using patient-specific phantoms to evaluate deformable image registration algorithms for adaptive radiation therapy," *Journal of Applied Clinical Medical Physics*, vol. 14, no. 6, pp. 177–194, 2013.

[11] P. Castadot, J. A. Lee, A. Parraga, X. Geets, B. Macq, and V. Grégoire, "Comparison of 12 deformable registration strategies in adaptive radiation therapy for the treatment of head and neck tumors," *Radiotherapy and Oncology*, vol. 89, no. 1, pp. 1–12, 2008.

[12] M. Foskey, B. Davis, L. Goyal et al., "Large deformation three-dimensional image registration in image-guided radiation therapy," *Physics in Medicine and Biology*, vol. 50, no. 24, pp. 5869–5892, 2005.

[13] S. Gao, L. Zhang, H. Wang et al., "A deformable image registration method to handle distended rectums in prostate cancer radiotherapy," *Medical Physics*, vol. 33, no. 9, pp. 3304–3312, 2006.

[14] A. Godley, E. Ahunbay, C. Peng, and X. A. Li, "Automated registration of large deformations for adaptive radiation therapy of prostate cancer," *Medical Physics*, vol. 36, no. 4, pp. 1433–1441, 2009.

[15] D. Yang, S. R. Chaudhari, S. M. Goddu et al., "Deformable registration of abdominal kilovoltage treatment planning CT and tomotherapy daily megavoltage CT for treatment adaptation," *Medical Physics*, vol. 36, no. 2, pp. 329–337, 2009.

[16] B. Rodriguez-Vila, F. Garcia-Vicente, and E. J. Gomez, "Methodology for registration of distended rectums in pelvic CT studies," *Medical Physics*, vol. 39, no. 10, pp. 6351–6359, 2012.

[17] J. F. Fowler, "The linear-quadratic formula and progress in fractionated radiotherapy," *British Journal of Radiology*, vol. 62, no. 740, pp. 679–694, 1989.

[18] H. Wang, L. Dong, M. F. Lii et al., "Implementation and validation of a three-dimensional deformable registration algorithm for targeted prostate cancer radiotherapy," *International Journal of Radiation Oncology Biology Physics*, vol. 61, no. 3, pp. 725–735, 2005.

[19] M. Avanzo, J. Stancanello, G. Franchin et al., "Correlation of a hypoxia based tumor control model with observed local control rates in nasopharyngeal carcinoma treated with chemoradiotherapy," *Medical Physics*, vol. 37, no. 4, pp. 1533–1544, 2010.

[20] R. M. Wyatt, A. H. Beddoe, and R. G. Dale, "The effects of delays in radiotherapy treatment on tumour control," *Physics in Medicine and Biology*, vol. 48, no. 2, pp. 139–155, 2003.

[21] S. Webb, "Optimum parameters in a model for tumour control probability including interpatient heterogeneity," *Physics in Medicine and Biology*, vol. 39, no. 11, pp. 1895–1914, 1994.

[22] J. F. Fowler, "Is there an optimum overall time for head and neck radiotherapy? A review, with new modeling," *Clinical Oncology*, vol. 19, no. 1, pp. 8–22, 2007.

[23] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[24] R. Castillo, E. Castillo, R. Guerra et al., "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Physics in Medicine and Biology*, vol. 54, no. 7, pp. 1849–1870, 2009.

[25] G. C. Sharp, N. Kandasamy, H. Singh, and M. Folkert, "GPU-based streaming architectures for fast cone-beam CT image reconstruction and demons deformable registration," *Physics in Medicine and Biology*, vol. 52, no. 19, pp. 5771–5783, 2007.

[26] S. Nithiananthan, K. K. Brock, M. J. Daly, H. Chan, J. C. Irish, and J. H. Siewerdsen, "Demons deformable registration for CBCT-guided procedures in the head and neck: convergence and accuracy," *Medical Physics*, vol. 36, no. 10, pp. 4755–4764, 2009.

[27] S. M. Bentzen and V. Gregoire, "Molecular imaging-based dose painting: a novel paradigm for radiation therapy prescription," *Seminars in Radiation Oncology*, vol. 21, no. 2, pp. 101–110, 2011.

[28] F. Z. Yetkin and D. Mendelsohn, "Hypoxia imaging in brain tumors," *Neuroimaging Clinics of North America*, vol. 12, no. 4, pp. 537–552, 2002.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 152693, 9 pages
http://dx.doi.org/10.1155/2015/152693

Review Article

Recent Development of Dual-Dictionary Learning Approach in Medical Image Analysis and Reconstruction

Bigong Wang[1,2] and Liang Li[1,2]

[1]*Department of Engineering Physics, Tsinghua University, Beijing 100084, China*
[2]*Key Laboratory of Particle & Radiation Imaging (Tsinghua University), Ministry of Education, Beijing 100084, China*

Correspondence should be addressed to Liang Li; lliang@tsinghua.edu.cn

Received 5 October 2014; Revised 12 January 2015; Accepted 6 April 2015

Academic Editor: Valeri Makarov

As an implementation of compressive sensing (CS), dual-dictionary learning (DDL) method provides an ideal access to restore signals of two related dictionaries and sparse representation. It has been proven that this method performs well in medical image reconstruction with highly undersampled data, especially for multimodality imaging like CT-MRI hybrid reconstruction. Because of its outstanding strength, short signal acquisition time, and low radiation dose, DDL has allured a broad interest in both academic and industrial fields. Here in this review article, we summarize DDL's development history, conclude the latest advance, and also discuss its role in the future directions and potential applications in medical imaging. Meanwhile, this paper points out that DDL is still in the initial stage, and it is necessary to make further studies to improve this method, especially in dictionary training.

1. Introduction

Compressive sensing (CS) is a novel theory in information acquisition and processing [1]. Since general signals are broadband, traditional signal reconstruction methods usually adopt Nyquist Sampling, requiring high sample rate and long processing time. However, CS theory offers a way to restore signal accurately with less measurement by solving an optimization problem in which signal is sparse, represented using a basis matrix, and the high-dimensional transformation is projected to a lower dimensional subspace. Therefore, CS theory has been widely recognized and applied in various fields.

Some groups focus on studies of CS applications and have developed various braches such as Bayesian CS and 1-Bit CS [2–4]. After it is applied in medical imaging reconstruction, CS theory is proven to be a method that effectively retains high image quality using undersampling measurement data in different imaging modalities including computed tomography (CT) and magnetic resonance imaging (MRI) [5–7]. Besides, CS theory shows great potential in multimodalities image reconstruction, one of the future directions of medical imaging.

Dictionary learning (DL) is a typical method of CS image reconstruction. In this method, sampled data is compressible in specific transform domain, and transformation coefficients are projected to a lower dimensional vector with essential image information retained well. As a result, complex reconstruction problem is simplified to an optimization problem. Usually, one should take three problems into consideration to solve image reconstruction problems using DL methods. First, design an overcompleted dictionary which can represent a signal sparsely. Second, get a measurement matrix strictly satisfied with isometry property. Third, develop a fast signal reconstruction algorithm with good robustness. The designed dictionary is important to the accuracy of CS image reconstruction. In DL method, the dictionary is self-adaptive and flexible; it is trained by particular image samples or group of images. Using different training methods, the image sparseness is quite different [8].

Though DL-based approach has been recognized in medical image reconstruction field, single dictionary applied in the whole image process brings out a limit in image quality. That means only one dictionary is far from enough as the prior information. In order to improve image quality, research scholars have optimized DL method to

dual-dictionary learning (DDL) which has more diverse prior information in imaging modalities like CT and MRI. DDL method was initially developed for image super-resolution. Lu et al. [9, 10] applied this method for CT reconstruction. Song et al. [11] used it in 3D MRI reconstruction. DDL shows a great potential in medical image reconstruction.

In this paper, we discuss the DL method in Section 2. Based on DL method, we review DDL's history and new development in Section 3, including its theory, feasibility demonstration, and the application in different fields. In Section 4, we discuss the use of DDL in medical image analysis. In the section of Discussion and Conclusion, we summarize algorithms and explore the future directions in medical image reconstruction.

2. Dictionary Learning (DL) Algorithm

2.1. DL Method and Theory. According to the CS theory, an undersampling image reconstruction problem is to solve an underdetermined system of linear equations $F_u x = y$ by minimizing the l_0 quasi norm (e.g., number of nonzeros) of the sparsified transform Ψx; it means the image x is sparse after a completed sparse transform $\Psi \in \Re^{M\times N}$. The corresponding optimization problem is

$$\min_x \|\Psi x\|_0 \quad \text{s.t. } F_u x = y. \tag{1}$$

In (1), x is the image to be reconstructed, F_u is the codebook for the given measurements y. Equation (1) is also known as a sparse coding problem, which is a NP-hard problem (nondeterministic polynomial). It can be solved by some greedy algorithms, for example, orthogonal matching pursuit (OMP) [12]. It is notable that if the l_0 norm is replaced with l_1 norm, the problem can be solved by linear programming in the real domain or second order cone programming in the complex domain.

Given an image X of size $N \times N$, it can be decomposed into some small patches of size $b \times b$, $b \ll N$. Each patch can be expressed as a $n = b^2$ dimensional vector $\mathbf{x} \in \Re^n$. All the patches are extracted from the object image X according to the patch size and the slide distance. A dictionary $D \in \Re^{n\times K}$ is a matrix that consists of K atoms $d_k \in \Re^{n=b\times b}$ which are the columns of the dictionary. As d_k is the patch vector from sample images, the initial dictionary constructed from the extracted patches is usually redundant or overcompleted; that is, $N \ll K$. Using specific atoms of initial dictionary D, each vector $\mathbf{x}$ in the image can be approximately represented as sparse coefficient [13]. Consider

$$\|\mathbf{x} - D\boldsymbol{\alpha}\|_2^2 < \varepsilon, \tag{2}$$

where $\varepsilon > 0$ for the error bound and $\boldsymbol{\alpha} \in \Re^K$ for the sparse representation vector which has few nonzero elements: $\|\boldsymbol{\alpha}\|_0 \ll N \ll K$, $i = 1, 2, \ldots, K$. To get the sparse representation of the vector $\mathbf{x}$, one can minimize the l_0 norm as

$$\min_\alpha \|\boldsymbol{\alpha}\|_0 \quad \text{s.t. } \|\mathbf{x} - D\boldsymbol{\alpha}\|_2^2 < \varepsilon. \tag{3}$$

If an image contains S patches, DL is to find a dictionary $\widetilde{D}$ in which all the patches should be sparsely represented as follows:

$$\min_{D,\alpha} \sum_{s=1}^{S} \left(\left\|\mathbf{x}_s - \widetilde{D}\boldsymbol{\alpha}_s\right\|_2^2 + \nu \|\boldsymbol{\alpha}_s\|_0 \right). \tag{4}$$

Usually, if ν is fixed by specific value, (3) is equivalent to solve the following problem:

$$\min_{D,\alpha} \sum_{s=1}^{S} \left(\left\|\mathbf{x}_s - \widetilde{D}\boldsymbol{\alpha}_s\right\|_2^2 \right) \quad \text{s.t. } \|\boldsymbol{\alpha}_s\|_0 < T_0. \tag{5}$$

2.2. Dictionary Construction. DL problem is NP-hard because it turns to a sparse coding problem when D and $\mathbf{x}$ are fixed. Currently, mainly four adaptive dictionary training algorithms were proposed to solve such a dictionary learning problem.

(1) Direct method (DM): DM is an original method that preserves all the details in the sample images because of a direct extraction process, and then a target image can be fully recovered as the patches are well chosen. Usually, this method is effective in super-resolution image reconstruction.

(2) Method of optimal directions (MOD): MOD fixes the coefficients corresponding to the dictionary vectors and then updates the atoms by minimizing the residuals between the training vectors and its representations. The main advantage of MOD is that it gives the optimal adjustment of the dictionary vectors in each iteration. Usually, it provides better convergence properties in ECG (electrocardiogram) signals [14].

(3) Generalized principal component analysis (GPCA): GPCA is a general method for modeling and segmenting some mixed data using a collection of subspaces. By introducing certain algebraic models and techniques into data clustering, traditionally a statistical problem, GPCA offers a new spectrum of algorithms for data modeling and clustering [15].

(4) K-means singular value decomposition (K-SVD): K-SVD is an iterative method updating the dictionary atoms to fit the data better. The method does SVD on the errors and updates the current dictionary atom and coefficient simultaneously with the item which has the minimum error. As the most widely used method to train the dictionary, K-SVD has an excellent convergence and sparsity [16].

Dictionary learning can be used to reconstruct image; a classic algorithm is summarized in Figure 1. Given an initial value x_0 (initial dictionary), do dictionary learning using appropriate training method and obtain the sparse representation, and then update x under specific transform (i.e., wavelet, Fourier) and output the result after several iterations at last.

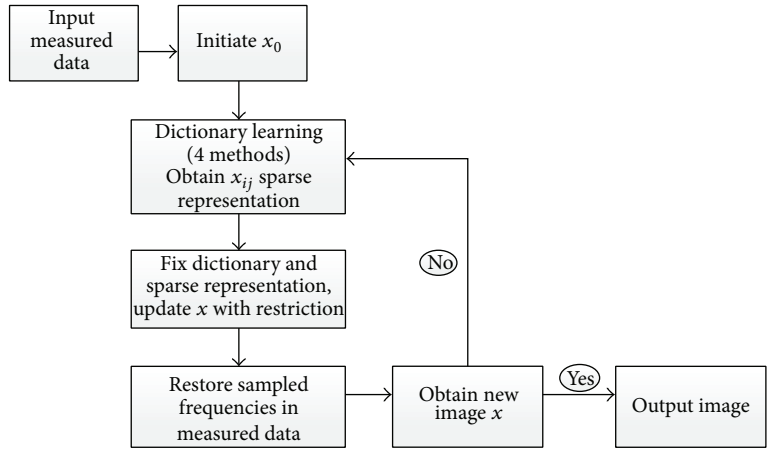


Figure 1: The algorithm block diagram of diction learning applied in image reconstruction.

3. DDL Algorithm in Image Analysis

3.1. From Single to Dual-Dictionary. DL method is widely used in image restoration [17–19], super-resolution reconstruction [20–23], image deblurring [24–26], denoising [27–32], medical image reconstruction [13, 33], image prediction [34], and image inpainting [35]. However, both dynamitic atoms in each iteration step and certain noise in measurement data would increase iteration time making DL method slow in most cases. As to improve DL's inefficiency, some come up with the solution that by introducing two or more dictionaries image quality would be further improved within less time. One of the improved methods is dual-dictionary learning (DDL).

DDL theory is first introduced by Curzion et al. as PADDL; it aimed to train a linear mapping in the case of a single dictionary. Note that this method is not using two different dictionaries but training one dictionary with its "dual" dictionary. In PADDL method, the essential concept is to update the dictionary $D = [d_1, \ldots, d_K] \in \mathfrak{R}^{d\times K}$ by means of its "dual" dictionary $C = [c_1, \ldots, c_K]^T \in \mathfrak{R}^{K\times d}$, as an auxiliary item. It aims to find an optimal pair of linear operators D by minimizing the following:

$$E(D, C, U) = \|X - DU\|_F^2 + \eta \|U - CX\|_F^2 + \tau \|U\|_1$$
$$\text{s.t. } \|d_i\|^2, \|c_i\|^2 \le 1, \tag{6}$$

where $X \in \mathfrak{R}^{d\times N}$ is the matrix to be trained and $U \in \mathfrak{R}^{K\times N}$ is the representation. The c_i can be treated as filters to approximate its optimal u. η is the weight parameters.

The result shows that this dual-dictionary training method can be applied well in calculating the sparse representations [36].

3.2. DDL in Super-Resolution Reconstruction. Zhang et al. proposed an efficient sparse representation method to solve image super-resolution reconstruction via DDL [37]. In this work, they assume that image patches with different resolution can share the same underlying sparse representation. Thus, given a dictionary pair $\{D_h, D_l\}$, where h stands for high resolution and l stands for low resolution, the sparse representation of x_i from low-resolution image X_l is similar as (3). Consider

$$\widehat{z}_i = \arg\min \|z_i\|_1 \quad \text{s.t. } \|D_l z_i - x_i\|_2^2 \le \varepsilon. \tag{7}$$

With the sparse representation vector z_i, the high-resolution patch can be approximately expressed as $x_i' = D_h z_i$. Put all the high-resolution patches back into corresponding positions and perform normalization. Finally we obtain the estimation of the high-resolution image x'.

The optimization model for learning coupled dictionaries with "dual" is as follows:

$$\{D_h, D_l, C_l, Z\} = \arg \min_{\{D_h, D_l, C_l, Z\}} \|X_c - D_c Z\|_F^2 + \eta \|Z - C_l X_l\|_F^2 + \lambda \|Z\|_1$$
$$\text{s.t. } \|D_i\|_2^2, \|C^i\|_2^2 \le 1. \tag{8}$$

$X_c = [(1/\sqrt{M}){X_h}^T, (1/\sqrt{N}){X_l}^T]^T \in \mathfrak{R}^{(M+N)\times L}$, in which M and N are the dimension of the high- and low-resolution patches. $D_c = [(1/\sqrt{M}){D_h}^T, (1/\sqrt{N}){D_l}^T]^T \in \mathfrak{R}^{(M+N)\times K}$. $C_l \in \mathfrak{R}^{K\times N}$ is the dual of D_l as mentioned in Section 2.1. After multiplying $z_i = C_l x_i$ by D_h, we acquire the high-resolution patch x_i'. In this method, D_l and D_h are treated as one dictionary and trained simultaneously with their dual, which refers to D_l and D_h.

With the approximate sparse coding procedure via model (8), the result shows that their method speeds up the overall super-resolution process significantly.

3.3. DDL in Image Restoration. Similar to HaiChaos' work, Wang et al. also applied DDL in image restoration [38].

They solved the problem of restoring the lost part of high-frequency detail information of images.

Wang et al. reconstructed the high-frequency (HF) details from the low-resolution images using the prior models. HF is decomposed into a combination of two components, main high-frequency (MHF) and residual high-frequency (RHF). Wang et al. restored MHF and RHF, respectively, with dual-dictionary and then added up MHF and RHF at last. For dictionary construction, K-SVD was used to train the two dictionaries. The experiment result reveals that the PSNR values are better than bicubic and sparse representation algorithm.

3.4. DDL in Human Pose Estimation. Ji and Su proposed a new method for robust 3D human pose estimation using DDL [39]. In their study, they constructed two dictionaries simultaneously including visual observation dictionary and body configuration dictionary. Both of the two dictionaries share with a same sparse representation with respect to every visual observation and its corresponding 3D body pose.

Since outline features are usually corrupted, the optimization model for robust human pose estimation is as follows:

$$\begin{aligned} &\min_{A,B,E,R} \|E\|_1 + \lambda \|R\|_1 \\ &\text{s.t.} \quad X = AR + E \\ &\qquad Y = BR, \end{aligned} \tag{9}$$

where $X \in \mathfrak{R}^{m\times n}$ for observation data matrix, $A \in \mathfrak{R}^{m\times d}$ for observation dictionary, $Y \in \mathfrak{R}^{k\times n}$ for 3D pose data matrix, and $B \in \mathfrak{R}^{k\times d}$ for body configuration dictionary. $R \in \mathfrak{R}^{d\times n}$ for common sparse representation of X and Y, and E is the corruption item to be minimized.

To solve problem (9), Hao and Fei used an inexact Augmented Lagrange Multiplier (IALM) method to update the two dictionaries. More details related to the IALM method can be learned from [29].

The experimental results show that their approach performs well in recovering outlines from corrupted data compared with other methods.

4. DDL Algorithm in Medical Image Reconstruction

Recently, DDL has gained attention in medical image reconstruction, which can improve image qualities and accelerate reconstruction process.

4.1. Method and Theory. Let u^l be a low-quality image and $D_l = [d_1{}^l, d_2{}^l, \ldots, d_K{}^l]$, and let $D_l \in \mathfrak{R}^{n\times K}$ be a low dictionary constructed from u^l. Similarly, let u^h be the high-quality counterpart of u^l and $D_h = [d_1{}^h, d_2{}^h, \ldots, d_K{}^h]$; $D_h \in \mathfrak{R}^{n\times K}$ constructed from u^h. As a corresponding relation between u^l and u^h, they can be connected with a general following model:

$$u^l = Qu^h + \varepsilon^l, \tag{10}$$

where ε^l is the noise and Q is the transform operator. For a specific u^h, we can assume that each patch u_i^h in u^h can be expressed as the linear combination of the atoms in the following dictionary D_h:

$$u_i^h = D_h\boldsymbol{\alpha}_i + \eta, \tag{11}$$

where η is the error; $\|\eta\|_2^2 < \varepsilon$. $\boldsymbol{\alpha}_i$ is sparse coefficient, $\|\boldsymbol{\alpha}_i\|_0 \ll K$. Combining (11) and (10) gives

$$\left\|u_i^l - D_l\boldsymbol{\alpha}_i\right\|_2^2 < \delta = \left\|u_i^l - QD_h\boldsymbol{\alpha}_i\right\|_2^2 < \delta. \tag{12}$$

According to the above derivations which are referred to as the Sparse-Land Model, the low-quality patch u_i^l can be sparse coded by the same vector $\boldsymbol{\alpha}_i$ under dictionary $D_l = QD_h$. Thus, given the dictionaries D_l and D_h with accurate one-to-one mapping atoms, we can approximately recover u_i^h simply by multiplying D_h and the sparse representation obtained from D_l as follows:

$$u_i^h = D_h\boldsymbol{\alpha}_i + \varepsilon_i. \tag{13}$$

The general workflow for DDL method in medical image reconstruction is summarized in Figure 2. Given two sets of measured data (high-resolution sample images and low-resolution sample images), we can obtain two dictionaries D_l and D_h using appropriate training methods (DM, MOD, GPCA, or K-SVD). When a measured data is input, we can obtain the sparse representation with D_l and then update the x using D_h.

4.2. DDL in CT Reconstruction. Computed tomography (CT) reconstruction is a process obtaining the tomographic image of human body from X-ray projection data. The reconstruction methods can be divided into two types, analytic and iterative methods. In recent years, CS-based iterative method was applied in 3D X-ray image reconstruction. It performs more flexible and accurate than analytic method in most of cases. Some typical topics include interior CT problem, low-dose imaging, and incomplete data reconstruction [40–44].

Lu et al. made a progress in few-view image reconstruction of CT images (SART-TV-DL) [9, 10] using DDL. Since each pair of corresponding sample images is reconstructed from the same object just different in view numbers of projection, a high-quality image and its low-quality counterpart have the relationship described in (10).

In their work, a set of high-quality images which were reconstructed with SART algorithm from adequate projection were used to construct a high-quality dictionary D_h; however, according to the pixel-to-pixel mapping rule, a low-quality dictionary D_l can be also generated from a set of blurry images which were reconstructed from under-sampled projection data. To solve the dictionary training problem, they used DM mentioned in Section 1 because it could reserve most details of the sample images. Moreover, this method can generate dictionaries easiest and fastest.

However, in a CT image, pixel values alone cannot reflect the relationship of the adjacent two pixels. Therefore, in

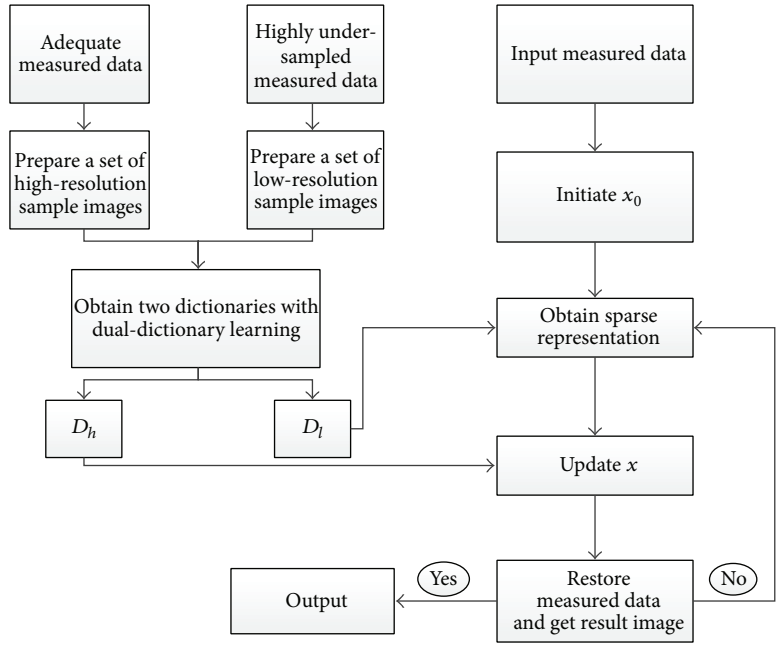


Figure 2: The general workflow for DDL method.

addition to DM, they used pixel values combined with its first-order gradient vector along x and y direction to provide more information of an image vector for each patch. That is, if an image patch is of size $\sqrt{n} \times \sqrt{n}$, the atom in the dictionary had $3n$ features because of the gradient. As the dictionaries were redundant or overcomplete, they reduced the redundancy of the dictionaries by means of setting a minimum Euclidean distances threshold.

The real data results demonstrate the potential of SART-TV-DL algorithm in CT image reconstruction with 30–50 views. It contributes to some preclinical and clinical applications such as C-arm, breast CT, and tomosynthesis.

Different from Lu's work, Cao and Xing applied DDL in CT limited angle reconstruction [45]. In his work, a two-dictionary learning (ART-TV-TDL) algorithm is proposed to remove the limited angle artifacts. The two dictionaries were, respectively, object dictionary D_o learned from a high-quality training image and artifact dictionary D_a from artifact image. A limited angle reconstruction $\widetilde{X}$, which could be divided into the object part X_o and the artifact part X_a, had the different sparse representation coefficients with D_o and D_a as follows:

$$\begin{aligned} \min_{\alpha_o} \left\| \widetilde{X} - D_o \alpha \right\|_2^2 \quad & \text{s.t. } \left\| \alpha_i \right\|_0 \le L_1, \\ \min_{\alpha_a} \left\| \widetilde{X} - D_a \beta \right\|_2^2 \quad & \text{s.t. } \left\| \beta_i \right\|_0 \le L_2. \end{aligned} \tag{14}$$

Here α and β are the sparse coefficient with L_1 and L_2 sparsity; the training method was K-SVD in this work. To get a better image with restrain artifacts, they combined these two representations for iterative reconstruction. Consider

$$X^{\text{next}} = \lambda_o D_o \alpha - \lambda_a D_a \beta + \lambda X^{\text{current}}, \tag{15}$$

where λ_o, λ_a, and λ are parameters to balance the effect. Their results show that the ART-TV-TDL method has smaller RMSE values in different limited angles (90 and 120) compared with ART-TV method.

4.3. DDL in 3D MRI Reconstruction. Song et al. proposed a novel method for multislice (3D) MRI reconstruction from undersampled k-space data using dual-dictionary learning (Dual-DL-MRI) [11].

For a high-resolution $M \times N \times H$ MRI images series S_h, one can represent them as one vector $s_{\text{high}} \in \Re^{MNH \times 1}$ of length MHN and get its undersampled k-space measurements y by Fourier transform $y = \widetilde{F}_u s_{\text{high}}$. $\widetilde{F}_u$ is a three-dimension undersampling Fourier matrix. Therefore, the corresponding series $s_{\text{low}} \in \Re^{MNH \times 1}$ can be reconstructed from undersampled k-space by inverse Fourier transform as follows:

$$s_{\text{low}} = \widetilde{F}^* y = \widetilde{F}^* \widetilde{F}_u s_{\text{high}} = Q s_{\text{high}}. \tag{16}$$

As we can see, (16) is one form of (10), which demonstrates the possibility of dual-dictionary in MRI reconstruction.

To construct dual-dictionary, they used K-SVD method to train the two dictionaries simultaneously to ensure the matching accuracy (one-to-one correspondence); D_l and D_h can be obtained by

$$\min_{D, \boldsymbol{\alpha}_i} \sum_i \left\| \boldsymbol{\alpha}_i \right\|_0 \quad \text{s.t. } \left\| \mathbf{s}_i - D \boldsymbol{\alpha}_i \right\|_2^2 \le \varepsilon_{\text{K-SVD}}, \quad \forall i, \tag{17}$$

where $S = \begin{bmatrix} s_{\text{low}} \\ s_{\text{high}} \end{bmatrix} = [s_1, s_2, \ldots, s_K]$ stands for two sample sets that are one-to-one matching; $D = \begin{bmatrix} D_l \\ D_h \end{bmatrix}$. It is worth noting that no more feature vectors are written in each dictionary atom except pixel values.

After updating the reconstruction result for each slice in the Fourier domain (restore the measured data), their work successfully reduce the PSNR of low-resolution MRI reconstruction images.

4.4. DDL in Multimodality Image Reconstruction. Multimodality biomedical imaging has found its increasing applications during the last decade and is becoming routine in clinical practice. Multimodality imaging is to integrate multiple imaging techniques into one instrument or fuse two or more imaging modalities such as CT, MRI, PET, and SPECT. This integration of structural, functional, and molecular information provides more accurate diagnoses. For example, MRI methods offer human soft tissue information with excellent clarity whereas CT depicts human hard tissue such as bone. Both of CT and MRI reveal important functional information. If these two modalities can be combined in one device, some small disease such as caducous blood clots could be exactly diagnosed. However, the imaging principles of MRI and CT are totally different, and how to build an accurate connection of these two modalities is an urgent problem.

In order to stylize the synergy between CT and MRI data sets from an object at the same time, Lu et al. try to investigate the possibility of CT-MRI unified imaging via dual-dictionary [46]. Figures 3(a) and 3(b) are, respectively, CT and MRI image; these two images are obtained from one layer of a patient's brain and are well registered. Figures 3(c) and 3(d) are the first-order gradient images of Figures 3(a) and 3(b) along x direction. Figure 3(e) is the subtraction of CT and MRI, and Figure 3(f) is the subtraction of their gradients. From Figures 3(c), 3(d), and 3(f), we can see that the interiors of CT and MRI are structurally correlated, especially the brain bone. Thus, it is possible to build a connection of CT and MRI using the structural information. With an MRI image as the a priori information, Lu tries to recover its corresponding CT image.

Since CT scan is totally different with MRI scan in physical principle, they use direct method to reserve as much information as possible to establish a knowledge-based connection between the two datasets. The two dictionaries are D_{MR} and D_{CT}; the former is derived from high-resolution MRI images, and the latter is from high-resolution CT images. The significant point of two dictionaries is that the patches in each dictionary are restricted one-to-one correspondence.

In reconstruction step, D_{MR} and D_{CT} are treated as D_l and D_h in (12), respectively. With dual-dictionary learning, a base CT image is first obtained just from a high-quality MRI image without corresponding CT data. Second, combined with base CT image and highly undersampled CT data, they reconstruct better resolution CT image using iterative method. The base CT image provides a better resolution and outline information, while highly undersampled CT image provides all the detailed information.

5. Discussion and Conclusion

In this paper, we discussed the recent advances of the DDL methods in medical imaging. Based on highly undersampled measured data, DDL algorithm has shown its great potential in reconstructing high-resolution images [47, 48].

Nowadays, MRI has become an indispensable medical modality of imaging diagnosis. However, during an MRI process, the scan time is usually up to fifteen minutes or even more. Patients might feel uncomfortable to keep motionless for a long time in the huge MRI gantry. Moreover, motion artifacts which reduce the images quality are always inevitable due to some organ movements such as heartbeat, pulse, and spasm. Researches demonstrated that the average displacement is over 0.35 mm within 100 seconds for one person lying on the cradle, while this number is up to 2.5 mm for a patient [42, 43]. Therefore, it has an important clinical significance to save the MRI scan time for better images quality and healthcare.

DDL method may be the future direction of fast MRI reconstruction. As mentioned in Section 4.4, the same slice of CT and MRI images from one object are structurally correlated. The advantage of CT is that the scanning time is short for some typical parts of body. Besides, the spatial resolution of CT is better than MRI. In the fast MRI, the measurement data is incomplete. Therefore, if the CT image data can be utilized as prior information in MRI reconstruction process, fewer measurement data (k-space) is required for high-resolution MRI image reconstruction. The essence of the reviewed DDL is establishing an appropriate relation between two spatial domains (e.g., different resolutions and different frequencies). One domain is for atom matching and the other domain is for image updating. Similarly, we may establish a quantitative relation between the two modalities using DDL. The relation can be a one-to-one mapping between the images boundaries which reflect the correlation between CT and MRI. In this way, DDL enables the fast MRI.

Overall, DDL method has shown its effective application in medical image reconstruction. With DDL method, we can reconstruct a high-resolution image with highly undersampling data. Inspired by its performances in one medical modality, DDL can be applied in structurally correlated image reconstruction problem, for example, multimodalities image reconstruction (CT-MRI).

However, the research work of DDL still remains in preliminary stage. For example, as discussed in the paper, reconstruction results may be relatively sensitive to the matching accuracy between the two dictionaries. Thus, how to establish closest connections between the images with different resolutions or even different modalities will be an important issue to be solved in the future. Also, the redundancy of dictionaries should be eliminated more reasonable to ensure better sparse representation.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was partly supported by the Grants from NNSFC 10905030 and 81427803, Beijing Natural Science Foundation

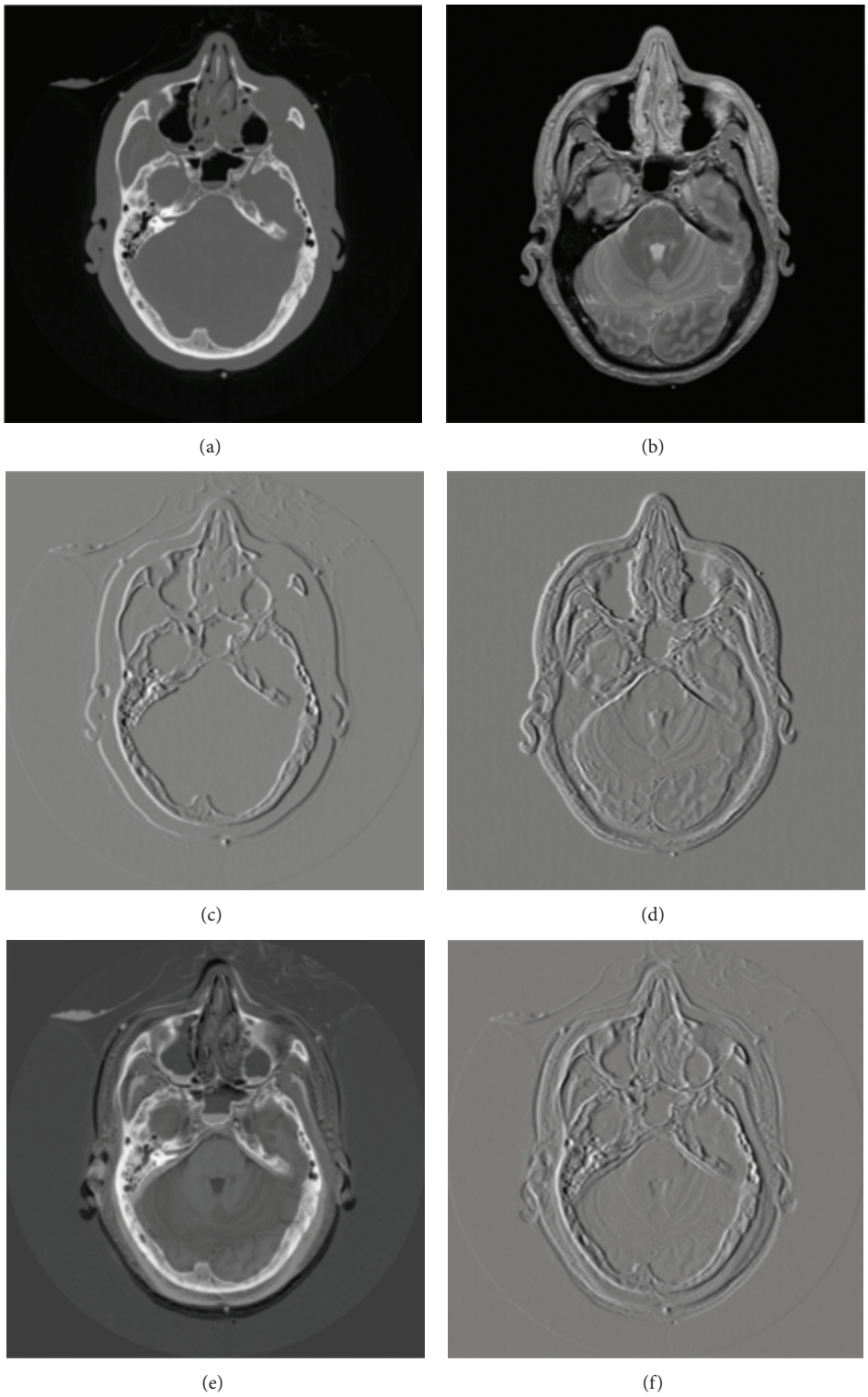


Figure 3: (a) CT image; (b) corresponding MRI image; (c) the first-order gradient of CT; (d) the first-order gradient of MRI; (e) CT and MRI images subtraction; (f) gradient images subtraction. (a) and (b) are obtained from *Visible Human Project* http://www.nlm.nih.gov/research/visible/visible_gallery.html.

(research on key techniques of medical cone-beam CT reconstruction from little data based on compressed sensing theory), and Beijing Excellent Talents Training Foundation (2013D0090040000004). Thanks are due to the supports by Visible Human Project (http://www.nlm.nih.gov/); CT and MRI figures are obtained from Visible Human Project.

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE Journal on Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.

[3] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[4] I. Daubechies, M. Defrise, and C. de Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[5] J. Qin and W. H. Guo, "An efficient compressive sensing MR image reconstruction scheme," in *Proceedings of the IEEE 10th International Symposium on Biomedical Imaging (ISBI '13)*, pp. 306–309, IEEE, San Francisco, Calif, USA, April 2013.

[6] F. A. Razzaq, S. Mohamed, A. Bhatti, and S. Nahavandi, "Locally sparsified compressive sensing for improved MR image quality," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC '13)*, pp. 2163–2167, 2013.

[7] F. A. Razzaq, S. Mohamed, A. Bhatti, and S. Nahavandi, "Non-uniform sparsity in rapid compressive sensing MRI," in *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC '12)*, pp. 2253–2258, October 2012.

[8] S. Ma, W. T. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, Alaska, USA, June 2008.

[9] Y. Lu, J. Zhao, and G. Wang, "Few-view image reconstruction with dual dictionaries," *Physics in Medicine and Biology*, vol. 57, no. 1, pp. 173–189, 2012.

[10] B. Zhao, H. Ding, Y. Lu, G. Wang, J. Zhao, and S. Molloi, "Dual-dictionary learning-based iterative image reconstruction for spectral computed tomography application," *Physics in Medicine and Biology*, vol. 57, no. 24, pp. 8217–8229, 2012.

[11] Y. Song, Z. Zhu, Y. Lu, Q. G. Liu, and J. Zhao, "Reconstruction of magnetic resonance imaging by three-dimensional dual-dictionary learning," *Magnetic Resonance in Medicine*, vol. 71, no. 3, pp. 1285–1298, 2014.

[12] H. Lee, D. S. Lee, H. Kang, B.-N. Kim, and M. K. Chung, "Sparse brain network recovery under compressed sensing," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1154–1165, 2011.

[13] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1028–1041, 2011.

[14] K. Engan, S. O. Aase, and J. H. Husoy, "Method of optimal directions for frame design," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '99)*, vol. I–VI, pp. 2443–2446, March 1999.

[15] R. Vidal, Y. Ma, and S. Sastry, "Generalized principal component analysis (GPCA)," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 12, pp. 1945–1959, 2005.

[16] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[17] Y. Li, "Dictionary learning based multitask image restoration," in *Proceedings of the 5th International Congress on Image and Signal Processing (CISP '12)*, pp. 364–368, October 2012.

[18] J. Zhang, D. B. Zhao, and W. Gao, "Group-based sparse representation for image restoration," *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3336–3351, 2014.

[19] J. Mairal, G. Sapiro, and M. Elad, "Learning multiscale sparse representations for image and video restoration," *Multiscale Modeling & Simulation*, vol. 7, no. 1, pp. 214–241, 2008.

[20] M. Tanaka, A. Sakurai, and M. Okutomi, "Across-resolution adaptive dictionary learning for single-image super-resolution," in *Digital Photography IX*, vol. 8660 of *Proceedings of SPIE*, February 2013.

[21] Y. Y. Fan, M. Tanaka, and M. Okutomi, "A classification-and-reconstruction approach for a single image super-resolution by a sparse representation," in *Digital Photography X*, vol. 9023 of *Proceedings of SPIE*, May 2014.

[22] L. Shang and Z. L. Sun, "Image super-resolution reconstruction based on two-stage dictionary learning," in *Intelligent Computing Methodologies*, vol. 8589 of *Lecture Notes in Computer Science*, pp. 277–284, Springer, Berlin, Germany, 2014.

[23] J. Zhang, C. Zhao, R. Q. Xiong, S. W. Ma, and D. B. Zhao, "Image super-resolution via dual-dictionary learning and sparse representation," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '12)*, pp. 1688–1691, IEEE, Seoul, Republic of Korea, May 2012.

[24] Q. G. Liu, D. Liang, Y. Song, J. H. Luo, Y. M. Zhu, and W. S. Li, "Augmented Lagrangian-based sparse representation method with dictionary updating for image deblurring," *SIAM Journal on Imaging Sciences*, vol. 6, no. 3, pp. 1689–1718, 2013.

[25] L. Ma, L. Moisan, J. Yu, and T. Zeng, "A dictionary learning approach for Poisson image Deblurring," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1277–1289, 2013.

[26] Z. Hu, J.-B. Huang, and M.-H. Yang, "Single image deblurring with adaptive dictionary learning," in *Proceedings of the 17th IEEE International Conference on Image Processing (ICIP '10)*, pp. 1169–1172, Hong Kong, China, September 2010.

[27] S. Beckouche, J. L. Starck, and J. Fadili, "Astronomical image denoising using dictionary learning," *Astronomy & Astrophysics*, vol. 556, article A132, 2013.

[28] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[29] R. Giryes and M. Elad, "Sparsity-based Poisson denoising with dictionary learning," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5057–5069, 2014.

[30] M. Razaviyayn, H.-W. Tseng, and Z.-Q. Luo, "Dictionary learning for sparse representation: complexity and algorithms," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '14)*, pp. 5247–5251, IEEE, Florence, Italy, May 2014.

[31] E. M. Eksioglu, "Online dictionary learning algorithm with periodic updates and its application to image denoising," *Expert Systems with Applications*, vol. 41, no. 8, pp. 3682–3690, 2014.

[32] X. Zhang, X. Feng, W. Wang, and G. Liu, "Image denoising via 2D dictionary learning and adaptive hard thresholding," *Pattern Recognition Letters*, vol. 34, no. 16, pp. 2110–2117, 2013.

[33] Q. Xu, H. Y. Yu, X. Q. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 31, no. 9, pp. 1682–1697, 2012.

[34] M. Türkan and C. Guillemot, "Dictionary learning for image prediction," *Journal of Visual Communication and Image Representation*, vol. 24, no. 3, pp. 426–437, 2013.

[35] G. H. Zhou, D. Z. Zhu, K. Wang, Q. Wu, X. C. Feng, and C. Wang, "Wavelet image inpainting based on dictionary learning with a beta process," in *Proceedings of the World Automation Congress (WAC '12)*, June 2012.

[36] C. Basso, M. Santoro, A. Verri, and S. Villa, "PADDLE: proximal algorithm for dual dictionaries LEarning," in *Artificial Neural Networks and Machine Learning—ICANN 2011*, vol. 6791 of *Lecture Notes in Computer Science Volume*, pp. 379–386, Springer, Berlin, Germany, 2011.

[37] H. Zhang, Y. Zhang, and T. S. Huang, "Efficient sparse representation based image super resolution via dual dictionary learning," in *Proceedings of the 12th IEEE International Conference on Multimedia and Expo (ICME '11)*, July 2011.

[38] X. Wang, Q. Ran, D. Chen, and F. Jiang, "Image restoration through dictionary learning and sparse representation," *Journal of Information and Computational Science*, vol. 10, no. 11, pp. 3497–3502, 2013.

[39] H. Ji and F. Su, "Robust 3D human pose estimation via dual dictionaries learning," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR '12)*, pp. 3370–3373, Tsukuba, Japan, November 2012.

[40] L. Li, K. J. Kang, Z. Q. Chen, L. Zhang, and Y. X. Xing, "A general region-of-interest image reconstruction approach with truncated Hilbert transform," *Journal of X-Ray Science and Technology*, vol. 17, no. 2, pp. 135–152, 2009.

[41] X. H. Duan, L. Zhang, Y. X. Xing, Z. Q. Chen, and J. P. Cheng, "Few-view projection reconstruction with an iterative reconstruction-reprojection algorithm and TV constraint," *IEEE Transactions on Nuclear Science*, vol. 56, no. 3, pp. 1377–1382, 2009.

[42] L. Li, Y. Xing, Z. Chen, L. Zhang, and K. Kang, "A curve-filtered FDK (C-FDK) reconstruction algorithm for circular cone-beam CT," *Journal of X-Ray Science and Technology*, vol. 19, no. 3, pp. 355–371, 2011.

[43] M. Chang, L. Li, Z. Q. Chen, Y. S. Xiao, L. Zhang, and G. Wang, "A few-view reweighted sparsity hunting (FRESH) method for CT image reconstruction," *Journal of X-Ray Science and Technology*, vol. 21, no. 2, pp. 161–176, 2013.

[44] Z. Chen, X. Jin, L. Li, and G. Wang, "A limited-angle CT reconstruction method based on anisotropic TV minimization," *Physics in Medicine and Biology*, vol. 58, no. 7, pp. 2119–2141, 2013.

[45] M. Cao and Y. X. Xing, "Limited angle reconstruction with two dictionaries," in *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC '13)*, pp. 1–4, Seoul, Republic of Korea, October 2013.

[46] Y. Lu, J. Zhao, T. G. Zhuang, and G. Wang, "Unified dual-modality image reconstruction with dual dictionaries," in *Developments in X-Ray Tomography VIII*, vol. 8506 of *Proceedings of SPIE*, October 2012.

[47] Z.-Q. Liu, L.-J. Bao, and Z. Chen, "Super-resolution reconstruction for magnetic resonance imaging based on adaptive dual dictionary," *Electro-Optic Technology Application*, vol. 28, no. 4, pp. 55–60, 2013.

[48] L. Feng, P. Wang, T.-F. Xu, M.-Z. Shi, and F. Zhao, "Dual dictionary sparse restoration of blurred images," *Optics and Precision Engineering*, vol. 19, no. 8, pp. 1982–1989, 2011.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 814104, 10 pages
http://dx.doi.org/10.1155/2015/814104

Research Article

Feature Selection Based on Machine Learning in MRIs for Hippocampal Segmentation

Sabina Tangaro,[1] Nicola Amoroso,[1,2] Massimo Brescia,[3] Stefano Cavuoti,[3] Andrea Chincarini,[4] Rosangela Errico,[2,5] Paolo Inglese,[1] Giuseppe Longo,[6] Rosalia Maglietta,[7] Andrea Tateo,[2] Giuseppe Riccio,[3] and Roberto Bellotti[1,2]

[1] *Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Via Orabona 4, 70125 Bari, Italy*
[2] *Dipartimento Interateneo di Fisica, Università degli Studi di Bari, Via Amendola 173, 70126 Bari, Italy*
[3] *Astronomical Observatory of Capodimonte, INAF, Salita Moiariello 16, 80131 Napoli, Italy*
[4] *Istituto Nazionale di Fisica Nucleare, Sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy*
[5] *Dipartimento Interateneo di Fisica, Università degli Studi di Genova, Via Dodecaneso 33, 16146 Genova, Italy*
[6] *Dipartimento di Fisica, Università degli Studi di Napoli Federico II, Complesso Universitario di Monte Sant'Angelo, Via Cintia 1, 80126 Napoli, Italy*
[7] *Istituto di Studi sui Sistemi Intelligenti per l'Automazione, CNR, Via Giovanni Amendola 122/D-I, 70126 Bari, Italy*

Correspondence should be addressed to Massimo Brescia; brescia@na.astro.it

Received 18 September 2014; Revised 9 January 2015; Accepted 15 January 2015

Academic Editor: Liang Li

Neurodegenerative diseases are frequently associated with structural changes in the brain. Magnetic resonance imaging (MRI) scans can show these variations and therefore can be used as a supportive feature for a number of neurodegenerative diseases. The hippocampus has been known to be a biomarker for Alzheimer disease and other neurological and psychiatric diseases. However, it requires accurate, robust, and reproducible delineation of hippocampal structures. Fully automatic methods are usually the voxel based approach; for each voxel a number of local features were calculated. In this paper, we compared four different techniques for feature selection from a set of 315 features extracted for each voxel: (i) filter method based on the Kolmogorov-Smirnov test; two wrapper methods, respectively, (ii) sequential forward selection and (iii) sequential backward elimination; and (iv) embedded method based on the Random Forest Classifier on a set of 10 T1-weighted brain MRIs and tested on an independent set of 25 subjects. The resulting segmentations were compared with manual reference labelling. By using only 23 feature for each voxel (sequential backward elimination) we obtained comparable state-of-the-art performances with respect to the standard tool FreeSurfer.

1. Introduction

The analysis of medical images such as magnetic resonance images (MRIs) is useful to investigate and identify the structural alterations in the brain, frequently associated with dementia or neurodegenerative diseases. In this context, the hippocampal segmentation is used to study and detect the correlation between the morphological anomalies of the hippocampus and the occurrence of the Alzheimer's disease. Hence its importance is strictly related to the early prediction of the dementia [1, 2]. Since the manual tracing is time-consuming and highly operator-dependent, it is important to make this process as much automatic as possible.

As discussed in [3], automatic image analysis and classification methods exist, which are able to recognize brain anomalies at the level of the single patient, which is more useful than at the level of groups or categories of individuals. Nonetheless they potentially require a large amount of parameters (vector of features) to properly manage all differences and specific features of the human brain among individuals, causing the parameter space to explode in terms of complexity, redundancy, and noise. To find a limited amount of features able to recognize patterns with a sufficient level of accuracy and without requiring a huge computational effort, would be indeed very helpful. This is especially true when the feature selection and classification are performed

by machine learning techniques, since the intrinsic self-organizing selection of important features and their cross-correlation remove any potential biased interpretability of the feature space.

Several approaches have been proposed to reach different levels of automation [4]. Among known methods, we quote just Morra et al. [5, 6], which suggest different automatic methods based on support vector machines (SVM) and hierarchical Adaboost, by considering about 18,000 voxel features, and FreeSurfer [7], a standard medical software tool for the analysis of cortical and subcortical anatomy, which performs a segmentation on cortical surface streams by constructing models of boundaries among white and gray matter.

Similarly, for an automatic hippocampal segmentation, we use a voxel-based approach by using 315 local features for each voxel included in a parahippocampal region larger than the hippocampus volume. Extracting 315 features for such a large number of voxels needs massive processing time and massive computational resources. For this reason, we consider crucial the issue of feature selection (FS) or reduction. The utility of feature selection is (a) to avoid overfitting, by minimizing the dimension of the parameter space and improving model performance, that is, prediction performance in the case of supervised classification and better cluster detection in the case of clustering, (b) to provide faster and more cost-effective models, (c) to gain a deeper insight into the underlying processes that generated the data, and (d) to optimize the processing time and massive computational resource.

There is a price to be paid for this advantage. To search for a subset of relevant features introduces in fact an additional layer of complexity in the modeling task: it needs to find the optimal model parameters for the optimal feature subset, as there is no guarantee that the optimal parameters for the full input feature set are equally optimal also for the best feature subset [8, 9].

By providing a small quantity of features, it may reduce the computational time as being proportional to the number of features. Furthermore, in some cases it allows to gain a better classification accuracy [10]. Also, the reduction of the feature's number is necessary when, to train the classifier, only a limited number of examples is available. In this regard, it is shown that, for the same error rate, a classifier requires a training whose duration grows exponentially with the number of variables [11–13].

Feature reduction, therefore, includes any algorithm that finds a subset of input feature set. A feature reduction capability is present also in more general methods based on transformations or combinations of the input feature set (feature extraction algorithms). An example being the well-known principal component analysis (PCA), which eliminates the redundancy of information by generating new features set by a combination of input features [14].

However, the best feature selection, by preserving the original semantics of features, permits also to maintain a coherent interpretability. The main goal of this study is to exemplify and demonstrate the benefits of applying FS algorithms in hippocampus segmentation field.

2. Materials

The database used to perform the described experiments is composed by thirty-five T1-weighted whole brain MR images and the corresponding manually segmented bilateral hippocampi (masks). All images were acquired on a 1.0 T scanner according to MP-RAGE sequence for magnetic resonance imaging of the brain [15–17].

The images are derived from the Open Access Series of Imaging Studies (OASIS). In particular we used 35 MP-RAGE MRI brain scans with a resolution of $1\,\text{mm}^3$ provided in occasion of the MICCAI SATA challenge workshop 2013 [18]. By using this homogeneous data sample it was possible to reduce the training image subsample without loss of generality and learning capabilities, giving the possibility to keep a sufficiently wide test set to perform a well-posed statistical analysis on the feature selection performances.

The image processing and classification were carried out blindly with respect to the subject status.

The first stage of our analysis chain requires an image preprocessing to standardize them both spatially and in gray intensity. This operation is obtained by registering the images on the Montreal Neurological Institute (MNI) standard template (ICBM152) using 12-parameter affine-registration and subsequent resampling on an isotropic grid with $1\,\text{mm}^3$ voxel size.

In order to reduce the computational time of the analysis, from the MRI, spatially standardized, two volumes containing the left and right hippocampus including the relevant parahippocampal regions are extracted using a new method FAPoD (fully automatic algorithm based on point distribution model) described in [19, 20].

We can then proceed with the feature extraction only in this identified region of interest: we approach a binary classification voxel-based problem, where the categories are *hippocampus* or *not-hippocampus*, that is, based on supervised pattern recognition systems. The features should contain information relevant to the classification task. Since manual segmentation of the hippocampus is based on local texture information, we adopted the related features. In the analysis presented here for each voxel a vector whose elements represent information about position, intensity, neighboring texture [21], and local filters was obtained.

Texture information was expressed using both Haar-like and Haralick features [6, 22].

The Haralick features were calculated from the normalized gray level cooccurrence matrices (GLCM) created on the $m \times m$ voxels projection subimages of the volume of interest; m defines the size of overlapping sliding-windows. For each voxel, values of m varying from 3 to 9 were used. Each element (k, p) of a cooccurrence matrix indicates the probability that two voxels, separated by a specified spatial angle and distance, have gray levels k and p, respectively.

A subset of Haralick features is sufficient to obtain a satisfactory discrimination. To establish which of the original 14 GLCM Haralick features gives the best recognition rate, several preliminary recognition experiments were carried out [23]. The resulting best configuration has been individuated in 4 features: energy, contrast, correlation, and inverse difference moment [20].

Table 1: The 315 features extracted from the 3D MRI images. Of each group of 66 Haralick features, 13 are the gradients along the 13 diagonals, 5 are the principal moments, and the rest are the three sets of 16 textural features, one set for each plane of the voxels. The gradients for each voxel are measured in all directions at one voxel distance and the relative 3D positions are included as features.

Number	Description
1	Position
1	Grey level
66	Haralick features for mask 3×3
66	Haralick features for mask 5×5
66	Haralick features for mask 7×7
66	Haralick features for mask 9×9
49	Haar-like 3D features

Finally, the gradients calculated in different directions and at different distances were included as additional features. The best analysis configuration, expressed by the highest values of statistical indicators (see Section 3), was obtained with 315 features, described in Table 1.

By summarizing, the knowledge base (KB) consisted of 35 regions of interest (ROI) extracted from as many images, each one composed of 7910 voxels, where each voxel is represented through a vector of 315 features. Therefore, the training set, including 10 randomly selected images, was formed by a total of 79100×315 entries. In quantitative terms, it can be considered a sufficiently wide dataset, qualitatively able to cover all feature types needed to perform a complete training, avoiding the useless redundancy of information not needed by machine learning methods [24] and leaving a sufficiently large amount of samples to be dedicated to the test sessions.

3. Methods

The FS techniques are usually counted in three categories, based on their internal combination between the selection and classification of the reduced parameter space. These categories are, respectively, named as wrapper, filter, and embedded methods [25].

Filter method is a technique based on the measurement of the importance of each single feature of the given parameter space [26]. The selected features are the most relevant to obtain a correct classification. This technique includes methods suitable for high-dimensional datasets, since they are computationally fast. Furthermore, they are independent from the classification algorithm and therefore their results can be used for all types of classifier. However, since each feature is considered separately from the others, their positive contribution based on the combined effect is neglected. The filter method used in our analysis is based on the Kolmogorov-Smirnov (K-S) test.

Wrapper methods basically integrate the two aspects of the workflow, that is, the model hypothesis and feature search [27]. This procedure involves the generation and evaluation of various subsets of features. Every generated feature subset is associated to a classification criterion (hence the name *wrapper*). Since the number of all possible feature subsets grows exponentially with the size of the dataset, some search heuristics can be adopted to reduce drastically the number of operations. They can be grouped into *deterministic* and *randomized* search methods. The advantage of these methods is the intrinsic best interaction among selected features and their classifiers, but with the downside of having a high computational cost and the risk of overfitting. The wrapper methods used in our analysis are, respectively, sequential forward selection (SFS) and sequential backward elimination (SBE).

Finally, in *embedded* methods the optimal feature subset search is directly nested into the classifier algorithm [28]. Such techniques can be interpreted in terms of a search within a combined parameter space, by mixing features and hypotheses. Analogously to wrapper methods, they include the interaction with classification algorithm but in a faster way. The embedded method used in our analysis is based on the Random Forest Classifier.

To recap, in our FS analysis we used the following:

(i) univariate filter method: Kolmogorov-Smirnov,

(ii) deterministic wrapper methods: sequential forward selection (SFS) and sequential backward elimination (SBE),

(iii) embedded method: Random Forest.

In addition, we have also used the PCA [29], being one of the most widely adopted feature reduction techniques, for comparison.

To estimate the goodness of the selected feature group we used the Näive Bayes Classifier [30], based on the simplified hypothesis that all attributes describing a specific instance on data are conditionally independent among themselves.

The FS analysis was performed in the 5-fold cross validation on 10 of 35 images in the database. The goodness of the selected group was tested on the remaining 25 images. As already discussed in Section 2, the selected training and test rates were considered sufficiently wide to ensure a well-posed training and the postprocessing statistical evaluation.

The k-fold cross validation is a technique able to avoid overfitting on data and is able to improve the generalization performance of the machine learning model. In this way, validation can be implicitly performed during training, by enabling at setup the standard leave-one-out k-fold cross validation mechanism [31]. The automatized process of the cross validation consists in performing k different training runs with the following procedure: (i) splitting of the training set into k random subsets, each one composed by the same percentage of the data set (depending on the k choice); (ii) at each run the remaining part of the data set is used for training and the excluded percentage for validation. While avoiding overfitting, the k-fold cross validation leads to an increase of the execution time estimable around $k-1$ times the total number of runs.

Furthermore, the combination of the Bayes rule with the above simplified assumption has a positive impact on the model complexity and its computational time. In particular, the latter property pushed us to choose this model as embedded classifier for the feature selection problem.

The agreement between an automated segmentation estimate and a manual segmentation can be assessed using overlap measures. A number of measures are available: (a) Dice index [20, 32]; (b) efficiency; (c) purity of a class; (d) completeness of a class; (e) contamination of a class.

At the base of the statistical indicators adopted, there is the commonly known *confusion matrix*, which can be used to easily visualize the classification performance [33]: each column of the matrix represents the instances in a predicted class, while each row represents the instances in the real class. One benefit of a confusion matrix is the simple way in which it allows seeing whether the system is mixing different classes or not.

We remark here that we were mostly interested in the feature analysis related to the classification of the *hippocampus* class voxels. Therefore, we considered as particularly relevant the Dice index, usually referred to as the *true positive* class (N_{AA} in our confusion matrix), which in our case corresponds properly to *hippocampus* class. Since, by definition, the Dice index does not take the true negative rate into account, the rate of *not-hippocampus* voxels is not involved within this indicator. A statistical evaluation of this latter class, corresponding to the background voxels, has been primarily included for completeness and for coherency with the full confusion matrix representation. The highest relevance given to the *hippocampus* class analysis represents also a common evaluation criterion in such context [6].

In terms of binary classification, we were more interested to perform a feature selection analysis, rather than to improve the classification performances. Therefore, we imposed a standard classification threshold to 0.5 at the beginning of the experiments and maintained unchanged all over the entire described process, by considering it as sufficient for our specific purposes.

More specifically, for a generic two-class confusion matrix, we consider

		OUTPUT	
	–	Class A	Class B
TARGET	Class A	N_{AA}	N_{AB}
	Class B	N_{BA}	N_{BB}

(1)

we then use its entries to define the following statistical quantities.

(i) Total Efficiency. te is defined as the ratio between the number of correctly classified objects and the total number of objects in the data set. In our confusion matrix example it would be

$$\mathrm{te} = \frac{N_{AA} + N_{BB}}{N_{AA} + N_{AB} + N_{BA} + N_{BB}}. \tag{2}$$

(ii) Purity of a Class. pc$N$ is defined as the ratio between the number of correctly classified objects of a class and the number of objects classified in that class. In our confusion matrix example it would be

$$\mathrm{pc}A = \frac{N_{AA}}{N_{AA} + N_{BA}},$$
$$\mathrm{pc}B = \frac{N_{BB}}{N_{AB} + N_{BB}}. \tag{3}$$

(iii) Completeness of a Class. cmp$N$ is defined as the ratio between the number of correctly classified objects in that class and the total number of objects of that class in the data set. In our confusion matrix example it would be

$$\mathrm{cmp}A = \frac{N_{AA}}{N_{AA} + N_{AB}},$$
$$\mathrm{cmp}B = \frac{N_{BB}}{N_{BA} + N_{BB}}. \tag{4}$$

(iv) Contamination of a Class. cnt$N$ is the dual of the purity; namely, it is the ratio between the misclassified objects in a class and the number of objects classified in that class; in our confusion matrix it example will be

$$\mathrm{cnt}A = 1 - \mathrm{pc}A = \frac{N_{BA}}{N_{AA} + N_{BA}},$$
$$\mathrm{cnt}B = 1 - \mathrm{pc}B = \frac{N_{AB}}{N_{AB} + N_{BB}}. \tag{5}$$

(v) Dice Index. Dice, known also with the name of F_1 *score*, is a frequent measure used in binary classification, which could be considered as a weighted average of the purity and completeness, reaching its best value at 1 and the worst at 0. By referring to our notation, we have the Dice defined as

$$\mathrm{Dice} = 2 \cdot \frac{\mathrm{pc}A * \mathrm{cmp}A}{\mathrm{pc}A + \mathrm{cmp}A} = 2 \cdot \frac{N_{AA}}{2N_{AA} + N_{BA} + N_{AB}}. \tag{6}$$

4. Results

By using Näive Bayes Classifier on all 315 input features, the goodness is estimated in 5-fold cross validation on 10 images. The results in terms of the statistics, derived from the confusion matrix, are shown in Table 2 and the Dice index is 0.60 ± 0.04.

The PCA applied to 315 input features returns the principal components (PCs) ordered by the amount of information they convey. The percentage of information contained in the first 98 PCs and in the first 197 PCs are, respectively, 90% and 99%.

Since our goal was to reduce the feature retaining the goodness in the classification, we considered the first 197 PCs containing 99.0% of the information. The results obtained are shown in Table 3 and the Dice index is 0.62 ± 0.07. As mentioned above, we used the Näive Bayes Classifier in 5-fold cross validation.

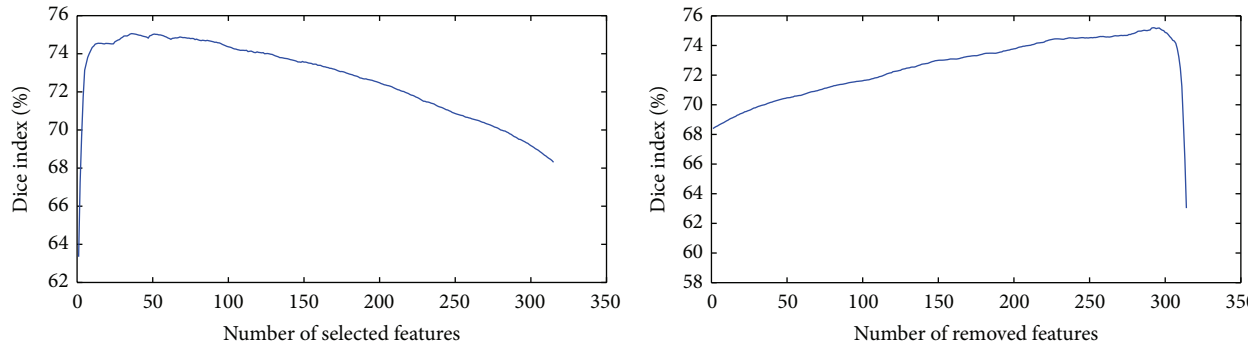


Figure 1: Best dice index of all the possible combinations of the relevant step in (a) sequential forward selection and in (b) sequential backward elimination methods.

Table 2: Classification result on all 315 input features using Näive Bayes Classifier in 5-fold cross validation based on confusion matrix.

315 input features	Completeness of a class	Purity of a class	Contamination of a class
Hippocampus	79%	62%	38%
Not-hippocampus	63%	80%	20%
Efficiency		**70%**	

Table 3: Classification result on the first 197 PCs using Näive Bayes Classifier using in 5-fold cross validation based on confusion matrix.

197 PCs	Completeness of a class	Purity of a class	Contamination of a class
Hippocampus	60%	68%	32%
Not-hippocampus	78%	72%	28%
Efficiency		**71%**	

Table 4: Classification result on all 315 PCs using Näive Bayes Classifier using in 5-fold cross validation based on confusion matrix.

315 PCs	Completeness of a class	Purity of a class	Contamination of a class
Hippocampus	86%	51%	49%
Not-hippocampus	36%	78%	22%
Efficiency		**58%**	

Compared to the use of all 315 original features, the values obtained with 197 PCs are on average 6% points lower in terms of Dice index. Therefore, to avoid loss of information, we considered all 315 PCs. The results are reported in Table 4 and the Dice index is 0.63 ± 0.03.

Even using all the PCs, the result was 5% points lower in terms of Dice Iindex. This result confirms what was already found by Golland et al. in [34]; that is, the selection of large-variance features performed by the PCA is not specifically suited for segmentation problems.

4.1. Kolmogorov-Smirnov Analysis. The K-S test provides an estimate of how much two distributions are related to each other. The K-S test allowed us to select only the features which have a correlation between the two *hippocampus* and *not-hippocampus* classes less than 5%, resulting in a total of 57 features.

Table 5: Classification result on 57 features selected through Kolmogorov-Smirnov test using Näive Bayes Classifier using in 5-fold cross validation based on confusion matrix.

57 features Kolmogorov-Smirnov test	Completeness of a class	Purity of a class	Contamination of a class
Hippocampus	84%	57%	43%
Not-hippocampus	52%	81%	19%
Efficiency		**66%**	

As mentioned above, we used the Näive Bayes Classifier in 5-fold cross validation. The results obtained are shown in Table 5 and the Dice index is 0.67 ± 0.04.

The K-S test results are comparable with the original parameters space based on 315 features.

4.2. Sequential Forward Selection and Backward Elimination. The two FS methods belonging to the wrapper category experimented in our case were SFS and SBE. In Figure 1(a) on the ordinate axis, the top value of Dice index achieved between all possible combinations related to the reference step depicted on the horizontal axis is shown. At each step, the feature achieving the best performance is chosen, when used in combination with the selected features in the previous step. The step number coincides with the number of selected features (SFS).

In Figure 1(b) on the ordinate axis, the top value of Dice index achieved between all possible combinations related to the reference step depicted on the horizontal axis is shown. At each step the feature without which the best performances are obtained is removed. The step number coincides with the number of eliminated features (SBE).

We observe that the SFS method reaches its highest Dice index, 0.75, at step 36. So it means that the best performance,

Table 6: Classification result on 36 features selected through forward selection method using Näive Bayes Classifier in 5-fold cross validation based on confusion matrix.

36 features Forward selection	Completeness of a class	Purity of a class	Contamination of a class
Hippocampus	82%	70%	30%
Not-hippocampus	73%	84%	16%
Efficiency		**77%**	

using the Näive Bayes Classifier, is obtained with only 36 selected features, listed in Table 7. In Figure 2 and Figure 3 a more detailed description of some features is shown.

The SBE method obtains its highest Dice index 0.75 at the step 292. Therefore, the best performance, evaluated with the Näive Bayes Classifier, is obtained by using the remaining 23 features (i.e., 315 − 292), listed in Table 9.

Tables 6 (with related Dice index is 0.75±0.03) and 8 (with related Dice index is 0.75 ± 0.02), respectively, show the relative performance of the peak value in Figure 1.

4.3. Random Forest Analysis. The Random Forest classification methodology allowed us to estimate the feature importance [35]. To select the best subset we have performed a study of classification with cross validation procedure based on the Näive Bayes Classifier, varying the threshold on the feature importance index. The optimal threshold was related to the maximum Dice Index value and achieved with 222 features. Also in this case we used the Näive Bayes Classifier in 5-fold cross validation to evaluate the features selected by the Random Forest. The result obtained is shown in Table 10 and the Dice index is 0.69 ± 0.04.

4.4. Random Selection Test. Furthermore, we performed an additional group of tests to evaluate whether randomly selected samples of 36 features among the original 315 might lead to Dice indexes greater than or comparable with the Dice value obtained with SFS (0.75). To do so, we estimate the empirical probability density function of Dice under the null hypothesis that any set S^* of 36 features provides a Dice value greater than or equal to the true Dice in predicting whether a voxel belongs to hippocampus or not. To test this hypothesis, 2000 sets S^* were generated, each composed of 36 features randomly drawn from the ones available and the corresponding Dice values were evaluated. The obtained results are shown in Figure 4.

5. Discussion and Conclusion

The main goal of this work was to verify the possibility to reduce the number of required voxel features without losing or better by enhancing the classification performances. Moreover the reduction of the number of voxel features could also improve the computational efficiency of the classification.

As clearly resulting from a recent review, [3], by now the feature selection has to be considered as an essential step within the field of neuroimaging approached by the machine

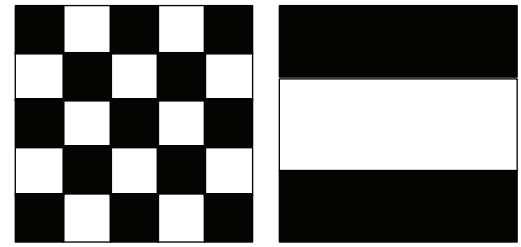


Figure 2: Haar-like template types 1 (a) and 2 (b) used in the experiments.

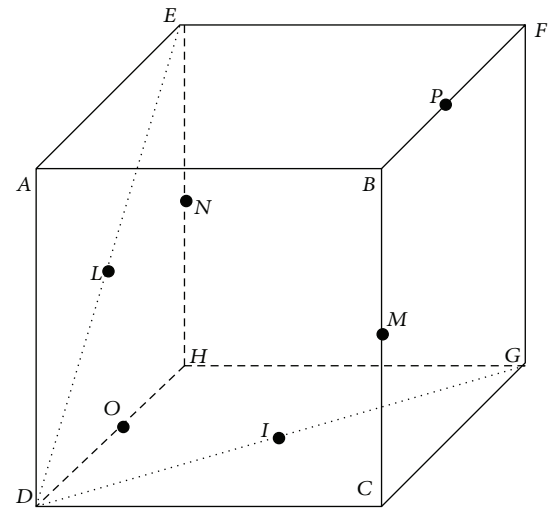


Figure 3: Representation of a generic cubic mask used for calculating the gradient features. The labeled points are either the vertexes of the cube or the median points of the segments.

learning paradigm. Its importance is also invariant to the specific technique used to extract and codify the features from MRIs regions of interest, whether it is based on standard n-dimensional feature vectors or on pairwise dissimilarity representation. In the present work we investigated the application of several feature selection methods.

The results obtained using different approaches are summarized in Table 11 and in Figure 5. We observe that by using these two selected subsets it is possible to obtain higher performances than using the entire input dataset.

By considering the percentage of random Dice values bigger than the best one with respect to the total number of random extractions, such value is zero. But, as it can be seen in Figure 4, in many cases it appears to obtain better performances by randomly extracting the feature sample rather than considering the complete set of 315 features.

Among the FS approaches presented in this work, the SFS and SBE show better performances.

We would underline that the results shown in Figure 5 have to be mainly interpreted as a comparison among the different methods of feature selection. What has to be stressed is that the performances are influenced by the feature information content and the image enhancement techniques

Table 7: Details of the 36 features resulting by the forward selection method using Näive Bayes Classifier.

36 features	Haralick features			Haar-like features	Statistical features	
Forward selection	Orientation	Coordinate	Mask size	Type	Mask size	Entry
Contrast*	135	Y	3			
Gradient*					5	$\overline{EC}$
Correlation	135	X	3			
Position*						Coordinates
Normalized gray level*						Value
Correlation*	45	X	5			
Gradient*					5	$\overline{DF}$
Correlation*	90	Y	9			
Correlation	45	Y	7			
Skewness*					7	
Homogeneity*	90	X	9			
Correlation	0	Y	5			
Correlation	90	Z	5			
Correlation*	45	X	3			
Correlation	135	Z	9			
Correlation	90	Y	5			
Correlation	135	Z	5			
Correlation	0	Z	7			
Correlation	90	Z	7			
Correlation	90	Z	9			
Correlation	0	Y	3			
Correlation	135	X	3			
Correlation	0	Z	9			
Template*				1		
Skewness*					5	
Correlation	90	Z	3			
Correlation	45	X	5			
Gradient					3	$\overline{MN}$
Template				2		
Correlation*	45	X	9			
Correlation	45	Y	5			
Correlation	90	Y	7			
Correlation	45	Z	5			
Gradient					9	$\overline{DF}$
Homogeneity	0	Z	9			
Correlation	0	Y	9			

The asterisk indicates the entries also present in the list of 23 SBE features. For Haralick features, the orientation in degrees, reference coordinate, and the size of the cubic mask used are reported. In case of Haar-like features, the entry value indicates the template type used (see Figure 2). For statistical/positional kind the size of the cubic mask used or the self-explained value is lister, depending on the specific feature type. In particular for gradients, the column named Entry indicates the segment of the reference diagonal as shown in Figure 3. All the features are listed in top-down order of their inclusion during the SFS procedure execution.

Table 8: Classification result on 23 features selected through backward elimination method using Näive Bayes Classifier in 5-fold cross validation based on confusion matrix.

23 features Backward elimination	Completeness of a class	Purity of a class	Contamination of a class
Hippocampus	83%	70%	30%
Not-hippocampus	73%	85%	15%
Efficiency		77%	

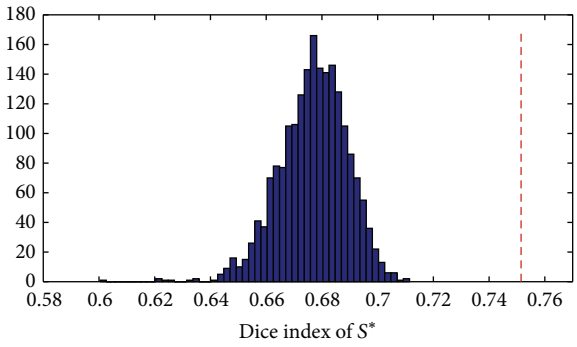


Figure 4: Distribution of 2000 random Dice values compared with true Dice (shown with the dashed red line) concerning 36 features obtained by the sequential forward selection.

employed. A quite simple method, such as the Näive Bayes Classifier, is able to reach state-of-the-art performances when preceded by a selection analysis on the feature space. A more detailed study of the classification methods and of the postprocessing technique which can be used to improve performances are presented in other studies [36, 37].

To test the goodness of the best feature selection methods presented in this paper we used the two selected sets formed, respectively, by 36 and 23 features on a blind test database composed of 25 MRIs (i.e., not used in training phase), in the algorithm cited in [36] (see Tables 7 and 9, resp.).

By analyzing the two subsets of selected features, it was obtained that 13 of the 23 extracted by the SBE method are also present in the sample of 36 features obtained by the SFS technique. Most of them are Haralick and Statistical features, except for the positional and Haar-like features, confirming the importance given by Haralick and Statistical types and a very low contribution of Haar-like type.

We remark that, by minimizing the presence of Haralick features, in particular the correlations, it allows improving the processing time and a better handling of the information content. In fact, among the three categories of features considered here, the Haralick type was the most *time-consuming* from the computational point of view.

The comparison of our FS methods with the widely used PCA demonstrates the very low performance of the PCA technique (as shown in Figure 5). This result is in agreement with the well-known downside of the method in presence of a very high nonlinearity of the feature correlations. It is also an indirect confirmation about the intrinsic difficulty to separate

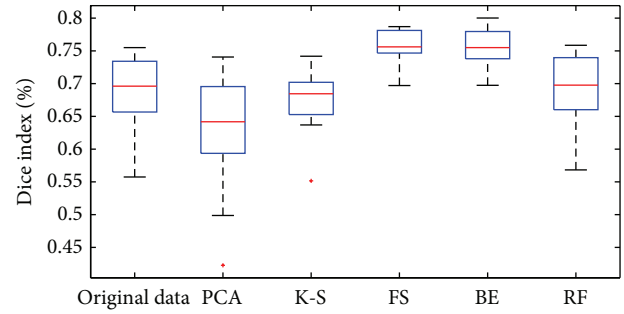


Figure 5: Dice index comparison for the following methods: original dataset (315 for each voxel); PCA (197 selected features); K-S test (57 selected features); SFS (36 selected features); SBE (23 selected features); Random Forest (222 selected features). Boxes have lines at the lower quartile, median, and upper quartile values, with whiskers extending to 1.5 times the interquartile range. Outliers are indicated by a plus sign.

the *hippocampus* versus *not-hippocampus* classes from MRI images.

We conclude that the SFS and SBE techniques are two promising methods allowing to reduce the input space size, with a very low loss of information, and permitting classification performances comparable or even better than the case with a larger amount of features.

In fact, in terms of feature space dimension comparison, Morra et al. [6] performs a voxel-based segmentation using about 18,000 features with the weighted voting method AdaBoost [38] tested on a different image data set. In addition, FreeSurfer [7], which is a not voxel-based method considered as standard benchmark for MRI segmentation experiments, reaches a Dice value of 0.76 ± 0.05.

In this work, we observed that the selected features from both SFS and SBE methods are related to the high frequency component of the image. So this result would suggest which kind of features are best suitable for high frequency classification problems such as edge recognition. In fact, these correlation features, being based on intensity differences, are able to capture local information based on discontinuity rather than similarity.

Besides, this result is a further suggestion for a future investigation which is to put in practice a preprocessing procedure to enhance the contours of the structures contained in the image and to assess the usefulness of these procedures in the diagnosis support systems.

Conflict of Interests

All the authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank the anonymous referee for extremely valuable comments and suggestions. Nicola Amoroso, Rosangela Errico, Paolo Inglese, and Andrea Tateo

Table 9: Details of the 23 features resulting by the backward elimination method using Näive Bayes Classifier.

23 features Backward elimination	Haralick features			Haar-like features	Statistical features	
	Orientation	Coordinate	Mask size	Type	Mask size	Entry
Position*						Coordinates
Normalized gray level*						Value
Correlation	0	Y	7			
Correlation*	45	X	3			
Correlation*	45	X	5			
Correlation	45	X	7			
Correlation*	45	X	9			
Correlation	45	Y	9			
Correlation	45	Y	5			
Correlation*	90	Y	9			
Homogeneity	135	Z	3			
Gradient*					5	$\overline{DF}$
Gradient					7	$\overline{DF}$
Contrast*	135	Y	3			
Gradient					9	$\overline{OP}$
Homogeneity*	90	X	9			
Gradient					3	$\overline{BH}$
Skewness*					7	
Gradient*					5	$\overline{EC}$
Gradient					3	$\overline{IL}$
Template*				1		
Skewness*					5	
Gradient					5	$\overline{MN}$

The asterisk indicates the entries also present in the list of 36 SFS features. For Haralick features, the orientation in degrees, reference coordinate, and the size of the cubic mask used are reported. In case of Haar-like features, the entry value indicates the template type used (see Figure 2). For statistical/positional kind, the size of the cubic mask used and/or the self-explained value is listed, depending on the specific feature type. In particular for gradients, the column named Entry indicates the segment of the reference diagonal as shown in Figure 3.

Table 10: Classification result on 222 features selected through Random Forest method using Näive Bayes Classifier in 5-fold cross validation based on confusion matrix.

222 features Random Forest	Completeness of a class	Purity of a class	Contamination of a class
Hippocampus	80%	62%	38%
Not-hippocampus	62%	80%	20%
Efficiency		**70%**	

Table 11: For each implemented method, size of selected group, mean Dice index (evaluated using Näive Bayes Classifier), and related σ are shown.

Method	Size of selected group	Dice index
Original dataset	315	0.69 ± 0.04
PCA selection	197	0.62 ± 0.07
K-S selection	57	0.67 ± 0.04
Forward selection	36	0.75 ± 0.02
Backward elimination	23	0.75 ± 0.02
Random Forest	222	0.69 ± 0.04

acknowledge funding by the Italian MIUR Grant PON PRISMA no. PON04a2_A.

References

[1] L. G. Apostolova, R. A. Dutton, I. D. Dinov et al., "Conversion of mild cognitive impairment to Alzheimer disease predicted by hippocampal atrophy maps," *Archives of Neurology*, vol. 63, no. 5, pp. 693–699, 2006.

[2] S. Tangaro, P. Inglese, R. Maglietta, and A. Tateo, "MIND-BA: fully automated method for computer-aided diagnosis of dementia based on structural MRI data," in *Proceedings of the MICCAI 2014*, Lecture Notes in Computer Science, Boston, Mass, USA, September 2014.

[3] E. Veronese, U. Castellani, D. Peruzzo, M. Bellani, and P. Brambilla, "Machine learning approaches: from theory to application in schizophrenia," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 867924, 12 pages, 2013.

[4] M. Chupin, A. Hammers, R. S. N. Liu et al., "Automatic segmentation of the hippocampus and the amygdala driven by hybrid constraints: method and validation," *NeuroImage*, vol. 46, no. 3, pp. 749–761, 2009.

[5] J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, A. W. Toga, and P. M. Thompson, "Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation," *IEEE Transactions on Medical Imaging*, vol. 29, no. 1, pp. 30–43, 2010.

[6] J. H. Morra, Z. Tu, L. G. Apostolova et al., "Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls," *NeuroImage*, vol. 43, no. 1, pp. 59–68, 2008.

[7] B. Fischl, "FreeSurfer," *NeuroImage*, vol. 62, no. 2, pp. 774–781, 2012.

[8] S. Tangaro, R. Bellotti, F. De Carlo et al., "MAGIC-5: an Italian mammographic database of digitised images for research," *La Radiologia Medica*, vol. 113, no. 4, pp. 477–485, 2008.

[9] S. Ciatto, D. Cascio, F. Fauci et al., "Computer-assisted diagnosis (CAD) in mammography: comparison of diagnostic accuracy of a new algorithm (Cyclopus, Medicad) with two commercial systems," *La Radiologia Medica*, vol. 114, no. 4, pp. 626–635, 2009.

[10] A. K. Jain and B. Chandrasekaran, "39 Dimensionality and sample size considerations in pattern recognition practice," *Handbook of Statistics*, vol. 2, pp. 835–855, 1982.

[11] L. Devroye, L. Györfi, and G. Lugosi, *A Probabilistic Theory of Pattern Recognition*, vol. 31 of *Stochastic Modelling and Applied Probability*, Springer, Berlin, Germany, 1996.

[12] G. L. Masala, B. Golosio, P. Oliva et al., "Classifiers trained on dissimilarity representation of medical pattern: a comparative study," *Nuovo Cimento della Societa Italiana di Fisica C*, vol. 28, no. 6, pp. 905–912, 2005.

[13] G. L. Masala, S. Tangaro, B. Golosio et al., "Comparative study of feature classification methods for mass lesion recognition in digitized mammograms," *Nuovo Cimento della Societa Italiana di Fisica C*, vol. 30, no. 3, pp. 305–316, 2007.

[14] A. Jain and D. Zongker, "Feature selection: evaluation, application, and small sample performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 153–158, 1997.

[15] J. P. Mugler III and J. R. Brookeman, "Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE)," *Magnetic Resonance in Medicine*, vol. 15, no. 1, pp. 152–157, 1990.

[16] Challenge Details, https://masi.vuse.vanderbilt.edu/workshop-2013/index.php/Segmentation.

[17] https://masi.vuse.vanderbilt.edu/workshop2013/index.php/NMMProtocol.

[18] https://masi.vuse.vanderbilt.edu/workshop2013.

[19] N. Amoroso, R. Bellotti, S. Bruno et al., "Automated shape analysis landmarks detection for medical image processing," in *Computational Modelling of Objects Represented in Images III: Fundamentals, Methods and Applications*, pp. 139–142, 2012.

[20] S. Tangaro, N. Amoroso, M. Boccardi et al., "Automated voxel-by-voxel tissue classification for hippocampal segmentation: methods and validation," *Physica Medica*, vol. 30, no. 8, pp. 878–887, 2014.

[21] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," in *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, pp. I-900–I-903, 2002.

[22] R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, 1979.

[23] R. Bellotti, F. de Carlo, S. Tangaro et al., "A completely automated CAD system for mass detection in a large mammographic database," *Medical Physics*, vol. 33, no. 8, pp. 3066–3075, 2006.

[24] G. Longo, M. Brescia, G. S. Djorgovski, S. Cavuoti, and C. Donalek, "Data driven discovery in astrophysics," in *Proceedings of the ESAESFRIN Conference on Big Data from Space*, Frascati, Italy, November 2014.

[25] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[26] N. Sanchez-Marono, A. Alonso-Betanzos, and M. Tombilla-Sanromán, "Filter methods for feature selection—a comparative study," in *Intelligent Data Engineering and Automated Learning—IDEAL 2007*, vol. 4881, pp. 178–187, Springer, 2007.

[27] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.

[28] V. N. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, New York, NY, USA, 1998.

[29] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[30] I. Rish, "An empirical study of the naive Bayes classifier," in *Proceedings of the Workshop on Empirical Methods in Artificial Intelligence (IJCAI '01)*, pp. 41–46, August 2001.

[31] S. Geisser, "The predictive sample reuse method with applications," *Journal of the American Statistical Association*, vol. 70, no. 350, pp. 320–328, 1975.

[32] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[33] F. Provost, T. Fawcett, and R. Kohavi, "The case against accuracy estimation for comparing induction algorithms," in *Proceedings of the 15th International Conference on Machine Learning*, pp. 445–553, Morgan Kaufmann, 1998.

[34] P. Golland, W. E. L. Grimson, M. E. Shenton, and R. Kikinis, "Detection and analysis of statistical differences in anatomical shape," *Medical Image Analysis*, vol. 9, no. 1, pp. 69–86, 2005.

[35] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[36] S. Tangaro, N. Amoroso, S. Bruno et al., "Active learning machines for automatic segmentation of hippocampus in MRI," in *Proceedings of the Industrial Conference on Data Mining Workshops (ICDM '13)*, pp. 181–191, 2013.

[37] R. Maglietta, N. Amoroso, S. Bruno et al., "Random forest classification for hippocampal segmentation in 3D MR images," in *Proceedings of the 12th International Conference on Machine Learning and Applications (ICMLA '13)*, vol. 1, pp. 264–267, IEEE, December 2013.

[38] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, part 2, pp. 119–139, 1997.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 486532, 16 pages
http://dx.doi.org/10.1155/2015/486532

Research Article

Hybrid Pixel-Based Method for Cardiac Ultrasound Fusion Based on Integration of PCA and DWT

Samaneh Mazaheri,[1] Puteri Suhaiza Sulaiman,[1] Rahmita Wirza,[1] Mohd Zamrin Dimon,[2] Fatimah Khalid,[1] and Rohollah Moosavi Tayebi[1]

[1]*Faculty of Computer Science and Information Technology, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia*
[2]*Cardiothoracic Unit, Surgical Cluster, Faculty of Medicine, 40450 Shah Alam, Selangor, Malaysia*

Correspondence should be addressed to Samaneh Mazaheri; mazaheri.samaneh@yahoo.com

Received 31 October 2014; Revised 26 January 2015; Accepted 8 February 2015

Academic Editor: Tianye Niu

Medical image fusion is the procedure of combining several images from one or multiple imaging modalities. In spite of numerous attempts in direction of automation ventricle segmentation and tracking in echocardiography, due to low quality images with missing anatomical details or speckle noises and restricted field of view, this problem is a challenging task. This paper presents a fusion method which particularly intends to increase the segment-ability of echocardiography features such as endocardial and improving the image contrast. In addition, it tries to expand the field of view, decreasing impact of noise and artifacts and enhancing the signal to noise ratio of the echo images. The proposed algorithm weights the image information regarding an integration feature between all the overlapping images, by using a combination of principal component analysis and discrete wavelet transform. For evaluation, a comparison has been done between results of some well-known techniques and the proposed method. Also, different metrics are implemented to evaluate the performance of proposed algorithm. It has been concluded that the presented pixel-based method based on the integration of PCA and DWT has the best result for the segment-ability of cardiac ultrasound images and better performance in all metrics.

1. Introduction

Image fusion is one of the major research fields in image processing. It is a procedure of combining the related info from two or several images, into a single image, without introducing any distortion which will be more informational and containing more details, and while it is more suitable for visual perception, it is complete in comparison with any of the inputs. Image fusion methods can enhance the quality and increase the application of input data [1]. For the purpose of the majority medical applications, medical image fusion aims to decrease ambiguity and minimize redundancy in result image when increasing the related info details [2, 3]. With the lately quick advances in the field of sensing technologies, multisensory systems have become a reality in medical imaging. As a result of using these technologies, we will have a huge increase in quantity of acquired data. Image fusion produces a useful way of decreasing that volume of information whilst, at the same time, extract all the valuable information from inputs. The goal of image fusion, except decreasing the quantity of info, is to build single enhanced image more appropriate for the purpose of human visual perception and for next image processing tasks like segmentation or feature detection in medical imaging.

Fusion also improves the capability for other applications by complementary information. In other words, the main condition for successful fusion is that "all" visible information in the input images should also appear visible in the fused image [4]. There are some needs of image fusion which are [5]

(i) extracting the whole desired information from the input images to get the relevant information,

(ii) not introducing distortions or inconsistencies that will amuse human observers,

(iii) robust and reliable to imperfections,

(iv) improving reliability.

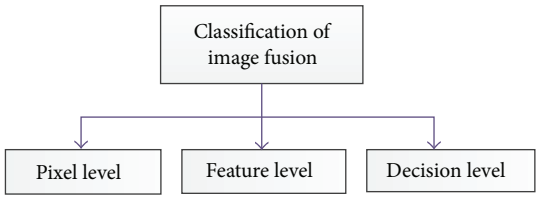


Figure 1: Classification of image fusion Algorithms.

This paper is organized as follows: Section 2 provides a background and reviews image fusion concepts and some related works; Section 3 outlines proposed method for echocardiography fusion and explain the proposed algorithm; experimental results and evaluation of the proposed algorithm and discussion on results are presented in Section 4; the paper finishes with concluding remarks in Section 5.

2. Background and Related Works

2.1. Classification of Image Fusion Algorithms. Fusion is a procedure of incorporating the applicable info from a set of images of the same view right into a one image and the resulting image could be more useful and beneficial than any of the inputs [6]. The actual fusion process can take place at different levels of information representation.

Different categories of image fusion methods are usually classified in various levels: pixel, feature and decision level (see Figure 1).

(1) Low-level or pixel-level: the pixel-level method works either in the transform domain or in spatial domain. They can directly work on the pixels of the images. Image fusion at pixel-level tries to incorporate low-level data, often in physical measurements like intensity [4] (see Figure 3).

(2) Middle-level or feature-level: the feature level methods perform on characteristics taken out from the input images. They originally divide the image into contiguous areas and combine the areas together using their properties. The characteristics employed may be computed individually from every image or they may be acquired through the simultaneous procedure from all the images.

(3) High-level or decision level: decision level fusion uses the results of initial object detection and classification as inputs to perform fusion as data integration [7].

Pixel-level image fusion represents the visual information of the same scene from numbers of images which can be obtained using different sensors [4]. A simple diagram of a system using pixel-level fusion is demonstrated in Figure 2.

Details of some pixel-level techniques are described here [6].

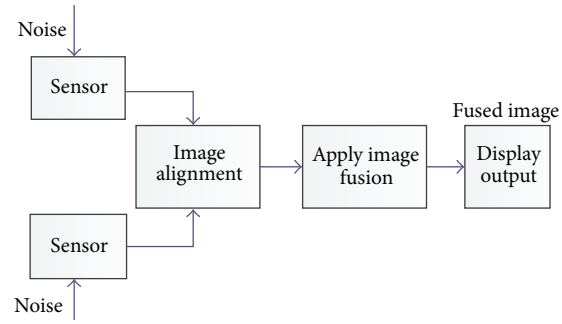


Figure 2: Pixel-level image fusion.

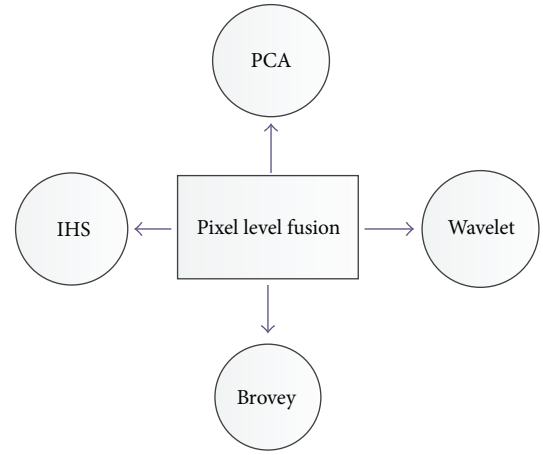


Figure 3: Some of well-known pixel-level image fusion techniques.

(a) Simple maximum: in this technique, the combined image is acquired through choosing the maximum intensity of related pixels from two inputs:

$$F(i,j) = \sum_{i=0}^{m} \sum_{j=0}^{n} \max A(i,j)\, B(i,j). \quad (1)$$

$A(i,j)$, $B(i,j)$ are inputs and $F(i,j)$ is the resultant one.

(b) Simple mnimum: in simple minimum technique, the combined image is acquired through choosing the minimum intensity of related pixels from two inputs:

$$F(i,j) = \sum_{i=0}^{m} \sum_{j=0}^{n} \min A(i,j)\, B(i,j). \quad (2)$$

$A(i,j)$, $B(i,j)$ are inputs and $F(i,j)$ is the resultant one.

(c) Simple average: here, the combined image is acquired through calculating the mean intensity of related pixels from two inputs:

$$F(i,j) = \frac{A(i,j) + B(i,j)}{2} \quad (3)$$

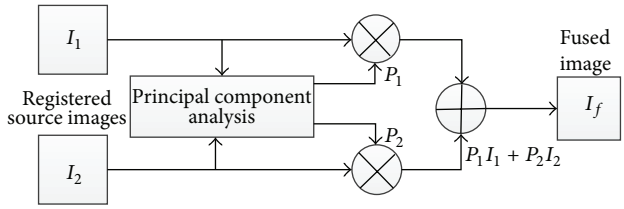


Figure 4: Flow diagram of information in first phase of proposed image fusion method.

$A(i, j)$, $B(i, j)$ are inputs and $F(i, j)$ is the resultant one.

(d) Weighted average: in weighted average technique, the combined image is acquired through calculating the weighted mean intensity of related pixels from two inputs:

$$F(i, j) = \sum_{i=0}^{m} \sum_{j=0}^{n} WA(i, j) + (1 - W) B(i, j) \quad (4)$$

$A(i, j)$, $B(i, j)$ are inputs and $F(i, j)$ is the resultant image and W is the weight component.

(e) Principal component analysis (PCA): PCA technique is a subspace one, which reduces the multidimensional data sets into lower dimensions for analysis. This method determines the weights for each source image using the eigenvector corresponding to the largest eigenvalue of the covariance matrix of each source image (see Figure 4).

(f) Discrete wavelet transform (DWT): these transforms have image decomposition tool that provide a variety of channels representing the image feature by different frequency subbands at multiscale. 2D discrete wavelet transformation (DWT) converts the image from the spatial domain to frequency domain (see Figure 6).

(g) Brovey transform (BT): Brovey transform also known as color normalized fusion is based on the chromaticity transform and the concept of intensity modulation. It is a simple method to merge data from different sensors, which can preserve the relative spectral contributions of each pixel but replace its overall brightness with the high spatial resolution image.

(h) Intensity hue saturation (IHS) method: the IHS method is a standard process in image fusion, with important restrictions in which only three bands are included. Basically, it was based on the RGB true color space. It provides the advantage that the separate channels outline certain color properties, namely, intensity (I), hue (H), and saturation (S). This special color space is usually selected because the visual cognitive system of human intends to treat these three elements as roughly orthogonal perceptual axes.

Finally, based on domains, image fusion methods are usually categorized into two groups.

(i) Spatial domain methods: with spatial domain approaches, we specifically deal with pixels of image. The pixel values are usually altered to acquire desired outcome.

(ii) Transform domain methods: in transform domain methods, image is transferred into frequency domain first.

Fusion methods like averaging, Brovey method, principal component analysis (PCA), and IHS based techniques are categorized under spatial domain methods. The discrete wavelet transform is categorized under frequency domain approaches.

2.2. Echocardiography Image Fusion. Image fusion has turned into a popular term employed in medical diagnostics and treatment [8]. Fused images may be produced from a number of images from the similar imaging modality [9] or by merging details from multiple modalities [10] like MRI, CT, PET, and SPECT. For precise diagnoses, radiologists should incorporate information from multiple image formats. Fused, anatomically consistent images are specifically helpful in diagnosing and treating [11]. Some of applications of image fusion techniques in medical images would be fusing CT and MRI images, computer assisted surgery, and spatial registration of 3D surface [1].

Echocardiography imaging is a widespread technique to acquire cardiac images; however, it suffers from artifacts, high noises, and a limited field of view [12]. The created images are degraded by an implicit distortion named "speckle," which comes from the destructive and constructive coherent summation of ultrasound echoes. The distortion caused by speckle can be defined as multiplicative noises that lead to granular look, degrade the contrast of the images, and decrease the ability to find information within the images. A technique to deal with these restrictions is using several images, selecting the best part from every image to provide a better quality result.

Some important requirements could be considered for fusion process: (a) the fusion process must keep the whole related info contained in the source images, (b) the fusion process should not produce any distortion or inconsistencies that would amuse the human observer or subsequent processing phases which can lead to a wrong diagnosis, and (c) irrelevant characteristics and noises need to be covered up to a maximum degree [13]. The problem that medical image fusion attempts to resolve is to fuse the information content from multiple images (or various imaging sensors) from the same view to achieve a fused image that contains the best possible details. Therefore, the fused image would produce improved superiority image compared to any of the original input images.

In this study, a pixel level based medical image fusion is introduced to display a fusion procedure creating a single fused image containing additional reliable information than individual input image. In spite of numerous attempts in

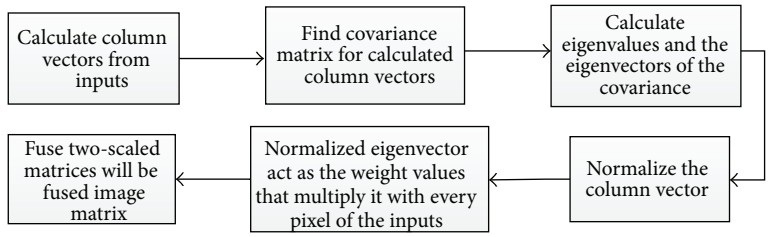


Figure 5: Stepwise procedure for first phase of proposed method.

direction of automation ventricle segmentation and tracking [14], the problem stay complicated study as a result of low quality characteristics of obtained images with missing anatomical details, or contains speckle noises, or restricted field of view. The simplest medical image fusion is to consider the mean of the gray level input images, pixel by pixel. However, using this technique on echocardiography images would produce several undesired effects and reduced feature contrast.

To overcome this problem, a new pixel based fusion method which integrates PCA and DWT is proposed in this paper. This fusion is achieved by weighted average of images to be fused. The weights for every input image are acquired from the eigenvector related to the largest eigenvalue of the covariance matrices of every input. Then, wavelet transform is used which presents directional details in decomposition levels and includes distinctive information at various resolutions. Performance metrics are used to evaluate the DWT, PCA, and the proposed algorithm performance as well [13].

3. The Proposed Method

A fusion approach which tries to provide a better quality image for improving the segment-ability of echocardiograms is proposed here. It is also capable of improving the contrast, decreasing the impact of echo artifacts, expanding the field of view, and improving the signal to noise ratio, as well. The proposed technique weights details of images, based on combinations of features of PCA incorporating with DWT between all the overlapping images to develop the fusion. The objective of this research is to implement spatial fusion on multiple echocardiography sequence and perform clinical examination in cooperation with cardiologists from UiTM, Sungai Bulah hospital, Malaysia.

The speckle should be filtered out without affecting salient characteristics in cardiac medical ultrasounds. Two main denoising approaches are employed: (a) fusion, which combines echocardiograms of the same area to provide a single decreased speckle image; (b) filtering.

This paper tries to optimize and continue the development of image fusion utilizing integration PCA with DWT, two well-known image fusion methods. PCA method can preserve a better resolution, but it distorts spectral features with various degrees as well. However, DWT will produce directional details in decomposition levels and includes special information at various resolutions. Additionally, DWT and PCA can keep more spectral details and spatial features, respectively. Also, PCA method can get salient characteristics to decrease redundancy. We present a spatial frequency (SF) motivated approach which integrates their advantages and improves image quality to avoid distortions and artifacts.

Lately, fusion echocardiography sequences are introduced through obtaining several echocardiography images with small changes in probe location and combining all of them together following alignment.

The particular idea of fusion echocardiography sequences allows us to enhance image quality as well as improving anatomical details which leads to better feature detection.

3.1. Principal Component Analysis. The PCA consists of a mathematic process which converts multiple correlated variables to several uncorrelated variables named principal components. Also as a linear transformation, it calculates a compressed and optimal explanation of dataset [13]. The 1st principal component accounts for the maximum amount of the variance in dataset as is achievable and every following component considers as the maximum amounts of the remaining variance as is available. First principal component is considered to be along the direction with the highest variance. The second principal component is restricted to put in the subspace perpendicular of the first one. Within this subspace, this component points the direction of maximum variance. The third principal component is considered in the maximum variance direction in the subspace perpendicular to the first two and so forth. PCA helps to minimize unnecessary information and emphasize the components with highest effect to improve signal to noise ratio metric. It is known as hoteling transform or the Karhunen-Loève transform as well [4].

3.2. Wavelet Transform. The PCA image fusion technique operates under spatial domain. However, the spatial domain fusion may produce spectral degradation. This is particularly crucial if the images are supposed to fuse such as echocardiography images were not acquired at the same time. Therefore, compared with the ideal result of the fusion, if this method is applied alone, it will produce poor and undesirable result for echocardiography images. Wavelet transforms are a decomposition tool for multiresolution images that offer a range of channels which represent the image characteristic through various frequency subbands. Since it has been found that wavelet fusion techniques outperform the standard fusion techniques in spectral quality, especially in minimizing

distortion, we use both PCA and wavelet transform to get their advantages and minimize their disadvantages.

The proposed method that combines the PCA method with wavelet transform provides outstanding outcomes compared to standard PCA or wavelet transform alone. Wavelet transforms are generally classified into three categories; continuous, discrete, and multiresolution based [4]. In discrete wavelet transform, while decomposition is applied, the estimation and information element can be different. 2D DWT transforms echocardiography image from spatial domain to frequency one. Input is separated to horizontal and vertical outlines and shows the DWT first order; then image is divided to 4 areas which are LL1, LH1, HL1, and HH1. When decomposition is performed, the L-L band provides the typical image info while other bands include directional information caused by spatial orientation. Higher complete wavelet coefficients value within the high bands related to important characteristics include lines or edges. As a result, in wavelet transform, the size of image is halved in spatial direction at every decomposition level of procedure, therefore ending to a multiresolution signal representation. The main phase for combination is the creation of combination pyramid.

3.3. The Proposed Algorithm. The proposed algorithm consists of two phases. Suppose the inputs are sorted in 2 column vectors. In first phase, steps to project this data into 2D subspaces are as follows.

(1) At first, the data should be organized in column vector. Suppose R is the result column vector of dimension $n \times 2$.

(2) After dividing the data into columns, the dimension of empirical average with each column is 1×2.

(3) Subtract average from every R column. The dimension of result matrix X is $n \times 2$.

(4) Calculate covariance matrix C from matrix X.

(5) Take into account first column of eigenvector V that is related to larger eigenvalue $V(1)$ to calculate normalized component P_1 and P_2 as

$$P_1 = \frac{V(1)}{\sum V}, \qquad P_2 = \frac{V(2)}{\sum V}. \tag{5}$$

In first phase, image fusion will be done by PCA. The information flowchart of the PCA algorithm is shown in Figure 5. The inputs $I_1(x, y)$ and $I_2(x, y)$ are ordered in 2 column vectors and their empirical averages are subtracted. The result vector has a dimension of $n \times 2$, in which n is length of every image vector. Calculate eigenvector and eigenvalues for the result vector and eigenvectors related to bigger eigenvalue acquired. The normalized components P_1 and P_2 (i.e., $P_1 + P_2 = 1$) using (5) are calculated from the acquired eigenvector. The result image will be

$$I_f(x, y) = P_1 I_1(x, y) + P_2 I_2(x, y). \tag{6}$$

There is not a restricted list of basic vectors such as wavelet, FFT, and DCT in PCA and it has its basic vectors rely on

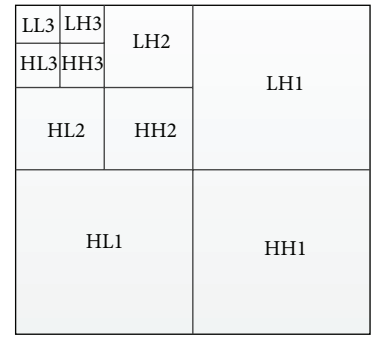


Figure 6: Wavelet decomposition in 3 levels.

dataset. Suppose X is a random D dimensional vector which has zero empirical average. Orthonormal projection matrix V will be like $Y = V^T X$ with the subsequent restrictions. The Y covariance, that is, $\operatorname{cov}(Y)$, is a diagonal one and also it is inverse of V which is equal to its transpose ($V^{-1} = V^T$). By using matrix algebra, we have

$$\begin{aligned} \operatorname{cov}(Y) &= E\left\{YY^T\right\} = E\left\{\left(V^T X\right)\left(V^T X\right)^T\right\} \\ &= E\left\{\left(V^T X\right)\left(X^T V\right)\right\} = V^T E\left\{XX^T\right\} V \\ &= V^T \operatorname{cov}(X)\, V. \end{aligned} \tag{7}$$

Multiplying each side of formula (7) by V, one gets

$$V \operatorname{cov}(Y) = VV^T \operatorname{cov}(X)\, V = \operatorname{cov}(X)\, V. \tag{8}$$

By writing V as $V = [V_1, V_2, \ldots, V_d]$ and $\operatorname{cov}(Y)$ as

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{d-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_d \end{bmatrix} \tag{9}$$

substituting (7) into (8) gives

$$\begin{aligned} &[\lambda_1 V_1, \lambda_2 V_2, \ldots, \lambda_d V_d] \\ &\quad = \left[\operatorname{cov}(X)\, V_1, \operatorname{cov}(X)\, V_2, \ldots, \operatorname{cov}(X)\, V_d\right]. \end{aligned} \tag{10}$$

It could be rewritten as

$$\lambda_i V_i = \operatorname{cov}(X)\, V_i, \tag{11}$$

where $i = 1, 2, \ldots, d$ and V_i as an eigenvector of $\operatorname{cov}(X)$.

In the second phase, in order to provide superior outcomes, we will apply wavelet fusion on images which were obtained from the first phase (see Figures 8 and 9). In this phase, while decomposition is applied, the estimation and

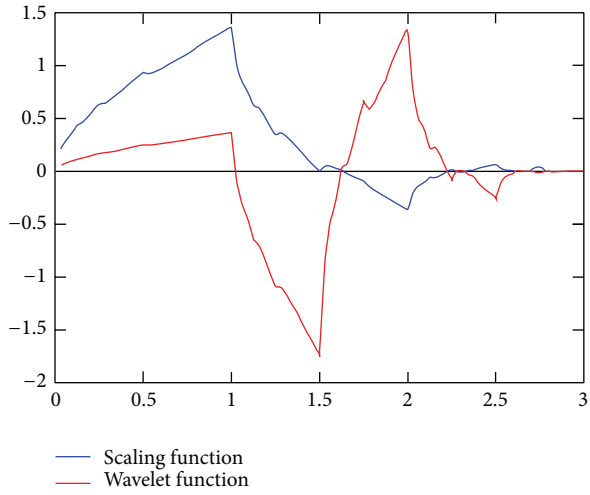


FIGURE 7: Daubechies 4 tap wavelet.

information element can be different. 2D DWT changes the echocardiography images domain from spatial to frequency. First, the image is separated to horizontal and vertical outlines, and then it shows DWT first order, through the image which is divided to 4 areas that are LL1, LH1, HL1, and HH1 (see Figure 6). A certain signal of specific energy is projected on a continuous family of frequency bands (or identical subspaces of the function space $L^2(\mathbb{R})$). For example, the signal might be displayed on each frequency band of the form $[f, 2f]$ for all constructive frequencies $f > 0$. Next, the main signal can be rebuilt by an appropriate integration overall the producing frequency components [15–17]. The frequency bands or subbands are scaled versions of a subspace at scale 1.

This subspace consequently is in the majority circumstances created by the shifts of one generating function $\psi L^2(R)$, the mother wavelet. For instance, for scale one frequency band $[1, 2]$ the function is given by

$$\Psi(t) = 2\sin c(2t) - \sin c(t) = \sin(2t) - \frac{\sin(t)}{t}. \quad (12)$$

Analyzing a signal utilizing all wavelet coefficients is extremely hard and computationally impossible, so maybe it is adequate to select a discrete subset of the higher half plane in order to construct a signal from the equivalent wavelet coefficients. For some actual variables $a > 1$, $b > 0$. The corresponding discrete subset of the half plane consists of all the points $(a^m, na^m b)$ with integers $m, n \in \mathbb{Z}^2$. The equivalent baby wavelet is given by

$$\psi_{m,n}(t) = a^{-m/2}\psi\left(a^{-m}t - nb\right). \quad (13)$$

An acceptable condition for construction of any signal x of specific energy through equation

$$x(t) = \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}}(x, \psi_{m,n})\cdot\psi_{m,n}(t). \quad (14)$$

The function $\{\psi_{m,n} : m, n \in \mathbb{Z}^2\}$ gives tight shape of $L^2(\mathbb{R})$.

In a discrete wavelet transform, there are simply a limited number of wavelet coefficients for every bounded rectangular area in the top half plane. However, every coefficient needs the analysis of an integral. To prevent this mathematical complexity, an auxiliary function is needed, the father wavelet $\phi \in L^2(\mathbb{R})$. Additionally, one has to limit a to be an integer. A typical selection is $a = 2$ and $b = 1$. The most popular two of father and mother wavelets is the Daubechies 4 tap wavelet (see Figure 7) [18–22]. We can construct the subspaces with the mother and father wavelets [15]:

$$\begin{aligned} V_m &= \operatorname{span}\left(\phi_{m,n} : n \in \mathbb{Z}\right) \\ \text{where } \phi_{m,n}(t) &= 2^{-m/2}\phi\left(2^{-m}t - n\right), \\ W_m &= \operatorname{span}\left(\psi_{m,n} : n \in \mathbb{Z}\right) \\ \text{where } \psi_{m,n}(t) &= 2^{-m/2}\psi\left(2^{-m}t - n\right). \end{aligned} \quad (15)$$

Let $S(n_1, n_2)$ be input data and its size is $n_1 \times n_2$; then the function of wavelet and scaling are

$$\begin{aligned} &w_\emptyset(j_0, k_1, k_2) \\ &\quad = \frac{1}{\sqrt{N_1 N_2}}\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} s(n_1, n_2)\,\emptyset_{j_0,k_1,k_2}(n_1, n_2), \\ &w_\epsilon(j_0, k_1, k_2) \\ &\quad = \frac{1}{\sqrt{N_1 N_2}}\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} s(n_1, n_2)\,\emptyset_{j_0,k_1,k_2}(n_1, n_2). \end{aligned} \quad (16)$$

The algorithm of second phase for the proposed technique is as follows.

(1) Implement discrete wavelet transform on image which is obtained from first phase to build lower wavelet decomposition.

(2) Combine every decomposition level.

(3) Hold inverse discrete wavelet transform on combined decomposed level, meaning rebuild the image, when the image rebuilt is F, the fused image.

4. Result and Discussion

In this section, it has been shown that the proposed approach can provide more satisfactory outcomes, compared to other techniques and algorithms in two aspects of visual effect and quantitative analysis. The evaluations are organized to employ well known image fusion approaches with four image quality measurements. Experiments have been done on the dataset which contains 10 subject cases. For every subject case, two echocardiographic phases have been considered which are different frames of heart cycle that have been obtained utilizing a Philips iE33 cardiac ultrasound scanning device (Philips Medical Devices, Sungai Bulah hospital, UiTM, Malaysia), under supervision of cardiac surgeon (see Figure 10). The presented integration of PCA and DWT fusion method is employed on the two mentioned phases of dataset. Also,

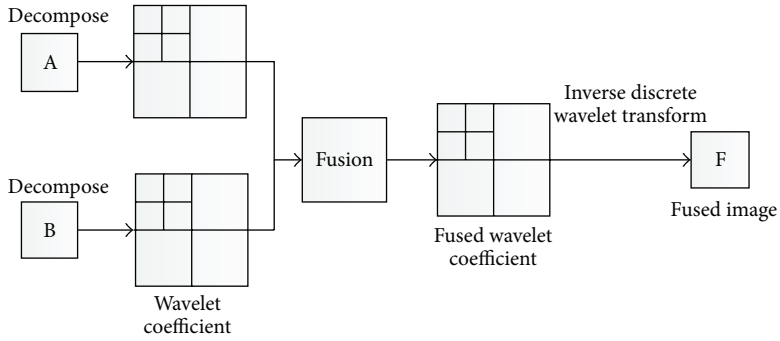


Figure 8: Second phase process for the proposed method.

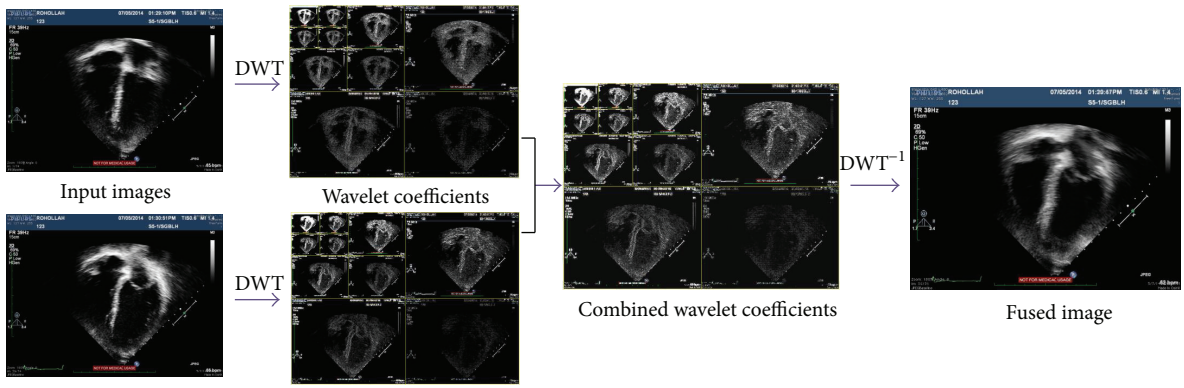
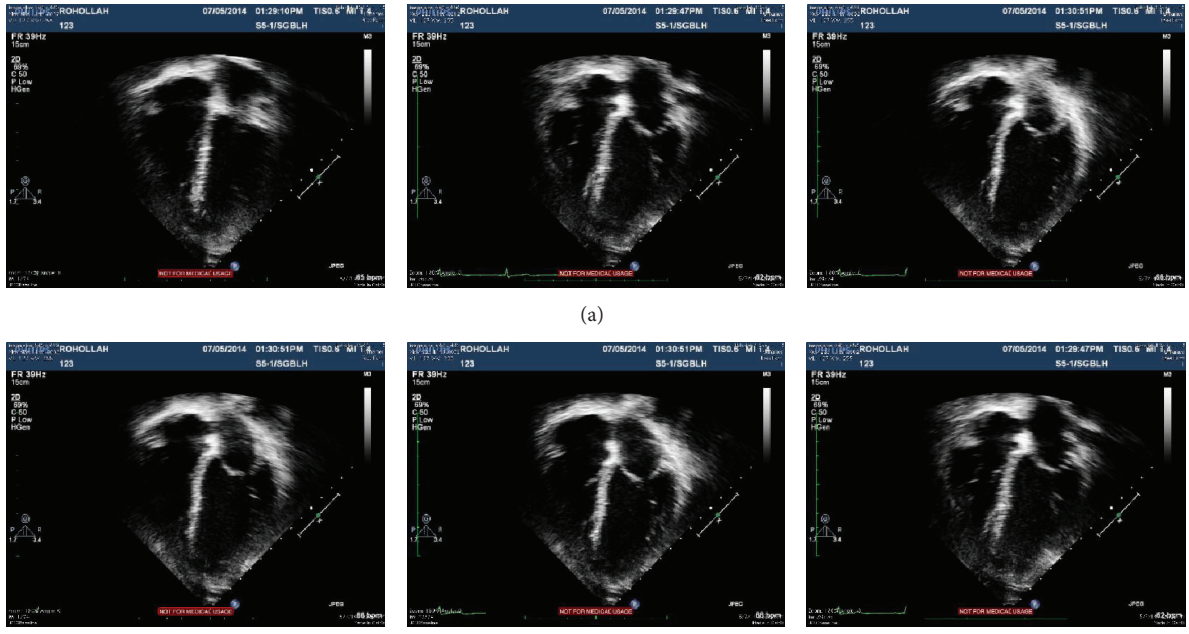


Figure 9: Applying second phase of proposed algorithm on two input images.



(a)

(b)

Figure 10: Original cardiac ultrasound images; (a) first row, dataset number 1, and (b) second row, dataset number 2.

PCA and DWT fusion techniques are employed on the data separately, and then the performance of every method is computed by applying four quantitative measurements. RMSE (root mean square error) is the appropriate measurement which shows how images are close together. Considering the possible artificial distortion over the combination procedure can also raise the entropy or spatial frequency measures of the combined image; the image quality index (IQI) is a reasonably trustworthy measurement for images with no reference image because IQI gets a value (in the range of 0 to 1) about how related is the result to both inputs.

4.1. Numerical Results. The typical requirement of a fusing procedure is to keep the whole useful info from the input images, whilst it should not produce any artifacts in resulting image. Efficiency measurements are employed to calculate the advantages of fusion and also are helpful to make comparison between results achieved with various techniques. Four measures are employed to examine image quality, including CC (correlation coefficient), IQI (image quality index), RMSE (root mean square error), and OCE (over-all cross entropy) [23–25]. The overall cross entropy is used to find the difference among two input images and the result one. Small measure is related to fine acquired result. The result of proposed integration of PCA and DWT method is compared with two results which are obtained by applying DWT and PCA alone. Table 2 demonstrates the experimental result. From the measurements, it can be seen that the CC and IQI are the biggest with the presented technique. The RMSE and OCE of the proposed technique are least in 2 sets. It is indicated that the presented technique has the best results for fused images.

4.2. Experimental Results. For evaluation, two types of comparison have been considered; first, result of proposed method is compared with result of other techniques, separately. Then, segmentation approach is used to evaluate the ventricle contours segment-ability on the fused image and input image both. To assess the proposed approach performance, two groups of echocardiograms are selected (as it is shown in Figure 11). All echocardiograms have the identical size of 1149×862 pixel.

4.2.1. Evaluation by Result Comparison. The result of PCA method for the two mentioned datasets is shown in Figure 11(a). It can be simply viewed that image fusion based on PCA alone provides blurred information of tissues especially in heart ventricles. Image fusion based on DWT alone produces a result without clear boundary for ventricles and walls as well (see Figure 11(b)). As it can be seen, the best image fusion outcome is acquired through employing our proposed integration of PCA and DWT fusion method, as it is shown in Figure 11(c).

The characteristics and precise details displayed in result of proposed approach are significantly better compared to other fusion images. The image information such as tissues is improved clearly. Additional valuable information such as ventricle borders and shape is nearly completely achieved.

4.2.2. Evaluation by Segment-Ability. This section provides experiments for the examination of ventricle border segment-ability on the fused image from three methods and input image as well. The ventricle segment-ability is particularly described in the following as the capability of the image for being effectively segmented utilizing a segmentation method. A level set segmentation method which is proposed in [26] is used to show the segmentation result on every image; so, comparison can be effectively done. This algorithm is a geometrically constrained level set segmentation which does not need a training or even prior shape approach and also uses intensity info within the particular image [26] (see Figures 13 and 14).

End systole and end diastole frames of every sequence were recognized by means of a specialist echocardiographer, based on the American Society of echocardiographic guideline [27]. Then the level set segmentation is applied on both phases and end systole and end diastole frames. First, the image was divided into four parts, and then the segmentation process began by putting some sort of elliptical shape of 10 pixels radius in the right bottom square or left bottom square. The variable values assigned for the right and left ventricles' segmentation approach are described in Table 3. The variables of feature detection, in other words edge indicator, as well as segmentation, were maintained identical between the fused image and input one, to examine the actual level of image quality sensitivity.

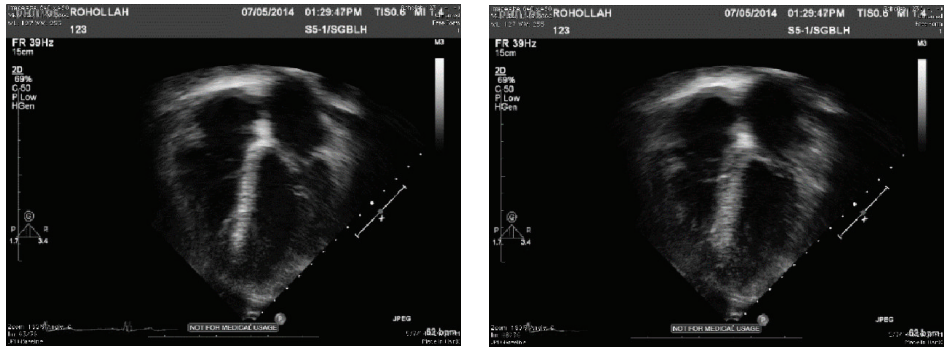
Effective convergence of the segmentation procedure, to achieve close to the left ventricle endocardial boundary, was quantified and the failing of the technique was categorized visually as a result of these probable reasons: (1) ventricle cavity speckle; (2) border loss because of inadequate border info; and (3) both loss and speckle.

In order to verify the measurements, the fused images are segmented manually by a specialist echocardiographer. Ventricles' endocardial boundaries are manually tracked in all planes of the end diastole as well as end systole. Left ventricle trabeculations with papillary muscles are integrated inside the left ventricle cavity, based on the American Society of echocardiographic guideline [27]. The actual result of left ventricle is regarded as source for assessment of segmented one, while employing the validation measurements will be explained in the following.

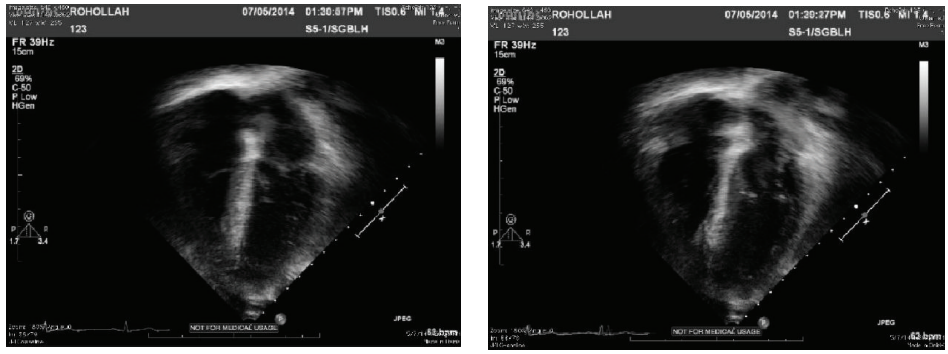
Validation Methods. The ventricles' endocardial acquired through geometrically constrained level set segmentation technique [26] is compared with contour extracted from manual delineation. The subsequent quantitative methods are applied to measure similarity among automated and source contours.

(1) DSI (dice similarity index) is calculated as a way of measuring the agreement among the contour (V) of automated technique and the source contour (V_{ref}), providing a rating value between 0 and 1 (0: no agreement, 1: full agreement). DSI is calculated as

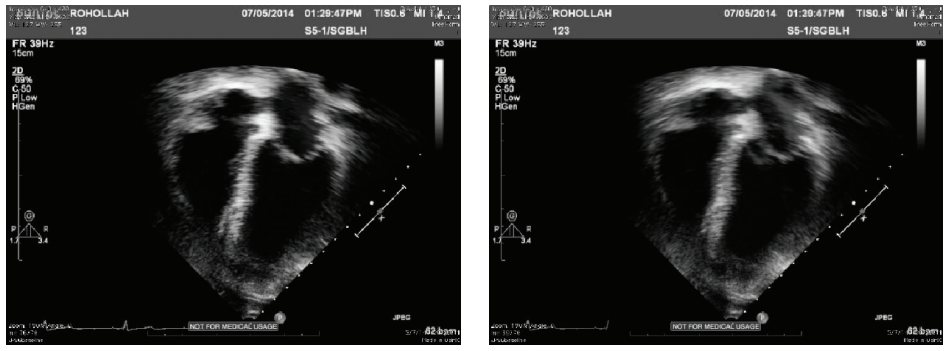
$$\text{DSI} = \frac{2\left(V \cap V_{\text{ref}}\right)}{V + V_{\text{ref}}}, \tag{17}$$



(a) Image fusion based on PCA alone, first image on dataset 1 and the second one on dataset 2



(b) Image fusion based on DWT alone, first image on dataset 1 and the second one on dataset 2



(c) The proposed integration of PCA and DWT fusion result, first image on dataset 1 and the second one on dataset 2

Figure 11: Results of applying PCA, DWT, and the proposed approach on dataset number one and dataset number two.

in which $\cap$ stand for the intersections between the 2 contours.

(2) Mean surface distance: this measure is defined as d_{mean}, among the surface (S) coming from automated technique and the source surface (S_{ref}) described as

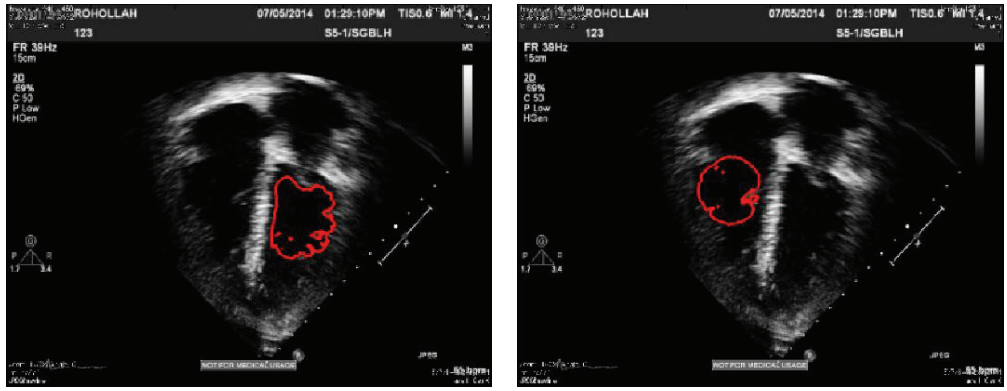
$$d_{\text{mean}} = \frac{1}{2}\left[d\left(S, S_{\text{ref}}\right) + d\left(S_{\text{ref}}, S\right)\right], \tag{18}$$

where $d(S, S_{\text{ref}})$ is the average of distances between every surface pixel in S and the closest surface pixel in S_{ref}, while $d(S_{\text{ref}}, S)$ is calculated in the same way.

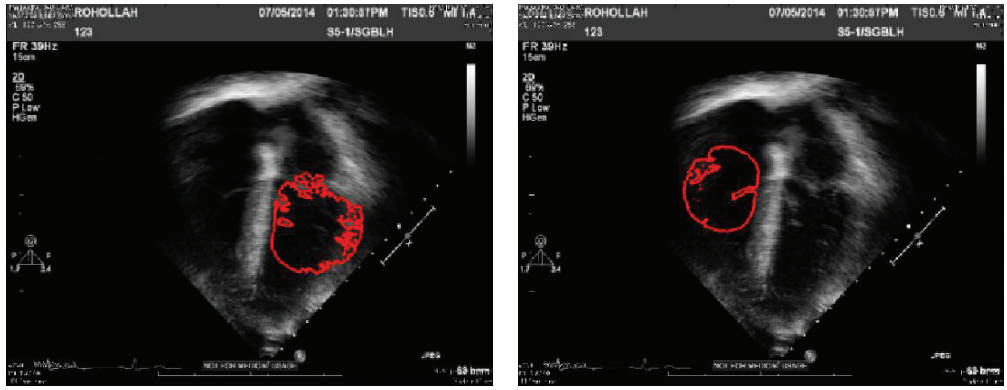
Figure 12 presents the graphic segmentation results for two sets of subject cases visualized on plane of images from fused echocardiography and input image both. It demonstrates a case of effective left ventricle and right ventriclesegmentation at end diastole on the fused image, with a failing on the input image. In this case, the failing is a result of lacking enough image data in addition to cavity noise which in turn leads to loss of the endocardial surface expanding exterior of the correct border.

As it is shown, there is a good result for the fused image, because of lots of noise in left ventricle cavity and a smaller amount observable anatomical info in input image than the fused one.

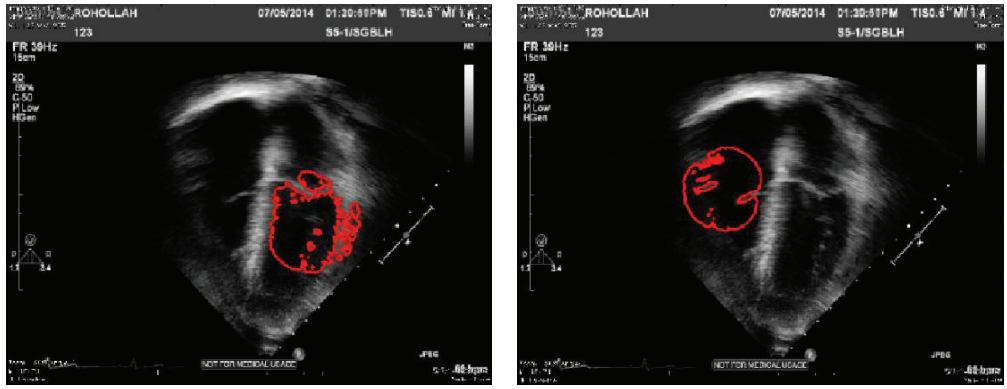
The segmentation result which was explained above shows that the segmentation technique acts more superior on the fused echocardiograms than on input images. This can



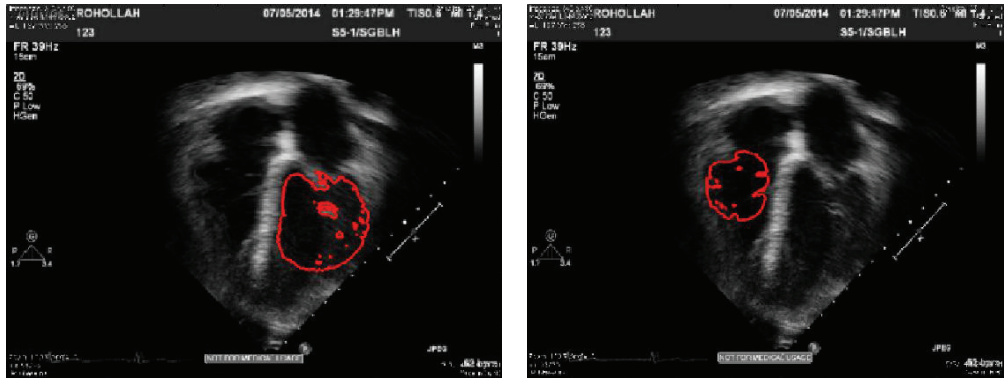
(a) Visual results of geometrically constrained level set segmentation algorithm on input image from dataset 1



(b) Results of Geo segmentation on fused images by DWT level 1 on dataset 1

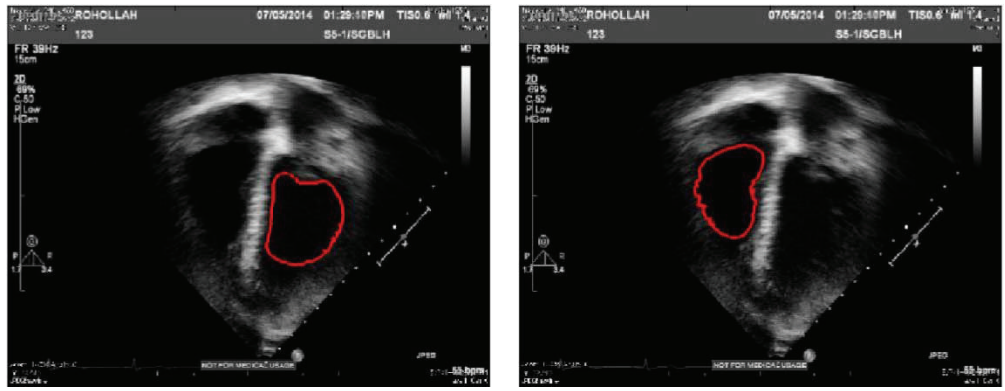


(c) Results of Geo segmentation on fused images by DWT level 2 on dataset 1

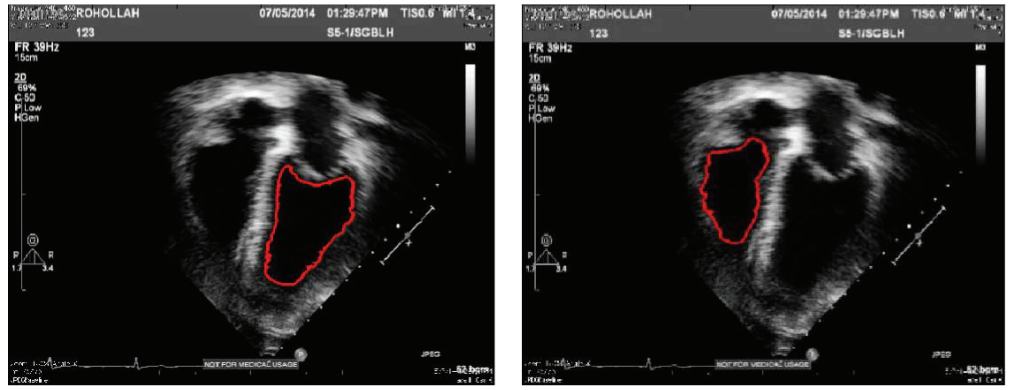


(d) Results of Geo segmentation on fused images by PCA on dataset 1

Figure 12: Continued.



(e) Results of Geo segmentation on fused images by PCA + DWT (level 1) on dataset 1



(f) Results of Geo segmentation on fused images by the proposed method on dataset 1

Figure 12: Visual results of segmentation on end-diastole frame for fused image and input image both; reference image segmentation results have some failure due to leakage.

be, essentially, a direct effect of enhancement in perfection of echocardiogram anatomical description and also improvement in image quality because of multiple image fusions. The final result indicates that the fused images are superior fitted to ventricles' endocardial segmentation qualitatively and quantitatively (Table 1).

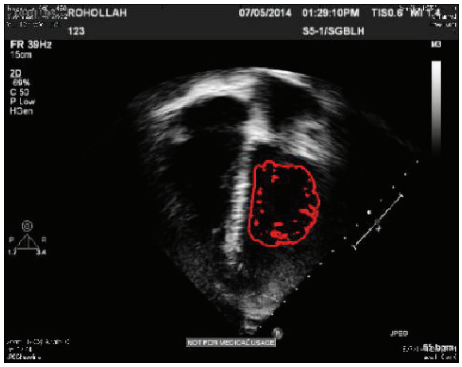
The particular amount of segmentation failures for every failing function is computed. Table 4 summarizes the percent of segmentation failures for fused images and input echocardiograms. For input image, the technique is unable in accurate segmentation in many instances at end diastole (87.3%) and over half the time at end-systole (61.7%). For fused images, the segmentation method is unable in accurate segmentation 24.6% of times at end diastole whilst there was just one failing at end systole (3.1%). This means that fusion results in enhanced image quality which consequently leads to effective ventricles segmentation.

It is seen that input echocardiograms possess greater quantity of ventricles cavity noise in accordance with fused images, as Table 4 shows that ventricles cavity noise is a significant cause of failure (27.8% in end diastole and 59% in end systole) on input images in comparison with absolutely no failures on the fused ones. Lastly, input images are far more impacted by the fused loss along with noise elements (57.8% at end diastole and 39% at end systole) compared to just one such situation (14.7% at end diastole) on fused ones.

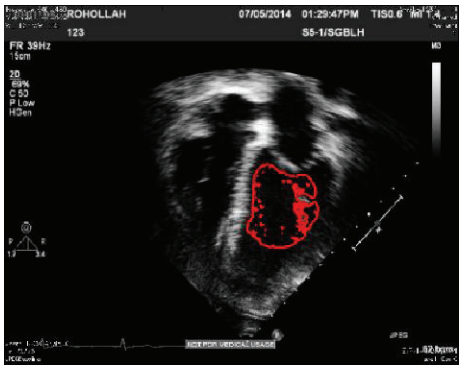
Table 5 presents the comparability among automated and manually delineated contours, at end systole and end diastole. The assessment was done on 8 end systole and 6 end diastole contours, based on the quantity of effective ventricle segmentation on fused echocardiograms, using the validation procedures explained before (Table 4).

Mean DSI measures of 0.91 and 0.79 in end diastole and end systole, respectively, indicate a superior overlap among the manually delineated and segmented contours. The distance failure among the automated and manual contour is smaller, pointed out through average distance of 2.26 mm and 1.64 mm at end diastole and end systole, respectively. Figure 15 demonstrates cases with the automated and manual contours superimposed on the input echocardiograms in order to demonstrate the similarity among them.

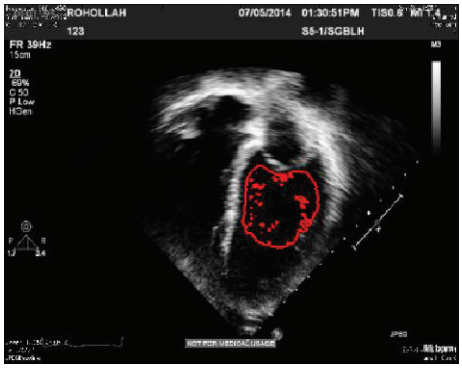
4.3. Accuracy Evaluation of Endocardial Contour. Previously, the analysis has concentrated on image quality, info for segmentation of ventricles, improved repeatability, and achievement endocardial segmentation after fusion. Nevertheless, a crucial problem should be solved: will fusion displace the positioning of endocardials or maintain it in any case? We made a comparison of the endocardial exterior among the fused and input echocardiograms. A subset consists of 6 subjects (fused and input echocardiogram for every subject)



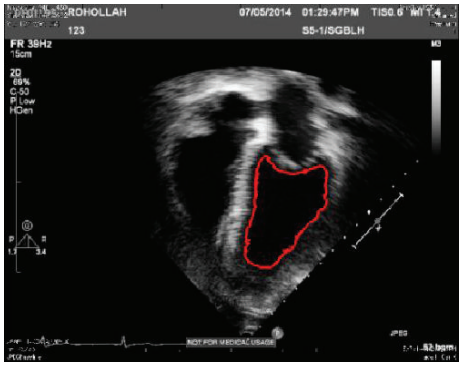
(a) Result on input image; first image from dataset 1



(b) Result on input image; second image from dataset 1



(c) Result on input image; third image from dataset 1



(d) Result on fused image; fused echocardiogram image from dataset 1

Figure 13: Results of segmentation of end systole frame; comparison between results on input images from dataset 1 by fused image from the same dataset.

Table 1: Advantages and disadvantages of the various image fusion methods.

Method	Domain	Advantages	Disadvantages
PCA (principal component analysis)	Spatial	(i) Fused images have high spatial quality (ii) Extract main features to minimize redundancy (iii) Preserve more spectral information contents (iv) Keep a better resolution	(i) Artifacts of spectral feature images and the reference low resolution images with different degrees (ii) Spectral degradation (iii) Not suitable for medical images
DWT (discrete wavelet transform)	Transform	(i) Have least spectral distortion (ii) Preserve more spatial features (iii) Produce directional info in decomposition level and include salient information at various resolution (iv) Give a good signal to noise ratio compare to pixel based approaches	(i) Less spatial resolution for final fused image (ii) Not suitable for medical images
Proposed method (integration of PCA and DWT)	Transform	(i) Integrate PCA and DWT advantages and powerfully improve fused image quality to avoid distortions (ii) Preserve more spatial features and more spectral information contents (iii) Extract main features to minimize redundancy (iv) Output image will contain high spatial resolution and also high quality spectral content	—

Table 2: Comparison of image fusion techniques result; IQI is image quality index; CC is correlation coefficient; RMSE is root mean square error; and OCE is overall cross entropy.

Input	Method	IQI	RMSE	CC	OCE
Dataset 1	**Proposed**	**0.1987**	**0.1041**	**0.9872**	**0.5987**
	PCA	0.1623	0.1252	0.9829	0.9170
	DWT	0.1527	0.1289	0.9819	0.8701
Dataset 2	**Proposed**	**0.2135**	**0.0926**	**0.9798**	**0.4528**
	PCA	0.1678	0.1268	09662	0.5821
	DWT	0.1542	0.1287	0.9654	0.6012

Table 3: Experimental variables of ventricles' segmentation technique.

Border indicator			Level set			Postprocessing procedures	
Gaussian smoothing standard deviation	σ	1.6 or 2[a]	Balloon force weight	α	1	radius of circle, Apex point selection	3
Edge contrast parameter	ν	0.06 or 0.1[a]	Regularization weight	β	0.64	Surface smoothness (Gaussian), standard deviation	0.3
Edge exponent parameter	λ	2	Advection weight	γ	3	—	

[a]Because of dissimilarities in quality throughout the database, a small subset of echocardiograms was prepared with unique variables intended for smoothing and also border indicator calculation. Typically, this is performed as a result of varying image quality in input echocardiography and less often for fused ones.

Table 4: Failure of segmentation technique on fused and input echocardiograms and its particular quantification. The outcomes are demonstrated for 10 subjects. As explained before, segmentation failing means the inability of the segmentation algorithm to achieve the actual endocardial border caused by border loss, noise of cavity, or both equally.

	End diastole		End systole	
	Input Image	Fused image	Input Image	Fused image
Total failures	87.3%	24.6%	61.7%	3.1%
Cavity noise	27.8%	0%	59%	0%
Leakage + noise	57.8%	14.7%	39%	0%

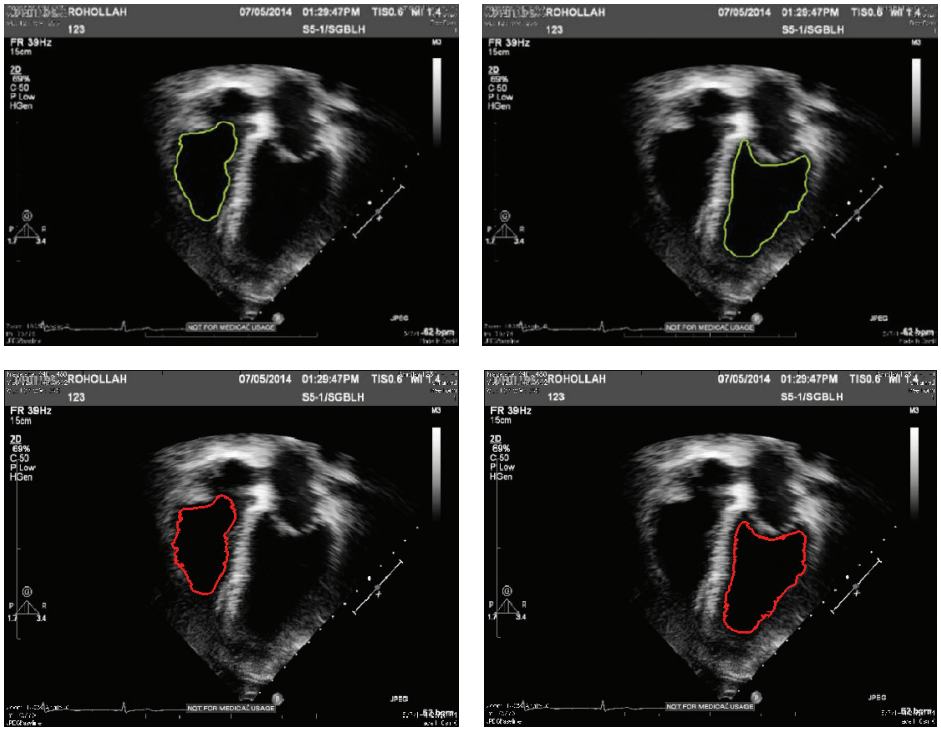


Figure 14: Comparison between manual delineation by expert and segmentation result from proposed method. Green contours: source segmentation through manual delineation done by expert. Red contours: acquired by employing geo segmentation technique on fused echo. The outputs indicate a near agreement among the automated technique and manual delineations.

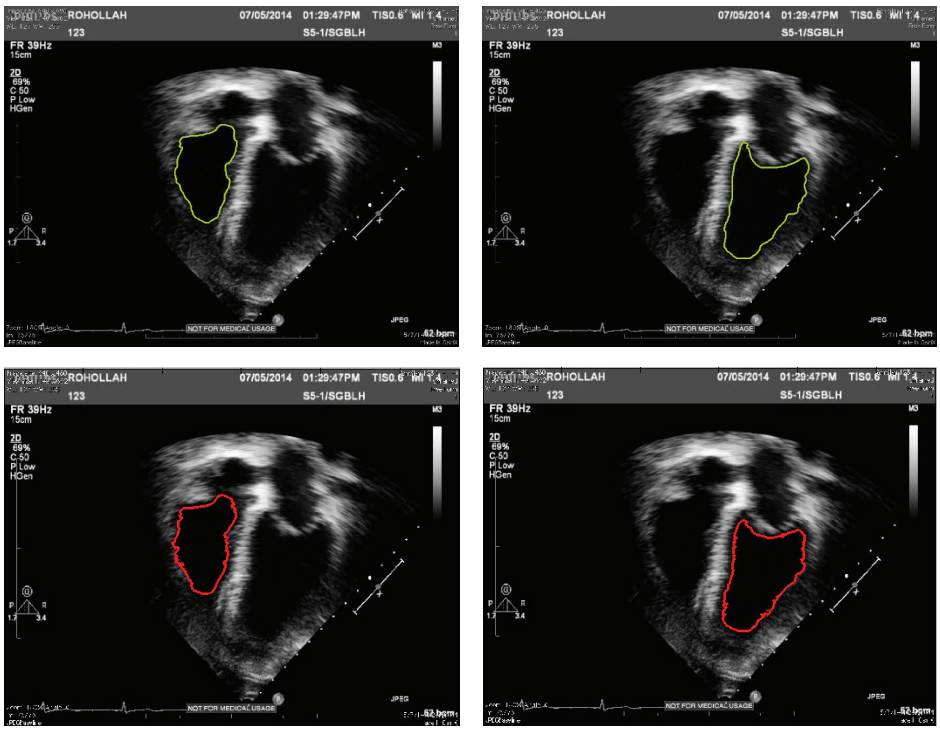


Figure 15: Comparison between manual delineation by expert and segmentation result from proposed method. Green contours: source segmentation through manual delineation done by expert. Red contours: acquired by employing geo segmentation technique on fused multiple view echo. The outputs indicate a near agreement among the automated technique and manual delineations.

Table 5: Failures for evaluation among surfaces from geometrically constrained image segmentation method and manual delineation at end diastole and end systole. DSI (dice similarity index) indicates that the quantity of agreement among the traced and source volumes. d_{mean} (surface distance mean) displays the standard distance, among the traced and source surfaces.

	DSI (mean ± std)	d_{mean} (mean ± std)
End diastole	0.91 ± 0.07	2.26 ± 0.78
End systole	0.79 ± 0.05	1.64 ± 0.45

Table 6: Comparison and evaluation of contours acquired from manual delineation of fused image and input image. The outcomes present a high similarity among them; it means endocardial contours stay stable after fusion and do not displaced through the combination procedure.

	Input image versus fused image
DSI [mean ± std]	0.89 ± 0.03
d_{mean} [mean ± std]	0.95 ± 0.34

and was chosen from database number 1. Those subjects are selected which demonstrated high quality on ventricle cover and anatomical description in single echocardiograms. The endocardial borders are delineated on fused and input echocardiograms both, by a well-experienced expert based on the standard protocol explained previously.

The extracted contours from delineation of fused image and input image are compared utilizing contour distance validation and DSI methods, explained before. The comparison results are presented in Table 6 showing a high agreement with 0.89 DSI and a modest distance failure about 0.95 mm.These outcomes demonstrate that there is high similarity among the endocardials delineated from fused image and input image. Also endocardials are maintained after fusion without obvious displacement in position.

4.4. Discussion. Evaluation has been performed utilizing volunteer and patient databases including ten series of echocardiography images. Result of the proposed method is compared with input (nonfused image) and with result of two other techniques. Experiments demonstrate that the proposed technique is able to get input images, degraded by artifacts, and provide a fused image with better quality. The visual evaluation is done by cardiologists and confirmed very good preference for the combined images regarding quality of image, expanded field of view, low cavity distortion, and high endocardial boundary description. Fusion echocardiographic images provide significant enhancement in anatomical details contained in the image and also in image quality. To analyze the result of this enhancement on automatic image analysis, this particular study described an organized process for evaluation the results using a geometrically constrained level set segmentation algorithm [26]. In this paper, the impact of increasing image info and quality in cardiac ultrasound images is explained objectively and quantitatively. The outcomes showed that fusion helps the automatic analysis in echocardiographic images, drastically.

5. Conclusion

In this study, a new fusion technique for echocardiography images has been presented based on integration of PCA and DWT. Experimental results indicated that the presented technique is effective in fusion echocardiography images and outperforms state-of-the-art developed approaches in quantitative and qualitative evaluation. Statistical and visual comparisons showed that the fusion result of the proposed method include more info, while artifacts are so small. In addition, the presented method can produce more acceptable outcomes, compared to other techniques in the two aspects of visual effect and quantitative analysis. Different metrics are employed to examine the performance of algorithm and it is shown that using discrete wavelet transform with higher level of decomposition incorporating with principal components analysis has better performance in all metrics.

This study mainly evaluated the quality as well as info of fused image and input image for segmentation. For segmentation, a geometrically constrained level set segmentation algorithm [26] was employed to show the result. This level set segmentation algorithm was employed on fused images and input images both. The qualitative and quantitative outcomes demonstrated that segmentation approach performed superiorly on the fused images more than input ones.

Further researches will try to examine ventricles' tracking and also evaluate the segment-ability of myocardial muscle which is more difficult than endocardial boundary. Furthermore, the motion approximation can be done for 3-dimensional strain evaluation, mainly because it basically offers a dense motions field. This study has shown the result of fusion echocardiography on automatic image segmentation for the specific echocardiographic images. The idea of fusion could be definitely placed on other fields of ultrasounds, for instance, fetal echocardiographic images [9]. Generally, we anticipate that the same developments of enhanced functionality in automatic analysis could be found in some other field of imaging of ultrasound. Therefore, fusion has a significant part to perform in ultrasound analysis to boost quantitative analysis.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This research was supported by Ministry of Science, Technology and Innovation, Malaysia, under science Fund no. UPM0007353.

References

[1] M. V. Parvatikar and G. S. Phadke, "Comparative study of different image fusion techniques," *International Journal of Scientific Engineering and Technology*, vol. 3, no. 4, pp. 375–379, 2014.

[2] C. He, G. Cao, and F. Lang, "An efficient fusion approach for multispectral and panchromatic medical imaging," *Biomedical Engineering Research*, vol. 2, no. 1, pp. 30–36, 2013.

[3] A. Ardeshir Goshtasby and S. Nikolov, "Image fusion: advances in the state of the art," *Information Fusion*, vol. 8, no. 2, pp. 114–118, 2007.

[4] T. Singh, M. Kaur, and A. Kaur, "A detailed comparative study of various image fusion techniques used in digital images," *International Journal of Advanced Engineering Technology*, vol. 4, no. 3, pp. 50–52, 2013.

[5] N. Al-Azzawi, H. A. M. Sakim, A. K. W. Abdullah, and H. Ibrahim, "Medical image fusion scheme using complex contour-let transform based on PCA," in *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '09)*, pp. 5813–5816, Minneapolis, Minn, USA, September 2009.

[6] K. Rani and R. Sharma, "Study of different image fusion algorithm," *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, no. 5, 2013.

[7] D. K. Sahu and M. P. Parsai, "Different image fusion techniques: a critical review," *International Journal of Modern Engineering Research*, vol. 2, no. 5, pp. 4298–4301, 2012.

[8] A. P. James and B. V. Dasarathy, "Medical image fusion: a survey of the state of the art," *Information Fusion*, vol. 19, no. 1, pp. 4–19, 2014.

[9] M. J. Gooding, K. Rajpoot, S. Mitchell, P. Chamberlain, S. H. Kennedy, and J. A. Noble, "Investigation into the fusion of multiple 4-D fetal echocardiography images to improve image quality," *Ultrasound in Medicine and Biology*, vol. 36, no. 6, pp. 957–966, 2010.

[10] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[11] H. Ketout and J. Gu, "Left ventricular endocardium tracking by fusion of biomechanical and deformable models," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 302458, 12 pages, 2014.

[12] C. Yao, J. M. Simpson, T. Schaeffter, and G. P. Penney, "Multi-view 3D echocardiography compounding based on feature consistency," *Physics in Medicine and Biology*, vol. 56, no. 18, pp. 6109–6128, 2011.

[13] V. P. S. Naidu and J. R. Raol, "Pixel-level image fusion using wavelets and principal component analysis," *Defence Science Journal*, vol. 58, no. 3, pp. 338–352, 2008.

[14] L. Yang, Y. Mu, L. A. Quaglia et al., "Effectiveness of myocardial contrast echocardiography quantitative analysis during adenosine stress versus visual analysis before percutaneous therapy in acute coronary pain: a coronary artery TIMI grading comparing study," *Journal of Biomedicine and Biotechnology*, vol. 2012, Article ID 806731, 7 pages, 2012.

[15] J.-J. Ding, "Time-Frequency Analysis and Wavelet Transform," http://djj.ee.ntu.edu.tw/index.php.

[16] I. Bankman, *Handbook of Medical Imaging: Medical Image Processing and Analysis*, Academic Press, 1st edition, 2000.

[17] R. M. Tayebi, P. S. B. Sulaiman, R. Wirza, M. Z. Dimon, S. Kadiman, and S. Mazaheri, "A fast and accurate method for automatic coronary arterial tree extraction in angiograms," *Journal of Computer Science*, vol. 10, no. 10, pp. 2060–2076, 2014.

[18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, USA, 2nd edition, 2002.

[19] D. G. Leckie, "Synergism of synthetic aperture radar and visible/infrared data for forest type discrimination," *Photogrammetric Engineering & Remote Sensing*, vol. 56, no. 9, pp. 1237–1246, 1990.

[20] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, USA, 1976.

[21] T. D. Garvey, J. D. Lowrance, and M. A. Fischler, "An inference technique for integrating knowledge from disparate sources," in *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81)*, vol. 1, pp. 319–325, August 1981.

[22] H. Kim and P. H. Swain, "A method for classification of multisource data using interval-valued probabilities and its application to HIRIS data," in *Proceedings of Workshop on Multisource Data Integration in Remote Sensing*, vol. 3099, pp. 27–38, NASA Conference Publication, Greenbelt, Md, USA, 1990.

[23] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002.

[24] D. Chen and Q. Li, "The use of complex contour-let transform on fusion scheme," *Proceedings of World Academy of Science, Engineering and Technology*, vol. 7, pp. 342–374, 2005.

[25] L. Yang, B. L. Guo, and W. Ni, "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," *Neurocomputing*, vol. 72, no. 1–3, pp. 203–211, 2008.

[26] S. Mazaheri, P. S. Sulaiman, R. Wirza, F. Khalid, M. Z. Dimon, and R. M. Tayebi, "A robust and accurate segmentation of echocardiographic images through a geometrically constrained level set," *Journal of Cardiothoracic Surgery*, In press.

[27] R. M. Lang, M. Bierig, R. B. Devereux et al., "Recommendations for chamber quantification: a report from the American Society of Echocardiography's guidelines and standards committee and the Chamber Quantification Writing Group, developed in conjunction with the European Association of Echocardiography, a branch of the European Society of Cardiology," *Journal of the American Society of Echocardiography*, vol. 18, no. 12, pp. 1440–1463, 2005.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 710326, 14 pages
http://dx.doi.org/10.1155/2015/710326

Research Article

Segmentation of Regions of Interest Using Active Contours with SPF Function

Farhan Akram,[1,2] Jeong Heon Kim,[2,3] Chan-Gun Lee,[2] and Kwang Nam Choi[2]

[1]*Department of Computer Engineering and Mathematics, Rovira i Virgili University, 43007 Tarragona, Spain*
[2]*Department of Computer Science & Engineering, Chung-Ang University, Seoul 156-756, Republic of Korea*
[3]*Korea Institute of Science & Technology Information, Daejeon 305-806, Republic of Korea*

Correspondence should be addressed to Kwang Nam Choi; knchoi@cau.ac.kr

Received 16 October 2014; Revised 10 January 2015; Accepted 31 January 2015

Academic Editor: Tianye Niu

Segmentation of regions of interest is a well-known problem in image segmentation. This paper presents a region-based image segmentation technique using active contours with signed pressure force (SPF) function. The proposed algorithm contemporaneously traces high intensity or dense regions in an image by evolving the contour inwards. In medical image modalities these high intensity or dense regions refer to tumor, masses, or dense tissues. The proposed method partitions an image into an arbitrary number of subregions and tracks down salient regions step by step. It is implemented by enforcing a new region-based SPF function in a traditional edge-based level set model. It partitions an image into subregions and then discards outer subregion and partitions inner region into two more subregions; this continues iteratively until a stopping condition is fulfilled. A Gaussian kernel is used to regularize the level set function, which not only regularizes it but also removes the need of computationally expensive reinitialization. The proposed segmentation algorithm has been applied to different images in order to demonstrate the accuracy, effectiveness, and robustness of the algorithm.

1. Introduction

The advanced imaging technologies have improved significantly the quality of medical care for patients. These technologies allowed a radiologist to make increasingly accurate diagnoses of suspicious regions like tumors, polyps, and blood rupture areas and helped physicians to render precise and measured modes of treatment [1]. It is usually a difficult task to identify significant information in a medical image because of intensity inhomogeneity and blurred object boundaries. Expert radiologists are needed to analyze the region of interest (ROI) in a particular image modality, which is costly and time consuming job. There we need an automated ROI system, which can prompt a particular region of interest and can help untrained personnel, saving both time and money. A region of interest analysis is a fundamental step in a computer-aided diagnosis (CAD) system for medical imaging, which helps early detection of cancer [2]. It prompts suspicious regions in different medical images such as magnetic resonance imaging (MRI), X-ray imaging, and computerized tomography (CT) imaging. The necessity of thoroughly examining a large number of image modalities in order to detect small number of cancers can cause high false positive, which can lead to unnecessary biopsies. Moreover, some of the salient regions can be missed due to radiologist's tiredness or distraction [3, 4]. A system which segments regions of interest can help radiologists by tapering their search only to the desired objects in an image. Use of this type of CAD system can increase the efficiency of correct detection of tumors and decrease number of false positives found by the radiologists [5, 6]. Numerous approaches have been developed to solve this tedious but necessary problem of ROI identification and segmentation in the postprocessing of cancer research and treatment. In late 1980s, Kass et al. introduced an active contour method, which is one of well-known techniques used to segment ROI in the computer vision and image processing applications [7]. In this method, a curve is evolved under a force by minimizing the energy until it stops at the object boundary. The energy functional is normally dependent on different characteristics

like curvature, image gradient, and image statistical information. The existing active contour models can be classified into two categories: edge-based models [7–10] and region-based models [11–19]. These two types of models have their own advantages and disadvantages, and the choice between them in applications depends on the different characteristics of images. The edge-based model builds an edge stopping function using image edge information, which enforces the evolution of contour towards the object boundary. A balloon force term is used in the contour evolution process, which helps the contour to shrink or expand. The selection of an accurate balloon force is main problem in this method [18]. Furthermore, for the images with intense noise or weak edges, the edge stopping function based on the image gradient can hardly stop at correct boundaries.

On the other hand, a region-based model uses statistical information to construct a region stopping function, which stops the contour evolution between different regions. Compared to an edge-based model, a region-based model performs better on images with weak or blurred edges. A region-based model is not sensitive to initialization of the level set function and can recognize object's boundaries efficiently. Therefore, region-based models, especially Chan and Vese (CV) model [11], have been widely applied for image segmentation. Although, a region-based model is better than an edge-based model in some aspects but it still has limitations. The traditional region-based models [11, 13] were proposed in the context of binary images with an assumption that input image has homogenous patterns throughout whole image domain. Therefore, such models cannot segment intensity inhomogeneous objects in an image. To solve this problem Zhang et al. proposed a new region-based active contour method [12], which uses the advantages of both CV and geodesic active contour (GAC) models.

Reinitialization, a technique used for occasionally reinitializing a level set function to a signed distance function (SDF) during the evolution, has been extensively used as a numerical remedy for maintaining stable curve evolution and ensuring desirable results. From a practical viewpoint, the reinitialization process can be quite convoluted and expensive and can have subtle side effects [20]. Zhang et al. proposed the active contours with selective local or global (ACSLG) segmentation method, which uses a Gaussian kernel to regularize the level function after each iteration step. It not only regularizes the level set but also removes the need of computationally expensive reinitialization. An edge-based active contour model gives very poor results for the images with intense noise and weak edges, while on the other hand a region-based model gives no satisfactory result for the images with the intensity inhomogeneity. In this paper, a new region-based model is proposed for image segmentation and contour maps computation by incorporating the advantages of algorithms the level set evolution without reinitialization (LSEWR) [9], CV [11], and ACSLG [12] models. The proposed model uses statistical information inside and outside of the contour to construct a region-based signed pressure force (SPF) function, which controls the direction of contour evolution. In the formulated energy function this SPF function substitutes the edge indicator function in LSEWR model.

The proposed SPF function having opposite signs across the object boundary helps level set to shrink and expand. Contour shrinks if the initial contour is outside the boundary of the object and it expands if initial contour is inside the object boundary. In the proposed SPF function a mask term is used to restrict the contour movement inwards. That mask term helps to select the inner region and discard the outer region during the contour evolution process. The proposed algorithm partitions an image into two subregions and then inner part of the subregions is partitioned further into two subregions iteratively and so on until a stopping condition is fulfilled. Subregions are detected through the minimization of a new energy model restricted to a characteristics function of a subregion. If n is the number of iterations in order to attain final segmentation result, then $n + 1$ would be total numbers of regions from initial to final contour.

The introduced model embeds an SPF function based on a traditional region-based model [11] to target images with intensity inhomogeneity. The traditional region-based model [11] cannot properly segment image with intensity inhomogeneity because it cannot differentiate small intensity differences between two consecutive regions and cannot detect weak object boundaries. In the proposed algorithm, the proposed SPF function by using a mask term can control the contour direction and contour stopping algorithm can control the stopping point between two consecutive contours. The proposed algorithm replaced an edge indicator function in LSEWR model with an SPF function to introduce a new region-based model to trace down high intensity ambiguous regions in medical images.

The proposed algorithm contemporaneously traces high intensity or dense regions, which are tumors, masses, or salient dense tissues in medical image modalities. The resulting representation establishes an analysis of the global structure of region of interest. The contour shrinkage depends on the intensity of the region. If intensity difference between background and desired object is high then contour will evolve quickly; otherwise, the contour evolution will take time. The proposed method is good at finding high intensity and dense regions in an image. Therefore, it can properly segment salient dense regions, tumors, polyps, and blood rupture veins. Segmentation of cancer tissues in mammograms, brain tumors in brain magnetic resonance (MR) images, and ruptured blood vessels analysis in the angiography are some of the applications in which proposed method can be used. The formulated algorithm has been applied to different real images in order to demonstrate the accuracy, effectiveness, and robustness of the algorithm. Furthermore, the proposed segmentation algorithm can also be used to produce an adaptive contour map for the topographic analysis of objects in medical images.

2. Related Work

In [13] Mumford and Shah formulated the image segmentation problems as follows: find an optimal piecewise smooth approximation function of u of image I, which varies smoothly within each subregion Ω_i of the image domain $\Omega_i \subset$

R^2 and rapidly or discontinuously goes across the boundaries of Ω_i. They proposed the following energy functional:

$$E_{\mathrm{MS}}(u, C) = \lambda \int_{\Omega} (I - u(x))^2 \, dx + \nu \int_{\Omega \setminus C} |\nabla u|^2 \, dx + \mu \mathrm{Length}(C), \tag{1}$$

where $|\mathbf{C}|$ is the length of the contour C and μ and $\nu \geq 0$ are fixed parameters. The unknown set C and the nonconvexity of the above energy functional make it difficult to be minimized. Some alternative methods have been proposed to simplify or modify the above functional, introduced as follows.

Chan and Vese [11] proposed an active contour method (ACM) based on the Mumford and Shah model [13]. Let $I : \Omega \rightarrow R$ be an input image and let C be a closed curve; the energy functional is defined by

$$E_{\mathrm{CV}}(\phi, c_1, c_2) = \lambda_1 \int_{\Omega} |I(x) - c_1|^2 H_\varepsilon(\phi(x)) \, dx + \lambda_2 \int_{\Omega} |I(x) - c_2|^2 (1 - H_\varepsilon(\phi(x))) \, dx + \mu \mathrm{Length}(C) + \nu \mathrm{Area}(\mathrm{in}(C)), \tag{2}$$

where $\mu \geq 0$, $\nu \geq 0$, λ_1, and $\lambda_2 > 0$ are fixed parameters. The Euclidean length term is used to regularize the contour. c_1 and c_2 are two constants that approximate the image intensities inside and outside of the contour C, respectively. Minimizing the above energy functional by using the steepest gradient descent method [21] and representing the contour C with zero level set, that is, $C = \{x \in \Omega \mid \phi(x) = 0\}$, we obtain the following variational formulation:

$$c_1 = \frac{\int_{\Omega} I(x) H_\varepsilon(\phi(x)) \, dx}{\int_{\Omega} H_\varepsilon(\phi(x)) \, dx}, \qquad c_2 = \frac{\int_{\Omega} I(x) (1 - H_\varepsilon(\phi(x))) \, dx}{\int_{\Omega} (1 - H_\varepsilon(\phi(x))) \, dx}, \tag{3}$$

$$\frac{\partial \phi}{\partial t} = \left(-\lambda_1 (I - c_1)^2 + \lambda_2 (I - c_2)^2 + \mu \, \mathrm{div}\left(\frac{\nabla \phi}{|\nabla \phi|} \right) - \nu \right) \delta_\varepsilon(\phi). \tag{4}$$

The data fitting term $-\lambda_1 (I - c_1)^2 + \lambda_2 (I - c_2)^2$ plays a key role in curve evolution, and λ_1 and λ_2 govern the trade-off between the first and second term. In most cases, we set $\lambda_1 = \lambda_2$ and $\nu = 0$. μ is a scaling parameter. If it is small enough, then small objects are likely to be extracted; if it is large, big objects can be detected [11]. Obviously, in (4), c_1 and c_2 are related to the global properties of the image contents inside and outside the contour, respectively. However, such global image segmentation is not accurate if the image intensity inside or outside the contour is inhomogeneous.

In [12] proposed a region-based active contour method. First a new region-based signed pressure force function is proposed, which efficiently stop the contours at weak or blurred edges. Second the exterior and interior boundaries are automatically detected with the initial contour being elsewhere in the image. Third the proposed ACM with selective binary and Gaussian filtering regularized level set has the property of selective local or global segmentation. It can segment not only the desired object but also the other objects. The computational cost for the traditional reinitialization is reduced. Finally, the proposed algorithm is efficiently implemented by the simple finite difference scheme.

Let Ω be a bounded open subset of R^2 and let $I : [0, a] \times [0, b] \rightarrow R^+$ be a given image. Let $C(q) : [0, 1] \rightarrow R^2$ be a parameterized planar curve Ω in image domain. The GAC is formulated by minimizing the following energy functional:

$$E_{\mathrm{GAC}}(C) = \int_0^1 g(|\nabla I(C(q))|) |C'(q)| \, dq. \tag{5}$$

Using calculus of variation, we get the following Euler Lagrange equation:

$$C_t = g(|\nabla I|) \kappa \vec{N} - (\nabla g \cdot \vec{N}) \vec{N}, \tag{6}$$

where κ is the curvature of the contour computed across the boundary of the system and $\vec{N}$ is the normal computed inwards to the computed contour. Normally a velocity term α is added in order to accelerate the evolution process of the contour. Then the above equation can be rewritten as

$$C_t = g(|\nabla I|)(\kappa + \alpha) \vec{N} - (\nabla g \cdot \vec{N}) \vec{N}. \tag{7}$$

The corresponding level set formulation will be as follows:

$$\frac{\partial \phi}{\partial t} = g |\nabla \phi| \left(\mathrm{div}\left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \alpha \right) + \nabla g \cdot \nabla \phi. \tag{8}$$

The SPF function defined in [22] is in the range $[-1, 1]$. It modulates the sign of pressure force inside and outside of the region of interest and it is used to shrink the contour when outside the object and expands it when inside the object. The mathematical formulation of the proposed SPF function is as follows:

$$\mathrm{spf}(I) = \begin{cases} \dfrac{I(x) - (c_1 + c_2)/2}{\max(|I(x) - (c_1 + c_2)/2|)}, & I(x) \neq 0 \\ 0, & I(x) = 0. \end{cases} \tag{9}$$

By replacing the edge indicator function with the sign pressure force function we get the following formulation:

$$\frac{\partial \phi}{\partial t} = \mathrm{spf}(I(x)) |\nabla \phi| \left(\mathrm{div}\left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \alpha \right) + \nabla \mathrm{spf}(I(x)) \cdot \nabla \phi. \tag{10}$$

In [12] a new regularization method is introduced for the level set regularization; therefore, curvature term can be removed.

In addition, the term $\nabla \text{spf} \cdot \nabla\phi$ in above equation can also be removed, because their model utilizes the statistical information of regions, which has a larger capture range and capacity of antiedge leakage. Finally, the level set formulation of the proposed model can be written as follows:

$$\frac{\partial \phi}{\partial t} = \text{spf}\left(I\left(x\right)\right) \cdot \alpha \left|\nabla\phi\right|. \tag{11}$$

In [19] a region-based active contour method is proposed, which utilizes an SPF function based on a traditional active contour method [11]. The proposed method uses both region and edge-based information in the implementation of the energy formulation. It is formulated by using edge-based model [9] as a base and by replacing the edge indicator function with a region-based SPF function, which was formulated using traditional active contour method [11].

Let $I : \Omega \rightarrow R$ be an input image and let C be a closed curve; the energy functional is defined by

$$E_{g,\lambda,\nu}\left(\phi\right) = \lambda L_g\left(\phi\right) + \nu A_g\left(\phi\right) + \mu P\left(\phi\right), \tag{12}$$

where L_g is length term of the level set curve, A_g is the area term which deals with the area across the object of interest, and P is the penalizing term used to remove the penalizing energy during the level set curve evolution process:

$$\begin{aligned} E_{g,\lambda,\nu}\left(\phi\right) = {} & \lambda \int_\Omega g\delta\left(\phi\right)\left|\nabla\phi\right| dx\, dy \\ & + \nu \int_\Omega gH\left(-\phi\right) dx\, dy \\ & + \mu \int_\Omega \frac{1}{2}\left(\left|\nabla\phi\right| - 1\right)^2 dx\, dy. \end{aligned} \tag{13}$$

By replacing the edge indicator function g with an SPF function based on a traditional region-based active contour method we get the following formulation:

$$\begin{aligned} E_{g,\lambda,\nu}\left(\phi\right) = {} & \lambda \int_\Omega g\delta\left(\phi\right)\left|\nabla\phi\right| dx\, dy \\ & + \nu \int_\Omega \text{spf}\left(I\right) H\left(-\phi\right) dx\, dy \\ & + \mu \int_\Omega \frac{1}{2}\left(\left|\nabla\phi\right| - 1\right)^2 dx\, dy. \end{aligned} \tag{14}$$

By the calculus of variations [21] the steepest descent process for minimization of the energy functional E is the following gradient flow:

$$\begin{aligned} \frac{\partial \phi}{\partial t} = {} & \lambda\delta\left(\phi\right) \text{div}\left(g\frac{\nabla\phi}{\left|\nabla\phi\right|}\right) + \nu\, \text{spf}\left(I\right)\delta\left(\phi\right) \\ & + \mu\left[\Delta\phi - \text{div}\left(\frac{\nabla\phi}{\left|\nabla\phi\right|}\right)\right], \end{aligned} \tag{15}$$

where Δ is the Laplacian operator used in the energy penalization step during the evolution of level set curve.

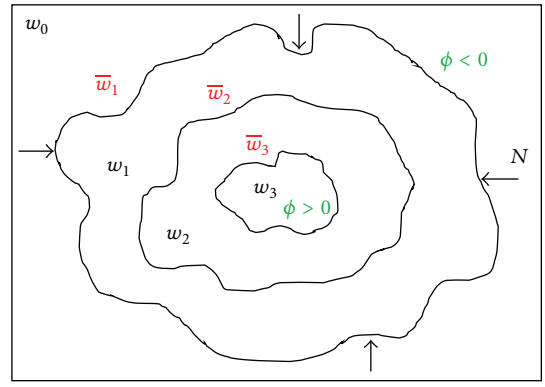


Figure 1: Calculated isocontour map from a given image w_0.

3. The Proposed Region-Based Active Contour Method for Segmentation of Region of Interest and Contour Mapping

A curve C in Ω is represented by a level set function $\phi : \Omega \rightarrow R$, which is zero $\phi = 0$ at object boundary in image I. Curve C partitions a subregion $W_k \subset \Omega$ into two subregions w, $\overline{w}$ with ϕ, such that

$$\begin{aligned} \text{inside}\left(C\right) &= w = \left\{x \in W_k : \phi\left(x\right) > 0\right\}, \\ \text{outside}\left(C\right) &= \overline{w} = \left\{x \in \Omega : \phi\left(x\right) < 0 \cup x \in \Omega \setminus W_k\right\}. \end{aligned} \tag{16}$$

In the proposed algorithm, the evolution of the level set starts from the initial level set function and goes on by moving the contour inwards. In order to compute the contour map of the image the initial contour is defined at the boundary of the image. To move the level set inwards, we calculate the inner subregion of zero level curve using $\phi > 0$.

Figure 1 shows an isocontour map image in which the inner regions w_1, w_2, and w_3 are calculated using $\phi > 0$, whereas the outer regions are calculated by subtracting the energy component of the current calculated inner subregion from the previous calculated inner region, as shown in the equations below:

$$\begin{aligned} & w_0 = w_1 + \overline{w}_1 \Longrightarrow \overline{w}_1 = w_0 - w_1, \\ & \overline{w}_2 = w_1 - w_2, \\ & \overline{w}_3 = w_2 - w_3. \end{aligned} \tag{17}$$

In general, the outer subregion can be calculated as

$$\overline{w}_k = w_{k-1} - w_k. \tag{18}$$

In image segmentation, active contours are dynamic curves that move toward the object boundaries. To achieve this goal, we explicitly define an external energy that can move the zero level set towards the object boundaries. We define energy functional for function ϕ as follows:

$$E_{\text{spf}}\left(\phi\right) = \lambda L_{\text{spf}}\left(\phi\right) + \nu A_{\text{spf}}\left(\phi\right), \tag{19}$$

where $\lambda > 0$ and v are constants, and the terms $L_{\text{spf}}(\phi)$ and $A_{\text{spf}}(\phi)$ are defined as below:

$$L_{\text{spf}}(\phi) = \int_{\Omega} \text{spf}(I)\, \delta_{\varepsilon}(\phi)\, |\nabla\phi|\, dx, \tag{20}$$

$$A_{\text{spf}}(\phi) = \int_{\Omega} \text{spf}(I)\, H_{\varepsilon}(-\phi)\, dx, \tag{21}$$

respectively. Here, spf(I) is the proposed SPF function defined in (29), while H_{ε} is the Heaviside function and $\delta_{\varepsilon} = H'_{\varepsilon}$ is the univariate Dirac delta function given in (30) and (31), respectively. The zero level curve C is driven into a smooth curve from a complicated curve to minimize the function $L_{\text{spf}}(\phi)$. The small energy of $A_{\text{spf}}(\phi)$ accelerates the evolution. The SDF satisfies the desirable property $|\nabla I| = 1$. The energy functional $E_{\text{spf}}(\phi)$ drives the zero level set toward the object boundaries. The coefficient v of $A_{\text{spf}}(\phi)$ in (21) can be positive or negative, depending on the relative position of the initial contour to the object of interest. For example, if the initial contours are placed outside the object, the coefficient v in the weighted area term should take a positive value, so that the contour can shrink faster. If the initial contours are placed inside the object, the coefficient v should take a negative value to speed up the expansion of the contours.

We define a new region-based energy functional based on CV [11] as shown in (2) with an additional mask term M^k to restrict the contour evolution inwards. We proposed E_{proposed} in order to formulate the SPF function that is used in (20) and (21). Let $I : \Omega \rightarrow R$ be an input image and let C be a closed curve; the energy functional E_{proposed} is defined by

$$\begin{aligned} E_{\text{proposed}} = {} & \int_{\Omega} |I(x) - c_1|^2 H_{\varepsilon}(\phi(x))\, M^k(x)\, dx \\ & + \int_{\Omega} |I(x) - c_2|^2 (1 - H_{\varepsilon}(\phi(x)))\, M^k(x)\, dx. \end{aligned} \tag{22}$$

Keeping ϕ fixed and minimizing the energy E_{proposed} with respect to c_1 and c_2, it is easy to express these constant functions of ϕ; we get c_1 and c_2 of regions w and $\overline{w}$, respectively, as follows:

$$c_1(\phi) = \frac{\int_{\Omega} I(x)\, H_{\varepsilon}(\phi(x))\, M^k(x)\, dx}{\int_{\Omega} H_{\varepsilon}(\phi(x))\, M^k(x)\, dx}, \tag{23}$$

$$c_2(\phi) = \frac{\int_{\Omega} I(x)\, (1 - H_{\varepsilon}(\phi(x)))\, M^k(x)\, dx}{\int_{\Omega} (1 - H_{\varepsilon}(\phi(x)))\, M^k(x)\, dx}. \tag{24}$$

Here, M^k is the characteristic function of subregion W_k, defined as

$$\begin{aligned} M^k(x) &= \phi > 0, \\ M^0 &: \Omega \longrightarrow -1. \end{aligned} \tag{25}$$

By the calculus of variations [21], the Gateaux derivative (first variation) of the functional $E_{\text{spf}}(\phi)$ in (19) can be written as

$$Q(\phi) = \left(\lambda \cdot \text{div}\left(\text{spf}(I) \cdot \frac{\nabla\phi}{|\nabla\phi|}\right) + v \cdot \text{spf}(I)\right) \delta_{\varepsilon}(\phi). \tag{26}$$

The function ϕ that minimizes this functional satisfies the Euler Lagrange equation $\partial E_{\text{spf}}/\partial\phi = 0$. The first term corresponds to the $L_{\text{spf}}(\phi)$ (weighted length term), which deals with curvature of the object boundary based on edge information, whereas the second term $A_{\text{spf}}(\phi)$ (weighted area term) is used to compute the area of regions of interest in image inside of the initial contour. If the SPF value is positive, then contour moves toward the high intensity region and vice versa.

A classical iterative process for minimizing the function $E_{\text{spf}}(\phi)$ is the gradient flow with artificial time t given as

$$\phi_{(t=0)} = \phi_0, \qquad \frac{\partial\phi}{\partial t} = Q(\phi). \tag{27}$$

After evolving the level set using (26) and (27) we regularize it by using $\phi^n = G_{\sigma} * \phi^n$. It not only regularizes the level set function but also eliminates the need of reinitialization, which is computationally very expensive. Regularization using Gaussian kernel has better smoothing results and no energy leakage as compared to the area smoothing and penalization terms used by Li et al. [9].

An SPF function is a mathematical expression whose value is in the range $[-1, 1]$ inside and outside of region of interest. Numerous ways have been devised to formulate an SPF function [12, 23, 24] some of them incorporate global intensity means and some use local intensity means in its construction. In this paper we propose a new SPF function, which uses global fitted image restricted with a mask term to enforce level set evolution inwards. If mask term is set to 1 then it modulates the signs of the pressure forces inside and outside the region of interest so that the contour shrinks when outside the object and expands when inside the object. A two-phase global fitted image model is for a level set function defined as follows:

$$I_{\text{GFI}} = c_1 H_{\varepsilon}(\phi) + c_2 (1 - H_{\varepsilon}(\phi)). \tag{28}$$

Using the global fitted model defined above we constructed the SPF function as follows:

$$\text{spf}(I) = \begin{cases} \dfrac{(I(x) - I_{\text{GFI}})\, M^k}{\max(|I(x) - I_{\text{GFI}}|)} & I(x) \neq 0, \\ 0 & I(x) = 0. \end{cases} \tag{29}$$

The terms $c_1(\phi)$, $c_2(\phi)$, and M^k are defined in (23), (24), and (25), respectively. The signs of SPF function are shown in Figure 2, which is negative for the inside region and positive for the outside region, considering an inside region with higher intensity than the outer one. The mask term M^k from (25) is to restrict the evolution of contour inside of the object boundary.

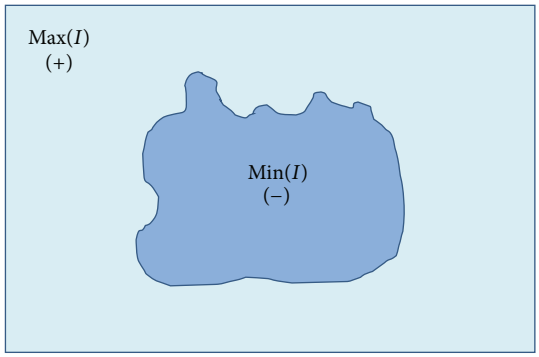


Figure 2: The signs of the SPF function inside and outside the object are opposite one another.

In the proposed work, the Dirac function $\delta_\varepsilon(z)$ and Heaviside function $H_\varepsilon(z)$ used in (23), (24), and (26) are the smoothed versions of the Dirac function and Heaviside function over the entire region. The approximations $H_\varepsilon(z)$ and $\delta_\varepsilon(z)$ as proposed in [11] are

$$H_\varepsilon(z) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\left(\frac{z}{\varepsilon}\right)\right), \tag{30}$$

$$\delta_\varepsilon(z) = \frac{\varepsilon}{\pi\left(z^2 + \varepsilon^2\right)}. \tag{31}$$

We use the regularized Dirac $\delta_\varepsilon(z)$ and Heaviside $H_\varepsilon(z)$ with $\varepsilon = 1.5$ for all the experiments in this paper, and the curvature term is computed using the central difference method.

In outmoded level set methods, it is essential to initialize the level set function ϕ as a signed distance function (SDF) ϕ_0. If the initial level set function is expressively different from the SDF, then the reinitialization schemes are unable to reinitialize the function to the SDF. In our formulation, not only is the reinitialization procedure completely eliminated, but the level set function ϕ also no longer needs to be initialized as SDF. The initial level set function ϕ_0 is defined as

$$\phi(x, t = 0) = \begin{cases} -\rho & x \in \Omega_0 - \partial\Omega_0, \\ 0 & x \in \partial\Omega_0, \\ \rho & x \in \Omega - \Omega_0, \end{cases} \tag{32}$$

where $\rho > 0$ is a constant and we use ϕ_0 with $\rho = 2$.

In order to stop the contour at certain point a stopping algorithm is used, which checks the similarity of pixels between two consecutive contours using stopping value. If

$$\sum_{i=0}^{\text{row}}\sum_{j=0}^{\text{col}} M_{i,j}^k < \left(\frac{\text{stopping value}}{100}\right)\sum_{i=0}^{\text{row}}\sum_{j=0}^{\text{col}} \text{old}M_{i,j}^k \tag{33}$$

then contour will stop moving any further, where old M^k is the mask term of the last computed contour, M^k is the mask term of the current computed contour, row is maximum number of rows, and col is maximum number of columns of the input image. The stopping value is always $98 <$ stopping value < 100 and computed by calculating the mean intensity value from the initial contour with the scale always in between 98 and 100. A threshold value T is used to remove small values while calculating stopping value. In the experiments related to mammograms we selected $T = 25$.

Finally, the principle steps of the algorithm can be summarized as follows.

(a) Initialize ϕ by $-\phi_0$ and M^k by $-M^0$, using (32) and (25), respectively, at $k = 0$.

(b) Compute $c_1(\phi)$ and $c_2(\phi)$ from (23) and (24), respectively.

(c) Calculate spf(I) using (29).

(d) Solve the partial differential equation (PDE) in ϕ from (26) and (27), to obtain ϕ^k.

(e) Calculate M^k using the previously calculated ϕ^k value (in next iteration, which is ϕ^{k-1}).

(f) Regularize the level set function by a Gaussian kernel; that is, $\phi^k = G_\sigma * \phi^k$, where σ is standard deviation.

(g) Check whether solution is stationary using stopping algorithm discussed above. If not, go to step (b), $k = k + 1$, and repeat.

If one of the subregions w, $\overline{w}$ is empty, then the formulation degenerates and the algorithm automatically terminates.

Although the main objective of the proposed method is to extract regions of interest from medical images it can also provide step by step contour information to build a contour map for the topographic analysis of a medical image. Using the contour map we can analyze the structure of a given image in more detail and extract the meaningful information more easily. Figure 3 shows how contour map helps in analyzing and pinpointing the salient region in a mammogram image with salient regions.

Figure 3(a) shows the original image with the initial contour. Figure 3(b) shows the contour map of 400 iterations with a step size of 50 iterations on a masked image and Figure 3(c) shows the computed contour map on the given image. In Figure 3(c) we can identify the salient regions with the area that contains thick contours. Or simply we can say that the salient regions are the regions in the contour map image where consecutive contour lines are dense or very close to each other. In Figure 3(a) salient regions are identified with arrows, where the top arrow identifies the pectoral muscle, which we need to ignore. Figure 3(c) verifies the dense contour lines for salient regions.

The computed contour map can further be used to form a contour tree which is a way to analyze topographic data in more depth. A contour tree is a graph that tracks contours of the level set as they split, join, appear, and disappear [25]. This provides a hierarchical representation of the enclosure relationship between isocontours. In the contour tree, each node represents a contour at a parent node spatially that encloses the contour at its descendent nodes. A node is classified as branching or nonbranching depending on its degree. A node with more than one intermediate descendent

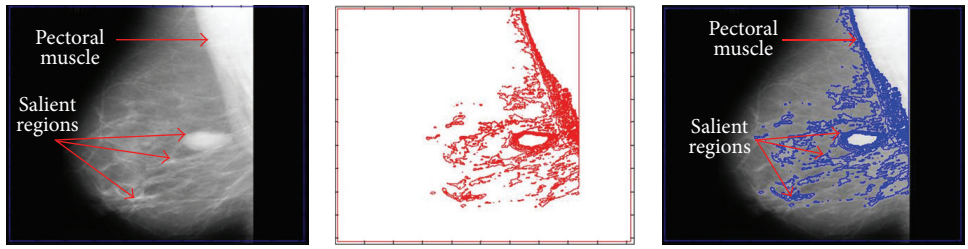


FIGURE 3: Isocontour map with step size of 50 iterations. (a) Initial contour nearby the boundary of image. (b) Isocontour map for 400 iterations. (c) Isocontour map on given mammogram image for 400 iterations.

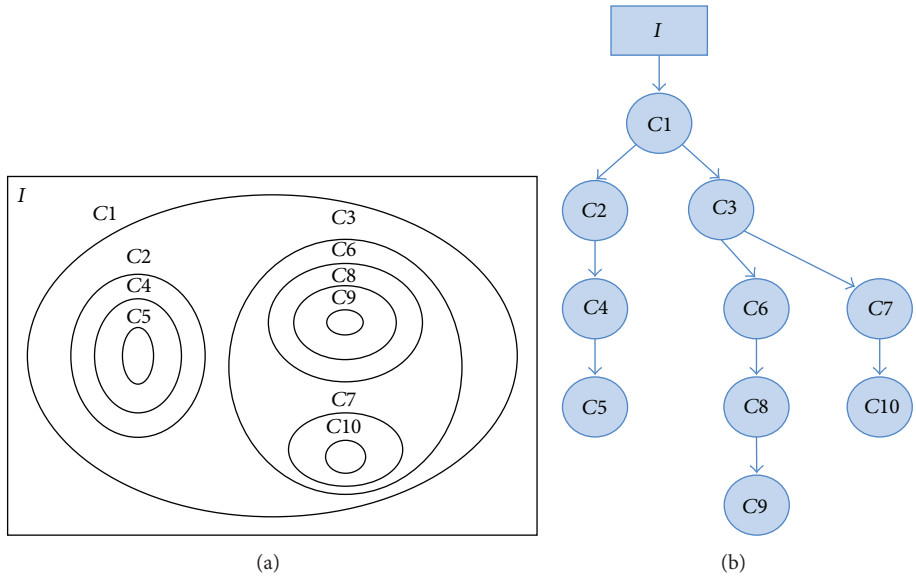


FIGURE 4: Schematic illustration of the inclusion tree. (a) Isocontour map containing branching contour and base contours. (b) Corresponding inclusion tree.

node is called a branching node, and a nonbranching node has only one or zero immediate descendent nodes. The contours at the branching nodes in the inclusion tree are called branching contours and the contours at their intermediate descendants are called base contours. In particular, contours at the terminal nodes that have no further child due to fulfilled stopping condition are called terminal contours [2]. The concept of inclusion tree is introduced here because one of the applications of the proposed algorithm is to create contour tree from the developed contour map. A schematic illustration of the isocontour map is shown in Figure 4(a), while Figure 4(b) shows the inclusion tree of the isocontour map; here, contours $C1$ and $C3$ are branching contours, contours $C2$, $C6$, and $C7$ are base contours, and contours $C5$, $C9$, and $C10$ are terminal contours.

We can measure the saliency of contour using minimum nesting depth from inclusion tree. The nesting depth for a contour is given by the number of contours from the innermost contour to the contour within the nesting structure. The base contours with higher minimum nesting depth correspond to the boundaries of distinctive regions with abrupt intensity changes.

4. Results

The proposed technique is applied to the mammogram images from the mini-MIAS database [26]. The range of intensities in all images is represented from 0 to 255, while the size in pixels (length × width) of the images is 1024 × 1024. The used experiment environment is Windows 7, Quad Core CPU 2.4 GHz, and 8 GB RAM. In the proposed method the following parameters are used: $K = 5$, $\sigma = 1$, $\lambda = 1.0$, $\nu = 15$, and $\Delta t = 1$, where Δt is a time step in the numerical implementation of (27). K is the width of the Gaussian kernel and σ is the standard deviation, which uses all points around the center point to make the level set function smooth.

In the proposed method parameters are selected manually. In order to achieve optimized values a small analytical experiment is conducted using a breast image (mdb005) from the mini-MIAS database. In the proposed method λ, σ and ν are the main values, which affect the evolution of contour and its accuracy. In order to find best values of λ, σ, and ν, F_1 score (F measure) is computed at different values of λ, σ, and ν. F_1 score is a measure to compute accuracy of a method, its value is close to 1 when method has high accuracy, and its value is

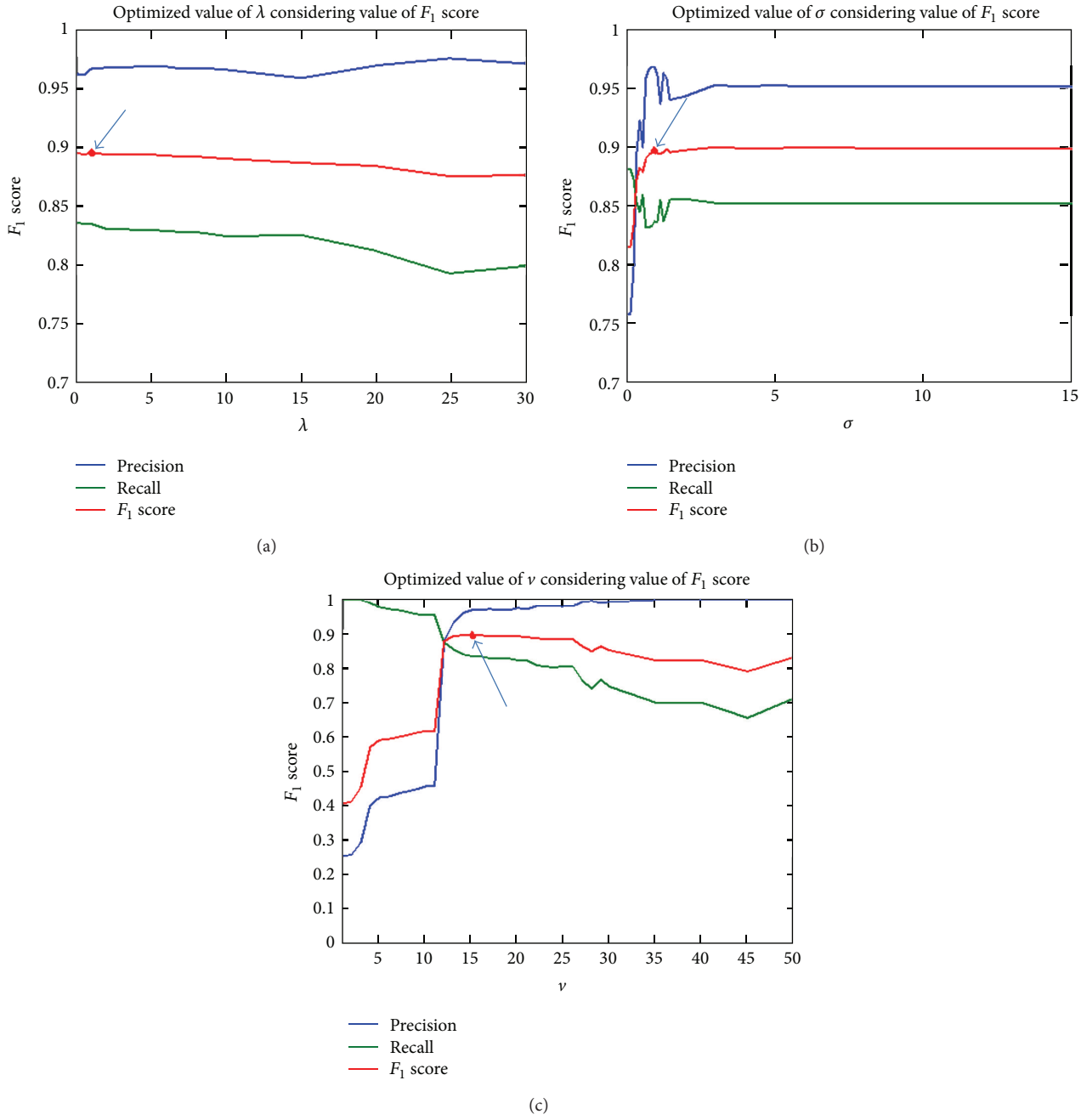
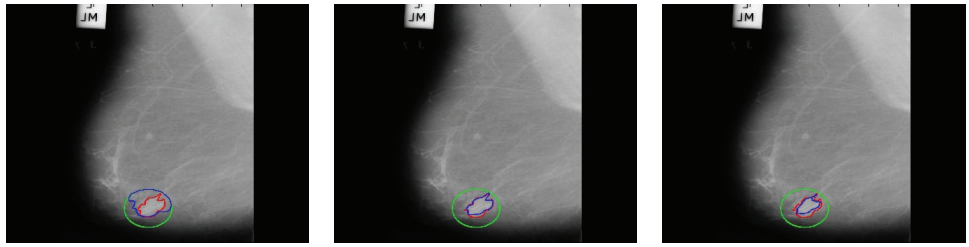


FIGURE 5: Finding optimal parameters for the proposed method using F_1 score analysis. (a) F_1 score at different value of λ (keeping σ and ν constant), (b) F_1 score at different value of σ (keeping λ and ν constant), and (c) F_1 score at different value of ν (keeping λ and σ constant).

close to 0 when method has low accuracy. The mathematical expression to compute F_1 score is as follows:

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{34}$$

where precision is the number of correct positive results divided by the number of all positive results and recall is the number of correct positive results divided by the number of positive results that should have been returned. The mathematical formulation of precision and recall is given in Section 5. Here, F_1 score is computed for all three parameters (λ, σ, and ν) one by one by keeping two parameters constant and varying one parameter on a certain scale to see on which parameter F_1 score is close to 1. Figure 5 shows precision, recall, and F_1 score of proposed method (applied on mdb005 from mini-MIAS database) at different values of λ, σ, and ν. Precision and recall should also be close to 1 for high accuracy. F_1 score is harmonic mean of precision and recall; therefore, if F_1 score is highest at some point then collective weight of precision and recall at that point is also high. Here, in Figure 5



(a) (b) (c)

Figure 6: Segmentation results at different ν. (a) At $\nu = 5$ with small step size, (b) at $\nu = 15$ with medium step size, and (c) at $\nu = 30$ with big step size.

we will only study F_1 score because its value is representing both precision and recall.

Figure 5(a) shows F_1 score of the proposed method at different values of λ (keeping σ and ν constant). It shows that at $\lambda = 1$ F_1 score has maximum value, which means it provides best result at $\lambda = 1$. From $0 \leq \lambda \leq 6$ contour is stable but for all values at $\lambda > 6$ contour evolution is unstable in first iterations because of oversaturation of energy in length term. Figure 5(b) shows F_1 score has maximum at $\sigma = 1$ and $\sigma = 1.4$. In all values at $\sigma > 0.8$ F_1 score is almost same. We choose $\sigma = 1$ because at small σ algorithm has less time complexity. If we increase σ window size of the smoothing kernel will also be increased, which will also add up the computations and increase the time complexity of the algorithm. Figure 5(c) shows that F_1 score has maximum value at $\nu = 15$. From the F_1 score analysis we found that at $\lambda = 1$, $\sigma = 1$, and $\nu = 15$ the proposed method gives high accuracy; therefore, we set the parameters to these values.

Here, ν is the most import parameter, which controls the scaling of the area term in the proposed algorithm. By changing the value of ν, the step size between two consecutive contours can be controlled, which can affect the segmentation accuracy of the region of interest. Figure 6 shows segmentation results using different values of ν. It shows that at $\nu = 5$ when step size is small contour stop evolving earlier than expected, while at $\nu = 30$ when step size is big contour stopped later than expected and at $\nu = 15$ more accurate segmentation results are achieved.

The proposed method is tested on 116 images with tumor tissues out of 315 total images in the database. No ground truths of tumor tissues are pregiven in the mini-MIAS database; only the information regarding location and diameter is mentioned. In the first step of our result evaluation we drew ground truth by ourselves with hand based marker using the location and diameter information given in database. Location of region of interest is kept in mind during initialization of initial contour. It is initialized outside the ground truth region (surrounding region of interest) and not far away in order to increase the accuracy. The proposed algorithm segmented all 116 images but some with accuracy problems. Figure 7 shows the PR curve computed using precision and recall from segmentation results of all 116 images. Visually, it shows overall accuracy of over 80%. In the computed segmentation results from 116 mammogram

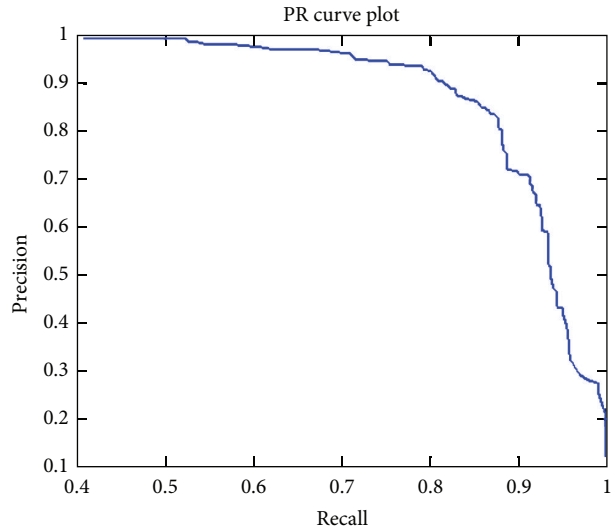


Figure 7: PR curve using precision and recall from segmentation results of 116 images with tumor tissues from mini-MIAS database.

images, 26.72% results have precision of 0.90 or more, 28.45% results have precision of 0.80 or more, 5.17% results have precision of 0.70 or more, 12.07% results have precision of 0.60 or more, 5.17% results have precision 0.50 or more, and 21.41% result have precision less than 0.49, while among 116 of total segmented mammogram images 43.10% results have recall of 0.90 or more, 41.38% have recall of 0.80 or more, 5.17% results have recall of 0.70 or more, 5.17% results have recall of 0.60 or more, 4.27% results have recall of 0.50 or more, and 0.86% have recall less than 0.49. Moreover, maximum achieved precision and recall from segmentation results are 0.99 and 1, respectively. While minimum achieved precision and recall are 0.12 and 0.41, respectively. The maximum scale for precision and recall values is 1. Figure 8 shows some of the well segmented tumor tissue results from the mini-MIAS database. In this figure initial contour is shown in green, ground truth contour is shown in red, and computed contour is shown in red. Note that the initial contour is always taken outside the region of interest because in the proposed algorithm contour always moves inwards discarding the outer contour.

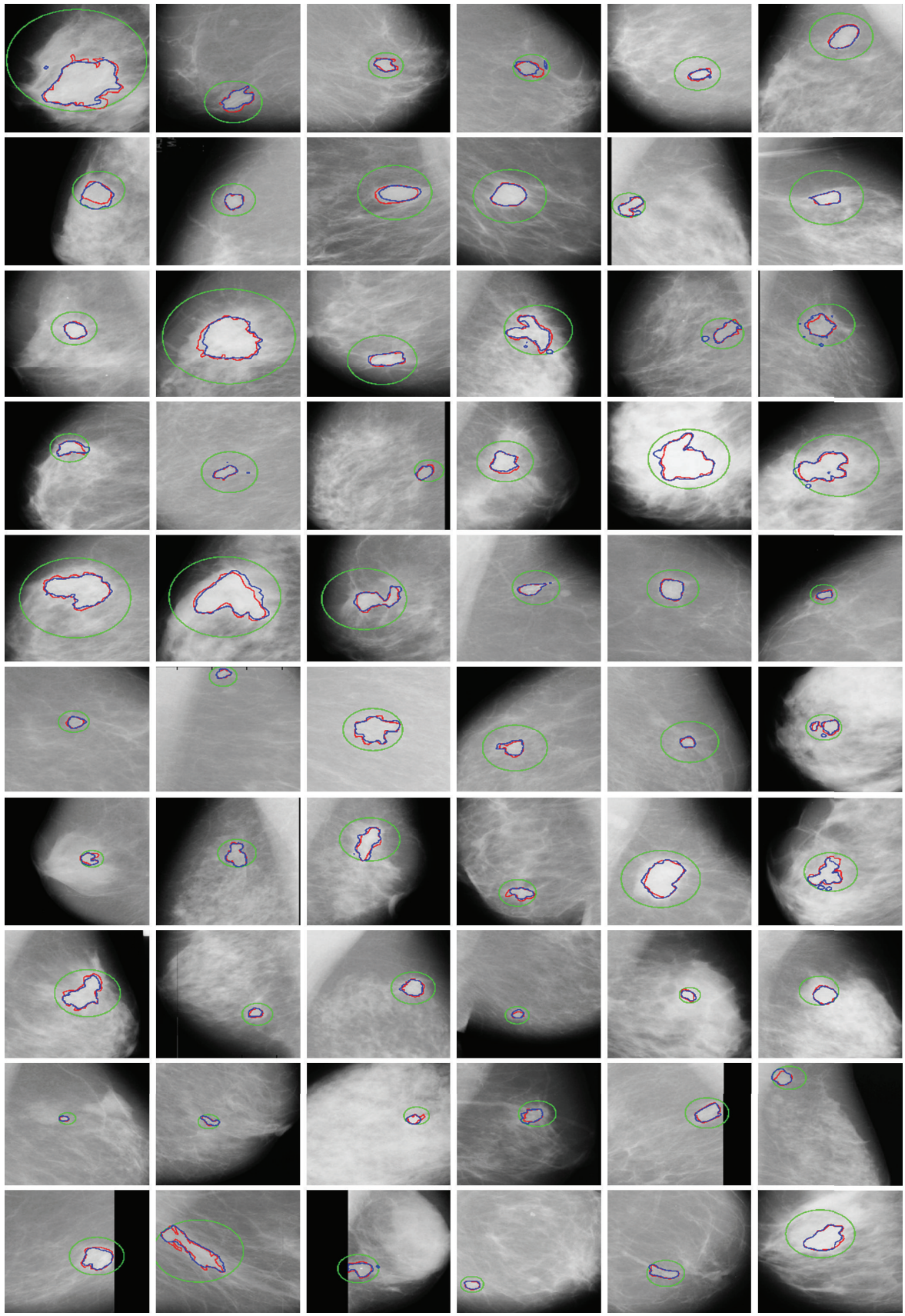


Figure 8: Examples of successful tumor tissue segmentation from mini-MIAS database. The initial contour is shown in green, the ground truth contour is shown in red, and the computed segmentation of region of interest is shown in blue.

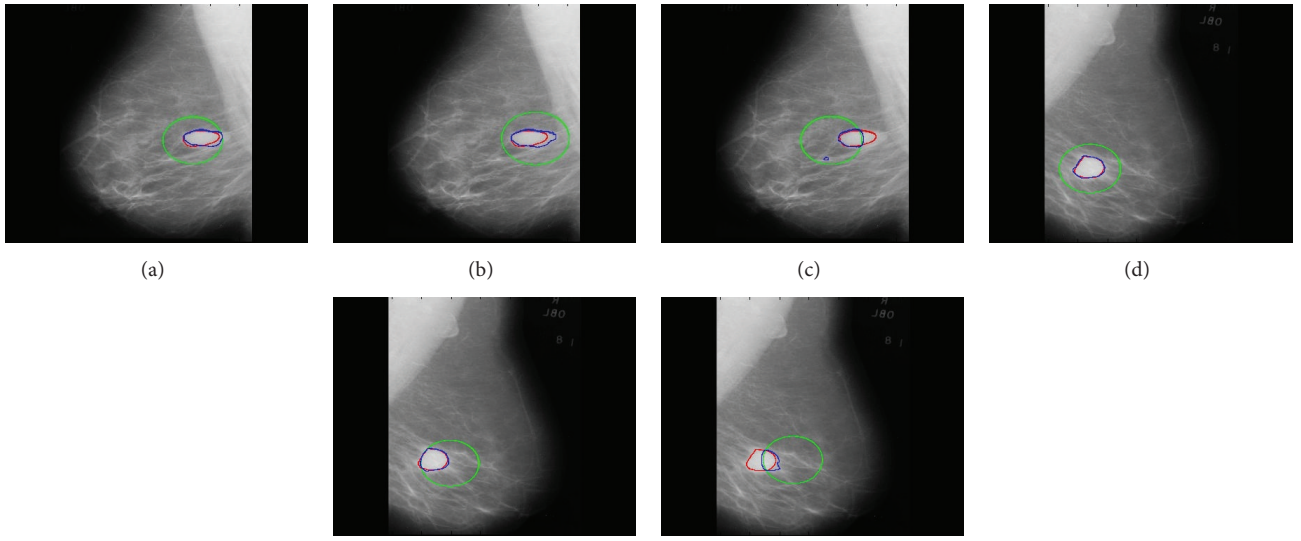


FIGURE 9: Effect of position of initial contour on segmentation of regions of interest.

Accuracy of segmentation results is not affected much if the center position of initial contour is changed but keep in mind that while changing its position it still must surround the region of interest completely. Figure 9 shows that the position of initial contour will not affect the accuracy of segmentation results if it contains the area of region of interest within it. But the segmentation accuracy will be affected if initial contour is out of region of interest or if it is partially surrounding the region of interest. The proposed method partitions an image using 2^n region method and discards all outer regions during the contour evolution process. If the initial contour surrounds partial region of interest and remaining part appears outside of it then the region of interest which lies outside the initial contour is considered useless information by the proposed algorithm and is discarded. Therefore, the proposed algorithm's accuracy is affected when initial contour does not surround whole region of interest. In Figures 9(a), 9(b), 9(d), and 9(e) the segmentation result is not affected much though the position of initial contour is changed because changed initial contour surrounded whole region of interest, whereas in Figures 9(c) and 9(f) segmented result is affected evidently because initial contour contained partial region of interest.

5. Discussion

In order to explain the advantages of the proposed method, we applied image segmentation test on four medical images shown in Figure 10. We compared the results of the proposed segmentation technique with Zhang et al. [12] and Jiang et al. [19] techniques in terms of computation time and accuracy of desired results. The parameters used for this comparison for Zhang et al. method are $\mu = 25$, $\rho = 1$, $\varepsilon = 1.5$, $\sigma = 1$, $K = 5$, and $\Delta t = 1$, while the parameters used for Jiang et al. method are $\mu = 0.04$, $\lambda = 3$, $\nu = 1$, $\rho = 2$, $\varepsilon = 1.5$, and $\tau = 5$. And the parameters used for the proposed method are same as mentioned in the Section 4.

Figure 10 shows a visual based comparison between the segmentation results of the proposed algorithm with Zhang et al. and Jiang et al. algorithms. Figures 10(a)–10(d) show the segmentation results produced by the proposed algorithm, Figures 10(e)–10(h) show the segmentation results produced by Zhang et al. algorithm, and Figures 10(i)–10(l) show the segmentation result computed by Jiang et al. algorithm. In Figure 10, four images from different image modalities are used with tumor tissues as regions of interests. We can see from the visual results the proposed algorithm provided well segmentation results of the desired region of interests in comparison with their respective ground truth, while the Zhang et el. and Jiang et al. methods which are in particular general segmentation method could not deliver well segmentation results of regions of interest as compared to their respective ground truths. Table 1 shows a quantitative analysis based on the computed segmentation results from Figure 10. First we computed false positive (fp), true positive (tp), false negative (fn), and true negative (tn) from segmented object and its ground truth analysis. In order to compare the accuracy of the segmentation we used the following statistical relationships:

$$\text{Precision} = \frac{\text{tp}}{\text{tp} + \text{fp}}, \qquad \text{Recall} = \frac{\text{tp}}{\text{tp} + \text{fn}},$$
$$\text{True Negative Rate} = \frac{\text{tn}}{\text{tn} + \text{fp}}, \tag{35}$$
$$\text{Accuracy} = \frac{\text{tp} + \text{tn}}{\text{tp} + \text{tn} + \text{fp} + \text{fn}},$$

The term precision, which shows how much of the segmented region is same as ground truth, in terms of true positive, is the most important among all. The term

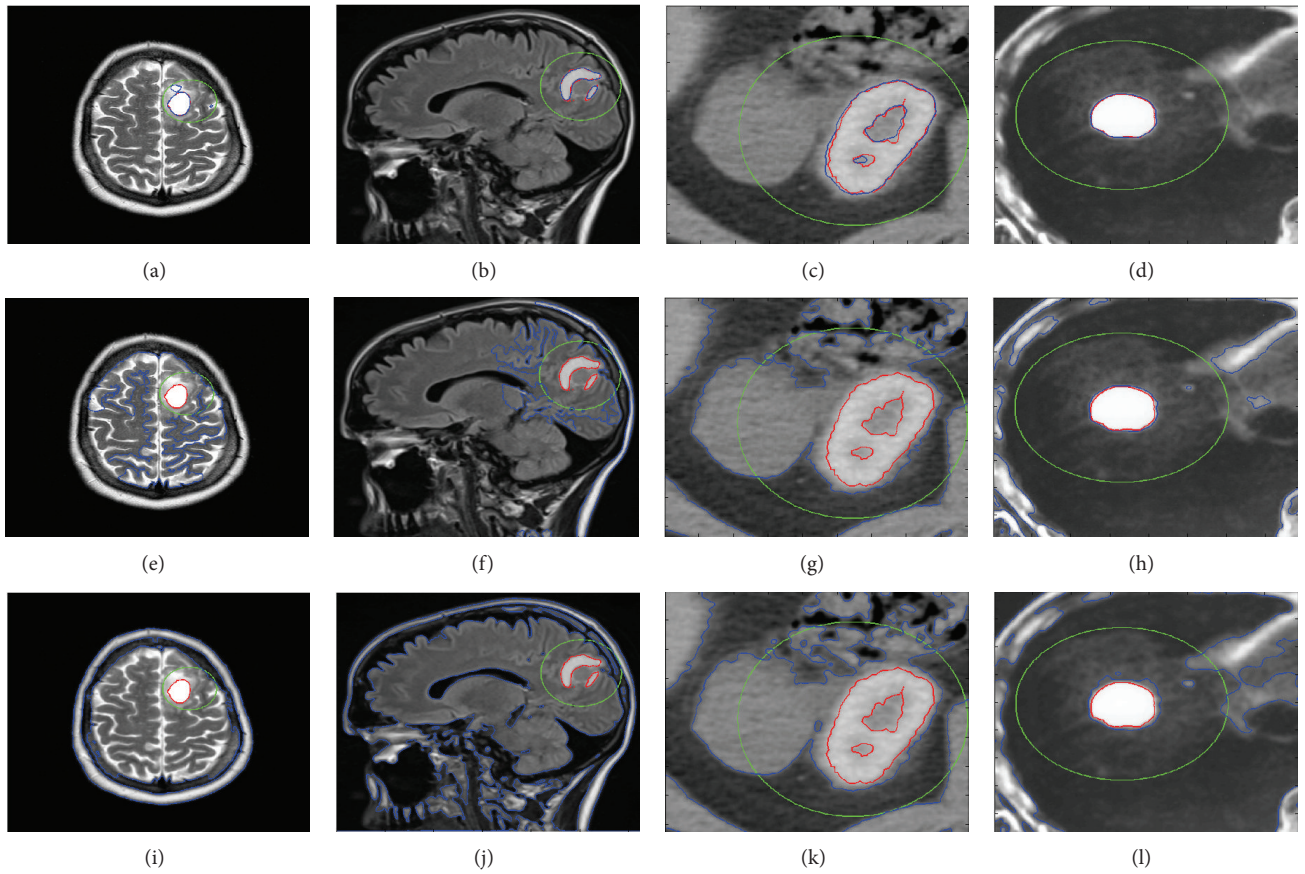


Figure 10: The visual results comparison between segmentation results of the proposed Zhang et al. and Jiang et al. algorithms. (a)–(d) shows initial contour in green, final segmentation results of four different images using proposed method in blue and ground truth in red. (e)–(h) shows initial contour in green, final segmentation results of four different images using Zhang et al. method in blue and ground truth in red. (i)–(l) shows initial contour in green, final segmentation results of four different images using Jiang et al. method in blue and ground truth in red.

recall shows whether the information that we obtain contains the information what we need, without considering false positives. The negative rate term tells us how much of true negative region is correctly ignored during segmentation. Accuracy term shows the accuracy of the segmentation result over whole image domain. The proposed method provides more precision rate, high true negative rate, high accuracy, and slightly less recall as compared to other methods. Table 1 shows the proposed method provided better precision and almost same recall as compared to Zhang et al. and Jiang et al. methods.

6. Conclusion

In this paper a region-based image segmentation technique is proposed in order to detect and segment regions of interest in medical image methodologies. A region of interest corresponds to distinctive areas that may include tumor, blood rupture area, breast boundary, masses, and other dense tissue regions.

The proposed segmentation algorithm is designed to partition an image into an arbitrary number of subregions. It starts by dividing the input image into two subregions and then one of the subregions is further divided into two subregions and so on until the stopping condition is fulfilled. It is implemented with the level set method proposed by Li et al., replacing the edge indicator function with a new region-based SPF function developed from the Chan-Vese (CV) energy functional model, which efficiently stops the contours at weak or blurred edges. A Gaussian kernel is used which not only regularizes the level set function but also eliminates the need of computationally expensive reinitialization.

Contour maps can also be produced using the proposed algorithm and with the help of contour maps we can built contour trees. These tree structures can help us to analyze the topological and geometrical relationships between different contours. With the help of contour tree we can analyze the regions of interest, and we can compute the saliency of salient regions using minimum nesting depth.

The experimental results show that the segmentation of the regions of interest in an image is highly dependent on the intensity, brightness and contrast of the objects and background. The greater the intensity difference between the salient object and the rest of the image is, the more accurate the segmentation results will be because intensity based active

Table 1: Quantitative analysis based on Figure 10.

Method	Figure	Precision	Recall	True negative rate	Accuracy	CPU time (s)	Number of iterations
Proposed method	Brain tumor 1	0.8536	0.9799	0.9991	0.9990	28.2500	42
	Brain tumor 2	0.9511	0.9326	0.9996	0.9991	32.8125	36
	Lung cancer	0.9113	0.9777	0.9901	0.9889	17.8750	39
	Liver cancer	0.9579	0.9979	0.9986	0.9986	13.3906	31
Zhang et al.	Brain tumor 1	0.0417	1	0.8773	0.8779	16.7188	100
	Brain tumor 2	0.0635	1	0.8778	0.8788	22.7656	100
	Lung cancer	0.1683	1	0.4866	0.5350	13.7813	100
	Liver cancer	0.2494	1	0.9047	0.9076	12.4688	100
Jiang et al.	Brain tumor 1	0.0183	1	0.7128	0.7143	50.9844	100
	Brain tumor 2	0.0172	1	0.5270	0.5309	68.2344	100
	Lung cancer	0.1758	1	0.5130	0.5588	32.3700	100
	Liver cancer	0.1473	1	0.8166	0.8223	30.7031	100

contour moves relatively slow on high intensity regions as compared to the low intensity regions.

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2010-0025512).

References

[1] R. Acharya, R. Wasserman, J. Stevens, and C. Hinojosa, "Biomedical imaging modalities: a tutorial," *Computerized Medical Imaging & Graphics*, vol. 19, no. 1, pp. 3–25, 1995.

[2] B.-W. Hong and B.-S. Sohn, "Segmentation of regions of interest in mammograms in a topographic approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 1, pp. 129–139, 2010.

[3] A. P. M. Forrest and R. J. Aitken, "Mammography screening for breast cancer," *Annual Review of Medicine*, vol. 41, pp. 117–132, 1990.

[4] S. M. Astley, C. R. M. Boggis, K. Walker et al., "An evaluation of a commercial prompting system in a busy screening centre," in *Digital Mammography: IWDM 2002—6th International Workshop on Digital Mammography*, pp. 471–475, Springer, Berlin, Germany, 2003.

[5] S. M. Astley, R. Zwiggelaar, C. Wolstenholme, and C. J. Taylor, "Prompting in mammography: how accurate must prompt generators be?" in *Proceedings of the 4th International Workshop on Digital Mammography*, pp. 347–354, 1998.

[6] Z. Huo, M. L. Giger, and C. J. Vyborny, "Computerized analysis of multiple-mammographic views: potential usefulness of special view mammograms in computer-aided diagnosis," *IEEE Transactions on Medical Imaging*, vol. 20, no. 12, pp. 1285–1292, 2001.

[7] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[8] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[9] C. Li, C. Xu, C. Gui, and M. D. Fox, "Level set evolution without re-initialization: a new variational formulation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 430–436, San Diego, Calif, USA, June 2005.

[10] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *International Journal of Computer Vision*, vol. 46, no. 3, pp. 223–247, 2002.

[11] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[12] K. Zhang, L. Zhang, H. Song, and W. Zhou, "Active contours with selective local or global segmentation: a new formulation and level set method," *Image and Vision Computing*, vol. 28, no. 4, pp. 668–676, 2010.

[13] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[14] J. Lie, M. Lysaker, and X.-C. Tai, "A binary level set model and some applications to Mumford-Shah image segmentation," *IEEE Transactions on Image Processing*, vol. 15, no. 5, pp. 1171–1181, 2006.

[15] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.

[16] D. Cremers, "A multiphase level set framework for motion segmentation," in *Scale Space Methods in Computer Vision: 4th International Conference, Scale Space 2003 Isle of Skye, UK, June 10–12, 2003 Proceedings*, vol. 2695 of *Lecture Notes in Computer Science*, pp. 599–614, Springer, Berlin, Germany, 2003.

[17] R. Ronfard, "Region-based strategies for active contour models," *International Journal of Computer Vision*, vol. 13, no. 2, pp. 229–251, 1994.

[18] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, "Implicit active contours driven by local binary fitting energy," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, pp. 1–7, Washington, DC, USA, June 2007.

[19] H. Jiang, R. Feng, and X. Gao, "Level set based on signed pressure force function and its application in liver image segmentation," *Wuhan University Journal of Natural Sciences*, vol. 16, no. 3, pp. 265–270, 2011.

[20] J. Gomes and O. Faugeras, "Reconciling distance functions and level sets," *Journal of Visual Communication and Image Representation*, vol. 11, no. 2, pp. 209–223, 2000.

[21] L. C. Evans, *Partial Differential Equations , Graduate Studies in Mathematics*, vol. 19, American Mathematical Society, Providence, RI, USA, 1998.

[22] C. Xu, A. Yezzi Jr., and J. L. Prince, "On the relationship between parametric and geometric active contours," in *Proceedings of the 34th Asilomar Conference on Signals, Systems and Computers*, pp. 483–489, November 2000.

[23] F. Akram, J. H. Kim, and K. N. Choi, "Active contour method with locally computed signed pressure force function: an application to brain MR image segmentation," in *Proceedings of the 7th International Conference on Image and Graphics (ICIG '13)*, pp. 154–159, July 2013.

[24] F. Akram, J. H. Kim, H. U. Lim, and K. N. Choi, "Segmentation of intensity inhomogeneous brain MR images using active contours," *Computational and Mathematical Methods in Medicine*, vol. 2014, Article ID 194614, 14 pages, 2014.

[25] H. Carr, J. Snoeyink, and U. Axen, "Computing contour trees in all dimensions," *Computational Geometry*, vol. 24, no. 2, pp. 75–94, 2003.

[26] The mini-MIAS database of mammograms, 2003, http://peipa.essex.ac.uk/info/mias.html.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 419826, 8 pages
http://dx.doi.org/10.1155/2015/419826

Research Article

Identification of Mitral Annulus Hinge Point Based on Local Context Feature and Additive SVM Classifier

Jianming Zhang,[1] Yangchun Liu,[1] and Wei Xu[2]

[1]*School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha 410114, China*
[2]*School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

Correspondence should be addressed to Jianming Zhang; jmzhang@csust.edu.cn

Received 19 September 2014; Accepted 12 February 2015

Academic Editor: Yi Gao

The position of the hinge point of mitral annulus (MA) is important for segmentation, modeling and multimodalities registration of cardiac structures. The main difficulties in identifying the hinge point of MA are the inherent noisy, low resolution of echocardiography, and so on. This work aims to automatically detect the hinge point of MA by combining local context feature with additive support vector machines (SVM) classifier. The innovations are as follows: (1) designing a local context feature for MA in cardiac ultrasound image; (2) applying the additive kernel SVM classifier to identify the candidates of the hinge point of MA; (3) designing a weighted density field of candidates which represents the blocks of candidates; and (4) estimating an adaptive threshold on the weighted density field to get the position of the hinge point of MA and exclude the error from SVM classifier. The proposed algorithm is tested on echocardiographic four-chamber image sequence of 10 pediatric patients. Compared with the manual selected hinge points of MA which are selected by professional doctors, the mean error is in 0.96 ± 1.04 mm. Additive SVM classifier can fast and accurately identify the MA hinge point.

1. Introduction

Congenital heart disease is one of the main reasons for death in children. The 3D shape and movement of the mitral apparatus are significant to analyze the function of left ventricular, diagnose mitral valve disease, and identify disorder of left ventricular [1–4]. There are some problems in surgical planning of mitral valve disease, including best surgical time and how to shape mitral valve. Therefore, analyzing the movement and shape of mitral apparatus with advanced computer technology and imaging technology has important clinical value and social value.

Precise positioning the hinge point of mitral value is helpful for modeling, motion tracking, and multimodalities registration of cardiac images. The vague and incomplete ventricular structure in ultrasound images due to the heavy noise, low resolution, and limited imaging range in real time echocardiography causes great difficulties to identify the mitral value manually and automatically. Nevo et al. [5] introduces an automated tracking algorithm using multidimensional dynamic programming which tracks the hinge point of mitral value leaflet in 2D echocardiographic images. Takemoto et al. [6] introduces an automated mitral annular tracking method. It adopts a partial shape constraint contour model to track and fit the ambiguous ventricular boundary and recognizes hinge point of mitral value using pattern matching algorithm. Veronesi et al. [7] tracks the mitral annular in 4D echocardiographic images combining optical flow method with block matching method. Due to the cumulative errors, the results require manual correction. Schneider et al. [8] use constrained optical flow combined with graph cut [9] and a valve state predictor to segment mitral annulus from four-dimensional ultrasound images.

This paper introduces a hinge point of mitral annulus identification algorithm based on additive SVM classifier [10, 11]. The optimized additive SVM classifier which can more quickly and efficiently classify test sample gets the same classification accuracy as classic SVM classifier. It is

difficult to design a feature for echocardiographic image because of the heavy noise and low resolution. The typical global feature such as local binary pattern (LBP) [12, 13] and histogram cannot specify the hinge point in the whole cardiac structure. Spatial relationships of atrium and ventricle are fixed in the echocardiography. Therefore, in this work, a local context feature is obtained for subsequent classification of MA hinge point candidates for intracardiac structures in echocardiography. Ideally, every pixel in the neighborhood can be put into context. However, this would generate a large feature space. This paper designs a local context feature which sparely samples the gray value of the context locations on eight directions in 45-degree intervals. Reasonable results will be achieved after applying the additive SVM classifier on this local context feature. The rest of this paper is organized as follows. Section 2 introduces additive SVM classifier. In Section 3, we describe the local context feature. Section 4 gives an improved method. In Section 5, we present the flow of classification. Section 6 shows some experimental results that demonstrate the effectiveness of this algorithm. Finally, we conclude in Section 7.

2. Additive SVM Classifier

SVM and boosted decision tree [14, 15] are the two main methods in target detection and multitargets recognition. Classifiers based on boosted decision trees have faster classification speed, but they are significantly slower to train. Furthermore, the complexity of training grows exponentially with the number of classes. The linear SVM is efficient during training on a given feature space. It can be used in real-time applications for low memory requirements and fast classification speed. Although the kernel trick is introduced to handle nonlinear problems in SVM, its complexity is much higher than linear SVM.

The linear SVM is more efficient, but many nonlinear kernels can get better results in pattern classification tasks due to the nonlinear distribution of features. Some popular nonlinear kernels which are based on histograms of low-level features like color and texture of the image use a kernel derived from histogram intersection or chi-squared distance to train a SVM classifier. To evaluate the classification function, the test histogram is compared with every support vector histogram. Maji et al. [10, 11] present a method which can efficiently compute the classification function based on histograms. This optimized method improves additive kernel SVMs significantly and can be used in any additive kernels.

2.1. Histogram Intersection Kernel SVMs. Given training set $\{(y_i, \mathbf{x}_i)\}_{i=1}^N$, class label $y_i \in \{-1, 1\}$; vector $\mathbf{x}_i \in \mathbf{R}^n$. To find the hyper plane to separate the sample set in linear problem, the minimization function can be written as

$$\phi(\mathbf{w}, \xi) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i, \tag{1}$$

where

$$y_i \times (\mathbf{w} \cdot \mathbf{x}_i) + b \geq 1 - \xi_i, \quad i = 1, 2, \ldots, N, \tag{2}$$

and $\xi_i \geq 0$ is the slack variable, $(1/2)\|\mathbf{w}\|^2$ is used to maximize the distance between support vectors and hyper plane, $C\sum_{i=1}^{N}\xi_i$ is used to minimize the error rate, $C > 0$ is the weight between maximized distance and slack constraints, $\mathbf{w}$ is the normal vector to the hyperplane, and b determines the offset of the hyperplane from the origin along the normal vector $\mathbf{w}$. The kernel $K(x, z) : R^n \times R^n \rightarrow R$ is the inner product $\varphi(\mathbf{x})\cdot\varphi(\mathbf{z})$ in high dimension. The maximized duality function is

$$W(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j), \tag{3}$$

where α is Lagrange multiplier, $0 \leq \alpha_i \leq C$, $\sum \alpha_i y_i = 0$. Decision function is $\mathrm{sgn}(h(x))$, where

$$h(\mathbf{x}) = \sum_{l=1}^{m}\alpha_l y_l K(\mathbf{x}_l, \mathbf{x}) + b. \tag{4}$$

And $\mathbf{x}_l, l \in \{1, 2, \ldots, m\}$ is the support vector. Therefore, classifying a test feature requires m times calculation of kernel function and stores m support vectors. Assuming that the complexity of decision function is $O(n)$, the complexity of classification of one test feature is $O(m*n)$. As to linear kernel function $K(\mathbf{x}, \mathbf{z}) = \mathbf{x}\cdot\mathbf{z}$, its decision function is $h(\mathbf{x}) = \mathbf{w}\cdot\mathbf{x} + b$ where $\mathbf{w} = \sum_{l=1}^{m}\alpha_l y_l \mathbf{x}_l$. So the complexity of linear SVM is $O(n)$.

With similarity of feature such as boundary, color can be represented as histogram which regularly uses histogram intersection as its evaluation of similarity. The histogram intersection kernel is

$$K_{\min}(\mathbf{x}, \mathbf{z}) = \sum_{i}^{n}\min(x_i, z_i), \tag{5}$$

where $\mathbf{x}, \mathbf{z} \in \mathbf{R}_+^n$ is histogram feature. The decision function is

$$\begin{aligned} h(\mathbf{z}) &= \sum_{l=1}^{m}\alpha_l y_l K_{\min}(\mathbf{z}, \mathbf{x}_l) + b \\ &= \sum_{l=1}^{m}\alpha_l y_l \sum_{i=1}^{n}\min(z_i, x_{l,i}) + b. \end{aligned} \tag{6}$$

The complexity of (6) is still $O(m * n)$. The key property of intersection kernels is that the order of summing can be exchanged. So, (6) can be converted to the following:

$$\begin{aligned} h(\mathbf{z}) &= \sum_{l=1}^{m}\alpha_l y_l \sum_{i=1}^{n}\min(z_i, x_{l,i}) + b \\ &= \sum_{i=1}^{n}\left(\sum_{l=1}^{m}\alpha_l y_l \min(z_i, x_{l,i})\right) + b = \sum_{i=1}^{n}h_i(z_i) + b. \end{aligned} \tag{7}$$

Function $h(\cdot)$ can be represented as the sum of 1D functions $h_i(\cdot)$, where

$$h_i(z_i) = \sum_{l=1}^{m}\alpha_l y_l \min(z_i, x_{l,i}). \tag{8}$$

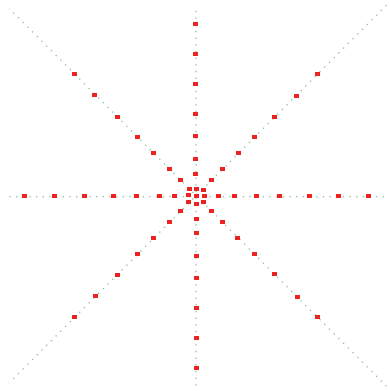


Figure 1: Local context feature.

So as to make the complexity of h_i to $O(\log m)$, let $\overline{x}_{l,i}$ denote the sorted $x_{l,i}$ in increasing order; the corresponding α_l and y_l are $\overline{\alpha}_l$ and $\overline{y}_l$. If $z_i < \overline{x}_{1,l}$, then $h_i(z_i) = z_i \sum_l \overline{\alpha}_l \overline{y}_l = 0$. Otherwise, let r be the largest integer satisfying $\overline{x}_{r,l} \le z_i$. The function $h_i(\cdot)$ is transformed to

$$\begin{aligned} h_i(z_i) &= \sum_{l=1}^{m} \overline{\alpha}_l \overline{y}_l \min(z_i, x_{l,i}) \\ &= \sum_{1 \le l \le r} \overline{\alpha}_l \overline{y}_l x_{l,i} + z_i \sum_{r < l \le m} \overline{\alpha}_l \overline{y}_l = A_i(r) + z_i B_i(r), \end{aligned} \tag{9}$$

where $A_i(r) = \sum_{1 \le l \le r} \overline{\alpha}_l \overline{y}_l x_{l,i}$, $B_i(r) = \sum_{r < l \le m} \overline{\alpha}_l \overline{y}_l$. Functions A_i and B_i which only depend on the support vectors and α can be computed after SVM model is trained. Binary search is adopted to get r and reduces the complexity.

3. Local Context Feature

The general image detection operator such as the Sobel operator and the Laplace operator cannot be applied to ultrasound images due to the heavy noise and blurred boundary. Context is the relationship with the neighbors which can be represented as a certain range of neighbors of a pixel in image processing. This paper introduces a local context feature which sparsely sample [16] the neighbor pixels of a pixel. Sampling sparsely can decrease the dimensions and shorten the training time. Because of the fixed spatial distribution of cardiac tissues in ultrasound images, the local context feature can quantify the position relationship between cardiac tissues.

As Figure 1 shows, red points represent the sampling points. This feature sampling at eight directions and the intervals between sample points largen as the distance to the center point lengthens. The points nearer to the center point contain more information about the center point, so most feature samples points are in near range. Taking the size of cardiac ultrasound image into account, the max sampling distance sets as 29 and the sampling position at each direction is $\{1, 3, 5, 9, 13, 17, 23, 29\}$. Because of the heavy noise in echocardiography, the gray value of sampling point will make training and classification features inaccurate, which will reduce the recognition rate. In order to eliminate the error caused by the noise, we apply mean filtering to the sampling points when sampling. Because the larger average template cannot make the sampling gray value accurately reflect the information of sampling points, we set the average of $3 * 3$ neighborhoods as the value of each sampling point. So one pixel can get a 65-dimension feature and the local feature is extracted fast.

4. Refine the Classification

We can get a good recognition result by adopting the local context feature and additive kernel SVM classifier, as Figure 2(a) shows. Red points are the candidate points given by the SVM classifier. The points that arrow 1 indicates get correct classification result. Due to the low resolution and heavy noise, the point that the arrow 2 indicates gets a misclassification result by SVM.

The SVM classifier is trained to get points like that arrow 1 indicates, so the majority of candidate points will be right. And density is a good feature to distinguish the right points and wrong ones. This paper applies a weighted template to each candidate point and gets a weighted density field. Figure 3 shows the weighted template obtained by block distance.

The density field function is

$$F(A) = \sum_{y \in C \cap N_{11}(A)} \left[10 - D_{\text{block}}(y, A)\right], \tag{10}$$

where A is any point in density field, C is the set of candidate points, $D_{\text{block}}(\cdot)$ is the block distance, $N_{11}(A)$ is the 11 neighborhoods of a point A. As Figure 2(c) shows, the right points are highlight in the density field and wrong points are dim. We can determine an adaptive threshold to exclude the dim points by binary search between the max and min values of density field. The following part shows the flow of the algorithm.

The Flow of Adaptive Thresholding. H = the max value in density field.

(1) If the number of continuous areas greater than H is exact two, go to step (4).

(2) If the number of areas is less than two, decrease H and go to step (1).

(3) If the number of areas is greater than two, increase H and go to step (1).

(4) Get the adaptive threshold H.

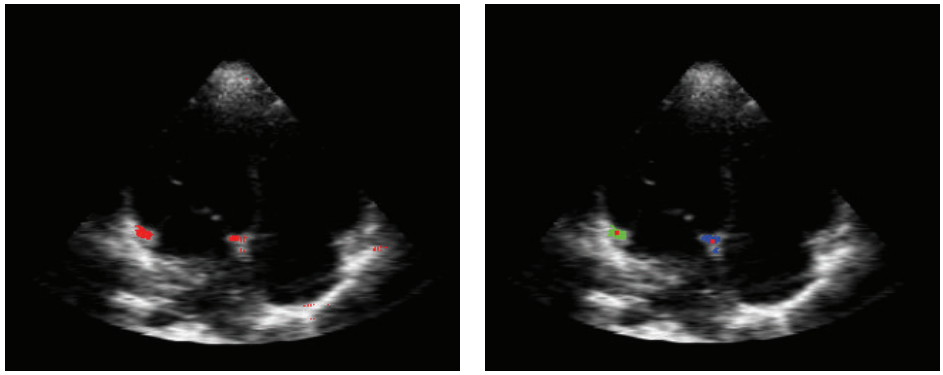
Figure 2(c) shows the result adopting adaptive threshold; we can see that it is so easy to separate the two blocks. Then the x and y average values of the two blocks can be calculated separately as the position of hinge point of mitral annulus. Another problem is that the adaptive threshold may also exclude some right points which lead to an unreliable result. In order to get an accurate result, we propose to use the two blocks centers as the initial clustering center of K-means to classify the whole candidate points in a certain scale. Due to the fixed size of mitral annulus, the certain scale can be

(a) Result of SVM classifier (b) Weighted density filed (c) Result of adaptive threshold

Figure 2: Classification result of additive SVM.

0								0
	...						...	
		6	7	8	7	6		
		7	8	9	8	7		
		8	9	10	9	8		
		7	8	9	8	7		
		6	7	8	7	6		
	...						...	
0								0

Figure 3: The weighted template obtained by block distance.



(a) Result of additive kernel SVM (b) Refined result using K-means

Figure 4: Refined classification result of K-means classifier.

obtained from experiments. The result is shown in Figure 4, the green and blue block are the results of K-means classifier, and the center red points in the two blocks are the accurate position of hinge point of mitral annulus.

5. Classification Flow

The flow of classification procedure can be integrated into three layers shown in Figure 5.

6. Experiments

The image data in the paper is from Sonos 7500 ultrasound image and the size of raw 3D image is $208 \times 160 \times 144$. Data is acquired from 10 children who are from 9 to 12 years old. The cardiac cycle has 9 to 24 frames. Experiment 1 shows how to confirm the sampling window of local context feature and how to choose the size of weighted template. Experiment 2 compares the $3 * 3$ average sampling and

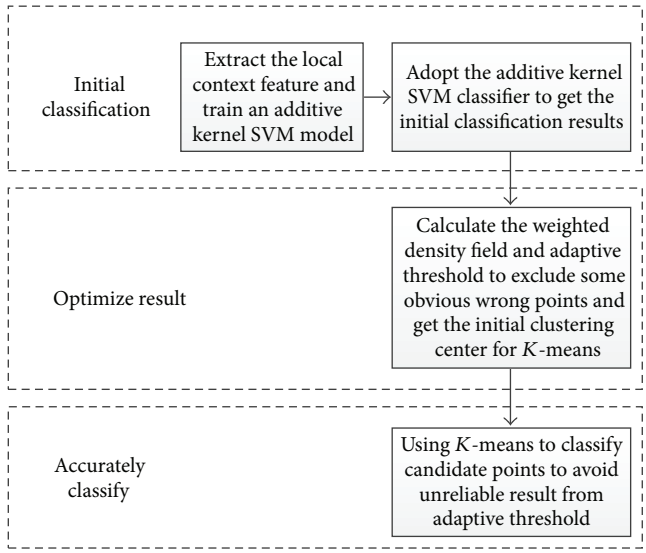


Figure 5: Flow of classification.

Table 1: Errors between our segmentation method and manual segmentation results.

	Septal				Lateral			
	x-axis		y-axis		x-axis		y-axis	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
mm	0.96	0.907	1.12	0.69	1.34	1.39	0.75	0.48

direct sampling result in the heavy noise ultrasound images. Experiment 3 compares the result of this algorithm and manual selected points which are selected by professional doctors. Experiment 4 shows various results using different kernel functions.

6.1. Experiment 1. Local context feature is to sample local structure of heart tissue. The specific size of sampling window and the size of weighted template are from experiments. And the experiment results show that different size has little influence to the results. Figure 6(a) is the result of parameters as $\{1, 3, 5, 9, 13, 17, 23, 29\}$ in local context feature and Figure 6(b) is the result of parameters as $\{1, 2, 4, 8, 12, 15, 22\}$. Figure 6(c) is the result of weighted template whose size is set to 10 and Figure 6(d) is the result of weighted template whose size is set to 8.

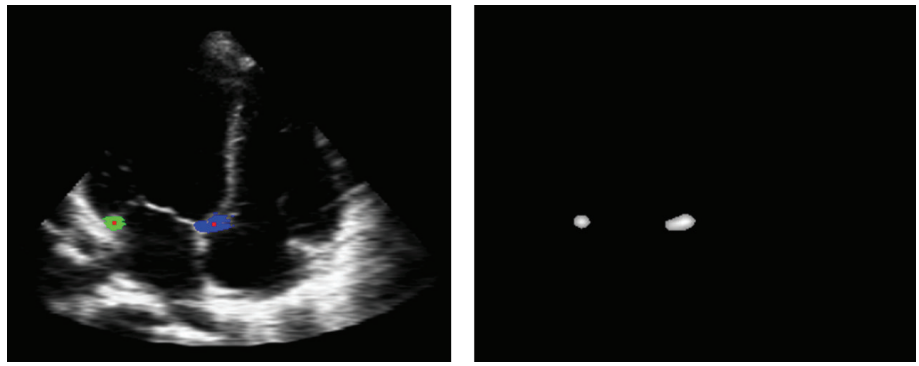
6.2. Experiment 2. Figure 7(a) shows the result of direct sampling method, and Figure 7(b) shows the result of $3 * 3$ average sampling. The experiment shows that the result of $3 * 3$ average sampling is better than direct sampling. Additionally, we can see that heavy noise and low resolution in ultrasound image affect the result seriously. This denotes that the regular image detection methods cannot be used in medical images. HOG or SIFT feature which detect the corner points also cannot be used in medical images especially ultrasound images.

6.3. Experiment 3. Table 1 shows the mean and variance of our results compared to manual select points. This result is obtained from cardiac cycle of 10 patients. Our mean error can be control in nearly 0.96 mm which is acceptable in medical diagnosis.

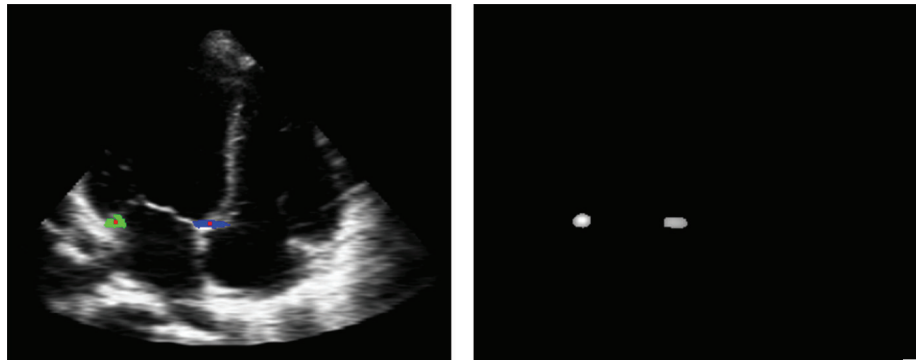
6.4. Experiment 4. Figure 8 shows various results using different kernel functions. In this figure, we can conclude that the choice of kernel function is very important for final result. Furthermore, we can see that intersection histogram kernel which matches the local context feature gets a more accurate classification result.

7. Conclusions

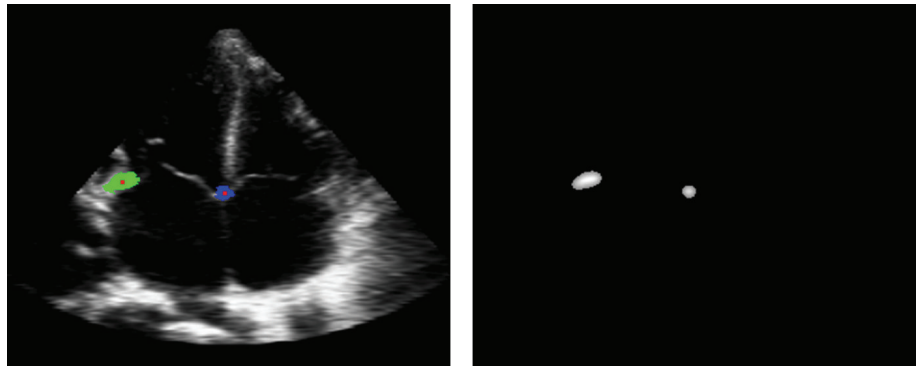
This paper introduced a hinge point of mitral annulus identification method using additive kernel SVM classifier and local context feature. Due to the classification errors, we design a weighted template to exclude the obvious wrong points. After refining the result, the mean and variance of error between automatic and manual result are controlled in 0.96 ± 1.04 mm. From the experiments, it is demonstrated that this algorithm can accurately locate the hinge point of mitral valve. For the fast feature extraction and accelerated classification procedure, this algorithm can be used in real-time applications.



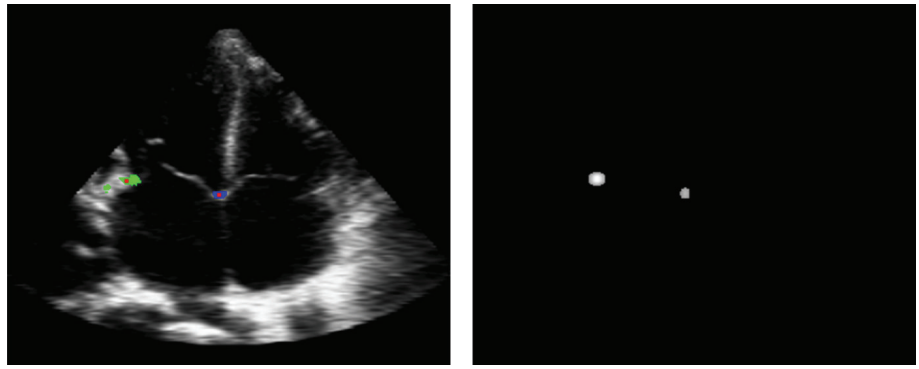
(a) Identification result



(b) Identification result

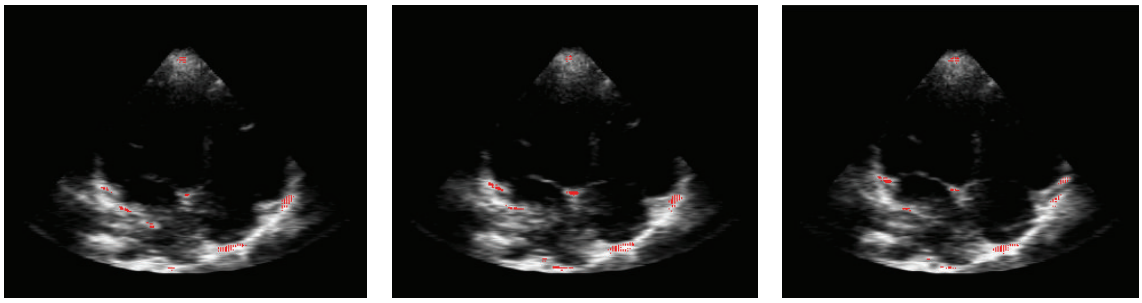


(c) Identification result

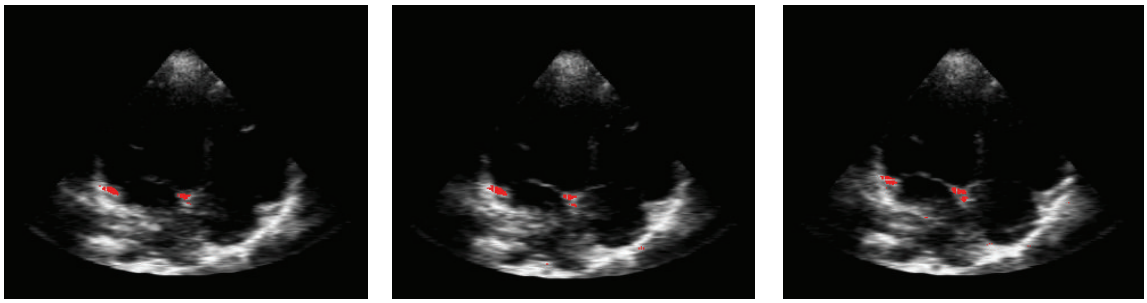


(d) Identification result

Figure 6: Identification results by different parameters.

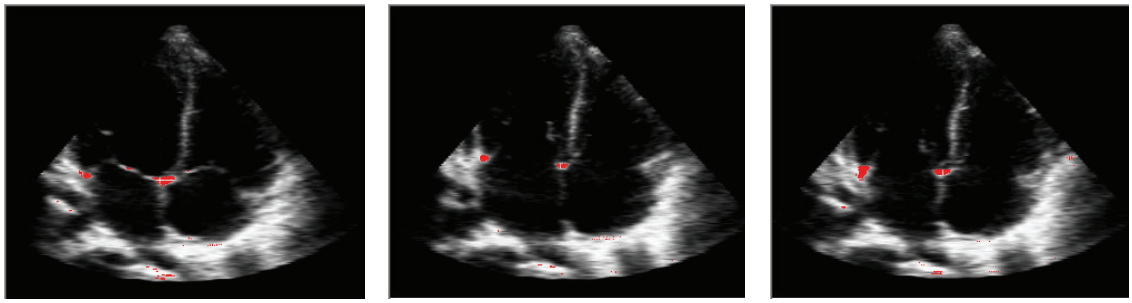


(a) Result of direct sampling

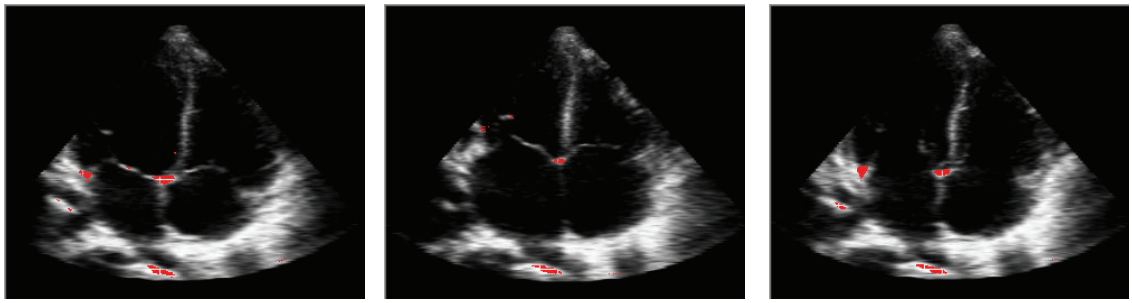


(b) Result of average sampling

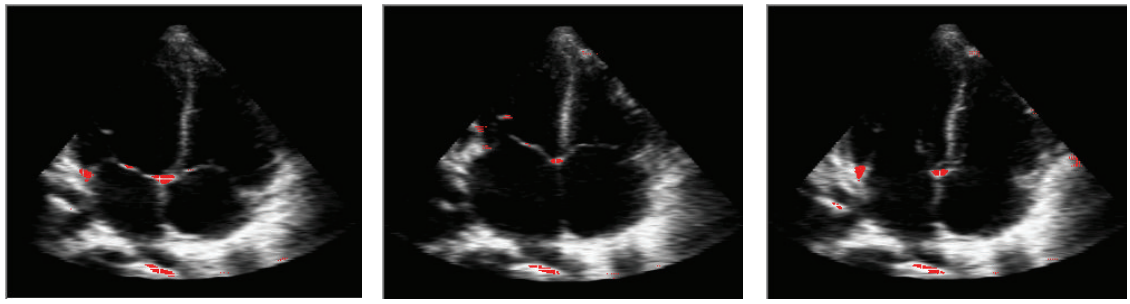
Figure 7: The classification result of different sample mode.



(a) Classification result using intersection kernel function



(b) Classification result using chi-square kernel function



(c) Classification result using the Jensen-Shannon kernel function

Figure 8: The classification result of different kernel function.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The work in this paper was partly funded by the Scientific Research Fund of Hunan Provincial Education Department (Grant No. 12B003), the Scientific Research Fund of Hunan Provincial Transportation Department (Grant No. 201334).

References

[1] S. J. Buss, D. Mereles, M. Emami et al., "Rapid assessment of longitudinal systolic left ventricular function using speckle tracking of the mitral annulus," *Clinical Research in Cardiology*, vol. 101, no. 4, pp. 273–280, 2012.

[2] N. P. Nikitin, P. H. Loh, R. de Silva et al., "Prognostic value of systolic mitral annular velocity measured with Doppler tissue imaging in patients with chronic heart failure caused by left ventricular systolic dysfunction," *Heart*, vol. 92, no. 6, pp. 775–779, 2006.

[3] J. Matos, I. Kronzon, G. Panagopoulos, and G. Perk, "Mitral annular plane systolic excursion as a surrogate for left ventricular ejection fraction," *Journal of the American Society of Echocardiography*, vol. 25, no. 9, pp. 969–974, 2012.

[4] W. Tsang, H. Ahmad, A. R. Patel et al., "Rapid estimation of left ventricular function using echocardiographic speckle-tracking of mitral annular displacement," *Journal of the American Society of Echocardiography*, vol. 23, no. 5, pp. 511–515, 2010.

[5] S. T. Nevo, M. van Stralen, A. M. Vossepoel et al., "Automated tracking of the mitral valve annulus motion in apical echocardiographic images using multidimensional dynamic programming," *Ultrasound in Medicine and Biology*, vol. 33, no. 9, pp. 1389–1399, 2007.

[6] Y. Takemoto, T. Hozumi, K. Sugioka et al., "Automated three-dimensional analysis of mitral annular dynamics in patients with myocardial infarction using automated mitral annular tracking method," *Echocardiography*, vol. 23, no. 8, pp. 658–665, 2006.

[7] F. Veronesi, C. Corsi, E. G. Caiani et al., "Semi-automatic tracking for mitral annulus dynamic analysis using real-time 3D echocardiography," in *Proceedings of the Computers in Cardiology (CIC '06)*, vol. 33, pp. 113–116, September 2006.

[8] R. J. Schneider, D. P. Perrin, N. V. Vasilyev, G. R. Marx, P. J. Del Nido, and R. D. Howe, "Mitral annulus segmentation from four-dimensional ultrasound using a valve state predictor and constrained optical flow," *Medical Image Analysis*, vol. 16, no. 2, pp. 497–504, 2012.

[9] R. J. Schneider, D. P. Perrin, N. V. Vasilyev, G. R. Marx, P. J. del Nido, and R. D. Howe, "Mitral annulus segmentation from 3D ultrasound using graph cuts," *IEEE Transactions on Medical Imaging*, vol. 29, no. 9, pp. 1676–1687, 2010.

[10] S. Maji, A. C. Berg, and J. Malik, "Efficient classification for additive kernel SVMs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 66–77, 2013.

[11] S. Maji, A. C. Berg, and J. Maliks, "Classification using intersection kernel support vector machines is efficient," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, June 2008.

[12] D.-C. He and L. Wang, "Texture features based on texture spectrum," *Pattern Recognition*, vol. 24, no. 5, pp. 391–399, 1991.

[13] M. Heikkilä, M. Pietikäinen, and J. Heikkilä, "A texture-based method for detecting moving objects," in *Proceedings of the 15th British Machine Vision Conference (BMVC '04)*, pp. 1–10, 2004.

[14] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor, "Boosted decision trees as an alternative to artificial neural networks for particle identification," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 543, no. 2-3, pp. 577–584, 2005.

[15] M. A. Kumar and M. Gopal, "A hybrid SVM based decision tree," *Pattern Recognition*, vol. 43, no. 12, pp. 3977–3987, 2010.

[16] Z. Tu and X. Bai, "Auto-context and its application to high-level vision tasks and 3D brain image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, pp. 1744–1757, 2010.

Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2015, Article ID 502573, 10 pages
http://dx.doi.org/10.1155/2015/502573

Research Article

Adaptive Ridge Point Refinement for Seeds Detection in X-Ray Coronary Angiogram

Ruoxiu Xiao,[1,2] Jian Yang,[1] Danni Ai,[1] Jingfan Fan,[1] Yue Liu,[1] Guangzhi Wang,[2] and Yongtian Wang[1]

[1] *Key Laboratory of Photoelectronic Imaging Technology and System, Ministry of Education of China, School of Optics and Electronics, Beijing Institute of Technology, Beijing 100081, China*
[2] *Department of Biomedical Engineering, School of Medicine, Tsinghua University, Room C249, Beijing 100084, China*

Correspondence should be addressed to Jian Yang; jyang@bit.edu.cn

Received 28 August 2014; Revised 15 January 2015; Accepted 16 January 2015

Academic Editor: Yi Gao

Seed point is prerequired condition for tracking based method for extracting centerline or vascular structures from the angiogram. In this paper, a novel seed point detection method for coronary artery segmentation is proposed. Vessels on the image are first enhanced according to the distribution of Hessian eigenvalue in multiscale space; consequently, centerlines of tubular vessels are also enhanced. Ridge point is extracted as candidate seed point, which is then refined according to its mathematical definition. The theoretical feasibility of this method is also proven. Finally, all the detected ridge points are checked using a self-adaptive threshold to improve the robustness of results. Clinical angiograms are used to evaluate the performance of the proposed algorithm, and the results show that the proposed algorithm can detect a large set of true seed points located on most branches of vessels. Compared with traditional seed point detection algorithms, the proposed method can detect a larger number of seed points with higher precision. Considering that the proposed method can achieve accurate seed detection without any human interaction, it can be utilized for several clinical applications, such as vessel segmentation, centerline extraction, and topological identification.

1. Introduction

Currently, vascular diseases are major threats to human health. Although a variety of imaging technologies exist, such as computed tomography angiography (CTA), magnetic resonance angiography (MRA), and ultrasound (US), X-ray angiography remains the gold standard for the interventional treatment of coronary artery diseases because of its high resolution and imaging speed. Foreshortening and overlapping are the major obstacles for the accurate identification of vascular structures because X-ray angiography is an integrated projection of the whole body in 3D space to 2D images. Vascular extraction technology aims to calculate the centerline, diameter, and direction vector of the vascular structure from X-ray angiograms; hence, it can provide the necessary reference for computer-aided diagnosis and treatment of vascular diseases.

To date, the widely used vascular extraction method in clinical practices is still the manual delineation method, which is very time-consuming and subjective. As its important clinical value, automatic vascular tree extraction method has been studied intensively in the past two decades, such are morphology based methods [1, 2], multiscale based methods [3, 4], edge detection based methods, and image registration based methods [5–7]. Among all methods, the tracking based methods propose to estimate centerline and diameter within the vascular boundaries, which do not need to scan the whole angiogram. Hence, the tracking based methods are usually with higher calculation efficiency than the other methods.

Generally, the tracking procedure proceeds from one or several manually delineated seed points. As the seed points are randomly selected from the angiogram, the reproducibility of the tracking algorithms are very much depended on the personal experience. Many researchers hence focus on improving the robustness of the tracking algorithm through seed optimization. Collorec and Coatrieux [8] detected seed points by scanning local extreme points and obtained a large set of seed points inside the vessels. However, extracted seed

points need to be refined because of the presence of noise. While Fritzsche et al. [9] combined the global threshold optimization for improving the robustness of the seed extraction, the global threshold may also lead to a large amount of false seed points in the background. Moreover, Boroujeni et al. [10] proposed an automatic seed point detection method by detecting edge points and checking the symmetric features in its neighboring regions. After all the boundaries are detected, the center line seed points can be calculated at the symmetric center of the edge points. All the above methods have greatly promoted the automatic seed detection methods.

In this paper, a novel adaptive ridge point refinement method is proposed for seed detection in coronary angiograms. First, based on the tubular feature distribution of Hessian matrix of the angiogram, vascular structures are enhanced according to the eigenvalue distribution in multiscale space. Second, the continuity property of eigenvalue and eigenvector of a Hessian matrix in multiscale space is theoretically analyzed. Third, based on theoretical analysis, the proposed theorem of ridge point existence is utilized to design the ridge discriminant function. And the candidate ridge points are extracted according to the predefined discriminant function. Afterwards, the detected points are refined according to a self-adaptive threshold that is calculated based on the order statistics of the detected ridge points.

2. Method

2.1. Characteristic of Ideal Vascular Topology. Let $I(M)$ represent the intensity of a point M in the image; then, the intensity distribution of the local feature around M can be calculated as follows [11]:

$$I(M+\delta M) = I(M) + M^T \nabla I(M) + M^T H(M) M + o(\|M\|), \tag{1}$$

where $\nabla I(M) = ((\partial/\partial x)I(M), (\partial/\partial y)I(M))$ is the gradient of the image I at the point M with respect to the x-axis and y-axis, while $(\partial/\partial x)I(M)$ and $(\partial/\partial y)I(M)$ are the first-order partial derivatives of $I(M)$ in the directions of x and y, respectively. And the Hessian matrix $H(M)$ of point M can be calculated as follows:

$$H(M) = \begin{pmatrix} \frac{\partial^2}{\partial x^2} I(M) & \frac{\partial^2}{\partial x \partial y} I(M) \\ \frac{\partial^2}{\partial y \partial x} I(M) & \frac{\partial^2}{\partial y^2} I(M) \end{pmatrix}, \tag{2}$$

where $(\partial^2/\partial x^2)I(M)$ and $(\partial^2/\partial y^2)I(M)$ denote the second partial differential of $I(M)$ in the direction of x and y, respectively, while $(\partial^2/\partial x \partial y)I(M)$ and $(\partial^2/\partial y \partial x)I(M)$ denote the second partial differential of $I(M)$. If the second-order differential of $I(M)$ is continuous, then we have $(\partial^2/\partial x \partial y)I(M) = (\partial^2/\partial y \partial x)I(M)$. Through singular value decomposition (SVD) decomposition, two eigenvalues λ_1 and λ_2 (suppose $\lambda_1 \leq \lambda_2$) and their corresponding eigenvectors v_1 and v_2 can be obtained from the Hessian matrix of each pixel of the angiogram. To be convenient for the followed analysis, in this study, λ_1 and λ_2 are named as the first eigenvalue and second eigenvalue of $H(M)$, and v_1 and v_2 are denoted by first eigenvector and second eigenvector of $H(M)$, separately.

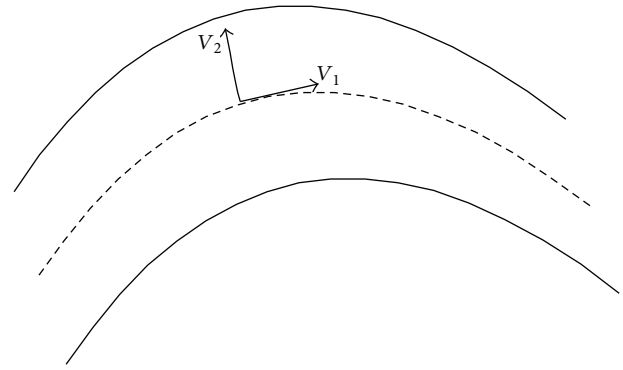


Figure 1: Relationship between the two eigenvectors of the Hessian matrix with respect to vascular topology.

Ideally, due to its tubular structure, the penetrating path of X-ray in the blood vessel decreases from the central axis to edge position. Therefore, the gray level distribution of the vessel in angiogram turns from dark to bright for its centerline to the edge [12]. If we look at the grey scale distribution of the blood vessel, the centerlines are rested on the ridge lines constituted by a series of ridge points. Generally, the ridge point is the local extreme with the direction vector perpendicular to the vascular direction on angiogram. Therefore, for the coronary artery in angiogram, we have the following equation [13]:

$$0 \approx |\lambda_1| \ll \lambda_2. \tag{3}$$

According to the definition of the ridge, the two eigenvectors of Hessian matrix can be denoted by the tangent direction (V_1) and the vertical direction (V_2), as can be seen in Figure 1.

2.2. Multiscale Vascular Enhancement. Since the disturbance of eigenvalue or eigenvector of the Hessian matrix is in the same order [14], if the disturbance of the Hessian matrix is ε, then the disturbances of the corresponding eigenvalue and eigenvector are $O(\varepsilon)$ because of the existence of noise in the angiogram. However, when zero-order disturbance is $\Delta\delta$, the second-order disturbance is amplified twice. In such condition, the use of eigenvalue or eigenvector will introduce a large amount of error for vessel enhancement. Hence, several researchers introduced multiscale operators, such as Gaussian scale transformation, to minimize error disturbance [3, 15, 16]. The Gaussian function has been proven to be the only kernel function in the linear scale space by Lindeberg [16, 17] and Florack et al. [18, 19]. According to the property of the Gaussian function, the multiscale space is not only linear but also satisfies several other properties, such as spatial shift invariance, noncreation of local extreme, rotational symmetry, and semigroup structure [20, 21].

In this paper, the Frangi filter [13] is utilized to enhanced vascular structure in the angiogram. For this filter, a single scale function can be defined as follows:

$$E(M,\sigma) = \begin{cases} 0 & \text{if } \lambda_2 < 0 \\ \exp\left(-\dfrac{\lambda_1^2}{2\alpha^2\lambda_2^2}\right)\left[1-\exp\left(-\dfrac{\lambda_1^2+\lambda_2^2}{2\beta^2}\right)\right] & \\ & \text{otherwise,} \end{cases} \tag{4}$$

where λ_1 and λ_2 are the first and the second eigenvalues, while α and β are controlling coefficient. To obtain the best vascular enhancement effect, the multiscale function should have the highest response of all the sampled scales, which can be formulized as follows:

$$E(M) = \max_{\sigma_{\min}\le\sigma\le\sigma_{\max}} E(M,\sigma), \tag{5}$$

where $[\sigma_{\min}, \sigma_{\max}]$ is the predefined scaling range. Typically, $\sigma_{\min}$ and $\sigma_{\max}$ correspond to minimum and maximum size of vessels on image.

The image enhanced by multiscale eigenvalue of Hessian matrix is usually referred to as a vesselness image. In a vesselness image, the background of the nonvascular region is suppressed, and the vessels appear brighter than that of the original image. Moreover, the pixels along the directions of vascular centerlines are strongly enhanced and appear brighter than the ones perpendicular to the vascular direction vector. Therefore, the ridge points on a vesselness image can be extracted to stand for the seed points from the extraction of blood vessels.

2.3. Continuity Analysis of Vesselness Image. In this study, the candidate seed points are extracted by detecting ridge points in a vesselness image based on the differential continuity of the image, which includes gradient and Hessian matrix together with its corresponding eigenvalue and eigenvector for each pixel. Usually, they are not located at integer coordinates and should be computed by interpolation. However, they should be continuous when being interpolated. Therefore, the continuity of the abovementioned differential information is very important for seed detection in the angiogram.

In traditional methods of differential information analysis, the differential information of the image is assumed to be continuous. Theoretically, an image can be described as a 2D continuous signal obtained from the optical sensor by light integration. As such, the zero-order gray scale information is a continuous function for the coordinates. To further utilize the differential information, this study theoretically analyzed the continuous property of gradient, eigenvalue, and eigenvector.

Suppose that $f(x, y; \sigma)$ represents an image convolved by a Gaussian function with kernel of σ; then we have the following equation:

$$\frac{\partial}{\partial x} f(x,y;\sigma) \otimes g(x,y,\sigma) = f(x,y;\sigma) \otimes \frac{\partial}{\partial x} g(x,y,\sigma), \tag{6}$$

since

$$\frac{\partial}{\partial x} g(x,y,\sigma) = -\frac{x}{\sigma^2} g(x,y,\sigma). \tag{7}$$

Then, the continuity of an image in Gaussian space will be transformed into the continuity of the image after Gaussian convolution. Therefore, we first discuss the continuity of the image after Gaussian smoothing in the following section.

According to Lemma A.1 (as can be seen in the section of appendix), we found that any 1D continuous function will be infinitely differentiable when it is convoluted with the Gaussian function. Similarly, the vesselness image, which is a two-dimensional continuous function, will be infinitely differentiable with the application of Gaussian smoothing. Essentially, if the vesselness image is at least differentiable in the second-order, its gradient vector and Hessian matrix are also continuous. As can be inferred from Lemma A.3, the two eigenvalues of Hessian matrix of vesselness image are single eigenvalue and the corresponding eigenvectors of these two eigenvalues are continuous according to Lemma A.2. Therefore, all the related differential terms, including gradient and Hessian matrix together with its corresponding eigenvalue and eigenvector of vesselness image, are continuous.

2.4. Seed Point Detection. In traditional methods, the detection of local maxima points in the image is a common method of seeding. However, the points on the centerline are essentially not the local maxima points in any direction. Instead, they are the local maxima points on the directions perpendicular to the centerline. In this paper, the ridge points are extracted as candidates of seed points. The definition of ridge point could be derived from the definition of local extreme point, which can be described as follows [13].

Definition 1. Let $f : \mathbb{R}^n \to \mathbb{R}$ represent a second-order continuous function. A point $x \in \mathbb{R}^n$ is a local extreme point for f if $(v\cdot\nabla) f(x) = 0$ for every direction v; that is, $\nabla f(x) = 0$. The extreme point can be classified as follows: (1) x is a local minimum point, if $(v \cdot \nabla)^2 f(x) > 0$ for every direction v; (2) x is a local maximum point, if $(v \cdot \nabla)^2 f(x) < 0$ for every direction v. The corresponding function value $f(x)$ at the extreme point x is named as the extreme value.

According to the Hessian matrix, the above definition of local extreme point can be described as follows.

Definition 2. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a second-order continuous function. A point $x \in \mathbb{R}^n$ is as follows: (1) a local minimum point for f if $\nabla f(x) = 0$ and the Hessian matrix of x is positive definite (all the eigenvalues are positive); (2) a local maximum point if $\nabla f(x) = 0$ and the Hessian matrix of x is negative definite (all the eigenvalues are negative).

A *n-d* type ridge point is the local maximum point in *n-d* orthogonal directions in *n*-dimensional space of which the definition is as follows.

Definition 3. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a second-order continuous function. A point $x \in \mathbb{R}^n$ is a *n-d* type ridge point if and only

if $[v_1, \ldots, v_d]^T \nabla I(x) = 0$ and $\lambda_d < 0$, where, $\lambda_1, \ldots, \lambda_n$ ($\lambda_1 \leq \cdots \leq \lambda_n$) are the eigenvalues of hessian matrix of point x and their corresponding eigenvectors are denoted by $v_1, \ldots, v_n$, and $1 \leq d \leq n$.

To extract ridge points on the image, we need to find all the points to meet the conditions that $[v_1, \ldots, v_d]^T \nabla I(x) = 0$ and $\lambda_d < 0$. For angiograms, images are two-dimensional data, ridge points are 1-type, and we need to detect the points that satisfy $(v_1)^T \nabla I(x) = 0$ and $\lambda_1 < 0$. According to the density of the real number, detecting all the ridge points within a limited time is impossible because ridge points are usually located on subpixel coordinates rather than on integer coordinates. In this paper, we obtain a sufficient number of ridge points based on the analysis of the gradient vector and the Hessian matrix at discrete pixels.

Ridge Point Existence Criterion. Assume that $I(x)$ is a vesselness image, for a point (x, y) and its neighbor points $(x + 1, y)$, $(x, y + 1)$, and $(x + 1, y + 1)$. There must be a ridge point (ξ, η) $(x \leq \xi \leq x + 1, y \leq \eta \leq y + 1)$ between them if the conditions are satisfied as follows:

$$\begin{aligned}
&\max \{v_1 \nabla I(x, y), v_1 \nabla I(x + 1, y), \\
&\qquad v_1 \nabla I(x, y + 1), v_1 \nabla I(x + 1, y + 1)\} > 0, \\
&\min \{v_1 \nabla I(x, y), v_1 \nabla I(x + 1, y), \\
&\qquad v_1 \nabla I(x, y + 1), v_1 \nabla I(x + 1, y + 1)\} < 0, \\
&\max \{\lambda_1(x, y), \lambda_1(x + 1, y), \\
&\qquad \lambda_1(x, y + 1), \lambda_1(x + 1, y + 1)\} < 0,
\end{aligned} \tag{8}$$

where $\nabla I(x, y)$ is the gradient vector at (x, y) and $\lambda_1(x, y)$ and $v_1(x, y)$ are the first eigenvalue and the first eigenvector of Hessian matrix at (x, y).

Proof. According to the Lemmas A.2 and A.3 and Theorem A.4, we know that since $I(x)$ is two-order continuous, $\nabla I(x, y)$ and $H(x, y)$ are continuous. $\lambda_1(x, y)$ is a single eigenvalue; therefore, $\lambda_1(x, y)$ and $v_1(x, y)$ are continuous for (x, y).

Thus

$$\begin{aligned}
&\max \{v_1 \nabla I(x, y), v_1 \nabla I(x + 1, y), \\
&\qquad v_1 \nabla I(x, y + 1), v_1 \nabla I(x + 1, y + 1)\} > 0, \\
&\min \{v_1 \nabla I(x, y), v_1 \nabla I(x + 1, y), \\
&\qquad v_1 \nabla I(x, y + 1), v_1 \nabla I(x + 1, y + 1)\} < 0.
\end{aligned} \tag{9}$$

According to the intermediate value theorem of continuous function, there is a point (ξ, η) $(x \leq \xi \leq x + 1, y \leq \eta \leq y + 1)$, which meets $v_1 \nabla(\xi, \eta) = 0$.

And since $\lambda_1(\xi, \eta)$ can be achieved by linear interpolation of $\lambda_1(x, y)$, $\lambda_1(x + 1, y)$, $\lambda_1(x, y + 1)$, and $\lambda_1(x + 1, y + 1)$,

$$\begin{aligned}
\lambda_1(\xi, \eta) = {} & \omega_3 [\omega_1 \lambda_1(x, y) + (1 - \omega_1) \lambda_1(x + 1, y)] \\
& + (1 - \omega_3) \\
& \cdot [\omega_2 \lambda_1(x, y + 1) + (1 - \omega_2) \lambda_1(x + 1, y + 1)], \\
& (0 \leq \omega_1, \omega_2, \omega_3 \leq 1).
\end{aligned} \tag{10}$$

We have

$$\begin{aligned}
&\max \{\lambda_1(x, y), \lambda_1(x + 1, y), \\
&\qquad \lambda_1(x, y + 1), \lambda_1(x + 1, y + 1)\} < 0.
\end{aligned} \tag{11}$$

Then $\lambda_1(\xi, \eta) < 0$. □

According to Definition 3, the point (ξ, η) is a ridge point. And according to the ridge point existence criterion, the ridge points can be detected by scanning the vesselness image line by line. In this paper, we take the pixel point (x, y) as the ridge point (ξ, η) to be a seed point. It can not only save the interpolation burden but also guarantee the extraction accuracy of the seed point.

2.5. Seed Point Refinement. A large amount of seed points located in vascular boundaries can be detected using the proposed method. However, a number of candidate seed points located on the background, which are denoted as pseudo seed points, can be observed because of the influence of noise. In this study, an automatic seed point refinement method is proposed based on the area gray scale distribution of the detected seed points. With P as the set of the sample points located in the area of the detected candidate seed points, a self-adaptive threshold function can be defined as follows:

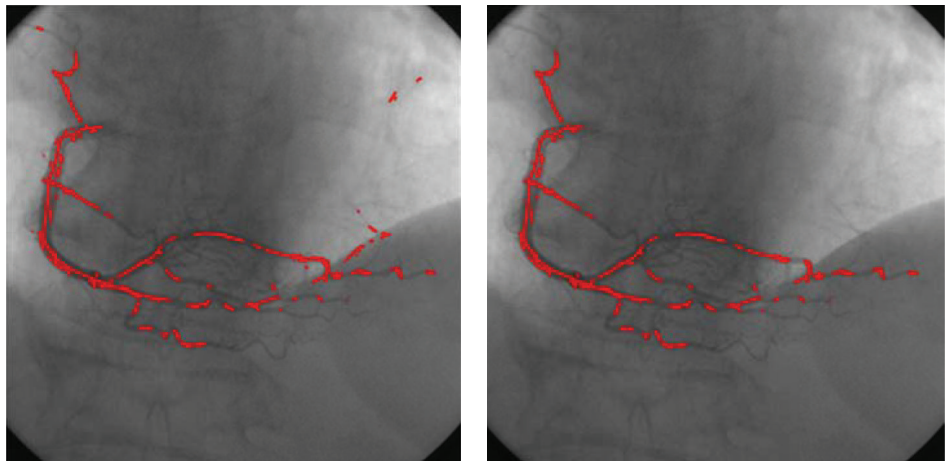
$$T = m(P) - \omega \cdot s(P), \tag{12}$$

where $m(P)$ is the median intensity value of arranged P and $s(P)$ is the median of the absolute value of all the points of P minus $m(P)$. ω is a weight factor, which controls the noise and sensitivity of the intensity.

Figure 2(a) shows the extracted candidate seed points on an angiogram based on the proposed ridge detection method. As shown in the figure, a large number of seed points are detected inside the vascular structures. However, some of them are still detected in the background. Figure 2(b) shows the refined results of the proposed method. As observed, the pseudo seed points located in the nonvascular region are removed, and most of the calculated seed points are inside vascular boundaries, whereas some of them are located near the vascular centerlines.

3. Experimental Results

To validate the performance of the proposed method, a series of coronary angiograms acquired from a Philips Digital



(a) (b)

Figure 2: Seed point calculation results in an angiogram. (a) Results of the ridge based method. (b) Results after seed refinement.

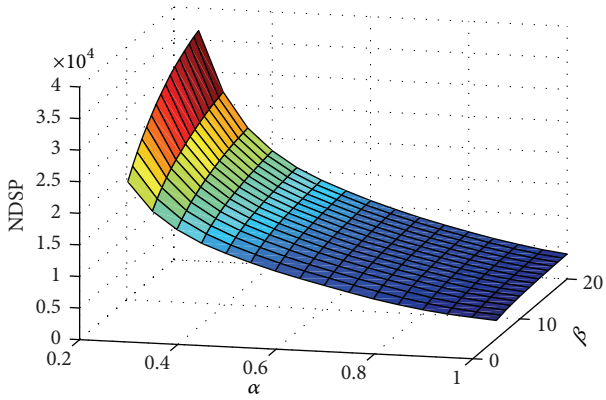


Figure 3: Relationship between NDSP and enhancement parameters α and β.

Imaging device at Beijing Chaoyang Rea-Cross Hospital were used. All angiograms have 512 × 512 resolution. The experiments were carried out on a desktop computer with an i7-2600 processor and 16 G memory, and the proposed method is compared with the other two traditional seed point detection algorithms.

Figure 3 was designed to evaluate the number of detected seed point (NDSP) with respect to the multiscale enhancement filtering parameters of α and β. For the experiment, the sampling ranges of α and β are set at $[0.25, 1]$ and $[5, 20]$, respectively, where the sampling step of α and β is set at 0.05 and 5. In this study, NDSP is defined as the number of detected seeds with respect to the enhancement response value larger than a predefined threshold of τ. In this experiment, the value of τ is set at 30. From the figure, if β is set constant, the NDSP values decrease quickly with increasing α. If α is set constant, the NDSP values increase slowly with increasing β. When β is comparatively small, a large amount of small enhanced noise appears in the angiogram. Hence, suppressing the noise while preserving the effect of enhancement during seed extraction is important.

To obtain the optimal enhancement parameters, the structures of coronary arteries in the angiogram were manually delineated from the background. As such, the number of pseudo seed points (NPSP) that rest in the background can be effectively quantified. Figure 4 demonstrates the relationship between NPSP and the enhancement parameters of α and β. From the figure, if β is set constant, NPSP first deceases rapidly and then slowly with increasing α. On the other hand, if α is set constant, NPSP increases slowly with increasing β.

To obtain the optimal enhancement parameters, the false detection ratio (FDR) is defined based on NDSP and NPSP as follows:

$$\mathrm{FDR} = \frac{\mathrm{NDSP}}{\mathrm{NPSP}}. \tag{13}$$

Figure 5 demonstrates the relationship between FDR and the enhancement parameters of α and β. From the figure, if β is set constant, FDR first deceases rapidly and then increases slowly with the increasing α. On the other hand, if α is set constant, NPSP first deceases and then increases slowly with increasing β. Therefore, there is a local minimum value of FDR corresponding to the optimal enhancement parameters of α and β.

To quantify the performance of the proposed method, the proposed seed point detection algorithm is compared with the methods proposed by Fritzsche et al. [9] and by Boroujeni et al. [10]. Figure 6 shows the seed extraction results from all the methods. The first to fourth columns correspond to the original angiograms, the results of Fritzsche, Boroujeni, and those of proposed method, respectively. The first to fifth rows correspond to five different data sets. For the Fritzsche method, only a few seed points are detected, which are mostly distributed in local parts of the vessels. Moreover, some detected points rest in the background. For the Boroujeni method, a number of detected seed points in the first four images are very small. Although more seed points are detected in the last image, detecting seed points in some of the major branches is difficult. Evidently, our proposed method can detect more seed points inside vascular

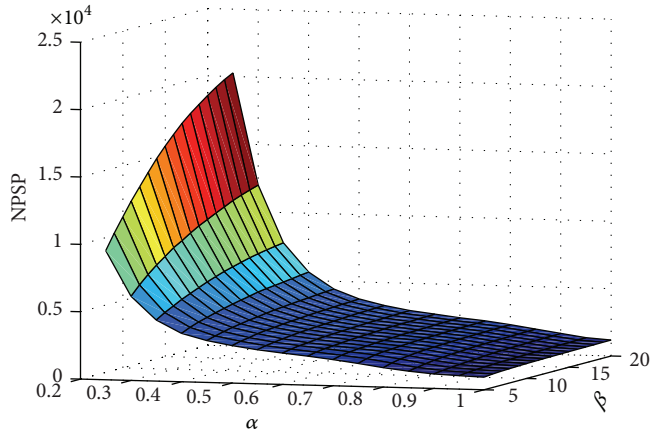


Figure 4: Relationship between NPSP and the enhancement parameters of α and β.

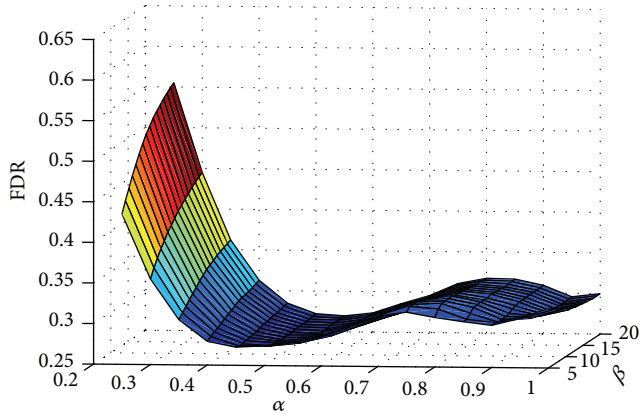


Figure 5: Relationship between the FDR and the enhancement parameters of α and β.

boundaries than the other two methods. Furthermore, the detected seed points are evenly distributed in the whole vasculature. Hence, it can be used more appropriately for the tracking procedures for centerline extraction.

To evaluate the proposed automatic seed point detection algorithm, three general measures of precision, recall, and f-measure are used in this paper. And the seed detection precision can be quantified by the percentage of correct seeds accounting for all the generated seeds; then we have the following equation:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{14}$$

where TP denotes the true positives (it is the total number of the detected seed points that are located inside the true vessels), while FP is the number of false positives (it is the number of the detected seed points that are located in the background).

The result of recall denotes the percentage of corrected seed points that can be detected by the proposed algorithm, and we have

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{15}$$

where FN denotes the number of false negatives; it is the total number of true seed points that are wrongly discarded by the refinement calculation procedure.

To balance between the precision and recall, the f-measure is proposed as follows:

$$\mathrm{FM} = \frac{2 \times R \times P}{R + P}. \tag{16}$$

Also, the total number of seed points and the number of vascular branches that can be detected are utilized to evaluate the preformation of algorithms, and they are denoted by $N1$ and $N2$, respectively. $N1$ shows the ability that how much seed points can be detected by the algorithm. The greater of $N1$ means more vascular point will be detected in the image; while the great of $N2$ means more vascular branches will be detected in the image.

Table 1 compares the seed point detection results of Fritzsche, Boroujeni, and the proposed method over five groups of data sets. The mean values of $N1$ of the Fritzsche, Boroujeni, and the proposed methods are 1127, 1845, and 2072, respectively. Obviously, the proposed method obtained more seed points than the other two methods. In the same manner, the mean values of precision of the Fritzsche, Boroujeni, and the proposed methods are 88.1%, 96.3%, and 98.2%, respectively. It indicates that only a few detected points of the proposed method are outside of the vascular boundaries, while the mean values of recall and f-measure of the Fritzsche, Boroujeni, and the proposed methods are almost the same, but the proposed algorithm is slightly higher than the other two methods. The mean values of $N2$ of the Fritzsche, Boroujeni, and the proposed methods are 5, 7, and 8, respectively. Clearly, the proposed method detects a larger number of vascular branches than the other two methods. It can be concluded that the Boroujeni method is

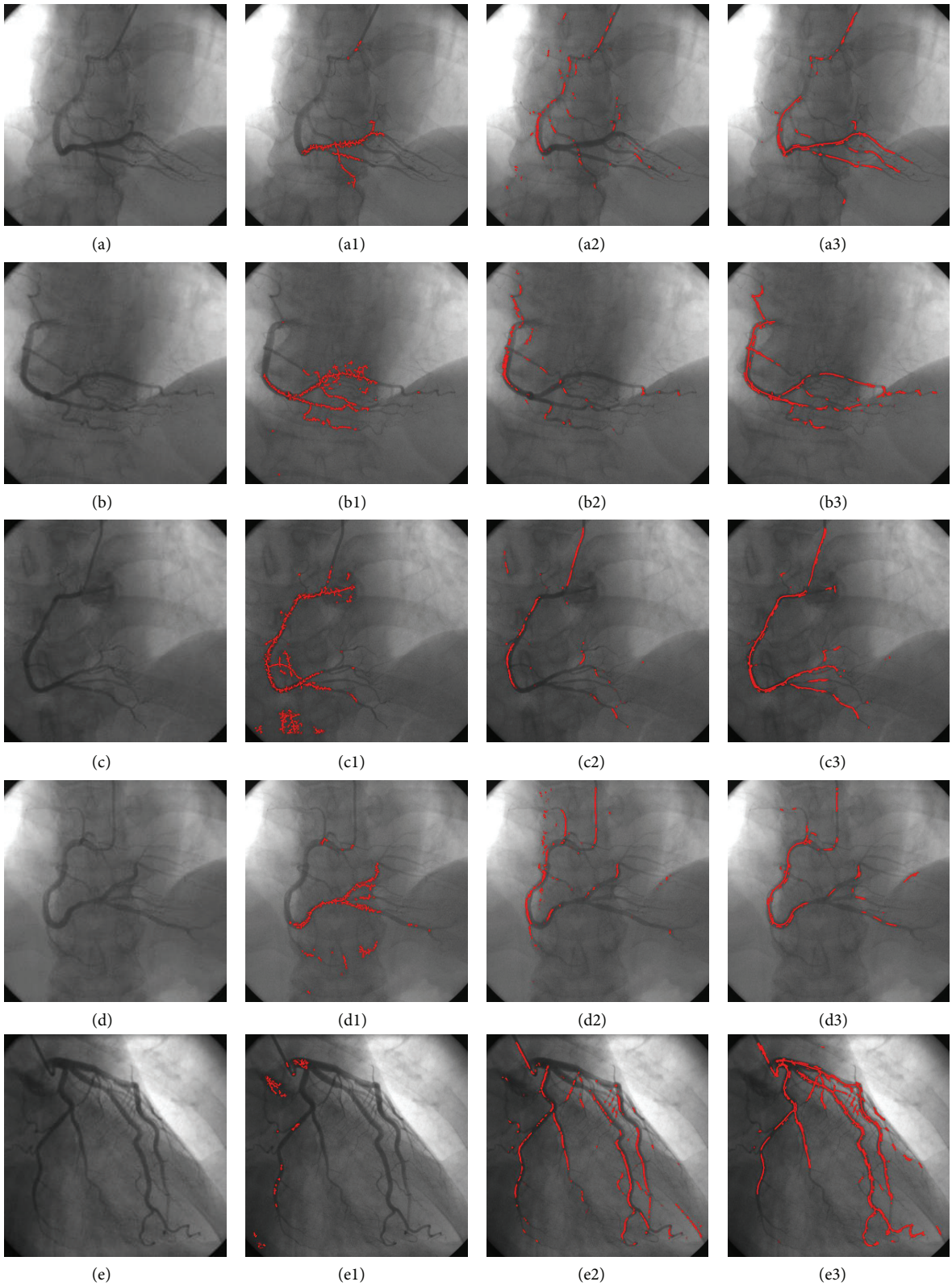


Figure 6: Seed point extraction results of five groups of data sets. The first to the fourth columns correspond to the source angiograms, segmentation results of Fritzsche, Boroujeni, and the proposed methods. The first to the fifth rows correspond to five different data sets.

Table 1: Comparison of the seed point detection results of the Fritzsche method, Boroujeni method, and the proposed method over five groups of data sets.

Data	Fritzsche method					Boroujeni method					Proposed method				
	P	R	FM	$N1$	$N2$	P	R	FM	$N1$	$N2$	P	R	FM	$N1$	$N2$
Data1	0.985	0.999	0.992	720	5	0.959	0.989	0.974	1272	9	0.994	0.999	0.996	1561	10
Data2	0.901	0.992	0.944	1488	7	0.973	0.998	0.985	1310	6	0.988	0.999	0.993	1760	9
Data3	0.912	0.981	0.945	1881	5	0.973	0.997	0.985	1131	5	0.990	0.999	0.994	1655	6
Data4	0.871	0.985	0.924	1129	4	0.936	0.989	0.962	1648	5	0.955	0.995	0.974	1079	6
Data5	0.736	0.948	0.829	416	2	0.976	0.992	0.984	3862	8	0.981	0.996	0.988	4306	11
Mean	0.881	0.981	0.927	1127	5	0.963	0.993	0.978	1845	7	0.982	0.998	0.989	2072	8

better than the Fritzsche method, while the proposed method outperforms the other two methods with respect to accuracy and the ability of branch detection.

4. Conclusion

This study proposes a novel automatic seed point detection method for X-ray angiographic images, which can be further utilized for vascular segmentation as well as centerline extraction. In study, the continuous properties of the eigenvalue and eigenvector are analyzed in depth. Based on the ridge point existence theorem, a novel discriminative function is proposed for candidate seed point detection from the multiscale Gaussian response of the angiographic image. The candidate seeds are refined according to the intensity distribution of neighboring pixels in the scanning lines. Furthermore, this study also discussed the optimal parameters for accurate seed detection. The study introduces five discrimination standards to quantify seed detection ability and evaluate the performance of different seed detection methods. The experiments demonstrate that the proposed method is very effective and robust for seed point detection in angiographic images with mean values of 98.2% and 2072 for the precision and number of detected seed points, respectively. Considering that the proposed method is fully automatic and with high detection ability, it can be utilized for fast centerline extraction as well as structure measurement for coronary arteries in clinical practice.

Appendix

Lemma A.1. *For any $f(x) \in C$, then $F(x) = f(x) * g(x) \in C^{\infty}$, where $g(x) = (1/\sqrt{2\pi})e^{-x^2/2}$.*

Proof. Consider

$$\begin{aligned} F(x) &= f(x) * g(x) \\ &= \int_{-\infty}^{+\infty} f(t)\, g(x-t)\, dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)\, e^{-(x-t)^2/2} dt \\ &= \frac{1}{\sqrt{2\pi}} \left[e^{-x^2/2} \int_{-\infty}^{+\infty} f(t)\, dt \right. \\ &\quad \left. + \int_{-\infty}^{+\infty} f(t)\, e^{xt} dt + \int_{-\infty}^{+\infty} f(t)\, e^{-t^2/2} dt \right] \\ &:= \frac{1}{\sqrt{2\pi}} \left[F_1 + F_2 + F_3 \right], \end{aligned} \tag{A.1}$$

where F_1 and F_2 are infinitely differentiable functions parameterized by x and F_3 is a constant. Therefore $F(x) = f(x) * g(x) \in C^{\infty}$. □

Lemma A.2. *Let $A, \Delta A \in C^{n\times n}$, where $A = \{a_{ij}\}$ and $\Delta A = \{\Delta a_{ij}\}$. If λ is a single eigenvalue of A and V is the corresponding eigenvector of λ, then for eigenvalue $\lambda(A + \Delta A)$ and its corresponding eigenvector $V(A + \Delta A)$ of $A + \Delta A$, it satisfied that $\lambda(A + \Delta A) \to \lambda$ and $V(A + \Delta A) \to V$, when $\Delta A \to 0$.*

Proof. Since the eigenvalue of the Hessian matrix is a continuous function for all the matrix elements and λ is a single eigenvalue of A, then for the eigenvalue $\lambda(A+\Delta A)$ of $A+\Delta A$, $\lambda(A + \Delta A) \to \lambda$ when $\Delta A \to 0$. And for a sufficiently small $\|\Delta A\|$, $\lambda(A + \Delta A)$ is a single eigenvalue of $A + \Delta A$.

Since λ is a single eigenvalue of A, according to the theorem of Jordan canonical form, we have $\operatorname{rank}(\lambda E - A) = \operatorname{rank}(A) - 1$, where E is an $n \times n$ identity matrix. Hence, we can find certain i and j, for which the $n-1$ order nonsingular matrix A^*_{n-1} can be obtained by removing the elements of the ith row and the jth column of A.

According to

$$AV = \lambda V, \tag{A.2}$$

where $V = (v_1, \ldots, v_n)^T$, the coefficient matrix of the system

$$-\sum_{\substack{m=1 \\ m\neq j}}^{m=n} (a_{km} - \delta_{km}\lambda)\, v_m = (a_{kj} - \delta_{kj}\lambda)\, v_j, \quad k \neq i, \tag{A.3}$$

is nonsingular, where

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases} \tag{A.4}$$

Without loss of generality, we assume that $v_j = 1$.

For a sufficiently small $\|\Delta A\|$, $\lambda(A+\Delta A)$ is a single eigenvalue of $A+\Delta A$, and $\operatorname{rank}(\lambda E-(A+\Delta A))=\operatorname{rank}(A+\Delta A)-1$. The $n-1$ order matrix $(A+\Delta A)^{*}_{n-1}$, which is obtained by removing the elements of the ith row and the jth column of $A+\Delta A$, is also nonsingular. Hence, the system of linear equations

$$-\sum_{\substack{m=1\\ m\neq j}}^{m=n}\left(a_{km}+\Delta a_{km}-\delta_{km}\lambda\left(A+\Delta A\right)\right)v_m\left(A+\Delta A\right)=\left(a_{kj}+\Delta a_{km}-\delta_{kj}\lambda\left(A+\Delta A\right)\right)v_j\left(A+\Delta A\right),\quad k\neq i, \tag{A.5}$$

have a unique solution $v_1(A+\Delta A),\ldots,v_{j-1}(A+\Delta A),v_{j+1}(A+\Delta A),\ldots,v_n(A+\Delta A)$, and they are continuous functions of ΔA. If we choose $v_j(A+\Delta A)=v_j=1$, there will be

$$V\left(A+\Delta A\right)=\left(v_1\left(A+\Delta A\right),\ldots,v_{j-1}\left(A+\Delta A\right),v_j\left(A+\Delta A\right),v_{j+1}\left(A+\Delta A\right),\ldots,v_n\left(A+\Delta A\right)\right)^{T}\longrightarrow V, \tag{A.6}$$

when $\Delta A\rightarrow 0$. □

Lemma A.3. *If the image $I(x,y)$ is second-order continuous, then the two eigenvalues λ_1 and λ_2 of Hessian matrix of the point (x,y) are unequal, and their corresponding eigenvalues v_1 and v_2 are orthogonal to each other.*

Proof. If $I(x,y)\in C^2$, then $I_{xy}=I_{yx}$; that is, the hessian matrix $H(x,y)$ of (x,y) is symmetrical. According to the properties of real symmetric matrices, the two eigenvectors of $H(x,y)$ are orthogonal.

The two eigenvalues can be obtained by solving the characteristic polynomial $\begin{vmatrix}\lambda-I_{xx} & -I_{xy}\\ -I_{yx} & \lambda-I_{yy}\end{vmatrix}=0$ as follows:

$$\lambda^2-\left(I_{xx}+I_{yy}\right)\lambda+I_{xx}I_{yy}-I_{xy}^2=0. \tag{A.7}$$

The two eigenvalues are equal if and only if

$$\left(I_{xx}+I_{yy}\right)^2-4\left(I_{xx}I_{yy}-I_{xy}^2\right)=0. \tag{A.8}$$

From (A.8) we have

$$\left(I_{xx}-I_{yy}\right)^2+4I_{xy}^2=0. \tag{A.9}$$

$H(x,y)$ has two equal eigenvalues if and only if

$$\begin{aligned}I_{xx}&=I_{yy},\\ I_{xy}&=I_{yx}=0.\end{aligned} \tag{A.10}$$

By solving (A.10), we have

$$\begin{aligned}I_x&=c_1,\\ I_y&=c_2,\end{aligned} \tag{A.11}$$

where c_1 and c_2 are constant. To satisfy (A.11), $I(x,y)\equiv c$, where c is constant; that is, the intensity of image $I(x,y)$ is constant, which is in conflict with our application. Therefore, for each point (x,y) on image $I(x,y)$, the two eigenvalues λ_1 and λ_2 of the Hessian matrix $H(x,y)$ are unequal. □

Theorem A.4. *If an image $I(x,y)\in\mathbb{C}^2$, for any point (x,y) on $I(x,y)$, its eigenvalues and eigenvectors of Hessian matrix are continuous.*

Proof. If $I(x,y)\in\mathbb{C}^2$, the Hessian matrix $H(x,y)$ is $\begin{pmatrix}I_{xx} & I_{xy}\\ I_{yx} & I_{yy}\end{pmatrix}$, where the elements I_{xx}, I_{xy}, I_{yx}, and I_{yy} are all continuous. According to the definition of eigenvalues, the two eigenvalues $\lambda_1(x,y)$ and $\lambda_2(x,y)$ are two real roots of the characteristic polynomial of H:

$$f(\lambda)=\begin{vmatrix}\lambda-I_{xx} & -I_{xy}\\ -I_{yx} & \lambda-I_{yy}\end{vmatrix}=0. \tag{A.12}$$

For continuous functions, their roots are also continuous; that is $\lambda_1(x,y)$ and $\lambda_2(x,y)$ are continuous. Therefore, the eigenvalues of Hessian matrix of image are continuous. □

Disclosure

Dr. Guangzhi Wang is a coauthor.

Conflict of Interests

The authors declare that they have no competing interests.

Acknowledgment

This work was supported by the National Basic Research Program of China (2013CB328806), the Key Projects in the National Science & Technology Pillar Program (2013BAI01B01), and the National Hi-Tech Research and Development Program (2013AA013703).

References

[1] C.-Y. Lin and Y.-T. Ching, "Extraction of coronary arterial tree using cine X-ray angiograms," *Biomedical Engineering—Applications, Basis and Communications*, vol. 17, no. 3, pp. 111–120, 2005.

[2] Y. Qian, S. Eiho, N. Sugimoto, and M. Fujita, "Automatic extraction of coronary artery tree on coronary angiograms by morphological operators," in *Proceedings of the Computers in Cardiology*, pp. 765–768, IEEE, Cleveland, Ohio, USA, September 1998.

[3] O. Wink, W. J. Niessen, and M. A. Viergever, "Multiscale vessel tracking," *IEEE Transactions on Medical Imaging*, vol. 23, no. 1, pp. 130–133, 2004.

[4] M. Sofka and C. V. Stewart, "Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures," *IEEE Transactions on Medical Imaging*, vol. 25, no. 12, pp. 1531–1546, 2006.

[5] S. Y. Tang, Y. T. Wang, and Y.-W. Chen, "Application of ICA to X-ray coronary digital subtraction angiography," *Neurocomputing*, vol. 79, pp. 168–172, 2012.

[6] Y. Bentoutou, N. Taleb, M. C. El Mezouar, M. Taleb, and L. Jetto, "An invariant approach for image registration in digital subtraction angiography," *Pattern Recognition*, vol. 35, no. 12, pp. 2853–2865, 2002.

[7] J. Yang, Y. Wang, S. Tang, S. Zhou, Y. Liu, and W. Chen, "Multiresolution elastic registration of X-ray angiography images using thin-plate spline," *IEEE Transactions on Nuclear Science*, vol. 54, no. 1, pp. 152–166, 2007.

[8] R. Collorec and J. L. Coatrieux, "Vectorial tracking and directed contour finder for vascular network in digital subtraction angiography," *Pattern Recognition Letters*, vol. 8, no. 5, pp. 353–358, 1988.

[9] K. Fritzsche, A. Can, H. Shen et al., "Automated model based segmentation, tracing and analysis of retinal vasculature from digital fundus images," in *State-of-the-Art Angiography, Applications and Plaque Imaging Using MR, CT, Ultrasound and X-Rays*, J. S. Suri and S. Laxminarayan, Eds., pp. 225–298, 2003.

[10] F. Z. Boroujeni, R. W. O. K. Rahmat, N. Mustapha, L. S. Affendey, and O. Maskon, "Automatic selection of initial points for exploratory vessel tracing in Fluoroscopic images," *Defence Science Journal*, vol. 61, no. 5, pp. 443–451, 2011.

[11] D. Eberly, R. Gardner, B. Morse, S. Pizer, and C. Scharlach, "Ridges for image analysis," *Journal of Mathematical Imaging and Vision*, vol. 4, no. 4, pp. 353–373, 1994.

[12] C. Smets, G. Verbeeck, P. Suetens, and A. Oosterlinck, "A knowledge-based system for the delineation of blood vessels on subtraction angiograms," *Pattern Recognition Letters*, vol. 8, no. 2, pp. 113–121, 1988.

[13] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Interventation—MICCAI'98*, vol. 1496 of *Lecture Notes in Computer Science*, pp. 130–137, Springer, Berlin, Germany, 1998.

[14] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, UK, 1965.

[15] M. Vlachos and E. Dermatas, "Multi-scale retinal vessel segmentation using line tracking," *Computerized Medical Imaging and Graphics*, vol. 34, no. 3, pp. 213–227, 2010.

[16] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 117–154, 1998.

[17] T. Lindeberg, *Scale-Space Theory in Computer Vision*, Springer, Boston, Mass, USA, 1994.

[18] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever, "Scale and the differential structure of images," *Image and Vision Computing*, vol. 10, no. 6, pp. 376–388, 1992.

[19] L. Florack, B. T. H. Romeny, M. Viergever, and J. Koenderevk, "The Gaussian scale-space paradigm and the multiscale local jet," *International Journal of Computer Vision*, vol. 18, no. 1, pp. 61–75, 1996.

[20] B. M. T. H. Romeny, "Introduction to scale-space theory: multiscale geometric image analysis," Tech. Rep., Utrecht University, Utrecht, The Netherlands, 1996.

[21] R. S. Eaton, M. R. Stevens, J. C. McBride, G. T. Foil, and M. S. Snorrason, "A systems view of scale space," in *Proceedings of the 4th IEEE International Conference on Computer Vision Systems (ICVS '06)*, p. 3, January 2006.