International Journal of Biomedical Imaging

Multimodal Imaging and Hybrid Scanners

Guest Editors: Haim Azhari, Robert R. Edelman, and David Townsend

Multimodal Imaging and Hybrid Scanners

International Journal of Biomedical Imaging

Multimodal Imaging and Hybrid Scanners

Guest Editors: Haim Azhari, Robert R. Edelman, and David Townsend

Contents

Multimodal Imaging and Hybrid Scanners, Haim Azhari, Robert R. Edelman, and David Townsend
Volume 2007, Article ID 45353, 2 pages

ESTERR-PRO: A Setup Verification Software System Using Electronic Portal Imaging, Pantelis A. Asvestas, Konstantinos K. Delibasis, Nikolaos A. Mouravliansky, and George K. Matsopoulos
Volume 2007, Article ID 61523, 10 pages

Towards a Noninvasive Intracranial Tumor Irradiation Using 3D Optical Imaging and Multimodal Data Registration, R. Posada, Ch. Daul, D. Wolf, and P. Aletti
Volume 2007, Article ID 62030, 14 pages

Multimodality Data Integration in Epilepsy, Otto Muzik, Diane C. Chugani, Guangyu Zou, Jing Hua, Yi Lu, Shiyong Lu, Eishi Asano, and Harry T. Chugani
Volume 2007, Article ID 13963, 12 pages

Adaptive Bayesian Iterative Transmission Reconstruction for Attenuation Correction in Myocardial Perfusion Imaging with SPECT/Slow-Rotation Low-Output CT Systems, Ji Chen, Ernest V. Garcia, Russell D. Folks, Aharon Peretz, and James R. Galt
Volume 2007, Article ID 18709, 7 pages

Improved Image Fusion in PET/CT Using Hybrid Image Reconstruction and Super-Resolution, John A. Kennedy, Ora Israel, Alex Frenkel, Rachel Bar-Shalom, and Haim Azhari
Volume 2007, Article ID 46846, 10 pages

Registration of Brain MRI/PET Images Based on Adaptive Combination of Intensity and Gradient Field Mutual Information, Jiangang Liu and Jie Tian
Volume 2007, Article ID 93479, 10 pages

Quantitative and O_2 Enhanced MRI of the Pathologic Lung: Findings in Emphysema, Fibrosis, and Cystic Fibrosis, Alfred Stadler, Leopold Stiebellehner, Peter M. Jakob, Johannes F. T. Arnold, Edith Eisenhuber, Isabella von Katzler, and Alexander A. Bankier
Volume 2007, Article ID 23624, 7 pages

Integration of Vibro-Acoustography Imaging Modality with the Traditional Mammography, H. Gholam Hosseini, A. Alizad, and M. Fatemi
Volume 2007, Article ID 40980, 8 pages

Multimodality Imaging of the Peripheral Venous System, Diana Gaitini
Volume 2007, Article ID 54616, 9 pages

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 45353, 2 pages
doi:10.1155/2007/45353

Editorial
Multimodal Imaging and Hybrid Scanners

Haim Azhari,[1] Robert R. Edelman,[2] and David Townsend[3]

[1] *Department of Biomedical Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel*
[2] *Department of Radiology, Evanston Northwestern Healthcare, 2650 Ridge Avenue, Evanston, IL 60201, USA*
[3] *Department of Medicine and Radiology, University of Tennessee Medical Center, 1924 Alcoa Highway, Knoxville, TN 37920, USA*

Received 19 April 2007; Accepted 19 April 2007

There is an old tale of three blind men who were brought to the zoo for the first time and allowed to touch the elephant. On their way home, they shared this exciting experience. "An elephant is a long, flexible, and cylindrical creature," said the man who had touched the elephant's trunk. "No! It is a thin and flat creature," said the man who had touched the elephant's ear. "No, no! An elephant is rough and rigid like a tree stem," disagreed the third person who had touched the elephant's leg. And the truth is that all three men were right!

When trying to see the invisible with our medical imaging systems, we are much like those three blind men. With each imaging modality, we shed some light on a different aspect of the general physiological picture. Although in some cases one modality may suffice to provide a definitive clinical answer, this is not the case in many other situations. Multimodal imaging (MMI) is needed for three basic reasons: (a) to acquire *complementary* information which may be needed to reach a definitive diagnostic conclusion, exclude certain pathologies, or obtain quantitative values (e.g., [1, 2]); (b) to create *synergism* by data fusion (i.e., to provide added information and new images which are more informative than the individual source images); (c) to plan *therapeutic* procedures and monitor treatment (e.g., [3, 4]).

An ideal MMI system or method should be capable of performing all three tasks mentioned above. Naturally, that requirement might be too demanding in terms of technological capabilities and operational considerations. Hence, diagnostic and therapeutic systems are commonly separated. However, one may see in the near future more system integration in the form of image-guided therapy.

There are several technical issues that are associated with MMI. A prerequisite is to obtain effective fusion and display of the data (e.g., [5–7]). Accurate spatial (and maybe also temporal) alignment is crucial for effective data fusion. There are basically two approaches for achieving coregistration. The first, which may be called the "hardware" approach, utilizes a hybrid design comprising two (or more) imaging modalities that are contained within a single device. The advantage of this approach is that the imaging modalities acquire data sequentially while the patient lies on the bed. The disadvantage is the need for dedicated MMI equipment which may be cumbersome or costly.

The second approach for achieving coregistration is the "software-" based approach. With this approach, image properties and tissue geometry and texture are used as clues for aligning the data sets. Alignment is thus achieved by manipulating the acquired data under certain optimization constraints or 3D model to achieve the best (most probable) match (e.g., [8–11]). Of course this approach is susceptible to noise and artifacts, but on the other hand it allows better versatility, and in many cases may be applied successfully to scans performed on different occasions and at different locations. Nevertheless, it is now widely recognized that the merger of information is more efficiently achieved by the hardware approach. The recent (2001) introduction of hybrid scanners has led to an expansion of this approach through the rapid adoption of the technology into the clinical arena.

One of the most promising examples of MMI hybrid systems that is currently demonstrating a significant clinical impact is the combination of CT with nuclear imaging, and specifically positron emission tomography (PET). Following the development of a prototype in the late 1990s [12], the first commercial combined PET/CT scanner was introduced in 2001 and since then, close to 2000 of such devices from different vendors have been installed in clinics worldwide. Both CT and PET technologies continue to advance and since 2006, new PET scanners are now only available in combination with CT. The MMI technology available clinically has demonstrated particular impact in staging malignant

disease [13, 14] and in monitoring response of the disease to therapy. The recent incorporation of high-speed, multislice CT scanners with PET also opens up the potential for applying this technology to cardiac disease.

Another attractive modality for MMI is MRI. Although MRI imposes severe restrictions on the imaging environment, it offers a broad spectrum of scan types and image contrast. Compared with CT, MRI offers greater soft-tissue contrast, better capability for quantitation of function (e.g., measurement of blood flow or tissue metabolism), and potentially new types of molecularly targeted contrast agents. Efforts for combining MRI with other modalities (e.g., PET/MRI and ultrasound/MRI) are currently under development.

Another aspect of MMI is the development of multimodal contrast enhancing materials. Such materials can be used in the form of a "fit-all" type of marker (e.g., [15]). Thus, their signals can be used as control points for 3D alignment. Alternatively, they can be used as standard contrast agents used for disease detection and characterization (e.g., [16]).

In conclusion, considering the current tends in radiology, it can be expected that MMI devices will become increasingly available in the clinical arena. PET/CT has already made an important clinical contribution to patient care for oncology, while the new combined SPECT/CT designs are enhancing SPECT applications and improving physicians' confidence with image interpretation. No doubt, new combinations of hybrid devices will appear in the clinical arena and in many situations. As demonstrated by PET/CT in the oncology field, they will become the primary imaging option. A PET/MR design for simultaneous acquisition of PET and MR has recently acquired the first patient images, and a combined PET and ultrasound device is also under development for breast imaging. For many reasons, therefore, hybrid imaging devices are finding widespread acceptance within the clinical environment and some are already contributing to patient care and management. There is little doubt that this trend will continue in the future with an increasing reliance on MMI devices for medical imaging, thereby ensuring that all involved can be satisfied that they will eventually obtain a true and consistent picture of the elephant.

Haim Azhari
Robert R. Edelman
David Townsend

REFERENCES

[1] H.-U. Kauczor, "Multimodal imaging and computer assisted diagnosis for functional tumour characterisation," *Cancer Imaging*, vol. 5, no. 1, pp. 46–50, 2005.

[2] D. J. Theodorou, K. N. Malizos, A. E. Beris, S. J. Theodorou, and P. N. Soucacos, "Multimodal imaging quantitation of the lesion size in osteonecrosis of the femoral head," *Clinical Orthopaedics and Related Research*, no. 386, pp. 54–63, 2001.

[3] P. Sabbah, H. Foehrenbach, G. Dutertre, et al., "Multimodal anatomic, functional, and metabolic brain imaging for tumor resection," *Clinical Imaging*, vol. 26, no. 1, pp. 6–12, 2002.

[4] R. Lopez, P. Payoux, P. Gantet, J. P. Esquerré, F. Boutault, and J. R. Paoli, "Multimodal image registration for localization of sentinel nodes in head and neck squamous cell carcinoma," *Journal of Oral and Maxillofacial Surgery*, vol. 62, no. 12, pp. 1497–1504, 2004.

[5] R. Stokking, I. G. Zubal, and M. A. Viergever, "Display of fused images: methods, interpretation, and diagnostic improvements," *Seminars in Nuclear Medicine*, vol. 33, no. 3, pp. 219–227, 2003.

[6] U. Pietrzyk, K. Herholz, A. Schuster, H.-M. von Stockhausen, H. Lucht, and W.-D. Heiss, "Clinical applications of registration and fusion of multimodality brain images from PET, SPECT, CT, and MRI," *European Journal of Radiology*, vol. 21, no. 3, pp. 174–182, 1996.

[7] D. Rubello, D. Casara, D. Fiore, P. Muzzio, G. Zonzin, and B. Shapiro, "An ectopic mediastinal parathyroid adenoma accurately located by a single-day imaging protocol of Tc-99m pertechnetate-MIBI subtraction scintigraphy and MIBI-SPECT-computed tomographic image fusion," *Clinical Nuclear Medicine*, vol. 27, no. 3, pp. 186–190, 2002.

[8] P. Hellier and C. Barillot, "A hierarchical parametric algorithm for deformable multimodal image registration," *Computer Methods and Programs in Biomedicine*, vol. 75, no. 2, pp. 107–115, 2004.

[9] J. L. Bernon, V. Boudousq, J. F. Rohmer, et al., "A comparative study of Powell's and Downhill Simplex algorithms for a fast multimodal surface matching in brain imaging," *Computerized Medical Imaging and Graphics*, vol. 25, no. 4, pp. 287–297, 2001.

[10] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A viscous fluid model for multimodal non-rigid image registration using mutual information," *Medical Image Analysis*, vol. 7, no. 4, pp. 565–575, 2003.

[11] A. Sarvazyan, "Model-based imaging," *Ultrasound in Medicine and Biology*, vol. 32, no. 11, pp. 1713–1720, 2006.

[12] T. Beyer, D. W. Townsend, T. Brun, et al., "A combined PET/CT scanner for clinical oncology," *Journal of Nuclear Medicine*, vol. 41, no. 8, pp. 1369–1379, 2000.

[13] R. Bar-Shalom, N. Yefremov, L. Guralnik, et al., "Clinical performance of PET/CT in evaluation of cancer: additional value for diagnostic imaging and patient management," *Journal of Nuclear Medicine*, vol. 44, no. 8, pp. 1200–1209, 2003.

[14] Z. Keidar, N. Haim, L. Guralnik, et al., "PET/CT using 18F-FDG in suspected lung cancer recurrence: diagnostic value and impact on patient management," *Journal of Nuclear Medicine*, vol. 45, no. 10, pp. 1640–1646, 2004.

[15] P. Bolan, M. Nelson, and M. Garwood, "A novel tissue marker for multimodal breast imaging," in *Radiological Society of North America (RSNA) 91st Scientific Assembly and Annual Meeting*, p. 624, Chicago, Ill, USA, November-December 2005, Abstract LPH01-03.

[16] M. A. McDonald and K. L. Watkin, "Small particulate gadolinium oxide and gadolinium oxide albumin microspheres as multimodal contrast and therapeutic agents," *Investigative Radiology*, vol. 38, no. 6, pp. 305–310, 2003.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 61523, 10 pages
doi:10.1155/2007/61523

Research Article

ESTERR-PRO: A Setup Verification Software System Using Electronic Portal Imaging

Pantelis A. Asvestas, Konstantinos K. Delibasis, Nikolaos A. Mouravliansky, and George K. Matsopoulos

Institute of Communication and Computer Systems, National Technical University of Athens, 9 Iroon Polytechniou Street, Zografos, Athens 15780, Greece

Received 26 January 2006; Revised 5 July 2006; Accepted 18 July 2006

Recommended by Haim Azhari

The purpose of the paper is to present and evaluate the performance of a new software-based registration system for patient setup verification, during radiotherapy, using electronic portal images. The estimation of setup errors, using the proposed system, can be accomplished by means of two alternate registration methods. (a) The portal image of the current fraction of the treatment is registered directly with the reference image (digitally reconstructed radiograph (DRR) or simulator image) using a modified manual technique. (b) The portal image of the current fraction of the treatment is registered with the portal image of the first fraction of the treatment (reference portal image) by applying a nearly automated technique based on self-organizing maps, whereas the reference portal has already been registered with a DRR or a simulator image. The proposed system was tested on phantom data and on data from six patients. The root mean square error (RMSE) of the setup estimates was 0.8 ± 0.3 (mean value $\pm$ standard deviation) for the phantom data and 0.3 ± 0.3 for the patient data, respectively, by applying the two methodologies. Furthermore, statistical analysis by means of the Wilcoxon nonparametric signed test showed that the results that were obtained by the two methods did not differ significantly (P value > 0.05).

1. INTRODUCTION

The effectiveness of radiation therapy depends on the patient setup accuracy at each radiation treatment session. A significant problem is to reproduce the intended position of the part of the patient that is irradiated with respect to the treatment beam(s) at each treatment session. It is a common clinical practice to verify the setup by comparing the portal image with a reference one which records the intended patient position. Typical reference images that can be used are simulator images, digitally reconstructed radiographs (DRRs), or another portal images. The introduction of electronic portal imaging devices (EPIDs) offers the potential for correcting inaccuracies in patient placement in a prospective manner, rather retrospectively as is done with conventional megavoltage films.

In general, setup errors are classified as random (or interfraction) and systematic errors [1]. The random errors are deviations between different fractions, during a treatment series, whereas the systematic errors are deviations between the intended patient position and the average patient position over a course of fractionated therapy [2]. Furthermore, random patient movement or periodic movements such as breathing can cause the so-called intrafraction error, which is defined as the deviation observed within a single fraction of fractionated therapy. However, these movements during a single fraction are usually insignificant for most patients and treatment sites, with a few exceptions (e.g., the lung).

A number of setup correction strategies aiming at improving target localization during radiation therapy treatments have been proposed. Most of these strategies are based on the matching of common anatomical features of portal images selected either manually or semiautomatically. In [3], the magnitude of the errors introduced into the registration between the rotated and the nonrotated phantom images and the reference DRR image was determined based on "match structures," which include the field edges and at least three anatomical landmarks manually selected on the reference image and matched with the corresponding anatomy in the portal images. An object-based registration method for portal images was developed in [4], which was based

on core analysis, a fundamental computer vision method, to define correspondence between common anatomical structures of the images manually selected and a curved-based matching algorithm (chamfer-matching) in order to determine the translation/rotation parameters of the image registration. Similar approaches for registering DRRs with portal images were also proposed using either the Pearson's correlation coefficient as a measure of match on selected anatomical features [5] or the template matching technique [6]. Other strategies for the verification of patient setup were based on the optimization of a similarity measure such as histogram matching technique [7], phase-only correlation [8], the minmax entropy [9], and the mutual information [10]. Furthermore, a comparative study of various similarity measures and optimization procedures had been performed on matching high-quality DRRs against portal images that were acquired right before treatment dose delivery [11].

Setup errors that are caused by out-of plane rotations can be estimated by means of three-dimensional techniques [5, 12]. In general, out-of-plane rotational errors that are smaller than 3° do not affect the projected anatomy in portal images significantly. However, when larger rotational errors are not taken into account, this causes a reduced accuracy in the measurement of the translational error [13]. In some cases, two-dimensional techniques can also provide reliable estimations of the out-of-plane rotational errors; out-of-plane rotation for an anterior image can be an in-plane rotation for a lateral image.

In this paper, an extended version of our paper in [14] is presented for the estimation of patient setup errors during radiotherapy treatments. According to the proposed methodology, the verification of patient setup consists of the following steps: (a) delineation of radiation field edges in the portal image in order to verify that the beam has the correct shape as well as to establish a common coordinate system with a previously delineated field edge from the reference image, and (b) matching of common anatomical structures within the two images in order to provide an estimation of the patient setup error relative to the field edges. Two registration methodologies are presented within the paper: (a) the registration of portal image of the current fraction of the treatment with the corresponding DRR image, used as a reference image and (b) the registration of portal images at different treatment sessions using a nearly automatic technique based on self-organizing maps (SOMs) to define automatic correspondence of common anatomical features of the portal images. Both registration methodologies have been incorporated towards the development of a software system, called ESTERR-PRO, for the estimation of patient setup errors as presented in the paper. A detailed description of the system in terms of registration methodologies is provided in Section 2. In Section 3, results of the performance of the system on phantom and real data are presented. Finally, in Sections 4 and 5, discussion of the results and concluding remarks are drawn, respectively.

2. MATERIALS AND METHODS

2.1. The proposed software system: "ESTERR-PRO"

The complete system incorporates the following features: (a) friendly user interface, (b) image processing tools, and (c) a database of patient records and images. The user of the system is able to access and display portal images and corresponding reference images (DRRs simulator images) for a patient selected from the database. The image processing tools can then be used to detect and objectively estimate the setup error, if present.

ESTERR-PRO runs on personal computers (PCs) under the Microsoft Windows operating system. The software was developed in the C++ language.

In order to review portal images, the user must first select a patient name from the database. Then all the available reference and portal images, sorted by date, are presented to the user. When the selection of the reference and the portal image is completed, the images are displayed next to each other (see Figure 1).

The image processing toolkit includes tools for (a) preprocessing such as brightness/contrast adjustment, contrast enhancement (histogram equalization [15], contrast and limited adaptive histogram equalization [16]), and smoothing (mean filtering, median filtering, morphological smoothing [15]), (b) patient setup verification.

2.2. Patient setup verification

As already mentioned, the patient setup verification comprises two steps: (a) delineation of radiation field edges in the portal image in order to verify that the beam has the correct shape as well as to establish a common coordinate system with a previously delineated field edge from the reference image, and (b) matching of common anatomical structures within the two images.

2.2.1. Verification of field shape

The radiation field edges in the portal image are delineated automatically as follows: a thresholding operation (with threshold level set to 5) of the gray-levels of the image is applied in order to obtain a rough approximation of the field contour. Then, the Canny edge detector [17] using a fast recursive implementation of the Gaussian kernel [18], applied on the original image in a band of width 15 pixels around the position of the initial contour, provides the final form of the field edges. The values for the standard deviation of the Gaussian kernel, the low threshold, and the high threshold for the nonmaximum suppression of the Canny edge detector were set to 2.0, 0.0, and 0.95, respectively. The field edges for each image are stored in the computer memory as a binary image, which is called *field edge map* and has the same size as the original image. A value 1 (0) in the field edge map indicates that the corresponding pixel belongs to the field contour (background).

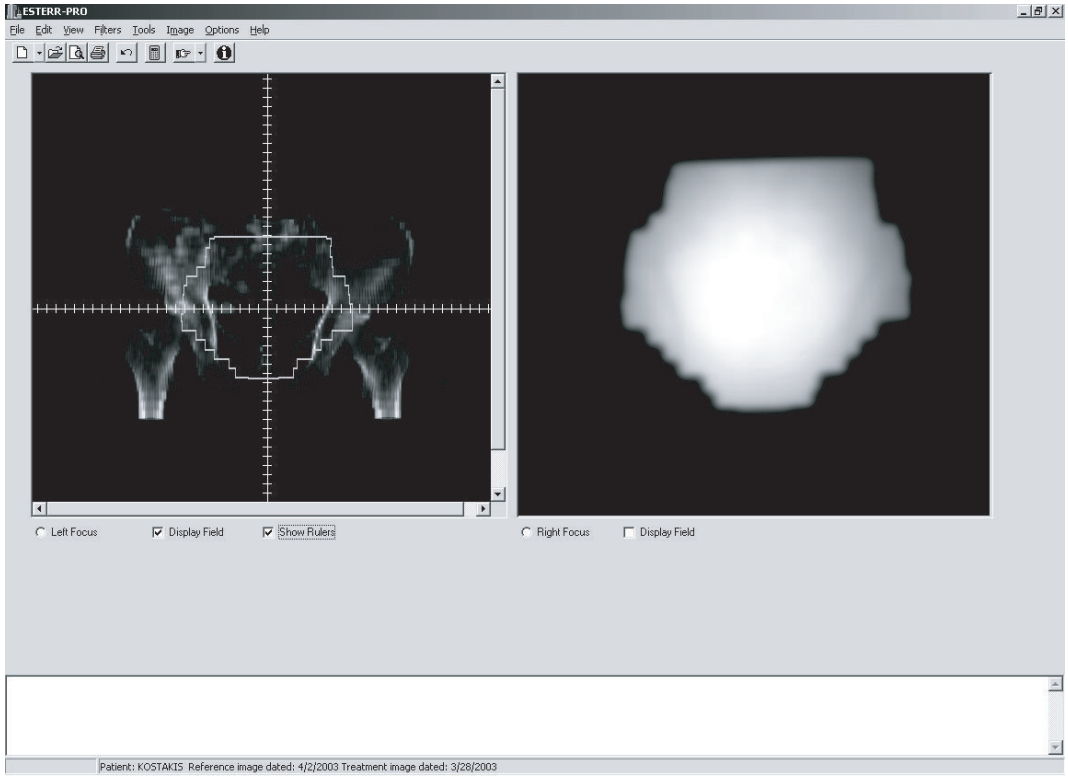


Figure 1: The ESTERR-PRO interface. The main screen is divided into two side-by-side panels. A portal image is displayed on the right panel and a reference image is displayed on the left panel.

The verification of the field shape is accomplished automatically as follows: first, the distance transformation of the reference field edge map is calculated [19]. The distance transformation provides the smallest distance of each pixel of the field edge map from the field edge. Then, an optimization process is applied in order to achieve the spatial coincidence of the two radiation field edge maps. The optimization process involves the minimization, with respect to the parameters of a rigid transform (namely displacement and rotation in the image plane) of the distance between the reference edge map and the transformed version of the edge map of the portal image, using the current values of the parameters of the rigid transform. The distance between the two edge maps is calculated by means of the distance transformation of the reference field edge map. In our implementation, the Powell's method [20] is used for the optimization process. After the end of the optimization process, the distance between the reference edge map and the transformed edge map of the portal image should be small enough. If this distance exceeds a predefined value, this means that the two field edges do not have the same shape and a warning is generated, which informs the user about this inconsistency. Additionally, if the reference image is another portal image then it is expected that the parameters of the rigid transform, obtained during the optimization process, to be nearly zero. If this is not the case, then a warning is also generated. It must be noted that the whole process is invoked automatically, immediately after the user selects the pair of the images.

2.2.2. *Matching of anatomical structures*

The choice of the procedure that is used for the matching of common anatomical structures between the reference and the portal image depends on the type of the reference image: (a) if the reference image is a DRR or a simulator image, a modified manual procedure is applied, (b) if the reference image is a portal image, a semiautomatic approach is followed.

In particular, when the reference image is a DRR or a simulator image, the following procedure is applied. The edges of each image are extracted automatically by the Canny edge detector. Trackbars for the adjustment of the values of the parameters Canny edge detector (namely, the standard deviation of the Gaussian kernel and the high threshold for the nonmaximum suppression) are available (see Figure 2) in order to provide the capability of the user to select only edges that correspond to the anatomical structures of interest.

When the user selects the edges that correspond to the same anatomical structures within the two images, a similar optimization procedure as the one used for the verification of the field shape is invoked. The outcome of the aforementioned procedure is the setup error in terms of horizontal displacement, vertical displacement, and angle of rotation around the axis that is perpendicular on the image plane.

When the reference image is another portal image, the matching procedure involves the definition of fiducial marks (points) on anatomical structures of the reference image that

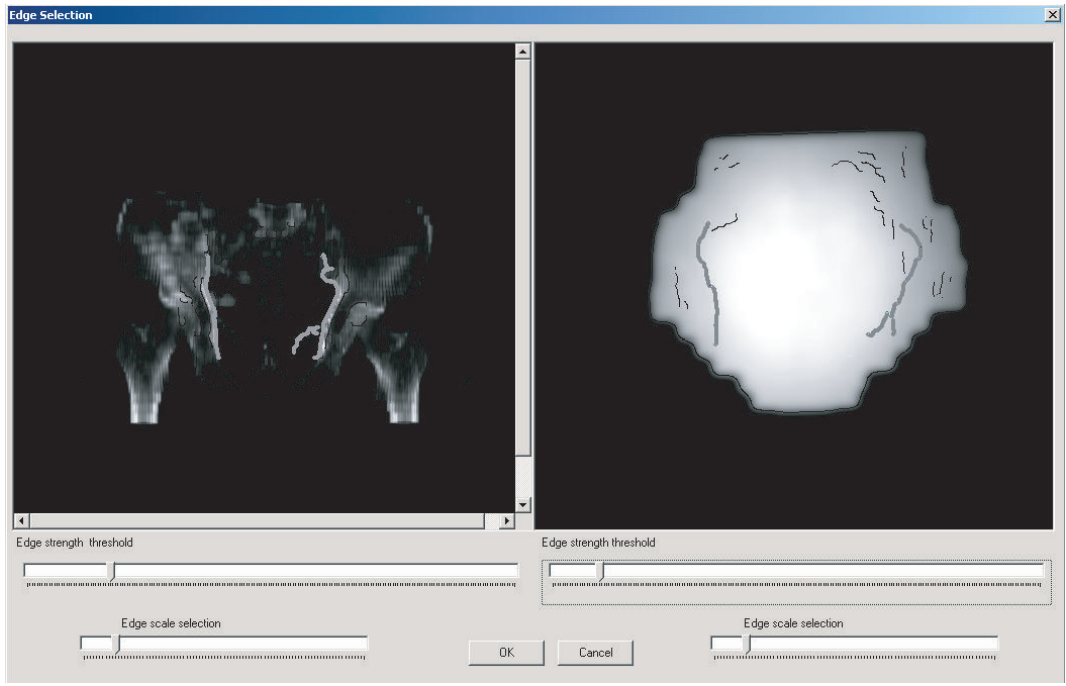


Figure 2: The edge selection window for matching the anatomical structures between a DRR and a portal image.

do not move with time, or with anatomical processes (the skeletal system is a good example, whereas the bladder or the intestines are counterexamples of placement of fiducial marks). The user-defined fiducial marks can be saved with the reference image to be utilized in subsequent fractions of the radiation treatment series. Let (x_i, y_i) $(i = 1, 2, \ldots, N)$ be the coordinates (in pixel units) of the user-defined fiducial marks. The next step of the procedure involves the maximization of a properly chosen function, f, with respect to the parameters of the setup error (horizontal displacement (dx), vertical displacement (dy), and angle of rotation (θ)) using self-organizing maps [21]. The self-organizing map (SOM) is a neural network algorithm, which uses a competitive learning technique to train itself in an unsupervised manner. Kohonen first established the relevant theory and explored possible applications [22]. The Kohonen model comprises a layer of neurons m, ordered usually in a one- or two-dimensional grid. The training of the network is performed in an iterative way. At each iteration k, a data point $\mathbf{x} \in \mathbb{R}^n$ is presented to the network; the neuron j with weight vector $\mathbf{w}_j \in \mathbb{R}^n$ is declared as the winning neuron, according to the following rule:

$$j = \arg\min_i \left(||\mathbf{x} - \mathbf{w}_i||\right). \tag{1}$$

The winning neuron j and its neighboring neurons i have their weight vectors modified according to the following rule:

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + h_{ij}(n)\left[\mathbf{x}(n) - \mathbf{w}_i(n)\right], \tag{2}$$

where $h_{ij}(n) = h(\|\mathbf{r}_i - \mathbf{r}_j\|, n)$ is a kernel defined on the neural network space as a function of the distance $\|\mathbf{r}_i - \mathbf{r}_j\|$ between the winning neuron j and its neighboring neurons i, as well as the iteration number n. This kernel has the shape of the "Mexican hat" function, which in its discrete form has maximum value at inter-neuron distance in the case of $i = j$ whereas its value drops in a Gaussian manner as the distance increases. The width of this function decreases monotonically with iteration number. In this way convergence to the global optimum is attempted during the early phases of the self-training process, whereas gradually the convergence becomes more local as the size of the kernel decreases.

Prior the description of the proposed method, some notations must be introduced. Let $\mu_A(I)$ denote the restriction of an image I to the region $A \subset \mathbb{R}^2$ and $T_{\mathbf{w}}(A) \subset \mathbb{R}^2$ is the rigid transformation, with parameters $\mathbf{w} = (dx, dy, \theta)$, of the region A, where dx, dy, and θ are the horizontal displacement, the vertical displacement, and the angle of rotation, respectively. Furthermore, $\mathrm{MoM}(I_1, I_2)$ denotes a measure of match between two images I_1 and I_2.

If I_R, I_F are the reference image and the image to be registered, respectively, then the implementation of the SOM network for registering the two images is as follows. The topology of the network is constructed by placing a neuron on each user-defined fiducial mark $\mathbf{P}_i = (x_i, y_i)$ $(i = 1, 2, \ldots, N)$ of the reference image. Each neuron is associated with a square area $A_i = [x_i - R, x_i + R] \times [y_i - R, y_i + R]$, of $(2R+1)^2$ pixels, centered at the position of the neuron. Additionally, a weight vector $\mathbf{w}_i = (dx_i, dy_i, \theta_i)$, which holds the parameters of a local rigid transformation, is assigned to each neuron.

The SOM network is trained as follows.

(1) For each neuron, the components of the weight vector are initialized to zero values, $\mathbf{w}_i(0) = (0, 0, 0)$, the quantities $\mathrm{MoM}_i(0) \equiv \mathrm{MoM}(\mu_{A_i}(I_R), \mu_{T_{\mathbf{w}_{i(0)}}(A_i)}(I_F))$ are

calculated, the variable $\mathrm{MoM}_{\mathrm{best}}$ is set to a very large (in magnitude) negative value, and the iteration variable, n, is set to 1.

(2) While n is less than $n_{\max}$,

(i) if the average value of the $\mathrm{MoM}_i(n-1)$, $\mathrm{MoM}_{\mathrm{ave}}(n-1)$, is better than $\mathrm{MoM}_{\mathrm{best}}$, then $\mathrm{MoM}_{\mathrm{best}} = \mathrm{MoM}_{\mathrm{ave}}(n-1)$ and the current weights are stored as $\mathbf{w}_i$;

(ii) an *input signal*, $\mathbf{s}(n) = (dx(n), dy(n), \theta(n))$, is generated randomly;

(iii) for every neuron, the quantity $\mathrm{MoM}_i(n) \equiv \mathrm{MoM}(\mu_{A_i}(I_R), \mu_{T_{\mathbf{s}(n)}}(A_i)(I_F))$ is calculated;

(iv) the *winning neuron*, k_n, in the current iteration, is defined as

$$k_n = \arg\max_i \{\mathrm{MoM}_i(n)\} \tag{3}$$

under the condition

$$\mathrm{MoM}_{k_n}(n) > \mathrm{MoM}_{\mathrm{ave}}(n-1); \tag{4}$$

(v) the weights of the neurons are updated according to the following equation:

$$\mathbf{w}_i(n) = \mathbf{w}_i(n-1) + h(k_n, n, i)\left[\mathbf{s}(n) - \mathbf{w}_i(n-1)\right], \tag{5}$$

where $h(k_n, n, i)$ $(i = 1, 2, \ldots, N)$ is given by the following equation:

$$h(k_n, n, i) = \begin{cases} L^{q(n)}, & \|\mathbf{P}_{k_n} - \mathbf{P}_i\| < \alpha^{q(n)} d_0, \\ 0 & \text{otherwise} \end{cases}$$
$$q(n) = \left\lfloor \frac{n}{p+1} \right\rfloor, \tag{6}$$

$L, a, d_0 \in \mathbb{R}$ and $p \in \mathbb{R}$ are parameters to be defined later, $\|\|$ denotes the Euclidean norm, and $\lfloor\ \rfloor$ is the floor function;

(vi) the iteration variable is increased by one.

(3) When the training is finished, the parameters of the affine transformation between the two portal images are calculated using a least squares method between the point sets $\{\mathbf{P}_i\}$ and $\{T_{\mathbf{w}_i}(\mathbf{P}_i)\}$ [20].

The selected measure of match was the gradient correlation coefficient, namely,

$$\begin{aligned} \mathrm{MoM}(I_R, I_F) &= GCC(I_R, I_F) \\ &= GCC_h(I_R, I_F) + GCC_v(I_R, I_F), \end{aligned} \tag{7}$$

where

$$GCC_h(I_R, I_F) = \frac{\left\{\sum_{x,y}\left[G_R^h(x,y) - \overline{G}_R^h\right]\left[G_F^h(x,y) - \overline{G}_F^h\right]\right\}^2}{\sum_{x,y}\left[G_R^h(x,y) - \overline{G}_R^h\right]^2 \sum_{x,y}\left[G_F^h(x,y) - \overline{G}_F^h\right]^2},$$

$$GCC_v(I_R, I_F) = \frac{\left\{\sum_{x,y}\left[G_R^v(x,y) - \overline{G}_R^v\right]\left[G_F^v(x,y) - \overline{G}_F^v\right]\right\}^2}{\sum_{x,y}\left[G_R^v(x,y) - \overline{G}_R^v\right]^2 \sum_{x,y}\left[G_F^v(x,y) - \overline{G}_F^v\right]^2}. \tag{8}$$

The subscript $R(F)$ refers to the reference (to be registered) image; the superscript $h(v)$ refers to the horizontal (vertical) direction in the image plane, and G denotes the first derivative. For example, $G_R^h(x, y)$ denotes the first derivative of the reference image along the horizontal direction estimated at pixel position (x, y). $\overline{G}$ refers to the mean value of the first derivative.

The rationale for selecting the aforementioned measure of match was that gradient measures concentrate the contributions on edge information, which intuitively appears sensible.

The following generator of random numbers is used for producing the input signals to the network:

$$\begin{aligned} s_j(n) &= w_{k_n,j} + \mathrm{sgn}\,(v_j - 0.5)\,\mathrm{TM}(n) \\ &\quad \times \left[\left(1 + \frac{1}{\mathrm{TM}(n)}\right)^{|2v_j - 1|} - 1\right](U_j - L_j) \quad (j = 1, 2, 3) \\ \mathrm{TM}(n) &= \begin{cases} 1, & n = 0, \\ \exp\left(-2(q(n))^{1/n}\right), & n > 0, \end{cases} \end{aligned} \tag{9}$$

where $s_1(n) = dx(n)$, $s_2(n) = dy(n)$, $s_3(n) = \theta(n)$, v_j is a uniformly distributed random variable in $[0, 1]$, and $U_j(L_j)$ denotes the maximum (minimum) allowed value for the corresponding component of the input signal. Although U_j and L_j are inputs to the matching process, for all pairs of images used in the current study, constant values were used (± 50 pixels for the displacement and $\pm 10^\circ$ for the angle of rotation).

It must be noted that (9) is a slightly modified version of the generator used in the very fast simulated annealing method [23] and provides random signals which in general lie in the range $[w_{k_n,j} - (U_j - L_j), w_{k_n,j} + (U_j - L_j)]$. When a generated signal is not in the allowed range $[L_j, U_j]$, then it is discarded and a new signal is produced until $s_j(n) \in [L_j, U_j]$. The parameter $\mathrm{TM}(n)$ controls how far from the weights of the current winning neuron the input signal can reach. As the iteration variable evolves, the magnitude of $\mathrm{TM}(n)$ falls exponentially and the generated input signals are more localized around the weights of the current winning neuron (see Figure 3). This is a desired property, since as the number of iterations grows, the weights of the current winning neuron get closer to the parameters of the solution of the matching problem.

The parameter d_0 provides the initial radius of a circular region around the winning neuron. Only neurons inside this region are updated. Usually, d_0 is set to the maximum distance between the fiducial marks. As can be seen from (6), this distance is reduced with geometric rate determined by the parameter α $(0 < \alpha \leq 1)$. A typical value for the parameter α is 0.995. The parameter L acts like a gain constant for the magnitude of the update that is applied to the weights of the neurons. This parameter also decreases geometrically as the iteration variable evolves. The range of values L is between 0.99 and 1.0; a typical value is 0.995. The parameter

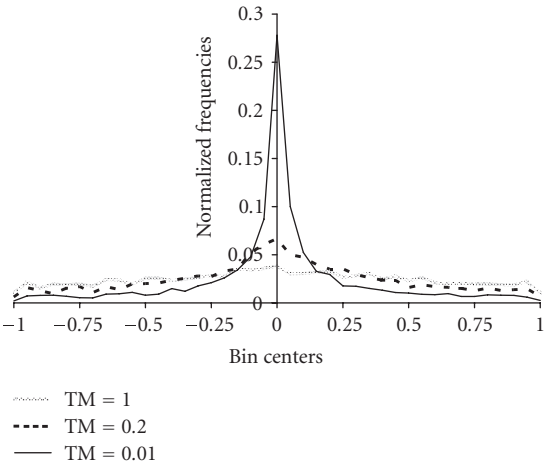


Figure 3: Normalized histogram of the values obtained by means of the random number generator described in (9) for different values of the parameter TM.

p is an integer that determines the rate of change of the parameters L and α. Practically, this parameter determines the number of iterations that are executed before an adjustment of the values for the parameters L, α, and $\mathrm{TM}(n)$ takes place. A typical value for this parameter is 200. The number of iterations is set to 5000 and the size of the square area associated with each neuron is 19 ($R = 9$).

Finally, since the transformed region $T_{\mathbf{s}}(n)(A_i)$ does not have integer coordinates, bilinear interpolation is used in order to calculate $\mathrm{MoM}_i(n)$.

In Figure 4, the results that were obtained from the application of the proposed methodology on real pelvic data are presented.

2.3. Data acquisition and evaluation protocol

The proposed system was tested on phantom and real data. All the portal images were acquired for gantry angle 0° using a CCD camera-based EPID (Beamview Plus v2.1, Siemens) of "HYGEIA" with a total dose of ∼ 20 MU/image; 6-MV X-rays were used. The size of the pelvic field was $12 \times 12\,\mathrm{cm}^2$ at the isocenter. The DRR images were obtained from a CT data set acquired using a Siemens Somaton Plus 4 scanner at 120 keV, with a 2 mm slice thickness and no gap between slices.

An anthropomorphic phantom (Alderson Rando phantom), commercially available system, was used. The evaluation method was as follows.

(a) A DRR and a portal image of the selected region of interest (head, lung, pelvis) of the phantom, with no setup error between them, were acquired. This was achieved by means of three fiducial markers, visible on the CT imaging, adhered on the phantom using laser alignment to define a reference point on both CT scanning and irradiation. This point was used as the isocenter during treatment planning (DRR) and irradiation (portal imaging). These DRR and portal images served as the reference images. Before the acquisition of CT phantom images and phantom irradiation, the accuracy of the lasers alignment in both the CT and the treatment room was checked and found within 1 mm. Moreover, possible introduced inaccuracies due to no-horizontal CT-couch motion or inaccuracies in the stated slice thickness must be excluded since an extensive quality control was performed prior to the use of the CT scanner.

(b) An expert from the "HYGEIA" hospital defined the fiducial marks on anatomical structures, (x_i, y_i) $(i = 1, 2, \ldots, N)$, on the portal reference image. These points were used for matching each portal image with the reference portal image.

(c) The treatment couch was moved along the horizontal (left-right) and/or vertical (head-foot) direction 2 mm, 4 mm, ..., 12 mm and a new portal image was acquired. Before phantom irradiation, the treatment couch (ZXT-Siemens) reading positions were checked and their accuracy was found within 1 mm.

(d) The setup verification tool of the system was invoked in order to estimate the setup error between the reference images and each new portal image, using the procedures described in Sections 2.2.1 and 2.2.2. The output of the system was the estimated values of the parameters of the simulated setup error, namely horizontal displacement (mm), vertical displacement (mm), and angle of rotation (degrees).

(e) A set of 50 test points, $\mathbf{P}_i$ $(i = 1, 2, \ldots, 50)$, was defined on the reference DRR. Since the setup error is known, the actual position of these points on the portal images (including the reference portal image) was identified. The position of these points on the portal images was also identified using the estimated values for the setup error. For each portal image, the root mean square error (RMSE) (in millimeters) between the actual and the estimated positions of this set points was calculated by the following equation:

$$\mathrm{RMSE} = \sqrt{\frac{1}{50}\sum_{i=1}^{50} \left|\left|(\mathbf{P}_i)_{\mathrm{act}} - (\mathbf{P}_i)_{\mathrm{est}}\right|\right|^2}, \tag{10}$$

where the subscript act (est) refers to the actual (estimated) positions of the test points on the portal images and $\|\cdot\|$ denotes the Euclidean distance.

For the real data, the aforementioned evaluation procedure was slightly modified, since it is not known the actual setup error. Therefore, a manual registration, to serve as ground truth, was carried out by two experts from "HYGEIA" hospital. The average value of the two obtained estimates was used as ground truth. A total of six subjects were investigated, who were recruited from patients referred to the "HYGEIA" hospital for prostate cancer treatment. For each patient, a DRR and thirty portal images were acquired correspondingly. The proposed evaluation protocol has been approved by the ethical committee of the "HYGEIA" hospital and the subjects gave informed consent to the work.

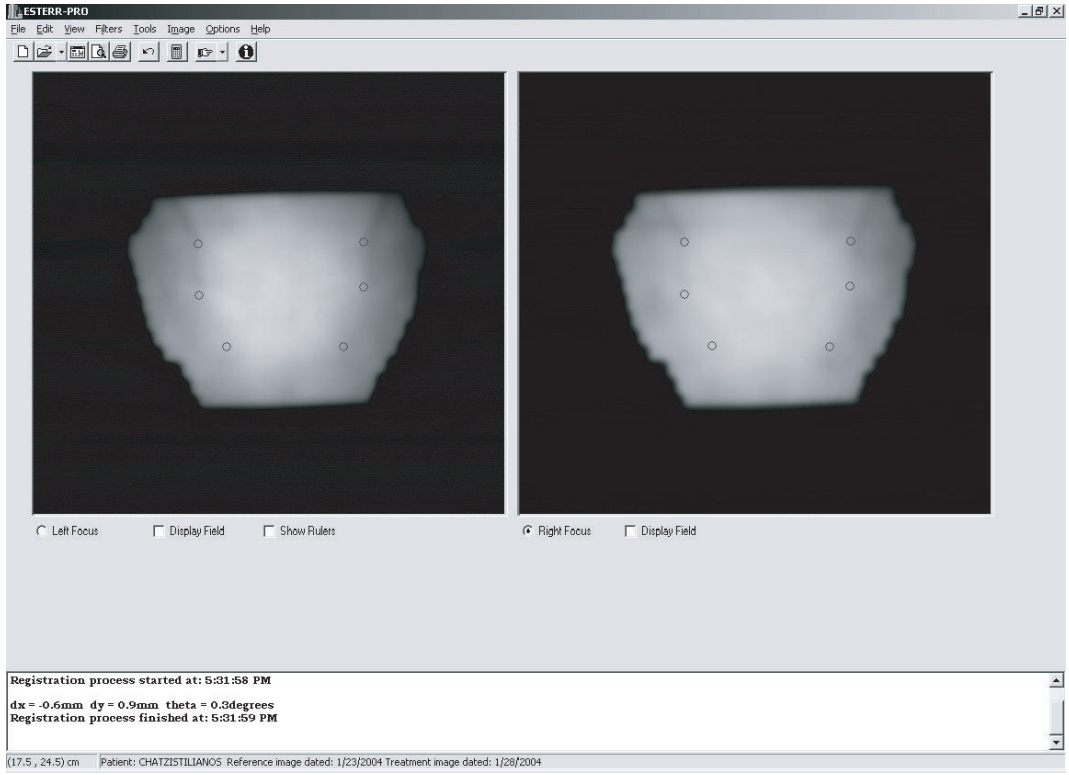


Figure 4: Example of matching a pair of portal images on patient data. The fiducial marks, defined by the user, are shown on the reference image (left panel). The corresponding points on the other portal image, after the end of matching procedure, are also shown.

2.4. Statistics

The statistical difference between the actual and estimated positions was assessed by means of the Wilcoxon signed nonparametric test, for both real and phantom data [24]. The null hypothesis was that the two methods (DRR—portal and portal—portal matching) did not differ as per the RMSE.

3. RESULTS

3.1. Phantom data

Setup error estimations (three parameters and RMSE) are shown in Tables 1 and 2, respectively. For the pelvic region, the RMSE was 0.8 ± 0.3 (mean value $\pm$ standard deviation) and 0.8 ± 0.4 when the reference image was the DRR and the portal image, respectively. For the cranial region, the RMSE was 0.8 ± 0.3 and 0.6 ± 0.3 when the reference image was the DRR and the portal, respectively. For both cases, the Wilxocon signed test showed that the null hypothesis could not be rejected at the 5% level (P value > 0.05).

3.2. Real data

For each patient, the average values of the setup error over the thirty portal images are shown in Table 3. The RMSE over all six patients was 0.3 ± 0.3 and 0.3 ± 0.3 when the reference image was the DRR and the portal corresponding to the first fraction of the treatment, respectively. The statistical analysis of the RMSE measurements of each patient showed that the null hypothesis could not be rejected at the 5% level (P value > 0.05).

4. DISCUSSION

The design and the development of the proposed system were based on several constraints. It should use the equipment available in the "HYGEIA" hospital: a portal imager and a CT scanner. The human intervention should be minimal, the results should be accurate and the execution time should be kept as low as possible.

In this framework, the estimation of the patient setup can be accomplished using two alternate processes: (a) the portal image corresponding to the current fraction of the treatment is matched directly with the DRR (or the simulator image). (b) The portal image is matched with the portal image acquired during the first fraction of the treatment (reference portal), whereas the reference portal has already been matched with the DRR image.

The rationale underlying the proposed system design was based on the following facts. In general, it is a very difficult task to achieve an accurate matching between a DRR and a portal image automatically, mainly due to the fact that

Table 1: Setup error estimations (horizontal displacement, vertical displacement, angle of rotation, and RMS error) for known setup error of the pelvic region of the phantom for DRR versus portal and portal versus portal.

Expected			DRR versus portal				Portal versus portal			
dx (mm)	dy (mm)	$\theta(^\circ)$	dx (mm)	dy (mm)	$\theta(^\circ)$	RMSE (mm)	dx (mm)	dy (mm)	$\theta(^\circ)$	RMSE (mm)
2	0	0	3.0	0.1	−0.4	1.2	3.1	0.2	0.1	1.2
4	0	0	4.4	−0.1	0.2	0.6	4.8	−0.2	−0.1	0.8
6	0	0	5.5	−0.2	0.4	0.9	5.6	−0.2	−0.2	0.7
8	0	0	8.6	−0.8	0.3	1.2	9.3	−0.6	0.3	1.6
10	0	0	10.9	−0.7	0.3	1.3	10.1	−0.4	0.4	0.9
12	0	0	12.2	−0.3	0.0	0.4	12.2	−0.8	0.2	0.9
0	2	0	0.0	1.7	−0.2	0.4	0.1	1.4	0.1	0.6
0	4	0	−0.3	5.2	0.2	1.2	−0.1	4.9	0.1	0.9
0	6	0	−0.2	6.2	0.2	0.5	−0.1	5.6	−0.1	0.5
0	8	0	0.0	8.2	0.4	0.8	0.1	8.2	−0.1	0.2
0	10	0	0.5	9.5	0.0	0.8	0.1	9.6	0.3	0.7
0	12	0	−0.1	12.5	-0.1	0.6	−0.3	11.9	0.1	0.3

Table 2: Setup error estimations (horizontal displacement, vertical displacement, angle of rotation, and RMS error) for known setup error of the cranial region of the phantom for DRR versus portal and portal versus portal.

Expected			DRR versus portal				Portal versus portal			
dx (mm)	dy (mm)	$\theta(^\circ)$	dx (mm)	dy (mm)	$\theta(^\circ)$	RMSE (mm)	dx (mm)	dy (mm)	$\theta(^\circ)$	RMSE (mm)
2	0	0	1.7	0.2	−0.3	0.7	1.8	−0.0	−0.1	0.3
4	0	0	3.9	0.2	0.2	0.4	4.7	−0.1	−0.1	0.7
6	0	0	6.4	−0.1	0.0	0.4	6.0	−0.0	−0.2	0.4
8	0	0	8.5	0.5	0.2	0.8	8.1	−0.2	0.1	0.3
10	0	0	10.5	0.3	−0.2	0.7	10.5	−0.2	−0.2	0.7
12	0	0	12.6	−0.2	−0.1	0.6	11.9	0.0	−0.1	0.2
0	2	0	−0.3	2.4	0.4	0.9	0.1	2.4	0.1	0.5
0	4	0	−0.1	4.6	0.3	0.8	0.2	4.5	0.2	0.7
0	6	0	0.1	6.8	0.1	0.8	0.6	6.8	0.2	1.0
0	8	0	0.6	8.5	0.25	0.9	0.2	8.4	−0.1	0.5
0	10	0	0.0	11.3	−0.1	1.4	0.2	10.9	0.1	0.9
0	12	0	0.8	12.6	0.4	1.3	0.5	12.8	0.2	0.9

the two images are acquired at totally different energies. It has been proposed to convert the DRR into a megavoltage DRR prior to the matching [5, 25, 26]. However, this was not possible in our case, due to software system installation into a dedicated computer platform, output incompatibility of the radiation treatment planning software and local network topology.

On the other hand, automated techniques based on segmentation should be excluded from the design since these methods rely heavily on the success of the segmentation step, which is a very difficult task due to the low inherent contrast of the portal images. Additionally, no segmentation technique can give satisfactory results for every anatomical region of interest. The very difficult task of portal image segmentation justified the development of other methods of research, such as intensity-based methods [5, 9, 27]. These methods assume there is a statistical relation between the gray level values of the pixels of the images to be matched and that this relation is at maximum when the images are matched. Although these methods seem to be promising, further work is required.

Another solution was the use of some kind of manual technique. However, pure manual methods depend on the accurate determination of homologous fiducial marks between the two images and furthermore are in general time consuming and prone to spatial inaccuracies. Therefore,

Table 3: Setup error estimations (horizontal displacement, vertical displacement, angle of rotation, and RMS error) for six subjects. The values listed are the mean values ± standard deviation calculated over a set of thirty portal images. The RMSE measurements for portal-portal versus DRR-portal did not show significant differences ($P > 0.05$, Wilcoxon signed test).

Patient	Setup error	dx(mm)	dy(mm)	$\theta(^\circ)$	RMSE(mm)
1	Expected	-1.8 ± 0.2	-1.6 ± 0.1	$-0.1 \pm .0.2$	
	DRR versus portal	-1.8 ± 0.12	-1.5 ± 0.1	$-0.2 \pm .0.2$	0.1 ± 0.1
	Portal versus portal	-1.6 ± 0.2	-1.8 ± 0.1	$0.2 \pm .0.2$	0.3 ± 0.1
2	Expected	2.0 ± 0.3	-1.0 ± 0.5	-0.2 ± 0.1	
	DRR versus portal	$2.0 \pm .0.4$	$-1.2 \pm .0.6$	$-0.2 \pm .0.2$	0.1 ± 0.1
	Portal versus portal	$1.9 \pm .0.2$	$-1.1 \pm .0.3$	$-0.2 \pm .0.1$	0.1 ± 0.1
3	Expected	-1.9 ± 0.5	-0.1 ± 0.0	$-0.1 \pm .0.2$	
	DRR versus portal	$-2.0 \pm .0.4$	$-0.1 \pm .0.1$	$-0.1 \pm .0.1$	0.8 ± 0.5
	Portal versus portal	$-2.0 \pm .0.4$	$-0.1 \pm .0.1$	$-0.1 \pm .0.2$	0.7 ± 0.5
4	Expected	$-0.1 \pm .0.1$	-3.5 ± 0.7	$-0.2 \pm .0.3$	
	DRR versus portal	$-0.1 \pm .0.1$	$-3.5 \pm .0.6$	$-0.2 \pm .0.2$	0.1 ± 0.0
	Portal versus portal	$-0.1 \pm .0.1$	$-3.8 \pm .0.6$	$-0.2 \pm .0.2$	0.2 ± 0.0
5	Expected	$0.5 \pm .0.1$	1.6 ± 0.6	$0.4 \pm .0.3$	
	DRR versus portal	$0.8 \pm .0.1$	$1.9 \pm .0.7$	$0.4 \pm .0.3$	0.6 ± 0.1
	Portal versus portal	$0.7 \pm .0.1$	$1.7 \pm .0.7$	$0.4 \pm .0.2$	0.3 ± 0.1
6	Expected	$2.0 \pm .0.3$	-1.5 ± 0.6	$0.2 \pm .0.2$	
	DRR versus portal	$2.4 \pm .0.4$	$-1.1 \pm .0.5$	$0.2 \pm .0.2$	0.3 ± 0.1
	Portal versus portal	$2.1 \pm .0.2$	$-1.4 \pm .0.4$	$0.1 \pm .0.1$	0.3 ± 0.0

regarding the match of DRR and a portal image, we have chosen a modified manual technique that automatically identifies candidate pairs of corresponding edges between the two images. Then, the user simply selects the proper pairs of edges that are going to be used for the matching. The results in Tables 1–3 indicate that the proposed methodology provides estimates of the setup error that are close enough to the expected ones. As can be observed, the values of patient displacements along the horizontal and vertical axis are smaller than those of the phantom, due to quality control processes adopted at the Radiotherapy Department of the "HYGEIA" Hospital.

As already mentioned, the estimation of the setup error can be also accomplished by means of a portal-to-portal matching method. This method requires the definition of a small number (four to seven) of fiducial marks only on the reference portal image. These fiducial marks are stored in the database and are automatically retrieved every time the specific patient is selected. The accuracy of the portal-to-portal matching lies within the limits imposed by the clinical routine. This approach introduces an additional error in the estimation of patient setup error (error due to the matching of DRR with the portal image of first fraction and error due to the matching of the portal of the current fraction with the portal image of the first fraction). However, statistical analysis showed that the two methods did not differ significantly as per the RMSE. Additionally, since a nearly automated method is used, the user is provided in return with a fast, reliable, robust, and user-friendly technique, which requires minimal user intervention.

5. CONCLUSIONS

An integrated software system has been presented for the calculation of patient setup errors in radiotherapy, using EPID images. The system handles both DRR-portal image pair as well as portal-portal pairs. The philosophy of the system was to achieve very fast execution time, increased robustness, considering parameter range as well as anatomic regions of the body and accuracy in the calculation of setup errors. The system has already been installed in the Radiotherapy Department of "HYGEIA" Hospital, Athens, and is fully operational in a clinical environment. The selected methods for image registration require minimal user intervention, achieve high accuracy, and have proven highly practical and popular among the physicians and physicists of the "HYGEIA" Hospital.

REFERENCES

[1] M. G. Herman, "Clinical use of electronic portal imaging," *Seminars in Radiation Oncology*, vol. 15, no. 3, pp. 157–167, 2005.

[2] C. W. Hurkmans, P. Remeijer, J. V. Lebesque, and B. J. Mijnheer, "Set-up verification using portal imaging; review of current clinical practice," *Radiotherapy and Oncology*, vol. 58, no. 2, pp. 105–120, 2001.

[3] J. Hanley, G. S. Mageras, J. Sun, and G. J. Kutcher, "The effects of out-of-plane rotations on two dimensional portal image registration in conformal radiotherapy of the prostate," *International Journal of Radiation Oncology, Biology, and Physics*, vol. 33, no. 5, pp. 1331–1343, 1995.

[4] D. S. Fritsch, E. L. Chaney, A. Boxwala, et al., "Core-based portal image registration for automatic radiotherapy treatment verification," *International Journal of Radiation Oncology, Biology, and Physics*, vol. 33, no. 5, pp. 1287–1300, 1995.

[5] L. M. Sirois, D. H. Hristov, and B. G. Fallone, "Three-dimensional anatomy setup verification by correlation of orthogonal portal images and digitally reconstructed radiographs," *Medical Physics*, vol. 26, no. 11, pp. 2422–2428, 1999.

[6] H. Lemnitzer, U. Wolf, G. Hildebrandt, and F. Kamprad, "Verification of electron field positioning," *Radiotherapy and Oncology*, vol. 52, no. 1, pp. 61–64, 1999.

[7] L. Dong and A. L. Boyer, "An image correlation procedure for digitally reconstructed radiographs and electronic portal images," *International Journal of Radiation Oncology, Biology, and Physics*, vol. 33, no. 5, pp. 1053–1060, 1995.

[8] J.-Z. Wang, L. E. Reinstein, J. Hanley, and A. G. Meek, "Investigation of a phase-only correlation technique for anatomical alignment of portal images in radiation therapy," *Physics in Medicine and Biology*, vol. 41, no. 6, pp. 1045–1058, 1996.

[9] R. Bansal, L. H. Staib, Z. Chen, et al., "A novel approach for the registration of 2D portal and 3D CT images for treatment setup verification in radiotherapy," in *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '98)*, pp. 1075–1086, Cambridge, Mass, USA, October 1998.

[10] S. Clippe, D. Sarrut, C. Malet, S. Miguet, C. Ginestet, and C. Carrie, "Patient setup error measurement using 3D intensity-based image registration techniques," *International Journal of Radiation Oncology, Biology, and Physics*, vol. 56, no. 1, pp. 259–265, 2003.

[11] A. Khamene, P. Bloch, W. Wein, M. Svatos, and F. Sauer, "Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy," *Medical Image Analysis*, vol. 10, no. 1, pp. 96–112, 2006.

[12] A. E. Lujan, J. M. Balter, and R. K. Ten Haken, "Determination of rotations in three dimensions using two-dimensional portal image registration," *Medical Physics*, vol. 25, no. 5, pp. 703–708, 1998.

[13] P. Remeijer, E. Geerlof, L. Ploeger, K. Gilhuijs, M. Van Herk, and J. V. Lebesque, "3-D portal image analysis in clinical practice: an evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients," *International Journal of Radiation Oncology, Biology, and Physics*, vol. 46, no. 5, pp. 1281–1290, 2000.

[14] G. K. Matsopoulos, P. A. Asvestas, K. K. Delibasis, et al., "Registration of electronic portal images for patient set-up verification," *Physics in Medicine and Biology*, vol. 49, no. 14, pp. 3279–3289, 2004.

[15] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley, Reading, Mass, USA, 2nd edition, 2002.

[16] S. M. Pizer, E. P. Amburn, J. D. Austin, et al., "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.

[17] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[18] L. J. van Vliet, I. T. Young, and P. W. Verbeek, "Recursive Gaussian derivative filters," in *Proceedings of the 14th International Conference on Pattern Recognition (ICPR '98)*, vol. 1, pp. 509–514, Brisbane, Australia, August 1998.

[19] G. Borgefors, "Distance transformations in arbitrary dimensions," *Computer Vision, Graphics, & Image Processing*, vol. 27, no. 3, pp. 321–345, 1984.

[20] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Minimization or maximization of functions," in *Numerical Recipes in C: The Art of Scientific Computing*, pp. 412–420, Cambridge University Press, Cambridge, UK, 2nd edition, 1993.

[21] G. K. Matsopoulos, P. A. Asvestas, N. A. Mouravliansky, and K. K. Delibasis, "Multimodal registration of retinal images using self organizing maps," *IEEE Transactions on Medical Imaging*, vol. 23, no. 12, pp. 1557–1563, 2004.

[22] T. Kohonen, *Self-Organizing Maps*, Springer, Berlin, Germany, 2000.

[23] L. Ingber and B. Rosen, "Genetic algorithms and very fast simulated reannealing: a comparison," *Mathematical Computer Modelling*, vol. 16, no. 11, pp. 87–100, 1992.

[24] S. Siegel and N. J. Castellan, *Nonparametric Statistics for the Behavioral Science*, McGraw-Hill, New York, NY, USA, 2nd edition, 1988.

[25] L. Dong and A. L. Boyer, "An image correlation procedure for digitally reconstructed radiographs and electronic portal images," *International Journal of Radiation Oncology, Biology, and Physics*, vol. 33, no. 5, pp. 1053–1060, 1995.

[26] D. S. Fritsch, S. Raghavan, A. Boxwala, et al., "Benchmark test cases for evaluation of computer-based methods for detection of setup errors: realistic digitally reconstructed electronic portal images with known setup errors," *International Journal of Radiation Oncology, Biology, and Physics*, vol. 37, no. 1, pp. 199–204, 1997.

[27] S. Clippe, D. Sarrut, C. Malet, S. Miguet, C. Ginestet, and C. Carrie, "Patient setup error measurement using 3D intensity-based image registration techniques," *International Journal of Radiation Oncology, Biology, and Physics*, vol. 56, no. 1, pp. 259–265, 2003.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 62030, 14 pages
doi:10.1155/2007/62030

Research Article

Towards a Noninvasive Intracranial Tumor Irradiation Using 3D Optical Imaging and Multimodal Data Registration

R. Posada,[1,2] Ch. Daul,[1] D. Wolf,[1] and P. Aletti[1]

[1] *Centre de Recherche en Automatique de Nancy (CRAN UMR 7039), Nancy-Université, CNRS, CAV, 2 avenue de la Forêt de Haye, 54516 Vandœuvre-Lès-Nancy, France*
[2] *Instituto Tecnologico de Orizaba, avenida oriente 9 no. 852, Colonia Emiliano Zapata, 94320 Veracruz, Orizaba, Mexico*

Received 28 April 2006; Revised 17 November 2006; Accepted 6 February 2007

Recommended by David Townsend

Conformal radiotherapy (CRT) results in high-precision tumor volume irradiation. In fractioned radiotherapy (FRT), lesions are irradiated in several sessions so that healthy neighbouring tissues are better preserved than when treatment is carried out in one fraction. In the case of intracranial tumors, classical methods of patient positioning in the irradiation machine coordinate system are invasive and only allow for CRT in one irradiation session. This contribution presents a noninvasive positioning method representing a first step towards the combination of CRT and FRT. The 3D data used for the positioning is point clouds spread over the patient's head (CT-data usually acquired during treatment) and points distributed over the patient's face which are acquired with a structured light sensor fixed in the therapy room. The geometrical transformation linking the coordinate systems of the diagnosis device (CT-modality) and the 3D sensor of the therapy room (visible light modality) is obtained by registering the surfaces represented by the two 3D point sets. The geometrical relationship between the coordinate systems of the 3D sensor and the irradiation machine is given by a calibration of the sensor position in the therapy room. The global transformation, computed with the two previous transformations, is sufficient to predict the tumor position in the irradiation machine coordinate system with only the corresponding position in the CT-coordinate system. Results obtained for a phantom show that the mean positioning error of tumors on the treatment machine isocentre is 0.4 mm. Tests performed with human data proved that the registration algorithm is accurate (0.1 mm mean distance between homologous points) and robust even for facial expression changes.

1. INTRODUCTION

1.1. Medical context

The goal of radiotherapy is to eradicate tumors while preserving the surrounding healthy organs as much as possible. Radiotherapy machines consist of X-ray sources turning around one axis and emitting ionizing beams destroying carcinogenic cells. One crucial task in radiotherapy is to know precisely the tumor position with respect to a 3D reference point called isocentre. During classical treatment, radiotherapists determine both the number and the distribution of the irradiation angles in order to control the energy distribution in the tumoral volume and to minimize the energy passing through the healthy regions. The more precise the patient placement is, the more efficient the radiotherapist's treatment protocols are.

Treatment protocols depend on the organ to be irradiated. This paper focusses on intracranial tumor treatment. For such tumors, the positioning is usually based on metallic frames screwed on the patient's skull. The frame-based method is also employed by the radiotherapists of the oncology centre (Centre Alexis Vautrin, Nancy, France) associated to this work. The therapy always starts with a computer tomography (CT) or another similar examination, the frame being already screwed on the patient's head. The tumor borders, manually delineated in each image, are used to compute the 3D target volume and the lesion localization with regard to a coordinate system $(O_f, \vec{x}_f, \vec{y}_f, \vec{z}_f)$ given by the frame. As shown in Figure 1, $(O_f, \vec{x}_f, \vec{y}_f, \vec{z}_f)$ is defined by orthogonal slots machined into the frame. The $\vec{x}_f$, $\vec{y}_f$, and $\vec{z}_f$ vector axes take O_f (frame centre) as origin and pass through slot intersections. As both the frame and the tumor are visible in the CT, the lesion can be localized in $(O_f, \vec{x}_f, \vec{y}_f, \vec{z}_f)$. As also illustrated in Figure 1, three laser beams sweep three orthogonal planes in the therapy room. The intersections of the three plane pairs support the orthogonal vectors of the

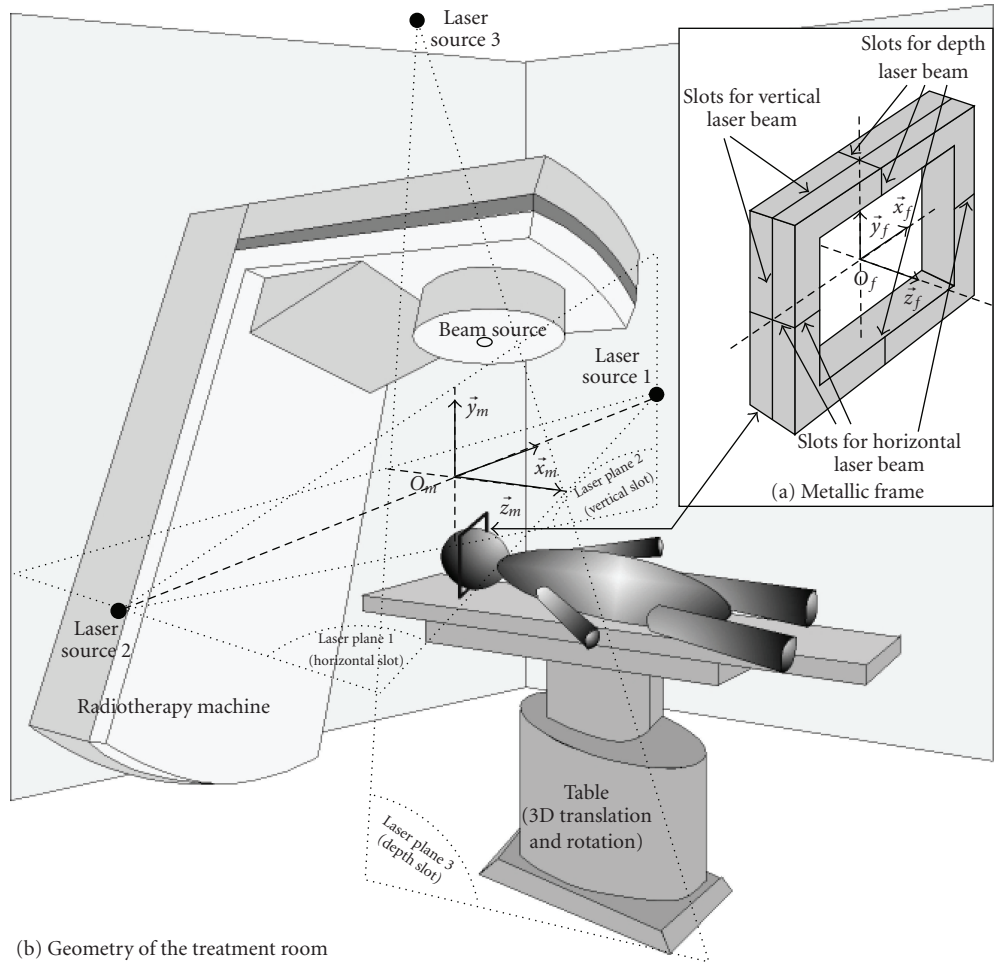


Figure 1: Principle of the frame-based method. (a) Geometry of the treatment room. (b) Frame geometry and coordinate system.

irradiation machine coordinate system $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$. O_m (isocentre) is the intersection point of the three planes. During the treatment, the patient's head is placed so that the laser planes fall into the slots. With this placement, $(O_f, \vec{x}_f, \vec{y}_f, \vec{z}_f)$ and $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$ are superimposed. Knowing the tumor localization with respect to the frame, the table on which the patient lies is displaced to bring the lesion to the isocentre.

One obvious drawback of the frame-based method lies in the fact that the treatment is traumatic for the patient (the frame is screwed on the head). Moreover, the frame can neither be fixed for a long time on the patient's head nor screwed and unscrewed several times. Consequently, the irradiation must be performed in one unique session. Meanwhile, fractioned treatment (irradiation in several sessions) is more efficient than treatment performed in one fraction. Notably, the healthy organs are less damaged in fractioned radiotherapy (FRT) than in one session irradiations. The mean positioning errors of the best invasive frame methods are 1 mm [1]. With these small errors, conformal radiotherapy (CRT) can be efficiently used. CRT is a technique which results in very accurate target volume irradiation.

1.2. Previous work

In the case of intracranial tumors, only few solutions improving the patient's positioning step of radiotherapy treatment were proposed in the literature. Noninvasive frames were conceived and tested, the screws being for example replaced by bands surrounding the head and maintaining the frame [2]. Devices fixed in the ears and on the nose were also used to maintain the frame on the patient's head [3, 4]. These devices allow radiotherapists to use FRT since the frames can be fixed several times. Meanwhile, historical results [5] have shown that these Noninvasive frames lead to a rather inaccurate positioning, the daily set up variability ranging in [1–3] mm. These positioning errors are too high when radiotherapists want to take advantage of the high irradiation accuracy of CRT.

The positioning problem in radiotherapy is to find the geometrical relationship between the coordinate systems of the therapy machine and the diagnosis device (CT, etc.). This problem lies in the fact that the two devices are usually placed in different rooms of a hospital. One way to solve this

problem is to place the diagnosis and treatment machines in the same room. For such solutions the geometrical relationship between the machines is known by construction and/or using calibration procedures. The known geometrical relationship is used either to displace the patient's table on rails [6] or with a robot [7]. These solutions lead to FRT and accurate positioning (1 mm error for [6]) but are usually far too expensive for most hospitals. For instance, a CT-scanner cannot always be dedicated to radiotherapy treatment only.

Another method employed for intracranial lesions [8] and prostate cancer [9, 10] is based on the use of portal images (PI) and digitally reconstructed radiographs (DRR) or simulated radiographs (SR). PI images are radiographs acquired during treatment. Since treatment involves high energy, the PI have poor contrast. DRR are artificial images computed with 3D CT data. The DRR are generated from the viewpoints of the PI. SR are radiographs acquired in simulation rooms having exactly the same geometry as treatment rooms, the irradiation sources being of low energy. The bone structures are the interesting information in the PI, DRR, and SR 2D planes. The disparity between the data of two modalities (IP and DRR [9, 10] or IP and SR [8]) is used to quantify the positioning quality. The bone structure segmentation and matching (registration) is done either visually or automatically. Such methods are not precise enough for CRT (1 cm error for [8] and 1.6 mm error for [9]). A noninvasive method was proposed by Meeks et al. [11] for intracranial tumors. The authors conceived a bite plate having on one of its extremities a molded part which is blocked by the patient's maxillary dentition. The bite plate supports aluminium spheres and infrared LEDs (ILEDs). Both the spheres and the tumors are visible in CT data. The tumor can be located in a coordinate system defined by the spheres. The ILEDs positions in the sphere coordinate system are obtained with a first calibration procedure. A 3D infrared sensor consisting of three cameras is fixed in the therapy room. The sensor position in the radiotherapy room is given by a second calibration. This sensor gives the ILEDs positions in the therapy room. Knowing the relative positions between the ILEDs and the spheres and the spheres and the tumors, it is possible to predict the tumor position in the treatment machine coordinate system. With this method, the mean positioning error is 1.11 mm. Among the Noninvasive solutions described in the literature, this method is one of the most accurate and can be used in CRT and FRT. Meanwhile, this accuracy was measured with respect to the results obtained for a classical frame-based method which was itself affected by errors. Moreover, the method is not suitable for people (small children and elderly people) who have missing teeth. A dedicated part (molded bite plate) must also be built for each patient.

Recently, Li et al. [12] proposed an interesting head positioning method based on 3D sensors fixed in the CT and therapy rooms. The algorithm principle can be divided into three parts consisting of a reference surface generation during CT-simulation, "controlled" patient face acquisitions in the therapy room, and data alignments providing the patient positioning parameters.

In the CT-room, the 3D sensor position is calibrated using a specially designed calibration plate. This calibration provides the geometrical link between the coordinate systems of the 3D sensor and of the CT-scanner (the 3D head surface and lesion positions are known with respect to a simulated isocentre and treatment machine coordinate system). During the CT-data acquisition, a 3D sensor is used to acquire points spread out over the patient's face. The corresponding 3D surface is placed in the planned (simulated) treatment position. The CT-face surface is not exactly the same as that given by the 3D sensor in the treatment room when face masks are used to immobilize the patient's head. Placing the 3D face surface acquired with 3D sensor in the simulation coordinate system (with the aim to replace the CT-surface) is one way to obtain a reference surface "comparable" to the face surface acquired in the treatment room. This placement is done with the calibration parameters.

Mandible or lip movements lead to nonnegligible changes in terms of facial expression. Li et al. project a light ray on the chin area and determine in real time skin/sensor distances. The mandible motions are small when the measured distances become stable (in such situations the authors verified that the acquired images were reliable). The treatment room sensor being calibrated with the same method as the CT-simulation sensor, the face point positions are known in the irradiation machine coordinate system.

The 3D surface obtained in the treatment room is then aligned with the reference surface using an iterative closest point algorithm. The geometrical parameters given by the alignment are used to adjust the head position.

Similar algorithms and sensors were used in [13] for breast lesion irradiation.

1.3. Objectives of the presented work

Considering the methods presented in the literature, the patient positioning algorithms proposed by Meeks et al. [11] and by Li et al. [12] are reference methods since they are Noninvasive and can be used in FRT. Meanwhile, the method of Meeks et al. is not suitable for people (small children and elderly people) having missing teeth. The method of Li et al. does not have this drawback. For this reason, a 3D sensor was chosen in the frame of our noninvasive patient positioning algorithm.

Facemasks are not always usable since some patients are allergic to masks or could not wear them because of a phobia. The positioning method must work with simple immobilization devices consisting of head supports and devices blocking the patient's forehead, ears, and/or mandibles. Thus, the smallest available reference face area for the positioning is the face region located between the bottom of the forehead and the bottom of the nose. With this constraint, cost and time related to the building of dedicated patient parts (face masks, dental supports, etc.) can be minimized.

Li et al. demonstrated that it is possible to position patients with submillimetre accuracy using 3D optical sensors. The aim of this contribution is to show that registration methods can lead to a robust patient positioning when using

3D sensors and simple immobilization devices. The method has to be precise even if the cutaneous face surface is not completely rigid (the surface shapes depend on facial expressions).

2. POSITIONING ALGORITHM

2.1. Algorithm principle

The difficulty relating to the patient positioning problem is due to the fact that the exact geometrical relationship between the CT coordinate system and that of the therapy machine is unknown. In other words, knowing only the tumor position in the CT coordinate system $(O_{CT}, \vec{x}_{CT}, \vec{y}_{CT}, \vec{z}_{CT})$ is not sufficient to determine the tumor position in the therapy machine coordinate system $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$.

In the case of the method used usually (invasive stereotactic frame), the relationship between $(O_{CT}, \vec{x}_{CT}, \vec{y}_{CT}, \vec{z}_{CT})$ and $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$ is known by using a third coordinate system related to the frame, namely, $(O_f, \vec{x}_f, \vec{y}_f, \vec{z}_f)$. The positioning problem can be solved because the frame ensures two functions. First, the frame provides a coordinate system in which the tumor can be localized in the therapy room. Second, the frame is also able to localize the machine coordinate system. That is the reason why the frame must be exactly in the same position on the patient's head during the whole treatment.

For the proposed method, the frame is replaced by two devices, each device having one of the two functions of the frame. The first device is a 3D sensor which is fixed in the therapy room above the patient's table. This sensor acquires the 3D surface of the patient's face. This data is used to localize the tumor in the sensor coordinate system $(O_{3DS}, \vec{x}_{3DS}, \vec{y}_{3DS}, \vec{z}_{3DS})$. The second device is a calibration piece. The geometry of this piece allows us to determine the mathematical relationship between $(O_{3DS}, \vec{x}_{3DS}, \vec{y}_{3DS}, \vec{z}_{3DS})$ and $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$. The two devices are used in the frame of a two step algorithm.

Step 1. The calibration piece (see Figure 2) was specially designed for classical therapy rooms equipped with the laser system described in Figure 1. The calibration piece consists of four spheres fixed onto a plate in which orthogonal slots were machined. The plate is positioned on the patient's table so that the laser beams fall into the slots. In this situation, the exact positions of the four sphere centres are known by construction in $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$. An image of the calibration piece is acquired with the 3D sensor and the sphere centre coordinates are computed in $(O_{3DS}, \vec{x}_{3DS}, \vec{y}_{3DS}, \vec{z}_{3DS})$. It is possible to find analytically the $\tilde{T}_{3DS,m}$ transformation linking $(O_{3DS}, \vec{x}_{3DS}, \vec{y}_{3DS}, \vec{z}_{3DS})$ to $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$ if, for a given calibration piece position, the sphere centre coordinates are known in the coordinate systems of both the 3D sensor and the therapy machine.

Step 2. During standard intracranial cancer treatment, the head borders are marked in the CT images so that the 3D head surface is systematically available. The registration of the 3D head data with the 3D face data acquired in the therapy room gives the $\tilde{T}_{CT,3DS}$ geometrical transformation linking $(O_{CT}, \vec{x}_{CT}, \vec{y}_{CT}, \vec{z}_{CT})$ to $(O_{3DS}, \vec{x}_{3DS}, \vec{y}_{3DS}, \vec{z}_{3DS})$.

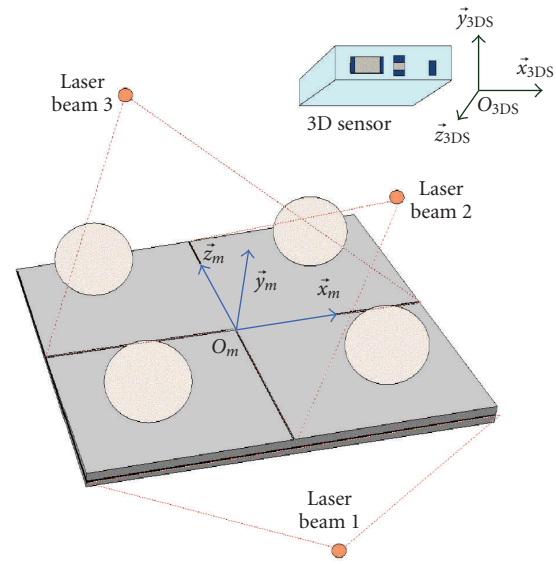


FIGURE 2: Calibration piece.

The $\tilde{T}_{3DS,m}$ and $\tilde{T}_{CT,3DS}$ transformations being matrices, the global transformation matrix $\tilde{T}_{CT,m} = \tilde{T}_{3DS,m} \times \tilde{T}_{CT,3DS}$ is sufficient to compute a given point position in $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$ with only its corresponding position known in $(O_{CT}, \vec{x}_{CT}, \vec{y}_{CT}, \vec{z}_{CT})$. Since both the 3D sensor and the CT-scanner provide data without spatial distortion and with the same isotropic scale factor of 1, $\tilde{T}_{CT,3DS}$ and $\tilde{T}_{3DS,m}$ are isometries (matrices containing only 3D translations and 3D rotations).

2.2. Data and 3D sensor description

In the CT-modality, data sets are typically represented by about 2000 points spread out over the whole cutaneous surface of the patient's head. The voxel size of the CT-scanner equals 0.313 mm × 0.313 mm × 2 mm.

The measurement principle of the 3D sensor[1] fixed in the therapy room is based on the structured light (visible light modality). The sensor is able to acquire data without any strong and particular constraints (no change in the lighting conditions, etc.). The face/3D sensor distance must only be approximatively 1 m. The typical data provided by the sensor is clouds of about 7000 points distributed over the patient's face. The field of view equals 210 mm × 320 mm for a depth of view of 100 mm. The sensor has a spatial resolution of 2 mm, 1 mm, and 0.2 mm for the $\vec{x}_{3DS}$, $\vec{y}_{3DS}$, and $\vec{z}_{3DS}$ axes, respectively.

[1] 3D flash! cam system from 3D metrics, Petaluma, CA 94954,USA.

2.3. 3D sensor calibration

The calibration starts with an acquisition of the calibration piece placed in such a way on the patient's table that the laser beams fall into the slots. In the first calibration step, the sphere centre coordinates are determined in $(O_{3DS}, \vec{x}_{3DS}, \vec{y}_{3DS}, \vec{z}_{3DS})$. The second step consists in the search for the analytical relationship $(\widetilde{T}_{3DS,m})$ between the sphere centre positions in the sensor coordinate system and the same positions in $(O_m, \vec{x}_m, \vec{y}_m, \vec{z}_m)$.

2.3.1. Sphere centre computation

For each 3D point, the sensor gives both position information and a color value. To take advantage of the color data, the calibration piece is put on black fabric. It is noticeable in Figure 2 that the color of the spheres is bright, while the plate is dark. With the color intensity information, it is easy to separate the sphere points from the other points (image background and plate points).

The geometry of the calibration piece is well known: 40 mm sphere diameters and 120 mm distances between neighbouring spheres (see Figure 3). These values, and all others relating to the calibration piece geometry, are known by construction with a 0.01 mm accuracy. During the calibration, the 3D points are sorted in four groups each corresponding to one sphere. The sorting is performed as follows: if the distance between the point currently treated and a point of a group is smaller or equal to 40 mm, then the current point is assigned to the tested point group.

For numerical reasons (the 3D sensor reconstructs the points with small errors), the points are not exactly located on a sphere. S^n is the nth sphere $(n = 1, 2, 3, 4)$ of radius r and has a centre C^n with coordinates $(x^n_{3DS}, y^n_{3DS}, z^n_{3DS})$ in $(O_{3DS}, \vec{x}_{3DS}, \vec{y}_{3DS}, \vec{z}_{3DS})$. If the ith point $p^{i,n}$ $(i \in [1, I_n]$, I_n point number of group $n)$, of coordinates $(x^{i,n}_{3DS}, y^{i,n}_{3DS}, z^{i,n}_{3DS})$, belongs to the sphere S^n, then (1) is verified:

$$\left(x^{i,n}_{3DS} - x^n_{3DS}\right)^2 + \left(y^{i,n}_{3-D} - y^n_{3DS}\right)^2 + \left(z^{i,n}_{3DS} - z^n_{3DS}\right)^2 = r^2. \tag{1}$$

The coordinates of centre C^n are determined by minimizing the functional ϵ_n given in (2),

$$\epsilon_n = \sum_{i=1}^{I_n} \left| \left(x^{i,n}_{3DS} - x^n_{3DS}\right)^2 + \left(y^{i,n}_{3-D} - y^n_{3DS}\right)^2 + \left(z^{i,n}_{3DS} - z^n_{3DS}\right) - r^2 \right|. \tag{2}$$

The initial value of the centre coordinates are given by the gravity centre of all the points of a group. The simplex [14] is used as optimization method since this algorithm is accurate and converges quickly towards the minimum when the solutions are close to the initial values.

2.3.2. Calibration matrix determination

The $\widetilde{T}_{3DS,m}$ homogeneous matrix, which provides the coordinates (x_m, y_m, z_m) of a point in the therapy machine coordinate system using the coordinates $(x_{3DS}, y_{3DS}, z_{3DS})$ of the same point in 3D sensor coordinate system, consists of nine r^{cal}_i rotation parameters and three t^{cal}_j translation parameters (see (3)):

$$\begin{pmatrix} x_m \\ y_m \\ z_m \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} r^{cal}_1 & r^{cal}_2 & r^{cal}_3 & t^{cal}_x \\ r^{cal}_4 & r^{cal}_5 & r^{cal}_6 & t^{cal}_y \\ r^{cal}_7 & r^{cal}_8 & r^{cal}_9 & t^{cal}_z \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\widetilde{T}_{3DS,m}} \begin{pmatrix} x_{3DS} \\ y_{3DS} \\ z_{3DS} \\ 1 \end{pmatrix}. \tag{3}$$

As shown in Figure 3, the three spheres S^1, S^3, and S^4 define the calibration piece coordinate system $(O_{cp}, \vec{x}_{cp}, \vec{y}_{cp}, \vec{z}_{cp})$. The fourth sphere S^2 is only used to check the calibration results consistency. The rotation parameters r^{cal}_i express the point with coordinates $(x_{3DS}, y_{3DS}, z_{3DS})$ in a rotated coordinate system having the same origin as the 3D sensor coordinate system but with axes parallel to those of $(O_{cp}, \vec{x}_{cp}, \vec{y}_{cp}, \vec{z}_{cp})$. The r^{cal}_i parameter values are given by (4) and depend on the sphere centre coordinates of (1). d_1, d_2 and $d_1 d_2$ are the norms of $\vec{x}_{cp}$, $\vec{y}_{cp}$, and $\vec{z}_{cp}$, respectively,

$$\begin{aligned}
r^{cal}_1 &= \frac{x^3_{3DS} - x^4_{3DS}}{d_1}, \qquad r^{cal}_2 = \frac{y^3_{3DS} - y^4_{3DS}}{d_1}, \\
r^{cal}_3 &= \frac{z^3_{3DS} - z^4_{3DS}}{d_1}, \qquad r^{cal}_4 = \frac{x^1_{3DS} - x^4_{3DS}}{d_2}, \\
r^{cal}_5 &= \frac{y^1_{3DS} - y^4_{3DS}}{d_2}, \qquad r^{cal}_6 = \frac{z^1_{3DS} - z^4_{3DS}}{d_2}, \\
r^{cal}_7 &= \frac{\left(y^3_{3DS} - y^4_{3DS}\right)\left(z^1_{3DS} - z^4_{3DS}\right)}{d_1 d_2} \\
&\quad - \frac{\left(y^1_{3DS} - y^4_{3DS}\right)\left(z^3_{3DS} - z^4_{3DS}\right)}{d_1 d_2}, \\
r^{cal}_8 &= \frac{\left(x^1_{3DS} - x^4_{3DS}\right)\left(z^3_{3DS} - z^4_{3DS}\right)}{d_1 d_2} \\
&\quad - \frac{\left(x^3_{3DS} - x^4_{3DS}\right)\left(z^1_{3DS} - z^4_{3DS}\right)}{d_1 d_2}, \\
r^{cal}_9 &= \frac{\left(x^3_{3DS} - x^4_{3DS}\right)\left(y^1_{3DS} - y^4_{3DS}\right)}{d_1 d_2} \\
&\quad - \frac{\left(x^1_{3DS} - x^4_{3DS}\right)\left(z^3_{3DS} - z^4_{3DS}\right)}{d_1 d_2}, \\
d_1 &= \left[\left(x^3_{3DS} - x^4_{3DS}\right)^2 + \left(y^3_{3DS} - y^4_{3DS}\right)^2 + \left(z^3_{3DS} - z^4_{3DS}\right)^2\right]^{1/2}, \\
d_2 &= \left[\left(x^1_{3DS} - x^4_{3DS}\right)^2 + \left(y^1_{3DS} - y^4_{3DS}\right)^2 + \left(z^1_{3DS} - z^4_{3DS}\right)^2\right]^{1/2}.
\end{aligned} \tag{4}$$

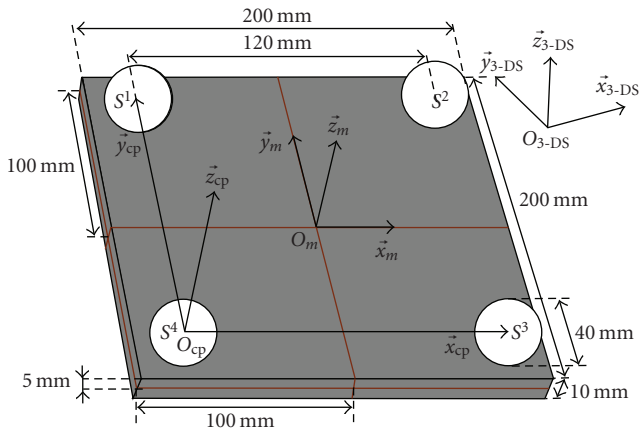


Figure 3: Calibration piece dimensions and coordinate systems.

As formulated in (5), two 3D translations define the global translation linking the 3D sensor and the therapy machine coordinate systems. The parameters of translation 1 are directly related to the coordinates $(x^4_{3DS}, y^4_{3DS}, z^4_{3DS})$ of the S^4 sphere centre position while translation 2 is completely defined by the calibration piece dimensions. Translation 2 gives the distances between the origins of the calibration piece and the therapy machine coordinate systems along the x-, y- and z-axes:

$$\begin{pmatrix} t_x^{cal} \\ t_y^{cal} \\ t_z^{cal} \end{pmatrix} = \underbrace{\begin{pmatrix} -x^4_{3DS} \\ -y^4_{3DS} \\ -z^4_{3DS} \end{pmatrix}}_{\text{translation 1}} + \underbrace{\begin{pmatrix} -60 \\ -60 \\ 25 \end{pmatrix}}_{\text{translation 2}}. \quad (5)$$

2.4. 3D data registration

An analysis of review papers dealing with medical image registration [15–17] shows that the superimposition of 3D CT data and 3D structured light data is an application that is hardly ever studied.

2.4.1. General considerations

Let us consider $I_m(x_m, y_m, z_m)$ and $I_t(x_t, y_t, z_t)$ as two 3D images containing homologous structures D_m and D_t extracted from the images with the segmentation algorithms f_m and f_t. The D_t data is transformed with the aim of superimposing it with the D_m model data. In other words, the registration procedure consists in finding the parameters θ of the $\widetilde{T}_{CT,3DS}$ transformation such as $D_m = \widetilde{T}_{CT,3DS}(D_t)$. The homologous structures are superimposed with an optimization method Ψ which minimizes a similarity measure S. The principle of the registration method providing the optimal $\widetilde{T}_{CT,3DS}$ transformation is mathematically formulated in (6):

$$\widetilde{T}_{CT,3DS} = \arg \min_{\theta \in \Theta | \Psi} S(T(\underbrace{f_t(I_t)}_{D_t}), \underbrace{f_m(I_m)}_{D_m}). \quad (6)$$

In our patient positioning application, D_m and D_t are point clouds directly provided by the sensors of the two modalities. No f_m and f_t segmentation algorithms are needed to extract the homologous structures. The advantage of our method is that the errors inherent in the segmentation algorithms are avoided.

It is noticeable that both the 3D point densities (see Section 2.2) and the 3D surface sizes are different for the two modalities. The model surface (D_m data set, patient's face of the visible light modality) is completely a part of the transformed surface (D_t data set, patient's head of the CT-modality) when the two data sets are registered.

The registration requires the definition of four mathematical entities, namely, the transformation type, the similarity measure S, the transformation space Θ giving the limits of the θ-parameters, and the search strategy (optimization Ψ).

2.4.2. Transformation type

As justified in Section 2.1, the transformation parameters are those of an isometry. The choice of the transformation type of $\widetilde{T}_{CT,3DS}$ was also realized on the assumption that a patient can make "similar enough facial expressions" during the CT-scan and the data acquisition with the 3D sensor (the impact of facial expression differences on the registration is discussed in Section 3.5).

The homogenous matrix $\widetilde{T}_{CT,3DS}$, used to determine the coordinates $(x_{3DS}, y_{3DS}, z_{3DS})$ of a point (tumor) in $(O_{3DS}, \vec{x}_{3DS}, \vec{y}_{3DS}, \vec{z}_{3DS})$ using the coordinates $(x_{3\,CT}, y_{3\,CT}, z_{3\,CT})$ of the same point in $(O_{3\,CT}, \vec{x}_{3\,CT}, \vec{y}_{3\,CT}, \vec{z}_{3\,CT})$, consists of the t_x^{reg}, t_y^{reg}, and t_z^{reg} translation parameters and of nine r_i^{reg} rotation parameters. The rotation parameters are defined with the Euler angles (for the Euler angles, the so-called "x-convention" is used: the first rotation is by an angle ψ about the $\vec{z}_{CT}$-vector, the second is by an angle $\theta \in [0, \pi]$ about the new $\vec{x}_{CT}$-vector, and the third is by an angle ϕ about the new $\vec{z}_{CT}$-vector).

2.4.3. Similarity measurement

During the registration of two surfaces, the similarity (superimposition degree) can be assessed by measuring a distance between the surfaces. In the case of surfaces represented by point clouds, the bottleneck distance [18], the Hausdorff distance (Hd) [19], the directed Hausdorff distance (dHd), or the combination of several of these distances [20] are often suitable. For a given application, a distance measure can be chosen according to the following criteria.

Data set type

A given measure is suitable or not depending on whether the two surfaces are represented by a similar or a different point number. The fact that the surfaces to be matched have the same size or not is another decision criterion.

Robustness against perturbations

Surfaces partially hidden, noise affecting the positions of all the points, or data sets with outliers, influence more or less the similarity measure correctness depending on the chosen measure.

Required transformation invariance

The measure has to exhibit appropriate properties according to the type of the geometrical transformation used in the registration scheme. For example, for isometries or affine transformations, $d(A,B) = d(T(A),T(B))$ must be verified, d being the distance between two data sets A and B.

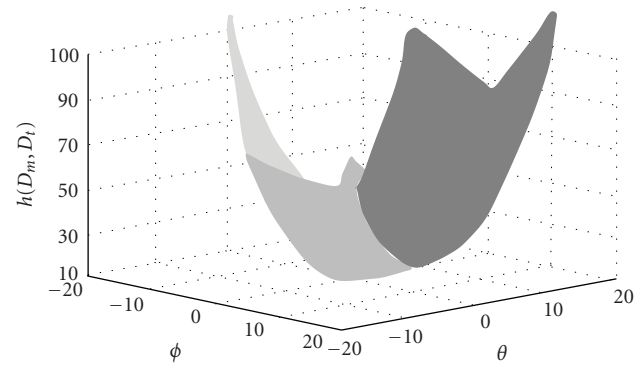
The bottleneck distance is suitable for data sets consisting of the same number of points. D_m and D_t being of different sizes, the bottleneck distance cannot be used in our application. Both the Hd and the dHd are suitable for data sets consisting in different point numbers. They are also invariant under isometries, and are robust against noise affecting the point positions. The dHd ($h(A,B)$ defined in (7)) is the greatest Euclidean distance chosen between all the smallest Euclidean distances from a point a of the data set A to all points b of the data set B. The Hd ($H(A,B)$ of (8)) is computed using the dHd,

$$h(A,B) = \max_{a\in A}\min_{b\in B}\|a-b\|, \tag{7}$$

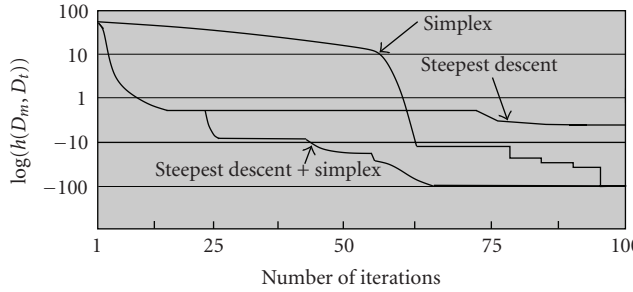
$$H(A,B) = \max\left(h(A,B),h(B,A)\right). \tag{8}$$

One advantage of the dHd, with respect to the Hd, lies in the fact that the $h(A,B)$ distance is more robust against occlusions than the $H(A,B)$ distance. In our positioning problem, the D_m model data set (patient's face) represents a smaller 3D surface than the D_t data set (patient's head). Indeed, the back of the patient's head is hidden for the 3D sensor fixed in the therapy room while the whole head is acquired in the CT-modality. Robustness against occlusions was the first criterion for choosing the dHd.

The second advantage of the dHd lies in the properties of the $h(A,B)$ and $H(A,B)$ distances. It is well known that the $H(A,B)$ distance is a metric. This means in particular that $H(A,A) = 0$ (identity) and that $H(A,B) + H(A,C) \geq H(B,C)$ (strong triangle inequality) are verified by the Hd. Symmetry ($H(A,B) = H(B,A)$) follows from the identity and strong triangle inequality. Symmetry is a propriety which is required in many matching problems. The strong triangle inequality is not verified by the dHd and consequently $h(A,B) \neq h(B,A)$. For the proposed application, if the two data sets are best registered then D_m is included in D_t. This means that for registered data, $h(D_t,D_m)$ is greater than zero and $h(D_m,D_t)$ equals 0 (in fact due to coordinate discretization, this latter value is small but never null). The dHd has also been chosen because it is interesting to have a similarity measure ($h(D_m,D_t)$) whose value is very small when the data is registered and which becomes monotonically greater when the surfaces move apart (in our application, the increasing of the similarity measure is not monotonic for the Hd).



(a) Parameter space for two rotation angles



(b) Similarity measure evolution

Figure 4: Appropriateness of the dHd. (a) Feature space for two rotation angles given in degrees. The D_m (3D sensor data) and D_t (CT modality) point sets were acquired for a phantom (plaster head). The $h(D_m,D_t)$ surface is not only convex for these two angles, but also for all other parameters of the isometry. (b) Similarity measure evolution. The decimal logarithm values of $h(D_m,D_t)$ are given for each iteration of the optimization. The combination of the steepest gradient and the simplex allows both a fast and accurate convergence.

2.4.4. Feature space limits

The interesting feature limits are those defining a parameter space Θ having a unique minimum and a convex similarity measure surface (see Figure 4(a)). The dHd measure is very robust against translations. Theoretically, there are no translation limits beyond which the surface convexity is affected. For the two D_m and D_t data modalities, it has also been verified experimentally that $h(D_m,D_t)$ decreases monotonically for rotation angles ranging between $[-20°, 20°]$. The patient's positions and the angles of view being approximately the same in the CT-scanner and on the radiotherapy table, only small rotation angles have to be considered for the registration. In this situation, the six-dimensional parameter space consisting of three translations and three rotations is effectively convex.

2.4.5. Minimization method

Experiments proved that the data of the two modalities lead to a quasiconvex hypersurface (instead of an ideal convex

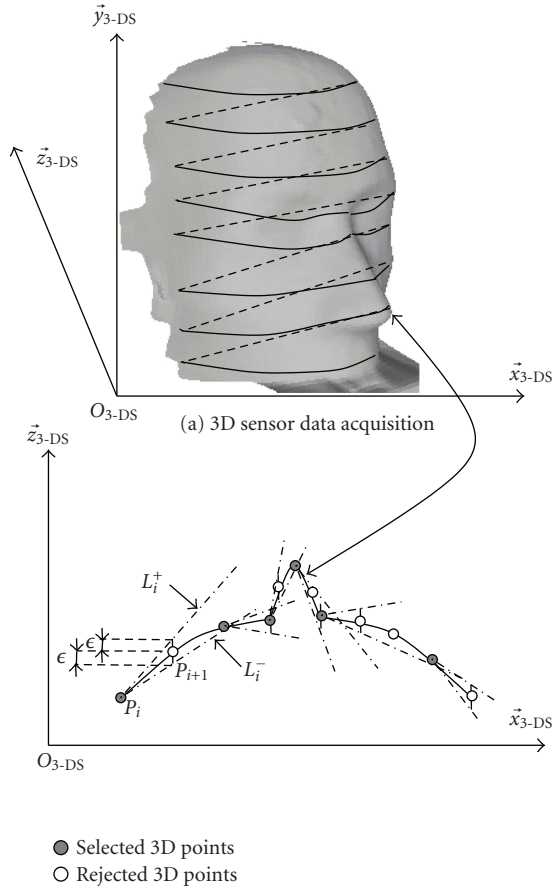


Figure 5: Data down-sampling algorithm.

surface) having one global minimum in the six-dimensional parameter space. Indeed, small local minima affect the hypersurface. A steepest descent algorithm is first used since this method converges quickly towards the solution whereas small local minima are avoided. As this algorithm only comes near to the global minimum (without reaching it), the simplex algorithm has been then used for obtaining the final $\widetilde{T}_{\text{CT,3DS}}$. The simplex algorithm is robust and accurate if the initialization is close to the solution (see Figure 4(b)).

2.4.6. Inherent accuracy of the registration algorithm

Data was acquired for a phantom (plaster head, see Figure 5(a)) with the 3D sensor in order to assess the inherent accuracy of the registration algorithm. A known $\widetilde{T}_{\text{test}}$ transformation was applied to this data set D_m, taken as model, to obtain the transformed data D_t. The registration algorithm was then used to superimpose D_t on D_m. For the $\widetilde{T}_{\text{CT,3DS}}$ matrix obtained in this way, one should ideally have $\widetilde{T}_{\text{CT,3DS}} = \widetilde{T}_{\text{test}}^{-1}$.

The parameter values of the $\widetilde{T}_{\text{test}}^{-1}$ transformation are given in the first column of Table 1. The second column of Table 1 gives the value differences between the parameters of $\widetilde{T}_{\text{CT,3DS}}$ and the corresponding ones of $\widetilde{T}_{\text{test}}^{-1}$. For the second column, the $\widetilde{T}_{\text{CT,3DS}}$ transformation was computed with the whole points of the D_m and D_t data sets (without point down-sampling, see Section 2.4.7). The greatest differences were about 1° and several hundredth of mm for, respectively, the three rotation angles (ψ, θ and ϕ) and the translations (t_x^{reg}, t_y^{reg} and t_z^{reg}). These differences lead to a mean registration error of 0.03 mm (mean Euclidian distance between homologous points of D_m and D_t transformed by $\widetilde{T}_{\text{CT,3DS}}$, namely, $\widetilde{T}_{\text{CT,3DS}}(\widetilde{T}_{\text{test}}(D_m))$). This test, performed with monomodal data, prove that the registration algorithm has high inherent accuracy.

2.4.7. Data down-sampling

The results obtained for the registration algorithm are satisfactory in terms of inherent accuracy but are not acceptable in the clinical case since the computation of $\widetilde{T}_{\text{CT,3DS}}$ requires about 4 hours on a PC with a 3.2 GHz Pentium IV processor with 2 gigabytes of RAM (the programs were written in C). This time is high since the application of the dHd to the two data sets consisting, respectively, of about 7000 3D points (visible light modality, D_m) and 2000 points (CT-modality, D_t) implies the computation of 14 million Euclidian distances. One solution to reduce the registration time is to diminish the point number of one modality. The visible light modality has been chosen since the D_m data set is the one with the most of the points.

The structured light-based sensor stores the 3D points camera line by camera line, each line having a constant $y_{3\text{DS}}$ value (see Figure 5(a)). A down-sampling algorithm (Fan algorithm [21]) whose principle is sketched in Figure 5(b) is used to eliminate points characterized by a low curvature. Two consecutive points (P_i and P_{i+1}) and a height value ϵ define two lines L_i^+ and L_i^- with a given aperture angle depending on ϵ. The selected points are P_i and the last point lying between the two lines. The last point becomes the new P_i and the algorithm is repeated until the last point on the line is reached.

By giving at ϵ the values of 0.01 (2014 remaining points for D_m), 0.1 (808 points), and 0.5 (406 points), the computation time of 4 hours ($\epsilon = 0$, whole data set D_m of 7060 points) falls, respectively, to 50 minutes, 12 minutes, and 2 minutes. The last time, also obtained with a Pentium IV processor, is acceptable in the frame of standard treatment protocols. Moreover, faster computers can be used if this time must still be reduced. It is noticeable in Table 1 that the data down-sampling with the ϵ values reported here had a very weak influence on the parameters of $\widetilde{T}_{\text{CT,3DS}}$ and on the inherent registration algorithm accuracy.

2.5. Tumor position in the therapy room

Finally, the $(x_m^{\text{tum}}, y_m^{\text{tum}}, z_m^{\text{tum}})$ tumor position in the therapy machine coordinate system can easily be computed with the $(x_{\text{CT}}^{\text{tum}}, y_{\text{CT}}^{\text{tum}}, z_{\text{CT}}^{\text{tum}})$ tumor position in the CT coordinate system

Table 1: Inherent accuracy of the registration algorithm according to the data down-sampling parameter ϵ.

$\widetilde{T}_{\text{test}}^{-1}$ parameter values	$\widetilde{T}_{\text{test}}^{-1} - \widetilde{T}_{\text{CT,3DS}}$ ($\epsilon = 0$)	$\widetilde{T}_{\text{test}}^{-1} - \widetilde{T}_{\text{CT,3DS}}$ ($\epsilon = 0.01$)	$\widetilde{T}_{\text{test}}^{-1} - \widetilde{T}_{\text{CT,3DS}}$ ($\epsilon = 0.1$)	$\widetilde{T}_{\text{test}}^{-1} - \widetilde{T}_{\text{CT,3DS}}$ ($\epsilon = 0.5$)
$\theta = 7°$	0.9	0.91	0.91	0.92
$\phi = 9°$	−0.6	−0.6	−0.62	−0.62
$\psi = -5°$	1.09	1.09	1.58	1.06
$t_x^{\text{reg}} = 50\,\text{mm}$	−0.02	−0.02	−0.03	−0.03
$t_y^{\text{reg}} = -40\,\text{mm}$	−0.01	−0.01	−0.01	−0.01
$t_z^{\text{reg}} = 50\,\text{mm}$	0.06	0.06	0.06	0.06

and the global transformation matrix $\widetilde{T}_{\text{CT,3DS}}$ (see (9)):

$$\begin{pmatrix} x_m^{\text{tum}} \\ y_m^{\text{tum}} \\ z_m^{\text{tum}} \\ 1 \end{pmatrix} = \underbrace{\widetilde{T}_{\text{3DS},m}\widetilde{T}_{\text{CT,3DS}}}_{\widetilde{T}_{\text{CT,3DS}}} \begin{pmatrix} x_{\text{CT}}^{\text{tum}} \\ y_{\text{CT}}^{\text{tum}} \\ z_{\text{CT}}^{\text{tum}} \\ 1 \end{pmatrix}. \tag{9}$$

3. EXPERIMENTS AND RESULTS

3.1. Simulation room description

During standard treatment, the patient positioning is first realized in a simulation room in order to assess the positioning accuracy and to check the dose distribution. The simulation room is geometrically identical to the treatment room. The two rooms are also equipped with the same devices. In particular, the simulation machine isocentre is also visualized by three laser beams.

However, between the two rooms there is a major difference related to the energy emitted by the irradiation sources. The linear accelerator of the treatment room is characterized by high energy whereas the source of the simulation machine is suitable to the realization of radiographic films (control radiographs). Such radiographs are generally taken for two well-defined viewpoints (see Figure 6).

As for the therapy machine coordinate system, the (O_{sm}, $\vec{x}_{\text{sm}}, \vec{y}_{\text{sm}}, \vec{z}_{\text{sm}}$) simulation machine coordinate system is completely defined and visualized by the laser beams. The two control radiographs are orthogonal since the first radiograph is parallel to the plane defined by the ($\vec{x}_{\text{sm}}, \vec{y}_{\text{sm}}$) axis pair, while the second radiograph is parallel to the ($\vec{z}_{\text{sm}}, \vec{y}_{\text{sm}}$) plane.

Moreover, a metallic cross is fixed in front of the X-ray source. The axis passing both through the X-ray point source and the 3D intersection point of the metallic cross is perpendicular to the radiograph planes, to the ($\vec{x}_{\text{sm}}, \vec{y}_{\text{sm}}$) plane of the first viewpoint and to the ($\vec{z}_{\text{sm}}, \vec{y}_{\text{sm}}$) plane of the second viewpoint. With this geometry, the projection of the axes $\vec{x}_{\text{sm}}$, $\vec{y}_{\text{sm}}$, and $\vec{z}_{\text{sm}}$ of the simulation machine coordinate system is visualized exactly by the projections of the metallic cross onto the radiographs.

The proposed positioning algorithm was tested in the simulation room.

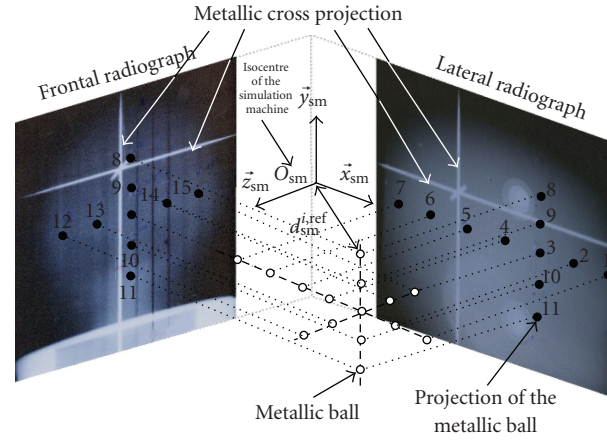


Figure 6: Simulation room. (a) Orthogonal control radiographs. (b) Room geometry.

3.2. Phantom description and CT-data

Tests were performed with a plaster head acting as phantom (see Figure 5). Fifteen metallic balls (simulating tumors) were included in the head. These radio-opaque balls, with a mean diameter of 5 mm, were regularly spaced and placed exactly on three orthogonal axes. Figure 6(a) gives the labels of these balls. It is noticeable that the balls are distributed into the whole head volume so that it can be checked if the positioning accuracy depends on the tumor localization.

Table 2: First positioning test results. The i-ball labels are those of Figure 6 (all values are given in millimeters).

i	$x_{sm}^{i,com}$	$y_{sm}^{i,com}$	$z_{sm}^{i,com}$	$d_{sm}^{i,com}$
1	0.13	−0.08	0.12	0.19
2	0.11	0.12	0.18	0.24
3	0.18	−0.02	0.04	0.19
4	0.47	0.11	0.35	0.59
5	−0.27	−0.10	−0.07	0.3
6	0.29	0.02	−0.18	0.34
7	0.6	0.29	0.3	0.74
8	−0.21	0.17	0.15	0.31
9	0.31	0.06	−0.18	0.36
10	0.18	−0, 31	0.32	0.48
11	0.26	−0.34	0.08	0.44
12	0.31	0,21	0.28	0.47
13	−0.16	0.21	0.42	0.5
14	0.15	0.23	−0.19	0.33
15	0.36	−0.17	0.28	0.49

A scan was performed with the plaster head placed in the CT-machine. The balls were spread out on several voxels of the CT. The mass centre positions $(x_{CT}^{i}, y_{CT}^{i}, z_{CT}^{i})$ were computed for each ball p_i ($i \in [1, 15]$).

3.3. First positioning test

The balls were successively placed at the simulation machine isocentre by superimposing the ball projections and the cross intersection projections viewed on the two control radiographs. This placement can be done very accurately by experienced radiotherapists. The laser positions on the plaster head were marked precisely for each ball placement on O_{sm}. Thus, the placement of the marks on the laser beams ensures a very accurate positioning of the balls on the isocentre. If a ball is placed on O_{sm}, then the positioning algorithm should ideally give $(0, 0, 0)$ as result for the ball coordinates in $(O_{sm}, \vec{x}_{sm}, \vec{y}_{sm}, \vec{z}_{sm})$. It is noticeable that this positioning experiment is conducted like a true patient positioning in the therapy room.

The sensor was fixed in the simulation room and its position was calibrated in the $(O_{sm}, \vec{x}_{sm}, \vec{y}_{sm}, \vec{z}_{sm})$ coordinate system. The surface of the plaster head given by the CT-scan was registered with the plaster head's face acquired in the simulation room. The balls p_i were all placed at the isocentre and positions of their centres were computed with (9). The $(x_{sm}^{i,com}, y_{sm}^{i,com}, z_{sm}^{i,com})$ ball coordinates in $(O_{sm}, \vec{x}_{sm}, \vec{y}_{sm}, \vec{z}_{sm})$ and their $d_{sm}^{i,com}$ distances to O_{sm} are given in Table 2. The mean and standard deviation values of the $d_{sm}^{i,com}$ distances are, respectively, 0.4 mm and 0.15 mm. With these results, several observations can be formulated.

The mean positioning error is very small and indicates a submillimetre accuracy.

No correlation can be established between tumor positions and positioning errors. In other terms, a weak variability affects the positioning accuracy when considering different head regions (head centre or skull region). This result is important since the lower this variability is, the more the positioning errors are predictable.

Table 3: Ball 3 positioning results for different acquisitions and 3D sensor viewpoints. $d_{sm}^{3,com}$ is the distance between the ball with coordinates $(x_{sm}^{3,com}, y_{sm}^{3,com}, z_{sm}^{3,com})$ and the isocentre (All values are given in millimeters).

Acquisition number	$x_{sm}^{3,com}$	$y_{sm}^{3,com}$	$z_{sm}^{3,com}$	$d_{sm}^{3,com}$
1	0.24	0.12	0.01	0.27
2	0.1	0.01	0.03	0.1
3	0.11	0.12	0.21	0.27
4	0.18	−0.02	0.04	0.19
5	0.14	0.13	0.22	0.29
6	−0.21	−0.07	0.12	0.25
7	0.03	0.15	−0.18	0.24
8	0.17	−0.08	0.11	0.22
9	0.08	−0.14	0.10	0.19

The voxel of the CT-modality having a size of 0.313 mm × 0.313 mm × 2 mm means that the $(x_{CT}^{i,com}, y_{CT}^{i,com}, z_{CT}^{i,com})$ centre coordinates of the balls (with a 5 mm diameter) are affected by errors. Theses errors have also an impact on the patient positioning accuracy. The positioning accuracy can still be improved with scanners (CT-modality or other modalities) delivering volume data with a higher resolution.

3.4. Second positioning test

The purpose of the second positioning test was the assessment of the variability of the positioning results with regards to the calibration data, the phantom data, and sensor viewpoint differences. Indeed, from one acquisition to another, the distributions of the 3D sensor points on the calibration piece spheres and on the plaster head are different, even if the point density remain quasiconstant. Concerning the viewpoint differences, acquisitions were performed for sensor/object distances ranging in [90, 110] cm and for angle deviations (from reference angles) belonging in $[-10°, 10°]$.

Each ball was acquired several times for different angles of view. Images of the calibration piece were also taken for each sensor position. For each ball, the mean distance and the standard deviation were computed for the $d_{sm}^{i,com}$ distances to the isocentre. The standard deviation, acting as first criterion for the assessment of the isocentre/ball distance variability, was smaller than 0.1 mm for the fifteen balls. The values given in Table 3 for ball 3 are representative of the positioning algorithm variability. The standard deviation with respect to the mean value of the $d_{sm}^{3,com}$ distances of ball 3 is 0.055 mm. The mean distance between the mean position of a ball and the different positions of the same ball is another criterion allowing the assessment of the positioning variability. The mean distance between the positions of ball 3 and the (0.093, 0.024, 0.073) mean position coordinates of ball 3 is 0.18 mm. It is noticeable that the mean position of ball 3 is

very close to the (0,0,0) isocentre coordinates. The small values obtained for the two algorithm variability criteria show that the positioning algorithm is relatively independent towards sensor position differences and different 3D point distributions. Moreover, it is noticeable that the sensor can be fixed once and for all in an optimal position in terms of patient positioning accuracy. Thus, the positioning accuracy dependency according to the sensor position is not a crucial problem.

3.5. Registration of human faces

The only step of the positioning algorithm which can lead to different results when human data is used instead of phantom data is the 3D surface registration. Two tests were carried out to assess the influence of the nonrigid cutaneous surface on the registration algorithm.

First registration test

A first image is acquired with the 3D sensor for a person who takes a neutral expression (eyes open in a natural way and closed mouth). This image simulates the CT-data. A second image was taken immediately after the first acquisition. Even if the person was asked to keep the same expression (the mouth remained closed), differences exist between the two images (eyelids more or less open, teeth more or less clenched, different point distribution over the face, etc.). In the second image, the data included in a window comprised between the bottom of the forehead and the bottom of the nose (see Figure 7(b)) was manually extracted (this face region can automatically be extracted by looking for the high curvature points corresponding to the nose and to the orbital arches). A transformation consisting of some decimetre translation components and of three rotation angles ranging each in $[-10°, 10°]$ is applied to the 3D surface extracted from the second image. The extracted data was then registered with the first image. This test was done for 15 women and men.

After registration, the distances between each point of the transformed surface (second image) and the corresponding computed points on the reference surface (closest points on the surface of the first image) are determined. The mean distance between these homologous points never exceeded 0.1 mm for all 15 people. It is noticeable that this mean value is only a little bit greater than the 0.03 mm inherent registration accuracy computed for the ideal phantom data (see Section 2.4.6). The 0.1 mm distances correspond to errors smaller than 1° and one tenth of millimetres for the angles and the translations, respectively. These results prove that the registration scheme based on the dHd is very robust and accurate, not only for phantom data, but also for human data. A high registration accuracy can be obtained since, in the considered window, the anatomical parts supporting the skin (orbital arches, nose and cheek-bone) are rigid surfaces. In this region, skin movements affect only slightly the 3D face shape. The mouth and the essential parts of the cheeks (nonrigid regions) are outside the window.

Second registration test

The aim of the second test is to assess the registration algorithm accuracy and robustness in more extreme situations. A first image is again acquired for people. For this reference image, the people systematically closed their eyes and their mouth (teeth slightly clenched). This face posture can easily be maintained. Other images were acquired for each person with different face configurations: closed mouth/open eyes, open mouth/closed eyes, and open mouth/open eyes. Transformations consisting of some decimetre translation components and of three rotation angles ranging each in $[-10°, 10°]$ are applied to these images. The latter are then registered with the reference image. The whole data of each image (no data extraction) was used during the registration.

Figure 7 illustrates typical results obtained with different men and women. Figure 7(c) allows a quantitative assessment of the registration quality of the 3D data represented by the images of Figures 7(a) and 7(b). The graphic of Figure 7(c) gives, for each point of the transformed image, the shortest distance to the surface of the model image. The distances between these homologous points vary greatly according to the face region. It was verified that the distances between homologous points located around the mouth or on the chin, on the cheeks, and on the regions close to the nose peak or orbital arches are, respectively, greater than 3 mm, range approximatively in $[0.3, 3]$ mm or are smaller than about 0.3 mm. These observations are coherent since

(i) if the images are well registered, big differences exist for the mouth and the chin due to unconscious movements,
(ii) some millimetre variability is normal for points located on cheeks which have a low rigidity, and
(iii) small errors for points located on the nose peak and on the orbital arches are predictable since these face parts are the most rigid (opening the mouth does not normally change the nose position).

If only the points on the nose peak and around the orbital arches are considered, the mean $\overline{d}_{\text{rigid}}$ distance between homologous points after registration is 0.09 mm. The $\overline{d}_{\text{all}}$ mean distance computed for all points is 1.6 mm. Meanwhile, the last measure is strongly influenced by the points located around the mouth and on the chin. Without these last points, the $\overline{d}_{\text{cheeks}}$ mean distance including the cheek points is 1.05 mm. However, the $\overline{d}_{\text{rigid}}$ measure is the most pertinent (since it is based on rigid face parts) and indicates that the registration had a submillimetre accuracy for the man of Figure 7. It is noticeable that the $\overline{d}_{\text{rigid}}$ value is close to the 0.1 mm mean distance obtained for the first registration test with the window.

The same observation can be made for the woman of Figures 7(e) and 7(f) with closed eyes (reference image) and open eyes (image to be transformed). After the registration the $\overline{d}_{\text{all}}$, $\overline{d}_{\text{cheeks}}$, and $\overline{d}_{\text{rigid}}$ mean distances are 1.2 mm, 0.99 mm and 0.11 mm, respectively. $\overline{d}_{\text{rigid}}$ indicates again a submillimetre registration accuracy. Similar results were obtained for all people, even if both the eyes and mouth were open.

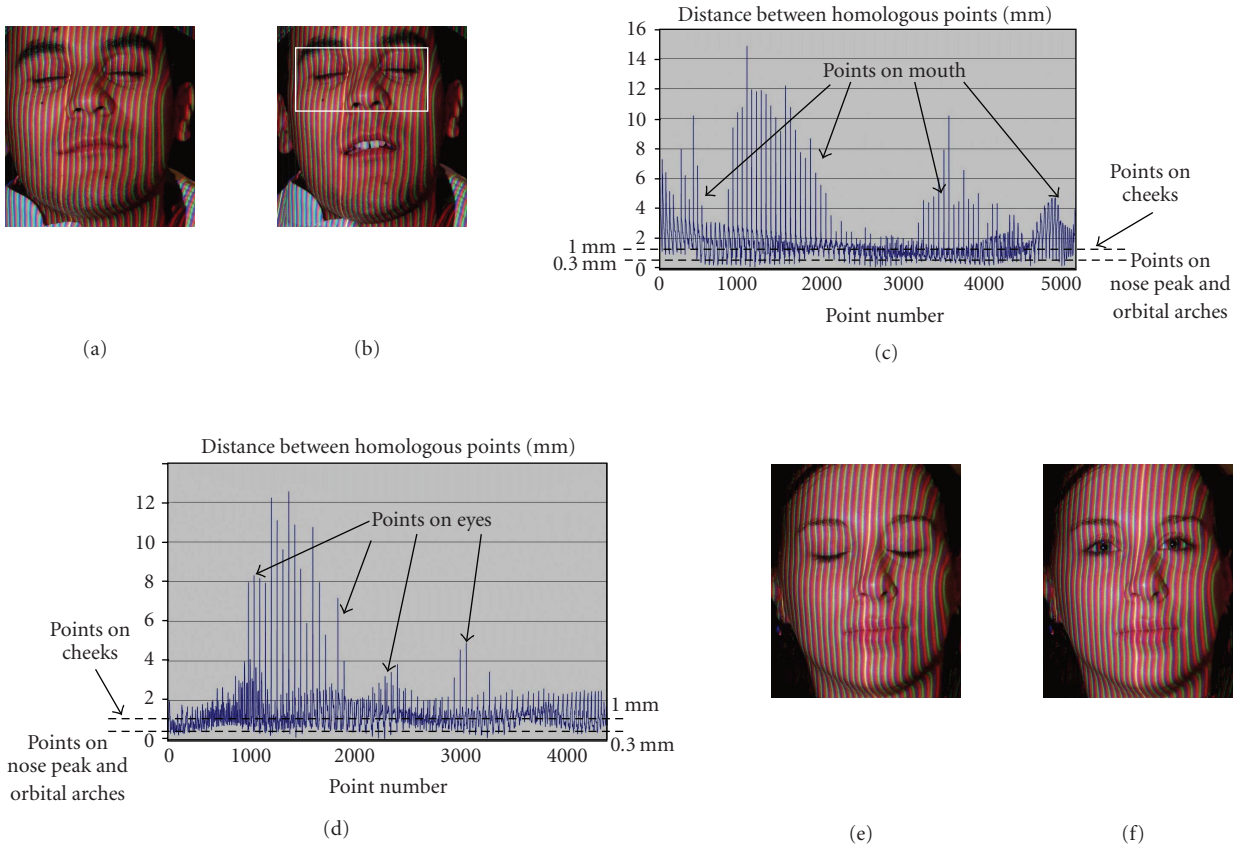


FIGURE 7: Registration results for humans. (a) Man with closed eyes and mouth. (b) Man with closed eyes and open mouth. The window in (b) indicates the data of the image to be transformed in the first registration test (the whole 3D points are used in the second test). The colour rays are the structured light information used for the 3D point reconstruction. (c) Computed distances between homologous points of (a) and (b) after registration. (d) Computed distances between homologous points of (e) and (f) after registration. (e) Woman with closed eyes and mouth. (f) Woman with open eyes and closed mouth.

Registration result discussion

The first tests presented here indicate that the 3D points between the bottom of the forehead and of the nose should systematically be extracted from the data set to be transformed before registering it with the model data (whole face points). It is recalled that this face part is always visible when the devices defined in section 1.3 are used. With this way to proceed, the distances between the homologous face points have a very small mean value (0.1 mm). The tests also proved that eyelid movements have a negligible impact on the registration accuracy. The tests with the phantom demonstrated that there is no correlation between the positioning accuracy and the lesion position in the head (Section 3.4). This fact indicates that if the lesion is close to the face surface (0.1 mm error) or in the head centre (two very different localizations), the lesion localization error due to the registration is always about 0.1 mm (or at least by far smaller than 1 mm). Thus, the first advantage of the dHd taken as similarity measure lies in the fact that small face surfaces lead to an accurate registration. The only condition is that face regions with enough geometrical information are included (regions with high curvatures like the nose or orbital arches) in the data. Another advantage of the dHd is its ability to register two surfaces of different sizes and point densities (this measure is often used when surface data of two different modalities must be registered).

The second tests proved that the dHd is able to register two surfaces presenting large geometrical differences while ensuring submillimetre alignment accuracy. In fact, the tests confirmed that the dHd can handle data containing outliers (points on the mouth or on the eyes in the case of very strong eyelid movement) without greatly affecting the registration accuracy.

The tests also proved that the registration algorithm converges in a robust way towards the solution, even for big head position differences between the two modalities. Moreover, neither an initial manual alignment nor an initial homologous point marking is required. According to the literature, the dHd leads to registration accuracies which are almost independent of the translation differences between surfaces. This fact was confirmed by the results. Orientation differences (around each axis) ranging in $[-10°, 10°]$ were always successfully treated by the registration algorithm. The tested

position differences are greater than those encountered in clinical situations. In fact, radiotherapists place the patient with initial errors of one centimetre (or a few millimetres) and some degrees in terms of translations and rotations. The proposed registration scheme is able to handle bigger differences in an automatic way.

4. CONCLUSION

The results presented in this contribution prove that the proposed algorithm is an important first step towards a patient positioning which allows for the association of CRT and FRT in the case of intracranial lesion irradiation. Tests with a phantom proved that the inherent accuracy of the whole positioning algorithm (sensor calibration and registration) is 0.4 mm. Registration tests with human data proved that the mean alignment errors are very small (about one tenth of millimetres). This registration accuracy leads us to think that the whole positioning method will also lead to a submillimetre accuracy for patient data. In fact, as suggested by Li et al. [12], if the calibration and the registration have each a submillimetre accuracy, the limitation in terms of precision is rather due to the precision of the patient immobilization devices than to positioning algorithm precision. The fact that Li et al. obtained a submillimetre accuracy with similar algorithm principles and sensors indicates that it is also possible to reach a submillimetre accuracy for patients. The next step of the positioning algorithm evaluation will consist in experiments conducted as follows. Patients will be positioned with the classical invasive frame-based method. The proposed algorithm will be used in parallel to obtain a second tumor coordinate set. The later coordinates will be compared to those given by the frame based method. Control radiographs will also be used to test the positioning accuracy of the algorithm with patient data.

The proposed method is noninvasive and no dedicated piece must be built for patients. Standard treatment protocols are not influenced by the algorithm. Moreover, only conventional and simple immobilization devices are required. The drawbacks relating to frames or face masks are avoided.

One of the main results of this contribution lies in the performances of the registration algorithm. The optimization method converges robustly and accurately towards the solution, even for large head position differences. Facial expression changes can also be processed by the algorithm.

Phantom-based tests proved that the positioning accuracy does not depend on the lesion position in the head. The fact that the irradiation must be done with well-known errors (at least submillimetre errors) explains why it is important for the positioning accuracy to be independent of the lesion localization.

ACKNOWLEDGMENTS

This work was partially sponsored by the ECOS-program, the CONACYT, the "Ligue contre le cancer," and the Centre Alexis Vautrin of Nancy. The authors would like to thank the DGEST for its participation in this work.

REFERENCES

[1] M. W. Gross, U. Spahn, and R. Engenhart-Cabillic, "Assessment of the accuracy of a conventional simulation for radiotherapy of head and skull base tumors," *Technology in Cancer Research and Treatment*, vol. 2, no. 4, pp. 345–351, 2003.

[2] S. S. Gill, D. G. T. Thomas, A. P. Warrington, and M. Brada, "Relocatable frame for stereotactic external beam radiotherapy," *International Journal of Radiation Oncology, Biology, Physics*, vol. 20, no. 3, pp. 599–603, 1991.

[3] M. I. Hariz, R. Henriksson, P.-O. Löfroth, L. V. Laitinen, and N. E. Säterborg, "A non-invasive method for fractioned stereotactic irradiation of brain tumors with linear accelerators," *Radiotherapy and Oncology*, vol. 17, no. 1, pp. 57–72, 1990.

[4] M. Delannes, N. Daly, J. Bonnet, J. Sabatier, and M. Tremoulet, "Fractioned radiotherapy of small inoperable lesions of the brain using a non-invasive stereotactic frame," *International Journal of Radiation Oncology, Biology, Physics*, vol. 21, no. 2, pp. 749–755, 1991.

[5] J. D. Graham, A. P. Warrington, S. Gill, and M. Brada, "A non-invasive, relocatable stereotactic frame for fractionated radiotherapy and multiple imaging," *Radiotherapy and Oncology*, vol. 21, no. 1, pp. 60–62, 1991.

[6] M. Uematsu, A. Shioda, A. Suda, et al., "Computed tomography-guided frameless stereotactic radiotherapy for stage I non-small cell lung cancer: a 5-year experience," *International Journal of Radiation Oncology, Biology, Physics*, vol. 51, no. 3, pp. 666–670, 2001.

[7] C. Mavroidis, J. Flanz, S. Dubowsky, P. Drouet, and M. Goitein, "High performance medical robot requirements and accuracy analysis," *Robotics and Computer-Integrated Manufacturing*, vol. 14, no. 5-6, pp. 329–338, 1998.

[8] E. N. J. Th. Van Lin, L. van der Vight, H. Huizenga, J. H. A. M. Kaanders, and A. G. Visser, "Set-up improvement in head and neck radiotherapy using a 3D off-line EPID-based correction protocol and a customised head and neck support," *Radiotherapy and Oncology*, vol. 68, no. 2, pp. 137–148, 2003.

[9] D. Sarrut and S. Clippe, "Patient positioning in radiotherapy by registration of 2D portal to 3D CT images by a contend-based research with similarity measures," in *Proceedings of 14th International Congress and Exhibition on Computer Assisted Radiology and Surgery (CARS '00)*, pp. 707–712, Elsevier Science, San Francisco, Calif, USA, June-July 2000.

[10] R. Bansal, L. H. Staib, Z. Chen, et al., "Entropy-based, multiple-portal-to-3D CT registration for prostate radiotherapy using iteratively estimated segmentation," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '99)*, vol. 1679 of *Lecture Notes in Computer Science*, pp. 567–578, Cambridge, UK, September 1999.

[11] S. L. Meeks, F. J. Bova, T. H. Wagner, J. M. Buatti, W. A. Friedman, and K. D. Foote, "Image localization for frameless stereotactic radiotherapy," *International Journal of Radiation Oncology, Biology, Physics*, vol. 46, no. 5, pp. 1291–1299, 2000.

[12] S. Li, D. Liu, G. Yin, P. Zhuang, and J. Geng, "Real-time 3D-surface-guided head refixation useful for fractionated stereotactic radiotherapy," *Medical Physics*, vol. 33, no. 2, pp. 492–503, 2006.

[13] D. Djajaputra and S. Li, "Real-time 3D surface-image-guided beam setup in radiotherapy of breast cancer," *Medical Physics*, vol. 32, no. 1, pp. 65–75, 2005.

[14] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer*, vol. 5, no. 7, pp. 308–313, 1965.

[15] T. Mäkelä, P. Clarysse, O. Sipilä, et al., "A review of cardiac image registration methods," *IEEE Transactions on Medical Imaging*, vol. 21, no. 9, pp. 1011–1021, 2002.

[16] P. A. van den Elsen, E. J. D. Pol, and M. A. Viergever, "Medical image matching-a review with classification," *IEEE Engineering in Medicine and Biology Magazine*, vol. 12, no. 1, pp. 26–39, 1993.

[17] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.

[18] A. Efrat and A. Itai, "Improvements on bottleneck matching and related problems using geometry," in *Proceedings of the 12th Annual Symposium on Computational Geometry*, pp. 301–310, Philadelphia, Pa, USA, May 1996.

[19] K. J. Kirchberg, O. Jesorsky, and R. Frischholz, "Genetic model optimization for Hausdorff distance-based face localization," in *Proceedings of the International ECCV 2002 Workshop on Biometric Authentication*, vol. 2359 of *Lecture Notes in Computer Science*, pp. 103–111, Copenhagen, Denmark, June 2002.

[20] P. Indyk and S. Venkatasubramanian, "Approximate congruence in nearly linear time," in *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 354–360, San Francisco, Calif, USA, January 2000.

[21] A. E. Pollard and R. C. Barr, "Adaptive sampling of intracellular and extracellular cardiac potentials with the fan method," *Medical and Biological Engineering and Computing*, vol. 25, no. 3, pp. 261–268, 1987.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 13963, 12 pages
doi:10.1155/2007/13963

Research Article

Multimodality Data Integration in Epilepsy

Otto Muzik,[1,2] Diane C. Chugani,[1] Guangyu Zou,[3] Jing Hua,[3] Yi Lu,[3] Shiyong Lu,[3] Eishi Asano,[4] and Harry T. Chugani[4]

[1] *Carman and Ann Adams Department of Pediatrics, Children's Hospital of Michigan, Detroit Medical Center, Wayne State University, Detroit, MI 48201, USA*
[2] *Department of Radiology, Children's Hospital of Michigan, Detroit Medical Center, Wayne State University, Detroit, MI 48201, USA*
[3] *Department of Computer Science, Wayne State University, Detroit, MI 48201, USA*
[4] *Department of Neurology, Children's Hospital of Michigan, Detroit Medical Center, Wayne State University, Detroit, MI 48201, USA*

Received 13 September 2006; Accepted 8 February 2007

Recommended by Haim Azhari

An important goal of software development in the medical field is the design of methods which are able to integrate information obtained from various imaging and nonimaging modalities into a cohesive framework in order to understand the results of qualitatively different measurements in a larger context. Moreover, it is essential to assess the various features of the data quantitatively so that relationships in anatomical and functional domains between complementing modalities can be expressed mathematically. This paper presents a clinically feasible software environment for the quantitative assessment of the relationship among biochemical functions as assessed by PET imaging and electrophysiological parameters derived from intracranial EEG. Based on the developed software tools, quantitative results obtained from individual modalities can be merged into a data structure allowing a consistent framework for advanced data mining techniques and 3D visualization. Moreover, an effort was made to derive quantitative variables (such as the spatial proximity index, SPI) characterizing the relationship between complementing modalities on a more generic level as a prerequisite for efficient data mining strategies. We describe the implementation of this software environment in twelve children (mean age 5.2 ± 4.3 years) with medically intractable partial epilepsy who underwent both high-resolution structural MR and functional PET imaging. Our experiments demonstrate that our approach will lead to a better understanding of the mechanisms of epileptogenesis and might ultimately have an impact on treatment. Moreover, our software environment holds promise to be useful in many other neurological disorders, where integration of multimodality data is crucial for a better understanding of the underlying disease mechanisms.

1. INTRODUCTION

With ever-improving imaging technologies and high-performance computational power, the complexity and scale of brain imaging data have continued to grow at an explosive pace. Recent advances in imaging technologies, such as molecular imaging using positron emission tomography (PET), structural imaging using high-resolution magnetic resonance (MR), as well as quantitative electrophysiological cortical mapping using electroencephalography (EEG), have allowed an increased understanding of normal and abnormal brain structures and functions [1–4] It is well understood that normal brain function is dependent on the interactions between specialized regions of the brain which process information within local and global networks. Consequently, there is a need to integrate the acquired multimodality data in order to obtain a more detailed understanding about process interaction in a complex biological system. In addition to integration, advanced computational tools need to be developed which allow quantitative analysis of a variety of functional patterns as well as the design of a software environment that allows quantitative assessment of relationships between diverse functional and anatomical features. Although this is an area of active research, current state-of-the-art technologies are still suboptimal with respect to multimodality integration and quantitative assessment of qualitatively distinct neuroimaging datasets in patients with structurally abnormal brains. This paper presents a multimodality database for neuroimaging, electrophysiological, as well as clinical data specifically designed for the assessment of epileptic foci

in patients with epilepsy, and can serve as a standard template for other diagnostic procedures which employ multiple imaging technologies.

The paper is organized as follows. In Section 2, we discuss previous work by other investigators and how it relates to the proposed framework. In Section 3, we introduce a novel image analysis method able to assess quantitatively the relationship between imaging and electrophysiological data. In Section 4, we describe the integration of diverse data sets into an extendable database structure and describe available queries for quantitative analysis of multimodality data. Section 5 details the implementation issues and shows the applications of the presented environment. Finally, in Section 6 we discuss the strength and limitations of the proposed system and conclude with potential future directions.

2. RELATED WORK

The accurate segmentation of anatomical brain structures from medical images and their compact geometrical representation in a rigorous computational framework is difficult due to the complexity and physiological variability of the structures under study. In the past, a large number of techniques have been applied to achieve spatial standardization of the cortical surface within a group of subjects, but only a few methods are currently available for the definition of homotopic volume elements within the whole brain. For example, Thompson et al. [5], Lohmann [6], and Vaillant et al. [7] used a manually labeled atlas brain, which was then warped to fit an individual subject's brain surface with subsequent transfer of labels onto the subject's cortical surface. Another related method is the 3D stereotactic surface projection (3D-SSP) technique [8], which is based on the extraction of subcortical functional data onto a predefined set of surface pixels following stereotactic anatomical standardization of an individual brain. More recently, conformal mapping techniques have gained a wider application in brain mapping [9]. For example, Hurdal and Stephenson [10] proposed a discrete mapping approach that uses spherical packing in order to produce a "flattened" image of the cortical surface onto a sphere yielding maps that are quasiconformal approximations to classical conformal maps. Moreover, Gu et al. [11] proposed optimization of the conformal parameterization method by composing an optimal Möbius transformation so that it minimizes the landmark mismatch energy. Based on the work of the above investigators as well as on our previous work with landmark-constrained conformal surface mapping [12], we have developed a landmark-constrained conformal mapping of the brain where the mapping accuracy is achieved through matching of various cortical landmarks (e.g., central sulcus, Sylvian fissure, interhemispheric fissure). These landmarks are manually defined in the subject's native space and subsequently aligned in the spherical domain using landmark-constrained conformal mapping. Finite cortical elements are defined geometrically in the spherical domain and subsequently reversely mapped into the subject's native space where all data analyses take place. As the position of cortical landmarks is affected by structural abnormalities, this method incorporates both physiological variations as well as structural abnormalities into the mapping process.

The creation of brain image databases which allow storage, organization, and the sharing of processed brain data between investigators with different scientific backgrounds in easily accessible archives is one of the main goals of the newly emerging field of neuroinformatics. This objective is becoming increasingly important for researchers in the brain imaging field as an unprecedented amount of brain data is now acquired using complementing modalities which need to be analyzed based on an overarching strategy [13–15]. Following the model of other scientific fields which were revolutionized through the implementation of powerful database structures such as the GenBank (http://www.ncbi.nlm.nih.gov) for the field of genomics or the Protein Data Bank (http://www.rcsb.org/pdb) for the field of proteomics, several neuroimaging databases were generated in the past decade in order to fully mine the information present in the acquired data. The Visible Human Project website http://www.nlm.nih.gov/research/visible/visible_human.html allows researchers to view high-resolution image sections of two human cadavers and is widely used in education and training. More comprehensive human brain databases which attempt to combine the expertise of computer programmers, statisticians, and basic researchers in an attempt to develop neuroinformatics tools for interdisciplinary collaboration are the ICBM database http://www.loni.ucla.edu which includes PET, MRI, fMRI, EEG, and MEG modalities [16] as well as the BrainMapDBJ http://www.brainmapDBJ.org and the ECHBD databases which both integrate PET and fMRI image data [17]. In comparison to these multiinstitutional large-scope archives, our database is specific to the task of understanding the mechanisms of epileptogenesis in patients with epilepsy who are evaluated for resective surgery. Whereas the above-mentioned archives integrate the data within the concept of a probabilistic atlas of the human brain, a major consideration of our database design was the analysis and integration of data in native space. As a consequence, our approach not only allows the integration of complementing modalities (PET, MRI, EEG) in native space of each individual patient, but also provides software tools able to quantify the relationship between these diverse modalities. Consequently, the developed framework has the potential to generate quantitative measures which characterize the state of the brain in more detail than would be possible with each individual modality.

3. INTEGRATION OF MULTIMODALITY DATA IN THE PRESURGICAL EVALUATION OF PATIENTS WITH EPILEPSY

The main objective of presurgical evaluation of patients with medically refractory epilepsy is to define the boundaries of epileptogenic brain regions to be resected. Towards this goal, definition of the epileptogenic cortex by intracranial subdural EEG recording remains the gold standard. However, the

accuracy of foci localization using subdural electrodes depends greatly on location of electrodes placed on the brain surface, and selection bias (i.e., area of cortex sampled) is a major limitation. In order to guide the placement of subdural electrodes, a combination of noninvasive anatomical and functional imaging, such as MR and PET imaging, is frequently used. These modalities provide the epilepsy surgery team with important information to guide placement of subdural electrodes over epileptogenic brain regions based on the position and extent of anatomical and functional abnormalities [18–20] as well as the seizure semiology (i.e., symptoms of seizures). Furthermore, after the intracranial EEG electrodes are implanted, each electrode records a number of electrophysiological parameters that need to be related to the imaging information in order to decide on the most effective course of surgical intervention. To this end, we designed an extendable multimodality database which allows assessment of relationships among various modalities across a large patient population. A focal point of our design was the extensibility of the database scheme in order to allow easy inclusion of data originating from newly developed PET tracers or the inclusion of emerging modalities such as diffusion tensor imaging (to incorporate a measure of brain connectivity) or susceptibility weighted imaging (to assess venous vasculature in brain tissue).

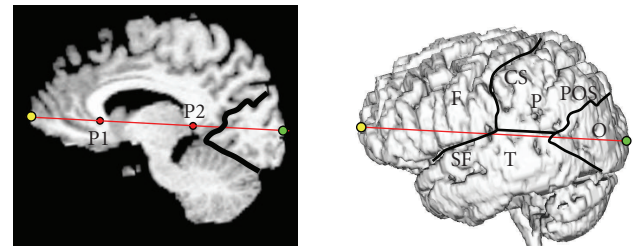
3.1. Data acquisition protocols

PET imaging using the tracers [F-18]deoxy-L-glucose (FDG) and [C-11]flumazenil (FMZ) was performed as part of the clinical management of patients undergoing evaluations for epilepsy surgery. PET studies were performed using the CTI/Siemens EXACT/HR scanner (Knoxville, Tenn, USA) with a reconstructed image resolution of about 5 mm FWHM. All subjects was fasted for 4 hours prior to the PET procedure and EEG was monitored using scalp electrodes during the whole study duration. Static PET images were obtained based on coincidence data acquired between 40–60 minutes post injection for the FDG and between 20–40 minutes post injection for the FMZ tracer.

MRI studies were performed on a GE 1.5 Tesla Signa unit (Milwaukee, Wis, USA). High-resolution T1-weighted images were performed using a fast spoiled gradient echo (SPGR) sequence. The SPGR technique generates 124 contiguous 1.5 mm sections of the entire head using a 4.6/1.3/450 (TR/TE/TI) pulse sequence, flip angle of 12 degrees, matrix size of 256×256, and FOV of 240 mm.

3.2. Image data processing

Following coregistration of the PET image volumes to the high-resolution MR image volume, all extracerebral structures were removed from the MR image volume using an in-house developed software package [3]. Following initial thresholding, the software performs multiple erosion and dilation operations in order to delete connections between the brain and the skull in subdural space. Finally, a connected component analysis is performed and all but the largest component (brain) is deleted, resulting in an MR image volume where all extracerebral structures are removed and cortical sulci and gyri become visible. The standard set of cortical landmarks consisted of the central sulcus, Sylvian fissure, and parieto-occipital sulcus, which were subsequently used to define the spatial extent of the four brain lobes (frontal, parietal, temporal, and occipital). In order to find the border of the occipital cortex on the lateral cortical surface, both the parieto-occipital sulcus as well as the medial border between the cerebellum and the occipital lobe were defined on the midplane and then projected normal to the midplane onto the lateral surface (Figure 1).



(a) (b)

Figure 1: Definition of the four brain lobes (F,P,T,O) on the lateral cortical surface based on user-defined landmarks such as the central sulcus (CS), Sylvian fissure (SF), and the parieto-occipital sulcus (POS). The locations of the frontal (yellow) and occipital (green) poles are obtained as the extension of the AC-PC line to the front and back of the brain. The horns of the corpus callosum (P1 and P2) as well as the POS are defined in the medial plane.

3.3. Quantitative assessment of bilateral PET abnormalities

We have recently developed a landmark-constrained conformal surface mapping technique [12, 21] which allows accurate and reproducible transformation of each patient's cortical surface to a canonical spherical domain (the conformal brain model (CBM), see the appendix). Once the cortical surface of a subject is mapped to the CBM, finite surface elements can be defined geometrically on the spherical surface at various resolution levels (8, 32, 128, 512, and 2048) (Figure 2).

The set of surface elements can be subsequently reversely mapped into native space where they represent homotopic surface elements in brains of individual patients. These surface elements scale proportionally to the size and shape of individual brains. Moreover, in order to integrate PET and MRI data, a normal fusion approach is applied in native space of each subject. The extracted MR brain surface is smoothed using a "roller-ball algorithm" based on alpha-shapes [22] where the parameter alpha is the radius of a ball rolled over the surface. Smoothing of the cortical surface is achieved as all points deep inside cortical sulci are recovered which

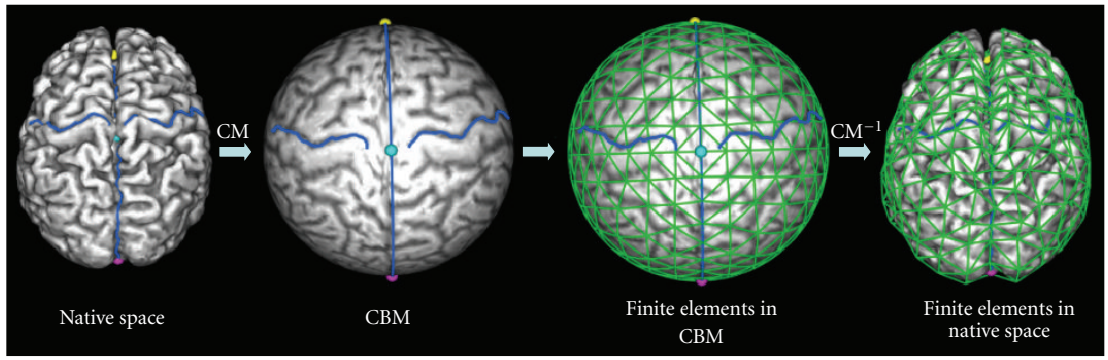


Figure 2: Landmark-constrained conformal surface mapping of the cortex onto a sphere and construction of finite surface elements. (a) Cortical surface of a brain with defined landmarks showing the bilateral central sulci. (b) Landmark-constrained conformal mapping (CM) of the cortical mantel to the surface of a sphere. Constraints are enforced during conformal mapping so that the midplane is mapped to the main circle of the unit sphere, and landmarks (e.g., central sulcus) are mapped to the same spatial location. (c) Geometric parcellation of the unit sphere into 512 surface finite elements. (d) Location of the same surface finite elements after reverse conformal mapping (CM^{-1}) into the subject's native space.

cannot be reached by the ball. Following triangulation of the brain surface, the normal vector to the surface is calculated in each surface voxel. By averaging the PET tracer concentration along the inverse normal vector in the coregistered PET image volume, the average PET tracer concentration within a 10 mm cortical mantel is calculated.

By comparing the PET tracer concentration in these cortical elements between a group of normal control subjects and an individual patient, functionally abnormal increases or decreases of PET tracer concentration can be determined in finite elements of the cortical mantel. Decision about the functional abnormality of a cortical element of a patient at a given location can be then made based on a severity index calculated as

$$\text{severity}_i = \frac{(m_i - \mu_i)}{\sigma_i}, \tag{1}$$

where m_i is the mean tracer concentration at cortical location i of a patient and μ_i and σ_i are the mean tracer concentration and standard deviation (SD) at cortical location i derived from a control group. Values of the severity index within ± 2 SD are assumed to represent normal cortex. The severity of detected abnormalities (outside ± 2 SD) is then color-coded and mapped onto the cortical surface allowing assessment of functional abnormalities relative to anatomical cortical landmarks.

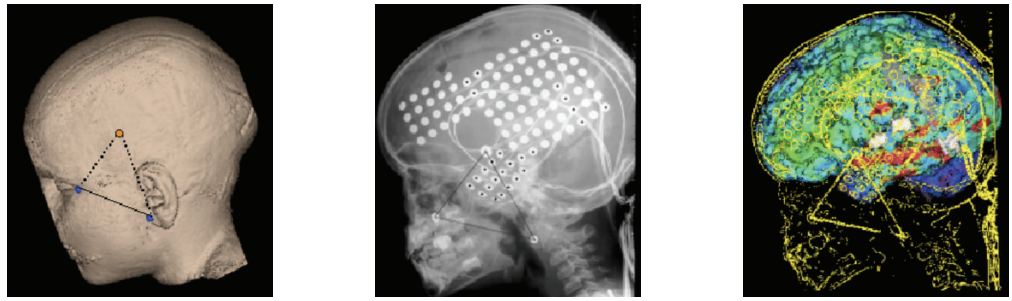
Our approach attempts to merge the advantages of voxel-based (such as SPM) and region-of-interest-based strategies, but at the same time tries to avoid the pitfalls associated with either of these methods. The semiautomated geometric parcellation procedure creates cortical elements that are much larger than individual voxels of the image volume, but circumvents the time consuming and subjective definition of large surface-based regions of interest. Moreover, as this method does not rely on asymmetry measures between homotopic cortical volume elements [23], it is well suited to detect bilateral cortical abnormalities. Finally, the most prominent advantage of our method is the fact that all data samplings and analyses are performed in the subject's native space. By avoiding spatial warping, we allow application of this method to brains that are very different from the normal adult brain, such as brains of patients with tuberous sclerosis (with a large number of tubers deep inside the brain) or children during various developmental stages.

3.4. Subdural EEG assessment

All patients with intractable focal epilepsy included in the present study underwent chronic EEG monitoring with subdural electrode grids as part of their presurgical evaluation. Subdural electrode placement was guided by seizure semiology, scalp EEG recordings, and cortical glucose metabolism abnormalities on PET. During the chronic subdural EEG monitoring, more than two habitual seizures were captured and analyzed using Stellate's SENSA 5.0 software [24], yielding for each grid electrode the mean interictal spike frequency as well as the normalized interictal spike frequency (normalized to the electrode with highest spike frequency). Identification of electrodes involved in seizure onset and seizure spread of habitual clinical seizures was determined by chronic subdural EEG monitoring (see, e.g., [8]).

3.5. Localization of electrodes on the cortical surface

In the past, we have implemented a method that allows the localization of subdural grid electrodes on the cortical surface [25, 26]. In short, this method relies on the accurate alignment of a lateral planar X-ray image (Figure 3(b)) of the patient's head with the electrode grid in place with a three-dimensional surface rendering of the head (Figure 3(a)). Upon completion, a surface view of the cortex is created which corresponds to the planar X-ray image and where the



(a) (b) (c)

Figure 3: (a) Anatomical information is used to define virtual markers corresponding to (b) fiducial markers placed on the patients's head prior to X-ray imaging. (c) The result is a surface view where the location of the four corner electrodes of each of the three EEG grids can be determined on the brain surface.

spatial coordinates of the four corner grid electrodes of each rectangular EEG grid can be determined on the brain surface (Figure 3(c)). The accuracy of this method was reported to be 1.24 ± 0.66 mm with a maximal misregistration of 2.7 mm [25].

3.6. Assessment of the spatial relationship between functional and electrophysiological data

Our previous results indicated that intracranial subdural electrodes showing electrophysiological ictal (seizure onset and spread) or interictal (spike frequency) abnormalities are either overlapping or in close proximity to functional abnormalities measured with PET imaging [27]. Because classical receiver operator characteristic (ROC) analysis is suboptimal in describing spatially related measures, we have previously developed a spatial proximity index (SPI) [28]. This index yields an overall quantitative measure of the spatial relationship between electrophysiological and functional PET image data and is calculated based on the position of intracranial grid electrodes relative to abnormal PET tracer concentration in cortical areas (Figure 4). The SPI is a continuous variable that equals zero for perfect overlap between seizure onset electrodes and PET-defined abnormal cortical elements and increases in value proportional to the distance between EEG-defined onset electrodes and PET-defined abnormal cortical elements.

The subdural EEG electrode array defines electrodes that are either EEG positive ($E+$) or negative ($E-$) for seizure onset. Moreover, electrodes located within the PET abnormality are designated as PET positive ($P+$) and those outside the PET abnormality are designated as PET negative ($P-$). Using these definitions, the (unitless) spatial proximity index (SPI) is computed as the ratio of the penalized total weighted distance between EEG positive and PET positive electrodes and the total number of seizure onset electrodes,

$$\mathrm{SPI} = \frac{\sum_{i=1}^{M} w_i\, d_i(E_i + \mid P_i +) + \sum_{i=1}^{K} P_i + (\neg E_i +)}{\sum_{i=1}^{M} (E_i +)}, \qquad (2)$$

where M is the number of EEG positive electrodes, K is the number of PET positive electrodes, and w_i is a weighting factor which accounts for the varying interictal spike frequency at different electrode locations. The weighting factor is either 1 or 0 for both seizure onset and spread, whereas for spike frequency the weighting factor is derived as the quotient between the spike frequency at a particular electrode location and the maximal spike frequency in the whole brain. In (2), the first term in the numerator represents the total weighted distance between all seizure onset electrodes and the nearest PET positive electrode, whereas the second term represents the number of all false positive PET electrodes (i.e., PET positive electrodes which are not EEG positive). Finally the denominator reflects the total number of EEG positive electrodes. Because subdural electrodes are always arranged in a rectangular lattice, the "city-block" metric [29] is used to calculate the distance between two electrodes. It is important to note that all characteristics of the Euclidian metric are preserved in the city-block metric. In this metric, the distance between two adjacent electrodes is 1 and the distance between two diagonal electrodes is 2. SPI values were calculated separately for seizure onset ($\mathrm{SPI}_{\mathrm{onset}}$), seizure spread ($\mathrm{SPI}_{\mathrm{spread}}$), and interictal spike frequency ($\mathrm{SPI}_{\mathrm{spike}}$).

In order to demonstrate the computation of SPI values, Figure 4 shows a (5×4) electrode grid with seizure onset electrodes (red) and electrodes which are not seizure onset electrodes (yellow). In Figure 4(a), the PET abnormality is adjacent to the seizure onset electrodes resulting in an $\mathrm{SPI}_{\mathrm{onset}}$ value of 2.66, while in Figure 4(b) the PET abnormality is remote to the seizure onset electrodes resulting in an $\mathrm{SPI}_{\mathrm{onset}}$ value of 5.66. In this example, the lower SPI value indicates a closer spatial proximity between (complementing) PET and EEG modalities and indicates that low SPI values are associated with situations in which PET imaging is successfully guiding the placement of an intracranial EEG grid (despite being nonoverlapping).

4. DATABASE DESIGN AND QUANTITATIVE ANALYSIS

4.1. Database design

Our database design is founded on the relational model depicted in Figure 5. Conceptually, the database is structured in

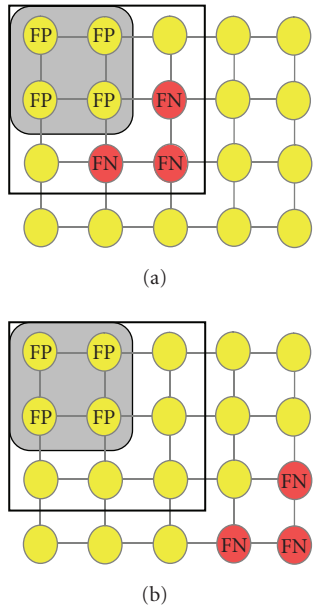


Figure 4: Schematic representation of an intracranial EEG grid placed on the surface of the brain. Electrophysiologically normal electrodes are yellow, whereas electrodes determined as abnormal (either onset, spread, or frequent spiking) are shown in red. The area with abnormal PET tracer concentration is depicted in grey. Electrodes overlaying the PET abnormality but which are found to be electrophysiologically normal represent false positive (FP) cases, whereas electrophysiologically abnormal electrodes not overlaying the PET abnormality represent false negative (FN) cases. (a) A PET abnormality in close spatial proximity to electrophysiologically abnormal electrodes prompts the insertion of an intracranial grid (black square) which will include the electrophysiologically abnormal area, whereas (b) a PET abnormality that is distant to the electrophysiologically abnormal area will trigger the insertion of intracranial grid electrodes likely to miss the electrophysiologically abnormal area.

three hierarchical levels: a patient level, a cluster level, and a cortical element level. At the highest level (patient level), clinical variables are stored such as the age of the patient, age at first seizure, seizure type, seizure severity, and seizure duration. An index of seizure severity was derived as the product of average monthly seizure number and average seizure duration. As the cortical distribution of each PET tracer (e.g., FDG, FMZ, etc.) is distinct, the database includes separate row for different PET tracers (Exam_Type in PET_Exam table). The lowest level is the cortical element level, which provides information with regard to the location of functional data in cortical elements. Cortical elements at different resolution levels are represented by the coordinates of the three vertices building the corners of a surface triangle (Corner_A, Corner_B, Corner_C in the Cortical_Element table). In addition, the PET table stores for each cortical element the Severity (Severity_index), and the SPI table stores SPI values for seizure onset, seizure spread, and interictal spiking for the whole brain at each resolution level. Moreover, the surface coordinates of each EEG grid electrode are stored at this lowest level in the EEG_grid table. This table stores for each of the potentially multiple electrode grids placed onto the cortex the location of the corner electrodes (Corner_I-Corner_IV in the EEG_grid table). The location of all other electrodes within the EEG grid can be then computed from the location of the corner electrodes given a 10 mm spacing between individual electrodes. Electrophysiological parameter such as the seizure category (onset, spread, spiking), interictal spike frequency, or normalized interictal spike frequency can be then mapped to a particular grid electrode. Given the location of each electrode, the software then automatically determines the corresponding cortical element at various resolution levels. Finally the intermediate cluster level (Cluster table) allows the grouping of cortical volume elements either into anatomical territories such as the prefrontal, motor-frontal, parietal, temporal, and occipital cortices, or into functional clusters according to PET or EEG data. To accommodate the flexibility of an evolving database schema, we applied an XML-based approach. In this way, the database allows graceful inclusion of additional modalities at each of the three levels. The XML-based approach does not only accommodate the incorporation of additional data structures into the database, but ensures also the scalibility of the system. This was achieved through following design criteria: (1) an efficient mapping algorithm that generates corresponding relational schemas from XML schemas in linear time [29], (2) the implementation of two efficient linear data mapping algorithms in order to store XML data in the database [30], and (3) the implementation of a linear XML subtree reconstruction algorithm able to reconstruct the XML subtree from the database [29].

4.2. Database initialization

The database was initially populated with control data determined in control subjects in order to establish a normative pattern. The control group consisted of 15 young adult controls (mean age 27.6 ± 4.5 years) who were not taking any medication, and had no history of neurological or psychiatric disorder. All adult controls had normal MRI scans.

4.3. Queries for quantitative analyses

The purpose of the database is to provide an integrated framework able to present relationships between functional and electrophysiological parameters in an organized and, for the user, easy-to-navigate fashion. A focus of our design was to merge an efficient database structure with a display module, so that users can grasp the results of their queries in an intuitive way. Figure 6 shows the main control panel allowing access to all data structures within the database and the link to an integrated 3D image display module.

In addition, the application provides various queries and advanced display options. For example, the user can request to "highlight" all EEG grid electrodes with spike frequency greater than 20% of the maximal spike frequency and to calculate the corresponding SPI values. This is highly relevant in light of our previous reports [19]

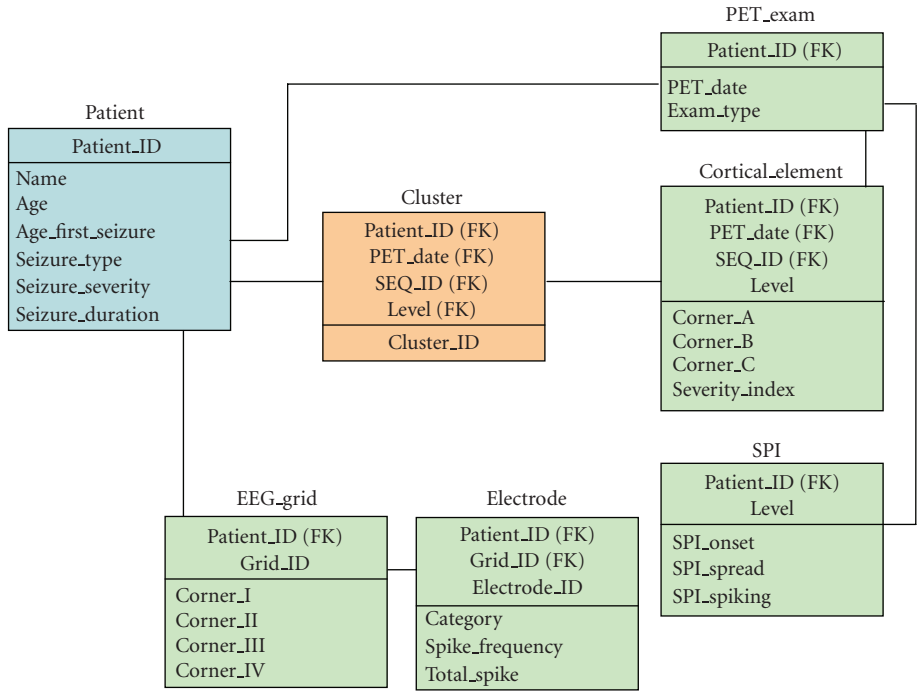


Figure 5: Relational model design of the database. On a conceptual level, the database is structured in three hierarchical levels. At the lowest level (green), PET and EEG information are stored, which is associated with distinct finite cortical elements. At the intermediate level (orange), functionally or electrophysiologically similar cortical elements are grouped into clusters. Finally, all clinical data characterizing the disease state are stored at the highest level (blue). For simplification of presentation, each three-dimensional coordinate is recorded as one attribute in tables, like Corner_A in table EEG-grid.

indicating that in patients with nonlesional neocortical epilepsy, seizures arose from areas adjacent to PET-defined abnormalities, rather than from within the PET-defined abnormalities. An SQL query answering the above request is Algorithm 1.

In essence, this query retrieves the electrophysiological data from table Electrode and the PET data from table Cluster, respectively, based on the input patient ID($PID). After the location of an electrode with (normalized) spike frequencies > 0.20 is determined, data representing the current electrode is integrated with data characterizing the distribution of abnormal PET clusters using the "join" condition Adjacent($PID, E.electrode, C.SEQ_ID). This expression will return the value "true" only in the case when the current electrode is adjacent to an abnormal cortical region. In addition, our design supports advanced graphical structure mining techniques in order to discover nontrivial associations between multiple electrophysiological parameters and their spatial relationship to the PET-defined abnormalities. One of these data mining techniques, the "association rule mining," is used to determine whether certain clinical patterns result in larger SPI values or whether different PET tracers produce larger SPI values given the same clinical pattern. This information is important in order to determine the clinical performance of newly developed PET tracers. Two measurements that are used in the association mining process are support and confidence defined as follows:

(i) support(condition c)
= # of patients satisfying c/total # of patients,
(ii) confidence(condition c1, condition c2)
= support(c1 & c2)/support(c1).

The association rule mining procedure returns all rules in the form of condition 1 → condition 2, with a support and a confidence greater than some user-specified thresholds. The two measurements in conjunction quantify the association strength between condition 1 and condition 2. One example of an association rule is

$$\begin{aligned}&(\#\text{ of patients }(\text{SeizureDuration } > 3 \text{ .AND. Seizure Severity} \\ &\qquad > 2 \text{ .AND. } \text{SPI}_{\text{onset}} < 1.5))/ \\ &(\#\text{ of patients }(\text{SeizureDuration} \\ &\qquad > 3 \text{ .AND. Seizure Severity } > 2)) > 0.8.\end{aligned} \tag{3}$$

This indicates that if a patient has had seizures for more than 3 years and the severity score of these seizures was on average higher than 2, then there is greater than 80% likelihood that the patient's $\text{SPI}_{\text{onset}}$ will be greater than 1.5.

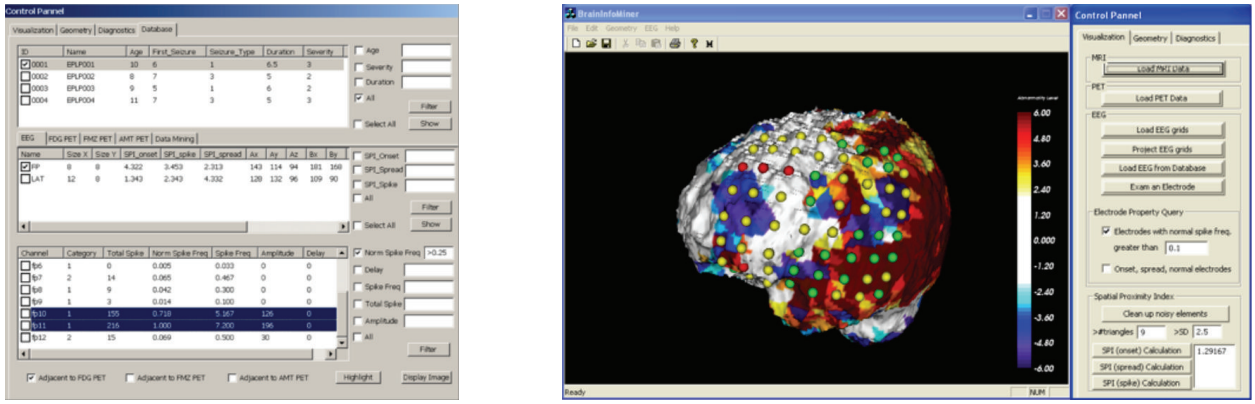


Figure 6: Main control panel which allows access to clinical, electrophysiological, and image data. The imaging module allows coding of EEG parameters (red = seizure onset, yellow = seizure spread, green = normal) and evaluation of each electrode with respect to its location in space by allowing real-time rotation and zoom of all objects in 3D. Furthermore, the SPI parameter which quantifies the spatial relationship between PET and EEG information can be saved in the database for further evaluation across patients.

```
SELECT    E.Electrode_ID
FROM      Electrode E, Cluster C
WHERE E.Patient_ID = $PID AND C.Patient_ID
                   = $PID AND E.Spike_frequency
                   > 0.20
AND       Adjacent($PID, E.Electrode, C.SEQ_ID).
```

Algorithm 1

5. IMPLEMENTATION AND RESULTS

We have utilized VTK/OpenGL for rendering, Oracle 9i for database applications, and C/C++ for implementation of computationally intensive algorithms. Furthermore, we separated the computational components from other functionality and created a standalone programming library consisting of the following modules: (1) conformal mapping; (2) objective definition of PET abnormalities; (3) EEG grid overlay; (4) calculation of SPI values; (5) extendable database component; and (6) graphical output module.

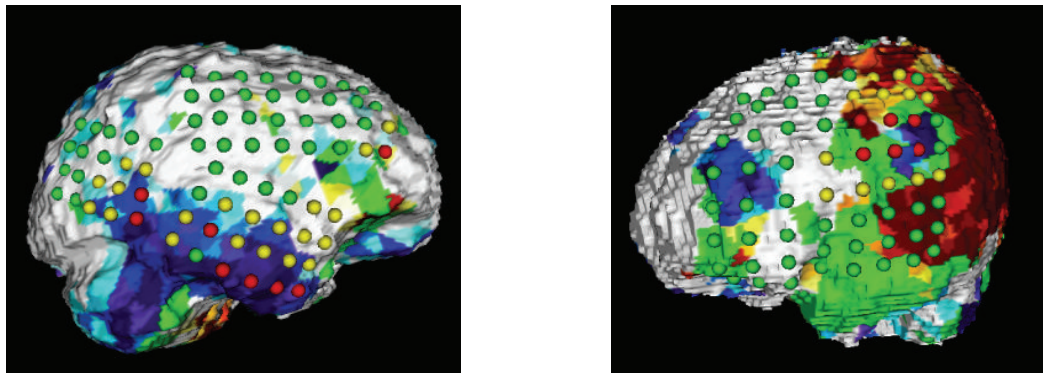
With the expected growth of knowledge about the evolution of epileptic foci as well as the emergence of new computational methods that require diverse information inputs, it is impractical to design a fixed database schema. As new requirements arise, it is critical that the database schema adjusts gracefully to these new requirements. In contrast to traditional database systems, where evolution of the database schema to new requirements is problematic for dynamic applications, we applied a flexible data model based on XML (eXtensible Markup Language) [29, 31, 32] due to its extensibility and flexibility nature. While in the past we have successfully developed an XML storage and query system, in which the database schema was automatically created from an XML schema [30], this technique needed to be extended to include support for the processing and storage of spatial information.

5.1. Patient population

Twelve children (mean age 5.2 ± 4.3 years, age range 1 to 14.8 years) with medically intractable partial epilepsy were analyzed. Patients were selected according to the following criteria. All children were diagnosed with unilateral seizure foci based on seizure semiology, scalp ictal, and intracranial EEG as well as FDG PET, which were performed as part of their presurgical evaluation. The region of epileptic focus as verified by EEG showed hypometabolism in all cases. No cortical or subcortical lesions on MRI scans were observed; however, patients with pure hippocampal atrophy were included. All studies were performed in accordance with guidelines stipulated by the Ethics Committee of Wayne State University.

5.2. Application number 1: multimodality display

Our application provides integration of anatomical, functional, and electrophysiological data as shown in Figure 7. Upon retrieval from the database, the software displays a surface rendering of the selected patient's brain with functional data derived from PET imaging mapped directly onto the cortical surface. In addition, each EEG electrode is color coded according to its electrophysiological characteristics (red = seizure onset, yellow = seizure spread, green = normal) and is displayed directly on the cortical surface. As the display is fully three-dimensional, the user can interactively rotate and zoom the brain in order to inspect the spatial relationship between EEG electrodes and PET abnormalities from various viewing angles.



(a) (b)

Figure 7: (a) Surface rendering of a patient's brain with an epileptic focus in the right temporal lobe. Finite cortical elements which represent abnormally decreased FDG PET tracer uptake are rendered as blue areas with dark blue representing the most severe decrease and with light blue representing the least severe decrease. The location of FDG PET abnormalities can be assessed with respect to the location of seizure onset electrodes (red) and seizure spread electrodes (yellow). The SPIonset for this patient was calculated as 2.75. (b) Surface rendering of a patient with tuberous sclerosis who had cortical tubers in the frontal and parietal lobes. Both tubers are associated with abnormally decreased FDG tracer uptake (blue areas). Furthermore, yellow, orange, and red areas represent various degrees of abnormally increased FDG tracer uptake, whereas green areas represent areas with abnormally homogeneous FDG tracer uptake and may indicate dysplastic tissue. Onset electrodes (red) were located near the tuber in the parietal lobe and this area was determined as epileptogenic. The relevance of abnormally increased FDG tracer uptake adjacent to epileptogenic tubers is at present unclear and warrants future studies.

5.3. *Application number 2: quantitative assessment of the spatial relationship between PET and EEG abnormalities*

In addition to the multimodality display, our software also allows calculation of SPI values in order to allow quantitative comparison between localizing information based on clinical readings (onset/spread) and localizing information obtained based on a semiquantitative analysis of interictal spike frequency. Figure 8 shows images determined in a patient studied with both FDG as well as FMZ PET, and reports SPI values which were calculated with regard to seizure onset electrodes and electrodes with high interictal spike frequency (>40% of maximal spike frequency). The user is able to select various thresholds of spike frequency and interactively assess the spatial relationship between frequently spiking cortical tissue, the location of seizure onset, and areas of functionally abnormal cortex. The spatial relationship between these areas is then encoded in SPI values and stored for further statistical assessment.

Usually, EEG abnormalities are located adjacent to extensive PET abnormalities and might be reversible after surgical intervention. In addition, remote areas of PET abnormality might exist, which possibly indicate secondary epileptic foci initially triggered by the primary focus but which might mature with time and become independent. This phenomenon is at present time poorly understood and possible mechanisms are discussed elsewhere [27]. There are strong indications, however, that the relationship between electrophysiology and molecular function is complimentary and that the EEG and PET modalities characterize different aspects of brain tissue epileptogenicity.

6. DISCUSSION

With ever-improving imaging technologies and boost in computational power, the medical imaging field has experienced a tremendous increase in the amount of information collected. Although medical imaging modalities provide diverse quantitative and qualitative information, each modality has its distinct strengths and limitations. While T1-weighted MR images yield high-resolution anatomical images, the obtained functional information is limited. In contrast, molecular imaging using a variety of PET tracers provides accurate functional information, unfortunately with significantly less anatomical detail as compared to MR imaging. Nevertheless, both modalities provide invaluable clinical information to the physicians even when utilized qualitatively. In the past, our group has successfully used data from PET, MR and EEG in the presurgical evaluation of epileptic children and showed a significant improvement in surgical outcome when using such multimodality approach [33, 34]. It appears that further improvement in the localization of epileptogenic brain tissue might be derived from a tighter integration of all presurgical data (anatomical, functional, electrophysiological, and clinical) within a comprehensive database. Once data from multiple imaging and nonimaging modalities are combined within a rigorous computational framework, the information entailed in one modality may be used to enhance or reinterpret information derived from a complementing modality. Thus, the information content in such a database is not simply additive, but is likely to have an amplifying effect. Large-scale database structures containing various brain disorders can be then constructed with a common frame of reference, allowing meta-analysis of data patterns distributed

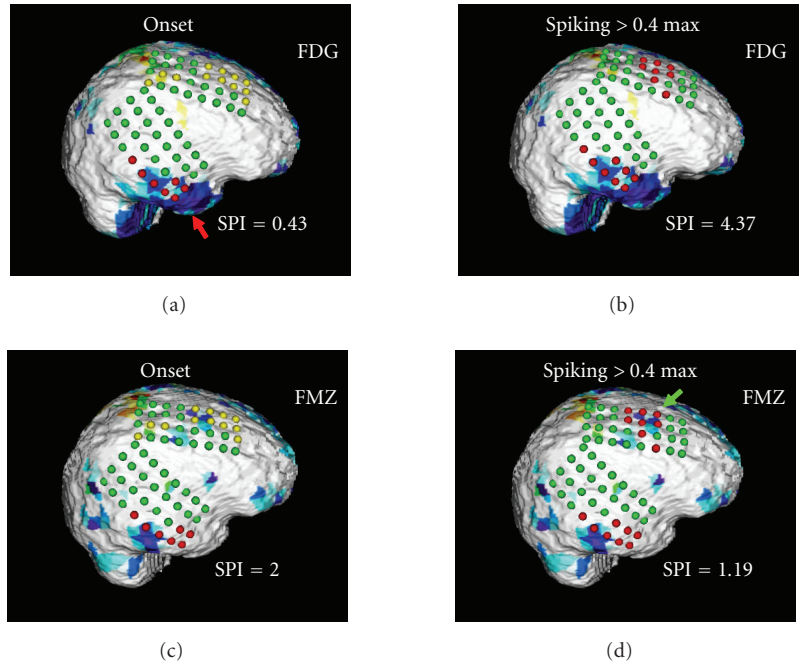


Figure 8: Cortical display of a patient studied with both FDG PET (panels (a) and (b)) and FMZ PET (panels (c) and (d)). Images on the left (panels (a) and (c)) show coding of the EEG grid according to the location of seizure onset electrodes (red) and seizure spread electrodes (yellow). The images indicate that the FDG PET abnormality corresponding to seizure onset electrodes in the temporal lobe (red arrow) is larger than the FMZ PET abnormality. Moreover, in the case of FDG PET, the seizure onset electrodes overlay a larger functionally abnormal area as compared to FMZ PET which is reflected in the lower SPI score for FDG (0.43 versus 2). The panels on the right (panels (b) and (d)) present the coding of the EEG grid according to a spike frequency threshold of 40% of the maximal spike frequency observed in this patient. Whereas FMZ PET is sensitive to the area of high spike frequency in the frontal lobe (green arrow), FDG PET fails to detect this abnormality. Again, this spatial relationship between functional and electrophysiological abnormalities is encoded in the SPI score (1.19 versus 4.37).

over several modalities. Obviously, such analyses require a well-developed visual interface which allows the researchers to grasp the relationship between data patterns in an intuitive way. We believe that such a "visual database" has the potential to provide physicians with additional important clues regarding the pathological state of cortical areas and is likely to be of relevance for the clinical management of epilepsy patients.

6.1. Methodological consideration

As with every method, a few limitations need to be considered when applying this method to images of children with intractable epilepsy. As indicated above, we used young adults to create a normal database, and this database was implicitly assumed to characterize accurately normal tracer concentration patterns in children. This approach is inevitable due to ethical guidelines which sanction only the study of children who may derive direct benefit from studies using administration of radioactive substances. Although the application of adult tracer concentration patterns to children represents a potential source of error, it appears that at least for FDG the tracer concentration patterns might be very similar between children and adults [35].

Secondly, integration of brain PET and MR image volumes requires their spatial alignment, a potential source of error. Because the brain is encased by the skull and can be therefore regarded as a rigid body, alignment between the MR and PET brain image volumes is likely to be accurate within a few millimeters, especially if anatomical landmarks inside the brain are used. Moreover, the localization of the electrode grid on the cortical surface might result in a maximal misregistration of about 3 mm [25]. Thus, despite the fact that misalignment effects associated with various alignment procedures might be additive, their total effect is likely to be less than 5 mm.

Finally, the initial processing step of our method requires the manual definition of cortical landmarks, guided by a detailed surface rendering of cortical sulci and gyri. Although high-resolution MR images provide detailed features of the cortex, reproducible definition of cortical landmarks is not trivial and might contribute to the misalignment of presumably homotopic cortical elements in different patients. In order to keep this source of error to a minimum, we chose only anatomically well-defined cortical features for our standard landmark set. Nevertheless, the definition of cortical landmarks contributes the largest error to possible misalignment between homotopic cortical elements

and future studies are warranted to assess their clinical relevance.

6.2. Conclusion

The creation of a multimodality extendable database structure accessible through an advanced 3D visualization interface holds promise of providing new insights into the formation and identification of epileptic foci. A better understanding of the relationship between functional and electrophysiological data might lead to new approaches in epilepsy surgery and will likely improve clinical management of a large number of patients suffering from intractable epilepsy.

APPENDIX

Because the brain surface is a genus zero surface and topologically equivalent to a sphere, we have previously developed a landmark-constrained conformal surface mapping technique [12, 21] that allows accurate and reproducible transformation of a subject's cortical surface to a canonical spherical domain (the conformal brain model (CBM)). The conformal mapping technique allows parametrization of a brain surface without angular distortion and is computed by minimizing the harmonic energy of the map. As an example, given two genus zero surfaces with maps M1 and M2, the transformation (f: M1→M2) is conformal if, and only if, f is harmonic. Based on this fact, one can compute the conformal mapping between genus zero surface by minimizing the harmonic energy [12]. In practice, we use a triangular mesh in order to approximate genus zero cortical surfaces. Once the cortical surface of a subject is mapped to the CBM, finite surface elements can be defined geometrically on the spherical surface at various resolution levels (see Figure 2).

ACKNOWLEDGMENTS

The authors thank Joel Ager and James Janisse from the Center for Health Care Effectiveness Research at Wayne State University for their invaluable help with statistical assessment of the data, as well as the PET Center staff at Children's Hospital of Michigan for their assistance in performing the studies. This work was partially supported by funding from Grants NIBIB R21DA015919, MTTC05-135/GR686 (O. Muzik), MTTC05-154/GR705 (J. Hua), NIH R01 NS 45151 (D. C. Chugani), and NINDS RO1 NS 34488 (H. T. Chugani).

REFERENCES

[1] P. M. Thompson and A. W. Toga, "A framework for computational anatomy," *Computing and Visualization in Science*, vol. 5, no. 1, pp. 13–34, 2002.

[2] P. J. Basser and C. Pierpaoli, "A simplified method to measure the diffusion tensor from seven MR images," *Magnetic Resonance in Medicine*, vol. 39, no. 6, pp. 928–934, 1998.

[3] O. Muzik, D. C. Chugani, C. Shen, et al., "Objective method for localization of cortical asymmetries using positron emission tomography to aid surgical resection of epileptic foci," *Computer Aided Surgery*, vol. 3, no. 2, pp. 74–82, 1998.

[4] M. Miller, A. Banerjee, G. Christensen, et al., "Statistical methods in computational anatomy," *Statistical Methods in Medical Research*, vol. 6, no. 3, pp. 267–299, 1997.

[5] P. M. Thompson, C. Schwartz, and A. W. Toga, "High-resolution random mesh algorithms for creating a probabilistic 3D surface atlas of the human brain," *NeuroImage*, vol. 3, no. 1, pp. 19–34, 1996.

[6] G. Lohmann, "Extracting line representations of sulcal and gyral patterns in MR images of the human brain," *IEEE Transactions on Medical Imaging*, vol. 17, no. 6, pp. 1040–1048, 1998.

[7] M. Vaillant, C. Davatzikos, and N. R. Bryan, "Finding 3D parametric representations of the deep cortical folds," in *Proceedings of the Workship on Mathematical Methods in Biomedical Image Analysis*, pp. 151–159, San Francisco, Calif, USA, June 1996.

[8] S. Minoshima, K. A. Frey, R. A. Koeppe, N. L. Foster, and D. E. Kuhl, "A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET," *Journal of Nuclear Medicine*, vol. 36, no. 7, pp. 1238–1248, 1995.

[9] H. A. Drury, D. C. Van Essen, M. Corbetta, and A. Z. Snyder, "Surface-based analyses of the human cerebral cortex," in *Brain Warping*, pp. 337–363, Academic Press, New York, NY, USA, 1999.

[10] M. K. Hurdal and K. Stephenson, "Cortical cartography using the discrete conformal approach of circle packings," *NeuroImage*, vol. 23, supplement 1, pp. S119–S128, 2004.

[11] X. Gu, Y. Wang, T. F. Chan, P. M. Thompson, and S.-T. Yau, "Genus zero surface conformal mapping and its application to brain surface mapping," *IEEE Transactions on Medical Imaging*, vol. 23, no. 8, pp. 949–958, 2004.

[12] G. Zou, J. Hua, X. Gu, and O. Muzik, "An approach for intersubject analysis of 3D brain images based on conformal geometry," in *Proceedings of International Conference on Image Processing (ICIP '06)*, pp. 1193–1196, Atlanta, Ga, USA, October 2006.

[13] J. D. Van Horn, J. S. Grethe, P. Kostelec, et al., "The Functional Magnetic Resonance Imaging Data Center (fMRIDC): the challenges and rewards of large-scale databasing of neuroimaging studies," *Philosophical Transactions of the Royal Society of London Series B Biological Sciences*, vol. 356, no. 1412, pp. 1323–1339, 2001.

[14] J. D. Van Horn, S. T. Grafton, D. Rockmore, and M. S. Gazzaniga, "Sharing neuroimaging studies of human cognition," *Nature Neuroscience*, vol. 7, no. 5, pp. 473–481, 2004.

[15] M. Barinaga, "Still debated, brain image archives are catching on," *Science*, vol. 300, no. 5616, pp. 43–45, 2003.

[16] A. W. Toga, "Neuroimage databases: the good, the bad and the ugly," *Nature Reviews Neuroscience*, vol. 3, no. 4, pp. 302–309, 2002.

[17] P. T. Fox and J. L. Lancaster, "Mapping context and content: the BrainMap model," *Nature Reviews Neuroscience*, vol. 3, no. 4, pp. 319–321, 2002.

[18] J. Engel Jr., T. R. Henry, M. W. Risinger, et al., "Presurgical evaluation for partial epilepsy: relative contributions of chronic depth-electrode recordings versus FDG-PET and scalp-sphenoidal ictal EEG," *Neurology*, vol. 40, no. 11, pp. 1670–1677, 1990.

[19] C. Juhász, D. C. Chugani, O. Muzik, et al., "Relationship between EEG and positron emission tomography abnormalities in clinical epilepsy," *Journal of Clinical Neurophysiology*, vol. 17, no. 1, pp. 29–42, 2000.

[20] E. L. So, "Role of neuroimaging in the management of seizure disorders," *Mayo Clinic Proceedings*, vol. 77, no. 11, pp. 1251–1264, 2002.

[21] G. Zou, Y. Xi, G. Heckenberg, Y. Duan, J. Hua, and X. Gu, "Integrated modeling of PET and DTI information based on conformal brain mapping," in *Medical Imaging: Physiology, Function, and Structure from Medical Images*, vol. 6143 of *Proceedings of SPIE*, pp. 9 pages, San Diego, Calif, USA, February 2006.

[22] H. Edelsbrunner and E. P. Mucke, "Three-dimensional alpha shapes," *ACM Transactions on Graphics*, vol. 13, no. 1, pp. 43–72, 1994.

[23] O. Muzik, E. A. Da Silva, C. Juhász, et al., "Intracranial EEG versus flumazenil and glucose PET in children with extratemporal lobe epilepsy," *Neurology*, vol. 54, no. 1, pp. 171–179, 2000.

[24] J. Gotman and P. Gloor, "Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 41, no. 5, pp. 513–529, 1976.

[25] A. Thiel, K. Herholz, H.-M. Von Stockhausen, et al., "Localization of language-related cortex with ^{15}O-labeled water PET in patients with gliomas," *NeuroImage*, vol. 7, no. 4, pp. 284–295, 1998.

[26] C. Juhász, D. C. Chugani, O. Muzik, et al., "Electroclinical correlates of flumazenil and fluorodeoxyglucose PET abnormalities in lesional epilepsy," *Neurology*, vol. 55, no. 6, pp. 825–834, 2000.

[27] C. Juhász, D. C. Chugani, O. Muzik, et al., "Is epileptogenic cortex truly hypometabolic on interictal positron emission tomography?" *Annals of Neurology*, vol. 48, no. 1, pp. 88–96, 2000.

[28] O. Muzik, S. Pourabdollah, C. Juhász, D. C. Chugani, J. Janisse, and S. Draghici, "Application of an objective method for localizing bilateral cortical FDG PET abnormalities to guide the resection of epileptic foci," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 9, pp. 1574–1581, 2005.

[29] S. Lu, Y. Sun, M. Atay, and F. Fotouhi, "A new inlining algorithm for mapping XML DTDs to relational schemas," in *Proceedings of the 1st International Workshop on XML Schema and Data Management, in Conjunction with the 22nd ACM International Conference on Conceptual Modeling*, vol. 2814 of *Lecture Notes in Computer Science*, pp. 366–377, 2003.

[30] M. Atay, Y. Sun, D. Liu, S. Lu, and F. Fotouhi, "Mapping XML data to relational data: a dom-based approach," in *Proceedings of the 8th IASTED International Conference on Internet and Multimedia Systems and Applications (ISMA '04)*, pp. 59–64, Kauai, Hawaii, USA, August 2004.

[31] A. Chebotko, D. Liu, M. Atay, S. Lu, and F. Fotouhi, "Reconstructing XML subtrees from relational storage of XML documents," in *Proceedings of the 2nd International Workshop on XML Schema and Data Management (XSDM '05), in Conjunction with 21st International Conference on Data Engineering Workshops (ICDE '05)*, p. 1282, Tokyo, Japan, April 2005.

[32] M. Rege, D. Liu, and S. Lu, "Querying XML documents from a relational database in the presence of DTDs," in *Proceedings of the 1st International Conference Distributed Computing and Internet Technology (ICDCIT '04)*, pp. 168–177, Bhubaneswar, India, December 2004.

[33] E. Asano, C. Juhász, A. Shah, et al., "Origin and propagation of epileptic spasms delineated on electrocorticography," *Epilepsia*, vol. 46, no. 7, pp. 1086–1097, 2005.

[34] K. Kagawa, D. C. Chugani, E. Asano, et al., "Epilepsy surgery outcome in children with tuberous sclerosis complex evaluated with α-[^{11}C]methyl-L-tryptophan positron emission tomography (PET)," *Journal of Child Neurology*, vol. 20, no. 5, pp. 429–438, 2005.

[35] H. T. Chugani, M. E. Phelps, and J. C. Mazziotta, "Positron emission tomography study of human brain functional development," *Annals of Neurology*, vol. 22, no. 4, pp. 487–497, 1987.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 18709, 7 pages
doi:10.1155/2007/18709

Research Article

Adaptive Bayesian Iterative Transmission Reconstruction for Attenuation Correction in Myocardial Perfusion Imaging with SPECT/Slow-Rotation Low-Output CT Systems

Ji Chen,[1] Ernest V. Garcia,[1] Russell D. Folks,[1] Aharon Peretz,[2] and James R. Galt[1]

[1] *Department of Radiology, School of Medicine, Emory University, 1364 Clifton Road, Atlanta, GA 30322, USA*
[2] *GE Healthcare, 39120 Haifa, Israel*

Received 27 January 2006; Revised 14 December 2006; Accepted 14 December 2006

Recommended by David Townsend

Objectives. SPECT/slow-rotation low-output CT systems can produce streak artifacts in filtered backprojection (FBP) attenuation maps, impacting attenuation correction (AC) in myocardial perfusion imaging. This paper presents an adaptive Bayesian iterative transmission reconstruction (ABITR) algorithm for more accurate AC. *Methods*. In each iteration, ABITR calculated a three-dimensional prior containing the pixels with attenuation coefficients similar to water, then used it to encourage these pixels to the water value. ABITR was tested with a cardiac phantom and 4 normal patients acquired by a GE Millennium VG/Hawkeye system. *Results*. FBP AC and ABITR AC produced similar phantom results. For the patients, streak artifacts were observed in the FBP and ordered-subsets expectation-maximization (OSEM) maps but not in the ABITR maps, and ABITR AC produced more uniform images than FBP AC and OSEM AC. *Conclusion*. ABITR can improve the quality of the attenuation map, producing more uniform images for normal studies.

1. INTRODUCTION

Attenuation correction (AC) has undergone extensive clinical investigation [1–4] and now is a recommended technique for improving the quality of myocardial perfusion imaging (MPI) with single-photon emission computed tomography (SPECT) [5]. Since the attenuating material in a patient's thorax is too varied to meet the constant attenuation coefficient approximation made in both Sorenson's and Chang's methods [6, 7], transmission imaging is required to obtain patient-specific attenuation maps for accurate AC in MPI [8–11]. It has been pointed out that high-quality transmission scans and sufficient transmission counts with low crosstalk from the emission radionuclide are essential to reduce the propagation of noise and error into the attenuation-corrected emission images [5].

Recently, hybrid X-ray CT and SPECT systems became available for MPI. Among the commercial SPECT/CT systems, Millennium VG/Hawkeye (GE Healthcare Technologies, Milwaukee, Wis) produces the lowest cost and dose to the patient, since this camera uses a single-slice, slow-rotation, low-output CT scanner. AC with these systems has shown to improve sensitivity, specificity, and predictive accuracy in detection of coronary artery disease [12]. Compared to the conventional radionuclide attenuation maps, the Hawkeye CT attenuation maps have higher resolution and contrast and do not have the crosstalk and low-count issues. With phantom models, it has been shown that Hawkeye AC is, to some extent, superior to AC given by the SPECT systems with radionuclide transmission sources [13]. However, in clinical patients streak artifacts are often observed in the attenuation maps produced by filtered backprojection (FBP) of the slow-rotation low-output CT scans (shown in Figure 1). These artifacts are produced by inconsistencies in the CT sinogram such as those caused by respiratory motion during the CT acquisition. It has become a serious concern that these streak artifacts may degrade the accuracy of AC, and essentially raise the probability of false-positive cases in clinical practice.

Bayesian techniques have been introduced to transmission reconstruction for AC using radionuclide transmission imaging [14, 15]. It has been validated that Bayesian techniques can better handle the low-count issue associated with radionuclide transmission imaging than FBP, so as to

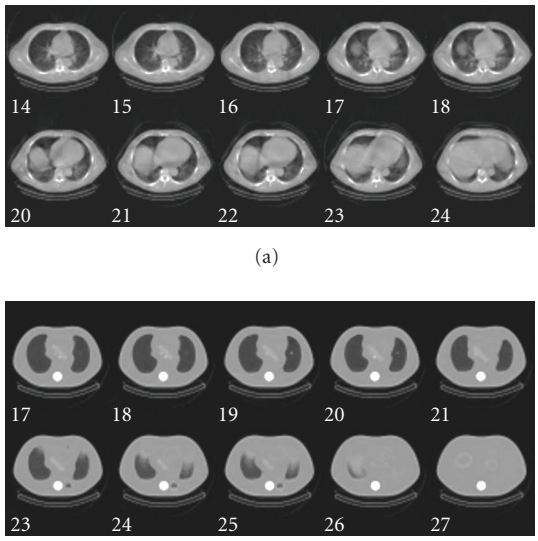


Figure 1: Example Hawkeye CT attenuation maps of (a) patient and (b) phantom.

enhance the quality of the reconstructed attenuation maps and improve the accuracy of AC [16, 17]. Although transmission noise is not an issue with the Hawkeye CT data, Bayesian technique may be useful in transmission reconstruction of the Hawkeye CT data for reducing the streak artifacts and improving the quality of the attenuation maps. Based on this hypothesis, this study developed an adaptive Bayesian iterative transmission reconstruction (ABITR) algorithm in order to remove the streak artifacts in the Hawkeye CT attenuation maps for more accurate AC.

2. MATERIALS AND METHODS

2.1. Phantom and patient studies

A data spectrum anthropomorphic torso phantom with cardiac, liver, lung, and spine components was used in this study. The cardiac insert consisted of a plastic chamber simulating the left ventricular chamber surrounded by a 1-cm-thick plastic chamber simulating the myocardium. Two smaller fillable chambers (90° and 45°, resp.), 1 cm thick and 2 cm long, were used to simulate hypoperfused defects (illustrated in Figure 2). The normal myocardium and defects were filled with 222 kBq/mL (6 μCi/mL) and 111 kBq/mL (3 μCi/mL) of Tc-99m, respectively, simulating typical patient uptakes of Tc-99m sestamibi or Tc-99m tetrofosmin. The liver and background were filled with 11.1 kBq/mL (0.3 μCi/mL) of Tc-99m.

This phantom was imaged four times with a GE Millennium VG/Hawkeye system using the following acquisition parameters:

(1) low-energy high-resolution collimators;
(2) 20% photopeak window centered at 140 keV;
(3) 12% scatter window centered at 118 keV;

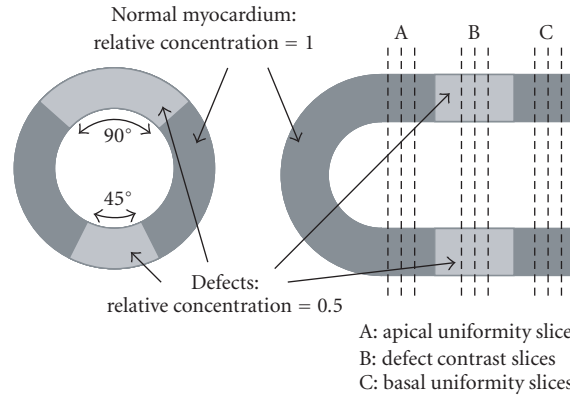


Figure 2: The cardiac insert configuration of the phantom.

(4) 180° acquisition, noncircular orbit;
(5) 60 projections, 30 seconds per projection.

The raw CT sinograms were preprocessed for crosstalk correction, nonlinearity correction, air and monitor normalization, and ring artifact correction by means of the manufacturer-provided tools. The preprocessed CT sinograms were rebinned from fan-beam geometry to parallel-beam geometry. In the rebinning, the pixel size of the CT sinograms was changed to the pixel size of the emission images. The rebinned sinogram was then submitted to ordered-subsets expectation-maximization (OSEM) (standard iterative reconstruction, without the Bayesian prior) and ABITR. Standard FBP (available on the system) was also used to reconstruct the attenuation maps for comparison. The FBP reconstruction was performed before the rebinning. The FBP attenuation maps were then rebinned to the same pixel size as the OSEM and ABITR attenuation maps.

Compton window subtraction [18] was performed on the tomographic projections as

$$P_{\text{scatter-compensated}} = P_{\text{photopeak}} - k \times P_{\text{scatter}}, \quad (1)$$

where $P_{\text{scatter-compensated}}$ is the scatter-compensated tomographic projection, $P_{\text{photopeak}}$ and P_{scatter} are tomographic projections in the photopeak window and scatter window, respectively, k is the scatter compensation scaling factor (the manufacturer-recommended value 1 was used). The scatter-compensated tomographic projections were submitted to emission reconstruction with AC. OSEM with 3 iterations and 6 subsets was used for the emission reconstruction. The reconstructed images then underwent three-dimensional (3D) Butterworth postfiltering (critical frequency of 0.4 cycles/cm and power of 10) followed by cardiac reorientation. Attenuation corrected short-axis images using the FBP, OSEM, and ABITR attenuation maps were then submitted to quantitative comparison and statistical analysis.

Four normal patients were used to compare OSEM AC and ABITR AC to FBP AC. These patients underwent a standard stress Tc-99m sestamibi rest Tl-201 dual-isotope protocol using the GE Millennium VG/Hawkeye system with the same acquisition parameters as those in the phantom

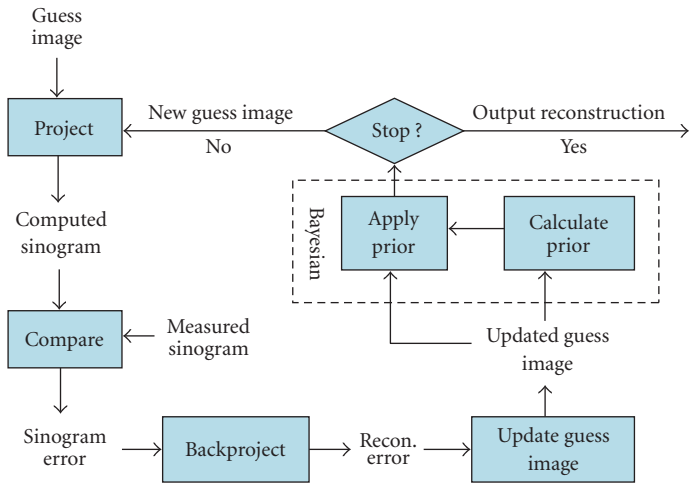


Figure 3: Illustration of the adaptive Bayesian iterative transmission reconstruction (ABITR) algorithm.

experiments. Short-axis images given by FBP AC, OSEM AC, and ABITR AC were obtained using the same reconstruction parameters as above and then submitted to quantitative comparison and statistical analysis.

2.2. Adaptive Bayesian iterative transmission reconstruction

ABITR is illustrated in Figure 3. The input of ABITR is a uniform image as initial guess image and a measured CT sinogram after preprocessing. A guess sinogram was generated from the guess image and then compared to the measured sinogram. The error of this comparison was backprojected, and then used to update the guess image. A 3D Bayesian prior image was calculated by image segmentation of the updated guess image. The image segmentation identified an object image which contained the pixels whose values were close to the water attenuation coefficients (±25%). Morphological processing was then applied to the object image to improve the results from image segmentation [19]. Morphological opening (removing small objects from an image while preserving the shape and size of large objects in the image) and closing (filling in small gaps inside large objects and smoothing the outer edges of large objects) were used to reduce the impact on image segmentation from the streak artifacts in the guess image. After the prior image was obtained, the pixels inside the object was updated using the following equation:

$$\text{Pixel}_{\text{new}} = (1 - b) \times \text{Pixel}_{\text{old}} + b \times \mu_{\text{water}}, \tag{2}$$

where $\text{Pixel}_{\text{old}}$ and $\text{Pixel}_{\text{new}}$ were the original and updated values of the pixels in the guess image that were inside the object, b is a factor controlling the strength of the Bayesian prior (with b equal to 0, the ABITR algorithm becomes a standard OSEM algorithm), and μ_{water} is the water attenuation coefficient, 0.153 cm^{-1} for 140 keV photopeak.

The guess image updated with the Bayesian prior was then submitted to forward projection to start a new iteration. ABITR stopped at a preset iteration number (30 iterations used in this study).

ABITR was performed on a PC with Pentium IV 2.4 GHz CPU and 512 Mb RAM. It was a totally automatic process and took less than 2 minutes to reconstruct a map with a 64×64 matrix and 25 slices.

2.3. Quantitative comparison and statistical analysis

Defect contrast and normal short-axis uniformity were used to compare the ABITR AC with the FBP AC in the phantom studies. Maximal-count circumferential profile (MCCP), extracted from a short-axis slice, was used to calculate the contrast and uniformity [20]. Three short-axis slices cutting through the two defects, as shown in Figure 2, were selected to calculate their defect contrasts, and their mean value represented the defect contract of that region. The defect contrast was defined as

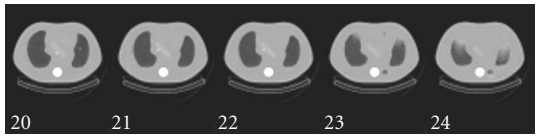
$$\text{contrast} = \frac{[\max(\text{MCCP}) - \min(\text{MCCP})]}{\max(\text{MCCP})}, \tag{3}$$

where max(MCCP) and min(MCCP) were maximal and minimal counts in the MCCP, respectively. The MCCPs extracted from three short-axis slices cutting through regions of homogeneous tracer distribution towards the base and towards the apex were selected to calculate their normal short-axis uniformities (shown in Figure 2), and their mean value represented the basal and apical uniformities, respectively. The uniformity was defined as

$$\text{uniformity} = \frac{\text{stdev}(\text{MCCP})}{\text{mean}(\text{MCCP})}, \tag{4}$$

where stdev(MCCP) and mean(MCCP) were the standard deviation and mean of the counts in the MCCP, respectively. The optimal defect contrast should be close to 50% and optimal uniformity should be close to 0.

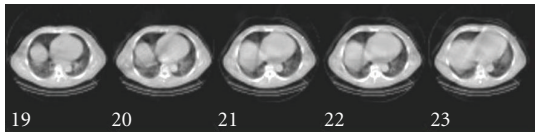
Short-axis uniformity, the same quantitative index as above, was used to compare the OSEM AC and ABITR AC



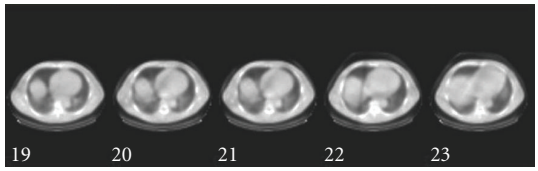
(a)

(b)

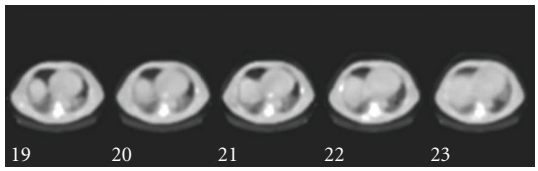
Figure 4: Phantom attenuation maps: (a) the FBP map (b) the ABITR map.



(a)



(b)



(c)

Figure 5: Patient (a) FBP, (b) OSEM, and (c) ABITR attenuation maps.

to the FBP AC in the patient studies for the apical, middle, and basal regions.

Paired t-test was used to compare the contrast and uniformity in the phantom studies and the uniformity in the patient studies.

3. RESULTS

Figure 4 shows the FBP and ABITR attenuation maps in phantom. Both maps looked similar. Figure 5 shows the FBP, OSEM, and ABITR attenuation maps in patient. Both FBP and OSEM maps had streak artifacts, not observed in the ABITR map.

The uniformity and defect contrast analyses of the phantom studies are shown in Table 1. No statistically significant difference was found between FBP AC and ABITR AC in the phantom studies. Table 2 shows the short-axis uniformity analysis of the patient studies. The ABITR AC images were barely significantly ($P = .0547$) more uniform at the basal region compared to the FBP AC images for the normal patient studies. No significant differences were obtained between the OSEM AC and FBP AC images, indicating that the Bayesian process played more important role in improving the quality of the attenuation map than iterative reconstruction.

Figure 6 shows the images of a normal subject. Bayesian AC produced images with less defect extent and severity, more similar to the AC normal database of the Emory Cardiac Toolbox than FBP AC.

4. DISCUSSION

This paper presents a transmission reconstruction technique, ABITR, for SPECT/slow-rotation low-output CT systems. Since the Hawkeye CT scanners use low-dose X-ray tubes, a typical transmission CT scan takes around 6 minutes to acquire 25 1-cm-thick slices. This slow rotation achieves good coregistration between the emission and transmission images; however, patient respiratory motion during the CT acquisition can result in inconsistency in the CT sinogram, and thus can create streak artifacts as we generally see in clinical practice. Due to the uncertainty in patient respiratory motion, it is difficult to manage the motion during the CT acquisition and to estimate where the streak artifacts will be present and how much they will impact AC. Nevertheless, the attenuation map with streak artifacts cutting through the myocardium may be of concern for clinicians and leads to questions regarding the possible impact of such attenuation maps on the quality of resulting AC images. From the patient studies presented in this paper, the ABITR technique showed to improve the quality of the slow-rotation low-output CT attenuation maps by eliminating the streak artifacts. It redistributed the pixel values in the attenuation maps to improve the soft-tissue uniformity while keeping the line integrals consistent to the data through iterative process. Compared to the FBP AC, ABITR AC yielded more uniform short-axis images for normal patient studies. In the phantom studies presented in this paper, ABITR and FBP yielded similar attenuation maps and AC images. This similarity indicated that ABITR did not create new artifacts in the attenuation map that can significantly impact AC. With phantom and patient studies, it is supported that ABITR can enhance the performance of the SPECT/slow-rotation low-output CT systems in AC in clinical MPI.

The original CT projections were in fan-beam geometry and had higher resolution (1-2 mm per pixel). It was rebinned to parallel-beam geometry with lower resolution (> 6 mm per pixel). The FBP reconstruction was done before the rebinning, whereas the ABITR and OSEM reconstructions were done after the rebinning. In other words, the ABITR AC had three differences from the FBP AC: (a) Bayesian processing, (b) iterative reconstruction; and (c) image resolution

Table 1: Short-axis uniformity and defect contrast analyses of the phantom studies.

Uniformity (%)	ABITR AC		FBP AC	
	Apical	Basal	Apical	Basal
Range	7.9–10.9	13.3–16.1	7.3–10.7	12.0–15.9
Mean	9.1	14.7	8.7	13.8
Std.	1.5	1.4	1.4	1.9
Mean Dif.*	0.4	0.9	—	—
Std. Dif.§	0.8	0.7	—	—
$P^{\dagger}$	.3888	.0857	—	—
Contrast (%)	Anterior	Inferior	Anterior	Inferior
Range	50.0–56.7	47.3–49.7	49.1–56.0	46.7–48.5
Mean	53.1	48.3	53.5	47.5
Std	2.8	10.4	3.0	9.2
Mean Dif	−0.4	0.8	—	—
Std Dif	1.4	1.6	—	—
P	.5864	.3732	—	—

* The mean differences between the ABITR AC and FBP AC studies.

§ The standard deviation of the mean differences.

† The P values were given by comparison between the ABITR AC and FBP AC studies using the paired t-test ($N = 4$). All of the P values in this table were greater than .05, indicating that there was no statistically significant difference between the ABITR AC and FBP AC images for the phantom studies.

Table 2: Short-axis uniformity analysis of the patient studies.

Uniformity (%)	ABITR AC			OSEM AC			FBP AC		
	Apical	Middle	Basal	Apical	Middle	Basal	Apical	Middle	Basal
Max.	9.2	8.1	21.1	9.0	8.2	22.4	8.8	7.9	22.0
Min.	4.4	5.2	12.7	4.8	5.3	13.7	5.0	5.9	12.9
Mean	7.2	6.6	15.8	7.3	7.0	16.5	7.1	7.1	16.3
Std.	2.0	1.2	3.7	1.8	1.3	4.0	1.6	1.0	4.1
Mean Dif.*	0.11	−0.43	−0.48	0.14	−0.16	0.21	—	—	—
Std. Dif.§	0.50	0.45	0.31	0.39	0.37	0.83	—	—	—
$P^{\dagger}$	.6771	.1510	.0547	.5327	.4536	.6480	—	—	—

* The mean differences between the ABITR AC and FBP AC studies and between the OSEM AC and FBP AC studies.

§ The standard deviation of the mean differences.

† The P values were given by comparison between the ABITR AC and FBP AC studies and between the OSEM AC and FBP AC studies using the paired t-test ($N = 4$). The P value of .0547 for the basal uniformities showed that the ABITR AC images had barely significantly better short-axis uniformity at the basal regions for the patient studies. No significant differences were obtained between the OSEM studies and FBP AC studies.

(FBP was used before rebinning, whereas ABITR was used after rebinning). It has been shown in Table 2 that there were no significant differences between the OSEM AC and FBP AC studies, indicating that reduction of the transmission image resolution by rebinning did not help much in improving the attenuation-corrected images. In Figure 5, the OSEM map appears smoother than the FBP attenuation map but less smooth than the ABITR attenuation map. It must be noted that the Bayesian processing (assigning similar attenuation coefficients for the tissue region) does have a smoothing effect on the image, but the smoothing is different from the smoothing we generally use to reduce image noise. There is actually very little noise in the CT transmission data. In summary, among the three differences between ABITR AC and FBP AC, the Bayesian processing was shown to be the major contributor in improving the quality of the attenuation map.

The major limitation of this study is the small number of patient studies. With the small sample size, it only showed the statistical trend but did not reach the statistical significance ($P < .05$). Another limitation of this study is that no abnormal patients were included in the preliminary evaluation of the ABITR AC. A gold standard that can accurately measure defect extent and severity has not yet been established for comparison between the FBP AC and ABITR AC

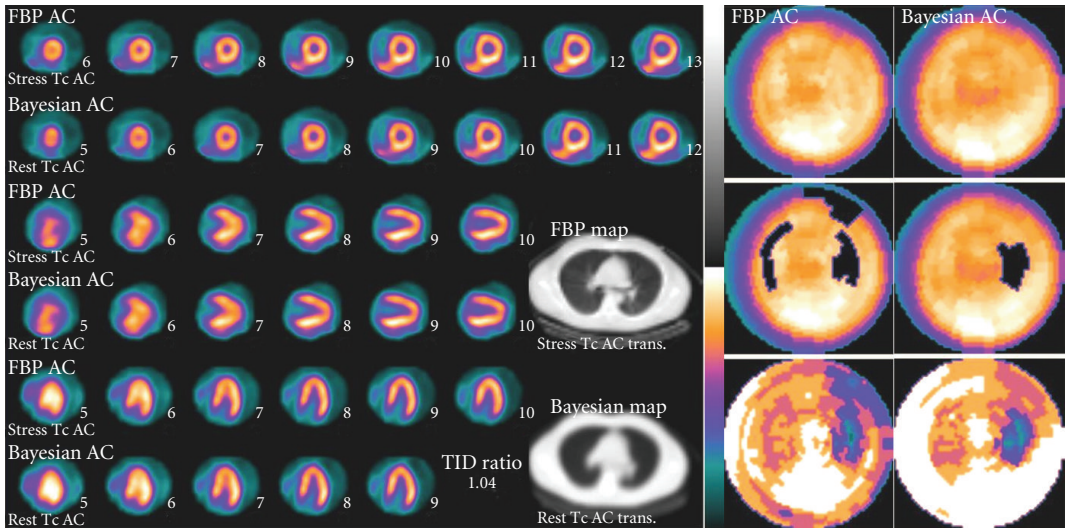


Figure 6: Images of a normal subject. FBP (top) and Bayesian (bottom) attenuation maps and AC images are shown on the left panel. FBP and Bayesian polar maps (three for each, shown as a column) are shown on the right panel. For each column, the images are raw polar map, defect extent, and defect severity given by the Emory Cardiac Toolbox and its attenuation correction normal file.

in abnormal patients. Nevertheless, the streak artifacts are more likely to create artifactual defects in the AC images of normal studies rather than to artifactually enhance the uniformity of the AC images on abnormal studies. The ABITR technique needs to be prospectively validated with a statistically sufficient sample size and with normal and abnormal patients and before it is implemented for clinical use. In addition, convergence test of the ABITR algorithm has not yet performed and a preset iteration number (30) was used in this study. It has been shown that iterative reconstruction of emission data converges around 20–50 iterations and then starts diverging when there is random noise in the emission data [21]. Since the transmission CT data has very little random noise, the ABITR algorithm is expected to converge quickly and to have very little divergence issues.

5. CONCLUSION

ABITR can remove the streak artifacts in the FBP attenuation maps caused by inconsistencies in the slow-rotation low-output CT sinogram such as those caused by patient respiratory motion during the acquisition. The improved quality of the ABITR attenuation map can yield more uniform attenuation-corrected images for normal subjects. ABITR can enhance the performance of SPECT/slow-rotation low-output CT systems in AC of clinical MPI. Prospective validation of this technique will be performed before the method is implemented for clinical use.

ACKNOWLEDGMENT

This work is supported in part by a grant from GE Healthcare.

REFERENCES

[1] R. C. Hendel, D. S. Berman, S. J. Cullom, et al., "Multicenter clinical trial to evaluate the efficacy of correction for photon attenuation and scatter in SPECT myocardial perfusion imaging," *Circulation*, vol. 99, no. 21, pp. 2742–2749, 1999.

[2] J. M. Links, L. C. Becker, P. Rigo, et al., "Combined corrections for attenuation, depth-dependent blur, and motion in cardiac SPECT: a multicenter trial," *Journal of Nuclear Cardiology*, vol. 7, no. 5, pp. 414–425, 2000.

[3] M. Shotwell, B. M. Singh, C. Fortman, B. D. Bauman, J. Lukes, and M. C. Gerson, "Improved coronary disease detection with quantitative attenuation-corrected Tl-201 images," *Journal of Nuclear Cardiology*, vol. 9, no. 1, pp. 52–61, 2002.

[4] J. M. Links, E. G. DePuey, R. Taillefer, and L. C. Becker, "Attenuation correction and gating synergistically improve the diagnostic accuracy of myocardial perfusion SPECT," *Journal of Nuclear Cardiology*, vol. 9, no. 2, pp. 183–187, 2002.

[5] R. C. Hendel, J. R. Corbett, S. J. Cullom, E. G. DePuey, E. V. Garcia, and T. M. Bateman, "The value and practice of attenuation correction for myocardial perfusion SPECT imaging: a joint position statement from the American Society of Nuclear Cardiology and the Society of Nuclear Medicine," *Journal of Nuclear Cardiology*, vol. 9, no. 1, pp. 135–143, 2002.

[6] J. A. Sorenson, "Quantitative measurement of radioactivity in vivo by whole body counting," in *Instrumentation in Nuclear Medicine*, J. H. Hine and J. A. Sorenson, Eds., pp. 311–348, Academic Press, New York, NY, USA, 1974.

[7] L.-T. Chang, "A method for attenuation correction in radionuclide computed tomography," *IEEE Transactions on Nuclear Science*, vol. 25, no. 1, pp. 638–643, 1978.

[8] B. M. W. Tsui, G. T. Gullberg, E. R. Edgerton, et al., "Correction of nonuniform attenuation in cardiac SPECT imaging," *The Journal of Nuclear Medicine*, vol. 30, no. 4, pp. 497–507, 1989.

[9] M. A. King, B. M. W. Tsui, and T.-S. Pan, "Attenuation compensation for cardiac single-photon emission computed tomographic imaging—part 1: impact of attenuation and methods of estimating attenuation maps," *Journal of Nuclear Cardiology*, vol. 2, no. 6, pp. 513–524, 1995.

[10] H. Zaidi and B. Hasegawa, "Determination of the attenuation map in emission tomography," *The Journal of Nuclear Medicine*, vol. 44, no. 2, pp. 291–315, 2003.

[11] E. G. DePuey, E. V. Garcia, and S. Borges-Neto, "Updated imaging guidelines for nuclear cardiology procedures—part 1," *Journal of Nuclear Cardiology*, vol. 8, no. 1, pp. G1–G58, 2001.

[12] Y. Liu, F. J. Wackers, D. Natale, et al., "Validation of a hybrid SPECT/CT system with attenuation correction: a phantom study and multicenter trial," *The Journal of Nuclear Medicine*, vol. 44, p. 290, 2003.

[13] M. K. O'Connor, B. Kemp, F. Anstett, et al., "A multicenter evaluation of commercial attenuation compensation techniques in cardiac SPECT using phantom models," *Journal of Nuclear Cardiology*, vol. 9, no. 4, pp. 361–376, 2002.

[14] J. A. Case, T.-S. Pan, M. A. King, D.-S. Luo, B. C. Penney, and M. Z. Rabin, "Reduction of truncation artifacts in fan beam transmission imaging using a spatially varying gamma prior," *IEEE Transactions on Nuclear Science*, vol. 42, no. 6, part 2, pp. 2260–2265, 1995.

[15] M. V. Narayanan, C. L. Byrne, and M. A. King, "An interior point iterative maximum-likelihood reconstruction algorithm incorporating upper and lower bounds with application to SPECT transmission imaging," *IEEE Transactions on Medical Imaging*, vol. 20, no. 4, pp. 342–353, 2001.

[16] J. A. Case, S. J. Cullom, J. R. Galt, F. V. Garcia, and T. M. Bateman, "Impact of transmission scan reconstruction using an iterative algorithm (BITGA) versus FBP: clinical appearance of attenuation-corrected myocardial perfusion SPECT images," *The Journal of Nuclear Medicine*, vol. 42, p. 51, 2001.

[17] S. J. Cullom, J. A. Case, T. M. Bateman, J. H. O. O'Keefe, and A. McGhie, "Reconstruction of attenuation maps from low-count Gd-153 transmission studies using an iterative Bayesian algorithm: clinical evaluation with simultaneous Tc-99m sestamibi SPECT," *The Journal of Nuclear Medicine*, vol. 41, no. 5, p. 134, 2000.

[18] R. J. Jaszczak, K. L. Greer, C. E. Floyd Jr., C. C. Harris, and R. E. Coleman, "Improved SPECT quantitation using compensation for scattered photons," *The Journal of Nuclear Medicine*, vol. 25, no. 8, pp. 893–900, 1984.

[19] E. R. Dougherty, *An Introduction to Morphological Image Processing*, SPIE Press, Bellingham, Wash, USA, 1992.

[20] E. V. Garcia, C. D. Cooke, K. F. Van Train, et al., "Technical aspects of myocardial SPECT imaging with technetium-99m sestamibi," *The American Journal of Cardiology*, vol. 66, no. 13, pp. 23E–31E, 1990.

[21] G. L. Zeng and G. T. Gullberg, "Unmatched projector/backprojector pairs in an iterative reconstruction algorithm," *IEEE Transactions on Medical Imaging*, vol. 19, no. 5, pp. 548–555, 2000.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 46846, 10 pages
doi:10.1155/2007/46846

Research Article

Improved Image Fusion in PET/CT Using Hybrid Image Reconstruction and Super-Resolution

John A. Kennedy,[1] Ora Israel,[2,3] Alex Frenkel,[2] Rachel Bar-Shalom,[2,3] and Haim Azhari[1]

[1] *Faculty of Biomedical Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel*
[2] *Department of Nuclear Medicine, Rambam Health Care Campus, Haifa 35245, Israel*
[3] *The Ruth and Bruce Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, Efron Street 1, P.O. Box 9649 Bat Galim, Haifa 31096, Israel*

Received 11 June 2006; Revised 3 September 2006; Accepted 17 October 2006

Recommended by David Townsend

Purpose. To provide PET/CT image fusion with an improved PET resolution and better contrast ratios than standard reconstructions. *Method*. Using a super-resolution algorithm, several PET acquisitions were combined to improve the resolution. In addition, functional PET data was smoothed with a hybrid computed tomography algorithm (HCT), in which anatomical edge information taken from the CT was employed to retain sharper edges. The combined HCT and super-resolution technique were evaluated in phantom and patient studies using a clinical PET scanner. *Results*. In the phantom studies, 3 mm ^{18}F-FDG sources were resolved. PET contrast ratios improved (average: 54%, range: 45%–69%) relative to the standard reconstructions. In the patient study, target-to-background ratios also improved (average: 34%, range: 17%–47%). Given corresponding anatomical borders, sharper edges were depicted. *Conclusion*. A new method incorporating super-resolution and HCT for fusing PET and CT images has been developed and shown to provide higher-resolution metabolic images.

1. INTRODUCTION

Positron emission tomography (PET) provides images of metabolic processes that are used increasingly in the clinical setting to obtain data on cancer and other pathological processes. In oncology, the diagnosis of cancer and the assessment of the extent of disease often rely on PET [1]. However, because PET images are relatively noisy and are limited by relatively poor spatial resolution, small lesions are difficult to detect [2] and the anatomical location of hypermetabolic regions can be difficult to determine in PET images [3].

The introduction of dual modality PET/CT scanners [4, 5] has addressed the latter issue by providing metabolic PET images registered with the anatomical information from CT. In these scanners, the patient lies still on a bed which is then translated through fixed mechanically aligned coaxial CT and PET gantries so that the data acquired are precisely coregistered in space. The PET acquisition typically occurs immediately after the CT acquisition to minimize the effects of patient motion. After reconstruction, the high-resolution anatomical images (from CT) are overlayed with the functional images (from PET) to provide precise localization of hypermetabolic regions. In oncology, such image fusion has been shown to improve the diagnostic reliability [6, 7].

In the interest of improving small lesion detectability, the objective of this study was to provide a new method for PET/CT image fusion with an improved resolution and better contrast ratio relative to standard reconstructions. First, a modified form of the super-resolution method of Irani and Peleg [8] shown to improve resolution in PET imaging (Kennedy et al. [9]) was employed for PET data acquisition and image reconstruction. In the super-resolution method, several acquisitions interspersed with subpixel shifts are combined in an iterative algorithm to yield a higher-resolution image, depicted schematically in Figure 1. Secondly, since the radiopharmaceutical distribution will often follow anatomical borders, the potential exists to combine the high-resolution border information from the CT image with the functional distribution from the PET image to yield a PET image with enhanced borders. The algorithm we used to incorporate CT data in PET images is called hybrid computed tomography (HCT). HCT was originally developed for

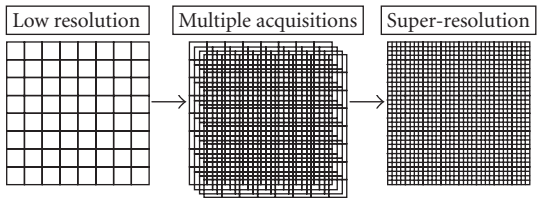


Figure 1: Super-resolution algorithms combine multiple low-resolution image acquisitions into a high-resolution image.

artifact reduction in ultrasonic computed tomography [10]. In regions not containing anatomical edges, HCT has been shown to provide noise reduction in PET images equivalent to the standard Gaussian filtering typically used [11]. In PET imaging, HCT provides sharper edges and improves contrast ratios [11].

In this paper, we demonstrate how a combination of a super-resolution acquisition and reconstruction combined with HCT filtering increases the contrast ratios of ^{18}F-FDG uptake in PET images while providing noise reduction equivalent to a standard Gaussian filter in regions without corresponding anatomical edges. Where corresponding anatomical edges are available, the technique enhances the edges of ^{18}F-FDG uptake. Through the combination of increased resolution and edge enhancement, the PET imaging of small features is improved.

2. MATERIALS AND METHODS

PET was performed using standard and super-resolution acquisitions [9]. Each type of acquisition was then filtered with one of two techniques: a standard Gaussian filter or the equivalent HCT filter [11] incorporating CT border information. Consequently, four methods of generating PET images were compared:

(a) standard acquisition and processing with Gaussian filtering;
(b) super-resolution acquisition and processing with Gaussian filtering;
(c) standard acquisition and processing with HCT filtering;
(d) super-resolution acquisition and processing with HCT filtering.

The degree of filtering was chosen to keep the level of noise constant among images compared.

2.1. Super-resolution and HCT

The term super-resolution refers here to a technique in which several low-resolution points of view (POVs) are combined iteratively to obtain a higher-resolution image. In the Irani and Peleg formulation of a super-resolution algorithm [8], the initial estimate of the high-resolution image, $f^{(0)}$, can be based on the average of the upsampled acquisitions shifted to a common reference frame:

$$f^{(0)} = \frac{1}{K}\sum_{k=1}^{K} T_k^{-1}(g_k \uparrow s), \tag{1}$$

where g_k is one of K acquisitions, T_k^{-1} is the geometric transformation to a common reference frame, and $\uparrow s$ is the upsampling operator from low-resolution to the high-resolution representation.

One could obtain the low-resolution measured data g_k from the "true" image f if the acquisition system was adequately modeled. The process would include shifting the image f to the kth POV, blurring to account for limited system resolution, downsampling to the system's sampling rate, and adding noise. For a given estimate of the image, $f^{(n)}$, the low-resolution data is modeled as in [8]:

$$\tilde{g}_k^{(n)} = (T_k(f^{(n)}) * h)\downarrow s, \tag{2}$$

where $*h$ is the blurring operation with the kernel h and $s\downarrow$ is the downsampling operator which averages the pixels to the lower resolution. The noise term is dropped. The original geometric transformation of the kth acquisition from the common reference frame is T_k. This is typically the physical shift between the object and the imager from the original position.

To obtain a better estimate of the image f, the previous estimate of the high-resolution image, $f^{(n)}$, is corrected by the difference between the low-resolution data g_k and the term $\tilde{g}_k^{(n)}$ that represents what the low-resolution data would have been, had the estimate, $f^{(n)}$, been correct. The next iteration $f^{(n+1)}$ of a high-resolution estimate is the following [8]:

$$f^{(n+1)} = f^{(n)} + \frac{1}{K}\sum_{k=1}^{K} T_k^{-1}\left(\left(\left(g_k - \tilde{g}_k^{(n)}\right)\uparrow s\right) * p\right). \tag{3}$$

Here, the differences between g_k and $\tilde{g}_k^{(n)}$ are upsampled, $\uparrow s$, to achieve the smaller super-resolution pixel size, moved to a common reference frame, T_k^{-1}, and averaged over K acquisitions. The symbol $*p$ is a sharpening kernel. This formulation of the super-resolution algorithm has been demonstrated to improve resolution in MRI imaging [12, 13] and in PET [9].

Although the blur and sharpening kernels can be set to unity [9, 12], in this study the blur kernel has been modeled as a Gaussian point spread function (PSF). In order to reduce the noise caused by sharpening, the upsampling procedure of Farsiu et al. [14] was used.

In addition to the super-resolution acquisition, a modified form of an iterative algorithm called hybrid computed tomography (HCT), implemented previously on ultrasonic CT data [10], was utilized here to fuse CT anatomical data with the PET functional data. The HCT algorithm is based on a two-dimensional (2D) Taylor series expansion of the gray levels which incorporates texture and edge information. The HCT algorithm utilizes edge information taken from CT to retain sharper edges while smoothing the PET data, which

often follow the anatomical borders. Thus, the resulting reconstructed image has reduced noise but sharp borders.

In HCT, each value of the image f at each pixel is expanded into neighboring pixels. Neglecting higher-order terms, the modified 2D Taylor expansion about pixel (a, b) has a value $f(x, y)$ at pixel (x, y) [10]:

$$f(x,y) = f(a,b) + \left[(x-a)\cdot \left.\frac{\partial f}{\partial x}\right|_{a,b} + (y-b)\cdot \left.\frac{\partial f}{\partial y}\right|_{a,b}\right] \cdot \beta(a,b), \tag{4}$$

where the function $\beta(x, y)$ has a zero value within homogeneous regions but is set to have a value of 1 at boundary points. In the PET/CT application, the function β can be obtained from the anatomical edge data of the CT scan. One method of modifying (4) to include discrete pixels and diagonal directions is to write it as

$$f(x,y) = f(a,b) + \left[\Delta_r \cdot \left.\frac{\Delta f}{\Delta_r}\right|_{a,b}\right] \cdot \beta(a,b), \tag{5}$$

where Δ_r is the step size in the direction $\vec{r} = [x-a \quad y-b]$ and $\Delta f = f(x, y) - f(a, b)$. Here, the expansion was limited to nearest neighbors, as depicted in Figure 2, so the step size was unity: $\Delta_r = 1$. In one HCT iteration, (5) is applied in a neighborhood of $f(x, y)$ and the results averaged, for each pixel (x, y) in the image. In the absence of a border, repeated iterations of (5) average a pixel value with its neighbors. If a 3×3 neighborhood is used, in regions distant from a border, it can be shown that n HCT iterations are equivalent to the application of a Gaussian filter with a full-width half-maximum (FWHM) of [11]:

$$\text{FWHM} = 4\sqrt{\frac{\ln(2)n}{3}} \text{ pixels.} \tag{6}$$

If the functional and anatomical boundaries do not match, HCT may introduce artifacts [11], but in the absence of border information the default of HCT is the standard Gaussian filtering.

For a simple HCT example, consider the 3×3 image in Figure 2. The central pixel f_{22} has an uptake indicated by the gray shading. In the first HCT iteration, the value assigned to f_{22} by (5) is determined by its nearest neighbors. If the thick solid line is the true border, β between the central pixel and the 3 gray pixels in the first column is set to 0 because there is no border among them and (5) sets the value of $f(x, y)$ to $f(a, b)$. However, when the index (a, b) falls on the other side of the border, β is set to 1 and $f(x, y)$ retains its original value. When applied to all 9 neighborhood pixels, the uptake in the central pixel is averaged with the uptake in those 3 gray pixels in the first column. Equation (5) generates a weighted average; in this case the center pixel is weighted at 6/9 and the 3 other pixels are weighted at 1/9 each. However, if the true border is between the central pixel f_{22} and f_{12}, as indicated by the dotted line, then β is set to 0 only among the pixels of the second and third columns. In the first iteration, the value of the central pixel is averaged with the 5 other pixels in the second and third columns which have no uptake (as indicated by white). Although the value of the central pixel is substantially reduced, the application of (5) to each of the other 5 pixels in turn effectively distributes this uptake among the 6 pixels in the second and third columns. Regardless of the position of the border, the application of (5) is an averaging operation; therefore HCT is a counts-preserving process.

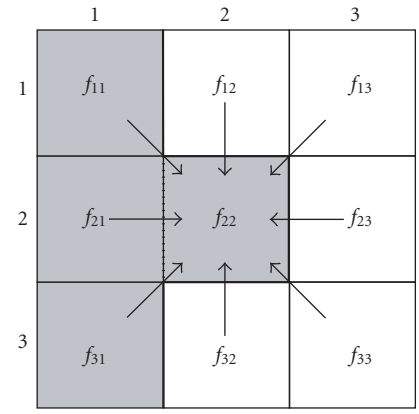


Figure 2: HCT applied to a 3×3 image. In the case that pixel f_{22} indicates a true uptake (gray), the solid line is the true border and HCT algorithm iteratively averages its value with the pixels in the first column. In the case that dotted line is the true border, the uptake in pixel f_{22} iteratively averages its value with the pixels in the second and third columns.

The combined technique (i.e., super-resolution and HCT) was evaluated in both phantom (3D brain-mode acquisition) and patient studies (2D whole-body mode acquisition), using a clinical PET scanner (GE Discovery LS, GE Healthcare Technologies, Milwaukee, WI).

2.2. Data acquisition and processing

The GE Discovery LS combines X-ray CT and PET scanners arranged such that the gantries are coaxial and a bed can automatically move through each gantry in order to provide images in both modalities that are coregistered. The PET portion of the scanner is similar to a GE Advance NXi described elsewhere [9, 15]. In a standard 2D whole-body PET acquisition, the septa between the 18 detector rings restrict the photons acquired to the transaxial plane. Transaxial images (35 per field of view, FOV) are typically reconstructed as 128×128 pixel images having a pixel size of 4 mm $\times$ 4 mm and a slice thickness of 4.25 mm. The axial FOV is 14.5 cm and the transaxial FOV, as reconstructed in this mode, is 50 cm. An ordered subsets expectation maximization (OSEM) algorithm [16] using 2 iterations and 28 subsets was used for reconstructing the 2D whole-body data from the PET sinograms (projections). Coronal and sagittal images are typically obtained by stacking the images of several axial FOVs into a three-dimensional (3D) data set and reslicing appropriately.

The 3D brain-mode acquisition is similar except that the septa are retracted to increase the number of photons detected. The data was rebinned into transaxial data sets using

Fourier rebinning [17] before being reconstructed with an OSEM algorithm using 5 iterations and 32 subsets. The pixel size is typically set to 2 mm × 2 mm reducing the reconstructed transaxial FOV width by a factor of 1/2. The slice thickness remains the same as in the 2D whole-body mode.

The CT provided 512 × 512 pixels transaxial images with a pixel size of 1 mm × 1 mm and a slice thickness of 4.25 mm which were coregistered with the PET images. A tube voltage of 140 kV and current of 90 mA was used. For attenuation corrected (AC) PET images, the CT images also served as the basis for an attenuation map by means of rescaling the Hounsfield units (HU) of the CT to attenuation coefficients appropriate for the higher energy of PET gamma rays [18–21].

In this study, the 2D whole-body mode data was reconstructed with a voxel size of 2 mm × 2 mm × 4.25 mm, similar to the 3D brain-mode acquisition. This gave transaxial PET images of 256 × 256 pixels for the 2D whole-body mode. This was the voxel size for all the standard acquisitions and for each low-resolution POV in the super-resolution acquisition data sets. After processing with the super-resolution technique, the pixel sizes obtained were smaller. When super-resolution was applied in the transaxial plane (see below), the resulting voxel size was 1 mm × 1 mm × 4.25 mm. When super-resolution was applied axially (see below), the resulting voxel size was 2 mm × 2 mm × 1 mm.

Unfiltered image data sets from standard and super-resolution acquisitions were then filtered with either a standard Gaussian filter or an HCT filter which could incorporate edge information while providing equivalent smoothing (6) in regions away from anatomical edges. The smoothing was set to maintain the same level of noise among the images obtained from the four processing methods (see below). In order to make effective use of the resolution of the border information provided by the CT [11], the filtering was applied after the images had been interpolated to a 0.25 mm × 0.25 mm pixel size for the 3D brain-mode PET/CT acquisitions and 0.5 mm × 0.5 mm for the 2D whole-body case using a piecewise cubic Hermite interpolation. The edges were extracted using a Canny edge detector algorithm [22] on CT images to which the scanner protocol's default contrast window had been applied (level: 40 HU, width: 400 HU). For edge extraction, the Gaussian smoothing employed on the CT by the Canny edge detector was 1.2 mm FWHM for the 3D brain-mode PET/CT acquisitions and 3.0 mm FWHM for the 2D whole-body case.

2.3. Phantom study

To evaluate image quality among the four processing methods implemented here, a specially designed phantom was used (Figure 3). The phantom provided cylindrical hotspots of ^{18}F-FDG solution with diameters of 1, 1.5, 2, 3, 4, 6, and 8 mm arranged in rows such that the separation between hotspots was equal to their diameters. The hotspots were created by drilling holes through a polycarbonate disk (diameter 9 cm, thickness 1.5 cm) and treating the disk with ozone to allow ^{18}F-FDG solution (130 kBq/mL) to flow freely through them when the disk was immersed in a fitted cup containing the solution. To a 1 cm depth, on each side of the disk, the cup contained just ^{18}F-FDG solution.

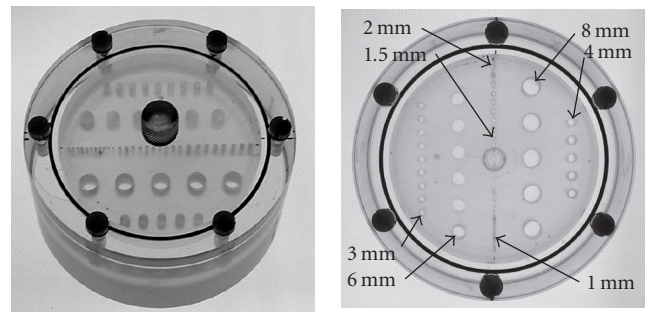


Figure 3: Phantom: a specially treated polycarbonate disk allowed ^{18}F-FDG solution to flow freely through holes of varying sizes when immersed in a cup of the solution.

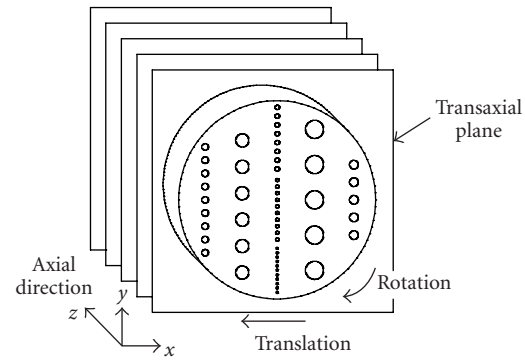


Figure 4: Geometry of phantom orientation for the 3D brain-mode PET acquisition. The phantom disk was aligned with the transaxial plane and translated and rotated within that plane between each of four separate POVs.

The phantom was placed in the scanner to obtain transaxial images in the plane of the disk using the 3D brain-mode acquisition protocol (Figure 4). A standard acquisition of 10-minute duration was followed by 4 acquisitions of 2.5 minutes each for the super-resolution acquisition. Each PET acquisition was accompanied by a CT scan to provide attenuation correction (AC) according to common practice with such PET/CT scanners [18]. Between the 4 acquisitions, the phantom was given a small displacement and rotation in the transaxial plane to provide the geometrical shifts needed by the super-resolution algorithm. The position of the initial acquisition was taken to be the common reference frame. In the case of the phantom trial, the size of the geometric shifts was tracked in the CT images using two 1 mm markers separated by 43 cm that had been fixed to the phantom in the transaxial plane. The shifts used are listed in Table 1. The initial CT image also provided the border information used by the HCT algorithm.

Table 1: Transaxial displacements and rotations from the initial position used in the 3D AC brain-mode acquisition phantom trial.

2.5-minute PET displacement acquisition	Horizontal displacement left (mm)	Vertical displacement up (mm)	Clockwise rotation (degrees)
Initial	0	0	0
Second	2.0	0.5	1.7
Third	5.0	1.2	3.9
Fourth	9.1	2.0	7.2

The geometry of the phantom and the method of super-resolution acquisition in the 3D brain mode is described elsewhere [9] in more detail.

For comparison purposes, each processing method was applied to achieve the same degree of noise reduction. As a measure of the noise, the variance in the PET signal was calculated in a region known to have a homogeneous uptake of ^{18}F-FDG solution. The transaxial slices of the cup of ^{18}F-FDG solution on either side of the polycarbonate disk contained no features except for the 9.0 cm diameter circular edge of the cup. A 5.0 cm diameter circular region of interest (ROI) was selected from one of these slices. Because such a region contains no edges from the CT, both HCT and Gaussian filtering provide the same degree of smoothing [11]. The FWHM (or HCT equivalent) of the smoothing was chosen so that the standard and super-resolution acquisitions and reconstructions had the same variance within this homogeneous ROI. The same filters were then applied to the phantom images containing the features of interest: the uptake in the holes of the polycarbonate disk.

As an indication of image quality, a contrast ratio was calculated for the phantom results. For each row of holes, the locations of the sources were known so they were masked and an average PET signal was calculated. The regions falling between holes were also masked and those pixel values were used to calculate an average background value for that row. The contrast ratio was taken to be the average PET signal to the average background, so that a contrast ratio of unity would indicate that the feature could not be distinguished. Because the level of noise as measured by the variance was kept constant, comparing these contrast ratios was equivalent to comparing a contrast to variance metric.

Three additional studies were performed to measure the PET resolution of this experimental arrangement in terms of a PSF of the data acquisition. A single 1 mm hole of the phantom disk was filled with 20 μCi (0.74 MBq) ^{18}F-FDG solution and capped in order to emulate a "point source" for transverse 3D brain-mode images that were acquired as above. The reference position for the source was 2.0 cm above the axial center line of the scanner. Additionally, to check axial resolution, the phantom was laid flat and fixed to the bed to emulate a "point source" in coronal images. Between each of 4 PET acquisitions, the bed was automatically shifted into the scanner in 1 mm increments, and the super-resolution technique was applied axially. The process was repeated for the 2D whole-body mode. These results have been reported elsewhere [9], but that study used a blurring-and-deblurring kernel of 1 pixel. Here, as a modification, the blur kernel was set to a Gaussian PSF with a FWHM chosen to minimize the FWHM of the "point source" and the blurring-and-deblurring procedure [14] described above was used. For the purpose of direct comparison, the same data set as the previous report [9] was used.

Anticipating the focus of the patient study below, the axial resolution of the 2D whole-body mode was also checked for 2 POVs with 2 mm axial shifts and 8 POVs with 0.5 mm axial shifts.

2.4. Patient study

The patient was injected with 370 MBq of ^{18}F-FDG after a 4 h fast and was then kept resting comfortably for 90 min before scanning. A 2D head-to-thigh PET/CT scan was acquired, including a CT scan followed by a PET scan consisting of 6 FOVs with an acquisition time of 4 min per FOV. During this standard PET acquisition, the CT was reviewed to identify an ROI suitable for employing the super-resolution technique. A FOV was chosen containing a suspected small lung lesion. After the standard PET scan, the patient was requested to remain still, the bed registration was maintained, and 4 additional POVs of the ROI were acquired, taking 4 min each. Each 4-minute acquisition interval was subdivided into 1-minute and 3-minute intervals so that four 1-minute-long POVs were available to check the case in which the total super-resolution acquisition time equaled the standard acquisition time. Between each subsequent POV, the bed was automatically moved 1 mm further into the scanner to provide 4 PET views differing by shifts which were subpixel since the slice thickness of a standard PET acquisition in the axial direction was 4.25 mm. The patient was not exposed to additional radiation since the X-ray CT scan was not repeated. Because registration was maintained, the initial X-ray CT scan could be used to provide border information for the HCT processing of both the standard and super-resolution PET images by matching the data from any transaxial PET slice with the data from the appropriate transaxial X-ray CT slice at the same location.

As in the phantom trial, the patient images were processed by the four methods. Nonattenuation corrected images were used because the pulmonary lesion was more evident than in the AC PET. The degree of image noise was measured by the variance. In the absence of a known region of homogeneous uptake, the variance was calculated from the nonzero pixel values excluding a 15 mm circular ROI around the lesion of interest in the coronal images. The degree of filtering in each of the four processing methods was chosen to keep the noise level the same, as measured by this variance.

In order to compare PET images in the patient study, target-to-background ratios were calculated as a measure of the intensity of the lesion's uptake for coronal, sagittal, and transverse slices through the lesion of interest. The precise target shape and location were unknown, so the masking method used for the phantom contrast ratio calculations was

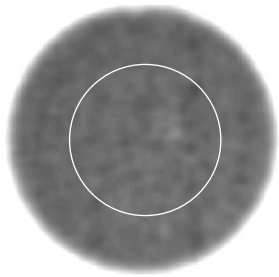


Figure 5: Transaxial 3D brain-mode PET image of a slice through the 9.0 cm diameter phantom cup. The 5.0 cm diameter ROI (white circle) was used to calculate the variance as a measure of image noise since it was known to contain a homogeneous distribution of ^{18}F-FDG solution.

inappropriate here. However, because the small lesion had substantially higher uptake than other tissues in each of the images, its location could be demarcated by setting a threshold. For each image, the target was defined as pixels having values greater than 60% of the maximum pixel value for that image. To exclude uptake erroneously assigned to regions known to be outside the body, a minimum threshold was set (5% of the maximum pixel value). The remaining nonzero pixels defined the background. The target-to-background ratio was calculated as the mean of the target pixel values divided by the mean of the background pixel values. A more intense, localized uptake would have a higher target-to-background ratio.

3. RESULTS

In order to establish that the phantom images had the same noise level, a transaxial slice adjacent to the polycarbonate disk was selected and an ROI used to measure noise was chosen in a region of homogeneous ^{18}F-FDG uptake (the white circle in Figure 5). To maintain a variance of $10.6 \pm 0.1\,\text{kBq}^2/\text{mL}^2$ in this ROI, the standard acquisitions were smoothed with a 1.8 mm FWHM Gaussian filter (equivalent to 15 HCT iterations; see (6)) and the super-resolution results were smoothed with a 3.0 mm FWHM Gaussian filter (equivalent to 41 HCT iterations). These filters were also applied on the transaxial images through the polycarbonate disk showing the features of interest (Figure 6).

In the phantom trial (Table 2), the super-resolution technique improved the concentration ratios of the 3 mm, 4 mm, 6 mm, and 8 mm features from an average of 1.9 (range: 1.1–2.9) for the standard acquisition to an average of 2.1 (range: 1.2–3.3). HCT filtering also improved the standard contrast ratios to an average of 2.1 (range: 1.3–3.1). Using the combined acquisition and processing technique of super-resolution and HCT, the PET contrast ratios were the highest (average: 2.8, range: 1.6–4.3). Using the super-resolution/HCT technique, 3 mm ^{18}F-FDG sources were more clearly resolved (Figure 6) than the standard image and the edges of the sources were more delineated. A plot of pixel value profiles through the 3 mm features of the phantom

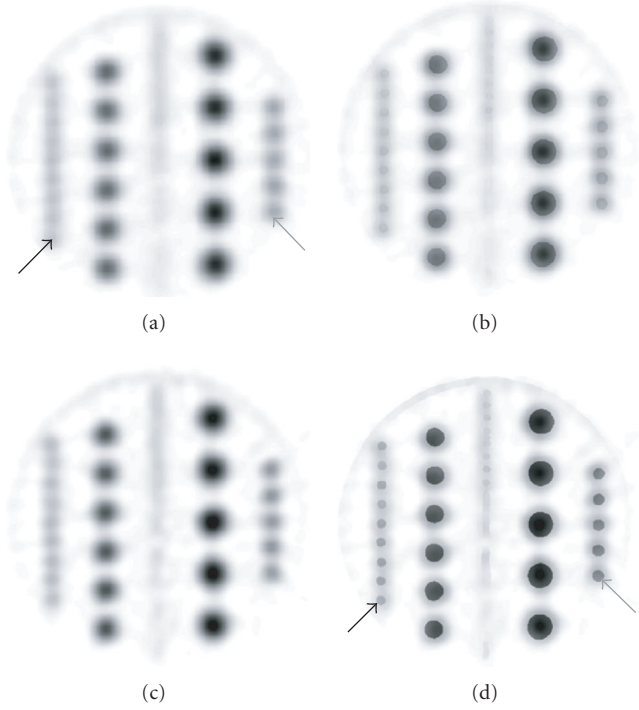


Figure 6: Transaxial PET images through the phantom disk using 3D brain-mode acquisition. (a) Standard processing. The nine hotspots in the row (black arrow) along the left are 3 mm in diameter and the five largest hotspots are 8 mm (gray arrow). (b) HCT result. (c) Super-resolution result. (d) Super-resolution/HCT result has the greatest contrast. The 3 mm sources (black arrow) are more clearly resolved than in the standard image. The 8 mm sources (gray arrow) show sharper edges than in the standard image.

Table 2: Contrast ratios for the PET signals in the 3D AC brain-mode acquisition phantom trial.

Image type	3 mm holes	4 mm holes	6 mm holes	8 mm holes
Standard	1.1	1.3	2.1	2.9
Super-resolution	1.2	1.5	2.4	3.3
HCT	1.3	1.5	2.4	3.1
HCT and super-resolution	1.6	2.2	3.2	4.3

(Figure 7) shows that the super-resolution profile (dashed line) and the HCT profile (dotted) both gave moderately better contrast than the standard method (dashed and dotted). The combination of HCT and super-resolution gave the best contrast of all the methods (Figure 7, solid black line).

The efficacy of including a Gaussian blur kernel in the super-resolution processing [14] was checked by measuring the PSF in the axial direction (2D whole-body mode and 3D brain mode) and transaxial directions (3D brain mode). In each type of image, the "point source" was provided by a cross section through a single 1 mm hole of the phantom which had been filled with ^{18}F-FDG and capped. Table 3

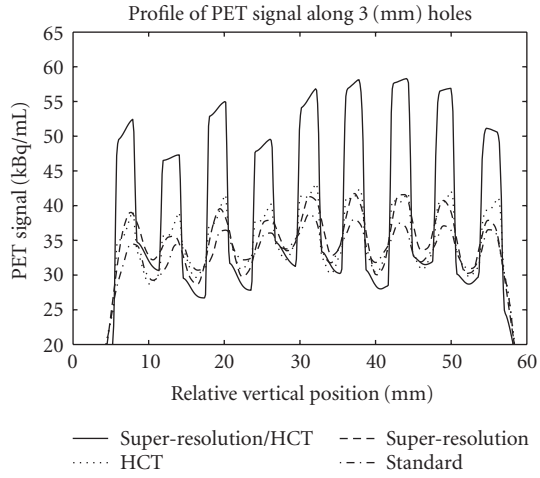


Figure 7: A plot of pixel values through the 3 mm features of the phantom images in Figure 6. The super-resolution (dashed line) and HCT (dotted) profiles give better contrast than the standard method (dashed and dotted). The combination of HCT and super-resolution gives the best contrast (solid black).

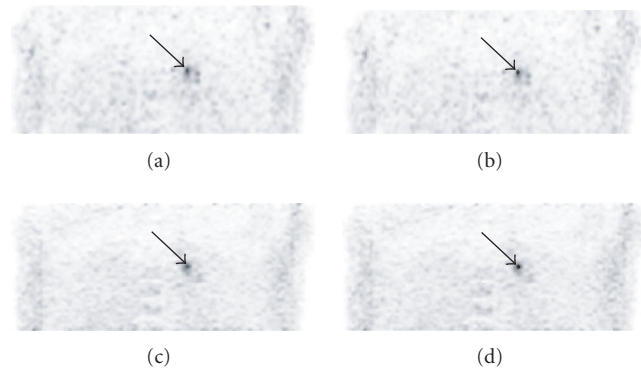


Figure 8: Coronal PET images of the patient through the pulmonary lesion. The black arrow marks the small lesion of interest. (a) Standard 2D whole-body mode acquisition. (b) HCT. The edge of the ^{18}F-FDG uptake is more delineated than in the standard image. (c) Super-resolution. The uptake is more localized than in the standard image. (d) Super-resolution and HCT. The uptake is the most localized in this image.

Table 3: Super-resolution point spread function FWHM values for phantom trials.

Acquisition mode	Axis	Blur kernel of 1 pixel[(a)] (mm)	Gaussian blur kernel of 3.0 mm FWHM (mm)
2D whole body	Axial	4.1	4.0
3D brain	Axial	4.8	4.6
3D brain	Radial	4.4	4.3
3D brain	Tangential	4.3	4.2

[(a)] Previously reported [9].

Table 4: Lesion target-to-background ratios for the PET signals in the 2D whole-body mode acquisition patient trial.

Image type	Transaxial	Coronal	Sagittal	Average
Standard filter (3.0 mm FWHM)	5.5	6.0	6.6	6.1
Super-resolution	6.3	6.3	5.9	6.2
HCT	7.6	7.7	7.4	7.6
HCT and super-resolution	8.1	8.3	7.7	8.0

shows that, using the same data, the inclusion of a Gaussian blur kernel improved the resolution by reducing the FWHM of the PSFs by a difference of 0.1 mm to 0.2 mm compared to previously reported results [9]. The value of the blur kernel used for Table 3 was set to 3.0 mm since this minimized the FWHM of the "point source."

In the 2D whole-body mode, when the number of axial shifts was decreased from 4 POVs (with 1 mm shifts) to 2 POVs (with 2 mm shifts), the axial resolution was degraded from 4.0 mm to 4.3 mm as measured by the FWHM of the axial PSF. The axial resolution of the 2D whole-body case did not improve when 8 POVs with 0.5 mm shifts were used; the FWHM of the axial PSF remained at 4.0 mm.

For the patient study in which the super-resolution acquisition time was the same as that of the standard (4 min total), the lesion of interest could not be resolved due to the low number of counts in each POV. By using a 4 min acquisition time for each POV (a total of 16 min), the super-resolution method clearly resolved the lesion as shown in Figure 8(a). In Figure 8, the filters were selected to achieve the same level of noise in the PET images. By smoothing the images with a 3.0 mm FWHM Gaussian filter (10 HCT iterations for the 0.5 mm $\times$ 0.5 mm pixel size; see (6)) a variance of 0.36 $\pm$ 0.01 kBq^2/mL^2 was maintained in the coronal images excluding a 15 mm diameter circular ROI around the lesion of interest. Table 4 shows that the lesion target-to-background ratios were higher with super-resolution (except for the sagittal image) when compared to the ratios for the standard images. The application of HCT further increased the target-to-background ratios. For the super-resolution acquisition that was processed with HCT, the target-to-background ratios were the highest. They improved to an average of 8.0 (range: 7.7–8.3) when compared to an average of 6.1 (range: 5.5–6.6) for the standard image. Sharper edges and more localized uptake were also depicted in the patient reconstructions using the combination super-resolution and HCT techniques when compared to the other images (Figure 8).

4. DISCUSSION

The super-resolution acquisition and reconstruction meets the goal of obtaining higher resolution in the PET acquisition. Super-resolution has been reported to improve the

axial resolution by 9% to 52% compared to a standard acquisition and by 14% to 16% compared to merely interleaving the acquired slices to the appropriate axial location [9]. As described above, modifying the Irani and Peleg method [8] to include a 3.0 mm blur kernel improves these results by a further 2% to 4% (Table 3), using the same data sets. Similarly, in the 3D brain-mode transaxial images, super-resolution has been reported to improve the resolution by at least 12% [9] and the modified method used here improves that result by a further 2%. The improved resolution due to the super-resolution technique compared to a standard acquisition is evident in the phantom image (Figure 6), in a pixel plot through its 3 mm features, and in the improved contrast ratios (Table 2). This improvement due to the super-resolution acquisition and processing holds true even when the super-resolution results require more smoothing than the standard images to achieve the same level of image noise, as in the phantom case.

In the phantom trial (Figure 6), the application of HCT filtering, as an algorithm for the fusion of PET and CT data, improves contrast ratios by an average of 14% (range: 7–18%) when compared to the standard Gaussian method (Table 2). This is similar to the improvement provided by super-resolution alone (average: 13%, range: 9–15%) and the pixel profiles through the 3 mm phantom features using these two methods roughly match (Figure 7). The application of both methods in tandem provides superior contrast ratios: an average of 54% (range: 45–69%) better than the standard processing method for images with the same level of noise. This increase in contrast is a combination of the reduction of partial volume effects provided by super-resolution [9] and the retention of uptake within established borders when the image is smoothed with HCT. Small features are most evident in the super-resolution/HCT image (Figure 6(d)) and pixel profile (Figure 7) when compared to the other three processing methods.

Although the improvement in the image due to the super-resolution technique and the HCT filtering can be demonstrated with the phantom, the same cannot be said for the patient trial since the true distribution of ^{18}F-FDG is unknown. However, in all but the sagittal image, super-resolution improved the lesion's target-to-background ratio (Table 4). HCT improved the target-to-background ratio by an average of 26% (range: 12–38%). The combined super-resolution/HCT procedure was superior and improved the target-to-background ratio by an average of 34% (range: 17–47%). In the super-resolution/HCT PET image, the uptake is more localized and delineated (Figure 8) as would be desired for small tumor detection.

Unlike the phantom case, in terms of acquisition time, the comparison between standard and super-resolution patient PET acquisitions is not one to one. The super-resolution acquisition and reconstruction for the patient required approximately four times the number of counts as the standard images. (The signal of the lesion of interest was lost due to the low-counting statistics when the total acquisitions times were kept the same.) Using four POVs of 4 min each, this super-resolution example demonstrates that these acquisitions are clinically feasible if restricted to one FOV of interest. When the total acquisition times were kept constant (as in the phantom case) the super-resolution data required more smoothing (Gaussian filters of 3.0 mm FWHM or their HCT equivalent) than the standard data (1.8 mm FWHM). In contrast, the super-resolution data for the patient did not require additional smoothing to obtain the same noise level as in the standard images (Gaussian filters of 3.0 mm FWHM or their HCT equivalent were used for both) because of the increased number of counts in the super-resolution case.

The choice of 4 POVs for the super-resolution technique in the patient case is reasonable. Since the automated bed motion readily provides increments of 0.5 mm, conceivably one could acquire 8 POVs for the super-resolution technique. However, at 4 min per POV the resulting long acquisition time may be prohibitive. On the other hand, keeping the total acquisition time constant renders the number of counts per position too low to be useful, as found in the four 1-minute POVs case. In general it could be stated that there is a minimal acquisition time required for each POV in order to obtain useful information. Hence, the number of POVs multiplied by that minimal acquisition time will determine the needed total acquisition time. The number of POVs used and their corresponding acquisition times has yet to be optimized.

It is worth reiterating from [9] that patient motion will further degrade the efficacy of the super-resolution technique because the registration of the POVs should be known to subpixel accuracy. Consequently, brain scans may be more suitable for the clinical application of super-resolution since the head is then firmly fixed and subject to little motion. Also, the application of this technique in the transverse direction would require a method of recording the geometric shifts of the patient in the transaxial plane. Conceivably, one could envision a new type of scanner with a rotating gantry, and perhaps even with some transaxial motion, that would be able to provide super-resolution without moving the patient.

Applying HCT in the axial direction as presented here is suboptimal since the slice thickness of the CT was automatically set by the scanner to be the same as that of standard PET images. However, the CT scanner can potentially provide thinner slice reconstructions. Using such images as the CT input would reduce partial volume effects and potentially further improve the results.

The improvement in resolution due to super-resolution acquisition and reconstruction and the improvement in contrast ratio using HCT filtering come at a considerable increase in computational time when applied together. Compared to standard processing, the super-resolution technique applied to PET increases processing times by a factor of 23 [9] and HCT filtering increases this by a factor of 8 [11]. On the Discovery-LS scanner used, the reconstruction time of AC PET is typically 2 to 3 min per FOV with most of the reconstruction being performed concurrent with a 20-to-30-minute acquisition of 5 to 7 FOVs per patient. Increasing processing times by factors greater than 8 could not be easily accommodated. Because of this prohibitive increase in computer processing time, the clinical application of the

combined super-resolution/HCT process would likely need suitable dedicated computer hardware or to be restricted to a suspicious region of interest to avoid spending computational resources sharpening the entire data set.

As an alternative to OSEM, one may consider the use of penalized-likelihood image reconstruction methods, as a complementary process to super-resolution. Penalized-likelihood iterative reconstruction algorithms include a penalty (regularization) term which discourages neighboring pixels from converging to widely disparate values [23]. With such an approach, edge information (obtained from another modality) may be introduced via the regularization term [24] or prior [25], and perhaps could replace the HCT processing stage. A disadvantage of using penalized-likelihood methods for emission tomography is that space-invariant penalties result in high-count regions tending to be smoothed more than low-count regions [26], but methods have been developed to give a more uniform spatial resolution [27]. Although not addressed by this paper, it would be worthwhile to try to achieve a similar improvement in resolution for a given variance by combining the super-resolution method with the penalized-likelihood reconstruction methods.

5. CONCLUSION

A new method incorporating two techniques, super-resolution and hybrid computed tomography (HCT), for fusing PET and CT images has been developed and evaluated. A super-resolution acquisition, modified to include a Gaussian blur kernel, has been shown to significantly improve the resolution of the PET acquisition. The feasibility of implementing the method in a clinical PET/CT scanner has been demonstrated by showing higher contrast ratios in a phantom study and higher target-to-background ratios in a small lesion from a patient study for images exhibiting the same level of noise. The resulting reconstructions provide higher resolution metabolic images with delineated edges where corresponding anatomical borders are available.

REFERENCES

[1] S. S. Gambhir, J. Czernin, J. Schwimmer, D. H. S. Silverman, R. E. Coleman, and M. E. Phelps, "A tabulated summary of the FDG PET literature," *Journal of Nuclear Medicine*, vol. 42, no. 5, supplement, pp. 1S–93S, 2001.

[2] M. B. Fukui, T. M. Blodgett, and C. C. Meltzer, "PET/CT imaging in recurrent head and neck cancer," *Seminars in Ultrasound CT and MRI*, vol. 24, no. 3, pp. 157–163, 2003.

[3] G. Tarantola, F. Zito, and P. Gerundini, "PET instrumentation and reconstruction algorithms in whole-body applications," *Journal of Nuclear Medicine*, vol. 44, no. 5, pp. 756–769, 2003.

[4] D. W. Townsend, T. Beyer, P. E. Kinahan, et al., "The SMART scanner: a combined PET/CT tomograph for clinical oncology," in *Proceedings of IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC '98)*, vol. 2, pp. 1170–1174, Toronto, Ontario, Canada, 1998.

[5] T. Beyer, D. W. Townsend, T. Brun, et al., "A combined PET/CT scanner for clinical oncology," *Journal of Nuclear Medicine*, vol. 41, no. 8, pp. 1369–1379, 2000.

[6] R. Bar-Shalom, N. Yefremov, L. Guralnik, et al., "Clinical performance of PET/CT in evaluation of cancer: additional value for diagnostic imaging and patient management," *Journal of Nuclear Medicine*, vol. 44, no. 8, pp. 1200–1209, 2003.

[7] Z. Keidar, N. Haim, L. Guralnik, et al., "PET/CT using ^{18}F-FDG in suspected lung cancer recurrence: diagnostic value and impact on patient management," *Journal of Nuclear Medicine*, vol. 45, no. 10, pp. 1640–1646, 2004.

[8] M. Irani and S. Peleg, "Motion analysis for image enhancement: resolution, occlusion, and transparency," *Journal of Visual Communication & Image Representation*, vol. 4, no. 4, pp. 324–335, 1993.

[9] J. A. Kennedy, O. Israel, A. Frenkel, R. Bar-Shalom, and H. Azhari, "Super-resolution in PET imaging," *IEEE Transactions on Medical Imaging*, vol. 25, no. 2, pp. 137–147, 2006.

[10] H. Azhari and S. Stolarski, "Hybrid ultrasonic computed tomography," *Computers and Biomedical Research*, vol. 30, no. 1, pp. 35–48, 1997.

[11] J. A. Kennedy, H. Azhari, A. Frenkel, et al., "PET/CT Image Data Fusion in Hybrid Scanners," Abstract, Society of Nuclear Medicine 52nd Annual Meeting Abstract Book, 2005 June; Toronto, Canada, p. 167P.

[12] H. Greenspan, G. Oz, N. Kiryati, and S. Peled, "Super-resolution in MRI," in *Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI '02)*, pp. 943–946, Washington, DC, USA, July 2002.

[13] H. Greenspan, G. Oz, N. Kiryati, and S. Peled, "MRI inter-slice reconstruction using super-resolution," *Magnetic Resonance Imaging*, vol. 20, no. 5, pp. 437–446, 2002.

[14] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.

[15] T. R. DeGrado, T. G. Turkington, J. J. Williams, C. W. Stearns, J. M. Hoffman, and R. E. Coleman, "Performance characteristics of a whole-body PET scanner," *Journal of Nuclear Medicine*, vol. 35, no. 8, pp. 1398–1406, 1994.

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

[17] M. Defrise, P. E. Kinahan, D. W. Townsend, C. Michel, M. Sibomana, and D. F. Newport, "Exact and approximate rebinning algorithms for 3-D PET data," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 145–158, 1997.

[18] E. Kamel, T. F. Hany, C. Burger, et al., "CT vs ^{68}Ge attenuation correction in a combined PET/CT system: evaluation of the effect of lowering the CT tube current," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 29, no. 3, pp. 346–350, 2002.

[19] T. Beyer, P. E. Kinahan, D. W. Townsend, and D. Sashin, "The use of x-ray CT for attenuation correction of PET data," in *Proceedings of IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS-MIC '95)*, vol. 4, pp. 1573–1577, San Francisco, Calif, USA, October-November 1995.

[20] P. E. Kinahan, D. W. Townsend, T. Beyer, and D. Sashin, "Attenuation correction for a combined 3D PET/CT scanner," *Medical Physics*, vol. 25, no. 10, pp. 2046–2053, 1998.

[21] C. Burger, G. Goerres, S. Schoenes, A. Buck, A. H. R. Lonn, and G. K. von Schulthess, "PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 29, no. 7, pp. 922–927, 2002.

[22] J. Canny, "A computational approach for edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[23] J. A. Fessler, "Penalized weighted least-squares image reconstruction for positron emission tomography," *IEEE Transactions on Medical Imaging*, vol. 13, no. 2, pp. 290–300, 1994.

[24] J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Regularized emission image reconstruction using imperfect side information," *IEEE Transactions on Nuclear Science*, vol. 39, no. 5, part 1, pp. 1464–1471, 1992.

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

[26] J. A. Fessler and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction: space-invariant tomographs," *IEEE Transactions on Image Processing*, vol. 5, no. 9, pp. 1346–1358, 1996.

[27] J. W. Stayman and J. A. Fessler, "Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 19, no. 6, pp. 601–615, 2000.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 93479, 10 pages
doi:10.1155/2007/93479

Research Article

Registration of Brain MRI/PET Images Based on Adaptive Combination of Intensity and Gradient Field Mutual Information

Jiangang Liu[1] and Jie Tian[1,2]

[1] *Medical Image Processing Group, Key Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Science, P.O. Box 2728, Beijing 100080, China*
[2] *Life Science Center, Xidian University, Xi'an, Shaanxi 710071, China*

Received 10 May 2006; Revised 13 January 2007; Accepted 22 January 2007

Recommended by Robert R. Edelman

Traditional mutual information (MI) function aligns two multimodality images with intensity information, lacking spatial information, so that it usually presents many local maxima that can lead to inaccurate registration. Our paper proposes an algorithm of adaptive combination of intensity and gradient field mutual information (ACMI). Gradient code maps (GCM) are constructed by coding gradient field information of corresponding original images. The gradient field MI, calculated from GCMs, can provide complementary properties to intensity MI. ACMI combines intensity MI and gradient field MI with a nonlinear weight function, which can automatically adjust the proportion between two types MI in combination to improve registration. Experimental results demonstrate that ACMI outperforms the traditional MI and it is much less sensitive to reduced resolution or overlap of images.

1. INTRODUCTION

Multimodal image registration plays a significant role in medical image processing such as medical diagnosis, therapeutic planning and assessment [1]. MI has proved to be an effective criterion for the multimodal medical image registration [1–3]. However, even with this method, the correct alignment cannot be guaranteed, especially when it is applied to images with low resolution or small overlapped area. MI function traditionally relies on only intensity information of images, lacking sufficient spacial information, so it has difficulty in accurately measuring the degree of alignment of two images. It is also apt to be influenced by intensity interpolation, therefore presents many local maxima which frequently lead to misregistration [4, 5].

Different tissues in human brain usually present different gray intensity no matter which imaging modality is applied to them. The intensity gradient at the transition of two tissues is steeper than the interior, where the gradient magnitude and phase lie on the imaging modality, and the spatial relative position is invariable. Therefore the gradient fields of two images can provide effective spatial information for their similarity measurement.

Some research introduced gradient or feature information into multimodal image registration to improve the performance of registration function. Butz and Thiran [6] performed the registration with MI based on feature space; Pluim et al. [7] integrated gradient information into mutual information to get a relatively smooth registration function; Haber and Modersitzki [8] presented an alternative similarity measurement based on normalized gradient field for multimodal image registration; Maintz et al. [9] showed that image intensity gradient was an effective multimodal image registration criteria. These methods were effective for the improvement of registration quality. Our work took full advantage of the gradient phase information and the relationship between intensity images and their gradient fields to further improve the performance of MI function.

Our current study proposes a technique for Multimodal image registration, namely adaptive combination of intensity and gradient field mutual information (ACMI). Unlike the intensity MI computed with original images, the gradient field MI is calculated with gradient code maps (GCM) which were obtained from corresponding original images by a spherical gradient coder. The intensity of each point in GCM is completely determined by thegradient vector of

corresponding point in original intensity image, so that the magnitude and phase information of spatial gradient field of original image is converted into intensity information of GCM. The properties of these two MI functions are complementary for each other and the ACMI is defined as the sum of products of each MI function and corresponding weighting function. The weighting function can be adaptively regulated to highlight the contribution of MI function with better performance to ACMI.

The simulated data experiment and the actual registration experiment were conducted to compare the performance of ACMI and traditional MI. The results of simulated data experiment showed that ACMI function was much smoother and more reliable than traditional MI. The statistical test for the results of actual registrations demonstrated that the registration quality with ACMI was significantly higher than that with traditional MI and it was much less sensitive to the reduction of resolution or overlapped region of images.

2. METHODS

2.1. Mutual information

Given reference image R and floating image F with their respective marginal intensity distributions p_R, p_F and joint intensity distribution p_{RF}, their MI is defined by means of Kullback-Leibler measure [3]:

$$I(R,F) = \sum_{i,j} p_{RF}(i,j) \log \frac{p_{RF}(i,j)}{p_R(i)p_F(j)}. \tag{1}$$

The MI criterion postulates that the images are geometrically aligned when $I(A,B)$ is maximal. However, this is not always true, because many local maxima are frequently announced and sometimes even worse, the global maximum does not correspond to the correct alignment [5].

MI is also defined by means of information theoretic notion of entropy. Given images R and F with their respective entropies $H(R)$ and $H(F)$ and their joint entropy $H(R,F)$, their MI $I(R,F)$ is defined as

$$I(R,F) = H(R) + H(F) - H(R,F). \tag{2}$$

In fact, MI is a measure method based on statistical notion, and its reliability depends on the number of voxels included in its computation. It is sensitive to image resolution or the overlapped area of two images. To solve this problem, some MI-related measures have been proposed such as entropy correlation coefficient (ECC) $E(R,F)$ [10]

$$E(R,F) = \frac{2I(R,F)}{H(R) + H(F)} \tag{3}$$

and normalized mutual information NMI $N(R,F)$ [11]

$$N(R,F) = \frac{H(R) + H(F)}{H(R,F)}. \tag{4}$$

The same artifact patterns as MI function are found in both ECC and NMI [4]. It does not guarantee an accurate and reliable alignment. Comparing (2) and (3), for fixed images R and F, their respective entropies $H(R)$ and $H(F)$ are approximately constant, so the EEC is in fact the product of $I(R,F)$ and a constant. As a result, ECC has the similar performance to MI except for its normalized value range [0, 1] [1]. Therefore in the following analysis, ECC is used in place of corresponding MI.

2.2. Spatial gradient field code

Given a 2D image F with intensity $f(x,y)$, its spatial gradient field $G_F(x,y)$ can be computed by

$$G_F(x,y) = \frac{\partial f(x,y)}{\partial x}\vec{i} + \frac{\partial f(x,y)}{\partial y}\vec{j}, \tag{5}$$

where the $\vec{i}$ and $\vec{j}$ are the unit vectors along x and y axes, respectively. If the horizontal and vertical derivatives, namely $\partial f(x,y)/\partial x$ and $\partial f(x,y)/\partial y$, are denoted by f_x and f_y, respectively, then the magnitude $\rho_{i,j}$ and phase $\theta_{i,j}$ ($[0, 2\pi)$) of gradient of voxel $f_{i,j}$ are calculated by

$$\begin{gathered}
\gamma_{i,j} = \sqrt{f_x^2 + f_y^2}, \\
\gamma_{\max} = \max\left(\gamma_{i,j}\right), \\
\rho_{i,j} = \begin{cases} \dfrac{\gamma_{i,j}}{\gamma_{\max}} & \gamma_{i,j} < \gamma_{\max}, \\ 0.999 & \gamma_{i,j} = \gamma_{\max}, \end{cases} \\
\theta_{i,j} = \begin{cases} \tan^{-1}\dfrac{f_y}{f_x} & f_x > 0,\ f_y > 0, \\ \tan^{-1}\dfrac{f_y}{f_x} + \pi & f_x < 0, \\ \tan^{-1}\dfrac{f_y}{f_x} + 2\pi & f_x > 0,\ f_y < 0, \\ \dfrac{\pi}{2} & f_x = 0,\ f_y > 0, \\ \dfrac{3\pi}{2} & f_x = 0,\ f_y < 0. \end{cases}
\end{gathered} \tag{6}$$

The ranges of $\rho_{i,j}$ and $\theta_{i,j}$ are [0,1) and $[0, 2\pi)$, respectively. The point $c_{i,j}$ in GCM C is obtained by coding the gradient vector $(\rho_{i,j}, \theta_{i,j})$ of corresponding point $f_{i,j}$ in F with gradient coder [12],

$$c_{i,j} = \begin{cases} \left\lfloor \dfrac{\rho_{i,j}}{\Delta_p} \right\rfloor N + \left\lfloor \dfrac{\theta_{i,j}}{\Delta_\theta} \right\rfloor & \rho_{i,j} \geq \text{Th}, \\ 0 & \rho_{i,j} < \text{Th}, \end{cases} \tag{7}$$

where $\lfloor \rho_{i,j}/\Delta_p \rfloor$ is the integer portion of $\rho_{i,j}/\Delta_p$ and Th is a prespecified threshold to ignore the point with low gradient magnitude. Δ_ρ and Δ_θ are, respectively, the magnitude and phase bin intervals of gradient coder, and N equals to $2\pi/\Delta_\theta$. Figure 1 illustrates a 2D gradient coder. It converts gradient difference of points in gradient field into intensity difference of points in GCM. Given two gradient vectors with equal magnitude, the one with more phase $(\theta_{i,j})$ will produce stronger intensity in GCM.Table 1 shows an example of

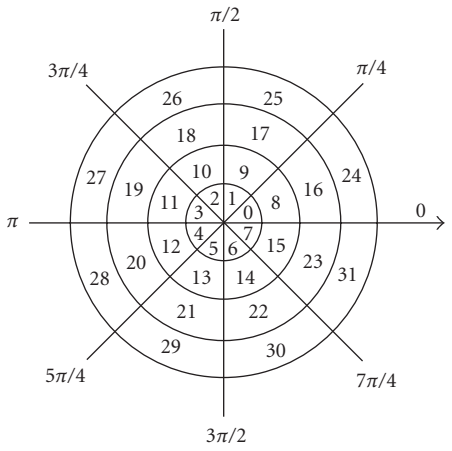


Figure 1: Illustration of a 2D gradient coder ($\Delta_\rho = 1/4$, $\Delta_\theta = \pi/4$).

Table 1: Mapping from gradient field of intensity image to GCM ($\Delta_\rho = 1/4$, $\Delta_\theta = \pi/4$).

Range	[0,0.25)	[0.25,0.5)	[0.5,0.75)	[0.75,1)
$[0, \pi/4)$	0	8	16	24
$[\pi/4, \pi/2)$	1	9	17	25
$[\pi/2, 3\pi/4)$	2	10	18	26
$[3\pi/4, \pi)$	3	11	19	27
$[\pi, 5\pi/4)$	4	12	20	28
$[5\pi/4, 3\pi/2)$	5	13	21	29
$[3\pi/2, 7\pi/4)$	6	14	22	30
$[7\pi/4, 2\pi)$	7	15	23	31

mapping from gradient field of intensity image to GCM. The most left column and the top row of Table 1 correspond to phase angle and magnitude of gradient vector, respectively. For example, two pixels with gradient vectors $(0.3, \pi/8)$ and $(0.3, 5\pi/8)$ will produce intensity values 8 and 10 in GCM, respectively. The gradient field map includes both magnitude and phase information, so it can provide more spatial information for the similarity measurement of two images.

This gradient coder can be easily extended to 3D images. Given a voxel $f_{i,j,k}$ of 3D image F with gradient vectors (f_x, f_y, f_z), the 3D gradient coder is defined as

$$c_{i,j,k} = \begin{cases} \left\lfloor \dfrac{\rho_{i,j,k}}{\Delta_p} \right\rfloor NK + \left\lfloor \dfrac{\phi_{i,j,k}}{\Delta_\phi} \right\rfloor K + \left\lfloor \dfrac{\theta_{i,j,k}}{\Delta_\theta} \right\rfloor & \rho_{i,j,k} \geq \text{Th}, \\ 0 & \rho_{i,j,k} < \text{Th}, \end{cases} \tag{8}$$

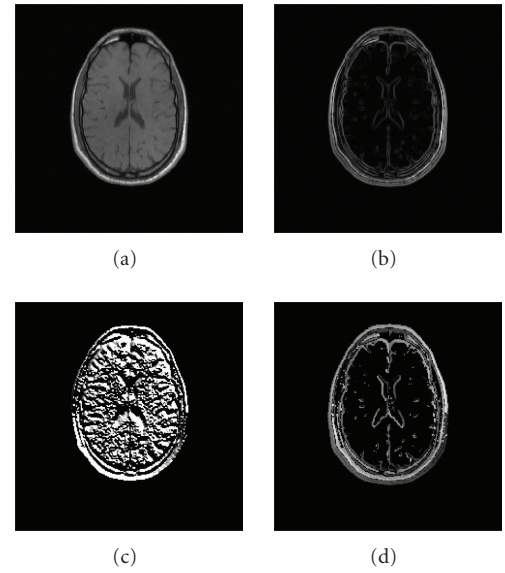


Figure 2: A slice of 3D GCM of MRI T1. (a) Original image, (b) gradient magnitude map, (c) gradient phase map, (d) GCM ($\Delta_\rho = 1/16$, $\Delta_\theta = \pi/8$, $\Delta_\phi = \pi/8$, and Th = 0.10).

with

$$\begin{aligned} \gamma_{i,j,k} &= \sqrt{f_x^2 + f_y^2 + f_z^2}, \\ \gamma_{\max} &= \max\left(\gamma_{i,j,k}\right), \\ \rho_{i,j,k} &= \begin{cases} \dfrac{\gamma_{i,j,k}}{\gamma_{\max}} & \gamma_{i,j,k} < \gamma_{\max}, \\ 0.999 & \gamma_{i,j,k} = \gamma_{\max}, \end{cases} \\ \phi_{i,j,k} &= \cos^{-1} \frac{f_z}{\gamma_{i,j,k}}, \\ \theta_{i,j,k} &= \begin{cases} \tan^{-1} \dfrac{f_y}{f_x} & f_x > 0,\ f_y > 0, \\ \tan^{-1} \dfrac{f_y}{f_x} + \pi & f_x < 0, \\ \tan^{-1} \dfrac{f_y}{f_x} + 2\pi & f_x > 0,\ f_y < 0, \\ \dfrac{\pi}{2} & f_x = 0,\ f_y > 0, \\ \dfrac{3\pi}{2} & f_x = 0,\ f_y < 0, \end{cases} \end{aligned} \tag{9}$$

where $\phi_{i,j,k}([0, \pi])$ and $\theta_{i,j,k}([0, 2\pi))$ are the polar angle and azimuthal angle, respectively. N and K equal to π/Δ_ϕ and $2\pi/\Delta_\theta$, respectively, where Δ_ϕ and Δ_θ are the polar angle and azimuthal angle bin intervals of 3D gradient coder, respectively. The other notations are defined similarly as in (6). With (8), spatial gradient field (magnitude and phase) information of original images is converted into intensity information of GCMs. Figures 2 and 3 show a slice of GCMs of 3D MRI T1 and PET images, respectively.

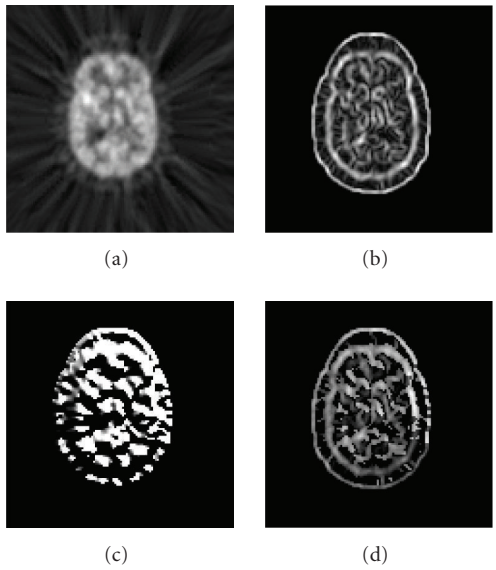


Figure 3: A slice of 3D GCM of PET. (a) Original image, (b) gradient magnitude map, (c) gradient phase map, (d) GCM ($\Delta_\rho = 1/16$, $\Delta_\phi = \pi/8$, $\Delta_\theta = \pi/8$, and Th = 0.16).

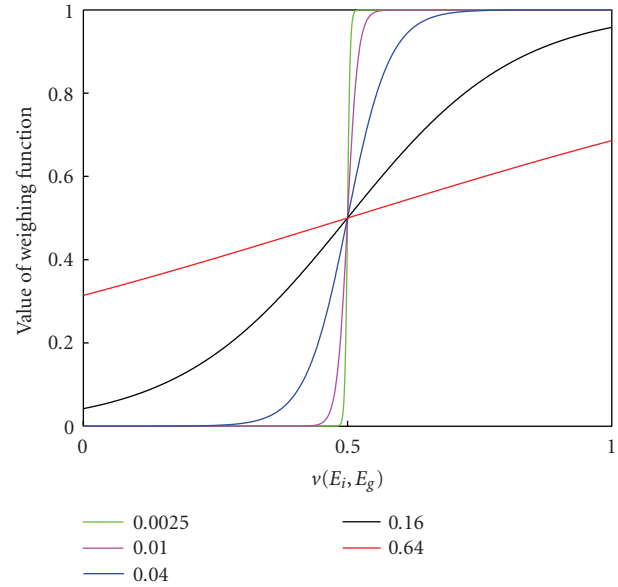


Figure 4: Weighing functions $f(v(E_i, E_g))$ with $T = 0.0025$ (green), $T = 0.01$ (magenta), $T = 0.04$ (blue), $T = 0.16$ (black) and $T = 0.64$ (red), respectively.

2.3. *Adaptive combination of intensity and gradient field mutual information (ACMI)*

With (3), the intensity ECC E_i is obtained from two original images and the gradient ECC E_g is computed from their GCMs. The ACMI E_a is defined as

$$E_a = f(v(E_i, E_g))E_i + (1 - f(v(E_i, E_g)))E_g, \tag{10}$$

with

$$f(v(E_i, E_g)) = \frac{1}{1 + \exp\left(-(v(E_i, E_g) - 0.5)/T\right)}, \qquad v(E_i, E_g) = \frac{E_i + E_g}{2} \quad 0 \le E_i,\ E_g \le 1. \tag{11}$$

As shown in Figure 4, the weighting function $f(v(E_i, E_g))$ actually is a logistic function with rightward half unit shift. This function has some expected properties [13]. The first is the saturation with the maximum of one and the minimum of zero. This property is very important for the weighting function because, as described by (10), the output of $f(v(E_i, E_g))$ presents a fraction whose value extends from zero to one. The second is differentiability which not only prevents introducing additional local maxima during combination of registration functions but also facilitates the optimization of ACMI with some derivative-needed techniques such as Gauss-Newton optimization method. The third, the most important one, is the nonlinearity. As indicated in Figure 4, the weighting function presents nonlinear characteristic in two terminal saturating parts but approximate linear characteristic in the middle nonsaturating part. Thus, according to (10), the ACMI is mostly determined by one of two ECC functions at each nonlinear terminal (gradient ECC for the left terminal and intensity ECC for the right), but equals to the sum of two ECC functions with similar weights in the middle linear part. The nonlinear degree of weighting function totally depends on the time constant T. If it is too large, for example 0.64 (Figure 4, the red line), the weighting function will present more linear characteristic. As a result, the unexpected middle linear part is lengthened and the weight-similar sum of registration functions can lead to severe roughness of ACMI. Decrease of time constant can shorten this unexpected middle linear part and lengthen the terminal saturating parts. On the other hand, the extremely small time constant, for example 0.0025 (Figure 4, green line), can impair the differentiability of weighting function and introduce additional local maxima when combining two registration functions. The optimal choice for T was obtained by experimental method (described in Section 3). In terms of these three properties, $f(v(E_i, E_g))$ is a desirable weighting function for combination of registration functions.

As shown in Figures 2(a) and 3(a), the original images contain abundant information. It has two effects on their ECC function. On the one hand, the similarity measurement of two images is more reliable because of abundant information and the ECC function presents a tendency of convergence to the global maximum which corresponds to the correct alignment. On the other hand, abundant information means strong nonuniformity of intensity across voxels, then the ECC function is easily influenced by intensity interpolation, and thereby presents many local maxima which can lead to inaccurate registration (see Figures 6(a), 6(d), and 6(g)).

Compared to the original images, as shown in Figures 2(d) and 3(d), GCMs contain less information (most voxels have zero intensity value except those at edges of some

tissues). This relatively higher-intensity uniformity reduces the effect of intensity interpolation on ECC function of two maps and therefore makes it smoother [4]. Additionally the edge information in GCMs can provide reliable and accurate spatial information for the similarity measurement of images. However, in the neighborhood of global maximum, the ECC function often presents plateaus or valleys, preventing convergence to global maximum (see Figures 6(b), 6(e), and 6(h)).

According to their complementary properties, these two ECC functions are combined by ACMI, using a nonlinear weighting function (10). In our study, the downhill simplex optimization [14] was used for the search of maximum in six-dimensional space (translations along x, y, and z axes, and rotations around x, y, and z axes). This algorithm is an efficient method for N-dimensional unconstrained minimization [15, 16]. It begins with $N + 1$ vertices which define a simplex in N-dimensional space and attempts to move them into the minimum. Given reference image R, floating image F, and transforming parameter vector (vertex) $x_j^{(k)}$ ($j = 1, 2, \ldots, 7$) in kth iteration, $TF(x_j^{(k)})$ denotes the transformed F with $x_j^{(k)}$, and $Ei_j^{(k)}$, $Eg_j^{(k)}$ and $Ea_j^{(k)}$ denote intensity ECC, gradient ECC and ACMI of R, and $TF(x_j^{(k)})$, respectively. The iterative procedure is the following [16].

Step 1. Initialize $x_j^{(k)}$ and calculate $Ea_j^{(k)}$.

Step 2. Order $x_j^{(k)}$ to satisfy $Ea_1^{(k)} \le Ea_2^{(k)} \le \cdots \le Ea_7^{(k)}$, and calculate the centroid of the six best ACMI values, $\overline{x}^{(k)} = \sum_{j=1}^{6} x_j^{(k)}/6$ and $\overline{Ea}^k$.

Step 3. $x_r^{(k)} = \overline{x}^{(k)} + (\overline{x}^{(k)} - x_7^{(k)})$ and calculate $Ea_r^{(k)}$.

Step 4. If $Ea_1^{(k)} \le Ea_r^{(k)} < Ea_6^{(k)}$, then $x_7^{(k)} = x_r^{(k)}$, $Ea_7^{(k)} = Ea_r^{(k)}$ and go to Step 9.

Step 5. If $Ea_r^{(k)} < Ea_1^{(k)}$, then $x_e^{(k)} = \overline{x}^{(k)} + 2(x_r^{(k)} - \overline{x}^{(k)})$, and calculate $Ea_e^{(k)}$. If $Ea_e^{(k)} < Ea_r^{(k)}$, then $x_7^{(k)} = x_e^{(k)}$, $Ea_7^{(k)} = Ea_e^{(k)}$, and go to Step 9; otherwise $x_7^{(k)} = x_r^{(k)}$, $Ea_7^{(k)} = Ea_r^{(k)}$, and go to Step 9.

Step 6. If $Ea_6^{(k)} \le Ea_r^{(k)} < Ea_7^{(k)}$, then $x_c^{(k)} = \overline{x}^{(k)} + 0.5(x_r^{(k)} - \overline{x}^{(k)})$, and calculate $Ea_c^{(k)}$. If $Ea_c^{(k)} \le Ea_r^{(k)}$, then $x_7^{(k)} = x_c^{(k)}$, $Ea_7^{(k)} = Ea_c^{(k)}$, and go to Step 9; otherwise go to Step 8.

Step 7. If $Ea_r^{(k)} \ge Ea_7^{(k)}$, then $x_{cc}^{(k)} = \overline{x}^{(k)} - 0.5(\overline{x}^{(k)} - x_7^{(k)})$ and calculate $Ea_{cc}^{(k)}$. If $Ea_{cc}^{(k)} < Ea_7^{(k)}$, then $x_7^{(k)} = x_{cc}^{(k)}$, $Ea_7^{(k)} = Ea_{cc}^{(k)}$, and go to Step 9; otherwise, go to Step 8.

Step 8. $x_j^{(k)} \leftarrow x_j^{(k)} + 0.5(x_1^k - x_j^{(k)})$.

Step 9. $c = \{(1/7)\sum_{j=1}^{7}[Ea_j^{(k)} - \overline{Ea}^{(k)}]^2\}^{1/2}$. If $c < 10^{-4}$, then the iteration exists; otherwise $k \leftarrow k + 1$, and go to Step 2.

In each iteration, the ACMI $Ea_j^{(k)}$ for each transforming parameter vector $x_j^{(k)}$ is calculated as follows.

Step 1. Transform F into $TF(x_j^{(k)})$ with transforming parameters vector $x_j^{(k)}$.

Step 2. Calculate intensity ECC $Ei_j^{(k)}$ and gradient ECC $Eg_j^{(k)}$ of R and $TF(x_j^{(k)})$.

Step 3. Adjust weighting function according to $Ei_j^{(k)}$ and $Eg_j^{(k)}$.

Step 4. Calculate ACMI $Ea_j^{(k)}$.

When alignment improves, ACMI uses an iterative algorithm to automatically adjust the proportion between intensity ECC and gradient ECC by changing the weighting function. Thus at the coarse registration stage, the ACMI depends mostly on gradient ECC due to the low sum of ECC values and presents a smooth property facilitating the convergence to the basin of global maximum. With the two images increasingly aligned, the $\nu(E_i, E_g)$ becomes larger due to the increase of the values of intensity ECC E_i and gradient ECC E_g (10). Accordingly, as indicated in Figure 4 (blue line), the weighting function shifts from the left saturating terminal to the right. At the fine stage where the gradient ECC varies slightly, the ACMI is determined mostly by intensity ECC for which the gradient ECC is a supplement. The higher the sum of ECCs is, the more reliable the intensity ECC is, therefore the more the ACMI depends on it than on gradient ECC. This coarse-to-fine and gradient-to-intensity strategy facilitates the convergence to global maximum which corresponds to correct alignment.

3. RESULTS

The brain image set used in the current study includes 35 brain MRI/PET image pairs (MRI T1, PD, T2, and their respective rectified versions versus PET). The brain image set and the standard transformations of each image pair were provided as the part of the project, "*Retrospective Image Registration Evaluation*," Vanderbilt University, Nashville, TN [17]. The accuracy of each registration was evaluated by bone-marker-based gold standard, and the registration error was defined as the error distances between the gold standard in the reference image and the centroid of volume of interest (VOI) in the floating image after alignment (see [17] for more details). To compare performance of traditional ECC and ACMI for image pairs with low resolution or small overlapped area, two types of image pairs were generated from each original image pairs, namely subsampled version (subsampled by a factor of two in three axes, resp.) and small-overlapped version (50% overlapped region of original pairs).

In our study, the threshold Th for each type image is 0.10 for MRI T1, 0.08 for MRI T2, 0.12 for MRI PD, and 0.16 for PET. The thresholds were obtained by the basic global thresholding mehod [18].

As indicated by (8), the smaller the bin intervals of magnitude and phase of gradient coder are, the more gray levels

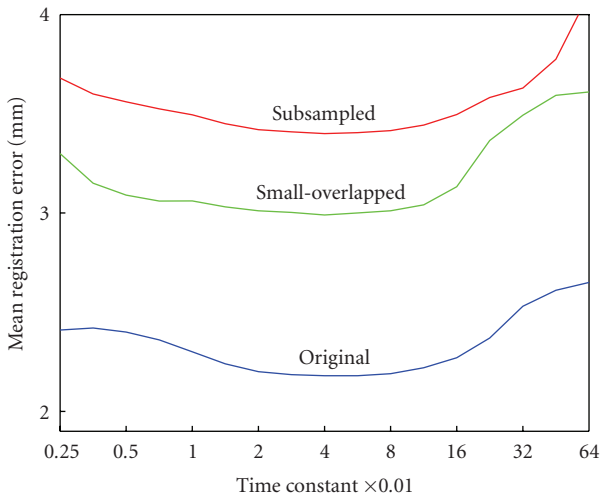


FIGURE 5: The mean registration error (RE) versus time constant for MRI/PET image pairs (blue), subsampled version (red) and small-overlapped version (green). The registration error is defined as the error distances between the gold standard in the reference image and the centroid of volume of interest (VOI) in the floating image after alignment. The horizontal axis is labeled with intervals of $\log_2$. For each image version, the registration error reaches the minimum near the time constant 0.04 where each is relatively insensitive to the changes of time constant.

the GCM has, and accordingly the more accurate registration result will be obtained. However, the increase of gray levels can decrease the statistic power of joint histogram from which the mutual information of two GCMs is calculated [4, 5]. Usually the overmuch gray levels of GCM are subsampled with a suitable bin interval width when calculating the joint histogram. So the extremely small bin intervals of magnitude and phase of gradient coder cannot improve the quality of registration. In our study, the magnitude, polar, and azimuthal angle bin intervals (Δ_ρ, Δ_ϕ, and Δ_θ) are 1/16, $\pi/8$, and $\pi/8$, respectively. Thus the gray level of GCM is 2048 ($16 \times 8 \times 16$) which is enough to identify the changes of magnitude and phase of intensity images, and it is subsampled to 128 gray levels in the calculation of joint histogram.

The optimal choice of time constant was obtained by "bootstrap" method. Fifty "bootstrap" data sets were created by randomly selecting 20 MRI/PET image pairs 50 times with replacement from the brain image set. For a given "bootstrap" data set, registration was applied to each image pair using ACMI with each of 17 different time constants (from 0.0025 to 0.64). For a given time constant T_j, the mean registration error across all "bootstrap" data sets was obtained by $\theta_j = (1/50)\sum_{i=1}^{50} m_{ji}$, where m_{ji} denoted the median registration error across image pairs in D_i for T_j. Figure 5 shows mean registration error across all "bootstrap" data sets versus time constant for MRI/PET image pairs (blue), and their subsampled version (red) and small-overlapped version (green). As presented by Figure 5, mean registration error for each image version reaches the minimum near the time constant 0.04 where it is relatively insensitive to changes in the time constant.

Figure 6 describes three types of registration functions of a PET/MRI T1 pair, namely intensity ECC (traditional ECC), gradient ECC, and their adaptive combination ACMI versus relative displacements between reference and floating images in horizontal and vertical orientations. The negative registration functions are displayed for visual convenience. For the original images, the ACMI performs slightly better than intensity and gradient ECC (top row of Figure 6). For the subsampled version, the intensity ECC presents a global maximum corresponding to the correct alignment, but it still has many local maxima, especially a secondary maximum near the global maximum (Figure 6(d)). The gradient ECC presents less local maxima, but a valley at the bottom (Figure 6(e)). For small-overlapped version, the intensity ECC is strongly rough, though it presents only a global maximum corresponding to the correct alignment (Figure 6(g)). The gradient ECC is smoother, but presents a plateau at bottom not including the correct registration point (Figure 6(h)). Even the global optimization method such as simulated annealing or genetic algorithm is applied to these versions, the correct alignment is not guaranteed. ACMI overcomes these problems (Figures 6(f) and 6(i)). It provides sufficient smoothness in the coarse registration stage due to the dominance of gradient ECC. On the other hand, it presents only one global maximum at coordinate (0, 0) of the graph (corresponding to the correct alignment) in the fine registration stage because of the combination of larger part of intensity information (intensity ECC).

For each image pair, the registrations were applied to its three types of versions with traditional ECC, gradient ECC, and ACMI, respectively. Table 2 summarizes the results of registration. The values in the left three columns labeled with "Accuracy" are the median/maximal registration error for intensity ECC, gradient ECC, and ACMI, respectively, and the values in the right columns labeled with "Number of Iterations" are the mean and standard deviation of number of iterations (Mean ± SD) for intensity ECC, gradient ECC, and ACMI, respectively.

Relative to intensity ECC and gradient ECC, ACMI provides 20.7% and 11.3% mean reduction in error for original version, 19.6% and 8.1% for subsampled version, and 22.5% and 10.8% for small overlapped version, respectively. For each of three image versions (original, subsampled, and small-overlapped version), a paired Student's t-test on ECC types reveals that the results of gradient ECC are significantly more accurate than those of intensity ECC ($P < .01$ for original version, $P < .0005$ for subsampled version and small-overlapped version) but less accurate than those of ACMI ($P < .01$ for original version, $P < .001$ for subsampled version and small-overlapped version). For each of ECC types (intensity ECC, gradient ECC, and ACMI), a one-way analysis of variance (ANOVA) on three image versions finds significant difference for intensity ECC ($P < .001$) and for gradient ECC ($P < .005$), but not for ACMI ($P > .05$). This reveals that ACMI function is much less sensitive to the reduction of resolution or overlapped area of images than intensity ECC and gradient ECC. Also, subpixel accuracy is obtained in all registrations with ACMI. As indicated by Table 2, gradient ECC

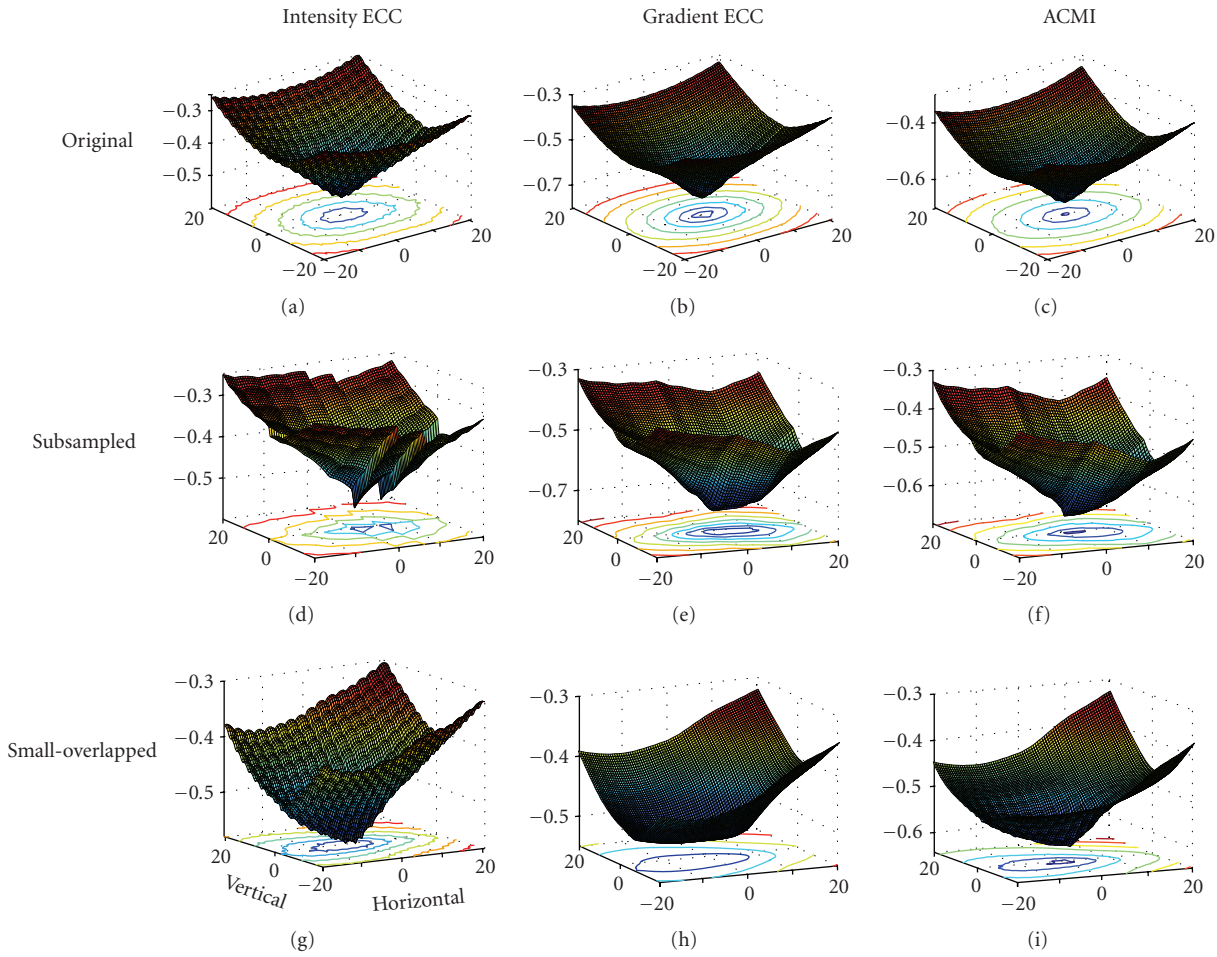


Figure 6: Registration function (PET/MRI T1) versus relative displacement between reference and floating images in horizontal and vertical orientations. For the original images, the ACMI performs slightly better than intensity and gradient ECC (top row). For the subsampled version, the intensity ECC presents a global maximum corresponding to the correct alignment, but also presents many local maxima, especially a secondary maximum near the global maximum (d). The gradient ECC presents less local maxima, but a valley at the bottom (e). For small-overlapped version, the intensity ECC is strongly rough, though it presents only a global maximum corresponding to the correct alignment (g). The gradient ECC is smoother, but presents a plateau at bottom not including the correct registration point (h). ACMI of each image version provides sufficient smoothness in the coarse registration stage and presents only one global maximum at coordinate (0, 0) of the graph (corresponding to the correct alignment) in the fine registration stage ((c), (f), and (i)).

performed better than intensity ECC, but it cannot achieve the optimal registration with absence of intensity information. In the fine stage of registration, the intensity information is required to further improve the registration quality. Table 2 also provides a comparison of number of iterations among intensity ECC, gradient ECC, and ACMI. As summarized by Table 2, the number of iterations was most for the intensity ECC, much more for ACMI, and least for the gradient ECC. Relative to intensity ECC and gradient ECC, ACMI provided 31.9% mean reduction but 69.8% mean increase in the number of iterations for original version, 21.4% mean reduction but 47.1% mean increase for subsampled version, and 31.9% mean reduction but 30.5% mean increase for small overlapped version, respectively. Taken together, ACMI outperforms intensity ECC in terms of accuracy and speed of registration. It can also provide a more accurate result but cost more processing time than gradient ECC. As for the fact that processing time of registration is not crucial due to highly developed computer, ACMI is preferred over gradient ECC except for a required online registration.

Figure 7 illustrates the registration results of three versions of MRI T1/PET image pairs with the intensity (traditional) ECC, gradient ECC, and ACMI. For the convenient display, the extracted brain of MRI T1 image (Figure 7 left of (a)) and the 50% transparent profile of extracted brain of PET image (Figure 7 right of (a)) are served as the underlying and the overlying, respectively. These examples clearly show that when the resolution or the overlapped area of images reduces, the intensity ECC and the gradient ECC can easily lead to misregistration, but the ACMI performs well (Figure 3).

Table 2: Accuracy and iterative number comparison among intensity ECC, gradient ECC, and ACMI.

Pair mode	Pairs	Accuracy (median/maximum mm)			Number of iterations (Mean ± SD)		
		Intensity ECC	Gradient ECC	ACMI	Intensity ECC	Gradient ECC	ACMI
Original images							
T1-PET	7	2.78/4.96	2.73/4.31	2.38/4.22	269.5 ± 84.8	111.3 ± 29.9	172.2 ± 44.0
T2-PET	7	1.91/6.37	1.52/5.84	1.39/5.51	288.0 ± 145.1	87.5 ± 45.8	161.7 ± 43.9
PD-PET	7	2.46/6.19	1.97/4.72	1.62/3.41	222.2 ± 108.7	93.2 ± 36.9	159.9 ± 58.6
T1 rec-PET	4	3.19/8.37	2.88/6.12	2.63/5.43	240.5 ± 92.2	102.4 ± 33.1	145.5 ± 49.7
T2 rec-PET	5	3.15/9.15	3.04/7.13	2.72/6.08	220.3 ± 97.8	93.6 ± 42.8	174.9 ± 37.3
PD rec-PET	5	3.07/8.13	2.65/7.41	2.37/6.74	224.7 ± 89.5	99.3 ± 41.9	183.1 ± 63.0
Subsampled version							
T1-PET	7	4.13/8.25	3.70/7.37	3.40/7.48	169.0 ± 65.0	93.7 ± 31.4	127.5 ± 36.4
T2-PET	7	4.14/9.18	3.52/6.18	3.06/5.10	189.0 ± 63.4	79.1 ± 35.9	138.8 ± 40.6
PD-PET	7	3.07/12.94	2.74/7.06	2.59/5.36	151.5 ± 63.3	82.1 ± 39.0	135.7 ± 32.8
T1 rec-PET	4	4.62/10.37	3.95/9.22	3.51/8.83	150.2 ± 59.3	88.6 ± 28.9	118.5 ± 26.3
T2 rec-PET	5	5.25/16.15	4.92/10.88	4.83/7.07	176.5 ± 59.1	89.7 ± 34.2	136.7 ± 35.6
PD rec-PET	5	4.17/11.82	3.39/8.17	3.03/6.33	144.5 ± 52.2	90.7 ± 40.6	113.7 ± 29.4
Small-overlapped version							
T1-PET	7	4.15/5.78	3.71/6.65	3.33/7.17	146.2 ± 93.2	94.9 ± 35.7	112.3 ± 29.9
T2-PET	7	3.34/8.37	2.44/6.07	2.20/4.94	207.5 ± 88.2	120.5 ± 61.3	153.0 ± 48.2
PD-PET	7	3.98/8.18	3.11/6.21	2.84/5.11	213.7 ± 83.1	106.8 ± 47.9	134.2 ± 51.4
T1 rec-PET	4	3.33/10.85	2.96/8.13	2.61/6.72	139.5 ± 64.1	68.6 ± 40.7	96.5 ± 43.6
T2 rec-PET	5	4.17/12.05	3.89/9.27	3.21/7.38	170.0 ± 79.3	63.9 ± 37.0	99.1 ± 46.5
PD rec-PET	5	3.39/9.10	3.30/7.58	3.16/6.91	193.0 ± 114.8	103.4 ± 58.0	133.1 ± 52.8

4. DISCUSSION AND CONCLUSION

Though MI method is a well-known effective criterion for Multimodal image registration, it still has some disadvantages which often make the alignment less than optimal.

First, MI is unreliable to measure the degree of alignment between two images. MI function includes only intensity information but little spatial information of images, so it usually either produces several global maxima or presents a global maximum which does not correspond to the correct alignment. Some research introduced spatial information such as gradient-based information [6–8] or feature-based information [19–21] to improve the quality of image registration. These methods were effective but they did not took full advantage of the phase information of gradient field or the relationship between intensity images and their gradient fields.

Second, MI function is easily influenced by the intensity interpolation and presents many local maxima to trap the optimization [4, 5], leading to the failure of registration. Various high-order interpolation methods [22, 23] and global optimization algorithms [6] were introduced to reduce the influence of local maxima. But these methods are computationally expensive [24, 25]. Moreover, these methods are meaningless if the similarity measurement is unreliable [26, 27].

Third, MI is sensitive to the reduction of resolution or overlapped area of images. MI is a similarity measurement method and its reliability depends on the statistical stability of samples. The reduction of resolution or the overlapped area decreases the sample size, then deteriorates the statistical stability of samples. As a result, MI presents a poor performance for the registration of images with low resolution or small overlapped area. NMI [11] and ECC [10] were introduced to solve this problem, but no significant improvement was observed [25, 28]. They are also sensitive to the reduction of resolution or overlapped area of images.

To overcome these disadvantages of MI, we propose a technique for Multimodal image registration, namely ACMI, based on adaptive combination of intensity and gradient field mutual information. We constructed GCM from which the gradient field mutual information of original intensity images is calculated. The GCM is obtained from corresponding original images by a spherical gradient coder and includes both magnitude and phase information of gradient field of original images. The gradient field mutual information provides sufficient spatial information for the similarity measurement of images, besides it is smoother due to the relatively higher intensity uniformity of GCMs. ACMI combines the advantages of intensity ECC and gradient ECC, and adopts a coarse-to-fine and gradient-to-intensity registration strategy, so it overcomes the nonsmoothness and unreliability of traditional MI function. Results of simulated data experiments and actual registration both demonstrate that ACMI function performs better than traditional MI and it is much less sensitive to the reduction of resolution or overlapped area of two images.

According to its advantages, ACMI function is suitable for the registration of low-resolution images or impaired images. One example is the registration with multiresolution

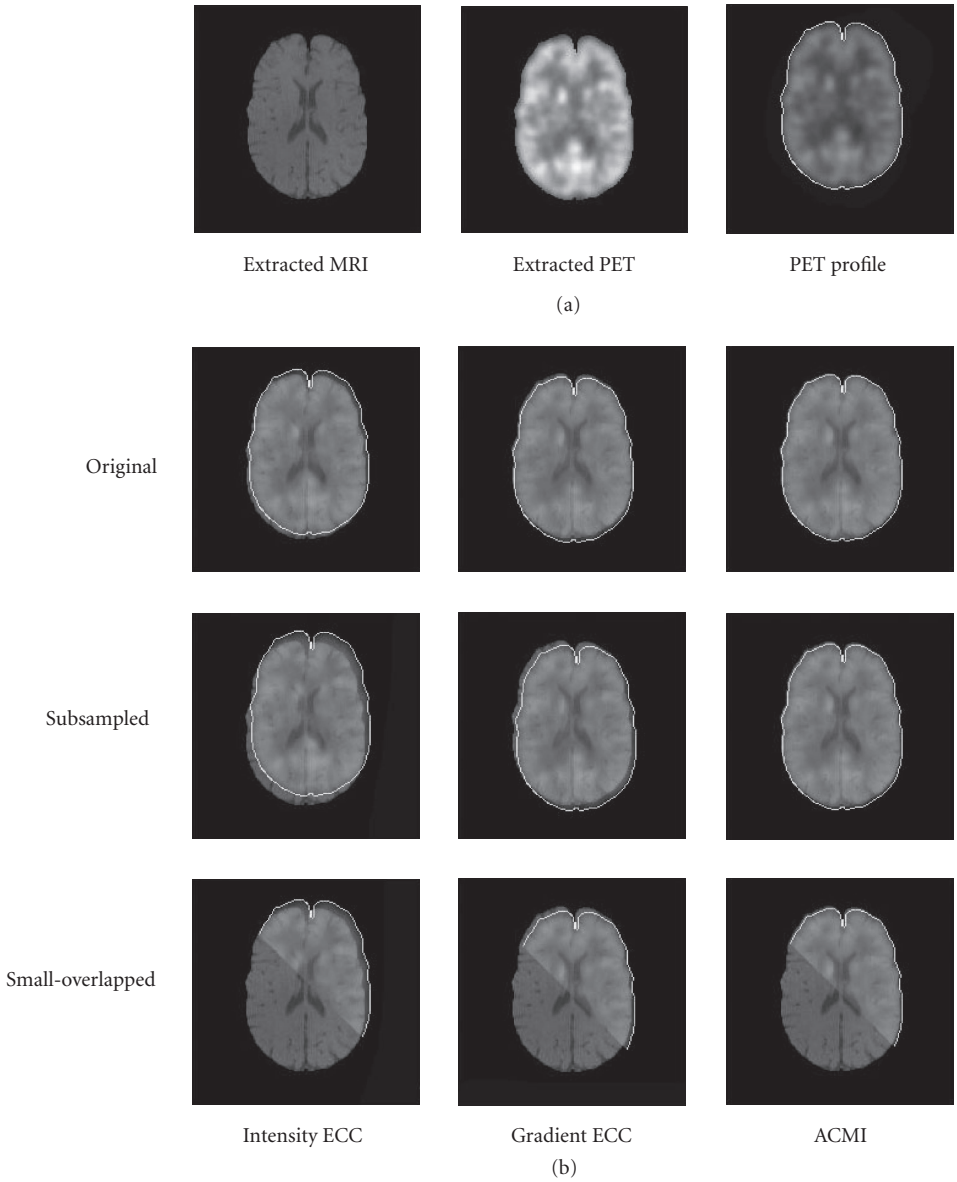


Figure 7: Illustration of registration results of PET/MRI T1 pair with intensity ECC, gradient ECC, and ACMI. (a) The extracted brain of MRI T1 image (left) and the 50% transparent profile of extracted brain of PET image (right) are served as the underlying and the overlying, respectively. (b) The first row: registration results of original version using intensity ECC, gradient ECC, and ACMI; the second row: the corresponding results of subsampled version; the third row: the corresponding results of small-overlapped version.

method whose object is to accelerate the registration speed without decreasing the registration accuracy and robustness. For low-resolution images, the multiresolution method with intensity ECC in fact prolongs the registration process, because the subsampling of these images deteriorates the smoothness of MI function, so that the convergence point of subsampled images is still a poor start point for the final images [29]. The ACMI can be used in multiresolution method for its insensitivity to the reduction of resolution.

In our study, the optimal value of time constant T was obtained using a "bootstrap" method. As shown in Figure 5, the mean registration error of each version is relatively insensitive to the changes of T near the optimal value, and extremely low or high T values can lead to relatively inaccurate registration. It is not clear whether this optimal T is suitable for other multimodality image pairs such as MRI/CT, MRI/SPECT. The registrations of these multimodality image pairs might present a similar pattern of "mean registration error versus time constant" to that of MRI/PET pairs (Figure 5). Extending ACMI to these multimodality image pairs will be our future work.

ACKNOWLEDGMENTS

The authors thank Dr. Ying Liu for her constructive comments on their manuscript. This study was supported by the

Project for National Science Fund for Distinguished Young Scholars of China under Grant no. 60225008, the Joint Research Fund for Overseas Chinese Young Scholars under Grant no. 30528027, Chair Professors of Cheung Kong Scholars Programme, PCSIRT, and CAS Hundred Talents Program.

REFERENCES

[1] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997.

[2] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," in *Proceedings of the 5th IEEE International Conference on Computer Vision (ICCV '95)*, pp. 16–23, Cambridge, Mass, USA, June 1995.

[3] F. Maes, D. Vandermeulen, and P. Suetens, "Medical image registration using mutual information," *Proceedings of the IEEE*, vol. 91, no. 10, pp. 1699–1721, 2003.

[4] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Computer Vision and Image Understanding*, vol. 77, no. 2, pp. 211–232, 2000.

[5] J. Tsao, "Interpolation artifacts in multimodality image registration based on maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 22, no. 7, pp. 854–864, 2003.

[6] T. Butz and J.-P. Thiran, "Affine registration with feature space mutual information," in *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '01)*, vol. 2208 of *Lecture Notes in Computer Science*, pp. 549–556, Utrecht, The Netherlands, October 2001.

[7] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Transactions on Medical Imaging*, vol. 19, no. 8, pp. 809–814, 2000.

[8] E. Haber and J. Modersitzki, "Intensity gradient based registration and fusion of multi-modal images," Tech. Rep. TR-2004-027-A, Department of Mathematics and Computer Science, Emory University, Atlanta, Ga, USA, June 2004.

[9] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "Comparison of edge-based and ridge-based registration of CT and MR brain images," *Medical Image Analysis*, vol. 1, no. 2, pp. 151–161, 1996.

[10] F. Maes, *Segmentation and registration of multimodal medical images: from theoty, implementation and validation to a useful tool in clinical practice*, Ph.D. thesis, Catholic University of Leuven, Leuven, Belgium, 1998.

[11] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.

[12] X. Wang and J. Tian, "Image registration based on maximization of gradient code mutual information," *Image Analysis and Stereology*, vol. 24, pp. 1–7, 2005.

[13] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, China Machine Press, Beijing, China, 2nd edition, 2004.

[14] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.

[15] F. Maes, D. Vandermeulen, and P. Suetens, "Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information," *Medical Image Analysis*, vol. 3, no. 4, pp. 373–386, 1999.

[16] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions," *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[17] J. West, J. M. Fitzpatrick, M. Y. Wang, et al., "Comparison and evaluation of retrospective intermodality brain image registration techniques," *Journal of Computer Assisted Tomography*, vol. 21, no. 4, pp. 554–566, 1997.

[18] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Publishing House of Electronics Industry, Beijing, China, 2nd edition, 2002.

[19] T. Hartkens, D. L. G. Hill, A. D. Castellano-Smith, et al., "Using points and surfaces to improve voxel-based non-rigid registration," in *Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '02)*, vol. 2489 of *Lecture Notes in Computer Science*, pp. 565–572, Tokyo, Japan, September 2002.

[20] P. Hellier and C. Barillot, "Coupling dense and landmark-based approaches for nonrigid registration," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 217–227, 2003.

[21] H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 450–461, 2002.

[22] P. Thevenaz and M. Unser, "A pyramid approach to sub-pixel image fusion based on mutual information," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '96)*, vol. 1, pp. 265–268, Lausanne, Switzerland, September 1996.

[23] E. H. W. Meijering, W. J. Niessen, and M. A. Viergever, "Quantitative evaluation of convolution-based methods for medical image interpolation," *Medical Image Analysis*, vol. 5, no. 2, pp. 111–126, 2001.

[24] R. He and P. A. Narayana, "Global optimization of mutual information: application to three-dimensional retrospective registration of magnetic resonance images," *Computerized Medical Imaging and Graphics*, vol. 26, no. 4, pp. 277–292, 2002.

[25] T. Yokoi, T. Soma, H. Shinohara, and H. Matsuda, "Accuracy and reproducibility of co-registration techniques based on mutual information and normalized mutual information for MRI and SPECT brain images," *Annals of Nuclear Medicine*, vol. 18, no. 8, pp. 659–667, 2004.

[26] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[27] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical Image Analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[28] C. Grova, A. Biraben, J. M. Scarabin, et al., "A methodology to validate MRI/SPECT registration methods using realistic simulated spect data," in *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '01)*, vol. 2208 of *Lecture Notes in Computer Science*, pp. 275–282, Utrecht, The Netherlands, October 2001.

[29] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual information matching in multiresolution contexts," *Image and Vision Computing*, vol. 19, no. 1-2, pp. 45–52, 2001.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 23624, 7 pages
doi:10.1155/2007/23624

Research Article

Quantitative and O_2 Enhanced MRI of the Pathologic Lung: Findings in Emphysema, Fibrosis, and Cystic Fibrosis

Alfred Stadler,[1] Leopold Stiebellehner,[2] Peter M. Jakob,[3] Johannes F. T. Arnold,[3] Edith Eisenhuber,[4] Isabella von Katzler,[1] and Alexander A. Bankier[1]

[1] *Department of Radiology, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria*
[2] *Department of Internal Medicine IV, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria*
[3] *Institute of Physics, Experimental Physics 5, University of Wuerzburg, 97074 Wuerzburg, Germany*
[4] *Department of Radiology, Otto Wagner Hospital, 1140 Vienna, Austria*

Received 23 November 2006; Accepted 25 February 2007

Recommended by Haim Azhari

Purpose: beyond the pure morphological visual representation, MR imaging offers the possibility to quantify parameters in the healthy, as well as, in pathologic lung parenchyma. Gas exchange is the primary function of the lung and the transport of oxygen plays a key role in pulmonary physiology and pathophysiology. The purpose of this review is to present a short overview of the relaxation mechanisms of the lung and the current technical concepts of T1 mapping and methods of oxygen enhanced MR imaging. *Material and Methods*: molecular oxygen has weak paramagnetic properties so that an increase in oxygen concentration results in shortening of the T1 relaxation time and thus to an increase of the signal intensity in T1 weighted images. A possible way to gain deeper insights into the relaxation mechanisms of the lung is the calculation of parameter Maps. T1 Maps based on a snapshot FLASH sequence obtained during the inhalation of various oxygen concentrations provide data for the creation of the so-called oxygen transfer function (OTF), assigning a measurement for local oxygen transfer. T1 weighted single shot TSE sequences also permit expression of the signal changing effects associated with the inhalation of pure oxygen. *Results*: the average of the mean T1 values over the entire lung in inspiration amounts to 1199 +/− 117 milliseconds, the average of the mean T1 values in expiration was 1333 +/− 167 milliseconds. T1 Maps of patients with emphysema and lung fibrosis show fundamentally different behavior patterns. Oxygen enhanced MRT is able to demonstrate reduced diffusion capacity and diminished oxygen transport in patients with emphysema and cystic fibrosis. *Discussion*: results published in literature indicate that T1 mapping and oxygen enhanced MR imaging are promising new methods in functional imaging of the lung and when evaluated in conjunction with the pure morphological images can provide additional valuable information.

1. INTRODUCTION

Beyond the pure morphological visual representation, MR imaging offers the possibility to quantify functional parameters in healthy, as well as, the pathologic lung tissue. The T1 time of biological tissue is one of the potential measurable parameters. Gas exchange is the primary function of the lung and oxygen uptake plays a major role in lung physiology and pathophysiology. Consequently, the representation and quantification of lung oxygen uptake provides important information regarding pulmonary function. Furthermore, image creation of pathologically altered oxygen exchange in the diseased lung is of great preclinical and clinical interest. Due to the paramagnetic properties of oxygen, MR imaging of oxygen transport is theoretically possible. The purpose of this review is to present a short overview of the relaxation mechanisms of the lung and the current technical concepts of T1 mapping and methods of oxygen enhanced MR imaging.

2. PULMONARY RELAXATION MECHANISMS

The first order approximation of the lung T1 relaxation time shows it to be monoexponential and considerably dependant upon tissue water content: the greater the water content the shorter the T1 time [1, 2]. In a majority of studies, alteration of T1 relaxation times in pathologically-changed lungs (i.e., pulmonary edema) could be qualitatively explained by the effect of water in the lungs [3], even though the measurements showed variable quantitative results [4]. The T1 alterations

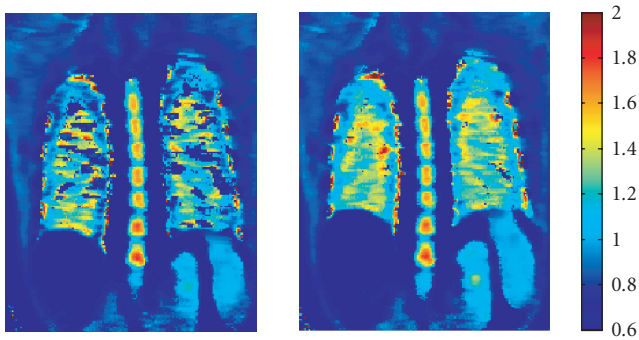


Figure 1: T1 map of a healthy subject: expiration (b) demonstrates a lengthening of the T1 times (seconds) as compared to inspiration (a).

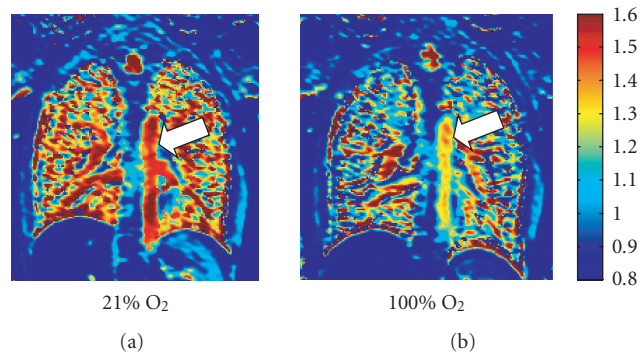


Figure 2: T1 map (seconds) of a healthy subject with inhalation of room air 21% oxygen (a), as well as inhalation of 100% oxygen (b). The increase of O_2 concentration shortens the T1 time of the lung parenchyma as well as the Aorta (arrow).

observed in animal experiments, or in ex vivo observations of various pathologies (i.e., Lung fibrosis) could not solely be explained by the effect of lung water [5–8].

The explanation of MR relaxation mechanisms in the presence of macromolecules, such as collagen, was simplified via the development of a theoretical model which is based on the interaction between water molecules and macromolecules [9]. This takes into account that the pulmonary T1 relaxation properties are determined by two balanced groups of water molecules: one compartment consists of free water and the other by water which is bound to macromolecules such as collagen. An alteration of the lung parenchyma, such as the increase in the quantity of macromolecules, leads to a change in the water-bound fraction and therefore a change in the proportion of water-bound molecules.

A significant percentage of signal generating protons are found in the blood. Parenchymal regional perfusion will, therefore, have an effect on the regional lung parenchymal T1 relaxation time. Lung perfusion and airspace oxygen content are dependent from the respiratory phase. However, in a recent publication, we illustrated that the T1 time in inspiration (1199+/−117 milliseconds) is shorter than in expiration (1333+/−167 milliseconds) [10] (Figure 1).

3. OXYGEN AS CONTRAST MEDIUM

Due to two unpaired electrons, molecular oxygen has weak paramagnetic properties with a magnetic moment of 2.8 Bohr magnetons. The idea to use oxygen as a paramagnetic contrast medium is not new and was initially discussed by Young et al. [11]. The potential advantages of oxygen as a contrast medium are obvious. Oxygen is inexpensive, readily available, and safe. Short-term inhalation is not associated with any adverse side-effects. Only after continuous inhalation of 100% oxygen for more than 24 hours does the possibility of permanent pathological lung changes arise [12]. The illustration of oxygen dispersion is of great physiological and pathophysiological interest in that oxygen transport represents the essential function of the lungs and not the distribution of nonphysiological substances which are measured in examinations such as a nuclear medicine-pulmonary function test. A portion of oxygen is bound to hemoglobin in the pulmonary capillary bed and a small fraction remains in soluble form. The hemoglobin bound oxygen is enclosed within the erythrocyte and therefore, the tissue water protons cannot engage in a spin-lattice interaction which leads to T1 relaxation [13]. Edelman et al. were the first to discuss the use of the paramagnetic properties of soluble oxygen to illustrate pulmonary oxygen transport [14]. Inhalation of pure oxygen increases the PaO_2 in the lungs. This increase in partial oxygen pressure leads to a shortening of the T1 time (Figure 2) and therefore, to a signal rise of the T1 weighted images [15]. Animal studies demonstrated a linear correlation between PaO_2 and lung parenchymal relaxivity (=1/T1) [16]. The difference in lung parenchymal acquisition signal intensities between inhaled room air and 100% oxygen is minimal, and visual representation is generally accomplished by image subtraction. Apart from the lungs, signal increases were also noted in other organs such as the aorta, spleen, and kidneys [15]. Besides the effect on T1, an elevated oxygen concentration also leads to a prolongation of lung parenchymal T2* time with only a minimal influence on the signal intensity [16].

The exact mechanism altering signal or the change in T1 time is not known. The influence of oxygen on T1 time seems to be played out on the pulmonary vein and parenchymal levels [17]. The molecular oxygen paramagnetic effect is in any case not measurable within the pulmonary gaseous spaces in that the oxygen only influences relaxation of water protons in its proximate surroundings and is itself not signal emitting. After inhalation of higher oxygen concentrations, the interplay of inhalation, diffusion, and perfusion influence the conventional acquisition signal and also alter the T1 time. It only follows that there is a measurable differences in signal behavior in the pathological lung after inhalation of oxygen which results from an alteration of one or more of these factor, thereby increasing interpretation difficulty.

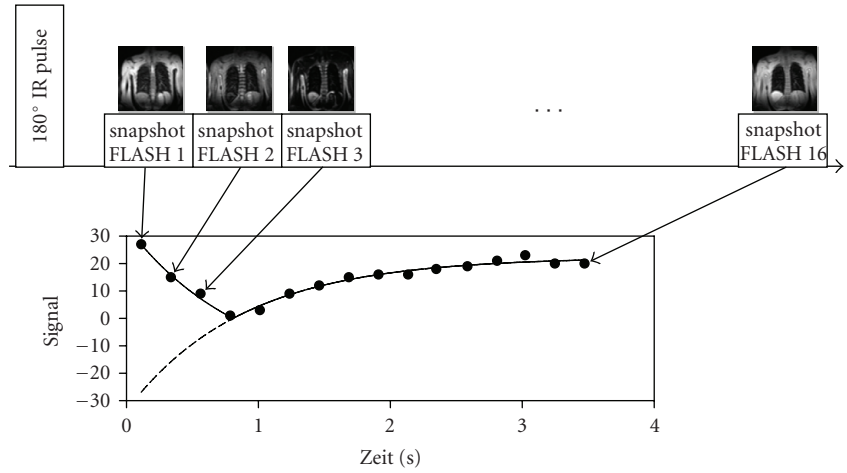


Figure 3: T1 Measurements: after a 180° inversion pulse, 16 Images were acquired at specific time points using a snapshot FLASH sequence. The individual measurements have a duration of 224 milliseconds, the complete slice 3,5 seconds. Due to the short acquisition time, patients with moderate dyspnea can be examined using this method. An exponential function was assigned to each Pixel covering the time span of the 16 measured signal intensities from which a Pixel T1 time can be calculated. The final T1 map results from the color coding of the T1 values.

Ohno et al. [18] used the absolute sum of the signal rise during a dynamic measurement as a parameter for lung diffusion capacity, Müller et al. [19] the slope of the rise.

Oxygen Transfer Function (OTF) was created to describe the oxygen transport T1 maps [20]. This was achieved by measuring T1 maps at various inspired oxygen concentrations. Increase in relaxivity is a measure for oxygen transfer of inhaled air into the blood stream. OTF describes the interplay between oxygen diffusion, ventilation, and perfusion.

4. TECHNIQUE

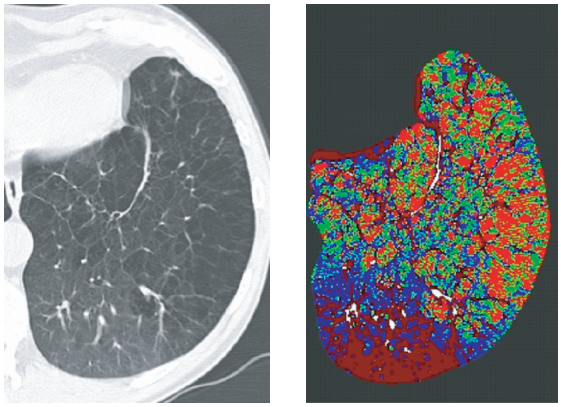
MRT examination of the pulmonary parenchyma exhibits a minor signal-to-noise ratio. On the one hand, this is due to minimal lung parenchymal proton density, and on the other, the multiple air/parenchymal surfaces cause susceptibility jumps which results in an extremely short T2 time of only a few milliseconds [21–23]. To achieve the highest possible signal, a short echo time (TE) is necessary. Therefore, the use of a single shot turbo spin-echo (TSE) technique has several advantages: a short TE, multiple 180° refocusing pulses to minimize susceptibility artifacts, a short inter echo time to minimize diffusion and perfusion effects, and a short acquisition time to reduce motion artifacts. In several studies [16, 17, 24, 25] a T1 weighted inversion recovery half-Fourier acquisition single shot turbo spin echo (HASTE) for oxygen enhanced pulmonary MRT was used.

The calculation of T1 maps allows for the elimination of T2 effects and the analysis of the effects of oxygen inhalation. In vivo T1 lung measurements are, as all other MR pulmonary parenchymal measurements, difficult. On the one side, low-proton densities cause a poor signal-to-noise ratio, and conversely, measurements are influenced by susceptibility artifacts as well as motion artifacts (cardiac pulsation and diaphragmatic motion). Most of the published studies examined an only small volume measurement that naturally covers only a minor portion of the lung. This results not only in problems of a statistical nature, but also varying T1 values are expected where it is then unclear if the entire pulmonary parenchyma is represented. In order to evaluate diffuse, but regionally inhomogeneous alterations in lung parenchyma, T1 maps of the entire lung are necessary where there is a corresponding T1 value for every pixel.

A possibility which is proposed in several projects, including work from our group, calculates T1 maps based on measurements from snapshot FLASH sequences [10, 26, 27]. These are based on the TOMROP sequence and consist of two elements: first, magnetization is inverted using a nonselective inversion pulse, then the return to original magnetization state occurs over longitudinal relaxation and an image is created using a series of measurements taken from a fast snapshot FLASH sequence [28]. The scan time for one slice at a specific point in time has an estimated duration of 200 milliseconds and is measured at 16 predefined points in time after the inversion pulse. Acquisition of an entire slice lasts about four seconds. The measurement provides 16 time-dependant signal intensities for each pixel. An exponential-fit over these data points yields the T1 relaxation time for each pixel. A color coded image of the T1 values for all the pixels produces the final T1 map (Figure 3).

5. EMPHYSEMA

WHO defines emphysema as irreversible enlargement of the air spaces distal to the terminal bronchiole with destruction of the elastic scaffolding without accompanying fibrosis [29]. The detection and the morphological grading of lung emphysema using computer tomography have been intensively explored over the last few years [30]. Furthermore, there are ongoing studies to evaluate the capacity of CT



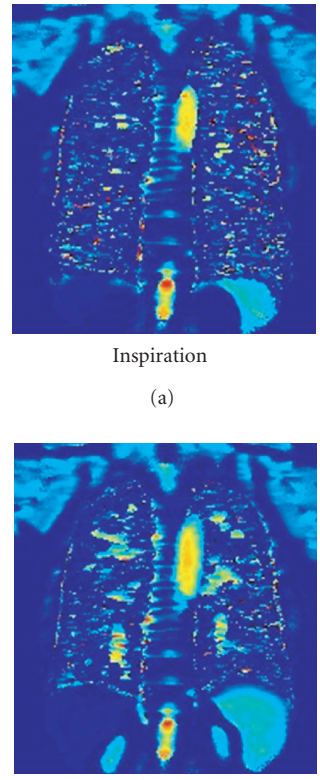
(a) (b)

Figure 4: Morphological CT (a) and perfusion map of an emphysematous lung show good correlation of severe emphysema in the anterior-lateral parts of the lung and decreased perfusion (green pixels) as compared to the normal posterior parts of the lung. The bright red pixels represent excluded, air containing parts of the lung.

to include functional information like perfusion (Figure 4). The inherent problems associated with pulmonary MR examinations, especially the low-proton density (and the poor signal-to-noise ratio) or the susceptibility artifacts are even more pronounced in the emphysema altered lung. Independent of emphysema pathogenesis, the total composition of macromolecules such as collagen and elastin are within normal limits, however the distribution and the organization are pathologically altered [31]. Signal behavior of the emphysemic lung is, apart from a total reduction in signal intensity which mirrors the pulmonary parenchymal destruction, difficult to foresee. It was recently demonstrated that the T1 relaxation time in emphysematous altered lung is significantly shorter than in healthy lung (Figure 5) [32]. The cause of this T1 time shortening possibly lies in vascular rarefaction or a redistribution of blood within the effected lung parenchyma. In any case, this has to be taken into consideration when performing inhalation examinations. At this time, there is no definitive data available relating to T1 relaxation time behavior after 100% oxygen inspiration.

A study from Ohno et al. compared healthy volunteers with emphysema patients [18], and healthy volunteers with patients diagnosed with bronchial carcinoma and Bronchial carcinoma patients without emphysema to those with [25], respectively. The time course of an acquisition was examined by taking sequential measurements using HASTE sequencing with inhalation of 21% (room air) and 100% oxygen.

The 100% oxygen signal rise in patients with pulmonary emphysema was significantly flatter and demonstrated excellent correlation with FEV1. Strong correlation between the maximum signal rise and CT emphysema scoring as well as pulmonary diffusion capacity was also noted. Müller was also able to show a reduction in diffusion capacity in the emphysemic lung segments when compared to healthy subjects through dynamic measurements of pulmonary signal behavior during inhalation of 100% oxygen. Contrary Ohno's results, it was demonstrated that inspiration of 100% oxygen provided a good correlation between the slope of the signal rise and the clinical measurement of diffusion capacity.



Inspiration

(a)

Expiration

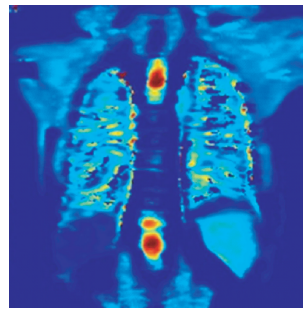
(b)

Figure 5: T1 map of a patient with lung emphysema: when compared to a healthy lung (Ill. 1), the T1 time is clearly shortened. Expiration (b) demonstrates no significant prolongation of the T1 time as compared with inspiration (a).

The existing studies show that oxygen enhanced pulmonary MRT can be a valuable complementary tool in evaluating pulmonary emphysema patients, especially in spatial encoded imaging of pulmonary diffusion capacity. Investigations in this field are, however, still in the preclinical trial phase.

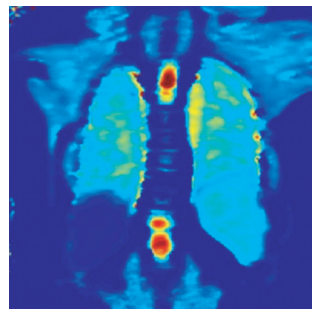
6. FIBROSIS

Independent of the etiology, pulmonary fibrosis is characterized by the deposition of newly-synthesized matrix molecules. In accordance with the above described simplified two compartment model, the relative increase in macromolecules should lead to a shortening of the T1 relaxation time. This assumption was partially verified by animal model measurements [7], and in some measurements there was no change in T1 time noted [6]. Our measurements demonstrated that patients with lung fibrosis, when compared to healthy volunteer subjects, showed a shortening in the T1 relaxation time which is less pronounced in expiration than



Inspiration

(a)



Expiration

(b)

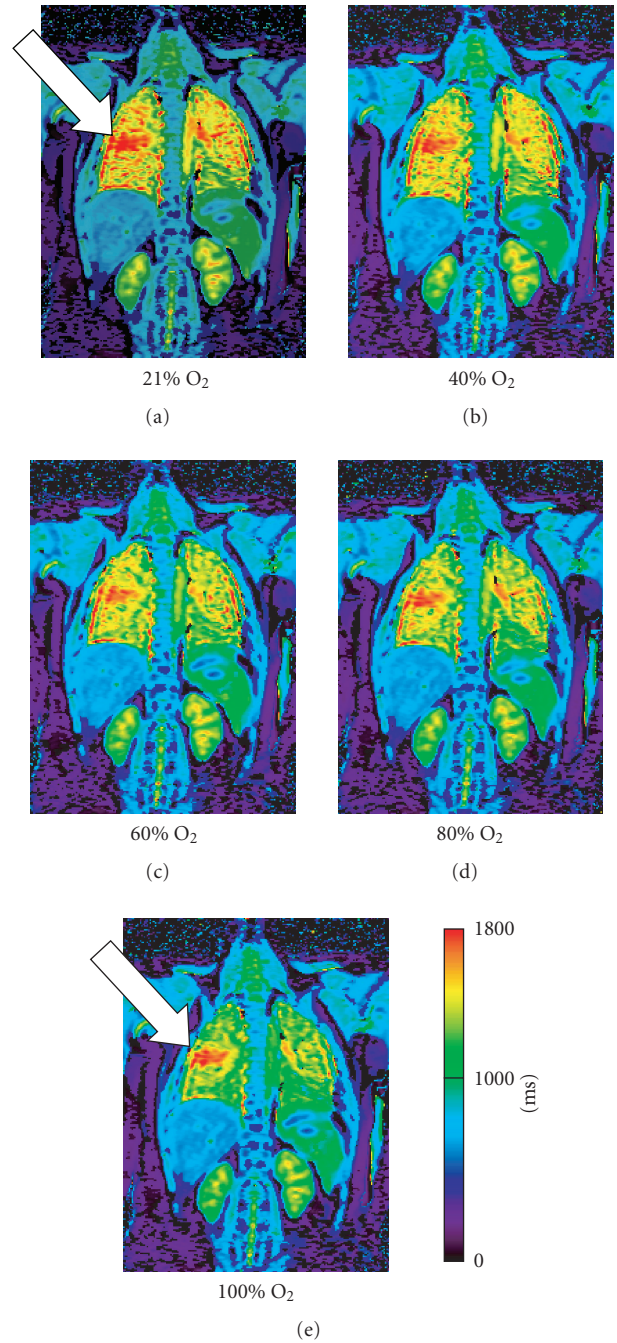
Figure 6: T1 map of a Patient with lung fibroses: in comparison to a healthy lung (Figure 2), the T1 time is shortened. Expiration (b) shows a significant T1 time prolongation as compared to inspiration (a).

in inspiration (Figure 6) [32]. However, there are no current publications discussing the behavior of the fibrotic lung under 100% oxygen inspiration.

7. CYSTIC FIBROSIS

Cystic fibrosis is an autosomal-recessive hereditary disease in which the pathological composition of the exocrine gland secretion leads to characteristic secondary changes in target organs. In the lungs, this disease has a homogeneous picture encompassing atelectasis, emphysema, microabscesses, bronchiectasis which ultimately leads to pulmonary fibrosis.

In a recent work, T1 maps of healthy volunteers were compared to those of cystic fibrosis patients [20]. The patient group showed an inhomogeneous distribution of T1 relaxation times where pathologically altered segments had shorter T1 times than the noneffected segments as well as the lungs of the healthy subjects (Figure 7). The OTF curve of the pathological lung segments exhibited a distinctly flattend pattern, consequently, a reduced dependence of pulmonary parenchymal relaxivity on the inhaled oxygen concentration (Figure 8). Reasons for this change can either be related to the limited pulmonary diffusion capacity or the alteration in ventilation or perfusion of the diseased segments, thus a supplementary MR pulmonary perfusion study is recommended.



21% O_2

(a)

40% O_2

(b)

60% O_2

(c)

80% O_2

(d)

100% O_2

(e)

Figure 7: T1 map (millisecond) of a patient with Cystic Fibrosis: inhalation of increasing oxygen concentrations demonstrates a steady shortening of T1 times in the healthy lung segments. The diseased tissue of the right mid lobe (arrow) distinguishes itself in that the T1 time remains relatively unchanged.

8. SUMMARY

T1 maps from healthy subjects and patients with emphysema and fibrosis reveal significantly different behaviors. These differences reflect the complex interaction of the structural and functional influences of the above-mentioned diseases.

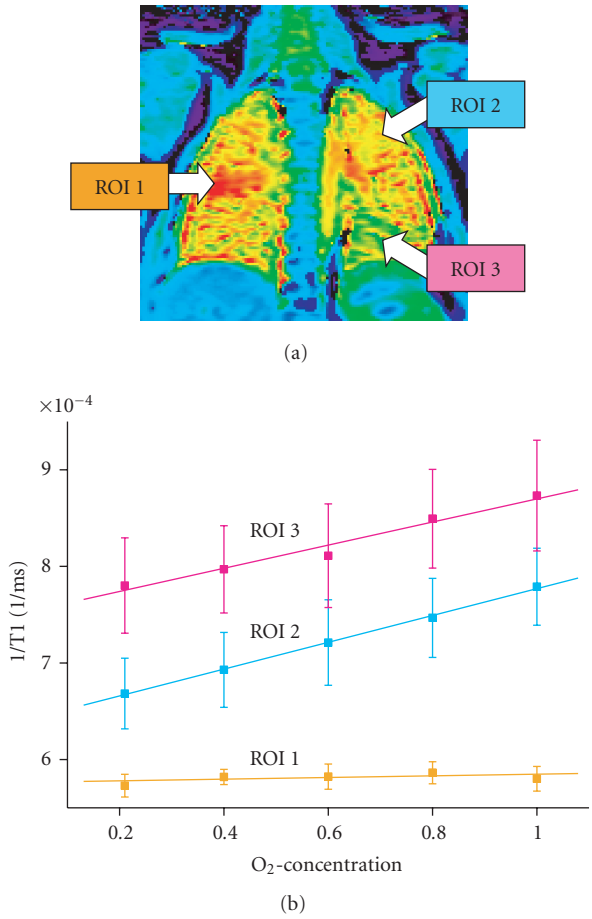


FIGURE 8: Oxygen transfer function (OTF) of a patient with Cystic Fibroses from Ill. 6. Healthy lung segments (ROI 2 and 3) demonstrate a linear increase in relaxivity (=1/T1) with increasing oxygen concentrations. The slope of the curve is a measurement of lung oxygen transport capacity. The diseased lung segment (ROI 1) shows no rise; there is therefore no oxygen transport.

Pulmonary relaxation mechanisms are still not fully understood. Further studies using T1 maps can advance the understanding of the relationship between lung structure and lung function. A basic knowledge of the T1 relaxation mechanisms is also the ground work for optimizing conventional MR pulmonary image sequencing. All published data, up to now, confirm that oxygen enhanced MRT has an important role in the imaging of diffusion capacity and oxygen transport which provides valuable information in the detection and understanding of the role that functional alterations have in lung diseases such as emphysema and cystic fibrosis.

In spite of the considerable technical difficulties, several publications confirm the potential that T1 maps and oxygen enhanced MRT have characterizing pathological changes in lung tissue. However, existing literature still cannot provide a final evaluation of the presented methods. The previously obtained results will allow for further informative insight into the functional changes of the pathological altered lung parenchyma.

REFERENCES

[1] P. A. Bottomley, C. J. Hardy, R. E. Argersinger, and G. Allen-Moore, "A review of ^{1}H nuclear magnetic resonance relaxation in pathology: are T_1 and T_2 diagnostic?" *Medical Physics*, vol. 14, no. 1, pp. 1–37, 1987.

[2] C. E. Hayes, T. A. Case, D. C. Ailion, et al., "Lung water quantitation by nuclear magnetic resonance imaging," *Science*, vol. 216, no. 4552, pp. 1313–1315, 1982.

[3] A. G. Cutillo, A. H. Morris, D. C. Ailion, C. H. Durney, and K. Ganesan, "Quantitative assessment of pulmonary edema by nuclear magnetic resonance methods," *Journal of Thoracic Imaging*, vol. 3, no. 3, pp. 51–58, 1988.

[4] P. W. Johnston, F. M. MacLennan, J. G. Simpson, and F. W. Smith, "Nuclear magnetic resonance imaging of pulmonary infarction and oedema in excised cadaver lungs," *Magnetic Resonance Imaging*, vol. 3, no. 2, pp. 157–161, 1985.

[5] M. Kveder, I. Zupancic, G. Lahajnar, et al., "Water proton NMR relaxation mechanisms in lung tissue," *Magnetic Resonance in Medicine*, vol. 7, no. 4, pp. 432–441, 1988.

[6] C. R. Taylor, H. D. Sostman, J. C. Gore, and G. Walker Smith, "Proton relaxation times in bleomycin-induced lung injury," *Investigative Radiology*, vol. 22, no. 8, pp. 621–626, 1987.

[7] S. Vinitski, M. G. Pearson, S. J. Karlik, et al., "Differentiation of parenchymal lung disorders with in vitro proton nuclear magnetic resonance," *Magnetic Resonance in Medicine*, vol. 3, no. 1, pp. 120–125, 1986.

[8] S. Vinitski, R. M. Steiner, H. R. Wexler, and M. Rifkin, "Assessment of lung water by magnetic resonance in three types of pulmonary edema," *Heart and Vessels*, vol. 4, no. 2, pp. 88–93, 1988.

[9] T. D. Scholz, S. R. Fleagle, T. L. Burns, and D. J. Skorton, "Tissue determinants of nuclear magnetic resonance relaxation times. Effect of water and collagen content in muscle and tendon," *Investigative Radiology*, vol. 24, no. 11, pp. 893–898, 1989.

[10] A. Stadler, P. M. Jakob, M. Griswold, M. Barth, and A. A. Bankier, "T1 mapping of the entire lung parenchyma: influence of the respiratory phase in healthy individuals," *Journal of Magnetic Resonance Imaging*, vol. 21, no. 6, pp. 759–764, 2005.

[11] I. R. Young, G. J. Clarke, D. R. Bailes, J. M. Pennock, F. H. Doyle, and G. M. Bydder, "Enhancement of relaxation rate with paramagnetic contrast agents in NMR imaging," *Journal of Computed Tomography*, vol. 5, no. 6, pp. 543–547, 1981.

[12] R. F. Lodato, "Oxygen toxicity," *Critical Care Clinics*, vol. 6, no. 3, pp. 749–765, 1990.

[13] R. A. Brooks and G. Di Chiro, "Magnetic resonance imaging of stationary blood: a review," *Medical Physics*, vol. 14, no. 6, pp. 903–913, 1987.

[14] R. R. Edelman, H. Hatabu, E. Tadamura, W. Li, and P. V. Prasad, "Noninvasive assessment of regional ventilation in the human lung using oxygen-enhanced magnetic resonance imaging," *Nature Medicine*, vol. 2, no. 11, pp. 1236–1239, 1996.

[15] E. Tadamura, H. Hatabu, W. Li, P. V. Prasad, and R. R. Edelman, "Effect of oxygen inhalation on relaxation times in various tissues," *Journal of Magnetic Resonance Imaging*, vol. 7, no. 1, pp. 220–225, 1997.

[16] Y. Ohno, Q. Chen, and H. Hatabu, "Oxygen-enhanced magnetic resonance ventilation imaging of lung," *European Journal of Radiology*, vol. 37, no. 3, pp. 164–171, 2001.

[17] Q. Chen, P. M. Jakob, M. A. Griswold, D. L. Levin, H. Hatabu, and R. R. Edelman, "Oxygen enhanced MR ventilation imaging of the lung," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 7, no. 3, pp. 153–161, 1998.

[18] Y. Ohno, H. Hatabu, D. Takenaka, M. van Cauteren, M. Fujii, and K. Sugimura, "Dynamic oxygen-enhanced MRI reflects diffusing capacity of the lung," *Magnetic Resonance in Medicine*, vol. 47, no. 6, pp. 1139–1144, 2002.

[19] C. J. Müller, M. Schwaiblmair, J. Scheidler, et al., "Pulmonary diffusing capacity: assessment with oxygen-enhanced lung MR imaging-preliminary findings," *Radiology*, vol. 222, no. 2, pp. 499–506, 2002.

[20] P. M. Jakob, T. Wang, G. Schultz, H. Hebestreit, A. Hebestreit, and D. Hahn, "Assessment of human pulmonary function using oxygen-enhanced T_1 imaging in patients with cystic fibrosis," *Magnetic Resonance in Medicine*, vol. 51, no. 5, pp. 1009–1016, 2004.

[21] C. J. Bergin, G. H. Glover, and J. M. Pauly, "Lung parenchyma: magnetic susceptibility in MR imaging," *Radiology*, vol. 180, no. 3, pp. 845–848, 1991.

[22] H. Hatabu, D. C. Alsop, J. Listerud, M. Bonnet, and W. B. Gefter, "T2* and proton density measurement of normal human lung parenchyma using submillisecond echo time gradient echo magnetic resonance imaging," *European Journal of Radiology*, vol. 29, no. 3, pp. 245–252, 1999.

[23] K. W. Stock, Q. Chen, H. Hatabu, and R. R. Edelman, "Magnetic resonance T2* measurements of the normal human lung in vivo with ultra-short echo times," *Magnetic Resonance Imaging*, vol. 17, no. 7, pp. 997–1000, 1999.

[24] H. Hatabu, J. Gaa, E. Tadamura, et al., "MR imaging of pulmonary parenchyma with a half-Fourier single-shot turbo spin-echo (HASTE) sequence," *European Journal of Radiology*, vol. 29, no. 2, pp. 152–159, 1999.

[25] Y. Ohno, H. Hatabu, D. Takenaka, S. Adachi, M. van Cauteren, and K. Sugimura, "Oxygen-enhanced MR ventilation imaging of the lung: preliminary clinical experience in 25 subjects," *American Journal of Roentgenology*, vol. 177, no. 1, pp. 185–194, 2001.

[26] P. M. Jakob, C. M. Hillenbrand, T. Wang, G. Schultz, D. Hahn, and A. Haase, "Rapid quantitative lung $^1HT_1$ mapping," *Journal of Magnetic Resonance Imaging*, vol. 14, no. 6, pp. 795–799, 2001.

[27] R. Löffler, C. J. Müller, M. Peller, et al., "Optimization and evaluation of the signal intensity change in multisection oxygen-enhanced MR lung imaging," *Magnetic Resonance in Medicine*, vol. 43, no. 6, pp. 860–866, 2000.

[28] A. Haase, D. Matthaei, R. Bartkowski, E. Duhmke, and D. Leibfritz, "Inversion recovery snapshot FLASH MR imaging," *Journal of Computer Assisted Tomography*, vol. 13, no. 6, pp. 1036–1040, 1989.

[29] G. Snider, J. Kleinermann, and W. Thurlbeck, "The definition of emphysema. Report of a National Heart, Lung, and Blood Institute, Division of Lung Diseases workshop," *American Review of Respiratory Disease*, vol. 132, no. 1, pp. 182–185, 1985.

[30] A. Madani, C. Keyzer, and P. A. Gevenois, "Quantitative computed tomography assessment of lung structure and function in pulmonary emphysema," *European Respiratory Journal*, vol. 18, no. 4, pp. 720–730, 2001.

[31] J. A. Pierce, J. B. Hocott, and R. V. Ebert, "The collagen and elastin content of the lung in emphysema," *Annals of Internal Medicine*, vol. 55, pp. 210–222, 1961.

[32] A. Stadler, L. Stiebellehner, P. M. Jakob, M. Griswold, and A. A. Bankier, "T1 mapping of the entire lung in patients with emphysema and fibrosis compared to normal individuals," in *Proceedings of the 91st Radiological Society of North America Conference (RSNA '05)*, p. 242, Chicago, Ill, USA, November-December 2005, Book of Abstracts.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 40980, 8 pages
doi:10.1155/2007/40980

Research Article

Integration of Vibro-Acoustography Imaging Modality with the Traditional Mammography

H. Gholam Hosseini,[1] A. Alizad,[2] and M. Fatemi[2]

[1] *School of Engineering, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand*
[2] *Deptartment of Physiology and Biomedical Engineering, Mayo Clinic College of Medicine, Rochester, MN 55905, USA*

Received 1 May 2006; Revised 29 November 2006; Accepted 12 December 2006

Recommended by Haim Azhari

Vibro-acoustography (VA) is a new imaging modality that has been applied to both medical and industrial imaging. Integrating unique diagnostic information of VA with other medical imaging is one of our research interests. In this work, we establish correspondence between the VA images and traditional X-ray mammogram by adopting a flexible control-point selection technique for image registration. A modified second-order polynomial, which simply leads to a scale/rotation/translation invariant registration, was used. The results of registration were used to spatially transform the breast VA images to map with the X-ray mammography with a registration error of less than 1.65 mm. The fused image is defined as a linear integration of the VA and X-ray images. Moreover, a color-based fusion technique was employed to integrate the images for better visualization of structural information.

1. INTRODUCTION

Screening X-ray mammography is continue to be the primary tool for early detection of breast cancer and therefore reduces the mortality rate of the disease. Over the past few years, the quality of mammography has improved significantly but the accuracy of image interpretation is still a remained challenge. Mammography interpretation depends on human factors which is difficult to quantify. Thus, ensuring accurate interpretation of mammography is important for women's health [1].

Vibro-acoustography (VA) is a new imaging modality based on ultrasound-stimulated acoustic emission which can be integrated with the mammography to enhance breast cancer diagnosis. The VA acoustic field in response to vibration of an object due to an applied cyclic force at each point is detected by a hydrophone and used to form the image of the object [2].

Vibro-acoustography has been tested as a noninvasive imaging tool to image excised human tissues, such as breast [3], liver [4], arteries [5], and prostate [6]. Vibro-acoustography has also been used as a nondestructive imaging tool to identify the structural flaws of materials by measuring changes in the mechanical response to vibration at a point of interest [7, 8].

Recently, we have developed a VA system for in vivo breast imaging [3]. This system is integrated with a clinical stereotactic mammogram machine. The combined system is designed to produce matching VA and mammography images of the breast. The dual modality system can serve two purposes. The mammogram is used as a reference image to evaluate and optimize VA performance. Secondly, it is anticipated that the VA and mammography images would provide complimentary information of the breast. Thus, by integrating the two images, the diagnostic value of the two-modality image would be more than the individual images.

While mammography is considered as an important diagnostic tool, particularly for screening microcalcification clusters and detecting malignancy [9–11], there are associated shortcomings which raise concern about the quality of image interpretation. For example, the efficacy of this modality heavily decreases in dense breast imaging [12]. Moreover, X-ray mammography does not contain information about the depth and thickness of the objects. Vibro-acoustography, on the other hand, is not hampered by tissue density [3].

The above argument further justifies integrating VA and mammography. Integration of multimodality imaging has been widely used for generating more diagnostic and clinical values in medical imaging. Proper image registration and multimodality image fusion techniques need to be employed for high quality image integration. On the other hand, inaccurate image registration and incorrect localization of region of interest risks a potential impact on patients. Integrating images of the same target generated with different modalities has been investigated for various clinical images [13–15].

In a study published by Behrenbruch et al. [13], fusion of the high-resolution structural information available from mammography with the functional data acquired from MRI imaging is proposed to offer a better pathological indicator such as calcifications. It has been reported that some tissue details that are not visible in contrast-enhanced MRI can be recognized in the fused images [13].

Prior to the fusion process, it is important to apply a robust registration technique to align images, from a single or from different modalities [14, 15]. By image registration, the correspondences between the images can be seen more easily and the clinicians can get maximum amount of information from the images [15].

This paper describes a reliable image registration technique for aligning VA images and X-ray mammogram. It also proposes principles of integrating VA and X-ray images after performing a reliable registration. As a result, a software-based image alignment tool was designed for integrating the two modalities and facilitating the diagnosis process. Assuming that these two completely different modalities should provide relatively independent information about the breast tissues, the ultimate aim of this research is to enhance the quality of image interpretation and further improving the effectiveness of breast cancer detection.

2. METHODS AND MATERIALS

2.1. Experiment setup

Figure 1 shows a schematic of combined vibro-acoustography-mammography system used for image generation. X-ray images are generated by Mammotest/Mammovision (Fischer Imaging Corporation's HF-X Mammography) system equipped with compression paddle to immobilize the target (breast). Vibro-acoustography transducer is mounted in a water tank attached to the mammography system (see Figure 1). This transducer is designed with two arrays (two compact transducers) driven by two continuous-wave or tone-burst signals at slightly different frequencies [16]. A window (104-by-80 mm) covered by a flexible membrane is mounted on water tank wall to allow both X-ray and the ultrasound beams to pass through to the target. The patient breast is covered by ultrasound coupling gel before it is placed in contact with the membrane. The imaging window for either imaging method is a 50-by-50 mm square. Within this area, the VA collects 256-by-256 points of the target by scanning the breast.

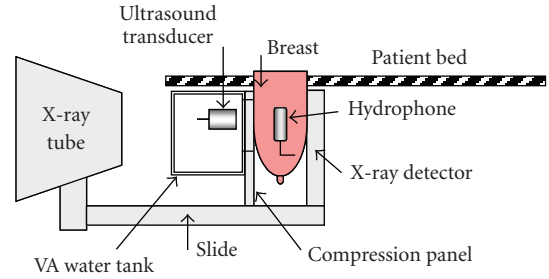


Figure 1: Schematic of combined vibro-acoustography-mammography system, "reproduced with permission from [3]."

2.2. Principles of vibro-acoustography

Vibro-acoustography is based on low-frequency vibrations induced in the object due to the radiation force of ultrasound. Radiation force is generated by a change in the spatial distribution of the energy density of an incident ultrasound beam. The change of energy density of the impinging ultrasound may be due to energy absorption, scattering, and reflection.

The magnitude of radiation force depends on a number of parameters, including the scattering and absorption properties of the object. For a planar object insonified with a plane wave, the radiation force is related to the power reflection coefficient of the object [3, 7].

In vibro-acoustography, two continuous wave (CW) ultrasound beams of slightly different frequencies, $f0 = 3\,\text{MHz}$ and $f0 + \Delta f$, are used with $\Delta f = 30\,\text{KHz}$. The two beams are focused at a joint focal point. At this point, the combined ultrasound field energy density is sinusoidally modulated. Therefore, the field generates a highly localized oscillatory radiation force at frequency Δf, when interacting with an object. The radiation stress is normally confined to a small volume of the object, which acts as an oscillating point force placed remotely inside the object [3].

The radiation force vibrates the object at frequency Δf and generates in a secondary acoustic field (acoustic emission) with the same frequency that propagates in the object. As the ultrasound beam is scanned across the object, an audio hydrophone can detect the acoustic emission. The hydrophone signal is recorded and its amplified amplitude is mapped into an image [3, 7].

Other VA system parameters are: the resolution of the VA setup = 0.7 mm in transverse plane and about 9 mm in axial direction, spatial sampling interval is 0.2 mm in both directions in the scan plane, the scan time is about 7 minutes for a 256×256 image, and ultrasound intensity at the focal point = $700\,\text{mW/cm}^2$ in compliance with the FDA recommendation for in vivo ultrasound. The audio hydrophone is covered with acoustic gel and placed in contact with the side of breast. The ultrasound transducer can scan the breast through a window. To take an X-ray, the water tank is emptied and the transducer is moved out of X-ray path [5].

The vibro-acoustogram illustrates two types of information about the object: (i) ultrasonic properties of the object, such as the scattering and power absorption characteristics; (ii) the dynamic characteristics of the object at frequency Δf, which also relates to the boundary conditions and the coupling to the surrounding medium. The acoustic emission is also influenced by the surrounding medium as it propagates from the focal point of the transducer to the hydrophone. As the transducer scans the object, the distance between the focal point and the hydrophone varies. However, because the attenuation of the tissue at Δf is not significant, variations of the acoustic emission due to variations of focal point-to-hydrophone distance is negligible. The ultrasonic properties are those that are also present in conventional ultrasound imaging. The dynamic characteristics at Δf, which are related to object stiffness, can be described in terms of object mechanical impedance at frequency Δf. Such information is not available from conventional ultrasonography [3].

Speckle is the snowy pattern which results from random interference of the scattered ultrasound field. Speckles reduce the contrasts of conventional ultrasound images and often limit detection of small structures, such as breast micro calcifications in tissue. Vibro-acoustography on the other hand uses the acoustic emission signal, which is at a low frequency, thus the resulting images are speckle free and have high contrast. This feature makes vibro-acoustography suitable for detection of breast micro calcifications [3].

2.3. *Theory of image registration*

The VA beams stay parallel as the object is scanned and can be used to generate 3D images of an object by integrating its image slices acquired at different depths. The VA beams scan across the object while focusing at a fixed depth. Various slices of the object are scanned at different depths and the corresponding VA images are formed. On the other hand, X-ray beams form a perspective projection image and generate a 2D image of the object without any depth/thickness information. Consequently, the VA images have a fixed magnification for targets at different depths but the X-ray images show variable magnifications depending on target depth. Figure 2 shows a schematic of beams and target coverage areas (filled ovals) in these two modalities. It is shown, in Figure 2(b), that the screened images of X-ray (large oval) and VA (small oval) of the same target are different. Therefore, for the purpose of overlaying the images, a geometric calibration is required to resize and shift one image to match with the other one.

Ideally, by registration, the size and position of VA images should be geometrically transformed to match exactly with mammography images pixel-by-pixel. In this work, we adopted an algorithm based on control points (CPs) and found an equation to adjust the registration transformation with a magnification factor (MF) for different depths. CPs can be identified in an image as pixel points related to user added markers or existing image spots [17]. The MF is defined by the ratio of W and Y_1 where W is the dimension of the image at the plane of the X-ray detector and Y_1 is the dimension of the target (see Figure 2(a)) by X-ray. Equation (1) expresses MF as the ratio of W and Y_1 and also versus SID and D_1 as follows:

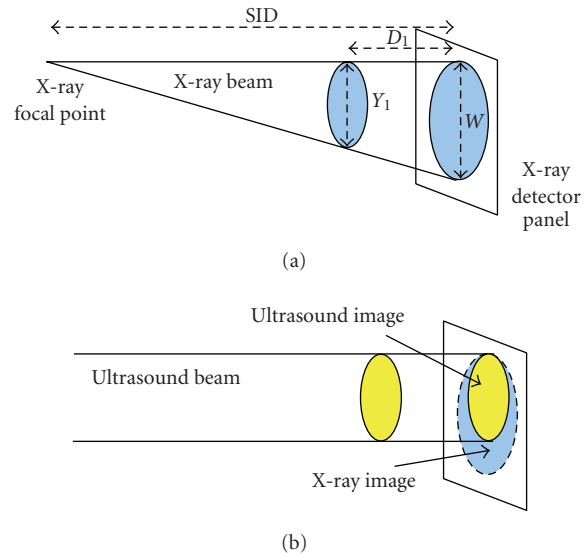


Figure 2: Schematic of beams and target coverage areas in (a) X-ray; (b) VA.

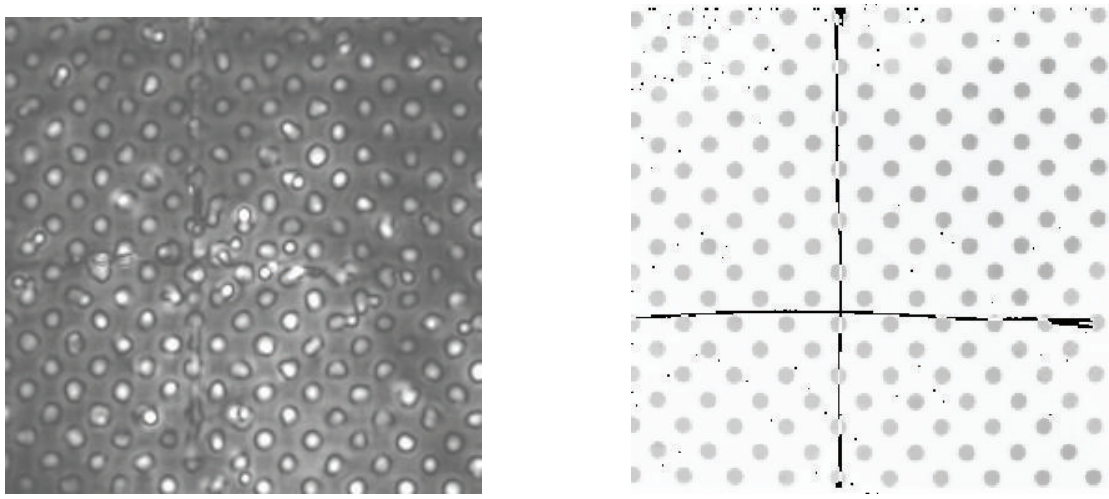
$$\mathrm{MF} = \frac{W}{Y_1} = \frac{\mathrm{SID}}{\mathrm{SID} - D_1}, \tag{1}$$

where SID is equal to 664 mm and D_1 is the distance of the target from the X-ray detector panel.

In Mammovision system, SID is defined as "source to image" or "focal spot to image receptor" distance which is equal to 664 mm in this case. For VA, the focal point is located at 70 mm distance from the transducer. The position of the transducer is changeable to focus at different depths in the object (target). As an example, any target located within a range of 10 mm to 100 mm distance from the X-ray detector panel can be scanned.

The sensitive part of the X-ray detector panel is a 50×50 mm square CCD screen. The stereotactic mammography system has target-focusing interface that can be used for measuring D_1. This system establishes depth of the target using stereo images taken by positioning the X-ray camera at +15 and −15 degree deviation related to its normal zeroth-degree position. The dept information is important in the understanding of the calibration system and error propagation.

An initial study of X-ray and VA images of a phantom was conducted to create a universal matrix of CPs. This matrix can be used in registering the breast images for the purpose of integrating two modalities. The phantom was fabricated from a perforated PVC board with two cross wires mounted as markers. It is used to mimic 2D targets and validate the registration method by calculating the target registration error (TRE) [18, 19]. The images acquired from the same phantom by VA and X-ray techniques are illustrated in Figures 3(a) and 3(b), respectively.



(a) (b)

Figure 3: Images of the phantom used to create the matrix of control points: (a) VA image; (b) X-ray image.

The CPs are selected along the image of two cross wires to ensure that the same points were selected in the X-ray and VA image. We used acoustic gel to couple the hydrophone to the PVC board to reduce bubbles which appear as white points in the VA images. The mid grey points on Figure 3(b) are part of the X-ray image and can be considered as pepper noise. The X-ray images are rescaled from resolution of 1024-by-1024 to 256-by-256 to match the number of pixels with the VA pixels. However, the positions of similar objects (such as CPs) are still different in the two images.

Figure 4 shows the normalized positions of similar points and the distance between two cross lines in two images. By assigning (1.00, 1.00) to the position of the very top-right pixel (256, 256) in Figure 4, we can normalize every pixel position of the image.

To establish correspondence between the phantom image obtained by VA with its X-ray image, we plotted a "black square" to illustrate the position of the X-ray frame (in Figure 4). Two cross wires mounted as markers are plotted by white cross lines for the VA image and blue lines for the corresponding X-ray. Figure 4 shows the overlap between the coverage areas of each modality. It shows that by repositioning the VA transducer, we can maximize the common coverage area by two modalities. In addition, due to different magnifications of these two methods still registration is required.

The CPs in X-ray image are selected as base points and similar points in the corresponding VA image are called input points. Eight CPs that can be clearly located in both images were selected to generate two 8-by-2 matrices of base points and input points. The number of CPs and their locations are flexible and can be optimized empirically. The locations of CPs which are the center of eight selected holes in the phantom are indicated by 4-point stars in Figure 4. These matrices are contained the X and Y coordinates of the selected CPs in the X-ray and the corresponding VA images.

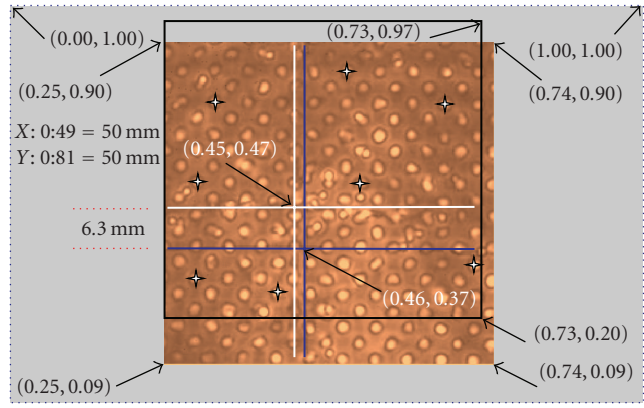


Figure 4: The normalized positions of similar points and the distance between two crossed lines in VA and X-ray images.

One of the key components of the registration process is to find a mathematical transformation to map the input image to the base image. A second-order polynomial, which is invariant to rotation and translation, was used to infer a spatial transformation of the X and Y pair of each pixel. For VA registration, this transformation can be applied to the base and input points to map any new grayscale VA image into its corresponding X-ray. The rotation and translation of VA images are mathematically assigned by the transformation, which is then used to create the fusion display of the original grayscale medical images.

The second-order polynomial transformation maps X_b and Y_b of the base-point matrix to X_i and Y_i of the input-points matrix according to (2) [20]:

$$[X_i, Y_i] = [1, X_b, Y_b, X_b * Y_b, X_b^2, Y_b^2] * \text{Inv}\, T, \tag{2}$$

where $*$ is the multiplication sign.

To specify all coefficients of Inv T with the size of 6-by-2, at least 6 CPs are required to solve the inverse of the second-order polynomial, Inv T [20].

We chose eight CPs and used normalized cross-correlation to adjust each pair of CPs to solve the second-order polynomial.

To adjust the X-ray image for different depths, it is scaled up by MF before applying the transformation of (2). Two-dimensional interpolation techniques such as nearest-neighbor, bilinear, and bicubic interpolation can be used to estimate the image value at a location in between image pixels.

Inaccurate image registration of VA and mammogram may cause incorrect localization of region of interest and may have a potential impact on interpretation of diagnostic information [18]. It is possible to find an adaptive transformation to correct the magnification problem and limit the TRE. This is a suggestion for future research.

2.4. Registration of VA breast images and X-ray mammography

The results of registration of the phantom images including the matrices of base points and input points were used to spatially transform in vivo breast VA images from volunteer. Figure 5(a) shows the coronal view X-ray mammography of the breast of a patient with a large calcification enclosed in a fibroadenoma region marked with arrows according to a radiologist diagnosis. The VA breast image obtained by scanning the same subject with focal point positioned at 4 cm from the skin (20 mm from the CCD screen) is shown in Figure 5(b).

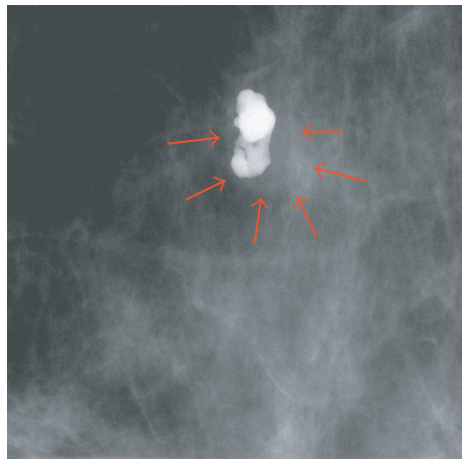
To map the VA image pixel-by-pixel on to the mammogram, a modified polynomial transformation was applied. The resultant registered image is shown in Figure 5(c). We imported the arrows from Figure 5(a) to this figure, at positions that matched the arrows in X-ray image to show the corresponding marking area in the X-ray image. The thick dark band at the top of registered VA image (see Figure 5(c)) is the area that was not covered by VA because the VA image frame did not match exactly. On the other hand, the bottom part of the target, which was scanned by VA, was not covered by X-ray. This error is caused due to misalignment of the X-ray and VA imaging windows. Figure 4 shows clearly the coverage area of each method. However, the rest of VA image is registered to match with its corresponding X-ray image.

2.5. Registration error

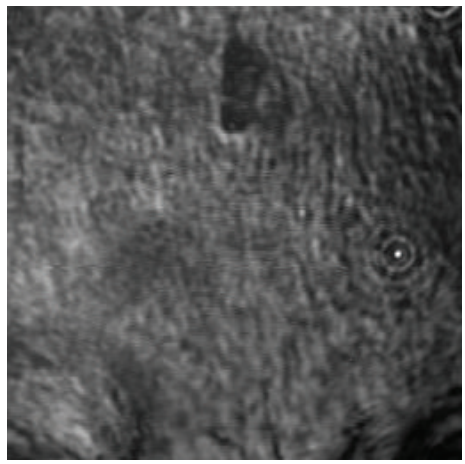
To evaluate the registration error, three different measures of registration accuracy can be defined as follows [21].

(i) Fiducial registration error (FRE), which is the value of the point-based registration cost function after registration.

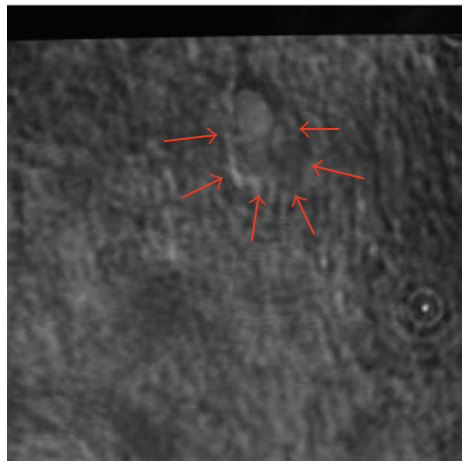
(ii) Surface registration error (SRE), which is the value of the surface-based registration cost function.



(a)



(b)



(c)

Figure 5: The breast image with fibroadenoma: (a) marked X-ray by an expert radiologist showing the fibroadenoma and the nearby calcification; (b) a $5 \times 5\,\text{cm}^2$ VA image of a breast including a large calcification and nearby fibroadenoma; (c) the resultant registered VA images with focal point at 20 mm from the CCD screen, "reproduced with permission from [3]."

Table 1: Results of MF_M calculations for different D_1.

D_1(mm)	MF	MF_m
0	1	0.9689
20	1.0311	1
40	1.0641	1.0330

(iii) Target registration error (TRE), which is defined as the distance between corresponding points other than those used to estimate the transformation parameters.

The TRE is a more objective measure of registration accuracy but it is difficult to quantify this error in an unmarked registered image. To measure this error, one marker needs to be assigned randomly on each patient as a target and four other markers as fiducials [21].

However, the FRE is a better parameter to measure the registration error in this study as no marker put on the patient. It measures the residual displacement between the points used for registration as the RMS (root mean square) error on the distance between the corresponding CPs of the registered VA and X-ray images,

$$\mathrm{FRE} = \frac{\sqrt{\sum_{n=1}^{N} d_n^2}}{N}, \tag{3}$$

where N is the total number of CPs and d_n is the minimum Euclidean distance between the nth CPs. The Euclidean distance, d, is the minimum distance between a base-point (X_b, Y_b) and input-point (X_i and Y_i) [22]:

$$d = \min\left\{\sqrt{(X_b - X_i)^2 + (Y_b - Y_i)^2}\right\}. \tag{4}$$

The FRE is an accumulation of different error components, such as MF, error due to the registration transformation, and error due to different characteristics of imaging techniques. It was found that the MF is the most dominate component of FRE. To minimize this error, we obtained matrices of CPs for the target located at D_1 equal to 20 mm, which is approximately the mean distance from the CCD screen to the middle of the breast. It is assumed that the maximum size of the compressed breast is less than 40 mm. Therefore, the MF, as expressed by (1), can be modified as follows:

$$\mathrm{MF}_m = (\mathrm{MF} - 0.0311). \tag{5}$$

MF_m is equal to one for the average distance (D_1 = 20 mm) and (MF = 1.0311). The results of MF_m calculations for the minimum (D_1 = 0 mm), average (D_1 = 20 mm), and maximum (D_1 = 40 mm) distance of the target are shown in Table 1. By selecting fixed CPs matrices for all target locations, the maximum error due to MF is 1.65 mm ($(1.033 - 1) \times 50 = 1.65$ mm) for a 50-by-50 mm^2 object.

2.6. Image fusion

Image fusion can be performed at three different levels: pixel level, feature level, and decision level [23]. We used pixel level image fusion techniques and for better visualization of the structural information contained in both images, it is decided to adopt a color-based method for fusing the registered images. The registered VA and X-ray images are assigned as the blue and red components of an RGB image, respectively. A zero matrix is assigned as the green-component of the RGB image. Figure 6(a) shows the resultant image of color-based fusion of the two primary images shown in Figure 5. This method generates an image with color-code information of each image which may be useful for diagnostic purposes.

To improve the quality of fused images, we used pixel level fusion with different ratio (R) of pixel values. Figure 6(b) shows the resultant image by fusing the two images using 50% of each image pixel value. The integrated image shows features of VA and X-ray in a single image. The calcification seen in the VA and X-ray mammography matches perfectly.

The proposed method for integrating of multimodality medical images allows extracting new information by fusing VA images at different depths with X-ray mammogram. User can select one VA image at a time from a file, containing VA images scanned at different depths of the object, and register it with a based mammogram. Finally, the registered VA image can be enhanced and fused with the base image of X-ray mammography.

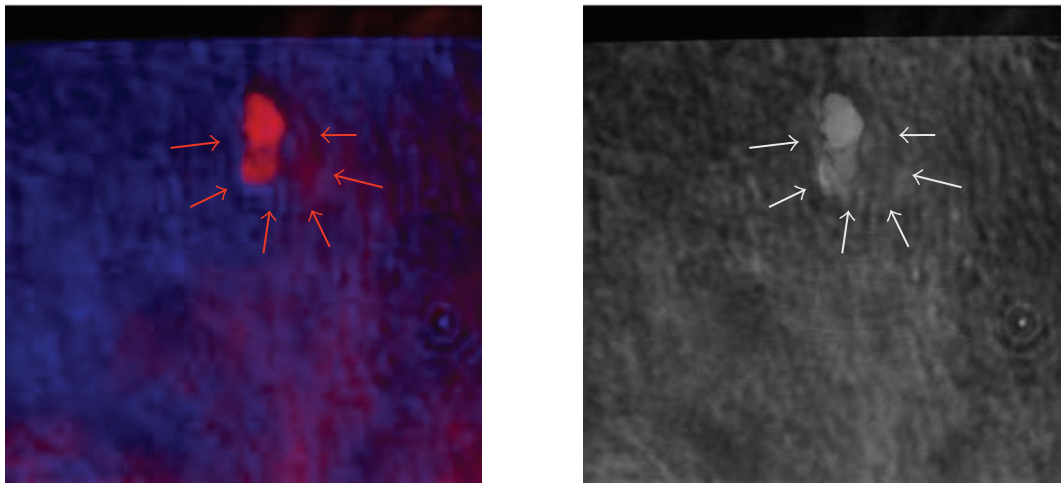
Figure 7(a) shows another scanned VA image taken at the depth of 5 mm of the CCD screen and Figure 7(b) shows the same image after fusing with the X-ray image which confirms the position of calcification. The VA image shows the fibroadenoma (marked by arrows) which is not clearly visible in the X-ray.

3. DISCUSSION AND CONCLUSIONS

Integrating images from two completely different modalities, VA and X-ray, using either color-based or pixel-value fusion techniques may generate more structural and diagnostic information. A color-based fusion technique may be more suitable for visualization of the structural information.

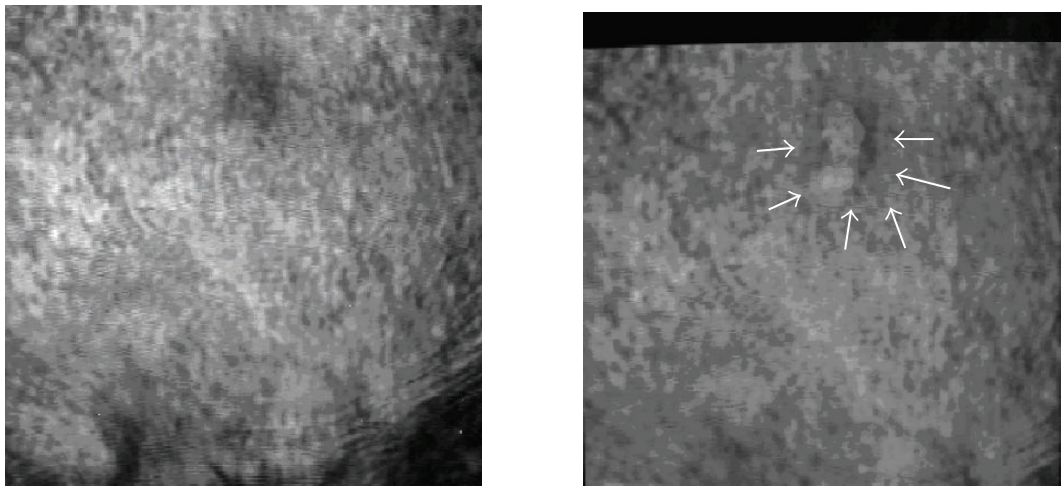
Here, we presented a method for integrating VA and X-ray (mammography). It is shown that, because X-ray image magnification varies with target depth, the registration transformation must be adjusted with a magnification factor for different target depths. In this work, we used a modified second-order polynomial, which leads to a scale/rotation/translation invariant paradigm for image registration. To validate the proposed registration method, we fused VA images at different depths with the X-ray mammogram and demonstrated that the detected classification area is located at the same position in both modalities.

In addition, a method of calculating the fiducial registration error (FRE) due to MF was proposed. By selecting the matrices of CPs for the target located at mid distance from the screen, the error of image registration can be limited to 1.65 mm. In most cases, the size of target is larger than the maximum error and it is positioned at around 20 mm from



(a) (b)

Figure 6: The fused breast image using the primary breast image with fibroadenoma as shown in Figure 5 by (a) color-based method; (b) pixel-level method with R = 50%.



(a) (b)

Figure 7: (a) The original VA image; (b) the fusion of VA with marked mammogram.

the CCD screen. Therefore, the registration error is not significant for such targets. We also note that the FRE of 1.65 is about twice the lateral resolution of the system (0.7 mm), indicating the possibility of one resolution cell error in image registration. However, for registering of small targets such as micro calcification, an adaptive transformation can be applied to reduce the FRE.

The fused image of two different modalities which is associated with an X-ray mammography, annotated by an expert radiologist, can be used to verify independent diagnostic information of the VA modality. Moreover, the aligned image would assist the users to gain maximum amount of information from X-ray mammogram and the VA modality.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge D. H. Whaley, for annotating the X-ray mammogram. Disclosure: some of the techniques presented here are patented by M. Fatemi. This work is supported by Grant no. EB00535 from NIH and Grant no BCTR0504550 from Susan G. Komen Breast Cancer Foundation.

REFERENCES

[1] S. Nass and J. Ball, "Improving breast imaging quality standards," in *Committee on Improving Mammography Quality Standards*, pp. 240 pages, National Research Council, National Academies Press, Washington, DC, USA, 2005.

[2] M. Fatemi and J. F. Greenleaf, "Vibro-acoustography: an imaging modality based on ultrasound-stimulated acoustic emission," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 12, pp. 6603–6608, 1999.

[3] A. Alizad, D. H. Whaley, J. F. Greenleaf, and M. Fatemi, "Potential applications of vibro-acoustography in breast imaging," *Technology in Cancer Research and Treatment*, vol. 4, no. 2, pp. 151–157, 2005.

[4] A. Alizad, L. E. Wold, J. F. Greenleaf, and M. Fatemi, "Imaging mass lesions by vibro-acoustography: modeling and experiments," *IEEE Transactions on Medical Imaging*, vol. 23, no. 9, pp. 1087–1093, 2004.

[5] A. Alizad, M. Fatemi, D. H. Whaley, and J. F. Greenleaf, "Application of vibro-acoustography for detection of calcified arteries in breast tissue," *Journal of Ultrasound in Medicine*, vol. 23, no. 2, pp. 267–273, 2004.

[6] M. Fatemi, A. Manduca, and J. F. Greenleaf, "Imaging elastic properties of biological tissues by low-frequency harmonic vibration," *Proceedings of the IEEE*, vol. 91, no. 10, pp. 1503–1519, 2003.

[7] M. Fatemi, L. E. Wold, A. Alizad, and J. F. Greenleaf, "Vibro-acoustic tissue mammography," *IEEE Transactions on Medical Imaging*, vol. 21, no. 1, pp. 1–8, 2002.

[8] M. Fatemi and J. F. Greenleaf, "Ultrasound-stimulated vibro-acoustic spectrography," *Science*, vol. 280, no. 5360, pp. 82–85, 1998.

[9] W. R. Hendee, "History and status of x-ray mammography," *Health Physics*, vol. 69, no. 5, pp. 636–648, 1995.

[10] R. Highnam and M. Brady, *Mammographic Image Analysis*, Kluwer Academic, Dordrecht, The Netherlands, 1999.

[11] H. Aichinger, J. Dierker, S. Joite-Barfuss, and M. Sabel, "Radiation exposure and image quality in x-ray diagnostic radiology," in *Physical Principles and Clinical Applications*, Springer, New York, NY, USA, 2004.

[12] A. Alizad, M. Fatemi, L. E. Wold, and J. F. Greenleaf, "Performance of vibro-acoustography in detecting microcalcifications in excised human breast tissue: a study of 74 tissue samples," *IEEE Transactions on Medical Imaging*, vol. 23, no. 3, pp. 307–312, 2004.

[13] C. P. Behrenbruch, K. Marias, P. A. Armitage, et al., "Fusion of contrast-enhanced breast MR and mammographic imaging data," *Medical Image Analysis*, vol. 7, no. 3, pp. 311–340, 2003.

[14] R. Kapoor, A. Dutta, D. Bagai, and T. S. Kamal, "Fusion for registration of medical images—a study," in *Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop (AIPR '03)*, pp. 180–185, Washington, DC, USA, October 2003.

[15] L. Hallpike and D. J. Hawkes, "Medical image registration: an overview," *Imaging*, vol. 14, no. 6, pp. 455–463, 2002.

[16] G. T. Silva, J. F. Greenleaf, and M. Fatemi, "Linear arrays for vibro-acoustography: a numerical simulation study," *Ultrasonic Imaging*, vol. 26, no. 1, pp. 1–17, 2004.

[17] W. Li and H. Leung, "A maximum likelihood approach for image registration using control point and intensity," *IEEE Transactions on Image Processing*, vol. 13, no. 8, pp. 1115–1127, 2004.

[18] B. C. Porter, D. J. Rubens, J. G. Strang, J. Smith, S. Totterman, and K. J. Parker, "Three-dimensional registration and fusion of ultrasound and MRI using major vessels as fiducial markers," *IEEE Transactions on Medical Imaging*, vol. 20, no. 4, pp. 354–359, 2001.

[19] D. J. Hawkes, "Registration methodology: introduction," in *Medical Image Registration*, J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, Eds., pp. 11–38, CRC Press, London, UK, 2001.

[20] A. Goshtasby, "Image registration by local approximation methods," *Image and Vision Computing*, vol. 6, no. 4, pp. 255–261, 1988.

[21] C. R. Maurer Jr., G. B. Aboutanos, B. M. Dawant, R. J. Maciunas, and J. M. Fitzpatrick, "Registration of 3-D images using weighted geometrical features," *IEEE Transactions on Medical Imaging*, vol. 15, no. 6, pp. 836–849, 1996.

[22] N. Paragios, M. Rousson, and V. Ramesh, "Non-rigid registration using distance functions," *Computer Vision and Image Understanding*, vol. 89, no. 2-3, pp. 142–165, 2003.

[23] A. Kapur, P. L. Carson, J. Eberhard, et al., "Combination of digital mammography with semi-automated 3D breast ultrasound," *Technology in Cancer Research and Treatment*, vol. 3, no. 4, pp. 325–334, 2004.

Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 54616, 9 pages
doi:10.1155/2007/54616

Research Article

Multimodality Imaging of the Peripheral Venous System

Diana Gaitini

Unit of Ultrasound, Department of Medical Imaging, Rambam Medical Center and Faculty of Medicine, Technion–Israel Institute of Technology, P.O. Box 9602, Haifa 31096, Israel

Received 17 September 2006; Accepted 28 September 2006

Recommended by Haim Azhari

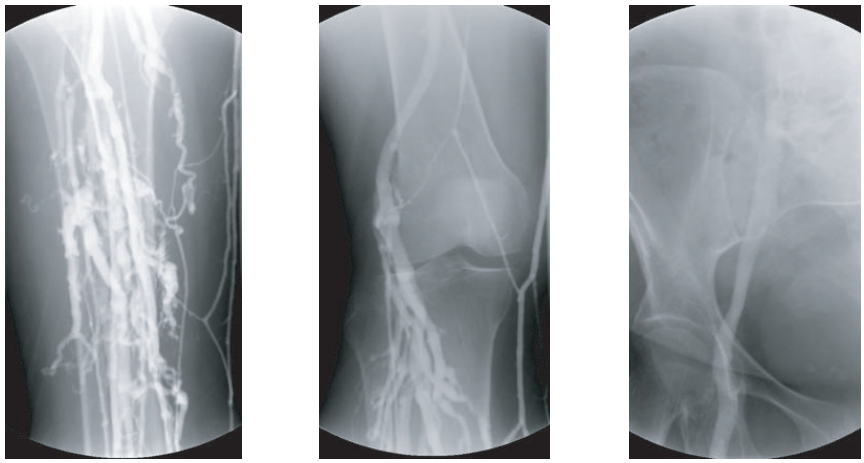
The purpose of this article is to review the spectrum of image-based diagnostic tools used in the investigation of suspected deep vein thrombosis (DVT). Summary of the experience gained by the author as well as relevant publications, regarding vein imaging modalities taken from a computerized database, was reviewed. The imaging modalities reviewed include phlebography, color Doppler duplex ultrasonography (CDDUS), computerized tomography angiography (CTA) and venography (CTV), magnetic resonance venography (MRV), and radionuclide venography (RNV). CDDUS is recommended as the modality of choice for the diagnosis of DVT. A strategy combining clinical score and D-dimer test refines the selection of patients. Phlebography is reserved for discrepant noninvasive studies.

1. INTRODUCTION

Deep venous thrombosis (DVT) of the lower limb is a common and life-threatening condition. The incidence in the United States is estimated at 70–100 000 new cases/year with as many as 200 000 hospitalizations/year. It carries a risk of pulmonary embolism (PE) and the development of postthrombotic syndrome. The incidence of PE is calculated at 600 000 cases/year, 100 000 of them are fatal [1–3]. Risk factors for lower extremity acute venous occlusion range from, prolonged immobilization to hypercoagulability syndromes, trauma, and malignancy. Venous thromboembolism (VTE) is a leading cause of morbidity and mortality during pregnancy and puerperium, and is second only to hemorrhage as the commonest cause of maternal death during pregnancy [4]. There is a fivefold increased risk of VTE in pregnant women compared with nonpregnant women of a similar age [5].

Malignancy and central venous lines are major risk factors of upper extremity thrombosis (UEDVT) with predicted poor survival. The increasing use of indwelling central venous catheters (CVC) for transparietal feeding, fluid administration, and chemotherapy has resulted in an increased prevalence of upper extremity venous thrombosis, although, the rate of catheter-associated thrombosis decreased in recent years thanks to improvement in biocompatibility and better insertion and maintenance techniques [6]. UEDVT may be asymptomatic or the clinical manifestations are not specific, presenting with arm or neck swelling or pain. In more than half of the cases objective methods of examination are negative for thrombosis. Pulmonary embolism secondary to UEDVT, sometimes a lethal complication, is not unusual and has been reported in a comparable prevalence to lower extremity thrombosis. Other significant complications of UEDVT are loss of vascular access, superior vena cava syndrome, and postthrombotic venous insufficiency [7–9].

The clinical complications (from postthrombotic syndrome to fatal pulmonary embolism) as well the risk of anticoagulant treatment require a precise diagnosis of DVT. The clinical diagnosis is unreliable: only 20–30% of symptomatic patients have proven DVT and 90% of fatal PE are asymptomatic for DVT [10]. Objective methods of examination are demanded to reach an accurate diagnosis. Phlebography, computerized tomography angiography followed by venography (CTA-CTV) and radionuclide venography (RNV) are invasive or semi invasive tests. Color Doppler duplex ultrasonography (CDDUS), and magnetic resonance venography (MRV) are noninvasive methods. This paper highlights the potential risks and benefits of each of these techniques and presents the advantages, disadvantages, and accuracy of the different imaging modalities. An appropriate imaging algorithm for the diagnosis of DVT is presented. The use of clinical pretest probability scoring and diagnostic algorithms can help identify patients



(a) (b) (c)

Figure 1: Normal phlebography. (a) Calf veins. (b) Popliteal vein. (c) Femoral vein at the groin and iliac vein at the pelvis.

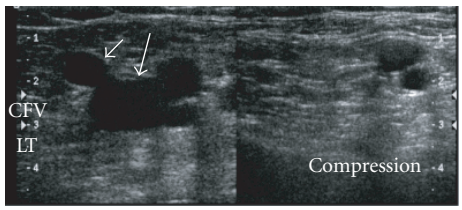
requiring further investigation for suspected venous thromboembolism (VTE).

2. PHLEBOGRAPHY

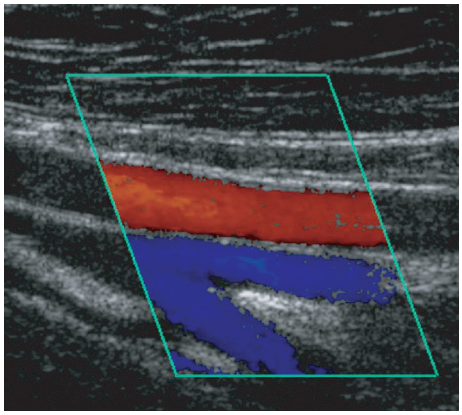
Phlebography (also called venography, ascending contrast phlebography, or contrast venography) is still considered the gold standard in the diagnosis of peripheral DVT; it is the most accurate test with a nearly 100% sensitivity and specificity [11]. This X-ray examination provides an image of the limb veins after contrast material is injected into a distal vein (Figures 1(a)–1(c)) [12]. Main phlebographic findings are persistent filling defect, abrupt interruption of contrast in a vein, lack of opacification in all or some deep veins, and flow diversion with opacification of collateral branches [13]. Venography relies on the anatomy of the venous system, lacking physiological information. It is painful; expensive, exposes the patient to a fairly high dose of radiation; and can cause complications related to nephrotoxicity and allergic reactions to iodinated contrast agents. It also carries a risk for post venographic phlebitis [14, 15]. In about 5% of cases, there are technical problems in conducting the test. Due to its invasive nature and the risk of complications, it cannot be used neither as a routine test for the diagnosis of symptomatic DVT nor as a screening tool in asymptomatic patients at high risk for DVT. Peripheral phlebography is performed when the noninvasive examination color Doppler US and duplex Doppler is doubtful or technically limited, such in suspected thrombosis of iliac vein, innominate vein, or superior vena cava [13].

3. COLOR DOPPLER DUPLEX ULTRASONOGRAPHY

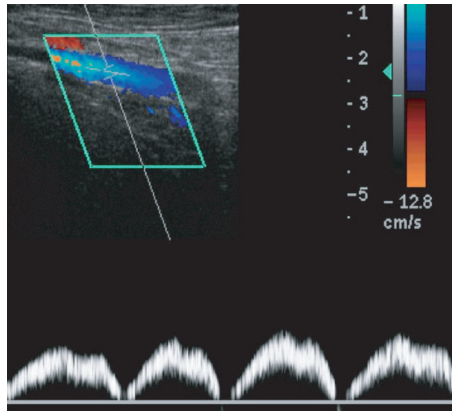
CDDUS is the initial test of choice for diagnosis of acute DVT due to its high accuracy, relatively low cost, portability, widespread, and lack of ionizing radiation [16]. B-mode ultrasound with Doppler color and duplex is the only noninvasive imaging test that combines anatomy and physiology of the veins by visualization of vein morphology and the map of flow velocity and direction. It is required as the primary instrumentation for peripheral venous testing according to the standards of the Intersocietal Commission for the Accreditation of Vascular Laboratories (ICAVL) [17]. CDDUS for the diagnosis of limb vein thrombosis uses a combination of gray-scale, compression, color, and spectral Doppler sonography. Color and spectral Doppler analysis are useful in the diagnostic evaluation of DVT but are best considered as adjuncts to the conventional compression ultrasound examination. The examination is performed by a high-resolution transducer of 7–10 MHz; a lower frequency-4–8 MHz is required for the obese patient, the edematous limb, and the pelvic veins. The veins scanned comprise the deep venous system—femoral vein at the groin and along the thigh, popliteal vein, and tibioperoneal trunk at the upper calf—and the confluence of the superficial great saphenous vein with the femoral vein. The deep calf veins are usually examined when localized pain or swelling is present. CDDUS findings of the normal vein are sonolucent lumen, easily compressible with a slight pressure exerted by the probe and centripetal nonpulsatile flow, with respiratory phasicity and augmentation after Valsalva performance (Figures 2(a)–2(c)). An echogenic lumen, depending on thrombus age, uncompressible and flow devoid is diagnostic of a thrombotic vein (Figures 3(a)–3(c)) [18–24]. The main aim of CDDUS is to confirm or exclude vein thrombosis. Further information includes thrombus extent and characterization—fresh or organized, free floating or attached, and partial or totally occlusive—that have prognostic value for the development of pulmonary embolism and post-thrombotic syndrome. Patients with proximal DVT tend to present a slower and incomplete resolution of thrombus and to develop a more severe post-thrombotic syndrome due to deep venous reflux [25]. Free floating thrombus carries



(a)

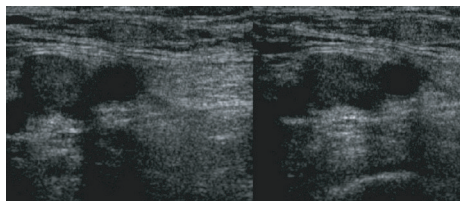


(b)

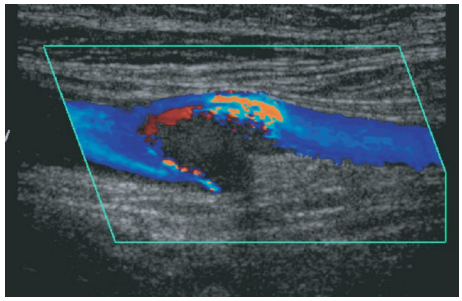


(c)

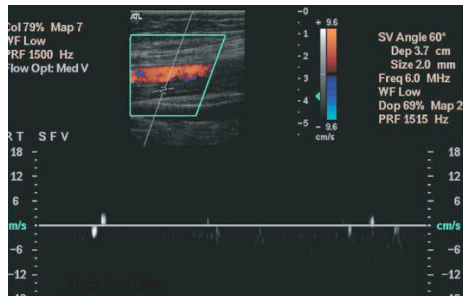
Figure 2: Normal vein. (a) Sonolucent lumen, easily compressible with a slight pressure exerted by the probe. Left side: before compression; right side: during compression; only the arteries remain visible. Large arrow: common femoral vein (CFV); short arrow: great saphenous vein. (b) Flow in femoral artery and veins at the level of the bifurcation. (c) Centripetal nonpulsatile flow in femoral vein, with respiratory phasicity.



(a)



(b)



(c)

Figure 3: Thrombotic vein. (a) Echogenic lumen, enlarged, noncompressible vein. Left side: before compression; right side: during pressure exerted by the probe, the vein does not collapse. (b) Thrombus at the bifurcation of the femoral vein, seen as color void and turbulent surrounding flow. (c) No flow demonstrated on duplex in a thrombotic femoral vein.

an increased risk of pulmonary embolism, although floating thrombus tends to attach to the vein wall or resolve, not warranting any specific therapeutic procedure [26]. Further diagnostic aims are to detect alternative disorders such as popliteal Baker's cyst, hematoma, aneurysm, pseudoaneurysm, lymphadenopathy, or other tumors, known as "pseudothrombophlebitis," mimicking DVT. The incidence of these alternative diagnoses is 11–18% [27]. A bilateral examination is indicated when high-risk patients are screened and in the workflow of suspected PE in patients with risk factors for DVT. Due to its high specificity, complete ultrasound examination of the proximal and distal veins at least down to the level of the popliteal trifurcation allows withholding anticoagulant therapy without the risk of major complications. Isolated calf vein thrombosis does not carry a significant adverse outcome; scanning the calf with localized symptoms or physical findings is cost-effective. A repeat examination is warranted if the clinical findings worsen; otherwise, a single examination is enough [28]. The sensitivity and specificity of USD for the diagnosis of DVT in symptomatic patients is very high. Compressibility under probe pressure (CUS) is the most accurate test; for proximal DVT, femoral, and popliteal veins, compression US reached a sensitivity of 97 to 100% and a specificity of 98 to 99%. For isolated calf DVT, the sensitivity dropped to 50–70% and the specificity to 60%. An echogenic lumen has a low sensitivity of about 50% for both proximal and calf DVT, due to the low echogenicity of the fresh thrombus [29–32]. In a meta-analysis of 100 cohort studies that compared Duplex US to contrast venography in patients with suspected DVT; the sensitivity for proximal DVT was 96.5%, for distal calf DVT, 71.2% and specificity of 94.3%; the sensitivity improved in the recent years probably due to equipment development, US technique used, and operator expertise [33].

Ultrasonography is the primary imaging modality also for the diagnosis of upper-extremity thrombosis (UEDVT). The veins examined include the deep system—internal jugular, subclavian, axillary, and brachial veins. The superficial veins—cephalic and basilica—are scanned in case of peripherally inserted catheter-related suspected thrombosis. The fresh clot may be not visualized and the diagnosis done on the presence of a vein enlarged and rigid, without changes on respiratory phases or respiratory maneuvers. Useful findings to rule-out thrombosis are an echo-free compressible vein, normal response to respiratory maneuvers-vein collapse on brief deep inspiration (sniff test), and enlargement on Valsalva test normal color Doppler and biphasic spectral display on duplex sonography [34–37]. The main obstacle for the diagnosis of UEDVT is the presence of overlying bones on the medial subclavian vein and centrally located veins, innominate and superior vena cava, that makes them difficult to visualize and impossible to directly assess by compression techniques.

Spectral Doppler abnormalities in the subclavian vein may be predictable for central occlusions. Flow void on color Doppler and a dampened nonpulsatile and nonphasic flow on duplex examination are diagnostic for a central venous thrombosis [38]. A reversed flow in the jugular vein may indicate thrombosis in the innominate vein with the internal jugular vein serving as a collateral pathway. Patel et al. [34] related a 100% positive predictive value and 91% negative predictive value for sonography in the diagnosis of complete central occlusions. Small nonobstructive thrombus may remain undiagnosed and large collateral veins misinterpreted as a normal vein, leading to false negative results. To overcome some of the limitations of US examination of the upper limb veins, a small footprint sector transducer from a supraclavicular or suprasternal approach may be of aid. CDDUS is a reliable method for diagnosing CVC-related thrombosis of the upper limb veins especially if several parameters are evaluated in combination [39]. High diagnostic accuracy of UEDVT was found in 6 prospective studies, with a sensitivity of 78–100% and a specificity of 82–100% [8, 40–44]. False positive results were unusual. A sensitivity of 100% and a specificity of 94% for compression US and color Doppler US for UEDVT using venography as the reference test ware reported by Prandoni et al. [44].

Chronic thrombosis in a patient with long-term catheterization is more challenging, as enlargement of the thrombotic lumen is not present. Color Doppler is even more useful in chronic thrombosis detecting collateral veins and an echogenic, flow void, and small caliber central vein. Large veins in an unusual anatomic position and without the accompanying artery must be recognized as enlarged collaterals and not be mistaken for the main vein. Aliasing due to high velocities and high pulsatility in the stenosed areas in comparison to dampened peripheral waveforms are additional diagnostic parameters. Frozen valve leaflets and echogenic synerchias may be seen as sequels of previous thrombosis [35, 37, 45]. In any case, the diagnosis of catheter-associated deep venous thrombosis may be difficult. Doppler ultrasound has a lower accuracy in this setting than it does in lower extremity venous thrombosis [46].

A particular different issue is acute on chronic thrombosis. The enlarged vein with hypoechoic lumen represents an acute process. Recurrent thrombosis is a challenging diagnosis for all imaging modalities. Comparison with a baseline examination may be helpful in these cases.

The clinical diagnosis of DVT is unreliable, but clinical prediction rules based on signs and symptoms do facilitate the categorization of patients into high, low, or medium risk categories [47]. A diagnostic strategy combining clinical score, D-dimer test, and compression US may refine the selection of patients. D-dimer assays have a high negative predictive value in patients with suspected VTE and can exclude the diagnosis. Based on clinical score and D-dimer test, venous US will be performed in patients with a high clinical score, an elevated D-dimer, or both (Figure 4).

Screening patients with plasma D-dimer and ultrasonography of the lower limbs may be the most cost-effective strategy. Ascending venography is reserved for patients with negative or equivocal CDDUS results and a high clinical probability of DVT [28, 48–50]. In the current state of the art, CDDUS is the modality of choice for the diagnosis of DVT. The appropriate examination is compression color duplex ultrasound of the complete venous system, including the

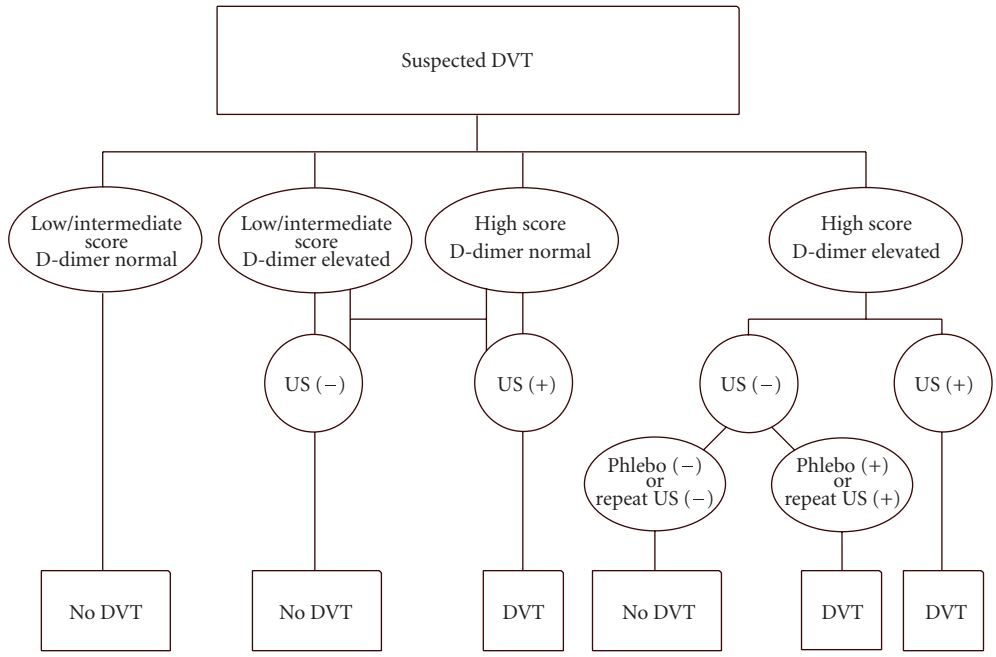


Figure 4: Algorithm for the diagnosis of DVT in symptomatic patients. By applying a diagnostic strategy based on the clinical score and D-dimer test; venous USD is performed in patients with a high clinical score, an elevated D-dimer, or both. The appropriate examination is compression color duplex ultrasound of the complete venous system, including the distal veins, when focal symptoms or physical findings are present and bilateral examination in the high-risk patient. Contrast venography is reserved for a minority of cases. Modified from Mantoni M. Ultrasound of limb veins. Eur Radio 11 : 1557-62, 2001 (with author's permission).

distal veins when focal symptoms or physical findings are present and a bilateral examination in the high-risk patient. It is an accurate examination and allows an early and safe diagnosis of thrombosis without straining the patients. It is the main diagnostic tool in symptomatic patients and in screening asymptomatic DVT in specific high-risk populations. Pitfalls and limitations of venous ultrasound are related to veins anatomy, flow changes, technical limitations, and operator expertise.

4. COMPUTERIZED TOMOGRAPHY ANGIOGRAPHY AND VENOGRAPHY

Multidetector CTA, combined with venous-phase imaging (CTA-CTV), can accurately diagnose a pelvic vein or inferior vena cava occlusion, sometimes the source of significant pulmonary emboli. Multidectector helical CT (MDCT) of the chest (100–140 mL of contrast medium injected at a rate of 3 mL/s) is followed by venous-phase imaging CT of the lower limbs without any additional contrast medium injection [51]. Indirect MDCT venography is acquired from the upper calves to the mid-abdomen. Thrombosis appears as a hypodense mass sometimes encircled by the hyperdense rim of contrast medium. The reported specificity and sensitivity compared with ultrasound is variable [52]. Coche et al. [51] compared the results of CT venography for diagnosing DVT with those of Doppler sonography and phlebography or repeated focalized sonography in case of discrepancy. Sensitivity and specificity of CTV were 93% and 97%, respectively (kappa = 0.88). CT venography in addition to CT pulmonary angiography is a relatively accurate method for evaluation of femoropopliteal venous thrombosis. In a comparative study between CTA-CTV and sonography, Garg et al. [53] found a 100% sensitivity, 97% specificity, 100% negative predictive value, and 71% positive predictive value for CTV. Satisfactory or good quality CT venography examination was obtained in 97% of the studies. Two CT venography studies had false-positive findings due to flow artifacts. The authors concluded that combined CT pulmonary angiography and CT venography may be more efficacious than sonography or two separate examinations in the selected patients. In another trial, CT venography had 93% accuracy compared with sonography in identifying deep venous thrombosis. However, the positive predictive value of CTV was only 67%, suggesting that sonography should be used to confirm the presence of isolated DVT before anticoagulation is initiated. CT venography interpretation should be performed with knowledge of certain pitfalls [54].

The prospective investigation of pulmonary embolism diagnosis II trial was conducted to investigate the accuracy of MDCTA alone and combined with venous-phase imaging (CTA-CTV) for the diagnosis of acute pulmonary embolism [55].

MDCTA alone had 83% sensitivity, 96% specificity, and positive predictive value with a concordantly high or low probability on clinical assessment. CTA-CTV for PE had 90% sensitivity and 95% specificity and was nondiagnostic with a discordant clinical probability like MDCTA alone. Missing

diagnoses were due to poor image quality of either CTA or CTV. According to this trial, MDCTA-CTV has a higher diagnostic sensitivity than does CTA alone with similar specificity in patients with suspected PE. The predictive value of both of them is high with a concordant clinical assessment, but additional testing is necessary when the clinical probability is inconsistent with the imaging results. According to Cham et al. [56], a substantial number of patients suspected to have PE had DVT in the absence of PE. The combined technique of pulmonary CTA-indirect CTV has been shown to identify DVT in up to 18% of patients with suspected PE who have no evidence of emboli on CTPA and thus could have a significant effect on patient care. Indirect MDCT venography is as accurate as sonography in the diagnosis of femoropopliteal DVT and can further reveal thrombus in large pelvis veins and the inferior vena cava, an important advantage over sonographic screening for DVT [57], although the technique is slightly more time consuming (up to 4 min delay after contrast injection) and has an increased radiation dose [58].

5. MAGNETIC RESONANCE VENOGRAPHY

Two-dimensional time-of-flight venography (TOF-MRV) is the technique of choice for magnetic resonance venography. Studies may be performed without contrast and can depict emboli as filling defects or directly detect the thrombus. MR direct thrombus imaging (MR-DTI) is a novel technique which detects metahemoglobin, allowing direct visualization of pulmonary emboli and simultaneous imaging of the legs without the need for intravenous contrast. This technique uses a T1-weighted gradient-echo sequence, with a preexcitation radio-frequency pulse to abolish fat signal, and an inversion recovery time chosen to nullify signal from flowing blood to maximize thrombus conspicuity. The technique is 98% sensitive and 96% specific for diagnosing DVT when compared with ultrasound and contrast venography. Early data suggest that MR-DTI is also highly accurate in detection of PE and the safety of withholding treatment on the basis of MR-DTI alone is currently being evaluated [59]. Acute occlusion of the pelvic veins and the inferior vena cava, often due to extension from the femoropopliteal system, represents a major risk for PE. Color flow Doppler imaging is often limited for the diagnosis of iliocaval thrombosis owing to obesity and bowel gas. Both CT scans and MR imaging can accurately diagnose acute pelvic vein or inferior vena cava occlusion and are as well helpful in diagnosing central chest vein occlusion. MRI is preferred because it is noninvasive, does not require contrast agent, carries no exposure to ionizing radiation, that is definitively demanded for pregnant women, and is highly accurate and reproducible [60].

Furthermore, MRV can differentiate an acute occlusion from chronic thrombus. In a study designed to evaluate the diagnostic value of MRV and color Doppler US in the assessment of DVT compared with contrast-enhanced venography, MRV was 100% sensitive and 100% specific in the diagnosis of DVT above the knee. Color Doppler imaging depicted 13 of 15 cases of DVT and 5 of 6 venous examinations that had normal results, yielding sensitivity and specificity of 87% and 83%, respectively. The differences in sensitivity and specificity between MRV and color Doppler US were not statistically significant [61]. In a recent meta-analysis to estimate the diagnostic accuracy of MRV for DVT, the pooled estimate of sensitivity was 91.5% (95% CI: 87.5–94.5%) and the pooled estimate of specificity was 94.8% (95% CI: 92.6–96.5%). Sensitivity for proximal DVT was higher than sensitivity for distal DVT (93.9% versus 62.1%) [62]. MR venography seems to be more accurate than color Doppler sonography in detecting the extension of deep venous thrombosis. Shankar et al. [63] performed two-dimensional gated inflow and phase contrast MRV in children with suspected upper extremity CVC-related thrombosis, to assess the extent of venous thrombosis and to locate patent veins for replacement central venous catheter. MRV was more accurate than Doppler ultrasonography and contrast studies for defining the extent of venous thrombosis. MRV correctly showed venous anatomy and patency for reinsertion of CVC. MRV is considered medically indicated for evaluation of venous thrombosis or occlusion in the large systemic veins (e.g., superior vena cava, subclavian, or other deep veins in the chest), for differentiation of tumor thrombus and blood clot and diagnosis of superior vena cava syndrome. The peer reviewed medical literature has not established MRV to be superior to duplex ultrasonography for diagnosis of deep vein thrombosis in the arms or legs. MRV has not been shown to be superior to US for lower limb DVT, except in imaging the deep femoral and hypogastric vessels. However, information about these vessels is not needed for management decisions, except in patients with pulmonary emboli where the source of the emboli has not been identified by ultrasonography [64]. MRV has the potential to be used as a stand-alone test for DVT but requires further evaluation. Therefore it is considered to be experimental and investigational for this application. Due to its high cost and limited availability, MRV should be reserved to diagnose DVT in patients for whom ultrasound examination is inappropriate or unfeasible [62] and to replace venography and CTV in pregnant women and patients with contraindications to iodinated contrast media injection.

6. NUCLEAR MEDICINE VENOGRAPHY

The radionuclide investigation of DVT includes such techniques as radionuclide venography and thrombus-avid scintigraphy. Although these methods have not been as thoroughly evaluated as compression ultrasound, studies thus far have indicated encouraging results, and further investigations are warranted [65]. Radionuclide venography of the upper extremity has been described as a reliable noninvasive procedure for early diagnosis of upper limb venous thrombosis associated with indwelling CVC. It is performed by injecting both arms with approximately 5 mCi of technetium pertechnetate followed by a normal saline flush. The dynamic images are acquired on a large field of view camera with a high-energy low-resolution collimator at the rate of two frames per second [66, 67]. (99 m)TC-MAA

radionuclide imaging is a useful method for noninvasive detection of DVT and PTE. Combined radionuclide venography and perfusion lung scan can also be performed in the same setting if Tc99m-MAA is used [68]. The radionuclide venogram appears accurate in the proximal veins and in excluding but not diagnosing distal venous thrombosis. The potential advantages of radionuclide venography versus contrast venography are low-volume and low-flow injection, no need to access a large peripheral vein, no adverse side effects, low radiation exposure (130 mrads), rapidity of execution, and no patient preparation. The disadvantage is the low anatomic detail [66, 67].

In summary, invasive testing for venous thromboembolism can be safely avoided in the majority of patients, using diagnostic strategies combining noninvasive tests.

Color and duplex ultrasound with manual compression (CDDUS) is the most sensitive and specific noninvasive test and is nowadays accepted as the modality of choice for the diagnosis of DVT. CT venous-phase imaging at the time of CT pulmonary angiography and MR venography is comparable with venous ultrasonography in the evaluation of femoropopliteal DVT. The iliac veins and vena cava, vessels poorly shown on ultrasonography but sometimes the source of significant pulmonary emboli, are also depicted by CT and MR venography. MRV can differentiate an acute occlusion from chronic thrombus. Due to its high cost and limited availability, MRV is not used for the routine diagnosis of DVT and should be reserved for the examination of inaccessible veins on ultrasonography and as a complementary test in nondiagnostic ultrasound studies for pregnant women and patients with contraindications to iodinated contrast media injection. Studies on venous scintigraphy have indicated encouraging results but further investigations are warranted. A diagnostic strategy combining clinical score, D-dimer test, and compression US can be used in a systematic way to reliably rule in or exclude venous thromboembolism.

7. CONCLUSIONS

Due to its high specificity, a negative examination may preclude anticoagulant treatment. A strategy combining clinical score and D-dimer test refines the selection of patients. Phlebography is the gold standard method but is invasive and carries risks of contrast media complications and ionizing radiation. CTV following pulmonary CTA and MRV is useful to detect iliocaval thrombosis. MRV can differentiate acute from chronic thrombosis and diagnose central obstructions. RNV has low-anatomic detail.

CDDUS is the modality of choice for the diagnosis of DVT. A diagnostic strategy combining clinical score, D-dimer test, and CDDUS is recommended. Phlebography is reserved for discrepant noninvasive studies.

REFERENCES

[1] E. E. Weinmann and E. W. Salzman, "Deep-vein thrombosis," *New England Journal of Medicine*, vol. 331, no. 24, pp. 1630–1641, 1994.

[2] P. A. Kyrle and S. Eichinger, "Deep vein thrombosis," *Lancet*, vol. 365, no. 9465, pp. 1163–1174, 2005.

[3] F. A. Anderson Jr., H. B. Wheeler, R. J. Goldberg, et al., "A population-based perspective of the hospital incidence and case-fatality rates of deep vein thrombosis and pulmonary embolism. The Worcester DVT Study," *Archives of Internal Medicine*, vol. 151, no. 5, pp. 933–938, 1991.

[4] A. F. Scarsbrook, A. L. Evans, A. R. Owen, and F. V. Gleeson, "Diagnosis of suspected venous thromboembolic disease in pregnancy," *Clinical Radiology*, vol. 61, no. 1, pp. 1–12, 2006.

[5] M. J. Bernstein, H. R. Roberts, S. Adel, et al., "Prevention of venous thrombosis and pulmonary embolism. National Institute of Health Consensus Development," *Journal of the American Medical Association*, vol. 256, no. 6, pp. 744–749, 1986.

[6] M. Levine and A. K. Kakkar, "Catheter-associated thrombosis: thromboprophylaxis or not?" *Journal of Clinical Oncology*, vol. 23, no. 18, pp. 4006–4008, 2005.

[7] M. C. Horattas, D. J. Wright, A. H. Fenton, et al., "Changing concepts of deep venous thrombosis of the upper extremity - report of a series and review of the literature," *Surgery*, vol. 104, no. 3, pp. 561–567, 1988.

[8] P. Prandoni, P. Polistena, E. Bernardi, et al., "Upper-extremity deep vein thrombosis: risk factors, diagnosis, and complications," *Archives of Internal Medicine*, vol. 157, no. 1, pp. 57–62, 1997.

[9] H. J. Baarslag, M. M. W. Koopman, B. A. Hutten, et al., "Long-term follow-up of patients with suspected deep vein thrombosis of the upper extremity: survival, risk factors and post-thrombotic syndrome," *European Journal of Internal Medicine*, vol. 15, no. 8, pp. 503–507, 2004.

[10] P. S. Wells, J. Hirsh, D. R. Anderson, et al., "Accuracy of clinical assessment of deep-vein thrombosis," *Lancet*, vol. 345, no. 8961, pp. 1326–1330, 1995.

[11] DeBakey, E. Michael, and A. M. Gotto Jr., "Invasive diagnostic procedures," in *The New Living Heart*, Adams Media, Holbrook, Mass, USA, 1997.

[12] K. Rabinov and S. Paulin, "Roentgen diagnosis of venous thrombosis in the leg," *Archives of Surgery*, vol. 104, no. 2, pp. 134–144, 1972.

[13] M. Salcuni, P. Fiorentino, A. Pedicelli, and C. Di Stasi, "Diagnostic imaging in deep vein thrombosis of the limbs," *Rays*, vol. 21, no. 3, pp. 328–339, 1996.

[14] R. Hull, J. Hirsh, D. L. Sackett, and G. Stoddart, "Cost effectiveness of clinical diagnosis, venography, and noninvasive testing in patients with symptomatic deep-vein thrombosis," *New England Journal of Medicine*, vol. 304, no. 26, pp. 1561–1567, 1981.

[15] D. P. Flanigan, J. J. Goodreau, S. J. Burnham, J. J. Bergan, and J. S. Yao, "Vascular-laboratory diagnosis of clinically suspected acute deep-vein thrombosis," *Lancet*, vol. 2, no. 8085, pp. 331–334, 1978.

[16] D. S. Katz and M. Hon, "Current DVT imaging," *Techniques in Vascular and Interventional Radiology*, vol. 7, no. 2, pp. 55–62, 2004.

[17] Intersocietal Accreditation Commisision, "ICAVL: essentials and standards for accreditation in noninvasive vascular testing. Part II. Vascular laboratory operations - peripheral venous testing," 200:1-8. http://www.intersocietal.org/intersocietal.htm.

[18] C. Kearon, J. A. Julian, T. E. Newman, and J. S. Ginsberg, "Noninvasive diagnosis of deep venous thrombosis. McMaster Diagnostic Imaging Practice Guidelines Initiative," *Annals of Internal Medicine*, vol. 128, no. 8, pp. 663–677, 1998.

[19] B. G. Birdwell, G. E. Raskob, T. L. Whitsett, et al., "The clinical validity of normal compression ultrasonography in

outpatients suspected of having deep venous thrombosis," *Annals of Internal Medicine*, vol. 128, no. 1, pp. 1–7, 1998.

[20] A. W. A. Lensing, P. Prandoni, D. Brandjes, et al., "Detection of deep-vein thrombosis by real-time B-mode ultrasonography," *New England Journal of Medicine*, vol. 320, no. 6, pp. 342–345, 1989.

[21] P. S. Wells, J. Hirsh, D. R. Anderson, et al., "Comparison of the accuracy of impedance plethysmography and compression ultrasonography in outpatients with clinically suspected deep vein thrombosis. A two centre paired-design prospective trial," *Thrombosis and Haemostasis*, vol. 74, no. 6, pp. 1423–1427, 1995.

[22] R. H. White, J. P. McGahan, M. M. Daschbach, and R. P. Hartling, "Diagnosis of deep-vein thrombosis using duplex ultrasound," *Annals of Internal Medicine*, vol. 111, no. 4, pp. 297–304, 1989.

[23] J. F. Polak, S. S. Culter, and D. H. O'Leary, "Deep veins of the calf: assessment with color Doppler flow imaging," *Radiology*, vol. 171, no. 2, pp. 481–485, 1989.

[24] D. Gaitini, J. K. Kaftori, M. Pery, Y. L. Weich, and A. Markel, "High-resolution real-time ultrasonography in the diagnosis of deep vein thrombosis," *RoFo Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin*, vol. 149, no. 1, pp. 26–30, 1988.

[25] A. M. Asbeutah, A. Z. Riha, J. D. Cameron, and B. P. McGrath, "Five-year outcome study of deep vein thrombosis in the lower limbs," *Journal of Vascular Surgery*, vol. 40, no. 6, pp. 1184–1189, 2004.

[26] D. Voet and M. Afschrift, "Floating thrombi: diagnosis and follow-up by duplex ultrasound," *British Journal of Radiology*, vol. 64, no. 767, pp. 1010–1014, 1991.

[27] M. Langsfeld, B. Matteson, W. Johnson, D. Wascher, J. Goodnough, and E. Weinstein, "Baker's cysts mimicking the symptoms of deep vein thrombosis: diagnosis with venous duplex scanning," *Journal of Vascular Surgery*, vol. 25, no. 4, pp. 658–662, 1997.

[28] D. Gaitini, "Current approaches and controversial issues in the diagnosis of deep vein thrombosis via duplex Doppler ultrasound," *Journal of Clinical Ultrasound*, vol. 34, no. 6, pp. 289–297, 2006.

[29] S. J. Theodorou, D. J. Theodorou, and Y. Kakitsubata, "Sonography and venography of the lower extremities for diagnosing deep vein thrombosis in symptomatic patients," *Clinical Imaging*, vol. 27, no. 3, pp. 180–183, 2003.

[30] G. M. Baxter, "The role of ultrasound in deep venous thrombosis," *Clinical Radiology*, vol. 52, no. 1, pp. 1–3, 1997.

[31] S. Mussurakis, S. Papaioannou, D. Voros, and T. Vrakatselis, "Compression ultrasonography as a reliable imaging monitor in deep venous thrombosis," *Surgery Gynecology and Obstetrics*, vol. 171, no. 3, pp. 233–239, 1990.

[32] A. Cogo, A. W. A. Lensing, M. W. M. Koopman, et al., "Compression ultrasonography for diagnostic management of patients with clinically suspected deep vein thrombosis: prospective cohort study," *British Medical Journal*, vol. 316, no. 7124, pp. 17–20, 1998.

[33] S. Goodacre, F. Sampson, S. Thomas, E. J. R. van Beek, and A. Sutton, "Systematic review and meta-analysis of the diagnostic accuracy of ultrasonography for deep vein thrombosis," *BMC Medical Imaging*, vol. 5, no. 6, 2005.

[34] M. C. Patel, L. W. Berman, H. A. Moss, and S. J. McPherson, "Subclavian and internal jugular veins at Doppler US: abnormal cardiac pulsatility and respiratory phasicity as a predictor of complete central occlusion," *Radiology*, vol. 211, no. 2, pp. 579–583, 1999.

[35] T. M. Kerr, K. S. Lutter, D. M. Moeller, et al., "Upper extremity venous thrombosis diagnosed by duplex scanning," *American Journal of Surgery*, vol. 160, no. 2, pp. 202–206, 1990.

[36] D. Gaitini, J. K. Kaftori, M. Pery, and A. Engel, "High-resolution real-time ultrasonography. Diagnosis and follow-up of jugular and subclavian vein thrombosis," *Journal of Ultrasound in Medicine*, vol. 7, no. 11, pp. 621–627, 1988.

[37] D. G. Longley, D. E. Finlay, and J. G. Letourneau, "Sonography of the upper extremity and jugular veins," *American Journal of Roentgenology*, vol. 160, no. 5, pp. 957–962, 1993.

[38] R. L. Falk and D. F. Smith, "Thrombosis of upper extremity thoracic inlet veins: diagnosis with duplex Doppler sonography," *American Journal of Roentgenology*, vol. 149, no. 4, pp. 677–682, 1987.

[39] C. Koksoy, A. Kuzu, J. Kutlay, I. Erden, H. Ozcan, and K. Ergin, "The diagnostic value of colour Doppler ultrasound in central venous catheter related thrombosis," *Clinical Radiology*, vol. 50, no. 10, pp. 687–689, 1995.

[40] H.-J. Baarslag, E. J. R. van Beek, M. M. W. Koopman, and J. A. Reekers, "Prospective study of color duplex ultrasonography compared with contrast venography in patients suspected of having deep venous thrombosis of the upper extremities," *Annals of Internal Medicine*, vol. 136, no. 12, pp. 865–872, 2002.

[41] A. Luciani, O. Clement, P. Halimi, et al., "Catheter-related upper extremity deep venous thrombosis in cancer patients: a prospective study based on Doppler US," *Radiology*, vol. 220, no. 3, pp. 655–660, 2001.

[42] S. Mustafa, P. D. Stein, K. C. Patel, T. R. Otten, R. Holmes, and A. Silbergleit, "Upper extremity deep venous thrombosis," *Chest*, vol. 123, no. 6, pp. 1953–1956, 2003.

[43] E. E. Chin, P. T. Zimmerman, and E. G. Grant, "Sonographic evaluation of upper extremity deep venous thrombosis," *Journal of Ultrasound in Medicine*, vol. 24, no. 6, pp. 829–839, 2005.

[44] P. Prandoni and E. Bernardi, "Upper extremity deep vein thrombosis," *Current Opinion in Pulmonary Medicine*, vol. 5, no. 4, pp. 222–226, 1999.

[45] T. L. Nack and L. Needleman, "Comparison of duplex ultrasound and contrast venography for evaluation of upper extremity venous disease," *Journal of Vascular Technology*, vol. 16, pp. 69–73, 1992.

[46] R. D. Bona, "Central line thrombosis in patients with cancer," *Current Opinion in Pulmonary Medicine*, vol. 9, no. 5, pp. 362–366, 2003.

[47] P. S. Wells, "Advances in the diagnosis of venous thromboembolism," *Journal of Thrombosis and Thrombolysis*, vol. 21, no. 1, pp. 31–40, 2006.

[48] M. Mantoni, "Ultrasound of limb veins," *European Radiology*, vol. 11, no. 9, pp. 1557–1562, 2001.

[49] J. J. Michiels, A. Gadisseur, M. van der Planken, et al., "Screening for deep vein thrombosis and pulmonary embolism in outpatients with suspected DVT or PE by the sequential use of clinical score: a sensitive quantitative D-dimer test and noninvasive diagnostic tools," *Seminars in Vascular Medicine*, vol. 5, no. 4, pp. 351–364, 2005.

[50] P. S. Wells, D. R. Anderson, M. Rodger, et al., "Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and D-dimer," *Annals of Internal Medicine*, vol. 135, no. 2, pp. 98–107, 2001.

[51] E. E. Coche, X. L. Hamoir, F. D. Hammer, P. Hainaut, and P. P. Goffette, "Using dual-detector helical CT angiography to detect deep venous thrombosis in patients with suspicion of pulmonary embolism: diagnostic value and additional findings," *American Journal of Roentgenology*, vol. 176, no. 4, pp. 1035–1039, 2001.

[52] P. A. Loud, D. S. Katz, D. A. Bruce, D. L. Klippenstein, and Z. D. Grossman, "Deep venous thrombosis with suspected pulmonary embolism: detection with combined CT venography and pulmonary angiography," *Radiology*, vol. 219, no. 2, pp. 498–502, 2001.

[53] K. Garg, J. L. Kemp, D. Wojcik, et al., "Thromboembolic disease: comparison of combined CT pulmonary angiography and venography with bilateral leg sonography in 70 patients," *American Journal of Roentgenology*, vol. 175, no. 4, pp. 997–1001, 2000.

[54] K. M. Duwe, M. Shiau, N. E. Budorick, J. H. M. Austin, and Y. M. Berkmen, "Evaluation of the lower extremity veins in patients with suspected pulmonary embolism: a retrospective comparison of helical CT venography and sonography," *American Journal of Roentgenology*, vol. 175, no. 6, pp. 1525–1531, 2000.

[55] P. D. Stein, S. E. Fowler, L. R. Goodman, et al., "Multidetector computed tomography for acute pulmonary embolism," *New England Journal of Medicine*, vol. 354, no. 22, pp. 2317–2327, 2006.

[56] M. D. Cham, D. F. Yankelevitz, D. Shaham, et al., "Deep venous thrombosis: detection by using indirect CT venography," *Radiology*, vol. 216, no. 3, pp. 744–751, 2000.

[57] K.-E. Lim, W.-C. Hsu, Y.-Y. Hsu, P.-H. Chu, and C.-J. Ng, "Deep venous thrombosis: comparison of indirect multidetector CT venography and sonography of lower extremities in 26 patients," *Clinical Imaging*, vol. 28, no. 6, pp. 439–444, 2004.

[58] S. Patel and E. A. Kazerooni, "Helical CT for the evaluation of acute pulmonary embolism," *American Journal of Roentgenology*, vol. 185, no. 1, pp. 135–149, 2005.

[59] J. Kelly, B. J. Hunt, and A. Moody, "Magnetic resonance direct thrombus imaging: a novel technique for imaging venous thromboemboli," *Thrombosis and Haemostasis*, vol. 89, no. 5, pp. 773–782, 2003.

[60] P. Haage, T. Krings, and T. Schmitz-Rode, "Nontraumatic vascular emergencies: imaging and intervention in acute occlusion," *European Radiology*, vol. 12, no. 11, pp. 2627–2643, 2002.

[61] J.-P. Laissy, A. Cinqualbre, A. Loshkajian, et al., "Assessment of deep venous thrombosis in the lower limbs and pelvis: MR venography versus duplex Doppler sonography," *American Journal of Roentgenology*, vol. 167, no. 4, pp. 971–975, 1996.

[62] F. C. Sampson, S. W. Goodacre, S. M. Thomas, and E. J. R. van Beek, "The accuracy of MRI in diagnosis of suspected deep vein thrombosis: systematic review and meta-analysis," to appear in *European Radiology*.

[63] K. R. Shankar, L. J. Abernethy, K. S. V. Das, et al., "Magnetic resonance venography in assessing venous patency after multiple venous catheters," *Journal of Pediatric Surgery*, vol. 37, no. 2, pp. 175–179, 2002.

[64] S. J. McRae and J. S. Ginsberg, "The diagnostic evaluation of deep vein thrombosis," *The American Heart Hospital Journal*, vol. 2, no. 4, pp. 205–210, 2004.

[65] M. B. Gotway, K. J. Edinburgh, V. A. Feldstein, J. D. Lehman, G. P. Reddy, and W. R. Webb, "Imaging evaluation of suspected pulmonary embolism," *Current Problems in Diagnostic Radiology*, vol. 28, no. 5, pp. 133–184, 1999.

[66] R. D. Dhekne, W. H. Moore, C. Peters, S. J. Dudrick, and S. E. Long, "Upper extremity radionuclide venography in the diagnosis and management of venous disease associated with intravenous catheters," *Angiology*, vol. 39, no. 10, pp. 907–914, 1988.

[67] D. A. Podoloff and E. E. Kim, "Evaluation of sensitivity and specificity of upper extremity radionuclide venography in cancer patients with indwelling central venous catheters," *Clinical Nuclear Medicine*, vol. 17, no. 6, pp. 457–462, 1992.

[68] J. Mangkharak, S. Chiewvit, W. Chaiyasoot, et al., "Radionuclide venography in the diagnosis of deep vein thrombosis of the lower extremities: a comparison to contrast venography," *Journal of the Medical Association of Thailand*, vol. 81, no. 6, pp. 432–441, 1998.