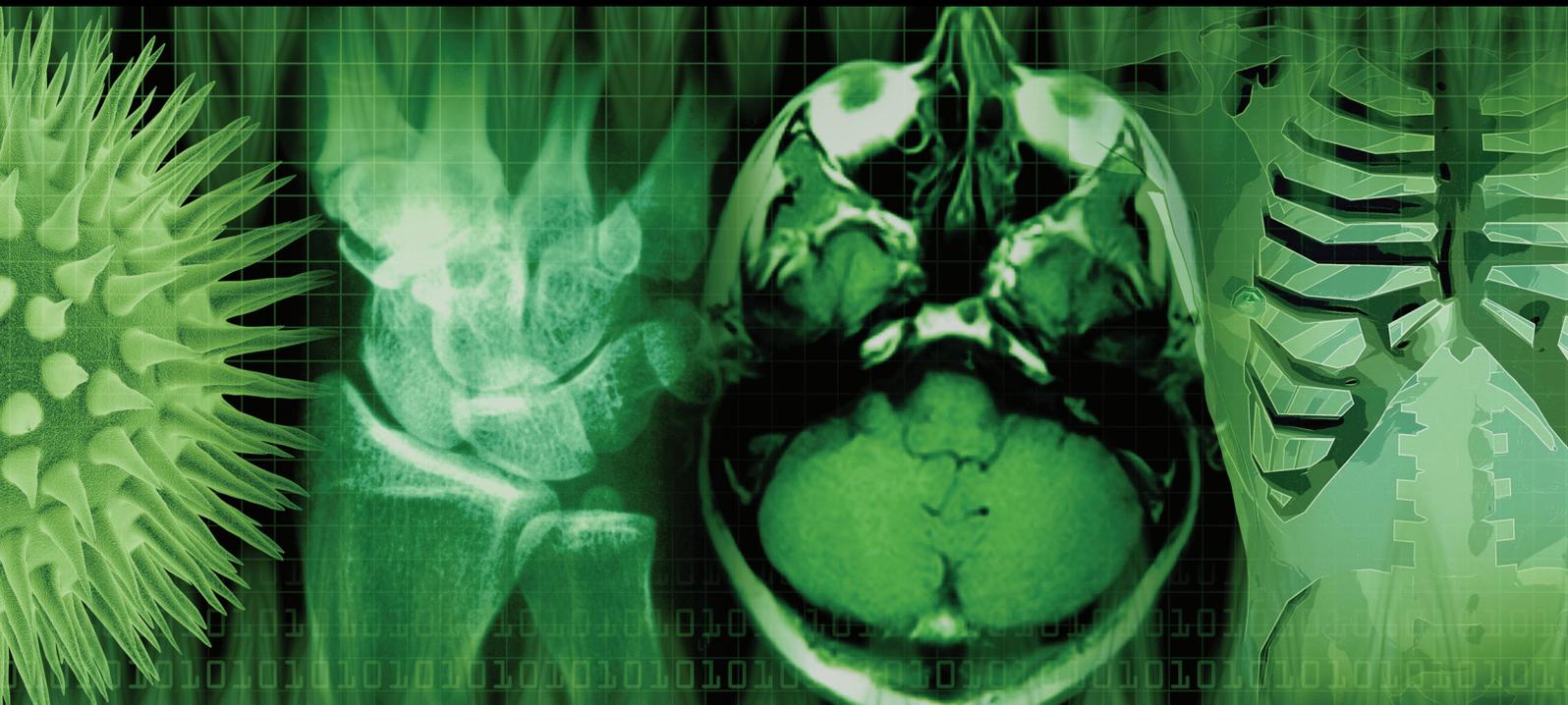


Image Processing Algorithms and Measures for the Analysis of Biomedical Imaging Systems Applications

Guest Editors: Karen Panetta, Sos Agaian, Jean-Charles Pinoli, and Yicong Zhou





**Image Processing Algorithms and Measures
for the Analysis of Biomedical Imaging Systems
Applications**

International Journal of Biomedical Imaging

**Image Processing Algorithms and Measures
for the Analysis of Biomedical Imaging Systems
Applications**

Guest Editors: Karen Panetta, Sos Aghaian, Jean-Charles Pinoli,
and Yicong Zhou



Copyright © 2015 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in “International Journal of Biomedical Imaging.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Haim Azhari, Israel
K. Ty Bae, USA
Richard Bayford, UK
Jyh-Cheng Chen, Taiwan
Anne Clough, USA
Carl Crawford, USA
Daniel Day, Australia
Eric Hoffman, USA
Tzung-Pei Hong, Taiwan
Jiang Hsieh, USA
Ming Jiang, China
Marc Kachelriess, Germany

Cornelia Laule, Canada
Seung W. Lee, Republic of Korea
A. K. Louis, Germany
Jayanta Mukherjee, India
Sushmita Mukherjee, USA
Scott Pohlman, USA
Erik L. Ritman, USA
Itamar Ronen, USA
Chikayoshi Sumi, Japan
Lizhi Sun, USA
Kenji Suzuki, USA
Jie Tian, China

Michael W. Vannier, USA
Ge Wang, USA
Yue Wang, USA
Guowei Wei, USA
D. L. Wilson, USA
Sun K. Yoo, Republic of Korea
Habib Zaidi, Switzerland
Yantian Zhang, USA
Jun Zhao, China
Yibin Zheng, USA
Tiange Zhuang, China
Yu Zou, USA

Contents

Image Processing Algorithms and Measures for the Analysis of Biomedical Imaging Systems Applications, Karen Panetta, Sos Aгаian, Jean-Charles Pinoli, and Yicong Zhou
Volume 2015, Article ID 926921, 1 page

Automatic Extraction of Blood Vessels in the Retinal Vascular Tree Using Multiscale Medialness, Mariem Ben Abdallah, Jihene Malek, Ahmad Taher Azar, Philippe Montesinos, Hamed Belmabrouk, Julio Esclarín Monreal, and Karl Krissian
Volume 2015, Article ID 519024, 16 pages

Transverse Strains in Muscle Fascicles during Voluntary Contraction: A 2D Frequency Decomposition of B-Mode Ultrasound Images, James M. Wakeling and Avleen Randhawa
Volume 2014, Article ID 352910, 9 pages

Comparison and Supervised Learning of Segmentation Methods Dedicated to Specular Microscope Images of Corneal Endothelium, Yann Gavet and Jean-Charles Pinoli
Volume 2014, Article ID 704791, 13 pages

Choosing the Optimal Spatial Domain Measure of Enhancement for Mammogram Images, Karen Panetta, Arash Samani, and Sos Aгаian
Volume 2014, Article ID 937849, 8 pages

Automated White Matter Hyperintensity Detection in Multiple Sclerosis Using 3D T2 FLAIR, Yi Zhong, David Utriainen, Ying Wang, Yan Kang, and E. Mark Haacke
Volume 2014, Article ID 239123, 7 pages

Nonreference Medical Image Edge Map Measure, Karen Panetta, Chen Gao, Sos Aгаian, and Shahan Nercessian
Volume 2014, Article ID 931375, 8 pages

Editorial

Image Processing Algorithms and Measures for the Analysis of Biomedical Imaging Systems Applications

Karen Panetta,¹ Sos Aгаian,² Jean-Charles Pinoli,³ and Yicong Zhou⁴

¹*Tufts University, USA*

²*University of Texas at San Antonio, USA*

³*Ecole Nationale Supérieure des Mines, France*

⁴*Computer and Information Science, University of Macau, Macau*

Correspondence should be addressed to Karen Panetta; karen@ieee.org

Received 25 December 2014; Accepted 28 December 2014

Copyright © 2015 Karen Panetta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The medical profession has changed dramatically in the past decade due to the advances in engineering innovations. Traditionally, medical doctors relied solely on their training and experience for diagnosis and interpretation of medical conditions captured through a variety of imaging techniques. Typically, these images are produced from CAT scan machines, MRI, or X-rays.

Today, computer aided techniques are designed to enhance the visual quality of biomedical images to better help humans discriminate regions of interest, such as the presence of cancer. Since these enhanced images are intended to assist the medical professional in making a diagnosis or to track the progress of a treatment, it is important that abnormalities presented in the images are actual objects and not artifacts resulting from the enhancement process that could lead to an inaccurate diagnosis.

Too much information presented can be distracting for the viewer and too little information presented could be disastrous for the patient, should a life-threatening abnormality be missed or filtered out. Therefore, the dilemma lies in answering the question: “What is the best image to use?”

The answer has always been subjective due to the perspective of the human observer and what constitutes the desired target object to be detected. For instance, some doctors looking at a scan may be interested in masses; another doctor may want to focus on calcifications or, even perhaps, the blood vessels.

In this special issue, articles are presented that address this question for several different biomedical image processing

applications. Several novel quantifiable evaluation methods that can rank the quality of a processed image in accordance with human visual perception are introduced. These metrics are rigorously compared to evaluations obtained from human evaluators using the “Mean Opinion Score” (MOS), along with other standard benchmarking techniques.

The innovative approaches presented in this issue are sure to lead the biomedical image processing field beyond one of an “assistive technology” to one that creates a reliable, accurate autonomous technology for use in distance and robot vision applications. Achieving these goals will forge new groundbreaking paths that will reach into and serve remote populations of the world.

Acknowledgment

Thanks are due to all my coeditors, Dr. Sos Aгаian, Dr. Jean-Charles Pinoli, and Dr. Yicong Zhou, for their contributions and expert insights into the topics.

*Karen Panetta
Sos Aгаian
Jean-Charles Pinoli
Yicong Zhou*

Research Article

Automatic Extraction of Blood Vessels in the Retinal Vascular Tree Using Multiscale Medialness

Mariem Ben Abdallah,¹ Jihene Malek,¹ Ahmad Taher Azar,² Philippe Montesinos,³ Hafedh Belmabrouk,¹ Julio Esclarín Monreal,⁴ and Karl Krissian⁴

¹Faculty of Sciences, Electronics and Microelectronics Laboratory, Monastir University, 5019 Monastir, Tunisia

²Faculty of Computers and Information, Benha University, Benha 13511, Egypt

³Institute of Mines and Ales, Laboratory of Computer and Production Engineering, 30319 Alès, France

⁴Imaging Technology Center (CTIM), Las Palmas-Gran Canaria University, 35017 Las Palmas de Gran Canaria, Spain

Correspondence should be addressed to Mariem Ben Abdallah; mariembenabdallah3@gmail.com

Received 11 April 2014; Accepted 12 November 2014

Academic Editor: Karen Panetta

Copyright © 2015 Mariem Ben Abdallah et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose an algorithm for vessel extraction in retinal images. The first step consists of applying anisotropic diffusion filtering in the initial vessel network in order to restore disconnected vessel lines and eliminate noisy lines. In the second step, a multiscale line-tracking procedure allows detecting all vessels having similar dimensions at a chosen scale. Computing the individual image maps requires different steps. First, a number of points are preselected using the eigenvalues of the Hessian matrix. These points are expected to be near to a vessel axis. Then, for each preselected point, the response map is computed from gradient information of the image at the current scale. Finally, the multiscale image map is derived after combining the individual image maps at different scales (sizes). Two publicly available datasets have been used to test the performance of the suggested method. The main dataset is the STARE project's dataset and the second one is the DRIVE dataset. The experimental results, applied on the STARE dataset, show a maximum accuracy average of around 94.02%. Also, when performed on the DRIVE database, the maximum accuracy average reaches 91.55%.

1. Introduction

For decades, retinal images are widely used by ophthalmologists for the detection and follow-up of several pathological states [1–5]. Fundus photographs, also called retinal photography, are captured using special devices called “Charged Coupled Devices” (CCD), which are cameras that show the interior surface of the eye [6–10]. These images directly provide information about the normal and abnormal features in the retina. The normal features include the optic disk, fovea, and vascular network. There are different kinds of abnormal features caused by diabetic retinopathy (DR) such as microaneurysm, hard exudate, soft exudate, hemorrhage, and neovascularization. An example of retinal images obtained by fundus photography is given in Figure 1, where two retinal images are shown. The first one does not show any DR sign

(Figure 1(a)) and the second one demonstrates advanced-DR signs indicated by color arrows (Figure 1(b)). However, the manual detection of blood vessels is very difficult since the blood vessels in these images are complex and have low level contrast [11]. Also, not all the images show signs of diabetic retinopathy. Hence, a manual measurement of the information about blood vessels, such as length, width, tortuosity, and branching pattern, becomes tedious. As a result, it increases the time of diagnosis and decreases the efficiency of ophthalmologists. Therefore, automatic methods for extracting and measuring the vessels in retinal images are needed to save the workload of the ophthalmologists and to assist in characterizing the detected lesions and identifying the false positives [12].

Several works have been proposed for detecting the 2D complex vessel network, such as single scale matched

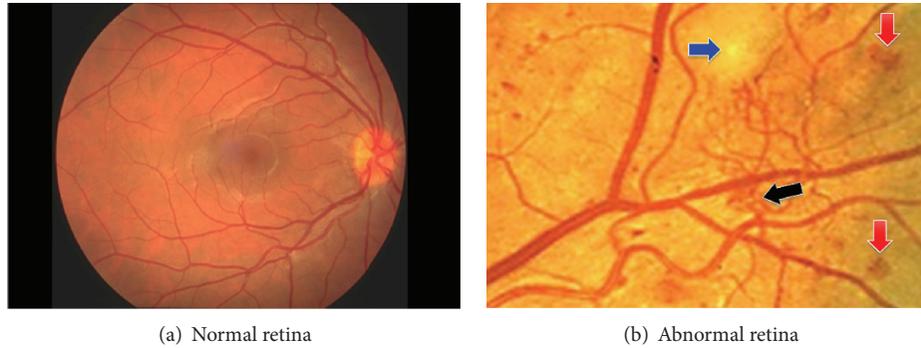


FIGURE 1: Retinal images [32].

filter [13–15], multiscale matched filter [16], adaptive local thresholding [17], single-scale Gabor filters [18], and multiscale Gabor filters [19]. Cinsdikici and Aydin [20] put forward a blood vessel segmentation based on a novel hybrid model of the matched filter and the colony algorithm, which extracts vessels perfectly but the pathological areas can affect the result. In [21–23] authors adapted another approach which applied mathematical morphological operators. The suggested method in [21] proved to be a valuable tool for the segmentation of the vascular network in retinal images, where it allowed obtaining a final image with the segmented vessels by iteratively combining the centerline image with the set of images that resulted from the vessel segments' reconstruction phase using the morphological operator. However, the inconvenience of this method is when a vessel centerline is missing, so the corresponding segmented vessel is normally not included in the final segmentation result. In [22], the authors proved that it was possible to select vessels using shape properties and connectivity, as well as differential properties like curvature. The robustness of the algorithm has been evaluated and tested on eye fundus images and on other images. Gang et al. [24] showed that the Gaussian curve is suitable for modeling the intensity profile of the cross section of the retinal vessels in color fundus images. Based on this elaboration, they proposed the amplitude-modified second-order Gaussian filter for retinal vessel detection, which optimized the matched filter and improved the successfulness of the detection. Staal et al. [25] explained a method for an automated segmentation of vessels in two-dimensional color images. The system was based on extracting image ridges that coincide approximately with vessel centerlines, where the evaluation was done using the accuracy of hard classifications and the values of soft ones. In [26], the authors presented a hybrid method for an efficient segmentation of multiple oriented blood vessels in colour retinal images. The robustness and accuracy of the method demonstrated that it might be useful in a wide range of retinal images even with the presence of lesions in the abnormal images. Dua et al. [27] presented a method for detecting blood vessels, which employs a hierarchical decomposition based on a quad tree decomposition. The algorithm was faster than the existing approaches. In the recent years, alternative approaches for an automated vessel segmentation have used the Hessian-based

multiscale detection of curvilinear structures, which has been effective in discerning both large and small vessels [28–31].

In this paper, we propose a multiscale response to detect linear structures in 2D images. We will use the formulation, which was suggested in [36, 37]. The presented detection algorithm is divided into two steps. First, we present a flux-based anisotropic diffusion method and apply it to denoise images corrupted by an additive Gaussian noise. In order to extract only the pixels belonging to a vessel region, we use a Gaussian model of the vessels for interpreting the eigenvalues and the eigenvectors of the Hessian matrix. Then, we compute the multiscale response from responses computed at a discrete set of scales. The method has been evaluated using the images of two publicly available databases, the DRIVE database [34] and the STARE database [33]. Prior to analysing fundus images, we have used the green channel alone, since it gives the highest contrast between the vessel and the background.

2. Methodology

2.1. Preprocessing Technique. In the ocular fundus image, edges and local details between heterogeneous regions are the most interesting part for clinicians. Therefore, it is very important to preserve and enhance edges and local fine structures and simultaneously reduce the noise. To reduce the image noise, several approaches have been proposed using techniques such as linear and nonlinear filtering. In linear spatial filtering, such as Gaussian filtering, the content of a pixel is given by the value of the weighted average of its immediate neighbors. This filtering not only reduces the amplitude of noise fluctuations but also degrades sharp details such as lines or edges, so the resulting images appear blurred and diffused [24, 38]. This undesirable effect can be reduced or avoided by designing nonlinear filters. The most common technique is median filtering. With it the value of an output pixel is determined by the median of the neighborhood pixels. This filtering retains edges but results in a loss of resolution by suppressing fine details [39]. In order to perform this task, Perona and Malik (PM) [18] developed an anisotropic diffusion method, a multiscale smoothing, and the edge detection scheme, which were a powerful concept

in image processing. The anisotropic diffusion was inspired from the heat diffusion equation by introducing a diffusion function, g , which depended upon the norm of the gradient of the image:

$$\frac{\partial u}{\partial t} = \text{div}(g(|\nabla u|) \cdot \nabla u), \quad (1)$$

where ∇ and $u(x, t)$ denote gradient operation and image intensity, respectively, div is the divergence operator, and $|\cdot|$ denotes the magnitude. The variable x represents the spatial coordinate, while the variable t is used to enumerate iteration steps in the discrete implementation. Perona and Malik suggested the following diffusion functions:

$$g(|\nabla u|) = \frac{1}{1 + (|\nabla u|/k)^2}, \quad (2)$$

$$g(|\nabla u|) = \exp\left[-\left(\frac{|\nabla u|}{k}\right)^2\right],$$

where k is a parameter of the norm gradient. In this method of anisotropic diffusion, the norm gradient is used to detect edges or frontiers in the image as a step of intensity discontinuity. To understand the relation between the parameter k and the discontinuity value $|\nabla u|$, $F(\nabla u)$ can be defined as the following product $F(\nabla u) = g \times \nabla u$, called the flow diffusion.

- (i) If $|\nabla u| \gg k$, then $g(|\nabla u|) \rightarrow 0$ and we have a filter pass-all.
- (ii) If $|\nabla u| \ll k$, then $g(|\nabla u|) \rightarrow 1$ and we obtain an isotropic diffusion filter (like a Gaussian filter), which is a low-pass filter that attenuates high frequencies.

The one-dimensional discrete implementation of (1) is given by

$$\begin{aligned} & \frac{\partial u}{\partial t}(x, t) \\ &= \frac{\partial}{\partial x} (g(x, t) \cdot \nabla (u)(x, t)) \\ &\approx \frac{\partial}{\partial x} \left(g(x, t) \cdot \frac{1}{dx} \left(u\left(x + \frac{dx}{2}, t\right) - u\left(x - \frac{dx}{2}, t\right) \right) \right) \\ &\approx \frac{1}{dx^2} \left[g\left(x + \frac{dx}{2}, t\right) \cdot (u(x + dx, t) - u(x, t)) \right. \\ &\quad \left. - g\left(x - \frac{dx}{2}, t\right) \cdot (u(x, t) - u(x - dx, t)) \right] \\ &\approx F_{\text{right}} - F_{\text{left}} \quad \text{if } dx = 1, \end{aligned} \quad (3)$$

where $F_{\text{right}} = F(x + (dx/2), t)$ and $F_{\text{left}} = F(x - (dx/2), t)$.

The above result is generalized in n -dimensional:

$$\frac{\partial u}{\partial t} \approx \sum_{i=1}^n F_{x_i^+} - F_{x_i^-} \quad (4)$$

if $\forall i, dx_i = 1$, $F_{x_i^+} = F_{x_i}(x + (dx_i/2), t)$ and $F_{x_i^-} = F_{x_i}(x - (dx_i/2), t)$.

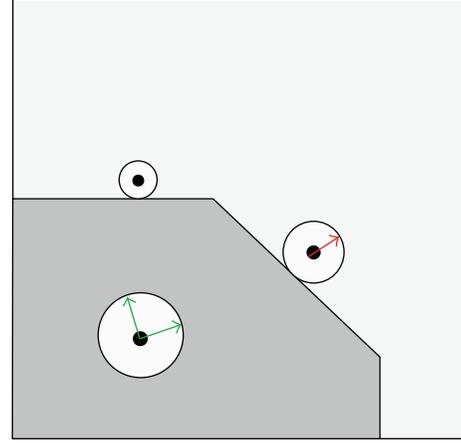


FIGURE 2: PM anisotropic diffusion.

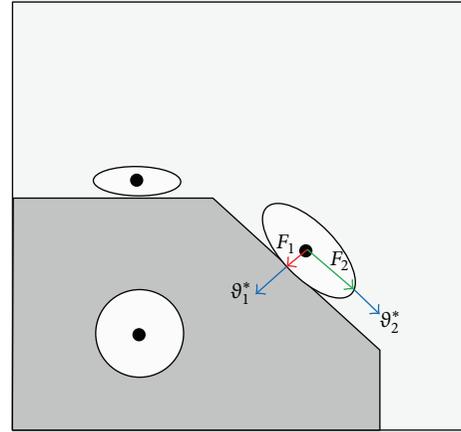


FIGURE 3: Directional anisotropic diffusion.

Up to now, the anisotropic diffusion has been defined as the case where the diffusivity is a scalar function varying with the location in the image. As described earlier, the PM diffusion (Figure 2) limits the smoothing of an image near the pixels with a high gradient magnitude (edge pixels). As the diffusion near an edge is very weak, the noise smoothing near the edge is also small. To address this, diffusions using matrices instead of scalars have been put forward [36, 40, 41], where the anisotropic diffusion allows the diffusion to be different along various directions defined by the local geometry of the structures in the image (Figure 3). Thus, the diffusion on both sides of an edge can be prevented while allowing the diffusion along the edge. This prevents the edge from being smoothed and then being removed during denoising.

The F flux of the matrix diffusion (MD) form can be written as

$$\text{div}(D\nabla u), \quad (5)$$

where D is a positive definite symmetric matrix that may be adapted to the local image structure, which can be written in

TABLE 1: Parameters and results of different filters for vessel image.

Filter	N	k	β	σ	dt	Neig.	PSNR (dB)	MSE
GF	—	—	—	2	—	9×9	37.7717	10.8620
MF	—	—	—	—	—	5×5	38.6364	8.9011
PM	13	7	—	—	0.15	—	39.6735	7.0103
DAD	50	7	0.05	0.8	0.05	—	40.4337	5.8845

terms of its eigenvectors v_1 and v_2 and eigenvalues λ_1 and λ_2 , as follows:

$$D = [v_1 \ v_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}. \quad (6)$$

Subsequently, the gradient vector field can be written as

$$\nabla u = u_{v_1} v_1 + u_{v_2} v_2. \quad (7)$$

Following the eigenvalues and eigenvectors that we have chosen, different matrix diffusions can be obtained [36, 41]. The diffusion matrix proposed by Weickert et al. [41] had the same eigenvectors as the structure tensor, with eigenvalues that are a function of the norm of the gradient [41, 42]. In our work, we have used a 2D basis (v_1^*, v_2^*) which corresponds, respectively, to unit vectors in the directions of the gradient and to the minimal curvature of the regularized (or smoothed) version of the image, which is the image convolved with a Gaussian filter with a standard deviation σ . This basis is of particular interest in the context of small, elongated structures such as blood vessels, where the minimal curvature holds for the axis direction orthogonal to the gradient. These directions are obtained as two of the eigenvectors of the Hessian matrix of the smoothed image: H_σ (further details are described in Section 2.3). Therefore, the eigenvectors are defined as follows:

$$\begin{aligned} v_1^* &\parallel \nabla u_\sigma, \\ v_2^* &\perp \nabla u_\sigma, \end{aligned} \quad (8)$$

where ∇u_σ is the gradient of the image convolved with a Gaussian filter with a standard deviation σ , v_2^* gives an estimation of the vessel direction, and v_1^* is its orthogonal. Also, we have used the eigenvalues in (6) as a diffusion function associated to each vector of the basis depending on the first order derivative of the intensity in this direction, instead of the traditional norm of the smoothed gradient. Furthermore, the diffusion can be decomposed as a sum of diffusions in each direction of the orthogonal basis and the divergence term can be written as [36]

$$\text{div}(F) = \text{div}\left(\sum_{i=1}^2 \phi_i(u_{v_i^*}) \cdot v_i^*\right) = \sum_{i=1}^2 \text{div}(\phi_i(u_{v_i^*}) \cdot v_i^*), \quad (9)$$

where $u_{v_i^*}$ and ϕ_i indicate the first order derivative of the intensity in the direction v_i and the i th diffusion function, respectively. Also, ϕ_1 can be chosen to be any of the diffusivity functions from the traditional nonhomogeneous isotropic

diffusion equation, which depends on the first order derivative of the intensity in this direction, as $\phi_1(u_{v_1^*}) = u_{v_1^*} e^{-(u_{v_1^*}/k)^2}$ and $\phi_2(u_{v_2^*}) = \alpha \cdot u_{v_2^*}$, with $0 < \alpha < 1$, being only a diffusing function to allow smoothing in a v_2^* direction. For further details, the reader could refer to [36, 43].

As in [36], we use a data attachment term with a coefficient β which allows a better control of the extent to which the restored image differs from the original image u_0 (at $t = 0$) and of the result of the diffusion process at convergence. The anisotropic diffusion equation becomes

$$\frac{\partial u}{\partial t} = \sum_{i=1}^2 \text{div}(\phi_i(u_{v_i^*}) \cdot v_i^*) + \beta(u - u_0). \quad (10)$$

In order to evaluate the denoising effects of the directional anisotropic diffusion (DAD), we have added a Gaussian white noise to each of the images in Figure 4. Once the diffusion method is applied to these noisy images, its effectiveness in reducing the noise is got by calculating the peak signal to noise ratio (PSNR) relative to the original image as follows:

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{d^2}{\text{MSE}}\right), \quad (11)$$

where $d = 255$ and MSE is the mean-squared error which is written as

$$\text{MSE} = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (I_{\text{original}}(i, j) - I_{\text{denoised}}(i, j))^2, \quad (12)$$

where I_{original} refers to the original image without noise and I_{denoised} is the image after the denoising process.

The higher the PSNR is, the better the effect of the denoising is. Note that this measure does not necessarily imply that an image with a higher PSNR is also more visually gratifying. However, based on our experiments using the three test images with an additive white Gaussian noise, we can draw some observations. First, all the techniques we have tried have several parameters that must be selected carefully to obtain the best results. Since we have a ‘‘clean’’ original image, as well as one with noise, we can use the increment in the PSNR value to guide our choice of the parameters. These parameters and the obtained results are indicated in Tables 1, 2, and 3, where we can observe that for the images corrupted with an additive Gaussian noise, the DAD method performs better than the PM method. It gains a higher PSNR (40.4337, 20.9045, and 33.3515) and a smaller MSE (5.8845, 527.9932, and 30.0557) than the aforementioned three methods.

Figure 4 represents some of the best results for the different methods (GF, MF, PM, and DAD) on the presented three

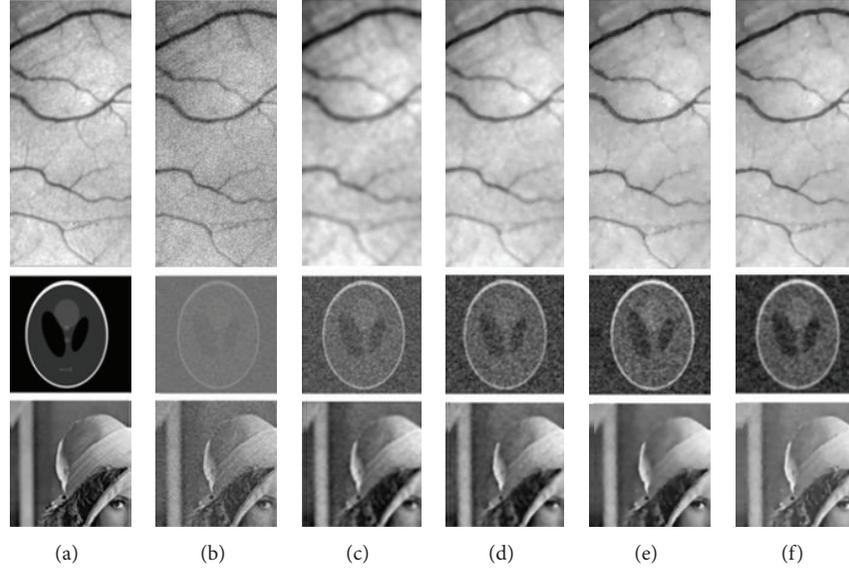


FIGURE 4: Original images (a) and the corresponding images with additive Gaussian noise (b); denoised images: best result with GF (c), best result with MF (d), best result with PM filter (e), and best result with directional anisotropic diffusion filter (f).

TABLE 2: Parameters and results of different filters for phantom image.

Filter	N	k	β	σ	dt	Neig.	PSNR (dB)	MSE
GF	—	—	—	2	—	5×5	18.8731	842.8924
MF	—	—	—	—	—	5×5	20.2437	614.7677
PM	20	3	—	—	0.15	—	20.8821	530.7294
DAD	75	2	0.05	0.8	0.05	—	20.9045	527.9932

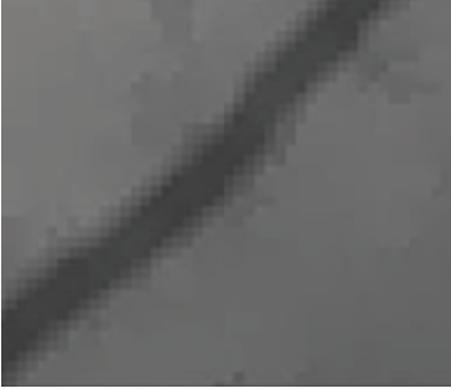
TABLE 3: Parameters and results of different filters for Lena image.

Filter	N	k	β	σ	dt	Neig.	PSNR (dB)	MSE
GF	—	—	—	2	—	5×5	31.4598	46.4621
MF	—	—	—	—	—	5×5	29.14504	79.1734
PM	10	7	—	—	0.15	—	32.9911	32.6562
DAD	20	7	0.05	0.8	0.05	—	33.3515	30.0557

test images (Vessels, phantom, and Lena). For instance, the results recorded after applying the DAD method show that this latter improves much more the visual rendering of the image compared to other methods. As shown in the images of the first row, a DAD filter can effectively improve the quality of a noisy image and also well enhance edges and preserve more details than other filters. Indeed, the Gaussian filter smooths very strongly the planar areas which causes loss of information regarding the fine structures of the image, and it blurs the image. The Median filter, compared to the Gaussian filter, preserves edges but loses details. Comparing the results of the DAD method to those obtained by the PM diffusion in Figures 5 and 6, we can derive several observations. The denoising of PM diffusion model is sensitive to the value of the conductance parameter k , and, therefore, smoothing is performed along ridges but not across a ridge line which causes enhancing the desired ridges as well as the noise. To be compared to the DAD diffusion filter, the diffusivity

is a tensor-valued function varying with the location and orientation of edges in an image. So, when this filter is applied to a ridge line smoothing is performed along ridges as across a ridge line while preserving the details.

2.2. Multiscale Medialness. The general approach of multiscale methods is to choose a range of scales between t_{\min} and t_{\max} (corresponding to σ_{\min} and σ_{\max}), which are discretized using a logarithmic scale in order to have more accuracy for low scales and to compute a response for each scale from the initial image [36, 43, 47]. The user specifies the minimal and maximal radius of the vessels to extract. Thus, the computation of the single scale response requires different steps. First, a number of points are preselected using the eigenvalues of the Hessian matrix. These points are expected to be near a vessel axis. Then, for each preselected point, the response is computed at the current scale σ . The response function uses eigenvectors of the Hessian matrix of the image

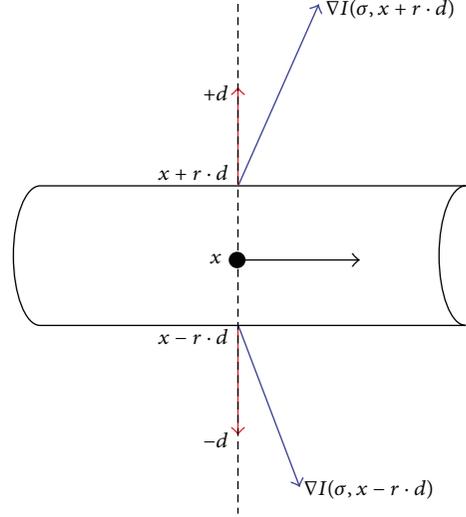
FIGURE 5: PM anisotropic diffusion ($k = 3, N = 100$).FIGURE 6: Directional anisotropic diffusion ($k = 3, N = 100, \alpha = 0.5$).

to define at each point an orientation $D(\sigma, x)$ orthogonal to the axis of a potential vessel that goes through M . From this direction, the two points located at an equal distance of M in the direction D , noted M_1 and M_2 (Figure 7). The response $R_\sigma(I)$ at M is taken as the maximum absolute value, among these two points, of the first derivative of the intensity in the D direction:

$$R_\sigma(x) = \max\{\nabla_\sigma I(\sigma, x + \sigma \cdot d) \cdot (+d), \nabla_\sigma I(\sigma, x - \sigma \cdot d) \cdot (-d)\}, \quad (13)$$

where d is the unitary vector of the direction D , that is, $d = \vec{v}_1$, and $\nabla_\sigma I$ is the gradient of the image at the scale σ . $\nabla_\sigma I$ is obtained by the convolution with the first derivative of a Gaussian function of the standard deviation σ , where multiplying the derivatives by σ ensures the scale invariance property and allows comparing the responses obtained from different scales. The gradient vector $\nabla_\sigma I$ can be computed by a bilinear interpolation for better accuracy, which is especially needed when looking at small vessels [37, 39].

A vessel of a radius r is detected at a scale t , so we use the scales corresponding to each radius for the multiscale processing. For a fixed scale t , we calculate a response image $R_t(I)$ where I is the initial image. Then we calculate

FIGURE 7: Representation of vesselness measure calculation (from the point x on the central line, d is the unit vector perpendicular to the main direction of the vessel and $r = \sigma$ is the current scale).

the multiscale response for the image $R_{\text{multi}}(I)$ which is the maximum of the responses over scales: for each point $x \in I$ and a range $[t_{\min}, t_{\max}]$ of scale:

$$R_{\text{multi}}(x) = \max_t \{R_t(x), t \in [t_{\min}, t_{\max}]\}. \quad (14)$$

This response $R_{\text{multi}}(x)$ can be interpreted as an indicator that the point x belongs to the center line of a vessel, and $R_t(x)$ can be interpreted as an indicator that the point x belongs to the center line of a vessel with a radius t . Finally, this response is normalized to give a multiscale response that combines interesting features of each single scale response.

One difficulty with the multiscale approach is that we want to compare the result of a response function at different scales, whereas the intensity and its derivatives are decreasing scale functions. So far, all considerations have been made at a single scale defined by the scale parameter σ . In his work, about scale space theory, Lindeberg and Fagerström [48] showed the need for a multiscale analysis to take the varying size of objects into account. He also showed the necessity of normalizing the spatial derivatives between different scales. Thus, the normalized vesselness response is obtained by the product of the normalization term σ^γ and the final vesselness:

$$R^*(\Sigma, \gamma, x) := \max_{\sigma \in \Sigma} \sigma^\gamma \cdot R(\sigma, x) = \max_{i=1, \dots, n} \sigma_i^\gamma \cdot R(\sigma_i, x). \quad (15)$$

The parameter γ can be used to indicate the preference for a particular scale (Figure 8). If it is set to one, no scale is preferred. Besides, the multiscale response is got by selecting the maximum response over a set of different scales between σ_{\min} and σ_{\max} .

2.3. Extraction of Local Orientations. The proposed model assumes that the intensity profile of the vessels in the cross section is Gaussian (Figure 9). This is a common assumption that it is employed in numerous algorithms [28, 35, 49].

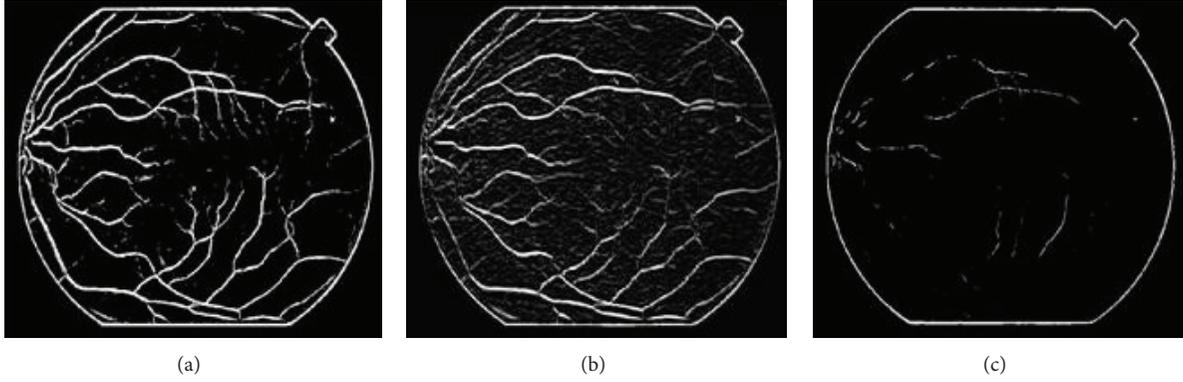


FIGURE 8: Influence of the normalization parameter γ on multiscale response; (a) $\gamma = 1$ is neutral; (b) $\gamma > 1$ favors large scales; finally, (c) $\gamma < 1$ favors small scales.

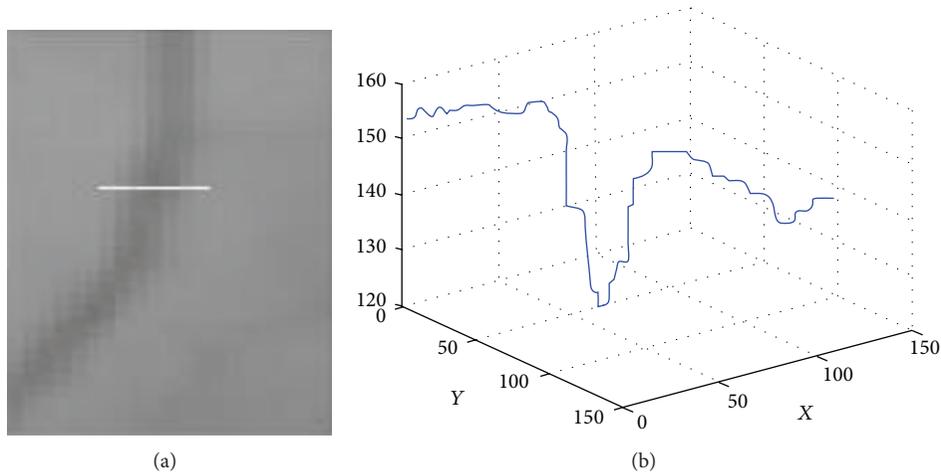


FIGURE 9: Example of cross sectional profile of blood vessel from gray scale 2D image (the gray intensities are plotted in a 3D view. The x , y axis is the position of the pixel in the 2D plane of the image, whereas the z -axis is the gray value or intensity of the pixel).

It is also commonly assumed that the intensity does not change much along vessels [49–51]. Recently, the Hessian matrix could be used to describe the local shape characteristics and orientation for elongated structures [35, 52]. The eigenvalues of this matrix, when the gradient is weak, express the local variation of the intensity in the direction of the associated eigenvectors. Subsequently, we assume that we want to characterize the dark vessels (low intensity) on a white background (high intensity).

Let us denote λ_1 and λ_2 as the eigenvalues of the Hessian matrix with $\lambda_1 \geq \lambda_2$ and \vec{v}_1, \vec{v}_2 being their associated eigenvectors (Figure 10). For a linear model with a Gaussian cross section, the vessel direction is defined by the eigenvector with the smallest eigenvalue at the center of the vessel, but it is less determined at the contours because both eigenvalues of the Hessian matrix are zero.

To summarize, for an ideal linear structure in a 2D image,

$$\begin{aligned} |\lambda_2| &\approx 0, \\ |\lambda_1| &> |\lambda_2|. \end{aligned} \quad (16)$$

In retinal images, some large vessels may have a white line in their center and some elongated and disjoint spots (Figures 11(a), 11(b), and 11(c)); accordingly, the vessels do not invalidate the Gaussian profile assumption. So, such lines are usually lost after the preselection of vessel pixels using the Hessian eigenvalue analysis and classified as background pixels. Therefore, the responses of the gradient magnitude are a task which is of particular importance in improving the detection vessels (Figure 11). The experimental results are demonstrated in Figure 11, which shows hand labeled “truth” images, and segmented images obtained, respectively, by the Hessian eigenvalue analysis and the gradient magnitude. From these results we can deduce that responses based on the gradient magnitude can availablely detect white lines as vessel pixels and removes some noise spots.

3. Results

In this section, the proposed method has been evaluated on two publicly available retinal image databases, the STARE database [33] and the DRIVE database [25]. The STARE

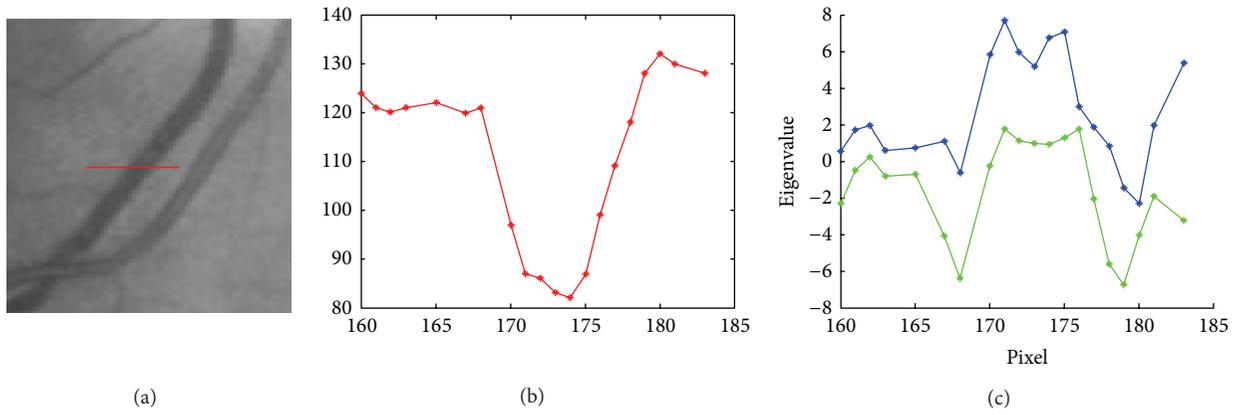


FIGURE 10: Eigenvalue analysis. (a) vessel cross section; (b) intensity distribution ($\sigma = 4.55$) vessel cross section; (c) corresponding eigenvalues.

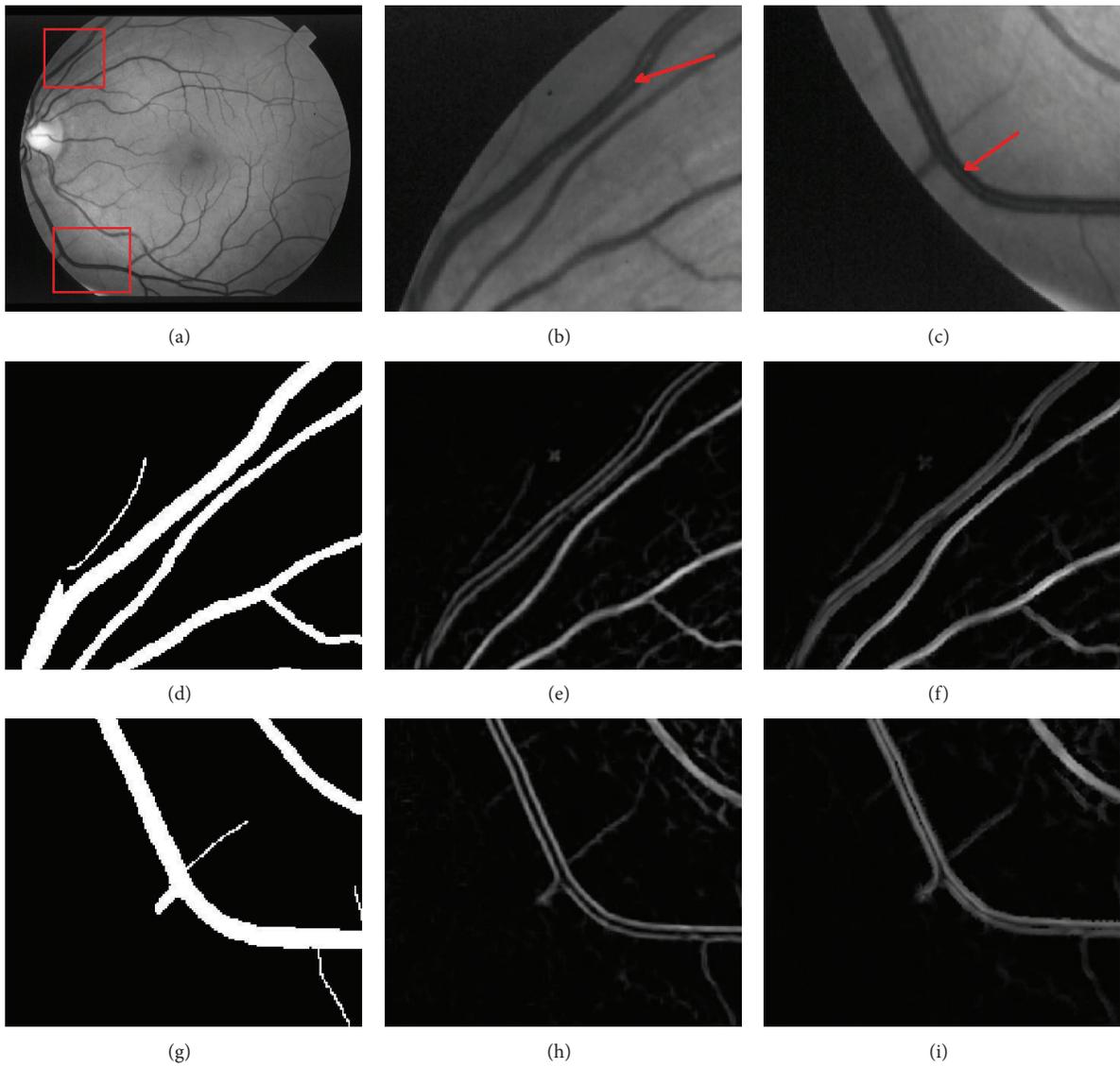


FIGURE 11: Retinal blood vessel detection. (a, b, and c) original images [33]; (d-g, e-h, and f-i) subimage of hand labeled image, vessel detection based Hessian eigenvalue analysis, and improved vessel detection with gradient magnitude.

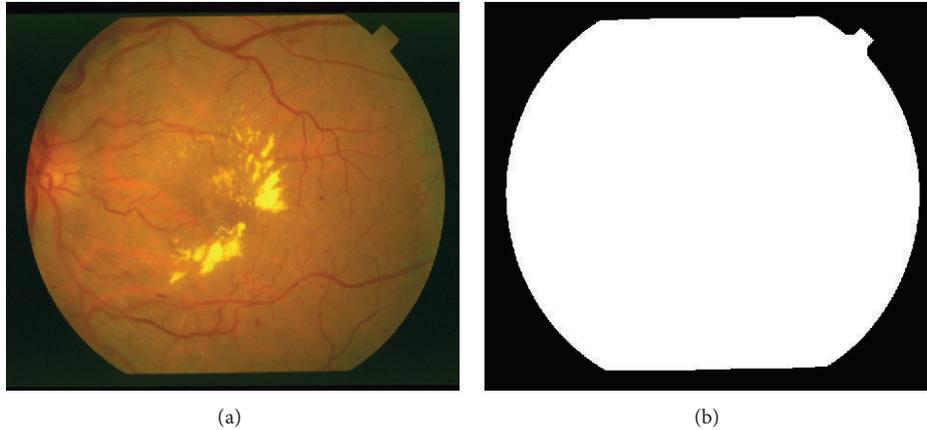


FIGURE 12: Binary mask of STARE project retinal image [33].

dataset contains twenty fundus colour retinal images, ten of which are from healthy ocular fundi and the other ten are from unhealthy ones. These images are captured by a Topcon TRV-50 fundus camera at a 35 Field Of View (FOV), which have digitized with a 24-bit gray-scale resolution and a size of 700×605 pixels. The dataset provides two sets of standard hand-labeled segmentations, which are manually segmented by two eye specialists. We create for this dataset a binary mask of the gray channel of the image using a simple threshold technique (Figure 12). We adapt the first eye specialist hand labelled as the ground truth to evaluate our vessel detection technique. The DRIVE dataset consists of 40 fundus ocular images, which have been divided into a training set and a test set by the authors of the database. These images are captured by the Canon CR5 camera at 45 FOV, which have been digitized at 24 bits with a resolution of 565×584 pixels. The dataset also gives two sets of standard hand-labeled segmentations by two human experts as a 9-ground truth.

The first expert hand labelled segmentation has been adapted as a ground truth to evaluate segmentation techniques on both STARE and DRIVE datasets. It is a common practice to evaluate the performance of retinal vessel segmentation algorithms using receiver operating characteristic (ROC) curves [25, 35]. An ROC curve plots the fraction of pixels correctly classified as vessels, namely, the true positive rate (TPR), versus the fraction of pixels wrongly classified as vessels, namely, the false positive rate (FPR), by varying the rounding threshold T from 0 to 1 (Figure 13). The closer the curve approaches the top left corner, the better the performance of the system. In order to facilitate the comparison with other retinal vessel detection algorithms, we have selected the value of the area under the curve (AUC), which is 1 for an ideal system.

To measure the performance of the proposed enhancement filter, we ran our multiscale analysis filter with the following set of parameters:

- (i) r_{\min} , r_{\max} , s , and the minimal and maximal radii used in this application are $r_{\min} = 1.25$ and $r_{\max} = 7$, discretized using $s = 4$ scales;

- (ii) the parameter γ set to one to indicate no scale is preferred;
- (iii) the value k is a constant threshold on the norm of gradient on the image;
- (iv) N is the number of iterations for the anisotropic diffusion filter.

The computing time of our algorithm for an image of the STARE database is about 64 seconds, including anisotropic diffusion filtering, and about the same time for the DRIVE database. The implementation of the filter has been done in MATLAB, on a personal computer with a 2.13 Intel Core Duo processor and 4 GB of memory. In the first experiment, we apply a preprocessing task such as filtering data with an anisotropic diffusion version, cited above, in order to remove or at least reduce noise. The DAD filter denoises the original image by preserving edges and details. To show that the segmentation works better with anisotropic diffusion, Figure 14 presents a segmentation result before and after the application of the anisotropic diffusion scheme. In this figure, we show the improvements provided by the DAD model, which tends to remove noise effects and, unfortunately, smaller objects. So, it preserves efficiently the vessels while making the background more homogeneous.

On the other hand, for computing the response, it is possible to retain the mean of the two calculated values (the gradient of the two points located at an equal distance from the current point), like in the 3D case proposed by [36], or the minimal calculated value in the 2D case [37]. We prefer retaining the maximum of these two values. Figure 15 shows a synthetic image which consists of 100×100 pixels with an 8-bit resolution. We have chosen this image because it contains an object close to the vessel form. The latter figure shows the segmentation results by maximum, average, and minimum response functions. We note that for the case of minimum or average responses, the ring is not completely detected like in the original image, since we can see it has been missing pixels belonging to the edges, in contrast to maximum case where the extraction of the ring is complete. Table 4 presents the AUC calculated with our method for the test set of the STARE

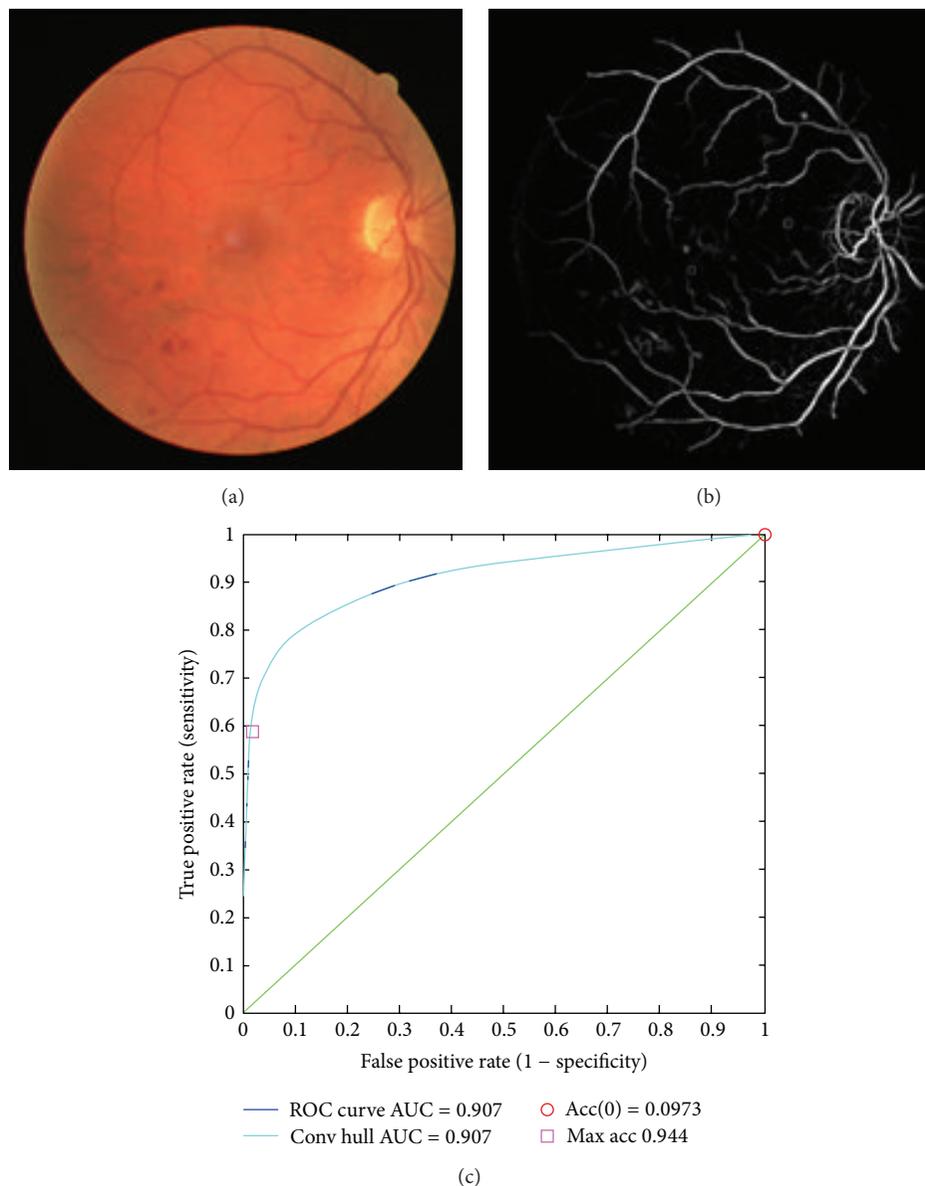


FIGURE 13: ROC curve of retinal image (06_test.tif) downloaded from DRIVE dataset [34]; (a) original image; (b) segmented image; (c) Roc curve.

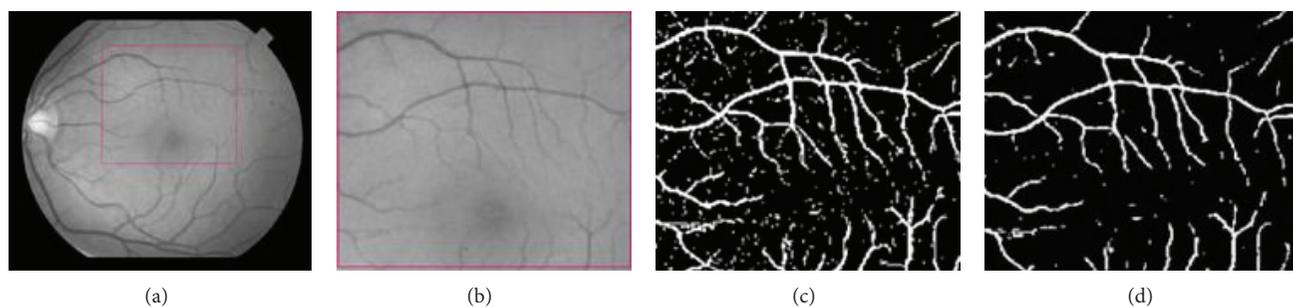


FIGURE 14: Effect of anisotropic diffusion. (a) Green channel of the original image downloaded from the STARE project dataset [33]. (b) Subimage of the original image, rescaled for better visualization, (c) segmentation without anisotropic diffusion, and (d) segmentation with anisotropic diffusion, $k = 1.25$, $\beta = 0.05$, and $N = 30$.

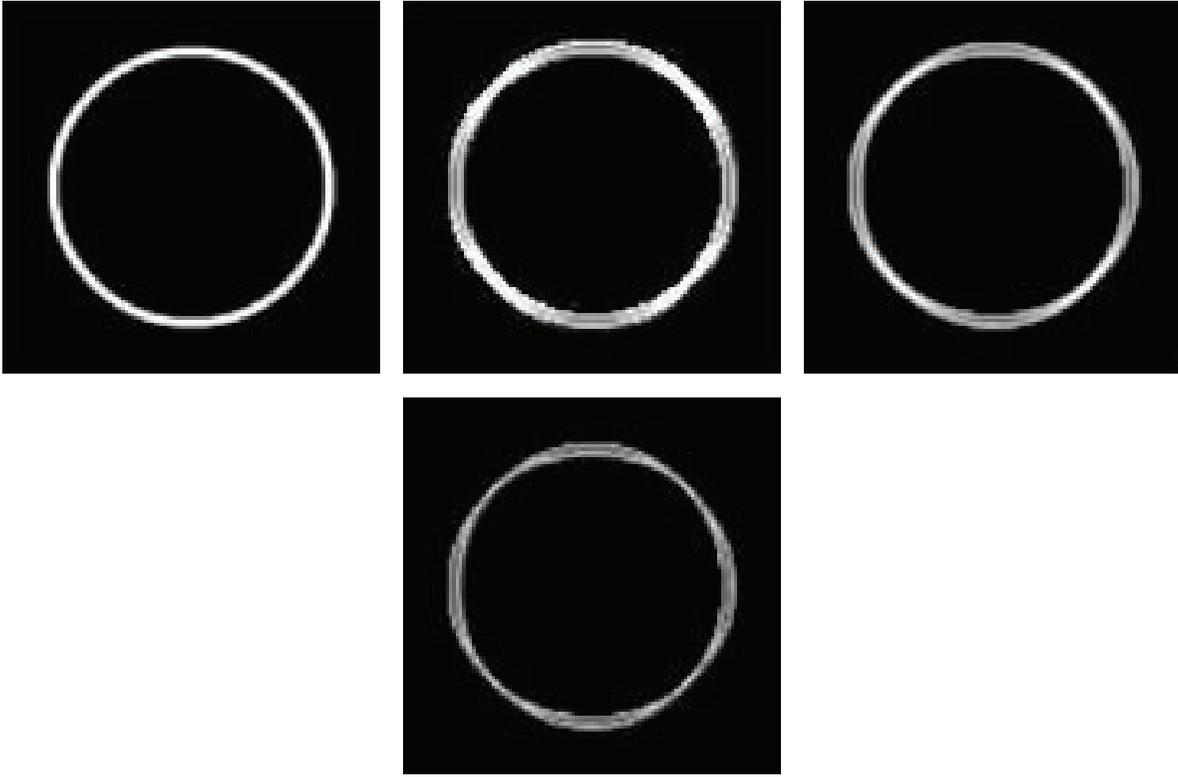


FIGURE 15: Original synthetic image, maximum response, average response, and minimum response $\sigma \in \{0.25, 0.5, 1, 2, 4\}$ (left to right-top to bottom).

TABLE 4: STARE project database [33].

	Mean	Min	Max
AUC	0.9329	0.9053	0.9445

database, using the green channel images. As given in the table, the experimental results show that the maximum model (AUC = 0.9445) performs much better than the average (AUC = 0.9329) or minimum model (AUC = 0.9053).

Figure 16 presents the obtained response image of a real retinal image, where four scales have been used for radii of vessels ranging from 1.25 to 7: {1.25, 2.22, 4, 7}. This figure shows that small and large vessels can be better distinguished in the maximum case than the minimum or average ones.

Although the contrast is not very high in the original figure (Figure 14(a)), the method detects most vessels, over a large size range. For example, in Figure 17, an image of the retinal tree vasculature is presented, where different responses recorded at increasing scales are represented. The last image shows a quite good performance of the vessel subtraction. Yet Figure 18 proves that it is possible to design a system that approaches the performance of human observers.

In order to evaluate the suggested method, experiment results of the 20-image sets of the STARE database are shown in Table 5. In Table 6, our method is compared to the most

TABLE 5: ROC curve analysis of STARE project database [33].

Number	MAA	TPR	FPR
1	0.9014	0.5537	0.0398
2	0.8740	0.1178	0.0045
3	0.9168	0.3819	0.0119
4	0.9286	0.5525	0.0135
5	0.9240	0.5678	0.0218
6	0.9414	0.5128	0.0139
7	0.9672	0.7626	0.0141
8	0.9683	0.7534	0.0149
9	0.9652	0.7366	0.0123
10	0.9420	0.6171	0.0182
11	0.9503	0.6379	0.0133
12	0.9655	0.7694	0.0105
13	0.9864	0.6992	0.0180
14	0.9480	0.6899	0.0162
15	0.9487	0.6882	0.0207
16	0.9226	0.6788	0.0215
17	0.9499	0.7099	0.0168
18	0.9484	0.6812	0.0102
19	0.9585	0.6058	0.0114
20	0.9345	0.6000	0.0172
	Av.MAA	Av.TPR	Av.FPR
	0.9402	0.6145	0.0162

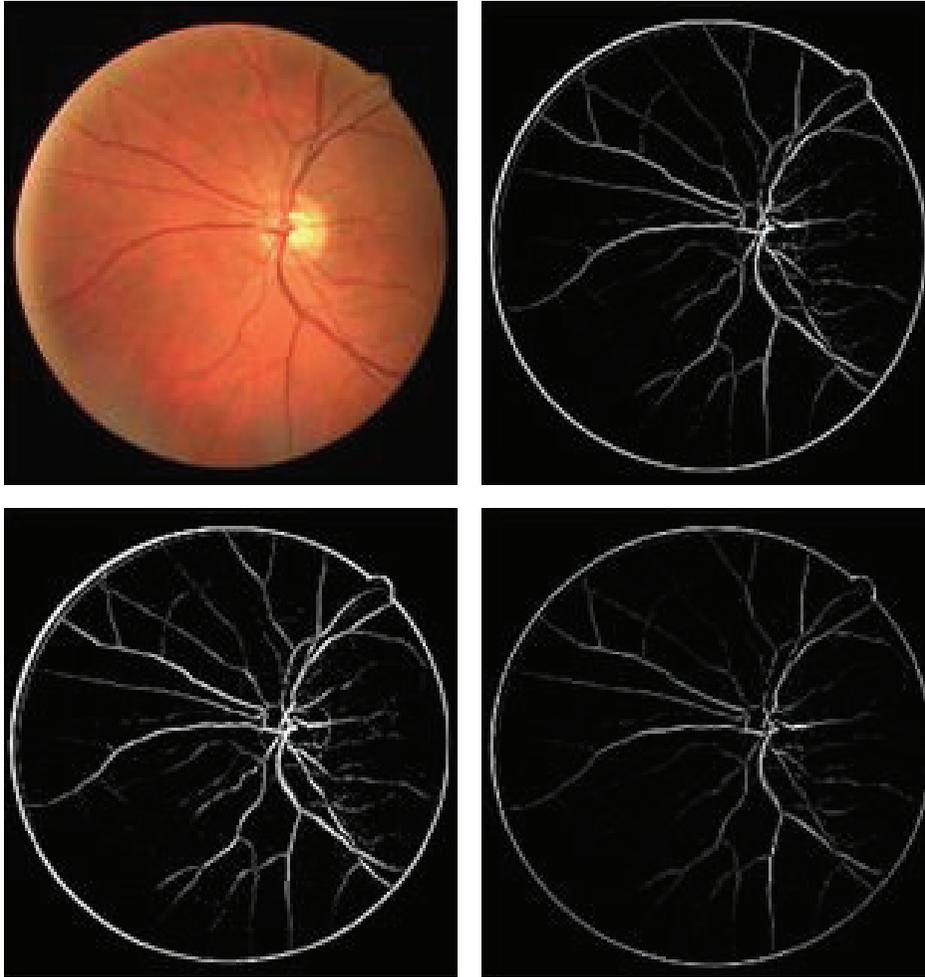


FIGURE 16: Real angiography image downloaded from DRIVE dataset [34], average response, maximum response, and minimum response (left to right-top to bottom).

TABLE 6: Comparison of vessel segmentation results on STARE project database [33].

Method	MAA	TPR	FPR
2nd human observer	0.9354	0.8949	0.0610
Hoover [33, 35]	0.9267	0.6751	0.0433
Mendonça (green) [21]	0.9440	0.6996	0.0270
Staal [25]	0.9516	0.6970	0.0190
Soares [44]	0.9480	0.7165	0.0252
Matched filter [13]	0.9384	0.6134	0.0245
Martinez-Perez [45]	0.9410	0.7506	0.0431
MF-FDOG [14]	0.9484	0.7177	0.0247
Proposed method	0.9402	0.6145	0.0162

recent methods in terms of TPR, FPR, and maximum accuracy average (MAA) where the maximal accuracy indicates

how to extract a binary image that matches the vessel images to a high degree. The accuracy is estimated by the ratio of the sum of the number of correctly classified foreground and background pixels, divided by the total number of pixels in the image. In this latest table, the performance measures of Staal et al. [25], Zhang et al. [14], Mendonça and Campilho [21], Chaudhuri et al. [13], Martinez-Perez et al. [45], and Hoover et al. [35] have been reported by their original papers. In addition, these performance results are the average values for the whole set of 20 images, except the method of Staal [25] which used 19 out of 20 images of the STARE images, among which ten were healthy and nine were unhealthy. Table 5 presents our results on all 20 images in the STARE database, estimated using the hand-labeled images set of the first human expert designated as a ground truth. The estimated experimental results are the average TPR = 0.6145 corresponding to an FPR of around 0.0162 and a maximum average accuracy MAA = 0.9402. The results show that our method has a competitive maximum average accuracy value where it performs better than the matched filter [13] and remains close to the others.

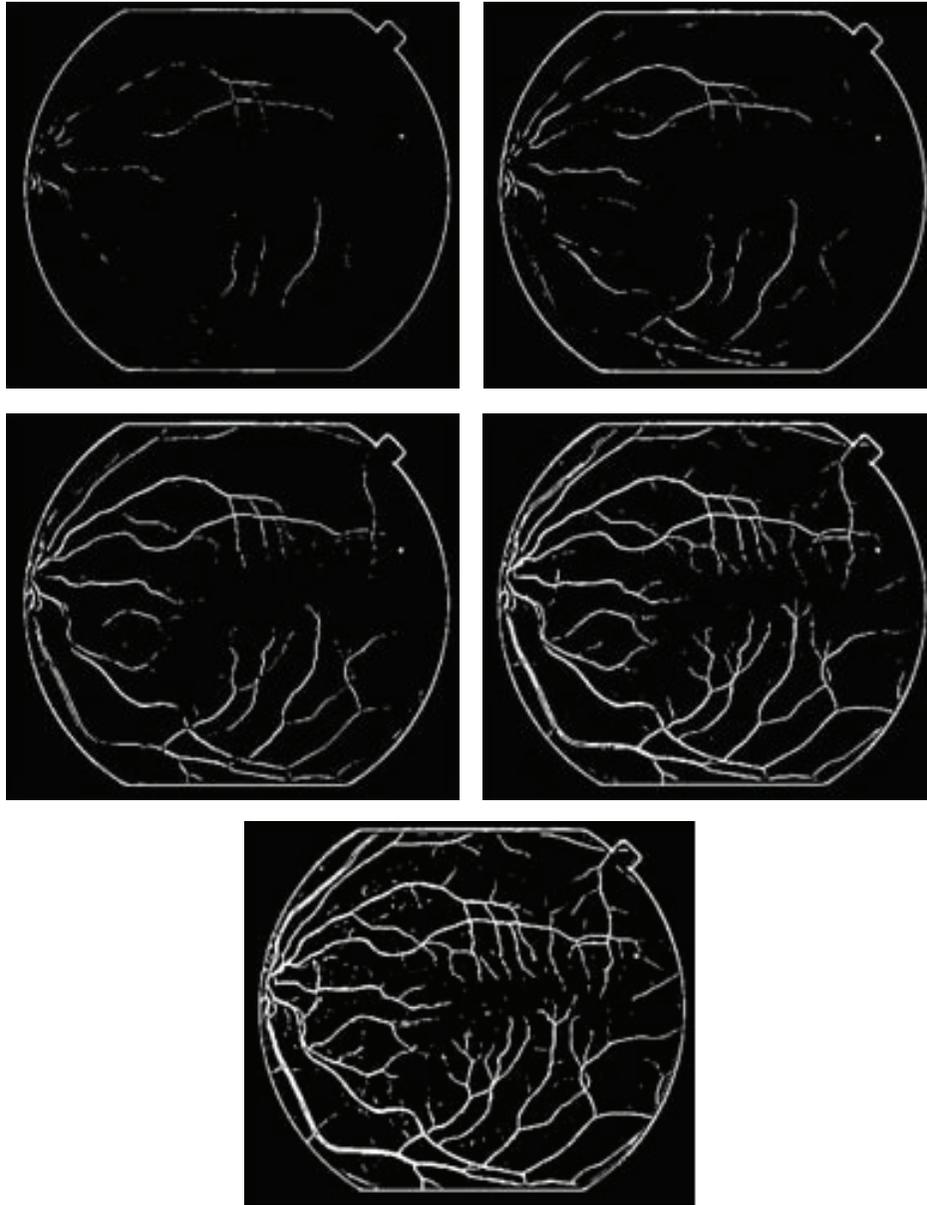


FIGURE 17: Different responses for different scales of Figure 14(a) (top to bottom); the first four images show the vesselness obtained at increasing scales. The last image is the result after the scale selection procedure (normalized image).

The results of the proposed method are also compared with those on twenty images from the DRIVE database, and the result is depicted in Table 7. The hand-labeled images by the first human expert have been used as ground truth. The experimental results show an MAA around of 0.9155. Also, we have compared the performance of the suggested technique with the sensitivities and specificities of the methods cited in Table 7. It has been found that for the DRIVE database the method has provided a sensitivity of 0.5879 and a specificity of 0.0166. We have shown that the proposed method performs well with a lower specificity even in the presence of lesions in the abnormal images.

4. Conclusion

The purpose of this work is to detect linear structures in real retinal images in order to help the interpretation of the vascular network. We put forward to combining an anisotropic diffusion filter to reduce the image noise with a multiscale response based on the eigenvectors of the Hessian matrix and on the gradient information to extract vessels from retinal images. The main advantage of this technique is its ability to extract large and fine vessels at various image resolutions. Furthermore, the directional anisotropic diffusion plays a vital role in denoising images and in decreasing

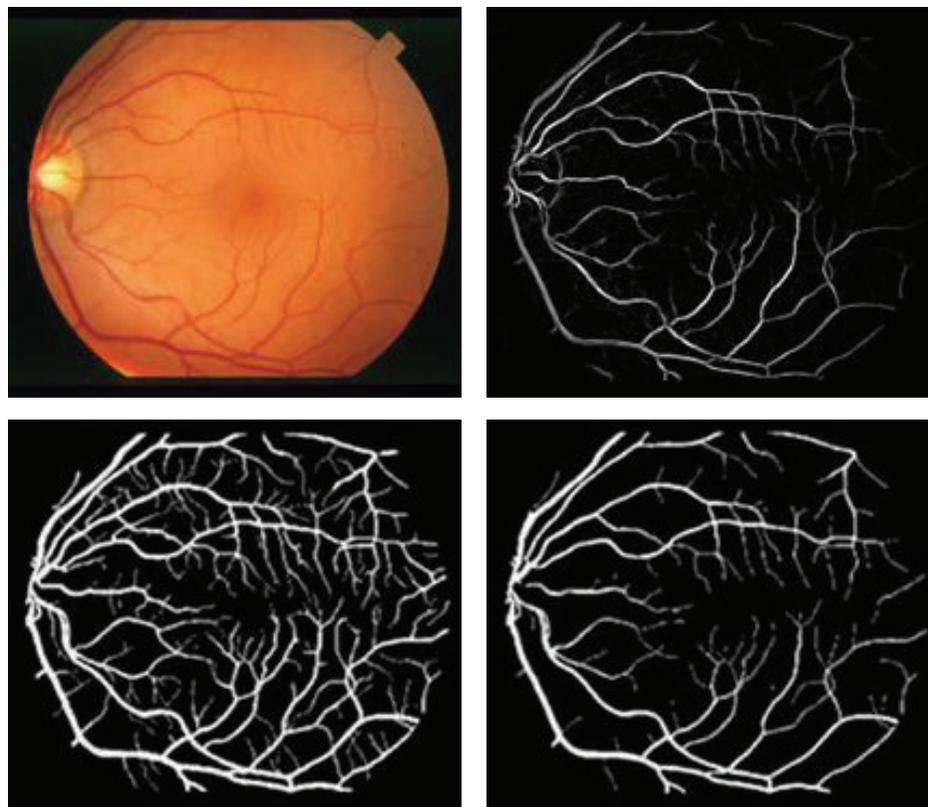


FIGURE 18: An image of a retina [35], the segmented image, and the hand labeled “truth” images (im0077.vk and im0077.ah) (left to right-top to bottom) [33].

TABLE 7: Comparison of vessel segmentation results on DRIVE database [34].

Method	MAA	TPR	FPR
2nd human observer [34]	0.9473	0.7761	0.0275
Martinez-Perez [45]	0.9344	0.7246	0.0345
Staal [25, 34]	0.9442	0.7194	0.0227
Mendonça [21]	0.9452	0.7344	0.0236
Matched filter [13]	0.9284	0.6168	0.0259
Niemeijer [34, 46]	0.9417	0.6898	0.0304
Proposed method	0.9155	0.5879	0.0166

the difficulty of vessel extraction especially for thin vessels. Our first results show the robustness of the method against noise as well as its applicability to detect blood vessels. The MAA is used as a performance measure, and the values achieved with our algorithm are competitive compared to the existing methods. Therefore, from the experimental results, it can be seen that the number of classified pixels has been slightly lower compared to the other methods using the same database mainly due to the weakness of blood vessels, causing missing vessels, and also because of lesions, resulting in a detection error. Also, the retinal images suffer from nonuniform illumination and have a poor contrast. Thus,

to avoid wrong classified pixels or miss classified ones, caused by an occasional false measurement, this system can very well be improved in the future with adding, for instance, some postprocessing tasks to reach more accurate measurement for blood vessels.

Conflict of Interests

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

References

- [1] R. Williams, M. Airey, H. Baxter, J. Forrester, T. Kennedy-Martin, and A. Girach, “Epidemiology of diabetic retinopathy and macular oedema: a systematic review,” *Eye*, vol. 18, no. 10, pp. 963–983, 2004.
- [2] R. Gupta and P. Kumar, “Global diabetes landscape-type 2 diabetes mellitus in South Asia: epidemiology, risk factors, and control,” *Insulin*, vol. 3, no. 2, pp. 78–94, 2008.
- [3] J. Malek, M. Ben Abdallah, A. Mansour, and R. Tourki, “Automated optic disc detection in retinal images by applying region-based active contour model in a variational level set formulation,” in *Proceedings of the International Conference on Computer Vision in Remote Sensing (CVRS '12)*, pp. 39–44, Xiamen, China, December 2012.
- [4] J. Malek and R. Tourki, “Blood vessels extraction and classification into arteries and veins in retinal images,” in *Proceedings*

- of the 10th International Multi-Conference on Systems, Signals & Devices (SSD '13), pp. 1–6, Hammamet, Tunisia, March 2013.
- [5] J. Malek and R. Tourki, "Inertia-based vessel centerline extraction in retinal image," in *Proceedings of the International Conference on Control, Decision and Information Technologies (CoDIT '13)*, pp. 378–381, Hammamet, Tunisia, May 2013.
 - [6] M. Al-Rawi, M. Qutaishat, and M. Arrar, "An improved matched filter for blood vessel detection of digital retinal images," *Computers in Biology and Medicine*, vol. 37, no. 2, pp. 262–267, 2007.
 - [7] M. E. Tyler, L. D. Hubbard, K. Boydston, and A. J. Pugliese, "Characteristics of digital fundus camera systems affecting tonal resolution in color retinal images," *The Journal of Ophthalmic Photography*, vol. 31, no. 1, pp. 1–9, 2009.
 - [8] T. W. Hansen, J. Jeppesen, S. Rasmussen, H. Ibsen, and C. Torp-Pedersen, "Ambulatory blood pressure and mortality: a population-based study," *Hypertension*, vol. 45, no. 4, pp. 499–504, 2005.
 - [9] T. Teng, M. Lefley, and D. Claremont, "Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy," *Medical and Biological Engineering and Computing*, vol. 40, no. 1, pp. 2–13, 2002.
 - [10] M. Ben Abdallah, J. Malek, R. Tourki, J. E. Monreal, and K. Krissian, "Automatic estimation of the noise model in fundus images," in *Proceedings of the 10th International Multi-Conference on Systems, Signals & Devices (SSD '13)*, pp. 1–5, Hammamet, Tunisia, March 2013.
 - [11] N. Patton, T. M. Aslam, T. MacGillivray et al., "Retinal image analysis: concepts, applications and potential," in *Progress in Retinal and Eye Research*, vol. 25, pp. 99–127, 2006.
 - [12] T. Walter, J.-C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina," *IEEE Transactions on Medical Imaging*, vol. 21, no. 10, pp. 1236–1243, 2002.
 - [13] 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.
 - [14] B. Zhang, L. Zhang, L. Zhang, and F. Karray, "Retinal vessel extraction by matched filter with first-order derivative of Gaussian," *Computers in Biology and Medicine*, vol. 40, pp. 438–445, 2010.
 - [15] J. Malek, A. T. Azar, and R. Tourki, "Impact of retinal vascular tortuosity on retinal circulation," *Neural Computing and Applications*, 2014.
 - [16] M. Sofka and C. V. Stewar, "Retinal vessel extraction using multiscale matched filters confidence and edge measures," Tech. Rep., Department of Computer Science, Rensselaer Polytechnic Institute, 2005.
 - [17] Y. Sato, S. Nakajima, N. Shiraga et al., "Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," *Medical Image Analysis*, vol. 2, no. 2, pp. 143–168, 1998.
 - [18] R. M. Rangayyan, F. Oloumi, F. Oloumi, P. Eshghzadeh-Zanjani, and F. J. Ayres, "Detection of blood vessels in the retina using Gabor filters," in *Proceedings of the 20th Canadian Conference on Electrical and Computer Engineering (CCECE '07)*, pp. 717–720, Vancouver, Canada, April 2007.
 - [19] T. Pock, C. Janko, R. Beichel, and H. Bischof, "Multiscale medialness for robust segmentation of 3D tubular structures," in *Proceedings of the 10th Computer Vision with Workshop*, The Austrian Science Fund (FWF) under the grants P17066-N04, February 2005.
 - [20] M. G. Cinsdikici and D. Aydin, "Detection of blood vessels in ophthalmoscope images using MF/ant (matched filter/ant colony) algorithm," *Computer Methods and Programs in Biomedicine*, vol. 96, no. 2, pp. 85–95, 2009.
 - [21] A. M. Mendonça and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," *IEEE Transactions on Medical Imaging*, vol. 25, no. 9, pp. 1200–1213, 2006.
 - [22] F. Zana and J.-C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1010–1019, 2001.
 - [23] M. M. Fraz, M. Y. Javed, and A. Basit, "Evaluation of retinal vessel segmentation methodologies based on combination of vessel centerlines and morphological processing," in *Proceedings of the 4th IEEE International Conference on Emerging Technologies (ICET '08)*, pp. 232–236, Rawalpindi, Pakistan, October 2008.
 - [24] L. Gang, O. Chutatape, and S. M. Krishnan, "Detection and measurement of retinal vessels in fundus images using amplitude modified second-order Gaussian filter," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 2, pp. 168–172, 2002.
 - [25] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, 2004.
 - [26] P. C. Siddalingaswamy, "Automatic detection of multiple oriented blood vessels in retinal images," *Journal of Biomedical Science and Engineering*, vol. 3, pp. 101–107, 2010.
 - [27] S. Dua, N. Kandiraju, and H. W. Thompson, "Design and implementation of a unique blood-vessel detection algorithm towards early diagnosis of diabetic retinopathy," in *Proceedings of the IEEE International Conference on Information Technology: Coding and Computing*, vol. 1, pp. 26–31, April 2005.
 - [28] A. Frangi, *Three-dimensional model-based analysis of vascular and cardiac images [Ph.D. thesis]*, Utrecht University, Utrecht, The Netherlands, 2001.
 - [29] C. Lorenz, I.-C. Carlsen, T. M. Buzug, C. Fassnacht, and J. Weese, "Multi scale line segmentation with automatic estimation of width, contrast and tangential direction in 2d and 3d medical images," in *Proceedings of the 1st Joint Conference Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery (CVRMed-MRCAS '97)*, pp. 233–242, Springer, London, UK, 1997.
 - [30] J. L. Federman, P. Gouras, H. Schubert et al., "Systemic diseases," in *Retina and Vitreous: Textbook of Ophthalmology*, S. M. Podos and M. Yano, Eds., vol. 9, pp. 7–24, Mosby, St. Louis, Mo, USA, 1994.
 - [31] W. Huang, *Automatic Detection and Quantification of Blood Vessels in the Vicinity of the Optic Disc in Digital Retinal Images*, University of Waikato, 2006.
 - [32] <http://reference.medscape.com/features/slideshow/retina>.
 - [33] A. Hoover, "STARE database," <http://www.ces.clemson.edu/~ahoover/stare/>.
 - [34] M. Niemeijer and B. van Ginneken, 2002, <http://www.isi.uu.nl/Research/Databases/DRIVE/results.php>.
 - [35] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical Imaging*, vol. 19, no. 3, pp. 203–210, 2000.

- [36] K. Krissian, "Flux-based anisotropic diffusion applied to enhancement of 3-D angiogram," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1440–1442, 2002.
- [37] C. Blondel, *Modelisation 3D et 3D + t des arteres coronaires a partir de sequences rotationnelles de projections rayons X [Ph.D. thesis]*, University of Nice Sophia Antipolis, Nice, France, 2004.
- [38] M. Ben Abdallah, J. Malek, R. Tourki, and K. Krissian, "Restoration of retinal images using anisotropic diffusion like algorithms," in *Proceedings of the International Conference on Computer Vision in Remote Sensing (CVRS '12)*, pp. 116–121, Xiamen, China, December 2012.
- [39] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Troussset, "Model-based detection of tubular structures in 3D images," *Computer Vision and Image Understanding*, vol. 80, no. 2, pp. 130–171, 2000.
- [40] G.-H. Cottet and L. Germain, "Image processing through reaction combined with nonlinear diffusion," *Mathematics of Computation*, vol. 61, no. 204, pp. 659–673, 1993.
- [41] J. Weickert, "Scale-space properties of nonlinear diffusion filtering with a diffusion tensor," Tech. Rep. 110, University of Kaiserslautern, Kaiserslautern, Germany, 1994.
- [42] J. Bigun, G. H. Granlund, and J. Wiklund, "Multidimensional orientation estimation with applications to texture analysis and optical flow," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 775–790, 1991.
- [43] M. Ben Abdallah, M. Jihene, K. Krissian, and R. Tourki, "An automated vessel segmentation of retinal images using multi-scale vesselness," in *Proceedings of the 8th International Multi-Conference on Systems, Signals & Devices*, IEEE, 2011.
- [44] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar Jr., H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," *IEEE Transactions on Medical Imaging*, vol. 25, no. 9, pp. 1214–1222, 2006.
- [45] M. E. Martinez-Perez, A. D. Hughes, S. A. Thom, A. A. Bharath, and K. H. Parker, "Segmentation of blood vessels from red-free and fluorescein retinal images," *Medical Image Analysis*, vol. 11, no. 1, pp. 47–61, 2007.
- [46] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog, and M. D. Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," in *Proceedings of the SPIE Medical Imaging*, M. Fitzpatrick and M. Sonka, Eds., vol. 5370, pp. 648–656, 2004.
- [47] X. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 1, pp. 131–137, 2003.
- [48] T. Lindeberg and D. Fagerström, "Scale-space with casual time direction," in *Proceedings of the 4th European Conference on Computer Vision (ECCV '96)*, pp. 229–240, Cambridge, UK, April 1996.
- [49] W. Changhua, G. Agam, and P. Stanchev, "A general framework for vessel segmentation in retinal images," in *Proceedings of the International Symposium on Computational Intelligence in Robotics and Automation (CIRA '07)*, June 2007.
- [50] X. Qian, M. P. Brennan, D. P. Dione et al., "A non-parametric vessel detection method for complex vascular structures," *Medical Image Analysis*, vol. 13, no. 1, pp. 49–61, 2008.
- [51] L. Wang, A. Bhalerao, and R. Wilson, "Analysis of retinal vasculature using a multiresolution hermite model," *IEEE Transactions on Medical Imaging*, vol. 26, no. 2, pp. 137–152, 2007.
- [52] A. P. Witkin, "Scale-space filtering: a new approach to multi scale description," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '84)*, pp. 150–153, March 1984.

Research Article

Transverse Strains in Muscle Fascicles during Voluntary Contraction: A 2D Frequency Decomposition of B-Mode Ultrasound Images

James M. Wakeling and Avleen Randhawa

Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC, Canada V5A 1S6

Correspondence should be addressed to James M. Wakeling; wakeling@sfu.ca

Received 8 April 2014; Accepted 3 September 2014; Published 28 September 2014

Academic Editor: Yicong Zhou

Copyright © 2014 J. M. Wakeling and A. Randhawa. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

When skeletal muscle fibres shorten, they must increase in their transverse dimensions in order to maintain a constant volume. In pennate muscle, this transverse expansion results in the fibres rotating to greater pennation angle, with a consequent reduction in their contractile velocity in a process known as gearing. Understanding the nature and extent of this transverse expansion is necessary to understand the mechanisms driving the changes in internal geometry of whole muscles during contraction. Current methodologies allow the fascicle lengths, orientations, and curvatures to be quantified, but not the transverse expansion. The purpose of this study was to develop and validate techniques for quantifying transverse strain in skeletal muscle fascicles during contraction from B-mode ultrasound images. Images were acquired from the medial and lateral gastrocnemii during cyclic contractions, enhanced using multiscale vessel enhancement filtering and the spatial frequencies resolved using 2D discrete Fourier transforms. The frequency information was resolved into the fascicle orientations that were validated against manually digitized values. The transverse fascicle strains were calculated from their wavelengths within the images. These methods showed that the transverse strain increases while the longitudinal fascicle length decreases; however, the extent of these strains was smaller than expected.

1. Introduction

Muscle fibres are nearly incompressible [1] and so must increase in girth as they shorten. This transverse expansion requires that the fibres in pennate muscle rotate to greater pennation angles during shortening to ensure they still pack together in close proximity [2, 3]. Due to these fibre rotations, the fibres shorten at a lower velocity than the muscle belly in a process known as muscle belly gearing [4–6], promoting a greater force and power production from the muscle. Understanding how muscle fibres change shape during contraction will elucidate how structural mechanisms affect the functional output of muscle.

It is commonly assumed that whole muscles maintain a constant volume during contraction, in a similar manner to their constituent fibres. Many muscle models are one-dimensional and thus assume constant thickness of the muscle

belly [3, 7–10]. Some studies have modelled muscle as 2D structures but have implemented a constant area assumption [6, 11–13], and other studies of 3D properties implement constant or nearly incompressible volumes [14–17]. However, whole muscle may change volume to a greater extent than fibres, due to a variable amount of blood that may pool and be pumped out from the muscle, through the action of the muscle contraction [18]. Additionally, recent imaging studies have shown that muscle varies in thickness in a complex and muscle-specific manner. Muscle thickness is not constant during contraction and can vary for both isometric [3] and dynamic [5, 6, 19] contractions.

The belly gearing within a muscle can vary according to the mechanical demands of the contraction [5, 6, 19], and so the fibres may exist at a range of different pennation angles for a given fibre length: it has been suggested that this variability is related to the load and stretch of connective

tissue such as aponeurosis [6, 15, 19]. Currently we do not know how the changes in transverse fibre dimensions relate to dynamic changes in fibre geometry, belly gearing, and thus the functional output of the muscle.

Recent developments in muscle imaging have allowed the longitudinal properties of muscle fibres to be determined. Diffusion-tensor MRI can identify the longitudinal direction of muscle fibres [20–24] allowing fibres to be tracked and their length calculated. However, MRI imaging necessitates prolonged, isometric contractions and thus is not suitable for dynamic studies of gearing. B-mode ultrasound imaging of muscle allows the fascicles to be imaged at faster rates, and a range of automated approaches have been used to quantify the fascicle lengths and even curvatures during dynamic contractions [25–28]. Neither these MRI nor ultrasound techniques have allowed the width of the fascicles, or a measure of their transverse expansion, to be quantified. However, muscle fascicles appear as nearly parallel, repeating bands within B-mode ultrasound images and the width of these bands can potentially be extracted by frequency decomposition of the images. The purpose of this study was to develop a method to extract information on the spatial frequencies of the fascicular structure within the muscle belly from B-mode ultrasound images and to relate this information to the fascicle size, shape, and orientations during muscle contraction.

2. Methods

2.1. Subjects. Six male subjects took part in this study (age 28.8 ± 5.5 years; mass 78.3 ± 6.2 kg; height 178 ± 2.3 cm; mean \pm SD). All subjects provided informed consent in accordance with requirements from the University Office of Research Ethics.

2.2. Data Collection. Images were acquired from the medial gastrocnemius (MG) and lateral gastrocnemius (LG) of the right leg during ankle plantarflexion contractions. Subjects were seated on a dynamometer (System 3, Biodex, New York, USA) with their knee held at 135° , their shank horizontal, and their foot secured to a footplate on the dynamometer. The central axis of the dynamometer was aligned to meet the axis through the medial and lateral malleoli. Subjects performed cyclic ankle extensions against isotonic loads, in time to the beat of a metronome. The dorsiflexion torque was limited to 0.5 N m, and three plantarflexion conditions were presented (5@0.42, 25@0.35, and 5@16: torque [N m] @ cycle frequency [Hz]); each condition had a 15° range of motion from 5 to 20° plantarflexion. Each trial consisted of 10 cycles of contraction, from which the middle 5 were analyzed.

The MG and LG muscle bellies were imaged using 128-element (60 mm width) linear array B-mode ultrasound probes (Echoblaster 128, Telemed, Lithuania), scanning at 40 Hz. The probes were aligned to the fascicle planes to obtain nearly continuous lines for the fascicles in each image. The probes were secured to the leg using custom mounts with adhesive and elasticized bandages. The probes were measured from the MG and LG simultaneously, and were synchronized to the position P and torque data T from the

dynamometer (recorded at 1000 Hz: USB-6229, National Instruments, Austin, TX, USA).

2.3. Image Analysis. Each ultrasound frame formed a square image $f(x, y)$ of $N = 512$ pixels per side with each greyscale pixel indexed by its x - and y -coordinate (Figure 1). Images were manually digitized (ImageJ software, NIH, Maryland, USA) to identify three coordinates on the superficial aponeurosis, three coordinates on the deep aponeurosis, and two coordinates on a representative fascicle. Aponeuroses were described using second-order polynomials that were fit to both the superficial and the deep coordinates using least-squares minimization, and the muscle belly thickness L_y was calculated as the mean distance between the aponeuroses. The fascicle inclination θ_d within the fascicle plane was given by the angle between the x -axis and the vector between the two digitized fascicle points. The fascicle length L_f was given by the length of the linear line passing through the fascicle coordinates that intersected the best-fit linear lines through the superficial and deep aponeuroses.

Each image was filtered using multiscale vessel-enhancement filtering. This method enhances the tubular structures in the image that are formed by the fascicles [29] and is capable of resolving tubular structures of different radii and has previously been applied to B-mode ultrasound images [25, 26, 30]. Here we followed Rana and coworkers [25] by using scales of 1.5, 2, 2.5, and 3. The region of interest was taken as the area of muscle tissue within the filtered image that was bounded by the aponeuroses, and a strip, 10 pixels wide, was removed inside the aponeuroses to ensure that the region of interest contained no features that were aligned with the aponeuroses.

Muscle fascicles appear as dark lines in the image and connective tissue between the fascicles appears as bright structures that parallel the fascicles [25]. The striped nature of the fascicles is enhanced and retained within the filtered image and is characterized by the spatial frequency of the stripes. The spatial frequencies $F(u, v)$ of the filtered image $f(x, y)$ were determined by a 2D discrete Fourier analysis of the region of interest, where

$$\begin{aligned} F(u, v) &= \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \\ &\times \left[\cos\left(\frac{2\pi(ux + vy)}{N}\right) + j \sin\left(\frac{2\pi(ux + vy)}{N}\right) \right], \\ j &= \sqrt{-1}. \end{aligned} \quad (1)$$

The amplitude spectra for the region of interest describe the amplitudes of the pixel intensities across a range of frequencies:

$$|F(u, v)| = \sqrt{\text{Re}^2(u, v) + \text{Im}^2(u, v)}. \quad (2)$$

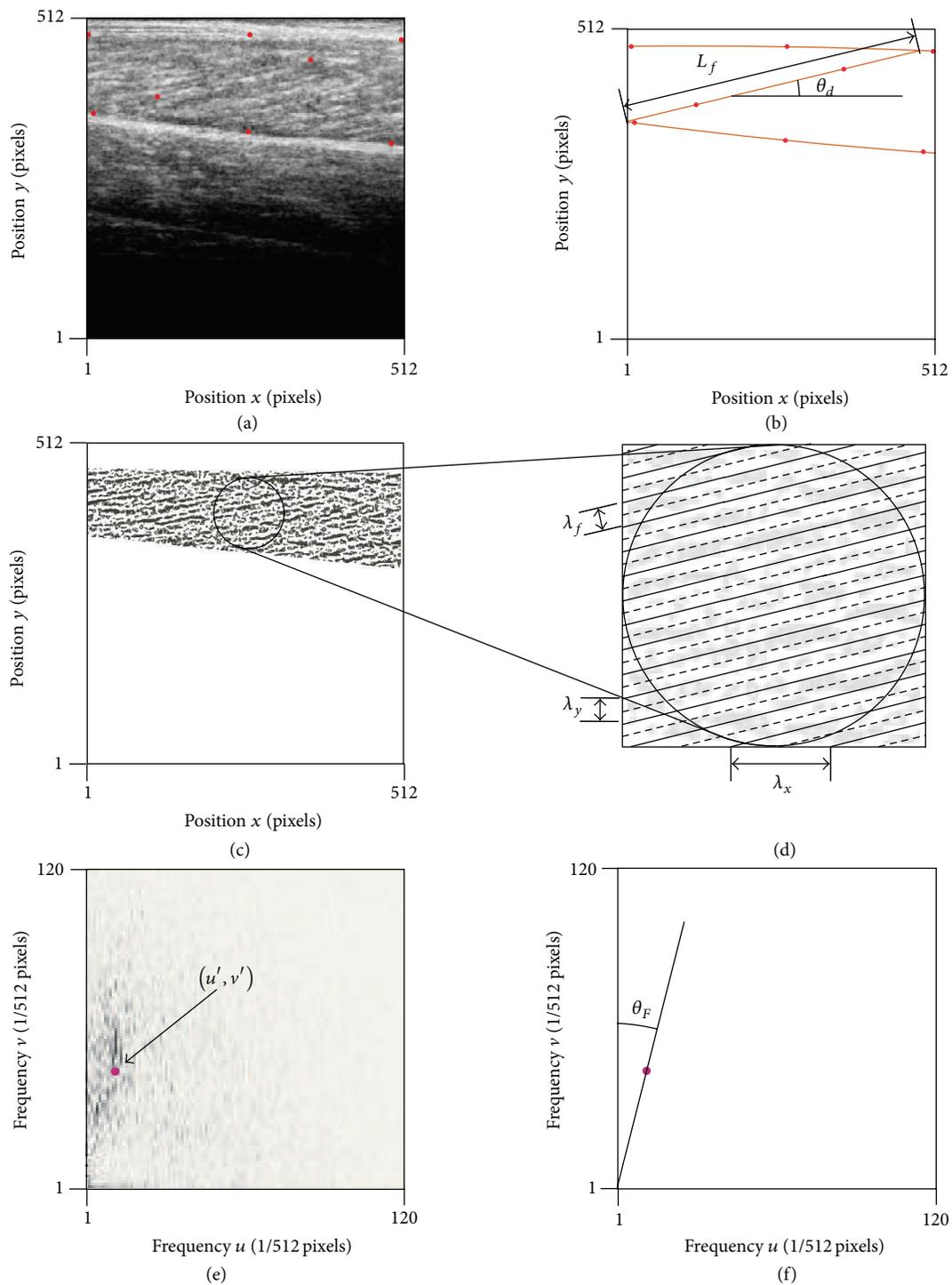


FIGURE 1: *The Analysis Method.* Coordinates on the superficial and deep aponeuroses and a representative fascicle were manually digitized (red points) for each ultrasound image (a) and used to calculate the fascicle length L_f and inclination angle θ_d relative to the x-axis (b). Fascicles were enhanced using multiscale vessel enhancement filters and the region of interest was determined between the aponeuroses (c). The repeating fascicular nature is described by its transverse wavelength λ_f , which can be resolved into wavelengths λ_u and λ_v that are in the x- and y-directions, respectively (d). The spatial frequencies u and v were determined using a discrete Fourier transform, and the major spatial frequencies characterized by the moment of frequency (u', v') (e). The inclination angle θ_F were determined from the major spatial frequencies (f), and the wavelengths λ_u and λ_v were calculated from these frequencies.

This is reduced to a single frequency value for each direction (u' and v') using the m th moment of frequency:

$$\begin{aligned} u' &= \frac{\sum_{u=b_1}^{b_2} \sum_{v=b_1}^{b_2} [|F(u, v)|^m u]}{\sum_{u=b_1}^{b_2} \sum_{v=b_1}^{b_2} |F(u, v)|^m}, \\ v' &= \frac{\sum_{v=b_1}^{b_2} \sum_{u=b_1}^{b_2} [|F(u, v)|^m v]}{\sum_{v=b_1}^{b_2} \sum_{u=b_1}^{b_2} |F(u, v)|^m}. \end{aligned} \quad (3)$$

The moments of frequency were calculated across the frequency range of $b_1 = 4$ to $b_2 = 120$, and this contained more than 99% of the power of the amplitude spectra.

The wavelengths for the fascicle stripes λ_x and λ_y were given by

$$\lambda_x = \frac{1}{u'}, \quad \lambda_y = \frac{1}{v'}. \quad (4)$$

The dominant repeating structure within the region of interest is given by the muscle fascicles. For large inclinations ($\approx 90^\circ$) λ_x would be small and λ_y would be large; conversely, for small inclinations ($\approx 0^\circ$) λ_x would be large and λ_y would be small. The fascicle inclination θ_F (relative to the x -direction) can be determined from the Fourier analysis as follows:

$$\theta_F = \text{ArcTan} \left(\frac{\lambda_y}{\lambda_x} \right). \quad (5)$$

The best value for m was determined by comparing θ_d and θ_F for $1 \leq m \leq 8$ (see Statistics section below).

The wavelengths λ_x and λ_y reflect the dominant characteristics of the repeated fascicles in the region of interest. They provide information not only on the fascicle inclination, but also on the wavelength of the ‘‘fascicle stripes’’ λ_f , that is, the wavelength of the stripes in a transverse direction across the fascicles:

$$\lambda_f = \lambda_x \text{Sin } \theta_F. \quad (6)$$

The muscle belly thickness L_y , fascicle length L_f , and the wavelength λ_f were normalized by their respective means that occurred throughout the five contraction cycles to yield normalized terms \hat{L}_y , \hat{L}_f , and $\hat{\lambda}_f$, respectively.

Pixel brightnesses in each ultrasound image are a measure of the echogenicity of the material being scanned. It is possible that, as the muscle expands in a transverse direction, there is an uneven expansion of fascicular and connective tissue. This would result in a change in the distribution of pixel brightness within the region of interest. This possibility was examined by quantifying the mean pixel brightness B_p within the region of interest.

2.4. Statistics. The best value for m was determined from the correlation coefficient r and the root-mean-square error (RMSE) between θ_d and θ_F for each contraction sequence. The effect of m on r and RMSE was determined with ANOVA with subject (random), muscle, and condition as factors (Minitab v16, Minitab Inc., State College, PA, USA).

For each condition the time ω was normalized to each contraction (0 to 360°), with 0° occurring at the midpoint of each dorsiflexion movement. The parameters P , T , \hat{L}_f , \hat{L}_y , θ_d , θ_F , $\hat{\lambda}_f$, and B_p were each described by a Fourier series of the form

$$c_1 + a_1 \text{Sin}(\phi_1 + \omega) + a_2 \text{Sin}(\phi_2 + 2\omega), \quad (7)$$

where the coefficients c_1 , a_1 , and ϕ_1 describe the mean value, the amplitude, and the phase for the first harmonic. The effect of subject (random), muscle, condition, parameter, and muscle-by-parameter interaction on these Fourier coefficients was determined using ANOVA.

Statistical tests were considered significant at the $\alpha = 0.05$ level. Mean values are reported as mean \pm standard error of the mean.

3. Results

Subjects performed a series of isotonic plantarflexions (Figure 2), with the ankle plantarflexor torque increasing during each plantarflexion. The fascicle length within the medial and lateral gastrocnemius shortened during each plantarflexion, and this coincided with an increase in the inclination angle of the fascicles. During fascicle shortening the thickness of both the fascicles and muscle belly increased, with the relative increases in the muscle belly thickness being greater than those for the fascicles. During each contraction cycle, the pixel intensity within the region of interest varied, with the lowest intensities occurring when the fascicles were shortest but thickest.

The estimates of the inclination angle based on the Fourier transform were dependent on the moment of frequency, m , selected. There was no significant effect of m on the correlation between the inclination angle determined by manual digitization, θ_d , and the inclination angle determined from the discrete Fourier transform, θ_F ; however, there was a significant effect of m on the root-mean-square error between these values (Figure 3). A value of $m = 5$ resulted to be close to the greatest correlation and the lowest RMSE and so was selected for further analysis. When considered across all subjects, muscles, and contraction conditions, the RMSE for $m = 5$ was 3.4° . The error between the two measures of inclination was partly due to the smaller amplitude of change in inclination for the θ_F than for θ_d .

The ANOVA showed there was a significant effect of the muscle, subject, parameter, and muscle-by-parameter interaction on the amplitude of the cyclic changes, a_1 . The main effects from the ANOVA (Figure 4) showed that the a_1 for MG was 0.41 greater than for LG. The interaction effect showed that a_1 for θ_d for the MG was greater than for the LG; however, this effect was not seen for θ_F . The magnitude of the parameter effects on a_1 can be seen in Tables 1 and 2.

The ANOVA showed there was a significant effect of the subject, parameter, and muscle-by-parameter interaction on the phase of the cyclic changes, ϕ_1 . The interaction effect showed that ϕ_1 for $\hat{\lambda}_f$ was slightly smaller and for \hat{L}_y was slightly larger for the MG than for the LG: in other words, there was a greater phase difference between the cycles of

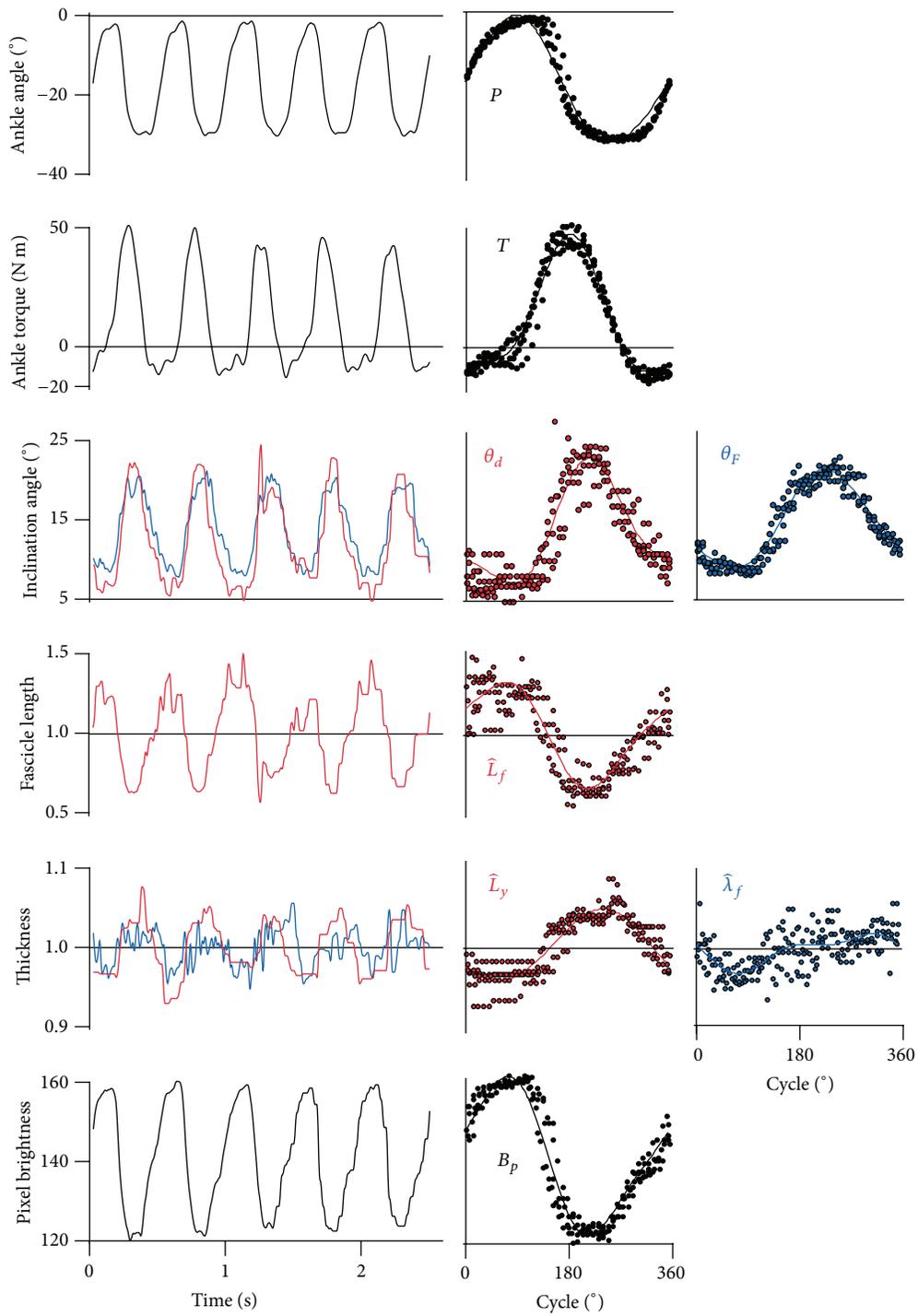


FIGURE 2: Muscle Structural Data for the Lateral Gastrocnemius during One Set of Ankle Plantarflexions. The first column shows the parameters changing over five contraction cycles. The middle and right columns show these data expressed as a percentage of the contraction cycle and show individual points as well as the smooth model calculated from Fourier series. Parameters determined from manual digitization are shown in red, and parameters determined following the discrete Fourier transform are shown in blue. $\hat{\lambda}_f$ was calculated with $m = 5$.

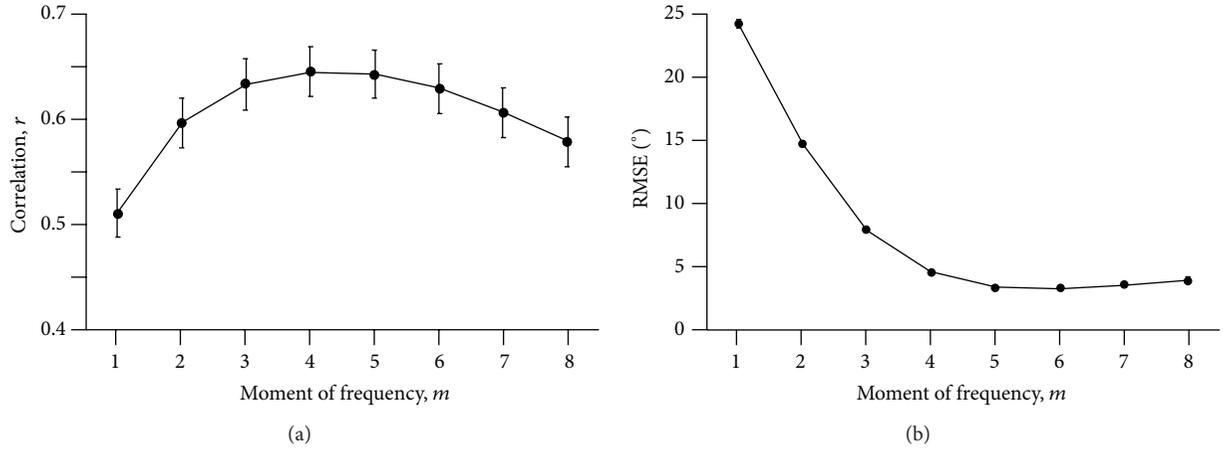


FIGURE 3: Error Analysis for the Selection of the Moment of Frequency, m . The correlation between the inclination angle determined by manual digitization, θ_d , and the inclination angle determined from the discrete Fourier transform, θ_F , is shown for different values of m (a). The root-mean-square error between these terms is shown in (b). Points show the mean \pm SEM ($n = 99$).

TABLE 1: Fourier coefficients for the medial gastrocnemius.

Parameter	MG		
	c_1	a_1	o_1
P	$-13.06 \pm 0.20^\circ$	$7.93 \pm 0.19^\circ$	$-1.4 \pm 2.1^\circ$
T	7.59 ± 1.05 N m	21.30 ± 1.91 N m	$252.7 \pm 3.4^\circ$
\widehat{L}_f	1.00 ± 0.00	0.22 ± 0.02	$29.7 \pm 2.9^\circ$
$\widehat{\lambda}_f$	1.00 ± 0.00	0.017 ± 0.002	$182.6 \pm 9.2^\circ$
θ_d	$17.10 \pm 0.45^\circ$	$5.34 \pm 0.46^\circ$	$215.0 \pm 2.8^\circ$
θ_F	$15.86 \pm 0.27^\circ$	$1.64 \pm 0.14^\circ$	$217.8 \pm 3.6^\circ$
\widehat{L}_y	1.00 ± 0.00	0.023 ± 0.003	$286.7 \pm 13.2^\circ$
B_p	146.27 ± 4.72	7.21 ± 0.48	$38.4 \pm 4.3^\circ$

Values are shown as mean \pm SEM ($n = 6$) and are for the first harmonic of the Fourier series.

fascicle thickness and muscle belly thickness for the MG than for the LG. The magnitude of the parameter effects on a_1 can be seen in Tables 1 and 2: there was no significant difference in the phase difference between \widehat{L}_y , θ_d , and θ_F . When the phase for fascicle length was offset by 180° ($o_1 + 180^\circ$) there was no significant difference between its phase and those for \widehat{L}_y , θ_d , and θ_F , and so the timing of fascicle length shortening exactly matches the increases in fascicle thickness.

4. Discussion

This study shows that there is information within B-mode images of the muscle bellies that has spatial frequencies that change in a cyclical manner during repeated contractions. These spatial frequencies are due to the fascicular (or vessel-like) structures within the muscle that were resolved by the multiscale vessel enhancement filtering [29]. 2D information from the images was retained by the 2D discrete Fourier transform of the images and allowed the inclination angle of the muscle fascicles to be determined, θ_F : this is a feature of the fascicles that could be validated against the manually determined inclination angles, θ_d (Figure 3).

TABLE 2: Fourier coefficients for the lateral gastrocnemius.

Parameter	LG		
	c_1	a_1	o_1
P	$-13.06 \pm 0.20^\circ$	$7.93 \pm 0.19^\circ$	$-1.4 \pm 2.1^\circ$
T	7.59 ± 1.05 N m	21.30 ± 1.91 N m	$252.7 \pm 3.4^\circ$
\widehat{L}_f	1.00 ± 0.00	0.19 ± 0.02	$36.2 \pm 5.4^\circ$
$\widehat{\lambda}_f$	1.00 ± 0.00	0.015 ± 0.001	$209.3 \pm 7.0^\circ$
θ_d	$9.23 \pm 0.57^\circ$	$2.54 \pm 0.34^\circ$	$219.7 \pm 5.4^\circ$
θ_F	$11.40 \pm 0.49^\circ$	$1.62 \pm 0.27^\circ$	$216.6 \pm 4.8^\circ$
\widehat{L}_y	1.00 ± 0.00	0.027 ± 0.003	$262.1 \pm 16.4^\circ$
B_p	148.62 ± 4.17	7.22 ± 0.96	$33.4 \pm 5.5^\circ$

Values are shown as mean \pm SEM ($n = 6$) and are for the first harmonic of the Fourier series.

A perfect match between θ_d and θ_F should not be expected. When ultrasound images are manually digitized, the inclination angles tend to reflect the dominant fascicle features within the image [25]. However, there is variation of fascicle orientations across each image [26] and sometimes nonfascicular features that may also occur in the image, and these features would influence the spatial frequencies determined by automated methods that consider the whole region of interest [25] such as the discrete Fourier transform as used in this study. The accuracy in these automated approaches can be maximized by careful selection of images that contain minimal nonfascicular structures [30] or by masking the undesired features within the region of interest. The inclinations θ_d and θ_F in this study measured the angles between the fascicles and the x -axis of the ultrasound images. By contrast, pennation angles measured in previous ultrasound studies are defined in different ways, for example, the angle between the fascicle and the superficial aponeurosis, or the deep aponeurosis or the mean direction of the superficial and deep aponeuroses [3, 31–33]. In this study, the θ_F was approximately 13° and 8° for the MG and LG, respectively, for an ankle plantarflexion angle of 5° (calculated from data

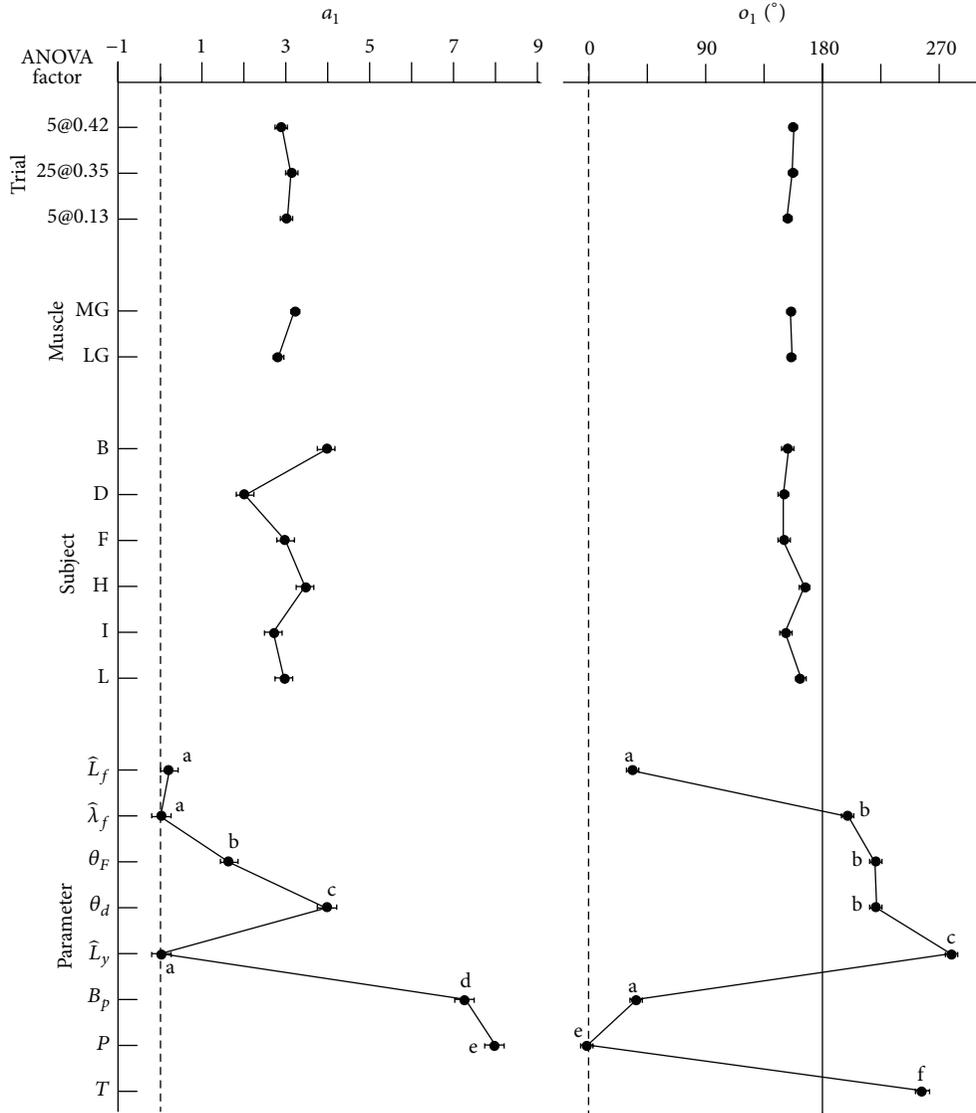


FIGURE 4: *Main Effects Determined by ANOVA.* The effects of the experimental factors on the coefficients a_1 and o_1 for the first harmonic in the Fourier series are shown. Points show the least-square mean \pm SEM as determined by the ANOVA. a_1 for torque was not tested due to its magnitude being much larger than the other variables. Letters next to the symbols indicate parameters that were not significantly different from each other, as calculated in a post hoc Tukey analysis.

in Tables 1 and 2), and these are approximately 5° smaller than the pennations reported for seated subjects in the same dynamometer [6].

The transverse wavelength of the fascicle strains $\hat{\lambda}_f$ changed in a cyclical manner, in time with the fascicle shortening (Tables 1 and 2; Figure 4). As the fascicles shorten their transverse strain increased. The mean pixel brightness also decreased as the transverse strain increased (Tables 1 and 2; Figures 2 and 4), indicating a greater proportion of darker elements in the image for higher $\hat{\lambda}_f$. Within the muscle bellies, the fascicles and connective tissue have different echogenicities, with the fascicles appearing darker. It is possible that the increase in B_p with increased $\hat{\lambda}_f$ indicates that the transverse strain in the muscle fascicles is underestimated by

$\hat{\lambda}_f$ (that includes elements from both fascicles and connective tissue).

The transverse strain was calculated by resolving λ_x and λ_y into the transverse direction. In theory, these values could be resolved into the longitudinal direction to provide a measure of longitudinal fascicle strain; however, this is practically not possible. In an ideal ultrasound image, all the fascicles would appear as continuous lines between the two aponeuroses: these would have a longitudinal spatial frequency of less than $1/512$ pixels and therefore would be beyond the resolution of the technique. In reality, fascicles appear as partial lines between the aponeuroses, and the exact length of each line segment is very sensitive to the exact orientation of the fascicles relative to the scanning plane. During contraction the fascicles can change their orientation

relative to the scanning plane [30], and thus fluctuations in line-length would reflect their 3D orientation as well as the fascicle length and therefore preclude measurements of fascicle length using these methods.

Assuming that the muscle fibres (and presumably the fascicles) maintain a constant volume during contraction [1], then they must increase in girth as they shorten. If an additional assumption is made that the increase in girth for the muscle fascicles is radially symmetrical then the muscle fascicles should have a Poisson ratio of 0.43 for a longitudinal strain of 0.2, where the Poisson ratio is the ratio of the transverse strain/longitudinal strain. The Poisson ratio ν can be calculated from this study as the ratio of (a_1 for $\hat{\lambda}_f$)/(a_1 for \hat{L}_f). The mean ν from this study was 0.09 ± 0.01 and was thus much smaller than expected. As discussed above, it is possible that $\hat{\lambda}_f$ is an underestimate, leading to low ν . An alternative estimate for the transverse strain for the fascicles can be calculated from the manually digitized parameters. If it is assumed that the entire muscle belly consists of fascicles that are parallel to each other, then the distance between the aponeuroses must equal the width of the fascicles acting in parallel, adjusted by their inclination; thus the normalized fascicle thickness will equal $\hat{L}_y / \cos \theta_a$. This alternative estimate of fascicle thickness yields a mean ν of 0.20 ± 0.02 that is still less than expected. Data from this study thus indicate that the increase in the transverse width of the fascicles does not meet that expected for isovolumetric muscle fibres that show radial symmetry in their expansion in girth. It will be necessary to investigate the fascicular expansion in the direction perpendicular to the scanning plane to identify the reasons for this discrepancy.

Changes to muscle belly thickness occur with changes in both fascicle thickness and fascicle rotations to different pennation angles [19], with the fascicle thickness and pennation angle being related to each other via intramuscular pressure, transverse forces, and compliance in connective tissues [15] such as aponeuroses and intramuscular connective tissue [34]. It is possible that the differences in whole-muscle bulging between MG and LG that have been reported in previous studies [3, 5, 6, 19, 32] may reflect differences in the direction of the transverse or perpendicular (out of plane) bulging of the fascicles, or due to differences in connective tissue properties and the tendency for the fascicles to rotate. In this study we found similar transverse expansion and Poisson ratio of the fascicles occurring in both the MG and LG (Tables 1 and 2), thus indicating that differences in the bulging of the muscle belly are caused more by differences in connective tissue properties and the tendency for the fascicles to rotate than by differences in the fascicle bulging *per se*.

This study describes a method to determine the transverse strain in the muscle fascicles during contraction and is the first study to describe these strains during dynamic and voluntary contractions. However, it should be noted that this methodological study has been constrained to a small set of contractions performed by male subjects: it will be important to understand how transverse bulging of the fascicles changes with both age and gender. Nonetheless, the results show that increases in transverse width are exactly timed with the

reductions in the longitudinal length of the fascicles. Surprisingly, the magnitude of the transverse strains, as imaged within the ultrasound scanning planes, appears smaller than expected. However, the imaging methods preclude the measurement of strains perpendicular to the ultrasound scans. Fully 3D studies are needed to explore the exact nature of shape changes to the fascicles during contraction and to relate these to the mechanisms of muscle contraction.

Conflict of Interests

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

Acknowledgments

Financial support was provided by an NSERC Discovery Grant to James M. Wakeling. The authors thank Eleanor Li for help during data collection.

References

- [1] R. J. Baskin and P. J. Paolini, "Volume change and pressure development in muscle during contraction," *American journal of physiology*, vol. 213, no. 4, pp. 1025–1030, 1967.
- [2] R. M. Alexander, *Animal Mechanics*, Blackwell Scientific, London, UK, 2nd edition, 1982.
- [3] C. N. Maganaris, V. Baltzopoulos, and A. J. Sargeant, "In vivo measurements of the triceps surae complex architecture in man: Implications for muscle function," *Journal of Physiology*, vol. 512, no. 2, pp. 603–614, 1998.
- [4] E. L. Brainerd and E. Azizi, "Muscle fiber angle, segment bulging and architectural gear ratio in segmented musculature," *The Journal of Experimental Biology*, vol. 208, no. 17, pp. 3249–3261, 2005.
- [5] J. M. Wakeling, O. M. Blake, I. Wong, M. Rana, and S. S. M. Lee, "Movement mechanics as a determinate of muscle structure, recruitment and coordination," *Philosophical Transactions of the Royal Society Series B: Biological Sciences*, vol. 366, no. 1570, pp. 1554–1564, 2011.
- [6] A. Randhawa, M. E. Jackman, and J. M. Wakeling, "Muscle gearing during isotonic and isokinetic movements in the ankle plantarflexors," *European Journal of Applied Physiology*, vol. 113, no. 2, pp. 437–447, 2013.
- [7] R. M. Alexander and A. Vernon, "The dimensions of knee and ankle muscles and the forces they exert," *Journal of Human Movement Studies*, vol. 1, pp. 115–123, 1975.
- [8] F. E. Zajac, "Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control," *Critical Reviews in Biomedical Engineering*, vol. 17, no. 4, pp. 359–411, 1989.
- [9] S. L. Delp and J. P. Loan, "A graphics-based software system to develop and analyze models of musculoskeletal structures," *Computers in Biology and Medicine*, vol. 25, no. 1, pp. 21–34, 1995.
- [10] A. J. Van den, D. Bogert, and D. Heinrich, "Implicit methods for efficient musculoskeletal simulation and optimal control," *Procedia IUTAM*, vol. 2, pp. 297–316, 2011.
- [11] C. J. Zuurbier and P. A. Huijting, "Influence of muscle geometry on shortening speed of fibre, aponeurosis and muscle," *Journal of Biomechanics*, vol. 25, no. 9, pp. 1017–1026, 1992.

- [12] M. Epstein and W. Herzog, *Theoretical Models of Skeletal Muscle. Biological and Mathematical Considerations*, John Wiley & Sons, Chichester, UK, 1998.
- [13] J. L. Van Leeuwen, "Muscle function in locomotion," in *Mechanics of Animal Locomotion*, R. M. Alexander, Ed., pp. 191–250, Springer, London, UK, 1992.
- [14] E. Otten and M. Hulliger, "A finite-element approach to the study of functional architecture in skeletal muscle," *Zoology*, vol. 98, pp. 233–242, 1995.
- [15] E. Azizi, G. B. Gillis, and E. L. Brainerd, "Morphology and mechanics of myosepta in a swimming salamander (*Siren lacertina*)," *Comparative Biochemistry and Physiology A: Molecular and Integrative Physiology*, vol. 133, no. 4, pp. 967–978, 2002.
- [16] C. W. J. Oomens, M. Maenhout, C. H. van Oijen, M. R. Drost, and F. P. Baaijens, "Finite element modelling of contracting skeletal muscle," *Philosophical Transactions of the Royal Society Series B: Biological Sciences*, vol. 358, no. 1437, pp. 1453–1460, 2003.
- [17] S. S. Blemker, P. M. Pinsky, and S. L. Delp, "A 3D model of muscle reveals the causes of nonuniform strains in the biceps brachii," *Journal of Biomechanics*, vol. 38, no. 4, pp. 657–665, 2005.
- [18] D. D. Sheriff and R. Van Bibber, "Flow-generating capability of the isolated skeletal muscle pump," *American Journal of Physiology—Heart and Circulatory Physiology*, vol. 274, no. 5, pp. H1502–H1508, 1998.
- [19] E. Azizi, E. L. Brainerd, and T. J. Roberts, "Variable gearing in pennate muscles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 5, pp. 1745–1750, 2008.
- [20] B. M. Damon, Z. Ding, A. W. Anderson, A. S. Freyer, and J. C. Gore, "Validation of diffusion tensor MRI-based muscle fiber tracking," *Magnetic Resonance in Medicine*, vol. 48, no. 1, pp. 97–104, 2002.
- [21] A. M. Heemskerk, T. K. Sinha, K. J. Wilson, Z. Ding, and B. M. Damon, "Quantitative assessment of DTI-based muscle fiber tracking and optimal tracking parameters," *Magnetic Resonance in Medicine*, vol. 61, no. 2, pp. 467–472, 2009.
- [22] J. H. Kan, M. Hernanz-Schulman, B. M. Damon, C. Yu, and S. A. Connolly, "MRI features of three paediatric intra-articular synovial lesions: a comparative study," *Clinical Radiology*, vol. 63, no. 7, pp. 805–812, 2008.
- [23] D. A. Lansdown, Z. Ding, M. Wadington, J. L. Hornberger, and B. M. Damon, "Quantitative diffusion tensor MRI-based fiber tracking of human skeletal muscle," *Journal of Applied Physiology*, vol. 103, no. 2, pp. 673–681, 2007.
- [24] D. I. W. Levin, B. Gilles, B. Mädler, and D. K. Pai, "Extracting skeletal muscle fiber fields from noisy diffusion tensor data," *Medical Image Analysis*, vol. 15, no. 3, pp. 340–353, 2011.
- [25] M. Rana, G. Hamarneh, and J. M. Wakeling, "Automated tracking of muscle fascicle orientation in B-mode ultrasound images," *Journal of Biomechanics*, vol. 42, no. 13, pp. 2068–2073, 2009.
- [26] A. I. L. Namburete, M. Rana, and J. M. Wakeling, "Computational methods for quantifying in vivo muscle fascicle curvature from ultrasound images," *Journal of Biomechanics*, vol. 44, no. 14, pp. 2538–2543, 2011.
- [27] J. Darby, E. F. Hodson-Tole, N. Costen, and I. D. Loram, "Automated regional analysis of B-mode ultrasound images of skeletal muscle movement," *Journal of Applied Physiology*, vol. 112, no. 2, pp. 313–327, 2012.
- [28] J. G. Gillett, R. S. Barrett, and G. A. Lichtwark, "Reliability and accuracy of an automated tracking algorithm to measure controlled passive and active muscle fascicle length changes from ultrasound," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 16, no. 6, pp. 678–687, 2013.
- [29] A. F. Frangi, W. J. Nissen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI'98*, vol. 1496 of *Lecture Notes in Computer Science*, pp. 130–137, Springer, Berlin, Germany, 1998.
- [30] M. Rana, G. Hamarneh, and J. M. Wakeling, "3D fascicle orientations in triceps surae," *Journal of Applied Physiology*, vol. 115, no. 1, pp. 116–125, 2013.
- [31] Y. Kawakami, Y. Ichinose, and T. Fukunaga, "Architectural and functional features of human triceps surae muscles during contraction," *Journal of Applied Physiology*, vol. 85, no. 2, pp. 398–404, 1998.
- [32] M. V. Narici, T. Binzoni, E. Hiltbrand, J. Fasel, F. Terrier, and P. Cerretelli, "In vivo human gastrocnemius architecture with changing joint angle at rest and during graded isometric contraction," *Journal of Physiology*, vol. 496, no. 1, pp. 287–297, 1996.
- [33] J. M. Wakeling, K. Uehli, and A. I. Rozitis, "Muscle fibre recruitment can respond to the mechanics of the muscle contraction," *Journal of the Royal Society Interface*, vol. 3, no. 9, pp. 533–544, 2006.
- [34] P. A. Huijting, "Muscular force transmission necessitates a multi-level integrative approach to the analysis of function of skeletal muscle," *Exercise and Sport Sciences Reviews*, vol. 31, no. 4, pp. 167–175, 2003.

Research Article

Comparison and Supervised Learning of Segmentation Methods Dedicated to Specular Microscope Images of Corneal Endothelium

Yann Gavet and Jean-Charles Pinoli

LGF UMR CNRS 5307, École Nationale Supérieure des Mines de Saint-Etienne, 158 Cours Fauriel, 42023 Saint-Etienne Cedex 2, France

Correspondence should be addressed to Yann Gavet; gavet@emse.fr

Received 12 February 2014; Accepted 12 August 2014; Published 22 September 2014

Academic Editor: Karen Panetta

Copyright © 2014 Y. Gavet and J.-C. Pinoli. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The cornea is the front of the eye. Its inner cell layer, called the endothelium, is important because it is closely related to the light transparency of the cornea. An *in vivo* observation of this layer is performed by using specular microscopy to evaluate the health of the cells: a high spatial density will result in a good transparency. Thus, the main criterion required by ophthalmologists is the cell density of the cornea endothelium, mainly obtained by an image segmentation process. Different methods can perform the image segmentation of these cells, and the three most performing methods are studied here. The question for the ophthalmologists is how to choose the best algorithm and to obtain the best possible results with it. This paper presents a methodology to compare these algorithms together. Moreover, by the way of geometric dissimilarity criteria, the algorithms are tuned up, and the best parameter values are thus proposed to the expert ophthalmologists.

1. Introduction

1.1. Human Eye and Cornea. The eye is the first sense organ responsible for human vision. The human eye functions like a camera to refract light and produce a focused image which stimulates neural responses transmitted to the brain vision centers. Crystalline lens is made of compacted protein fibers and is anchored in place by muscles attached to the wall of the eyeball. Contraction of these muscles causes the lens to change its shape and curvature, thus improving the focusing power. Refracted light passes through the eye cavity and strikes the inner surface at the back, known as the retina. The retina contains the specialized nerve cells called rods and cones that detect the intensity and the frequency of the incoming light. Light stimulates the rods and cones, which creates neural impulses that are transmitted to the brain through a network of nerve cells bunched together to form the optic nerve that exits from the back of the eyeball and passes to the brain.

1.1.1. The Human Cornea. The cornea is the transparent, spherical surface covering the front of the eye. It is a powerful refractive surface, providing about 2/3 of the eye's focusing power. Healthy cornea has no blood vessel, which accounts for its clarity. But it is rich in nerve endings and so it is extremely sensitive to pain. The tears and aqueous humor, a watery fluid circulating in the cavity behind it that contains glucose and several electrolytes, nourish the cornea. The cornea is a highly organized tissue consisting of cells and protein arranged in three main layers:

- (i) epithelium: this is the outermost layer comprising about 10% of the total thickness. Along with the tear film that bathes the outer surface of the eye, it provides a protective function preventing the entry of foreign material into the eye;
- (ii) stroma: it makes up to 90% of the corneal thickness. It consists primarily of water (78%) and layered collagen fibers (16%) that give the cornea its strength, elasticity,

and shape. It also contains cells scattered between the fibers that produce the stromal constituents. The lattice-like arrangement and uniform spacing of the collagen fibers are essential for corneal transparency;

- (iii) endothelium: this is the innermost layer facing the aqueous and consists of a single layer of hexagonal cells. It pumps water out of the cornea and hence plays a vital role in keeping it in a dehydrated state. Without this pumping action, the stroma would accumulate water and become hazy and finally opaque (corneal oedema) leading to loss of vision.

1.1.2. Physiology of the Human Corneal Endothelium. The cornea must remain transparent to refract light properly and the corneal endothelium ensures the integrity and transparency of the cornea.

The corneal endothelium consists of a single layer of closely packed, flat, hexagonally shaped cells covering the back surface of the cornea. In the human cornea at birth, there are more than 4000 cells/mm². With age, the number of endothelial cells gradually decreases, but because they cannot regenerate, neighboring cells spread out to fill the gap leading to an alteration of cell shape (pleomorphism) and size (polymegathism). The mean endothelial cell density (ECD) in adults is generally between 500 and 3500 cells/mm². Cell density, as well as variation in size and shape, can be examined by specular microscopy in living human subjects. These methods permit early diagnosis of any damage of the corneal endothelium.

1.2. Principles of Specular Microscopy. Since the cornea is transparent, cornea cells can easily be observed in vivo with a specular microscope. This technology comes from the early 1980s. Those optical microscopes can acquire an image of the cells on a very little surface (0.08 mm² compared to the endothelium surface of about 100 mm²; see Figure 5). The image is then analyzed by a computer software (embedded in the microscope) to provide both cell density and morphometry. The physical principle of this type of microscope is the specular reflection (i.e., the same reflection as for a mirror). As the light goes through the several different layers of the cornea, it is reflected at each interface [1, 2]. The deeper the layer is in the cornea, the darker it appears on the image. This explains the intensity variations within the images, which physically corresponds to the superposition of several different layers.

1.3. Quality Criteria of the Corneal Endothelium. It is necessary to evaluate the quality of the human corneal endothelium in several circumstances (for example, after accidents, surgery, or trauma). The main cause is corneal grafting.

The two criteria required for the evaluation are

- (i) the endothelial cell density (ECD, in cells/mm²): there are several threshold values: for example, an ECD lower than 400 cells/mm² does not enable maintaining the cornea transparency. An ECD lower

than 1000 cells/mm² is a contraindication for using intraocular lens implants.

- (ii) the morphometry of endothelial cells: their size regularity (called the polymegathism, i.e., the variation of areas of the cells) and their shape regularity (called the pleomorphism, i.e., the percentage of hexagon-like cells) induce a good quality of the cornea.

2. Three Image Segmentation Methods

Different methods exist to perform the segmentation of images of endothelial cells. Among those, three methods give the better results [3]. The present paper first recalls their related algorithms and will then compare their results with regard to several criteria.

The presented algorithms have a common structure. First, they filter the original image. Second, they aim to find some markers of the cells, and then they perform a morphological operation (a watershed; see [6]) to get closed contours for each cell. Notice that these three algorithms make important use of mathematical morphology operators (see, e.g., [7]).

2.1. Vincent and Masters' Method. This method has been proposed in [4]. It is based on the fact that cell borders' intensities are lower than the cell interiors' intensities and represent somehow a local maximum of intensity that is retrieved by a morphological operation called a *h*-maxima [7]. To avoid the problem of noise, a first filtering process is performed by the way of a morphological alternate sequential filter. It involves two parameters.

- (i) *h* is a value for the *h*-maxima operation (an intensity) that gives the final markers of the cells.
- (ii) *o* is the order of the morphological alternate sequential filter.

The algorithm is summarized in Algorithm 1.

2.2. Angulo and Matou's Method. This method is more recent than the previous one [5]. The cell markers are detected by the mean of a distance map after some filtering process (performed by a morphological opening and a morphological top-hat). The difference with Vincent and Masters' algorithm is that a first segmentation of the contours is performed and then is improved by the use of the watershed [6]. Algorithm 2 details the algorithm. Five parameters are required. The three first ones, *s*₁, *s*₂, and *g*, are used for the filtering process. The last two ones, *t* and *h*, are used to get the markers that will be used by the watershed.

2.3. Gavet and Pinoli's Method. This is the most recent method proposed in [3]. The improvements to the other methods come from the fact that the first segmentation of the borders of the cells better performs due to the elimination of nonlinear parts of the skeleton (see details of Algorithms 2 and 3). The Gavet and Pinoli's method requires five parameters, and its algorithm is summarized in Algorithm 3. After a filtering process (parameter *o*), the contours are first

Data: $Input \leftarrow$ Greyscale image of cornea endothelial cells.
 $o \leftarrow$ Order of alternate sequential filter.
 $h \leftarrow$ Value for the h -maxima.
Result: S : Segmentation of the cornea endothelial cells.

- (1) **begin**
- (2) $ASF \leftarrow$ Alternate Sequential Filter of order o of $Input$;
- (3) $M \leftarrow h$ -maxima of ASF and binarization (threshold 0);
- (4) $W \leftarrow$ Watershed constrained by M on $Input$ image;
- (5) $S \leftarrow$ Pruning of W to eliminate the lines touching the borders of the image;
- (6) **end**

ALGORITHM 1: Vincent and Masters's algorithm for detecting the human corneal endothelium cells [4].

Data: $Input \leftarrow$ Greyscale image of cornea endothelial cells.
 $s_1, s_2 \leftarrow$ Structuring element sizes.
 $g \leftarrow$ Size of Gaussian filter.
 $t \leftarrow$ Threshold value.
 $h \leftarrow$ Value for the h -maxima.
Result: S : Segmentation of the cornea endothelial cells.

- (1) **begin**
- (2) $G \leftarrow$ Gaussian filter of $Input$;
- (3) $O \leftarrow$ Opening of G with s_1 ;
- (4) $T \leftarrow$ Top-Hat of O with s_2 ;
- (5) $B \leftarrow$ Binarization of T with threshold t ;
- (6) $S \leftarrow$ Skeletonization of T ;
- (7) $DM \leftarrow$ Distance Map from $skel$;
- (8) $M \leftarrow h$ -maxima of DM (markers);
- (9) $W \leftarrow$ Watershed constrained by the markers M in the inverse of DM ;
- (10) $S \leftarrow$ Pruning of W to eliminate the lines touching the borders of the image;
- (11) **end**

ALGORITHM 2: Angulo and Matou's algorithm for detecting the cells [5].

Data: $Input \leftarrow$ Greyscale image of corneal endothelial cells.
 $o \leftarrow$ Order of alternate sequential filter.
 $t_s \leftarrow$ Length of segments.
 $c \leftarrow$ Structuring element size for a closing operation.
 $s \leftarrow$ Structuring element size for an opening operation.
Result: Segmentation of the cornea endothelial cells.

- (1) **begin**
- (2) $ASF \leftarrow$ Alternate Sequential Filter of size o of $Input$;
- (3) $R \leftarrow (Input - ASF) < 0$ (residu computation);
- (4) $U \leftarrow$ Supremum (union) of Openings of R by segments of length t_s ;
- (5) $C \leftarrow$ Closing of U (size c);
- (6) $SK \leftarrow$ Skeletonization of C ;
- (7) $DM \leftarrow$ Distance Map of SK ;
- (8) $M \leftarrow$ Regional Maxima of DM ;
- (9) $M_2 \leftarrow$ Opening of M of size s ;
- (10) $W \leftarrow$ watershed constrained by markers M_2 on inverse of DM ;
- (11) $S \leftarrow$ Pruning of W to eliminate the lines touching the borders of the image;
- (12) **end**

ALGORITHM 3: Gavet and Pinoli's algorithm for detecting the cornea endothelial cells.

TABLE 1: Summary of the control parameters of the three presented image segmentation algorithms.

Method	Parameters p	Description
Vincent and Masters Algorithm 1	$\{o, h\}$	o : order of filter h : h -maxima
Angulo and Matou Algorithm 2	$\{s_1, g, s_2, t, h\}$	s_1 : opening g : Gaussian filter s_2 : top-hat t : threshold h : h -maxima
Gavet and Pinoli Algorithm 3	$\{c, t_s, o, s\}$	c : closing t_s : length of segments o : order of filter s : opening

roughly detected by the use of the openings by segments (parameters t_s and then o , c , and e). This operators intend to detect linear cell contours. Then, these contours are corrected by the mean of a distance map and a constrained watershed (parameter s), as proposed by the two previous methods.

2.4. Summary of the Control Parameters. Table 1 summaries the control parameters used by the three presented image segmentation methods. The main problem here is how to select the values of these parameters. The following section will try to answer this problem by using two comparison criteria (namely, the dissimilarity criteria ϵ and fom) and by using them on an image database to find the best parameter values.

3. Image Segmentation Evaluation

The evaluation of a segmentation quality is a common problem encountered when developing a segmentation method. Like the segmentation methods themselves, the image segmentation evaluation criteria can be classified into region-based or contour-based approaches, although they usually can be adapted from one class to the other. The segmentation processes of the corneal endothelium result in the contours of the cells, but the proposed comparison methods are also suitable for segmented regions.

This paper deals with supervised segmentation evaluation, that is, involving a criterion that compares the result of the segmentation process to a ground truth image (usually manually segmented by an expert of the application field). This is usually preferred to unsupervised evaluation (where some kind of intraregion homogeneity is involved), but the bias introduced by the expert does not have to be neglected (see Section 3.4).

3.1. Basic Notations. The following notations are first introduced: I_R and I_X represent two binary images that correspond to a reference segmentation method R and the evaluated

segmentation method X , respectively. Both I_R and I_X are considered as sets of contours. In this paper, R and X may be employed for I_R and I_X in order to alleviate the notations and more deeply for emphasizing the geometrical problems. A point $p \in I_X$ or $p \in I_R$ means a point present in the related segmented binary image.

3.2. Classical Dissimilarity Criteria. This paper will not present an exhaustive view of supervised evaluation of segmentation criteria. The reader can have a look at [8] for a more complete presentation and a comparison.

The two detailed criteria have been chosen because they are tolerant towards spatial variations. One could also use other frequently used criteria proposed in the literature [9–11], like the Hausdorff distance, the Dice coefficient (also known as the Jaccard index), or classification criteria (specificity, sensibility). The main drawback of these criteria is that a small noise (like a misdected point) implies a high comparison value. This is why people introduced a distance ponderation, like the p th point in the Hausdorff distance, or the following figure of merit.

The figure of merit [12, 13] is defined by

$$\text{fom}_R(X) = 1 - \frac{1}{\max\{\#(M), \#(X)\}} \sum_{p \in X} \frac{1}{1 + d^2(p, R)}, \quad (1)$$

where $d(p, R)$ is the Euclidean distance from the pixel $p \in X$ to the closest pixel of R , and $\#$ is the number of pixels of the considered segmentation result R or result X (which are nonempty images, at least for R).

3.2.1. Partitioning. An image segmentation process refers to the action of partitioning the spatial domain of an image into adjacent regions, each of them preserving a certain homogeneity following a given criterion. Thus, a computer program is able to answer the following binary question: is this pixel inside the region of interest or not?

To formalize this mathematically, let \mathcal{S} be a binary image resulting from a segmentation process, defined by the number of regions (number of labels L , $K = [1; L]$) that partitions the spatial domain D and by the set \mathcal{R} of adjacent regions \mathcal{R}_i that fulfill D :

$$\begin{aligned} \forall (i, j) \in K, i \neq j, \quad \mathcal{R}_i \cap \mathcal{R}_j = \emptyset, \\ \mathcal{R} = \bigcup \mathcal{R}_i, \quad i \in K. \end{aligned} \quad (2)$$

This paper deals with the case where contours are detected and the segmentation result is a binary image; that is, $L = 2$, $K = [1; 2]$ (label 1 stands for the background and label 2 for the detected contours).

3.3. The ϵ Dissimilarity Criterion. The ϵ dissimilarity criterion is based on the symmetric difference Δ of sets, but this latter lacks some tolerance, which is introduced by the Minkowski addition.

3.3.1. *Symmetric Difference Set.* First, let us recall that the symmetric difference set between two segmentations R and X ($R \subset D$ and $X \subset D$), denoted by $\Delta(R, X)$, is defined by

$$\begin{aligned} \Delta(R, X) &= (R \cup X) \setminus (R \cap X) \\ &= (R \setminus X) \cup (X \setminus R). \end{aligned} \quad (3)$$

3.3.2. *Minkowski Addition.* The Minkowski addition [14] defines an algebraic operation between sets in the n -dimensional Euclidean space \mathbb{R}^n . It will be used to spatially “enlarge” the segmentations X or R to be compared in order to be less sensitive to small spatial variations or noises.

If X is a set (segmentation result) and B is another set (generally the unit ball), the Minkowski sum of X and B , denoted by $X \oplus B$, is then defined by

$$\begin{aligned} X \oplus B &= \{x + b \mid x \in X, b \in B\} \\ X \oplus B &= \bigcup_{b \in B} \{x + b \mid x \in X\}, \end{aligned} \quad (4)$$

where \oplus is the Minkowski addition symbol. In the field of mathematical morphology [7], it is equivalent to the dilation, and B is called a structuring element (for example, the unit ball).

3.3.3. *Definition.* In [15], each pixel in the segmentation result is attributed a distance value to the reference segmentation, and a histogram of these distances is thus computed. Then, a statistical analysis of this histogram is performed. In the same spirit, we propose a dissimilarity criterion that is tolerant towards small spatial variations. The ϵ dissimilarity criterion with the tolerance ρ applied to segmented images is defined in the case of discrete images (R is the reference segmentation result and X is the considered segmented image) by [8]

$$\epsilon_R^\rho(X) = \frac{\# \{(X \setminus (R \oplus \rho N)) \cup (R \setminus (X \oplus \rho N))\}}{\# \{R \oplus \rho N\}} \quad (5)$$

with N being the structuring element of radius 1 (typically the unit ball) and $\#$ designating the number of pixels in the set ($\#$ is the cardinal operator, counting the number of nonzero valued pixels in the set X or set R). Practically, ρ is the radius of the ball used to dilate the binary images, thus forming a tolerance tube around the original set X or set R . This paper will propose a way of selecting the right value for ρ .

The main properties of ϵ are

- (i) $\epsilon_R^\rho(R) = 0$, which means that when R is compared to itself, the numerical result is 0;
- (ii) $\epsilon_R^\rho(R) \xrightarrow{\rho \rightarrow \infty} 0$, which means that if the tolerance increases, the numerical value tends to 0;
- (iii) ρ is the tolerance value; thus, ϵ is tolerant towards small spatial variations (like translations, rotations, and over- and undersegmentations).

3.3.4. *Discussion about the Notion of Metric.* The usual concept to compare mathematical objects is the metric notion,

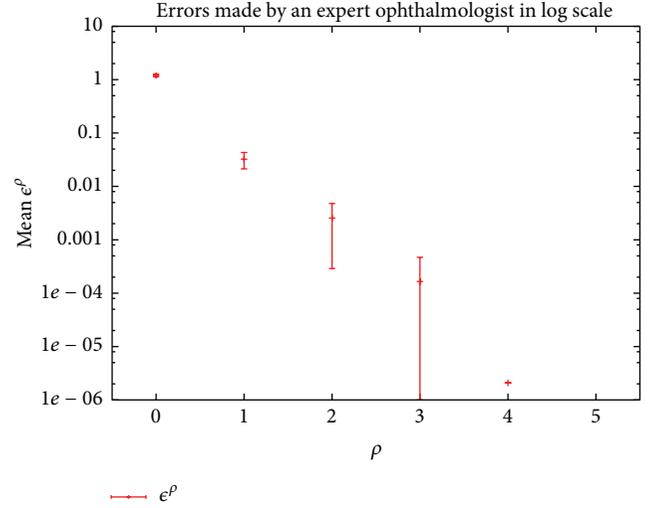


FIGURE 1: Method for fixing the tolerance parameter. In this example, ρ is in pixels, and there is a strong gap between no tolerance ($\rho = 0$) and a tolerance of one pixel ($\rho = 1$). The scale is logarithmic.

defined by four axioms (identity, separation, symmetry, and triangle inequality; see [16]). If a metric has important mathematical properties, it has been proved that the human visual system does not follow them [17, 18]. For example, the human visual system does not always consider two distinct objects as different (the separation property is thus not verified). This is also true for the triangle inequality and the symmetry property [8]. It is important to notice that ϵ is not a metric: separation, symmetry, and triangle inequality are not verified. This is why it is called a dissimilarity criterion.

3.4. *Bias in Experts Manual Segmentation: Choice of the Tolerance Value.* The problem of the experts reference segmentation is crucial because subject to variations between experts and sometimes also for one expert. To deal with this problem, some articles use an average result, like [19]. Some others do not take these into account and use only one reference segmentation as an absolute truth. The proposed ϵ dissimilarity criterion deals with this problem by the choice of the tolerance parameter ρ . The ρ value will in fact reflect the precision of the manual segmentation of the expert.

For one original gray-tone image, the experts have manually drawn their segmented image several times, and the ϵ dissimilarity criterion has been used to compare every manually segmented image to the others. The mean value of the ϵ dissimilarity criterion is represented in Figure 1. The reader can consider that an expert will always (try to) draw the contours at the same location within a certain spatial tolerance (i.e., within the tolerance tube), depending on the image size and the precision of the drawing tool. In the Figure 1, if an error is fixed at a maximum of $\epsilon = 0.05$, the application should then use a tolerance value of $\rho = 2$.

Thus, the ϵ dissimilarity criterion is able to deal with the bias in the experts reference segmentation as well with the noises present in the segmentation results themselves. The

next section will focus on cornea endothelium images and their segmentations.

4. Image Segmentation Method Tuning

The different segmentation algorithms presented in the previous sections require to setup the values of the so-called control parameters. The choice of the control parameter values for a specific application issue is generally not trivial, especially for nonimage analysis experts. This section explains the generic way of selecting the best parameters in average for the considered three image segmentation methods.

4.1. Method

4.1.1. Definitions. Let p be the control parameters set of a given algorithm A . For example, for Algorithm 1, $p = \{s, h\}$. Then, $A_p(I)$ is the result of the segmentation process by algorithm A with the parameter set p on the input image I .

Let C denote the criterion used to compare the segmentation results with the reference. In this paper, C will be either the dissimilarity criterion ϵ or criterion fom.

Let Q be an evaluation function of the quality of the segmentation, depending on the considered criterion C , defined as follows:

$$Q_C(A_p, I) = C(A_p(I), R), \quad (6)$$

where R is the reference segmentation of the image I . To simplify the notations, R will be used instead of $R(I)$.

4.1.2. Best Parameter Set. In the following, we consider an image database of \mathcal{N} gray-tone images, each being associated with a reference segmented image. What we are looking for is the best parameter set, that is, the parameter set that will result in the best segmented images considering the reference R and a specific comparison criterion (among the two considered criteria, fom and ϵ).

Let \hat{p} be the best parameter set regarding the mean of all quality values on the \mathcal{N} gray-tone images of the database, yielding to

$$\begin{aligned} \hat{p}(A, C) &= \arg \min_p \left\{ \frac{1}{\mathcal{N}} \sum_I Q_C(A_p, I) \right\} \\ &= \arg \min_p \left\{ \text{mean}_I Q_C(A_p, I) \right\}. \end{aligned} \quad (7)$$

Let \widehat{Q} be the minimal mean value of Q on the \mathcal{N} images of the database, yielding

$$\begin{aligned} \widehat{Q}(A, C) &= \min_p \left\{ \frac{1}{\mathcal{N}} \sum_I Q_C(A_p, I) \right\} \\ &= \text{mean}_I Q_C(A_{\hat{p}}, I). \end{aligned} \quad (8)$$

This way of finding the best parameter set is also called leave-one-out cross validation.

4.1.3. Trimmed Mean. Some noise may be present in the computed values (mainly because of a too poor image quality). To be more tolerant towards these perturbations, the trimmed mean (sometimes called truncated mean) is also employed: in the addressed application issue, given parts of the sample are discarded at the high end.

If $k \in [0; 0.5]$ is the percentage of discarded values, then

$$\begin{aligned} \widehat{p}_k(A, C) &= \arg \min_p \left\{ \text{Tmean}_{I,k} Q_C(A_p, I) \right\}, \\ \widehat{Q}_k(A, C) &= \min_p \left\{ \text{Tmean}_{I,k} Q_C(A_p, I) \right\}. \end{aligned} \quad (9)$$

Notice that the trimmed mean corresponds to the classical mean for $k = 0$; namely,

$$\begin{aligned} \widehat{p}_0 &= \hat{p} \\ \widehat{Q}_0 &= \widehat{Q}. \end{aligned} \quad (10)$$

4.1.4. Median. The median of Q , denoted by $\text{med} Q$, is also a classical way to avoid noise perturbations in such measurements, yielding to

$$\begin{aligned} \tilde{p}(A, C) &= \arg \min_p \left\{ \text{med}_I Q(A_p, I) \right\} \\ \widetilde{Q}(A, C) &= \min_p \left\{ \text{med}_I Q(A_p, I) \right\}. \end{aligned} \quad (11)$$

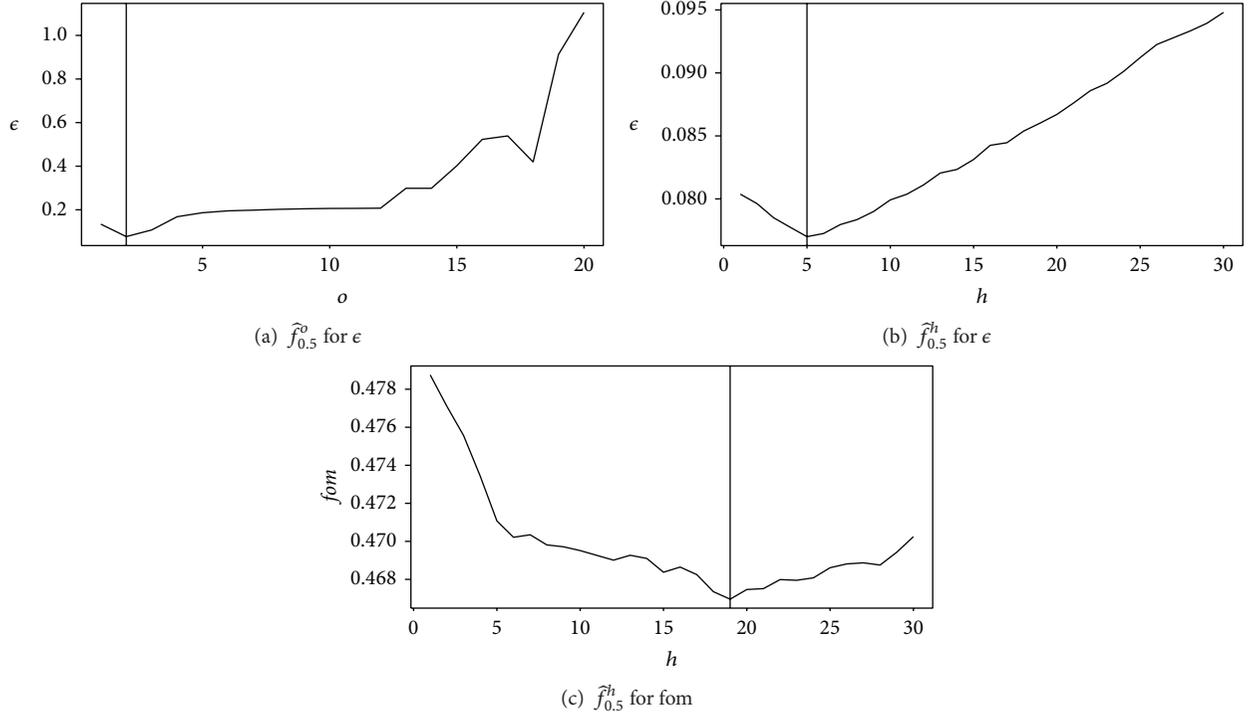
4.1.5. Projection. In order to observe the influence of one control parameter in the segmentation results, it is interesting to fix every control parameter but the considered one, and see if there is an impact on the quality of the segmentation. Let \mathcal{P} be a parameter of the set p . Let $\widehat{Q}_k^{\mathcal{P}}$ be the (trimmed) mean evolution of Q when the parameter \mathcal{P} is varying and the other parameters are fixed at values of \widehat{p}_k . The parameters are chosen among those defined in Table 1.

4.1.6. K-Fold Cross Validation. The K -fold cross-validation consists in validating the learning process by splitting the database into K -folds, using $K - 1$ -folds as the learning database and the last one as the test database [20]. This is repeated K times such that each fold will be used as the test database. The result is a discrepancy value that reflects the pertinence of the learning. It will be noticed CV in the different result tables, which is the mean result value over each test partition.

The results will be presented in Tables 5, 6, and 7 in detail. For each method and for each partition, the learned parameter values are presented as well as the corresponding result of the criterion value for the test partition. Let Q_i^{CV} be the value of the criterion for the partition i , and $Q^{CV} = \text{mean}_i \{Q_i^{CV}\}$.

TABLE 2: Results for Algorithm 1, method of Vincent and Masters, where the control parameters are $p = \{o, h\}$. See Figure 2 for illustrations.

Criterion	Optimal parameters		Trimmed mean ($k = 0.5$)		Median		Q^{CV}
	\hat{p}	\hat{Q}	$\hat{p}_{0.5}$	$\hat{Q}_{0.5}$	\tilde{p}	\tilde{Q}	
ϵ	{2, 12}	0.11	{2, 5}	0.08	{2, 5}	0.10	0.115
fom	{2, 19}	0.52	{2, 19}	0.47	{2, 23}	0.50	0.520

FIGURE 2: Projection on the optimal values of the control parameters o and h for the method of Vincent and Masters. See Table 2 for numerical values.

For comparison purposes, we also provide the best criterion value that could have been obtained on the test partition, denoted by \tilde{Q}_i^{CV} (for partition i) and by \tilde{Q}^{CV} for the mean on all partitions. The value of \tilde{Q}^{CV} should be only a little smaller than Q^{CV} .

4.2. Quantitative Comparison Results. This section presents the results for the three aforementioned image segmentation methods.

An image database of $\mathcal{N} = 30$ gray-tone images of the human corneal endothelium acquired with a specular microscope is employed to evaluate the segmentation processes realized by the different algorithms. This image database (see Figure 5) contains gray-tone images and also the related experts' segmented images (manually performed).

4.2.1. Vincent and Masters' Method. The summary of the optimal control parameters values is presented in Table 2.

It appears that ϵ and fom do not provide the same results for the optimal value of parameter h . One shall notice that the value of $\tilde{Q}_{0.5}^h$ does not vary a lot for both ϵ and fom criteria

(see Figures 2(b) and 2(c)). This means that the choice of h appears as not crucial.

4.2.2. Angulo and Matou's Method. The results are presented in Table 3. Both ϵ and fom give the same results. In addition, the projections (Figures 3(e) and 3(c)) show that s_2 and h are useless. Thus, the top-hat transform can be avoided, and the h -maxima operation can be replaced by the computation of the maxima of the distance map.

4.2.3. Gavet and Pinoli's Method. The summary of the optimal control parameter values is presented in Table 4.

For both fom and ϵ criteria, the optimal control parameter values are identical. The filtering parameter o used in the alternate sequential filter must be higher than 4. This parameter is linked to the length of an edge of a cell, and, thus, some corneal endothelium with big cells would get a better result with a higher value of o (this explains the almost constant values that can be observed in Figure 4(c)).

The control parameter s used for filtering markers must be chosen with a low value (less than 7). In this case, the effect is not really noticeable.

TABLE 3: Results for Algorithm 2, method of Angulo and Matou, where the control parameters are $p = \{s_1, g, s_2, t, h\}$. See Figure 3 for illustrations.

Criterion	Optimal parameters		Trimmed mean ($k = 0.5$)		Median		Q^{CV}
	\hat{p}	\hat{Q}	$\hat{p}_{0.5}$	$\hat{Q}_{0.5}$	\tilde{p}	\tilde{Q}	
ϵ	{4, 4, 1, 28, 1}	0.15	{3, 4, 1, 22, 1}	0.08	{4, 3, 1, 16, 1}	0.11	0.150
fom	{3, 4, 1, 21, 1}	0.54	{3, 4, 1, 21, 1}	0.46	{3, 4, 1, 8, 1}	0.51	0.525

TABLE 4: Results for Algorithm 3, method of Gavet and Pinoli, where the control parameters are $p = \{c, t_s, o, s\}$. See Figure 4 for illustrations.

Criterion	Optimal parameters		Trimmed mean ($k = 0.5$)		Median		Q^{CV}
	\hat{p}	\hat{Q}	$\hat{p}_{0.5}$	$\hat{Q}_{0.5}$	\tilde{p}	\tilde{Q}	
ϵ	{2, 7, 4, 6}	0.10	{2, 7, 4, 6}	0.06	{2, 7, 4, 6}	0.08	0.099
fom	{2, 7, 4, 7}	0.50	{2, 7, 4, 7}	0.45	{2, 7, 4, 8}	0.49	0.506

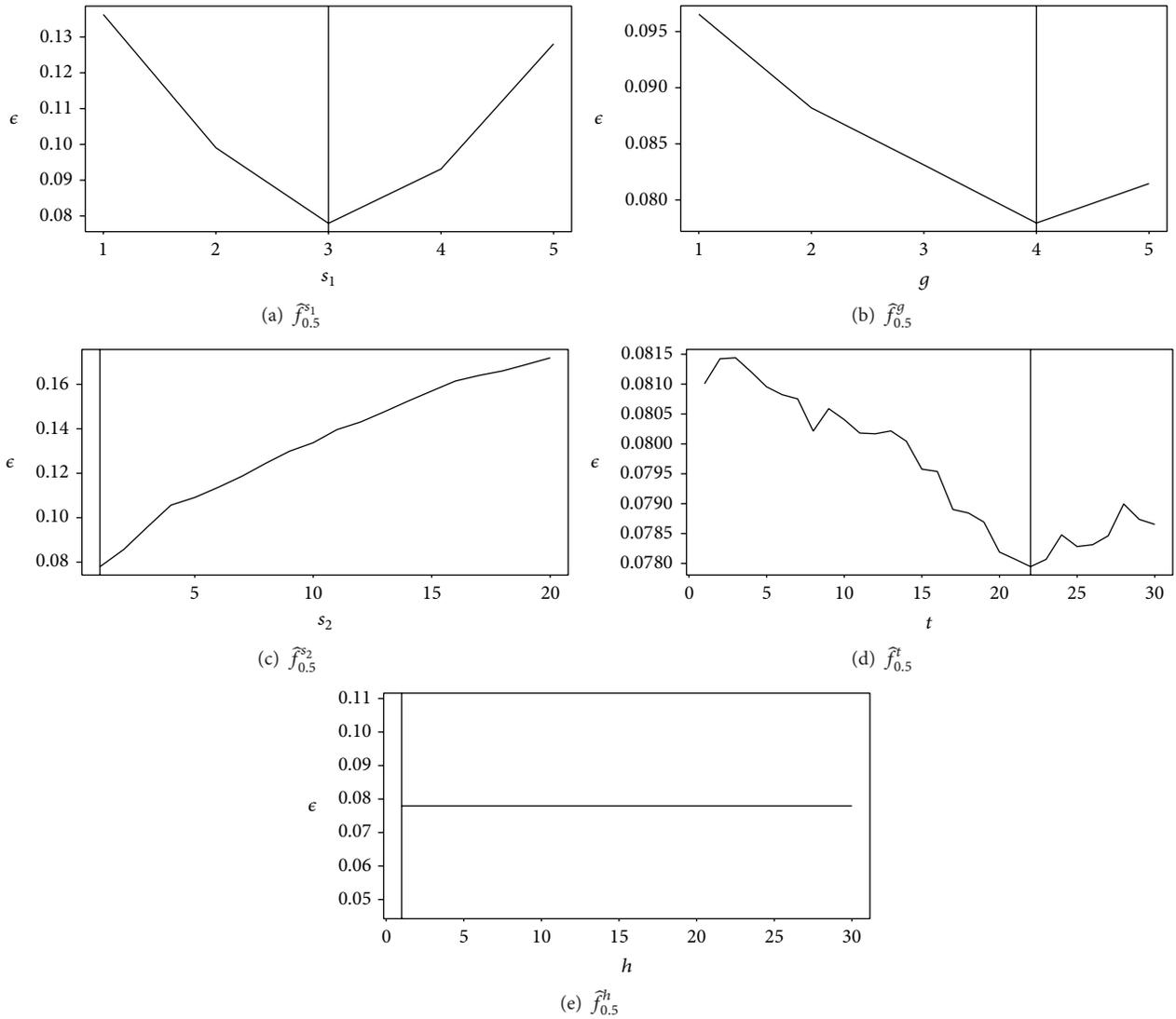


FIGURE 3: Projection on the optimal values of the control parameters for the method of Angulo and Matou. See Table 3 for numerical values.

TABLE 5: Cross-validation information for Algorithm 1, method of Vincent and Masters, with ϵ criterion.

Partition number	Optimal parameter values	Q_i^{CV}	\bar{Q}_i^{CV}
1	{2, 8}	0.146	0.118
2	{2, 12}	0.099	0.096
3	{2, 9}	0.110	0.109
4	{2, 12}	0.117	0.115
5	{2, 12}	0.104	0.104
Mean		$Q^{CV} = 0.115$	$\bar{Q}^{CV} = 0.108$

TABLE 6: Cross-validation information for Algorithm 1, method of Vincent and Masters, with fom criterion.

Partition number	Optimal parameter values	Q_i^{CV}	\bar{Q}_i^{CV}
1	{2, 19}	0.512	0.512
2	{2, 18}	0.524	0.510
3	{2, 19}	0.490	0.488
4	{2, 19}	0.533	0.533
5	{2, 19}	0.540	0.540
Mean		$Q^{CV} = 0.520$	$\bar{Q}^{CV} = 0.516$

TABLE 7: Cross-validation information for Algorithm 2, method of Angulo and Matou, with ϵ criterion.

Partition number	Optimal parameter values	Q_i^{CV}	\bar{Q}_i^{CV}
1	{4, 5, 1, 25, 1}	0.162	0.148
2	{4, 4, 1, 28, 1}	0.157	0.155
3	{4, 4, 1, 28, 1}	0.141	0.141
4	{4, 4, 1, 28, 1}	0.109	0.109
5	{4, 4, 1, 30, 1}	0.180	0.158
Mean		$Q^{CV} = 0.150$	$\bar{Q}^{CV} = 0.142$

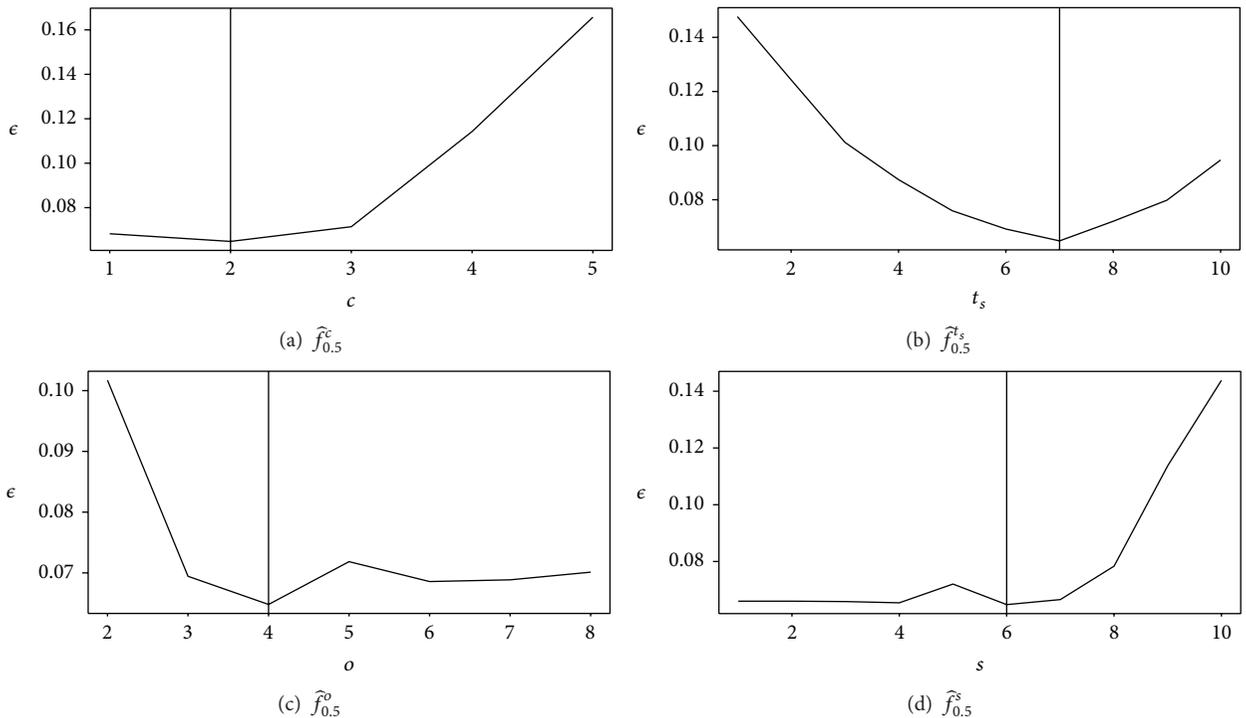


FIGURE 4: Projection on the optimal values of the control parameters for the method of Gavet and Pinoli. See Table 4 for numerical values.

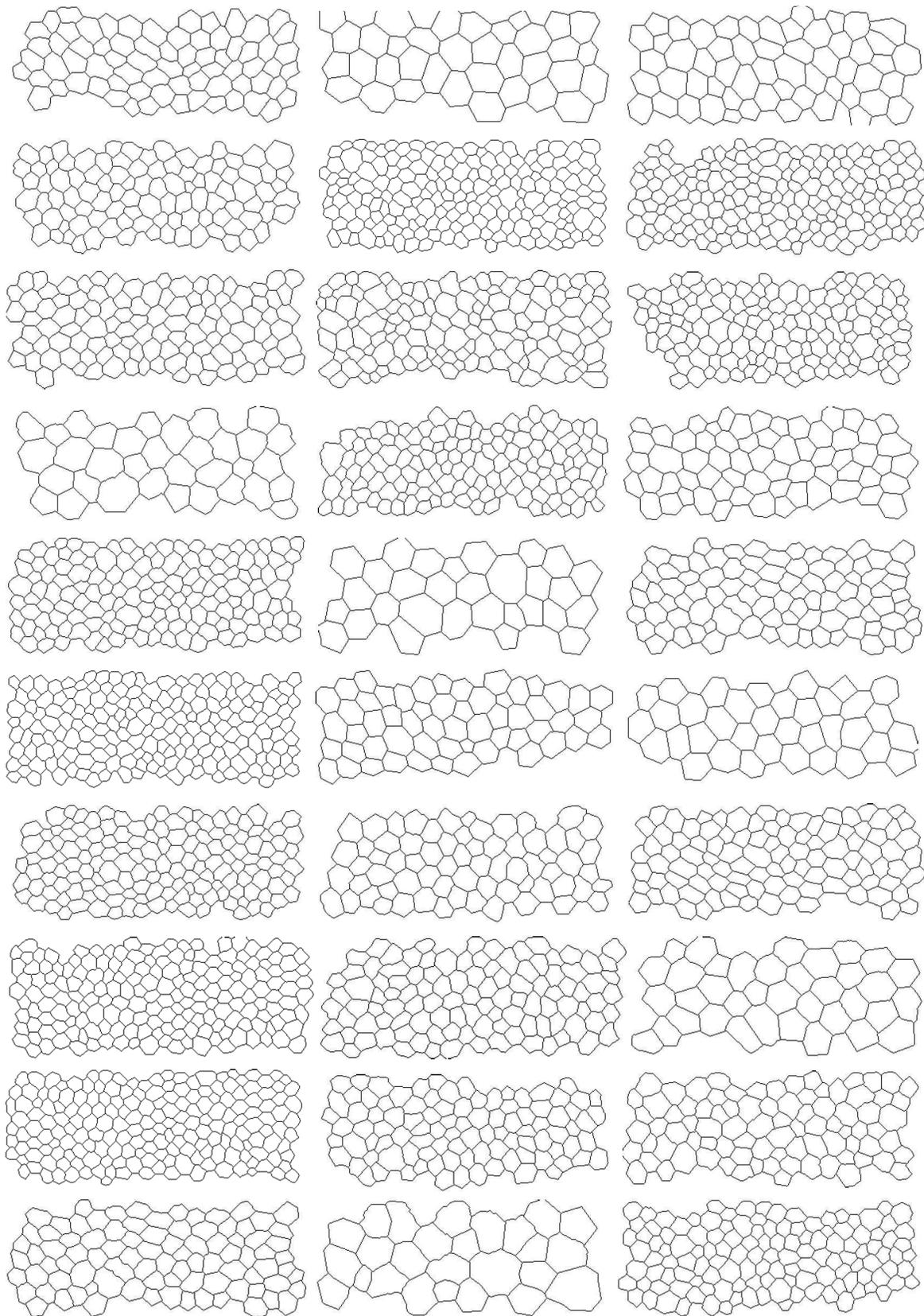


FIGURE 5: Table of the 30 reference segmented images of the database. They have been manually drawn by an expert ophthalmologist from a human corneal endothelium image database (see Figure 6). These images come from [8].

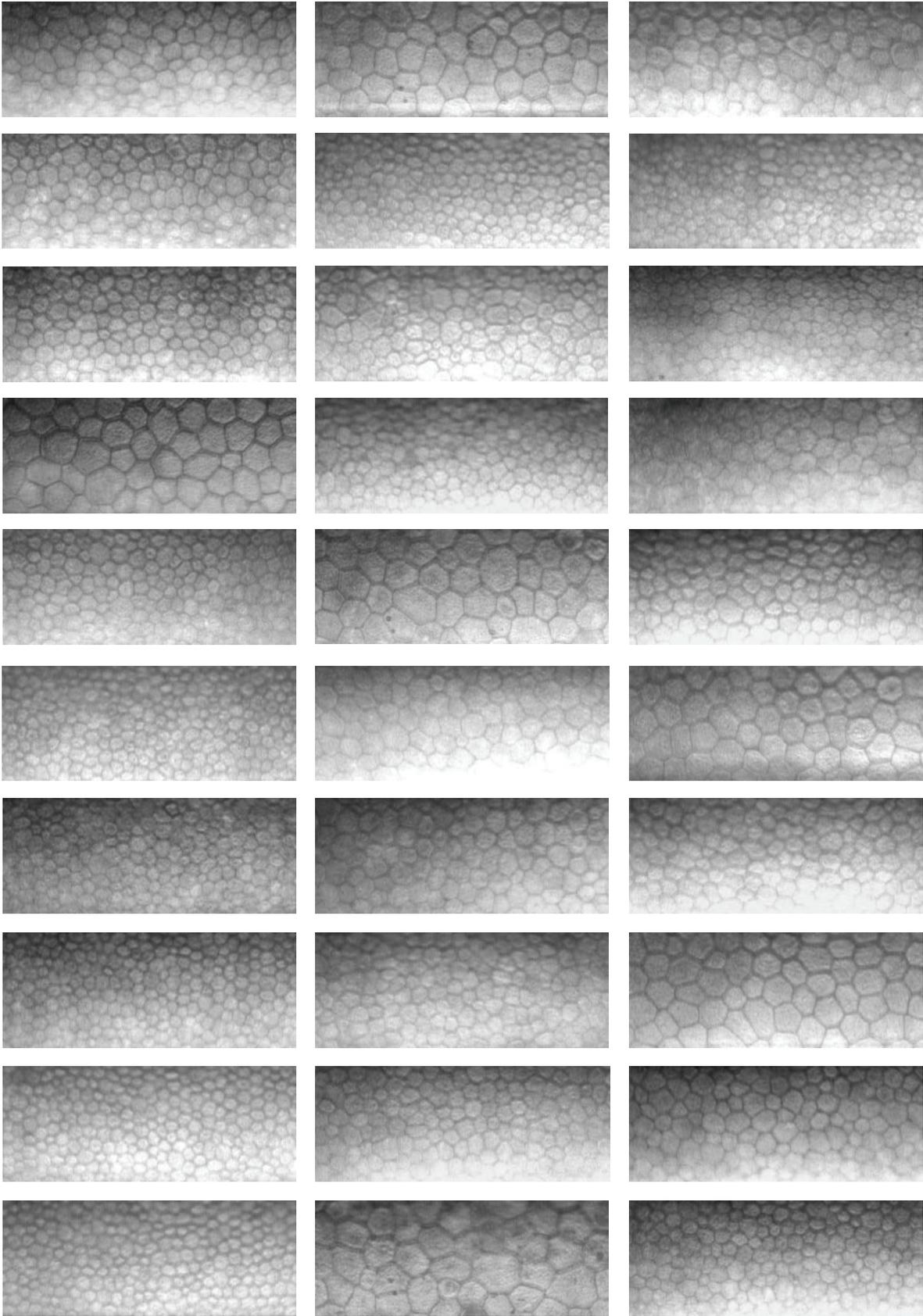


FIGURE 6: Table of the 30 specular microscopy images of corneal endotheliums of the database. They are segmented by the proposed method and by an ophthalmologist (see Figure 5). These images come from [8].

TABLE 8: Cross-validation information for Algorithm 2, method of Angulo and Matou, with fom criterion.

Partition number	Optimal parameter values	Q_i^{CV}	\bar{Q}_i^{CV}
1	{3, 4, 1, 21, 1}	0.510	0.510
2	{3, 4, 1, 21, 1}	0.565	0.558
3	{3, 4, 1, 21, 1}	0.624	0.598
4	{3, 4, 1, 21, 1}	0.448	0.443
5	{3, 4, 1, 21, 1}	0.479	0.479
Mean		$Q^{CV} = 0.525$	$\bar{Q}^{CV} = 0.518$

TABLE 9: Cross-validation information for Algorithm 3, method of Gavet and Pinoli, with ϵ criterion.

Partition number	Optimal parameter values	Q_i^{CV}	\bar{Q}_i^{CV}
1	{2, 7, 4, 6}	0.078	0.078
2	{2, 7, 4, 6}	0.100	0.099
3	{2, 7, 4, 5}	0.111	0.093
4	{2, 7, 4, 6}	0.109	0.101
5	{2, 7, 4, 7}	0.100	0.091
Mean		$Q^{CV} = 0.099$	$\bar{Q}^{CV} = 0.092$

TABLE 10: Cross-validation information for Algorithm 3, method of Gavet and Pinoli, with fom criterion.

Partition number	Optimal parameter values	Q_i^{CV}	\bar{Q}_i^{CV}
1	{2, 7, 4, 8}	0.506	0.502
2	{2, 7, 4, 7}	0.490	0.486
3	{2, 7, 4, 8}	0.486	0.480
4	{2, 7, 4, 7}	0.484	0.477
5	{2, 7, 4, 7}	0.562	0.548
Mean		$Q^{CV} = 0.506$	$\bar{Q}^{CV} = 0.498$

The control parameter t_s is strictly related to the length of the cells edges (its purpose is to retrieve the linear parts of the detected borders and thus eliminate the noise).

4.3. *Comments.* According to the numerical values, the method of Gavet and Pinoli outperforms the methods from Vincent and Masters, and Angulo and Matou (see the values of Q^{CV} , \bar{Q} , $\bar{Q}_{0.5}$, or \bar{Q}). Clearly, the superiority of this method comes from the extraction of the linear parts of the cell borders.

The K -fold cross-validation (see Tables 5, 6, 7, 8, 9, and 10) shows that the learning is not far from the optimal value, in terms of ϵ or fom, as well as in terms of optimal parameter values. This can be concluded from comparing the mean ϵ or fom value from the test partitions (Q^{CV}) to the best possible value (\bar{Q}^{CV}): there is only a small difference between these values.

Moreover, the optimal parameter values obtained for the different training partition do not vary a lot and are really similar to those proposed in Tables 2, 3, and 4.

5. Conclusion and Perspectives

In this paper, three segmentation methods suitable for binarizing the optical specular microscopy gray-tone images of human corneal endotheliums have been presented. These methods involve different control parameters. This is always a hard problem for the user because he has no time to manually

tune up his computer softwares (and especially his image segmentation softwares). Two dissimilarity criteria have been employed (ϵ dissimilarity criterion [8] and Pratt's figure of merit fom [12]) to tune up the segmentation algorithms in regard to the expert manual segmentation. As a result, this paper proposes the optimal control parameter values to use for these images. It also proposes to avoid some operations since their parameters do not really influence the segmentation results. More generally, this paper highlights the relevance of the ϵ dissimilarity criterion to a spatial tolerance, suitable to handle the problem of the bias in the reference segmentation. This ϵ dissimilarity criterion is adapted to compare binary contour images as well as binary sets, for 2D or even 3D images. In a near future, we expect to report such a criterion to compare gray-tone images.

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 thank Pr. Gain and Pr. Thuret from the University Hospital of Saint-Etienne for providing them with their images and their expertise in the field of ophthalmology. This work was partially supported by the ANR Project TecSan 2012, CorImMo 3D.

References

- [1] W. M. Bourne and J. M. Enoch, "Some optical principles of the clinical specular microscope," *Investigative Ophthalmology & Visual Science*, vol. 15, no. 1, pp. 29–32, 1976.
- [2] R. A. Laing, M. M. Sandstrom, and H. M. Leibowitz, "Clinical specular microscopy," *Archives of Ophthalmology*, vol. 97, no. 9, pp. 1714–1719, 1979.
- [3] Y. Gavet and J.-C. Pinoli, "Visual perception based automatic recognition of cell mosaics in human corneal endothelium microscopy images," *Image Analysis and Stereology*, vol. 27, no. 1, pp. 53–61, 2008.
- [4] L. Vincent and B. Masters, "Morphological image processing and network analysis of cornea endothelial cell images," in *3rd Image Algebra and Morphological Image Processing*, P. D. Gader, E. R. Dougherty, and J. C. Serra, Eds., vol. 1769 of *Proceedings of SPIE*, pp. 212–226, San Diego, Calif, USA, June 1992.
- [5] J. Angulo and S. Matou, "Automatic quantification of in vitro endothelial cell networks using mathematical morphology," in *Proceedings of the 5th IASTED International Conference on Visualization, Imaging, and Image Processing (VIIP '05)*, pp. 51–56, Benidorm, Spain, September 2005.
- [6] S. Beucher and C. Lantuejoul, "Use of watersheds in contour detection," in *Proceedings of the International Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, Rennes, France, 1979.
- [7] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer, New York, NY, USA, 2nd edition, 2003.
- [8] Y. Gavet and J.-C. Pinoli, "A geometric dissimilarity criterion between Jordan spatial mosaics. Theoretical aspects and application to segmentation evaluation," *Journal of Mathematical Imaging and Vision*, vol. 42, no. 1, pp. 25–49, 2012.
- [9] S. Ghose, A. Oliver, J. Mitra et al., "A supervised learning framework of statistical shape and probability priors for automatic prostate segmentation in ultrasound images," *Medical Image Analysis*, vol. 17, no. 6, pp. 587–600, 2013.
- [10] M. Schaap, C. T. Metz, T. van Walsum et al., "Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms," *Medical Image Analysis*, vol. 13, no. 5, pp. 701–714, 2009, Includes Special Section on the 12th International Conference on Medical Imaging and Computer Assisted Intervention.
- [11] H. A. Kirişli, M. Schaap, C. T. Metz et al., "Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography," *Medical Image Analysis*, vol. 17, no. 8, pp. 859–876, 2013.
- [12] I. E. Abdou and W. K. Pratt, "Quantitative design and evaluation of enhancement/thresholding edge detectors," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 753–763, 1979.
- [13] K. C. Strasters and J. J. Gerbrands, "Three-dimensional image segmentation using a split, merge and group approach," *Pattern Recognition Letters*, vol. 12, no. 5, pp. 307–325, 1991.
- [14] H. Minkowski, "Volumen und Oberfläche," *Mathematische Annalen*, vol. 57, no. 4, pp. 447–495, 1903.
- [15] B. Moretti, J. M. Fadili, S. Ruan, D. Bloyet, and B. Mazoyer, "Phantom-based performance evaluation: application to brain segmentation from magnetic resonance images," *Medical Image Analysis*, vol. 4, no. 4, pp. 303–316, 2000.
- [16] M.-M. Deza and E. Deza, *Dictionary of Distances*, Elsevier, 2006.
- [17] A. Tversky, "Features of similarity," *Psychological Review*, vol. 84, no. 4, pp. 327–352, 1977.
- [18] A. Tversky and I. Gati, "Similarity, separability, and the triangle inequality," *Psychological Review*, vol. 89, no. 2, pp. 123–154, 1982.
- [19] K. Hameeteman, M. A. Zuluaga, M. Freiman et al., "Evaluation framework for carotid bifurcation lumen segmentation and stenosis grading," *Medical Image Analysis*, vol. 15, no. 4, pp. 477–488, 2011.
- [20] T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning*, vol. 1, Springer, New York, NY, USA, 2001.

Research Article

Choosing the Optimal Spatial Domain Measure of Enhancement for Mammogram Images

Karen Panetta,¹ Arash Samani,¹ and Sos Agaian²

¹ Department of Electrical and Computer Engineering, Tufts University, Medford, MA, USA

² Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX, USA

Correspondence should be addressed to Arash Samani; samani@ieee.org

Received 11 April 2014; Accepted 19 June 2014; Published 6 August 2014

Academic Editor: Yicong Zhou

Copyright © 2014 Karen Panetta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Medical imaging systems often require image enhancement, such as improving the image contrast, to provide medical professionals with the best visual image quality. This helps in anomaly detection and diagnosis. Most enhancement algorithms are iterative processes that require many parameters be selected. Poor or nonoptimal parameter selection can have a negative effect on the enhancement process. In this paper, a quantitative metric for measuring the image quality is used to select the optimal operating parameters for the enhancement algorithms. A variety of measures evaluating the quality of an image enhancement will be presented along with each measure's basis for analysis, namely, on image content and image attributes. We also provide guidelines for systematically choosing the proper measure of image quality for medical images.

1. Introduction

Mammography is noninvasive imaging that uses a low-dose X-ray to photograph breast tissue. The result of mammography is a mammogram, which is used as a screening test for breast cancer. Mammography is a valuable screening procedure that can detect breast cancer early, as long as two years before a lump can be felt. Mammography also is used to help clarify whether a suspicious breast lump is a cyst or a tumor and whether a tumor is more likely to be benign or malignant. Mammography misses breast cancer about 5% to 10% of the time, but the rate can be as high as 30% for women with dense breast tissue [1]. The X-ray hardware limitation and the high density breast tissue will affect the quality of mammograms that can contribute to misdiagnosis of breast cancer. Enhancing the contrast of mammogram images could improve the results of mammography significantly [2, 3].

There are many different techniques for enhancing the quality of an image [2–7], which allow the observer to better perceive the desirable information in the image. These techniques often have parameters to control the enhancement process and its outcome. To optimize these parameters, feedback is required in the enhancement process, which

is a quantitative assessment of image quality, known as measure of enhancement or image quality measure. There have been different definitions of an adequate measure of performance based on contrast [8–10]. Contrast based measure of enhancement methods can be categorized as either spatial or transform based domain measures [10]. The spatial domain measures are calculated based on luminance of pixels in different portions of an image, but the transform domain measures work based on the Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), or Discrete Wavelet Transform (DWT) of the image [11]. Most spatial domain measures are derivatives of the Weber-Fechner law, Michelson contrast measure [12], or Contrast Ratio (CR), using statistical analysis to better evaluate the image contrast enhancement. EME, EMEE, AME, AMEE, LogAME, and LogAMEE are examples of such spatial domain measures [4, 5, 13].

Performance of spatial domain measures is highly dependent on image attributes such as image content, lighting, uniform versus nonuniform background, texture, periodic patterns, randomness, single versus multiple targets, noise, and distortions. The choice of a measure of enhancement method could affect the outcome of image enhancement

algorithms. If a measure is not designed to handle a specific image attribute properly, that measure will not provide a good metric for properly evaluating the contrast enhancement for that specific image.

In this paper, we explain the relationship between the image properties and the commonly used spatial domain measures of image quality. We will also provide guidelines of how to choose the measure of enhancement based on the image attributes and how to select the optimal operating parameters that are used by these measures. Finally, recommendations are presented for determining which spatial domain measure is best suited to assess the quality of mammogram images.

2. Materials and Methods

When enhancing a digital image, there is a need for a quantitative measure to evaluate the visual performance of the enhancement process. In iterative automated enhancement algorithms, one goal of the enhancement measure is to provide a statistic to help optimize the enhancement algorithm's parameters. These measures are based on luminance contrast statistics of the image and generally have two components. The first is comparing the difference between the two luminance components in question (e.g., of a symbol and its background). If the state of adaptation of the visual system stays constant, larger luminance differences produce larger brightness differences (higher brightness contrast). The second component of any luminance contrast statistic is some measure describing the adaptation state of the eye. A luminance that produces a large brightness on a dim background will produce a smaller brightness difference on a brighter background due to visual adaptation. To capture this behavior, designers of luminance contrast statistics generally divide a numerator that describes the luminance change by a denominator that describes the average luminance to which the eye is adapted:

$$\text{Luminance Contrast} = \frac{\text{Luminance Change}}{\text{Adaptation Descriptor}}. \quad (1)$$

The variety of popular statistics for luminance contrast mostly reflects the fact that the adaptation state of the eye is affected differently by different kinds of stimulus patterns. In the following section, we provide the fundamental measures used to evaluate contrast.

2.1. Basic Measures of Image Contrast

2.1.1. Michelson Contrast. This measure is commonly used for patterns where both bright and dark features are equivalent and cover similar fractions of the area, for example, periodic patterns such as a sinusoidal grating. Michelson contrast is defined as

$$C_M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (2)$$

with I_{\max} and I_{\min} representing the highest and lowest luminance. The denominator represents twice the average

of the luminance [12]. For simple periodic patterns (e.g., textures) there is no large area of uniform luminance that dominates the user's brightness adaptation. The denominator of C_M is twice the mean of the maximum and minimum luminance; that is, the adaptation luminance estimate is based on the space-average luminance.

2.1.2. Weber-Fechner Law. The Weber-Fechner law implies a logarithmic relationship between physical luminance and subjectively perceived brightness. Weber contrast measure assumes a large uniform luminance background with a small test target:

$$C_W = \frac{I_s - I_b}{I_b}, \quad (3)$$

where I_s is the luminance of the target and I_b is the luminance of the immediately adjacent background. It is commonly used in cases where small features are present on a large uniform background; that is, the average luminance is approximately equal to the background luminance. When the background is lighter than the target C_W is negative and ranges from zero to -1 . When the background is darker than the target C_W is positive and ranges from zero to potentially very large numbers.

The Weber-Fechner measure is used in cases where the average luminance is approximately equal to the background luminance.

2.1.3. Contrast Ratio. This measure has often been applied to the stimulus class in either linear or logarithmic form:

$$CR = \frac{I_s}{I_b}, \quad \log(CR) = \frac{\log(I_s)}{\log(I_b)}. \quad (4)$$

This measure, C_R , does not mathematically compare with the Weber contrast since the numerator is not the luminance difference between the target and the background.

2.1.4. Entropy. Entropy is calculated from the histogram of an image and is calculated over the entire image. It is a scalar value representing the entropy of an intensity image, a statistical measure of randomness that can be used to characterize the texture of the image:

$$\text{Entropy} = - \sum p * \ln(p), \quad (5)$$

where p is the histogram count for a segment of image.

Since entropy is calculated over the entire image, rearranging segments of the image would not change this contrast measure. Also increasing the contrast in one part of an image and decreasing it in another part may result in similar entropy as the original image.

2.2. Complex Measures of Contrast. These measures are based on the basic contrast measures with additional optimization parameters. Since the basic measures of contrast are highly sensitive to image contents and attributes such as noise, periodicity, texture, randomness, uniform background, and

target size, often a combination of two basic measures in a new quality measure could extend the application of this new complex measure. The image contents should be visually inspected to choose the optimal measure. Figure 1 shows examples of different image attributes.

2.2.1. *EME*. Let an image $x(n, m)$ be split into $k_1 k_2$ blocks of $W_{k,l}(i, j)$; the EME is defined as

$$\text{EME}_{k_1 k_2} = \frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} \left[20 \ln \left(\frac{I_{\max;k,l}}{I_{\min;k,l}} \right) \right], \quad (6)$$

where $I_{\min;k,l}$ and $I_{\max;k,l}$ are the minimum and maximum of the image inside the block $W_{k,l}$. As defined in (4), for each block of (k, l) in (k_1, k_2) blocks, the Contrast Ratio is

$$\text{CR}_{k,l} = \frac{I_{\max;k,l}}{I_{\min;k,l}}. \quad (7)$$

If $k_1 k_2 = 1$ (image is divided into one block), this method will return result similar to Contrast Ratio (C_R):

$$\text{EME}_{k_1 k_2} = \frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} [20 \ln (\text{CR}_{k,l})]. \quad (8)$$

Dividing the image into $k_1 k_2$ blocks will turn the complex image into simple blocks assuming that there is only one target per block and I_s turns out to be I_{\max} and the background (I_b) carries I_{\min} .

The EME measure of enhancement is suitable for images with the following properties:

- (i) noncomplex segments (CR property);
- (ii) uniform background in segments (Weber property);
- (iii) small targets in segments (Weber property);
- (iv) nonperiodic pattern in segments (Weber property);
- (v) little to no randomness in segments (no entropy in measure).

The EME measure would not be a good choice for images with these attributes:

- (i) complex segments within an image or complex images with large block size (CR property);
- (ii) images with nonuniform background in each segment or images with large segment size (Weber property);
- (iii) large target in segments the way that the segment is equally divided into high and low luminance areas; this attribute also extends to periodic patterns and images with random texture (Weber property).

2.2.2. *EMEE*. Let an image $x(n, m)$ be split into $k_1 k_2$ blocks of $W_{k,l}(i, j)$; the EMEE is defined as

$$\text{EMEE}_{\alpha k_1 k_2} = \frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} \left[\alpha \left(\frac{I_{\max;k,l}}{I_{\min;k,l}} \right)^\alpha \ln \left(\frac{I_{\max;k,l}}{I_{\min;k,l}} \right) \right]. \quad (9)$$

Similar to entropy (5), for each block $W_{k,l}$ of the image, the ratio

$$p = \left(\frac{I_{\max;k,l}}{I_{\min;k,l}} \right)^\alpha \quad (10)$$

represents the number of intensity levels in that block if block is normalized by $I_{\min;k,l}$ if $\alpha = 1$. For each block, the Contrast Ratio is defined as in (7); therefore the EMEE measure yields to

$$\begin{aligned} \text{EMEE}_{\alpha k_1 k_2} &= \frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} \left[\alpha (\text{CR}_{k,l})^\alpha \ln (\text{CR}_{k,l}) \right] \\ &= \frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} [\text{Entropy} (\text{CR}^\alpha)]. \end{aligned} \quad (11)$$

Therefore, the EMEE is the entropy of the Contrast Ratio for each block $W_{k,l}$ scaled by α , averaged over the entire image. This association makes EMEE measure of enhancement suitable for images with the following properties:

- (i) noncomplex segments (CR property);
- (ii) nonperiodic patterns in segments (Weber property);
- (iii) being able to handle randomness in texture (because of added entropy compared to EME);
- (iv) using a larger “ α ” parameter will help to handle more randomness in image texture by emphasizing the entropy term.

EMEE measure is not a good choice to handle

- (i) images with complex segments or if segment size is chosen too large to create a complex segment (CR property);
- (ii) periodic images or images where the high and low luminance are equally spread in segments (Weber property).

2.2.3. *AME*. Let an image $x(n, m)$ be split into $k_1 k_2$ blocks of $W_{k,l}(i, j)$; the AME is defined as

$$\text{AME}_{k_1 k_2} = -\frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} \left[20 \ln \left(\frac{I_{\max;k,l} - I_{\min;k,l}}{I_{\max;k,l} + I_{\min;k,l}} \right) \right]. \quad (12)$$

From the definition of Michelson contrast in (2) we have

$$\text{AME}_{k_1 k_2} = -\frac{1}{k_1 k_2} \sum_{l=1}^{k_2} \sum_{k=1}^{k_1} [20 \ln (C_{M,k,l})]. \quad (13)$$

This means that AME measure of enhancement for an image is an average of Michelson contrast for each block $W_{k,l}$, in a logarithmic form, over the entire image. The AME measure is suitable for images with the following properties:

- (i) periodic patterns in segments (Michelson property);
- (ii) no randomness in texture (lack of entropy).

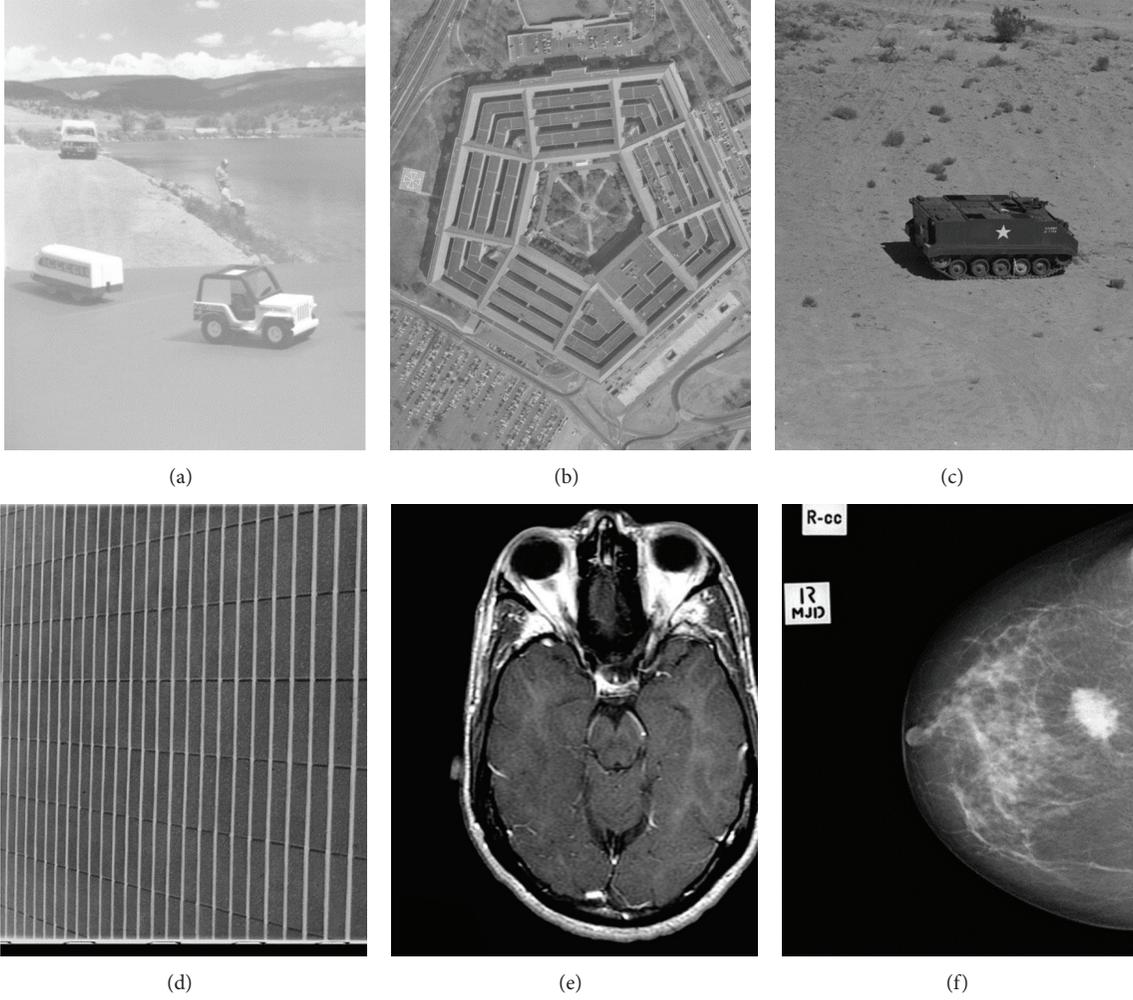


FIGURE 1: Examples of different image attributes. (a) Large uniform background, small targets, no randomness, and no periodic patterns; (b) semiperiodic, nonuniform background; (c) large uniform background, textured in upper part of image, nonperiodic; (d) periodic, nonuniform background; and (e), (f) random texture, nonperiodic, nonuniform background.

This measure lacks ability to analyze images with the following attributes:

- (i) images with uniform background (Michelson property);
- (ii) areas of large uniform luminance in segments (Michelson property);
- (iii) randomness in image texture (entropy property).

2.2.4. AMEE. Let an image $x(n, m)$ be split into $k_1 k_2$ blocks of $W_{k,l}(i, j)$; the AMEE is defined as

$$\begin{aligned} \text{AMEE}_{\alpha k_1 k_2} &= -\frac{1}{k_1 k_2} \\ &\times \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \left[\alpha \left(\frac{I_{\max;k,l} - I_{\min;k,l}}{I_{\max;k,l} + I_{\min;k,l}} \right)^\alpha \ln \left(\frac{I_{\max;k,l} - I_{\min;k,l}}{I_{\max;k,l} + I_{\min;k,l}} \right) \right]. \end{aligned} \quad (14)$$

Using the definition of Michelson contrast (2) and entropy (5) we have

$$\begin{aligned} \text{AMEE}_{\alpha k_1 k_2} &= -\frac{1}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \left[\alpha (C_{M,k,l})^\alpha \ln (C_{M,k,l}) \right] \\ &= \frac{1}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} [\text{Entropy}(C_{M,k,l}^\alpha)]. \end{aligned} \quad (15)$$

In comparison with the relationship between EMEE and EME methods, the AMEE is simply the entropy-based measure of AME. In other words, AMEE is the average of entropy of the Michelson law for each block $W_{k,l}$ over the entire image, scaled by α .

The AMEE measure is suitable for images with the following properties:

- (i) periodic patterns in segments (Michelson property);
- (ii) no area of large uniform luminance in segments (Michelson property);

- (iii) being able to handle additional randomness in texture (entropy).

The AMEE measure does not perform well for the images with the following properties:

- (i) images with large uniform background (Michelson property).

2.2.5. *LogAME*. Let an image $x(n, m)$ be split into $k_1 k_2$ blocks of $W_{k,l}(i, j)$; the LogAME is defined as

$$\begin{aligned} \log\text{AME}_{k_1 k_2} &= \frac{1}{k_1 k_2} \otimes \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \left[\frac{1}{20} \otimes \ln \left(\frac{I_{\max;k,l} \ominus I_{\min;k,l}}{I_{\max;k,l} \oplus I_{\min;k,l}} \right) \right]. \end{aligned} \quad (16)$$

This measure is similar to AME measure, based on Michelson contrast for each block $W_{k,l}$, in a logarithmic form, over the entire image. However, in this measure the arithmetic operations ($*$, $+$, and $-$) were replaced by the PLIP arithmetic operations \otimes , \oplus , and \ominus [5]. The coefficient changes and the sign change are to provide a comparable numeric return to AME method. Using the log and the PLIP operations will put more emphasis on areas with low luminance.

If we define a version of Michelson contrast that uses the PLIP operators as

$$\text{Log}C_M = \frac{I_{\max} \ominus I_{\min}}{I_{\max} \oplus I_{\min}}, \quad (17)$$

we can rewrite the LogAME measure as

$$\log\text{AME}_{k_1 k_2} = \frac{1}{k_1 k_2} \otimes \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \left[\frac{1}{20} \otimes \ln \left(\text{Log}C_{M,k,l} \right) \right]. \quad (18)$$

The LogAME measure is suitable for images with the following properties:

- (i) periodic patterns in segments (Michelson property);
- (ii) unlike AME, LogAME can better handle areas with large uniform luminance in blocks or between blocks (PLIP property).

For images with small differences between target and background luminance, LogAME will behave similarly to AME.

The LogAME measure will not be the best choice for measure of enhancement for images with the following attributes:

- (i) images with small targets and a large background (Michelson property);
- (ii) images with small difference between the background luminance and the target luminance (PLIP property);
- (iii) images with randomness in texture (entropy property).

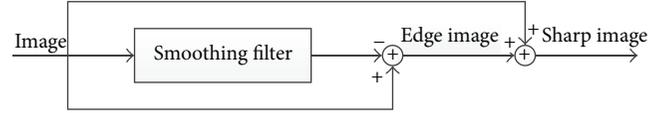


FIGURE 2: Unsharp masking contrast enhancement.

2.2.6. *LogAMEE*. Let an image $x(n, m)$ be split into $k_1 k_2$ blocks of $W_{k,l}(i, j)$; the LogAMEE is defined as

$$\begin{aligned} \log\text{AMEE}_{\alpha k_1 k_2} &= \frac{1}{k_1 k_2} \otimes \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \left[\left(\frac{I_{\max;k,l} \ominus I_{\min;k,l}}{I_{\max;k,l} \oplus I_{\min;k,l}} \right) \right. \\ &\quad \left. * \ln \left(\frac{I_{\max;k,l} \ominus I_{\min;k,l}}{I_{\max;k,l} \oplus I_{\min;k,l}} \right) \right]. \end{aligned} \quad (19)$$

This measure is similar to AMEE measure; it is an entropy-base measure of AME which is the average of entropy of the Michelson law for each block $W_{k,l}$ over the entire image, but the arithmetic operations were replaced by the PLIP arithmetic operations. Using the log and the PLIP operations will put more emphasis on areas with low luminance.

Combining the same techniques that we used in LogAME and EMEE measures we get

$$\log\text{AMEE}_{\alpha k_1 k_2} = \frac{1}{k_1 k_2} \otimes \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \left[\text{entropy} \left(\text{Log}C_{M,k,l} \right) \right]. \quad (20)$$

The LogAMEE measure is suitable for images with the following properties:

- (i) periodic patterns in segments (Michelson property);
- (ii) unlike AMEE, it can better handle areas with large uniform luminance in segments (PLIP property);
- (iii) in comparison with LogAME, it can handle additional randomness in texture (entropy property);
- (iv) for images with small differences between target and background luminance, LogAMEE will behave similarly to AMEE.

The LogAMEE measure is not a good measure for images with these properties:

- (i) images with small targets and a large background (Michelson property);
- (ii) images with small difference between the background luminance and the target luminance (PLIP property).

2.3. *Choosing the Measure for Mammograms*. In a mammogram image, the large black area beside the breast is not considered the image background (the adaptation state of the eye); instead we consider the breast itself as the background and the malignant tissues, cysts, and calcifications as the target. Hence, the image does not include a large uniform

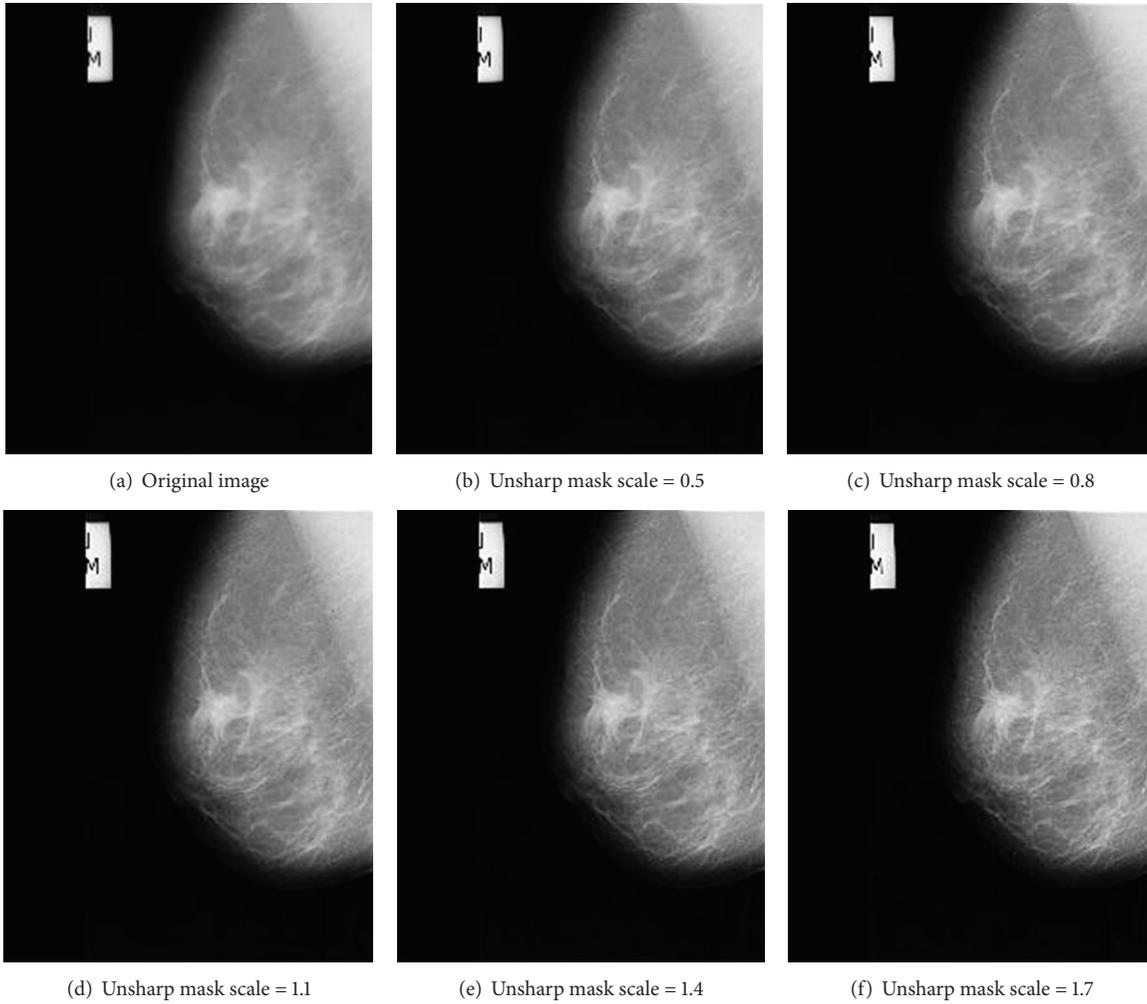


FIGURE 3: Image enhancement using unsharp masking. Higher unsharp mask scale results in higher contrast in enhanced image.

background, which makes EME and EMEE measures inappropriate quality measures to use for these images. The lack of the presence of a large uniform background makes Michelson contrast base measures (AME, AMEE, LogAME, and LogAMEE) good candidates for mammograms. Considering the small difference between the luminance of breast tissue and the luminance of the area containing the abnormality, it is not recommended to use the LogAME and LogAMEE measures, which are best suited for images with the large luminance difference between the target and the background. Also, because of the display of soft tissues in X-ray images, mammograms tend to be textured and have a fair amount of randomness in the target area of the image. This attribute aligns with the properties of entropy-base measures, which can handle images with randomness and texture better. Therefore, considering the above argument about the nature of the mammograms, we recommend using AMEE measure of image quality when evaluating these images.

In the next section, we compare the performance of these measures on a database of mammogram images.

3. Results

For our study, we used mammograms from “the mini-MIAS database of mammograms” [14]. To evaluate the performance of different measures, each mammogram was enhanced using unsharp masking contrast enhancement (Figure 2), alpha-rooting, CLAHE, and Lee’s enhancement algorithm [15, 16]. As demonstrated in Figure 3, increasing the unsharp mask scale will result in enhancement of image contrast. We used this scale as the ground truth for studying the performance of different quality measures. The enhanced images from [14] were tested using different quality measures.

Figure 4 shows an example of images from Figure 3 evaluated by the AMEE measure using different segment sizes. As the scaling constant increases, the image will have higher contrast. A lower AMEE measure also shows higher image contrast. For an optimal segment size, $k > 16 \times 16$ pixels, a monotonic decrease in the AMEE indicates that the measure is correctly evaluating the image quality. Controlling the parameter α in the AMEE measure can help to better address

TABLE 1: Average Pearson correlation for all 1932 images from MIAS database of mammograms (322 images with 6 enhanced versions of each image). The AMEE measure shows the best performance between the measures we tested.

Quality measure	EME	EMEE	AME	AMEE	LogAME	LogAMEE
Average correlation with enhancement	0.0781	0.0222	0.7854	0.9974	0.7851	0.3254

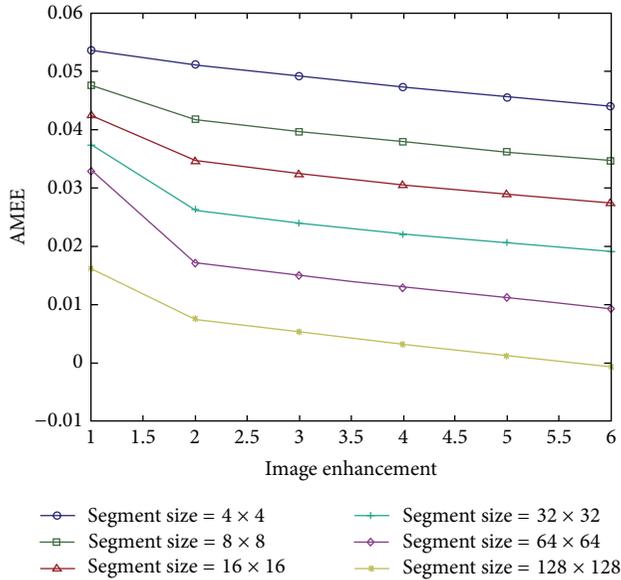


FIGURE 4: Enhanced image from Figure 3 was tested by AMEE measure. Higher image index (x -axis) means higher image contrast; also lower AMEE measure (y -axis) indicates more enhanced image.

w randomness in images. In our study, we set the parameter α to 0.1.

To quantify the overall performance of each quality measure over the entire MIAS database [14], we computed the Pearson correlation between each enhanced image and each quality measure and averaged over the entire database. There are 322 mammogram images in MIAS database [14] and we introduced 6 enhanced versions for each mammogram. For each measure of enhancement, we averaged the Pearson correlation for the 1932 images and results are shown in Table 1. When the average correlation is close to 1, it means that the measure and the enhancement are closely agreeing with each other, and a negative correlation indicates that the measure was not able to correctly evaluate the image enhancement. As shown in Table 1, the AMEE measure exhibits the best performance by predicting the image enhancement correctly in over 99.7% of images.

4. Conclusions

Like most medical imaging systems, mammography requires enhancement of low quality images due to the X-ray hardware limitations. Almost all image enhancement processes require reliable evaluation of the image quality to help with the parameter selection and optimization of enhancement process. In this study, we examined several image quality measures against a database of mammograms using different

enhancement processes. We provided our recommendation for the optimal image quality assessment and its parameters for this category of images.

Conflict of Interests

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

References

- [1] R. D. Rosenberg, B. C. Yankaskas, L. A. Abraham et al., "Performance benchmarks for screening mammography," *Radiology*, vol. 241, no. 1, pp. 55–66, 2006.
- [2] Z. Yicong, K. Panetta, and S. Agaian, "Human visual system based mammogram enhancement and analysis," in *Proceedings of the 2nd International Conference on Image Processing Theory Tools and Applications (IPTA '10)*, pp. 229–234, 2010.
- [3] K. Panetta, Y. Zhou, S. Agaian, and H. Jia, "Nonlinear unsharp masking for mammogram enhancement," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 6, pp. 918–928, 2011.
- [4] S. S. Agaian, B. Silver, and K. A. Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy," *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 741–758, 2007.
- [5] K. A. Panetta, E. J. Wharton, and S. S. Agaian, "Human visual system-based image enhancement and logarithmic contrast measure," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, no. 1, pp. 174–188, 2008.
- [6] B. Silver, S. Agaian, and K. Panetta, "Contrast entropy based image enhancement and logarithmic transform coefficient histogram shifting," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 1–5 of *Speech Processing*, pp. 633–636, March 2005.
- [7] E. J. Wharton, K. A. Panetta, and S. S. Agaian, "Human visual system based image enhancement," in *Mobile Multimedia/Image Processing for Military and Security Applications*, S. S. Agaian and S. A. Jassim, Eds., vol. 6579, April 2007.
- [8] G. Chen, K. Panetta, and S. Agaian, "No reference color image quality measures," in *Proceedings of the IEEE International Conference on Cybernetics (CYBCONF '13)*, pp. 243–248, Lausanne, Switzerland, June 2013.
- [9] K. Panetta, G. Chen, and S. Agaian, "No reference color image contrast and quality measures," *IEEE Transactions on Consumer Electronics*, vol. 59, pp. 643–651, 2013.
- [10] A. Samani, K. Panetta, and S. Agaian, "Transform domain measure of enhancement—TDME—for security imaging applications," in *Proceedings of the IEEE International Conference on Technologies for Homeland Security (HST '13)*, pp. 265–270, 2013.
- [11] S. S. Agaian, K. Panetta, and A. M. Grigoryan, "Transform-based image enhancement algorithms with performance measure," *IEEE Transactions on Image Processing*, vol. 10, no. 3, pp. 367–382, 2001.

- [12] A. A. Michelson, *Studies in Optics*, The University of Chicago Press, Chicago, Ill, USA, 1927.
- [13] E. Wharton, K. Panetta, and S. Agaian, "Human visual system based similarity metrics," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '08)*, pp. 685–690, Suntec, Singapore, October 2008.
- [14] J. Suckling, *The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica*, International Congress Series 1069, 1994.
- [15] L. Jong-Sen, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, no. 2, pp. 165–168, 1980.
- [16] R. Gonzalez, *Digital Image Processing*, 2007.

Research Article

Automated White Matter Hyperintensity Detection in Multiple Sclerosis Using 3D T2 FLAIR

Yi Zhong,^{1,2} David Utriainen,² Ying Wang,³ Yan Kang,¹ and E. Mark Haacke^{1,2,3,4}

¹ School of Sino-Dutch Biomedical and Information Engineering, Northeastern University, Shenyang, Liaoning 110004, China

² Magnetic Resonance Innovations Inc., 440 E. Ferry Street, Detroit, MI 48202, USA

³ Department of Biomedical Engineering, Wayne State University, Detroit, MI 48201, USA

⁴ Magnetic Resonance Imaging Institute for Biomedical Research, 440 E. Ferry Street, Detroit, MI 48202, USA

Correspondence should be addressed to E. Mark Haacke; nmrimaging@aol.com

Received 11 April 2014; Revised 9 July 2014; Accepted 10 July 2014; Published 22 July 2014

Academic Editor: Sos Agaian

Copyright © 2014 Yi Zhong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

White matter hyperintensities (WMH) seen on T2WI are a hallmark of multiple sclerosis (MS) as it indicates inflammation associated with the disease. Automatic detection of the WMH can be valuable in diagnosing and monitoring of treatment effectiveness. T2 fluid attenuated inversion recovery (FLAIR) MR images provided good contrast between the lesions and other tissue; however the signal intensity of gray matter tissue was close to the lesions in FLAIR images that may cause more false positives in the segment result. We developed and evaluated a tool for automated WMH detection only using high resolution 3D T2 fluid attenuated inversion recovery (FLAIR) MR images. We use a high spatial frequency suppression method to reduce the gray matter area signal intensity. We evaluate our method in 26 MS patients and 26 age matched health controls. The data from the automated algorithm showed good agreement with that from the manual segmentation. The linear correlation between these two approaches in comparing WMH volumes was found to be $Y = 1.04X + 1.74$ ($R^2 = 0.96$). The automated algorithm estimates the number, volume, and category of WMH.

1. Introduction

Multiple sclerosis (MS) is considered an autoimmune inflammatory demyelinating disease affecting the central nervous system. It manifests as white matter hyperintensities (WMH) as seen on T2-weighted imaging (WI) using magnetic resonance imaging (MRI). The high signal intensity lesions are a hallmark of MS and are believed to represent inflammation associated with the disease. It has been reported that WMH lesion volume and brain atrophy are independent risk factors for conversion to MS [1, 2]. After diagnosis, MS patients are followed longitudinally and often receive MR imaging multiple times to monitor lesion development. This necessitates the need for a tool which radiologists, neurologists, and MS researchers can use to quantify parenchymal WMH accurately and efficiently.

Although T2-weighted imaging remains important in imaging MS patients, its ability to delineate WMH is usually hampered by the fact that WMH and cerebral spinal fluid

(CSF) are both bright. This drawback is overcome by using T2-weighted fluid attenuated inversion recovery (FLAIR) which suppresses the CSF signal and yet maintains good contrast between the lesions and the white matter (WM). Despite this advantage, several challenges remain for quantifying WMH with FLAIR images, including (1) decreased contrast between gray matter (GM) and WM especially in elderly patients; (2) the major spatial variations in the MR images caused by variable radiofrequency response across the brain (this is referred to as the bias field); and (3) background noise which makes it difficult to separate lesions from white matter. Our goal in this paper is to create an automated white matter lesion detection algorithm which is capable of estimating the volume of WMH accurately, allowing for future ease of use in a clinical setting.

Many automatic WMH quantification techniques for MS lesion detection are reported in the literature [3–7]. Most involve the use of various MRI techniques including T1WI, T2WI, spin density WI, and 2D FLAIR. Although using

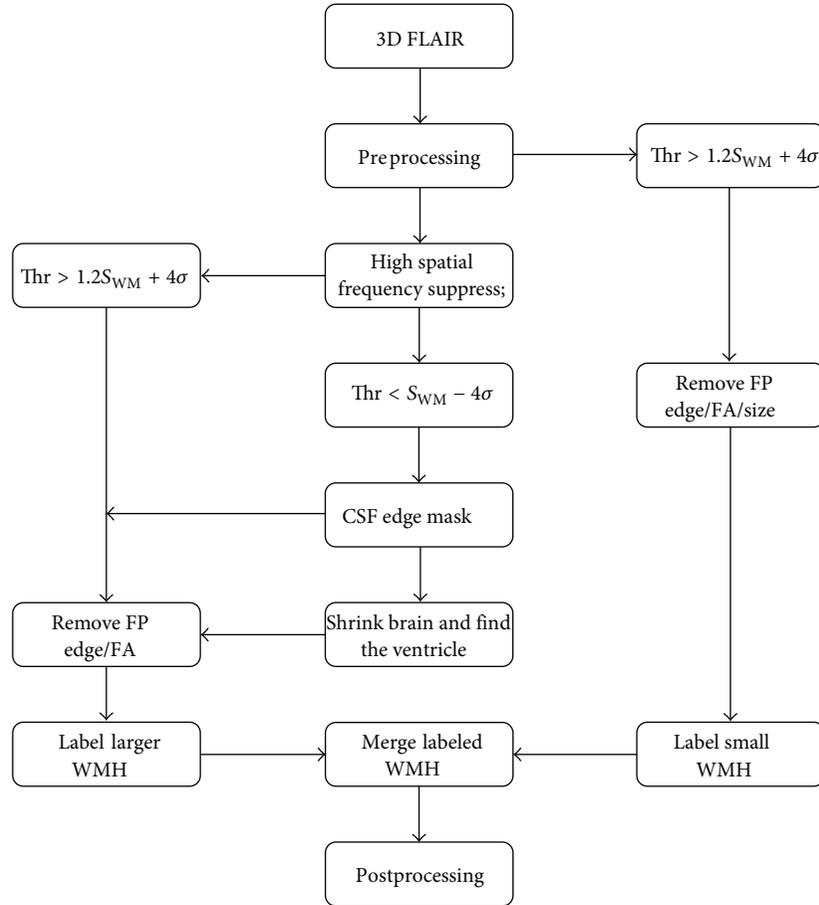


FIGURE 1: Overall work flow of the automatic quantification of WMH using 3D FLAIR images.

multiple data sets provides more objective information to identify the WMH, it requires more time and co-registration processes are needed. Comparing 2D FLAIR image and 3D FLAIR images, the latter provide the best resolution, signal-to-noise, and contrast-to-noise and minimize partial volume effects for detecting WMH [8, 9]. Our goal is to use only 3D T2 FLAIR images to automate the extraction of brain volume, CSF volume, ventricle volume, WMH count, and WMH volume and to automate the categorization of the WMH based on their locations into either deep WM hyperintensities (DWMH) or periventricular hyperintensities (PVH).

2. Materials and Methods

2.1. Data Acquisition. Twenty-six (26) MS patients (age range: 27–51 years, obtained from Synergy Health Concepts, CA) with a diagnosis of relapsing-remitting multiple sclerosis (RRMS) and twenty-six (26) age matched healthy controls (obtained from Wayne State University, MI) were imaged using 3T Siemens scanners (Siemens, Erlangen, Germany). A TRIO scanner was used at Synergy Health Concepts while a VERIO scanner was used at Wayne State University. Institutional review board approval was obtained for this study from both locations and all subjects signed informed consent. MS patient and volunteer data were collected using a

12-channel head and neck coil arrangement at both sites. Both scanners used the same 3D T2 FLAIR sequence parameters. The imaging parameters for the 3D T2 FLAIR sequence were repetition time (TR) = 6000 ms, echo time (TE) = 396 ms, inversion time (TI) = 2200 ms, flip angle (FA) = 120°, $N_x \times N_y = 512 \times 512$, field-of-view (FOV) = 256 mm \times 256 mm, in-plane resolution = 0.50 mm \times 0.50 mm, and slice thickness = 1.0 mm.

2.2. Data Processing. Two novel concepts are introduced to separate WM from the surrounding tissue. One is a high spatial frequency suppression method to remove false GM boundary contributions and the other is a shape-dependent approach to segment WMH from other confounding contrast changes, since conventional threshold segmentation methods contain numerous false positives. The overall workflow, as illustration of Figure 1, our proposed method consisted of seven steps: (1) skull stripping for the removal of the nonbrain tissue; (2) bias field correction for RF field inhomogeneity; (3) CSF sulcus and ventricle segmentation; (4) removal of high spatial frequency GM areas (with the receiver operating characteristic [ROC] curve being used to determine the Hanning high pass window size and other relevant parameters) for WMH segmentation; (5) small lesion detection,

for lesions less than 0.025 mL; (6) estimating the lesions volume; and (7) automatic labeling of the lesions as DWMH or PVL. Those lesions with a distance less than 3.0 mm to the edge of the ventricle were considered PVH, while all others were considered DWMH. We describe below the detailed implementation of each step.

2.2.1. Preprocessing: Skull Stripping. In order to apply skull stripping to the 3D T2 FLAIR images, we developed a moment of inertia structure tensor method, which is used for brain surface detection, combined with local morphology to remove nonbrain tissue (previous work in [10]).

2.2.2. Preprocessing: Bias Field Correction. For the bias field correction, a modified homodyne high pass filter approach similar to that used in phase processing for susceptibility weighted imaging [11, 12] and later proposed by Axel et al. for magnitude imaging [13] was applied via

$$u(x) = \frac{v(x)}{b(x)} = v(x) \cdot \frac{\kappa}{\text{LPF}(v'(x))}, \quad (1)$$

where $u(x)$ is the inhomogeneity free image, $v(x)$ is the original image, $b(x)$ is the bias field estimate image, and $v'(x)$ is a filtered, dilated version of $v(x)$ as described next. LPF stands for low pass filter, while κ is a constant based on the most populated signal (the mode) in the brain. The idea is to use the LPF version of $v(x)$ to estimate low spatial frequency bias field effects. However, this approach itself produces edge effects wherever there are high contrast regions in the brain or at the boundary of the brain. To alleviate this problem, we modified the original image $v(x)$ to smooth the image. First, we used dilation to extrapolate the brain $v(x)$ in order to reduce high spatial frequency edge artifacts. In order to get smooth edge expansion, we used a 3×3 template to do grayscale dilation over 20 iteration for the brain edge expansion. This new image is then referred to as $v'(x)$. Finally, the bias field image was estimated by low pass filtering of $v'(x)$. The original image $v(x)$ is then normalized to the LPF $v'(x)$, which is then multiplied by κ as shown above.

2.2.3. CSF and Ventricle Extraction. Once the brain has been isolated, the predominant three tissues remaining (in the order of decreasing signal intensities) are GM, WM, and CSF. If there is sufficient contrast between WM and CSF, then we can use the Otsu [14] algorithm to determine statistically the best threshold to remove the pixels with low CSF signal. A binary mask is generated for any value lower than the threshold and the resulting image is then considered as the CSF mask image. Because of the fact that CSF occupies both the subarachnoid space and the ventricle, based on the anatomic connectivity of the subarachnoid and ventricular systems, we applied a large scale 3D erosion of the brain image and multiplied it by the CSF mask image; for the 3D erosion template size we use a ball radius of 50 pixels about 25 mm to make sure we remove most CSF surrounding the brain. The largest remaining connected volume was labeled as the ventricular CSF.

2.2.4. WMH Segmentation with High Spatial Frequency Suppression. White matter lesions have higher intensity and generally clear boundaries relative to the surrounding WM. Unfortunately, in the 3D T2 FLAIR images, the GM can also present with high signal intensity, especially in elderly people. This leads to numerous false positives when attempting to isolate WMH. However, GM structures tend to be narrow; thus they have predominantly high spatial frequency components. If we removed these components, which often tend to mimic lesions in their signal intensities, it may be possible to eliminate many of these false positives. To accomplish this, we created a new low pass filtered image defined via

$$u'(x) = u(x) - \lambda \cdot \text{HP}(u(x)), \quad (2)$$

where $u(x)$ is the original image, $\text{HP}(u(x))$ is the high pass filtered version of $u(x)$, and λ is a constant to be determined. There are three parameters that affect the lesion segmentation result: the window size of the high pass filter, the constant λ , and a WMH threshold value. This threshold is chosen to be $1.2 \cdot S_{\text{WM}} + 4 \cdot \sigma_{\text{WM}}$, where S_{WM} is the mean value of the WM over the entire brain and σ_{WM} is the standard deviation of the noise. The factor of 1.2 raises the baseline of the WM signal closer to that of the GM signal. This is a very conservative threshold designed to avoid false positives. A typical patient case from the patient's group (with many lesions and a total WMH volume more than 10 mL) was chosen to analyze the receiver operator characteristic (ROC) curve to determine the high pass filter window size and λ which minimize the number of false positives. The segmentation results of the candidate lesions were labeled in 3D, and their distance from the ventricular CSF and their fractional anisotropy (FA) values were calculated. Here we use FA value to denote the shape character of the candidate lesion, using the moment of inertial structure tensor [15] instead of the diffusion tensor to calculate FA value. The candidate lesions signal intensity was weighted to calculate the eigenvectors in three orthogonal directions. Finally, candidate lesions were removed if they had either a very small distance to the cortical CSF (2.0 mm) in the sulci or a very thin prolate spheroid shape (high FA value over 0.6).

2.2.5. Small Lesion Segmentation. The high frequency suppression step not only removes the GM boundary false positives, but also tends to remove small lesions (since they are intrinsically high spatial frequency in nature). Here, we define a lesion with less than 0.025 mL (100 pixels) as a small lesion. These can be detected using a higher threshold ($1.2 \cdot S_{\text{WM}} + 4 \cdot \sigma_{\text{WM}}$) in the bias field corrected image. Then the distance from CSF, FA characteristic and ventricle information were utilized to remove the false positives.

2.2.6. Lesion Volume Estimation. Since the edges of bigger lesions were removed in step 4, and only the highest components of the smaller lesions were kept in step 5, we applied a region growing algorithm (dynamic programming) [16] to return the candidate lesions to their actual size. From this point, the number of lesions and the size and intensity of information of each lesion can be measured.

2.2.7. Categorizing the Lesions. Finally, the ventricle position information can be used to determine the category of the lesions. Those lesions with an edge that is less than 3.0 mm to the ventricles are classified as PVH while the rest are classified as DWMH.

2.3. Human Interfacing and Correction Schemes. The total number of lesions and the volume of the lesions were measured through manual segmentation to generate a gold standard for lesion number and volume. To assess the agreement between the gold standard and the proposed method, we used the similarity index (SI) measures:

$$SI = \frac{2 \cdot (V_{\text{auto}} \cap V_{\text{GS}})}{V_{\text{auto}} + V_{\text{GS}}}, \quad (3)$$

where V_{auto} denotes the lesion area obtained from automated segmentation and V_{GS} is the gold standard obtained from the processor's manual segmentation (two processors reviewed each other). The SI value ranges from zero to unity, with zero for total disagreement and unity for a hundred percent agreement of the two methods. The algorithm was implemented in C++ and integrated into our in-house software SPIN (Signal Processing in NMR, Detroit, Michigan).

3. Results

In order to evaluate the method as a whole, the high spatial frequency suppression parameters were first optimized. This was done using the ROC curve approach as shown in Figure 2 where we plot the similarity index as a function of both filter size and constant value λ used to enhance the reduction of high spatial frequency or GM edge information. These results suggest that the best choice of high pass window size is 130 pixels with a λ value of 1.2. With these variables fixed, the algorithm can be run automatically and compared to a manual segmentation approach. An example set of processed images is shown in Figure 3. The DWMH and PVH were denoted by black contour and white contour. The total processing time for each case takes less than 5 minutes (Intel i7 2.8 GHz, 8 G RAM).

The scatter plot of the WMH volume obtained using both the manual segmentation and the automated method is shown in Figure 4. The correlation between these two approaches in comparing WMH volumes for all 26 MS cases was found to be $Y = 1.04 \cdot X + 1.74$ ($R^2 = 0.96$). Clearly, the age matched healthy controls show much less WMH than the MS patients. The results for the similarity index for the WMH volume reveal that the larger lesions were well detected by the proposed method (Figure 5). A review of three random cases showed that lesions as large as 0.15 mL could be missed if the contrast-to-noise in that area was poor but most lesions missed were closer to 0.012 mL.

4. Discussion

We have presented an effective method for automatic segmentation of WMH using 3D T2 FLAIR images. FLAIR imaging is one of the most popular protocols for MS. In

particular, it has been shown that 3D FLAIR reveals more lesions than conventional axial T2WI [17]. Using 3D FLAIR also avoids the need for registration or collection of different datasets such as TIWI, T2WI, and spin density WI when brain volume and lesion load are needed. This algorithm is easy to set up for batch processing of data collected from the same site with the same imaging parameters for each patient. The output of this algorithm includes lesion number, volume, and lesion type such as PVH or DWMH. A particular strength of this approach is that all the image processing steps are integrated within a single interface, with a clear work flow interface that allows the user to change options as desired. Several semiautomated and manual assist methods are also provided to allow the user to override some of the automated components.

Other methods to segment WMH have been presented in the literature [5]. Most other approaches require more than one type of imaging technique. For example, Khayati et al. [18] used 2D FLAIR images with low resolution and the partial volume effect reduced the accuracy of the segmentation results. More recently, Simões et al. [6] developed an automatic segmentation using only 3D FLAIR images but required BET and FAST for the preprocessing and 3D Slicer for evaluation. To the best of our knowledge, our approach is the only integrated software tool using only 3D FLAIR images.

The choice of λ also had an effect on the scaling factor used to raise the WM baseline level to that of the GM when the threshold $p \cdot S_{\text{WM}} + 4 \cdot \sigma_{\text{WM}}$ is used. For $\lambda = 1.2$, p was 1.2 while for $\lambda = 1.4$, p was 1.0. We chose the former based on its higher similarity index. The factor of $4 \cdot \sigma_{\text{WM}}$ appears in the thresholding for small lesions because the background variation of WM was such that the usual choice of $2 \cdot \sigma_{\text{WM}}$ or $3 \cdot \sigma_{\text{WM}}$ still captured false positives.

This study has several limitations. First, the false positives caused by the gray matter boundaries are problematic. We have made major strides in dealing with this by using our high pass filter approach, but there still remained a few false positives that escaped automatic removal. Figure 5 shows the relationship between the similarity index and the candidate WMH volumes. The cases with total lesion volume over 0.5 mL achieved the highest similarity index values (average SI = 0.77). Second, the optimal data processing parameters are dependent on the imaging parameters. If the FLAIR imaging parameters are changed significantly, it may be necessary to redo the ROC analysis to find the optimal parameter settings. Third, we use the location of the lesions relative to CSF in the sulci to remove false positive lesions from GM. However, some of these cannot be easily detected. In addition, the lesions in the GM may also be falsely removed, if they are too close to the CSF. The type of lesion was classified by its distance from the nearest ventricle edge. When the ventricles could not successfully be extracted then the lesions could not be classified correctly.

In conclusion, we have developed an algorithm that automatically estimates the volume and category of WMH using a 3D T2 FLAIR series of images. The automated quantitative algorithm has been shown to correlate well with the manual segmentation result. This approach should make

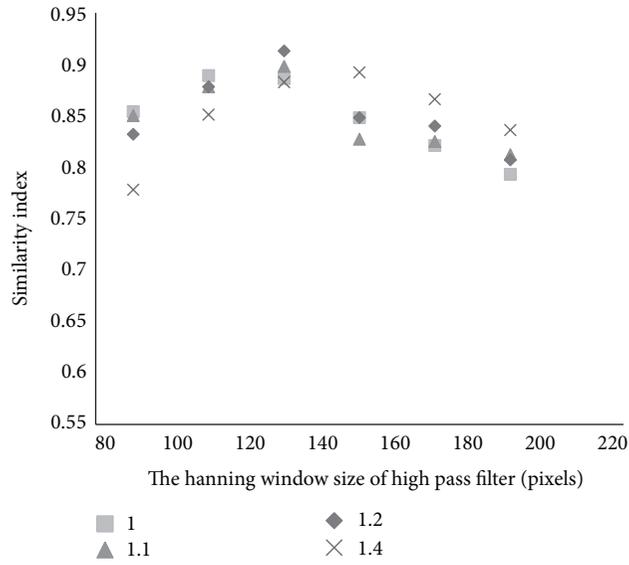


FIGURE 2: Receiver operating characteristic plot to determine the parameters of the high spatial frequency suppression filter.

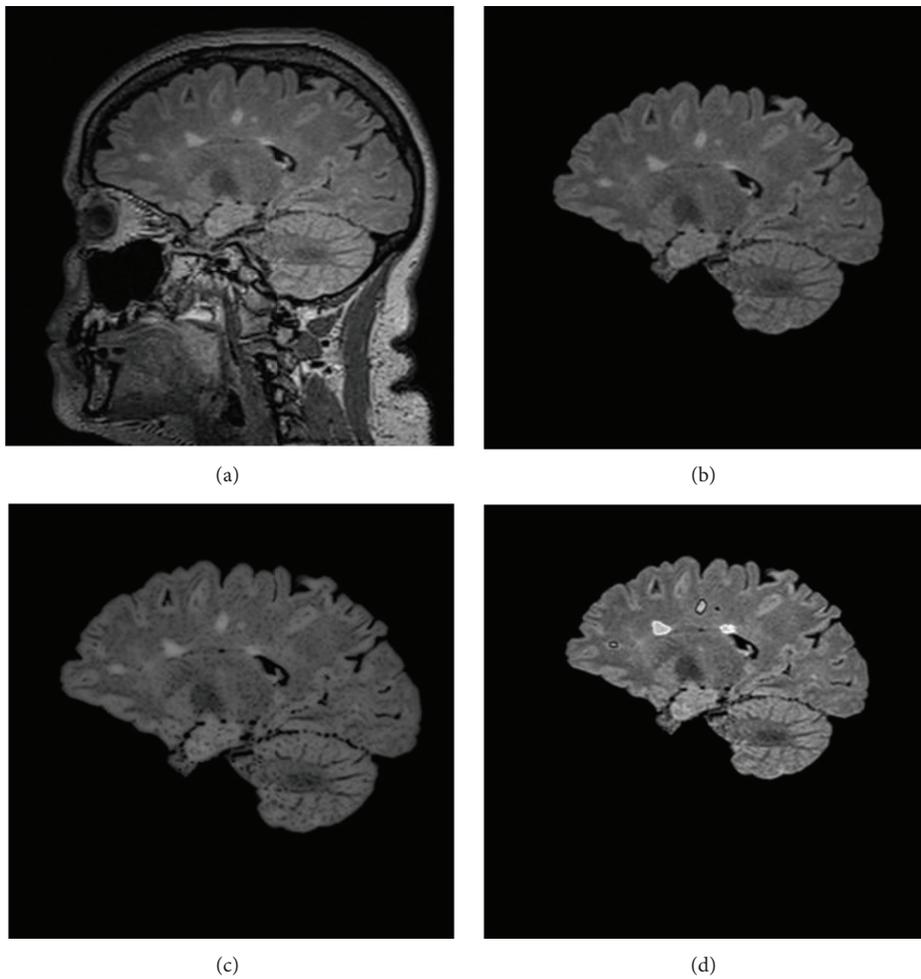


FIGURE 3: A 3D sagittal view FLAIR brain image. (a) Original images; (b) after skull stripping and bias field correction; (c) after high pass suppression of GM edges; and (d) the final segmentation of lesions and assignment of their locations (black contour for DWMH, white contour for PVH).

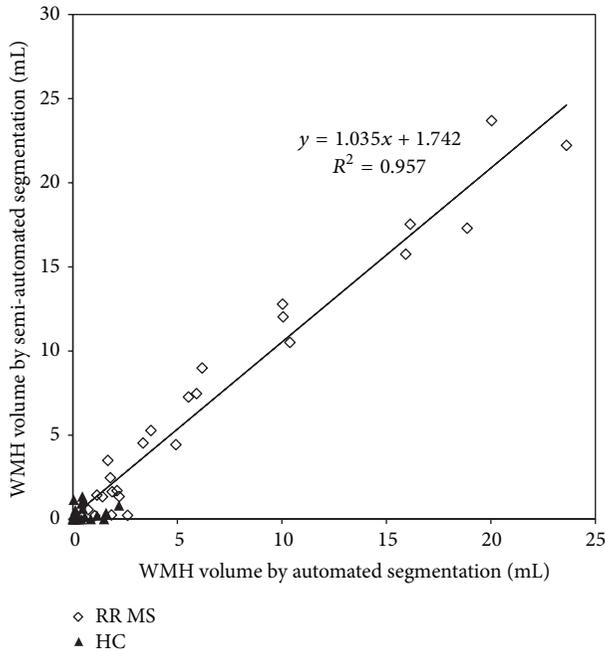


FIGURE 4: Correlation of the manual segmentation and automated segmentation methods. The open diamonds represent the MS patients and the solid triangles represent normal controls.

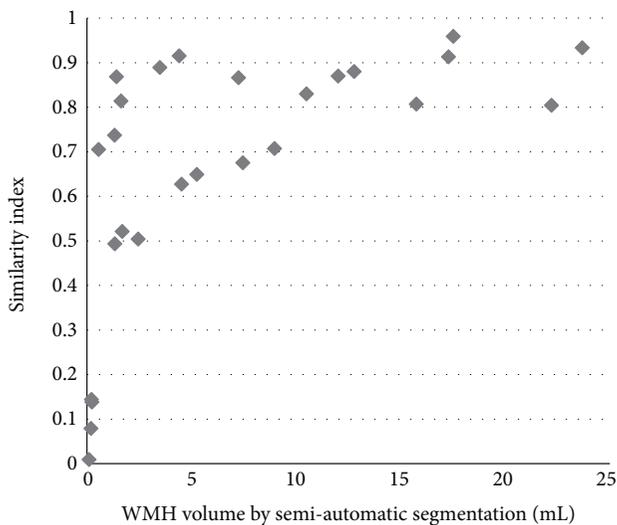


FIGURE 5: Similarity index for all 26 patients plotted over lesion volumes as derived from manual tracing.

it easier to study how the lesion load relates to other factors in MS and to monitor the number and volume of lesions over time.

Conflict of Interests

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

Acknowledgment

This work was supported in part by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award no. R42HL112580. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- [1] D. T. Chard, C. M. Griffin, G. J. M. Parker, R. Kapoor, A. J. Thompson, and D. H. Miller, "Brain atrophy in clinically early relapsing-remitting multiple sclerosis," *Brain*, vol. 125, part 2, pp. 327–337, 2002.
- [2] M. J. Magraner, I. Bosca, M. Simó-Castelló et al., "Brain atrophy and lesion load are related to CSF lipid-specific IgM oligoclonal bands in clinically isolated syndromes," *Neuroradiology*, vol. 54, no. 1, pp. 5–12, 2012.
- [3] D. Mortazavi, A. Z. Kouzani, and H. Soltanian-Zadeh, "Segmentation of multiple sclerosis lesions in MR images: a review," *Neuroradiology*, vol. 54, no. 4, pp. 299–320, 2012.
- [4] T. Magome, H. Arimura, S. Kakeda et al., "Automated segmentation method of white matter and gray matter regions with multiple sclerosis lesions in MR images," *Radiological Physics and Technology*, vol. 4, no. 1, pp. 61–72, 2011.
- [5] P. Schmidt, C. Gaser, M. Arsic et al., "An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis," *NeuroImage*, vol. 59, no. 4, pp. 3774–3783, 2012.
- [6] R. Simões, C. Mönninghoff, M. Dlugaj et al., "Automatic segmentation of cerebral white matter hyperintensities using only 3D FLAIR images," *Magnetic Resonance Imaging*, vol. 31, no. 7, pp. 1182–1189, 2013.
- [7] D. García-Lorenzo, S. Francis, S. Narayanan, D. L. Arnold, and D. L. Collins, "Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging," *Medical Image Analysis*, vol. 17, no. 1, pp. 1–18, 2013.
- [8] I. L. Tan, P. J. W. Pouwels, R. A. van Schijndel, H. J. Adèr, R. A. Manoliu, and F. Barkhof, "Isotropic 3D fast FLAIR imaging of the brain in multiple sclerosis patients: initial experience," *European Radiology*, vol. 12, no. 3, pp. 559–567, 2002.
- [9] G. H. Chagla, R. F. Busse, R. Sydnor, H. A. Rowley, and P. A. Turski, "Three-dimensional fluid attenuated inversion recovery imaging with isotropic resolution and nonselective adiabatic inversion provides improved three-dimensional visualization and cerebrospinal fluid suppression compared to two-dimensional flair at 3 tesla," *Investigative Radiology*, vol. 43, no. 8, pp. 547–551, 2008.
- [10] Y. Zhong, S. Qi, Y. Kang, W. Feng, and E. M. Haacke, "Automatic skull stripping in brain MRI based on local moment of inertia structure tensor," in *Proceeding of the IEEE International Conference on Information and Automation (ICIA '12)*, pp. 437–440, Shenyang, China, June 2012.
- [11] E. M. Haacke, Y. Xu, Y. C. Cheng, and J. R. Reichenbach, "Susceptibility weighted imaging (SWI)," *Magnetic Resonance in Medicine*, vol. 52, no. 3, pp. 612–618, 2004.
- [12] J. R. Reichenbach, R. Venkatesan, D. J. Schillinger, D. K. Kido, and E. M. Haacke, "Small vessels in the human brain: MR venography with deoxyhemoglobin as an intrinsic contrast agent," *Radiology*, vol. 204, no. 1, pp. 272–277, 1997.

- [13] L. Axel, J. Costantini, and J. Listerud, "Intensity correction in surface-coil MR imaging," *American Journal of Roentgenology*, vol. 148, no. 2, pp. 418–420, 1987.
- [14] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975.
- [15] Y. Zhong, S. Qi, Y. Kang, W. Feng, and E. M. Haacke, "Automatic skull stripping in brain MRI based on local moment of inertia structure tensor," in *Proceedings of the IEEE International Conference on Information and Automation (ICIA '12)*, pp. 437–440, Shenyang, China, June 2012.
- [16] J. Jiang, E. M. Haacke, and M. Dong, "Dependence of vessel area accuracy and precision as a function of MR imaging parameters and boundary detection algorithm," *Journal of Magnetic Resonance Imaging*, vol. 25, no. 6, pp. 1226–1234, 2007.
- [17] R. H. Hashemi, W. G. Bradley Jr., D. Chen et al., "Suspected multiple sclerosis: MR imaging with a thin-section fast FLAIR pulse sequence," *Radiology*, vol. 196, no. 2, pp. 505–510, 1995.
- [18] R. Khayati, M. Vafadust, F. Towhidkhah, and M. Nabavi, "Fully automatic segmentation of multiple sclerosis lesions in brain MR FLAIR images using adaptive mixtures method and Markov random field model," *Computers in Biology and Medicine*, vol. 38, no. 3, pp. 379–390, 2008.

Research Article

Nonreference Medical Image Edge Map Measure

Karen Panetta,¹ Chen Gao,¹ Sos Aгаian,² and Shahan Nercessian³

¹ Department of Electrical and Computer Engineering, Tufts University, Medford, MA 02155, USA

² Department of Electrical and Computer Engineering, The University of Texas at San Antonio, San Antonio, TX 78249, USA

³ MIT Lincoln Laboratory, Lexington, MA 02421, USA

Correspondence should be addressed to Chen Gao; cgchengao@gmail.com

Received 11 April 2014; Accepted 21 June 2014; Published 15 July 2014

Academic Editor: Yicong Zhou

Copyright © 2014 Karen Panetta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Edge detection is a key step in medical image processing. It is widely used to extract features, perform segmentation, and further assist in diagnosis. A poor quality edge map can result in false alarms and misses in cancer detection algorithms. Therefore, it is necessary to have a reliable edge measure to assist in selecting the optimal edge map. Existing reference based edge measures require a ground truth edge map to evaluate the similarity between the generated edge map and the ground truth. However, the ground truth images are not available for medical images. Therefore, a nonreference edge measure is ideal for medical image processing applications. In this paper, a nonreference reconstruction based edge map evaluation (NREM) is proposed. The theoretical basis is that a good edge map keeps the structure and details of the original image thus would yield a good reconstructed image. The NREM is based on comparing the similarity between the reconstructed image with the original image using this concept. The edge measure is used for selecting the optimal edge detection algorithm and optimal parameters for the algorithm. Experimental results show that the quantitative evaluations given by the edge measure have good correlations with human visual analysis.

1. Introduction

Edge detection is an essential preprocessing step for early cancer detection and diagnosis in medical image processing such as medical image segmentation, registration, and reconstruction. For example, accurate edge detection algorithms can be used to track the size of a tumor and this information can help to monitor whether the treatment is effective or not. Traditional edge detection algorithms can be grouped into two categories: one utilizes the first order gradient information and the other uses second derivative zero crossing information. Some popular algorithms include Sobel, Robert, Prewitt, Laplacian, LoG, and Canny algorithm [1]. Some other state-of-the-art edge detection algorithms also include Partial Derivatives of Boolean Functions [2] and the Alpha Weighted Quadratic Filter [3]. Although the performance of most of these detectors is acceptable for simple noise free images, the case is dramatically different for medical images subjected to noises from the acquisition systems [4]. Unfortunately, medical images usually suffer from low contrast or poor resolution due to the limitation

of hardware systems or the exposure time. Therefore, it is necessary to have a reliable evaluation method to measure the performance of different edge detection algorithms and help select the optimal algorithm for specific medical applications.

Many edge measures have been proposed including the full reference edge measure [5], the nonreference measure [6, 7], and the subjective evaluation. The full reference edge measure requires a ground truth image as a reference and compares the similarity between candidate edge maps and the ground truth edge map. However, the ground truth image is not available for medical images. Subjective evaluation ratings by medical experts are the most widely accepted evaluation method in medical image processing. This approach avoids the use of ground truth edge maps. However, it is impossible to remove all the bias and the results can still be inconsistent. Furthermore, subjective evaluation is expensive with respect to time and resources; thus it is difficult to be automated.

The nonreference based method does not require a ground truth and it can be automated. Unfortunately, the existing nonreference edge measures are still far from ideal.

A nonreference based edge map evaluation should only use the information from the resultant edge map and the original image itself to make an evaluation. Yitzhaky and Peli proposed a probabilistic based nonreference measure using the receiver operating characteristics (ROC) [7]. Their method first estimates a ground truth edge map using automatic statistical analysis of the correlation of detection results produced using different detector parameters. Then, the edge map which is most similar to the estimated ground truth is selected as the optimal edge detection results. This method balances specificity and sensitivity. However, this method suffers from its bias regarding the generation of the estimated ground truth, because the candidate edge maps used can directly affect the estimated ground truth. Therefore, if the majority of the edge maps used are not of adequate quality or fail to extract certain features, this will be reflected in the derived estimated ground truth. Also, since the original image data is not used, there is no way to indicate how well the best determined edge detector output from this approach corresponds to the original image.

In this paper, we present a reconstruction based nonreference edge measure. The theoretical basis of this method is that a good edge map captures the essential structures and details of the original images. Therefore, the reconstruction using the pixel information on a better edge map would be more similar to the original image. In our method, the edge measure is composed of two components: the first is the gradient based structural similarity measure between the original image and the reconstructed image, and the second component is the penalty factor. For instance, the reconstructions from edge maps with the most edge pixels have the greatest similarity measure. However, they utilize more information from the original image. To compensate for this, a penalty factor which is inversely proportional to the number of edge pixels is also included in the formulation of the measure. In other words, we want a measure that chooses the optimal edge map as the one that shows the structural details in the image with minimal information and minimal false positives.

The rest of this paper is organized as follows. Section 2 reviews existing reconstruction methods and the similarity measures. More details of Yitzhaky's edge measure are also reviewed in this section. Section 3 presents a new nonreference edge measure (NREM). Section 4 presents the experimental results of using the NREM on choosing the optimal edge detection algorithm and optimal operating parameters for medical images. The comparisons against Yitzhaky's edge measure are also presented in this section. The conclusions are discussed in Section 5.

2. Background

The new NREM is a reconstruction based edge measure. In this section, the existing reconstruction methods are reviewed. Similarity measures are used to compare the correlation between the reconstructed image and the original image. Varieties of similarity measures are also reviewed in this section. Lastly, as a comparison to the NREM, theoretical

analysis and the basic steps of Yitzhaky's measure are shown in this section.

2.1. Reconstruction. Interpolation has been widely used to obtain the missed pixels from the original image. In the context of reconstruction, the pixels along edges are used to predict the pixel values in the smooth areas. One of the linear interpolation methods [8] can be described as follows. For each pixel location $(i, j) \notin e_D$, the algorithm searches in the four horizontal and vertical directions and four diagonal directions for the nearest pixel in the given direction that $\in e_D$. The inverse of the distances of the first pixel encountered in each direction from the given pixel d_k is then used as the weights for the weighted average of their respective image intensity values t_k , yielding the reconstructed intensity value for the given pixel. Thus, reconstruction is carried out for each pixel location $(i, j) \notin e_D$ by the following:

$$r(i, j) = \frac{\sum_{k=1}^8 (1/d_k) t_k}{\sum_{k=1}^8 (1/d_k)}. \quad (1)$$

An improvement of (1) is using a weighted median instead of weighted mean to make it more robust to noise. Another modification utilizes the central weighted median. The central weighted median of a sequence x with weights w is given by (2), where the weights are inversely related to the distance (3) and \diamond is the replication operator representing the fact that intensity value t_i is repeated w_i times in the sequence of median calculation in the following:

$$r(i, j) = \text{median}(t_1 \diamond w_1, t_2 \diamond w_2, \dots, t_8 \diamond w_8), \quad (2)$$

$$w_k = \text{round}\left(\frac{100}{d_k}\right). \quad (3)$$

Another type of reconstruction methods is based on the partial differential equation (PDE) discretization proposed by Ballester et al. [9]. In such methods, high order PDEs are designed to restore smooth regions as well as thin structures. These reconstruction based methods have clear advantages for effectively incorporating the original image information on the edge pixels. In nonreference measures this information is essential because no ground truth exists.

2.2. Similarity Measures. To compare the similarity between two images, the most commonly adopted methods are the statistical methods including the pixel-wised mean square error (MSE) and mean absolute error (MAE). The MSE and MAE between two images x and y are defined as shown in (4). In (4), x and y represent the two images for comparison and i and j represent the pixel locations. These statistical methods have clear physical meanings and are straightforward. Under

these definitions, two images with more similarity have lower MSE or MAE:

$$\begin{aligned} \text{MSE} &= \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [x(i, j) - y(i, j)]^2, \\ \text{MAE} &= \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x(i, j) - y(i, j)|. \end{aligned} \quad (4)$$

However, these statistical methods do not take into consideration the human visual system (HVS) properties. Therefore, they are inappropriate to be used as reliable measures for medical images. Bovik's structural similarity measure (SSIM) [10] is based on the hypothesis that human visual system (HVS) is highly adapted for extracting structural information. The SSIM measure defines the similarity of two images as a function of luminance, contrast, and structure, where the luminance, contrast, and structure are defined as

$$\begin{aligned} l(x, y) &= \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \\ c(x, y) &= \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \\ s(x, y) &= \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}. \end{aligned} \quad (5)$$

Given two images x and y , the μ_x and μ_y represent the means, σ_x and σ_y represent the standard deviation of the x and y image, respectively, and σ_{xy} represent the covariance of x and y . C_1 , C_2 , and C_3 represent constant values. SSIM is a combination of luminance, contrast, and structure measure and it is defined as shown in (6). The SSIM is applied on nonoverlapping windows. Thus the mean of the SSIM values over the entire image (MSSIM (7)) is used to indicate the similarity between two images:

$$\text{SSIM}(x, y) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right), \quad (6)$$

$$\text{MSSIM}(x, y) = \frac{1}{N} \sum_{i=1}^N \text{SSIM}(x_i, y_i). \quad (7)$$

2.3. Yitzhaky and Peli's Edge Measure. Yitzhaky and Peli proposed an objective edge detection evaluation method in [7]. In this paper, the new measure is compared with this method. Yitzhaky's edge measure is briefly reviewed in the following section.

Yitzhaky's edge measure [7] is a nonreference edge measure. This method performs statistical analysis of the correspondence of detection results produced using different detector parameters. The statistical measures are used to estimate the ground truth edge map considering the tradeoff between the true and false edges and to extract the best detector's parameter sets. In Yitzhaky's method, first, an estimated ground truth is automatically constructed by examining the

corresponding threshold receiver operating characteristics (CT-ROC) curve, given a range of detection results obtained from different detection parameter sets. Then, the single parameter set that provides the most similar edge map to the estimated ground truth edge map is identified. The major steps of Yitzhaky's edge measure are summarized in Table 1.

The original Yitzhaky's edge measure performs a general automatic self-evaluation and parameter selection within a range of examined detector parameters. It assumes that the best detection of a certain edge detector in a given image is that which is most consistent with the variety of detection outputs that can be produced by the detection algorithm when different parameters are used. Therefore, by adopting N edge detection results using N sets of parameters of one edge detection algorithm, this algorithm can be used to select the optimal operating parameters. If the N edge detection results are obtained from N edge detectors, this algorithm can be used to compare performances between detection approaches.

3. New Nonreference Edge Map Measure

A generalized block diagram and intermediate results of the established nonreference reconstruction based edge measure (NREM) are shown in Figure 1. It consists of three major steps: grayscale edge map generation, reconstruction, and similarity measure.

(1) Grayscale Edge Map Generation. The edge detection algorithm to be evaluated is applied in this step. Regular edge detection algorithms such as Canny, Sobel, Log, and Roberts can be used. Furthermore, some state-of-the-art edge detection algorithms designed specifically for the medical image applications such as the CLF [11], morphological gradient operator [12], nonlinear diffusion [13], and mathematical morphology [14] can also be used in this step.

After the edge map is subtracted, a morphological dilation is applied to generate a continuous edge map. The dilated edge map is then multiplied with the original image, yielding a grayscale edge map. In this way, the pixels on the dilated edge map contain information from the original image and these pixels are used to predict the pixel intensity in the smooth area.

(2) Reconstruction. In the previous section, four major interpolation based reconstruction methods were reviewed. The weighted average (1) method utilizes all the information from the eight neighbors but is sensitive to noise. Unfortunately, noise commonly exists in medical image applications. To be more robust to noise, the weighted median and central weighted median (2) can be used. In this way, only one of the neighbors is used to predict the new pixel value. This replication solves the noise problem but also results in another problem. That is, in some areas with a low gradient change, such as the breast tissue in the mammogram image, this reconstruction may mistakenly yields a large uniform region.

To get a good balance, a weighted alpha trimmed mean can be used in the reconstruction. Each of the eight neighbors

TABLE 1: Basic steps in Yitzhaky's NR edge measure.

Yitzhaky's edge measure
(1) Generate N edge detection results $D_i, i = 1, \dots, N$ using N combinations of parameters
(2) Generate N potential ground truth $PGT_i, i = 1, \dots, N$
(3) Calculate the average true positive (TP), true negative (TN), false positive (FP), and false negative (FN) rate for each potential ground truth
(4) Construct the correspondence threshold ROC curve (CT-ROC)
(5) Extract the estimated ground truth using either a diagnosis line or the chi-square estimation
(6) Select the best edge map which gives the best match to the estimated ground truth

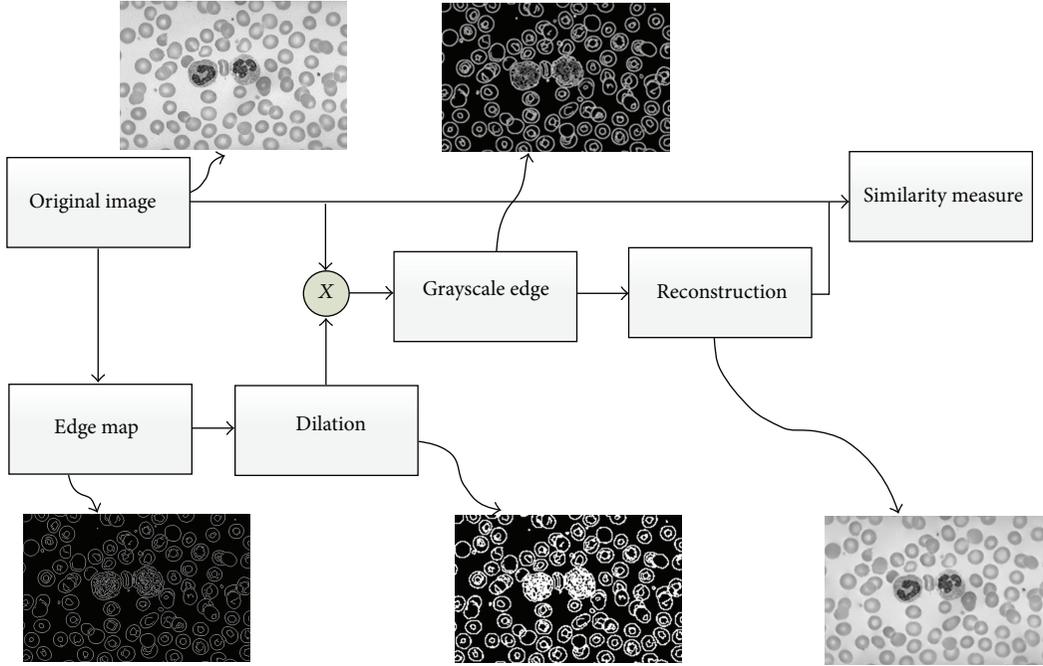


FIGURE 1: Flow and the intermediate results of the new nonreference edge measure.

is assigned with the weighted intensity $x_k = (1/d_k)t_k$, where t_k is the actual edge pixel intensity and d_k is the distance between the pixel to be predicted and edge pixel on a specific direction. Then sort values of all the neighbors in ascending order such that $x_1 \leq x_2 \leq \dots \leq x_K$. Let $T_\alpha = \lceil \alpha K \rceil$ (the nearest integer greater than or equal to αK) be the number of the smallest and largest pixel values to be trimmed or discarded from the sorted sequence, x_1, x_2, \dots, x_K . The alpha trimmed mean [15] is defined by

$$X_\alpha = \frac{1}{K - 2T_\alpha} \sum_{i=T_\alpha+1}^{K-T_\alpha} x_i. \quad (8)$$

The alpha trimmed mean will be different when the parameter α changes. For example, it will be the mean value of the image for $\alpha = 0$ and the median value of the image if α is close to 0.5. In this way, the parameters can be tuned for different applications. In this paper, the results are obtained

by discarding the maximum and minimum neighbors in the calculation of alpha weighted mean.

(3) *Similarity Measure NREM.* The reconstructed image is then compared to the original image using a similarity measure, which is then used as an assessment of the edge map. The SSIM [10] is widely used in clean images, but it is noted that the performance of the SSIM index degraded substantially when assessing Gaussian blurred images. The noise in medical images is very prevalent and difficult to model. In the experiments, the SSIM does not perform well when the images are subjected to other distortions such as low contrast and blurring effects. To measure the similarity between the original image and the reconstructed image, the GSSIM [16, 17] is used. GSSIM suggests that the gradients of the images to be compared be integrated into the image similarity assessment to penalize dissimilarity in image contours and edges. Thus, the GSSIM index makes comparisons between both x and y and the gradients of x and y . The gradients of x and y specifically indicate the

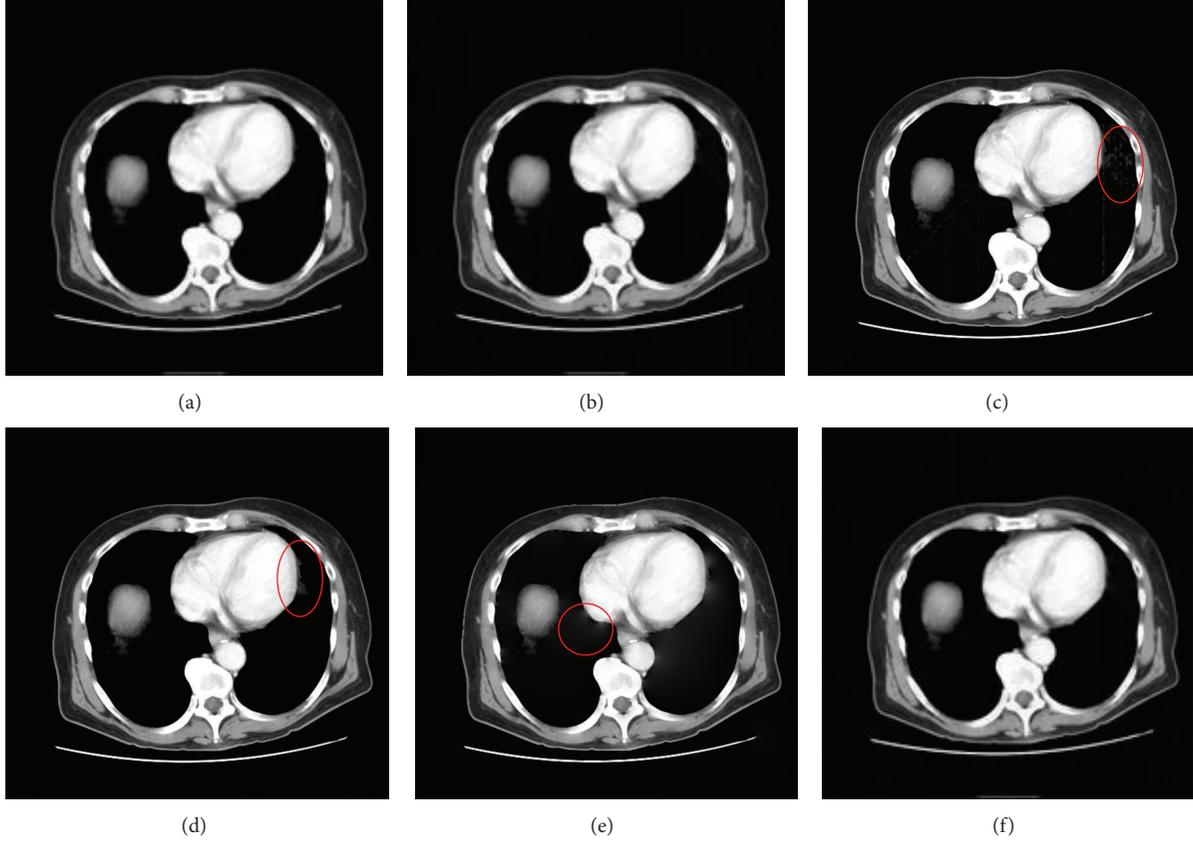


FIGURE 2: Reconstruction results using different interpolation methods. (a) Original CT kidney image. (b)–(f) Reconstructed results from (b) weighted average, (c) weighted median, (d) central weighted median, (e) PDE, and (f) alpha trimmed weighted average.

similarity between edges. In the expression of GSSIM, the contrast and structure terms of SSIM are modified as in (9), where $\sigma_{x'}$ and $\sigma_{y'}$ represent the standard deviation of the gradient magnitude of x and y , respectively:

$$\begin{aligned} c(x', y') &= \frac{2\sigma_{x'}\sigma_{y'} + C_2}{\sigma_{x'}^2 + \sigma_{y'}^2 + C_2}, \\ s(x', y') &= \frac{\sigma_{x'y'} + C_3}{\sigma_{x'}\sigma_{y'} + C_3}. \end{aligned} \quad (9)$$

Using similar methods to fuse luminance, gradient contrast, and gradient structure together, the GSSIM over sub-blocks can be shown in the following:

$$\text{GSSIM}(x, y) = [l(x, y)]^\alpha [c(x', y')]^\beta [s(x', y')]^\gamma. \quad (10)$$

Therefore, the mean of GSSIM over the entire image can be used to indicate the similarity between the reconstructed image and the original image:

$$\text{MGSSIM}(x, y) = \frac{1}{N} \sum_{i=1}^N \text{GSSIM}(x_i, y_i). \quad (11)$$

It is worth noting that the similarity measure itself is not accurate enough to measure the reconstruction performance.

The reason is that when more edge pixels exist in an edge map, more information from the original image is used in the reconstruction which will definitely yield more similar results. Therefore, a penalty factor f_p which is formulated as a decreasing function of the total number of edge pixels is introduced in (12), where N_e represents the total number of edge pixels in the dilated edge map and MN represents the total number of pixels in the original image:

$$f_p = \frac{1}{1 + N_e/MN}. \quad (12)$$

The final nonreference based edge measure (NREM) is comprised as the alpha weighted product of these two terms. In this paper, the results are shown with $\alpha = 1$ and $\beta = 3$, which were obtained experimentally:

$$\text{NREM}(i, e) = (\text{MGSSIM})^\alpha f_p^\beta. \quad (13)$$

4. Experimental Results

Edge detection plays an important role in medical image processing as it determines the structure of objects in images. In this section, we demonstrate some applications of the new nonreference edge measures NREM on medical image processing. The testing images are obtained from the Frederick National Library for Cancer Research Database [18].

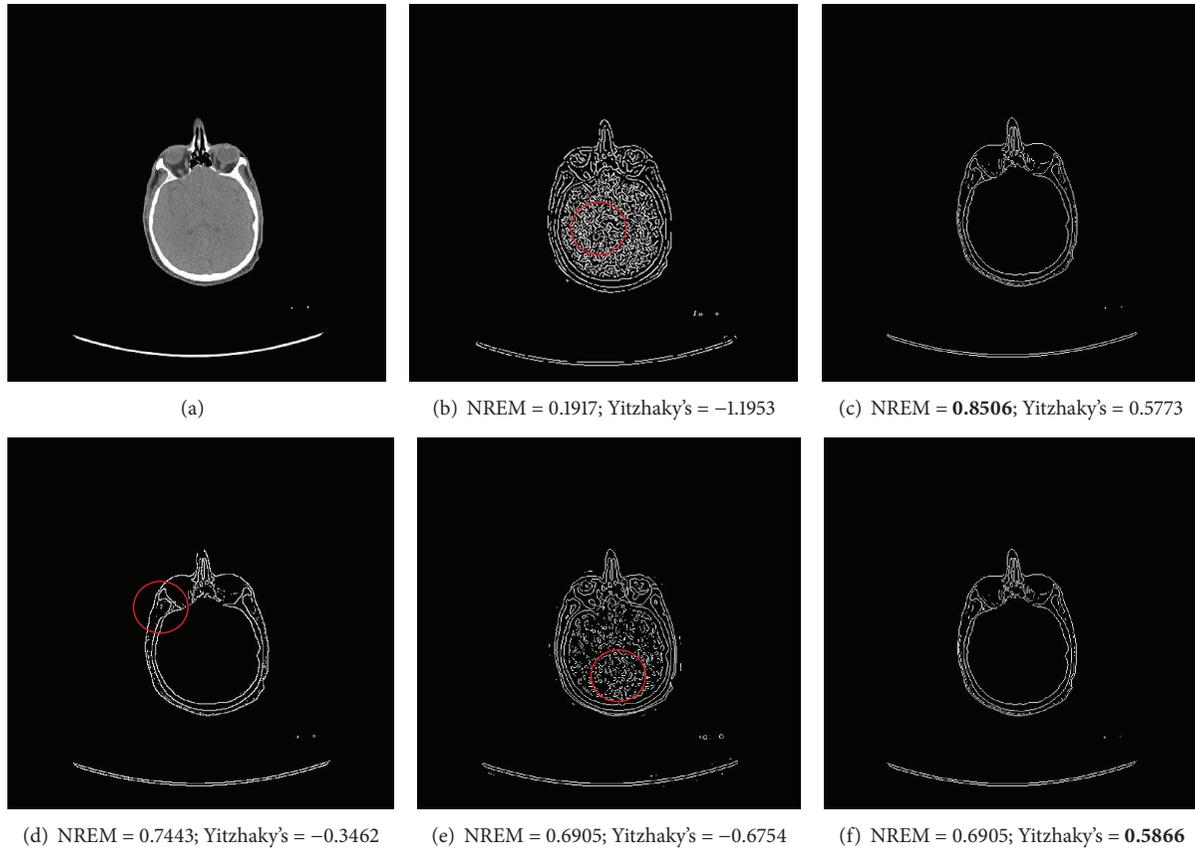


FIGURE 3: Using measures to select optimal edge detection algorithms. (a) Original CT abdomen image. (b)–(f) Edge detection results from (b) Canny, (c) Sobel, (d) Roberts, (e) LoG, and (f) Prewitt. The optimal edge maps with highest NREM value and Yitzhaky's measure value are in bold.

The first example compares the reconstruction results using the weighted mean, weighted median, central weighted median, PDE, and alpha trimmed weighted mean based interpolation. Figure 2 shows the original CT kidney image and five reconstructions. It is seen from the results that, for the low quality medical images, the mean based reconstruction is not as sharp as the median based reconstruction. However, the median based reconstruction introduces some artificial lines. As analyzed before, the alpha trimmed mean can be converted to mean or median filter with different parameter alpha. In Figure 3, the alpha trimmed weighted mean reconstruction achieves a good balance. The central weighted median retrieves details but also introduces false details especially around the edge of the real tissues. The PDE based painting method suffers from severe blurring effects.

The second example is using the nonreference edge measures selecting the optimal edge detection algorithm. In Figure 3, multiple edge detection results from the Canny, Sobel, Roberts, Log, and Prewitt for a CT abdomen image are shown in Figures 3(b)–3(f). These edge detection algorithms shown in Figure 3 are commonly used edge detection algorithms and each has its advantages and disadvantages. For example, gradient based edge detection algorithms such as Sobel and Prewitt are simple but sensitive to noise. The Canny edge detector improves the signal to noise ratio by

smoothing the image; however, the smoothing may lead to loss of corners and detection of double edges. Therefore, it is necessary to have a reliable edge measure that can help to decide the optimal edge detection algorithm for a specific image. The NREM selects the Sobel edge detection result as the optimal and Yitzhaky's method selects the Prewitt. These two edge detection results agree with the visual assessment. As a comparison, the Canny and LoG edge subtract all the soft tissues inside the abdomen, while the Roberts edge has the disconnection problem on the key edges.

Another example of using the edge measure as a means of selecting optimal parameter values is shown in Figure 4. In this experiment, the Sobel edge detection algorithm with different threshold ranging from 0.01 to 0.08 is used. The testing image is an X-Ray chest image which suffers from low contrast. Therefore, lower threshold values tend to keep more soft tissue or other noise components in the edge map, while higher threshold values discard essential edges. The proposed measure selects the optimal parameter at threshold = 0.03 and achieves the best tradeoff between noise removal and feature extraction. In contrast, Yitzhaky's method selects threshold = 0.06 which losses some ribs in the edge map. Figure 4(k) also illustrates the need for the edge pixel density function in the formulation of NREM as the use of MGSSIM alone results in the discussed edge pixel bias.

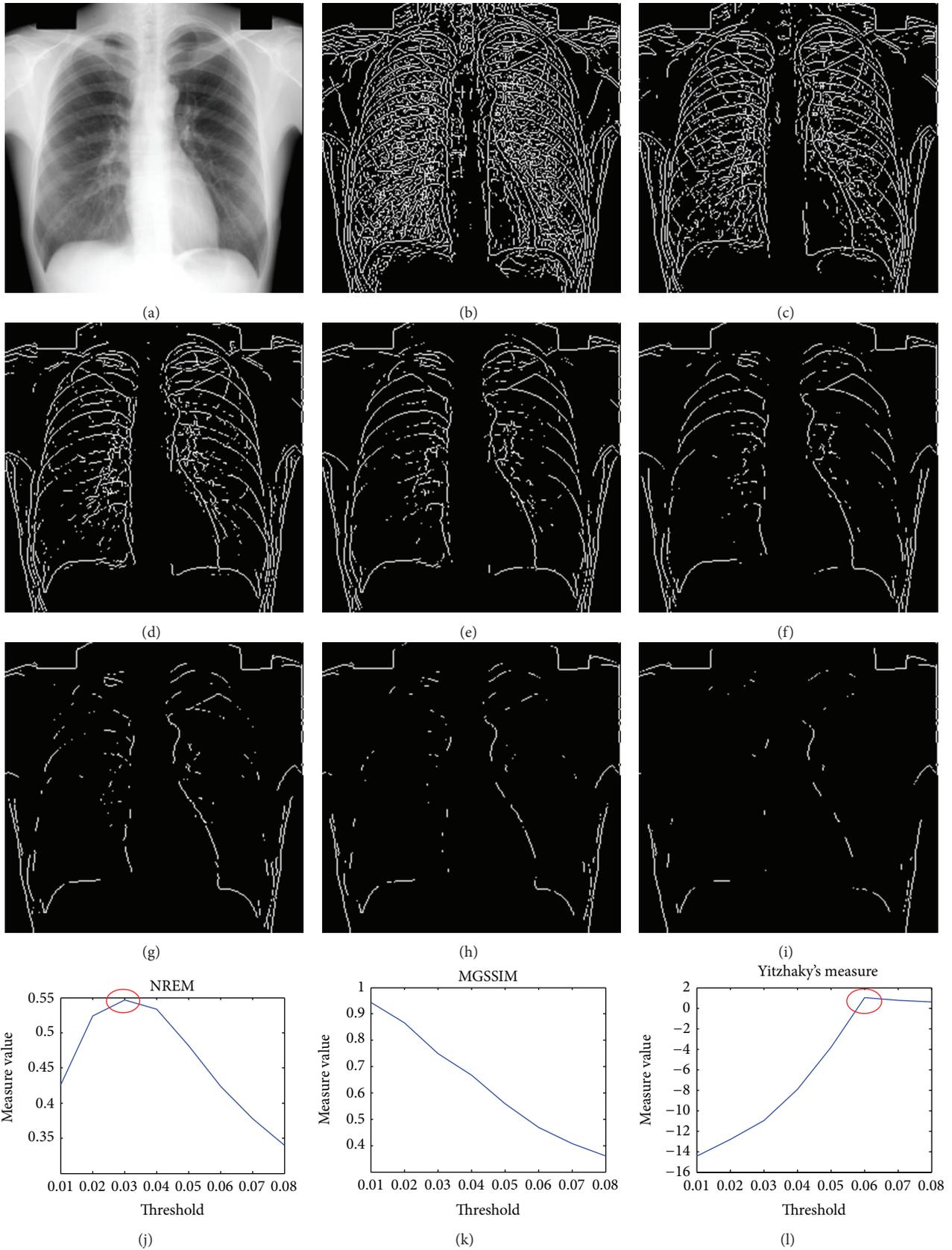


FIGURE 4: Applying edge measures in assisting in selecting optimal operating parameters. (a) Original X-ray chest image. (b)–(i) Edge detection results using the Sobel edge detector with the threshold ranging from 0.01 to 0.08. (j) Presented measure plot indicating $T = 0.03$ as the optimal parameter value. (k) Measure plot of using MGSSIM alone. (l) Yitzhaky's method indicating $T = 0.06$ as the optimal parameter value.

5. Conclusions

Nonreference edge measure is very useful in medical image segmentation, registration, and reconstruction. A new non-reference edge map evaluation NREM for medical applications is proposed in this paper. This measure is based on the fact that the best edge map results consist of the least number of edge pixels at their correct locations needed to characterize all the relevant structures in the reconstruction image. Comparison with state-of-the-art nonreference edge detection measure shows the advantages of the new measure: the NREM utilizes the information from the original image and thus can achieve better performance. Experimental results on using the NREM on selecting the optimal edge detection algorithm and optimal operating parameters show that the measure coincides with subjective evaluation, validating the usefulness of the measure.

Conflict of Interests

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

References

- [1] D. Ziou and S. Tabbone, "Edge detection techniques: an overview," *Pattern Recognition and Image Analysis C/C of Raspoznaniye Obrazov I Analiz Izobrazhenii*, vol. 8, pp. 537–559, 1998.
- [2] E. E. Danahy, S. S. Aгаian, and K. A. Panetta, "Directional edge detection using the logical transform for binary and grayscale images," in *Mobile Multimedia/Image Processing for Military and Security Applications*, vol. 6250 of *Proceedings of SPIE*, pp. 1–12, 2006.
- [3] C. Gao, K. Panetta, and S. Aгаian, "New edge detection algorithms using alpha weighted quadratic filter," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC'11)*, pp. 3167–3172, Anchorage, Alaska, USA, October 2011.
- [4] M. Bennamoun, "Edge detection: problems and solutions," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3164–3169, October 1997.
- [5] S. C. Nercessian, S. S. Aгаian, and K. A. Panetta, "A new reference-based measure for objective edge map evaluation," in *Mobile Multimedia/Image Processing, Security, and Applications*, vol. 7351 of *Proceedings of SPIE*, April 2009.
- [6] S. Nercessian, K. Panetta, and S. Aгаian, "A non-reference measure for objective edge map evaluation," in *Proceeding of the IEEE International Conference on Systems, Man and Cybernetics (SMC'09)*, pp. 4563–4568, San Antonio, Tex, USA, October 2009.
- [7] Y. Yitzhaky and E. Peli, "A method for objective edge detection evaluation and detector parameter selection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 1027–1033, 2003.
- [8] B. Govindarajan, K. A. Panetta, and S. Aгаian, "Image reconstruction for quality assessment of edge detectors," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '08)*, pp. 691–696, October 2008.
- [9] C. Ballester, M. Bertalmio, V. Caselles, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1200–1211, 2001.
- [10] 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.
- [11] J. Quintanilla-Domínguez, M. G. Cortina-Januchs, A. Jevtić, D. Andina, J. M. Barrón-Adame, and A. Vega-Corona, "Combination of nonlinear filters and ANN for detection of microcalcifications in digitized mammography," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '09)*, pp. 1516–1520, October 2009.
- [12] C. Santhaiah, G. A. Babu, and M. U. Rani, "Gray-level morphological operations for image segmentation and tracking edges on medical applications," *International Journal of Computer Science and Network Security*, vol. 9, pp. 131–136, 2009.
- [13] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM Journal on Numerical Analysis*, vol. 29, no. 1, pp. 182–193, 1992.
- [14] Y. Zhao, W. Gui, Z. Chen, J. Tang, and L. Li, "Medical images edge detection based on mathematical morphology," in *Proceedings of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (IEEE-EMBS '05)*, pp. 6492–6495, September 2005.
- [15] J. B. Bednar and T. L. Watt, " α -trimmed means and their relationship to median filters," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 1, pp. 145–153, 1984.
- [16] G.-H. Chen, C.-L. Yang, and S.-L. Xie, "Gradient-based structural similarity for image quality assessment," in *Proceedings of the IEEE International Conference on Image Processing (ICIP'06)*, pp. 2929–2932, Atlanta, Ga, USA, October 2006.
- [17] S. Nercessian, S. S. Aгаian, and K. A. Panetta, "An image similarity measure using enhanced human visual system characteristics," in *Mobile Multimedia/Image Processing, Security, and Applications*, 806310, vol. 8063 of *Proceedings of SPIE*, SPIE Defense, Security, and Sensing, Orlando, Fla, USA, April 2011.
- [18] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Proceedings of the 9th European Conference on Computer Vision (ECCV'06)*, pp. 430–443, Springer, 2006.