30th Annual International IEEE EMBS Conference
Vancouver, British Columbia, Canada, August 20-24, 2008

Adaptive Mesh Optimization and Nonrigid Motion Recovery Based Image Registration for Wide-Field-of-View Ultrasound Imaging

Chaowei Tan[†], Bo Wang[†], Paul Liu[‡], and Dong Liu[†]

Abstract— **Wide field of view (WFOV) imaging mode obtains an ultrasound image over an area much larger than the real time window normally available. As the probe is moved over the region of interest, new image frames are combined with prior frames to form a panorama image. Image registration techniques are used to recover the probe motion, eliminating the need for a position sensor. Speckle patterns, which are inherent in ultrasound imaging, change, or become decorrelated, as the scan plane moves, so we pre-smooth the image to reduce the effects of speckle in registration, as well as reducing effects from thermal noise. Because we wish to track the movement of features such as structural boundaries, we use an adaptive mesh over the entire smoothed image to home in on areas with feature. Motion estimation using blocks centered at the individual mesh nodes generates a field of motion vectors. After angular correction of motion vectors, we model the overall movement between frames as a nonrigid deformation. The polygon filling algorithm for precise, persistence-based spatial compounding constructs the final speckle reduced WFOV image.**

Index Terms— **AOS gaussian smoothing, adaptive mesh, structure tensor, nonrigid motion recovery, linear transformation, compounding accuracy improvement, polygon filling**

I. Introduction

Ultrasound imaging is well established in medicine as a diagnostic tool that provides real-time anatomical and physiological information in two-dimensional slice format. However, the field of view of individual images produced by conventional 2D display format is limited to the narrow scan region of typical transducers. As the probe moves along the body, the previous displayed image is replaced with the new scan, and tissue information in regions outside of the current field of view are discarded. This presents potential diagnostic difficulties in interpreting a wide region of interest. The wide field of view (WFOV) imaging mode [1] addresses this problem by compounding successive image frames as the probe is moved, forming a panorama image of a region much larger than the conventional scan mode. A modified searching algorithm for image registration, and an angle-based correction method to reduce errors in the results of searching are introduced in [2]. The position sensor may also provide exact position information, however, the electromagnetic interference and additional cost limit its use in the hospital, so the development of image-based WFOV methods is necessary.

Each slice of the whole series in [1] [2] is divided into small blocks, whose motion between frames are individually calculated. This reduces the problem of overall motion between two frames into that of two point sets representing the locations of each block. References [1] [2] are limited to a rigid transformation of the image blocks and hence unable to account for more general tissue motion such as shear or non-linear deformations. In this paper, we introduce an adaptive mesh method over images pre-processed by a fast smoothing operation to optimize motion estimation. Furthermore, we incorporate a more flexible motion model by solving the least square optimal linear transformation between two weighted point sets, where inaccurate block displacements are given less weight. To compound the registered images and reduce speckle, we use persistence method with the polygon filling algorithm for improving accuracy.

The proposed image registration technique consists of a pre-smoothing described in Section II, feature tracking using an adaptive mesh and block motion estimation discussed in Section III, angle correction in Section IV, recovering overall image motion by linear transformation in Section V, and compounding in Section VI. Results and conclusions follow in Sections VII, VIII respectively.

II. Pre-smoothing

Because of the presence of speckle, it is quite difficult to directly use common registration methods. To improve the accuracy of registration, we should reduce the influence of speckle. Here, we use a fast and flexible method to do preprocessing. It is Semi-implicit Additive Operator Splitting (AOS) based Gaussion smoothing [3]. Because Gaussian Smoothing is a kind of isotropic diffusion [4], it is,

$$\frac{u^{k+1}-u^k}{\tau} = A(u^k)u^{k+1} \tag{1}$$

where u^k and u^{k+1} are the image gray value vectors of the current and next time steps respectively, obtained by stacking each column vector of the K by L image on top of another into one long column vector of length KL, τ is the time step size (TSS), it controls the extent of blurring. $A(u^k)$ is a tridiagonal matrix, with $A(u^k) = [a_{ij}(u^k)]$ and

$$a_{ij}(u^k) \triangleq \begin{cases} 1/h^2 & j \in N(i) \\ -\sum_{n\in N(i)} 1/h^2 & j = i \\ 0 & \text{else} \end{cases} \tag{2}$$

where $N(i)$ is the set of two neighbors of pixel i (boundary pixels have only one neighbor). The 2D AOS scheme Gaussion smoothing equation is

$$u^{k+1} = \frac{1}{2}\sum_{l=1}^{2}[I - \tau A_l(u^k)]^{-1}u^k \tag{3}$$

[†]Chaowei Tan, Bo Wang, Dong Liu are with the College of Computer Science, Sichuan University, Chengdu, P.R. China. wangbo.1222@gmail.com, chaoweitan@gmail.com
[‡]Paul Liu is with Saset Healthcare Inc., Chengdu, P.R. China.

Above linear system is tridiagonal and diagonally dominant, we could use the $O(KL)$ Thomas Algorithm [3] to solve it where KL is the length of u^k. Compared with spatial domain filtering such as Gaussian and box filters whose time consumption will increase as we increase the extent of blurring, the first advantage of AOS scheme based Gaussian filtering is that its time consumption is constant for a large range blurring. Although the AOS scheme based Gaussian filtering is unconditionally stable numerically, if we use a too large TSS to process a image, there will be many artifacts on the blurred image. So we said that it is constant for a large range, and the range is large enough for us to do preprocessing. Another advantage of AOS scheme based Gaussian filtering is that it is very flexible to control the extent of blurring, we do not need to consider the boundary condition problem as we use spatial domain filters.

III. Adaptive Mesh and Block Motion Estimation

In order to have more accurate matching information, we propose a new rectangular adaptive mesh over the entire smoothed image to home in on areas with feature such as structural boundaries, which will lead better motion estimation. Traditionally, some literatures [6] calculated additional feature energy map and used 2D searching method(such as gradient decent) to position mesh nodes. However, the time consumption of this method is unacceptable for fast processing. Hence we develop a faster structure matrix based method. Structure matrix had been used in speckle reduction [5], it can be expressed as follows:

$$J(I(i,j)) = \begin{pmatrix} I_x^2(i,j) & I_x(i,j)I_y(i,j) \\ I_x(i,j)I_y(i,j) & I_y^2(i,j) \end{pmatrix} \quad (4)$$

where $I(i,j)$ is the intensity of pixel (i,j) and I_x and I_y are the x and y partial derivatives respectively. However, directly discretizing (4) is too susceptible to noise, here, we use a modified structure tensor to detect local coherence and direction of maximum variations. Defining $\hat{I}_{x+}(i,j) = \sum_{k=0}^{n-1} I(i+k+1,j)$, $\hat{I}_{x-}(i,j) = \sum_{k=0}^{n-1} I(i-k,j)$, $\hat{I}_{y+}(i,j) = \sum_{k=0}^{n-1} I(i,j+k+1)$, $\hat{I}_{y-}(i,j) = \sum_{k=0}^{n-1} I(i,j-k)$, we discretize (4) as

$$\hat{J}(I) = \begin{pmatrix} \left(\hat{I}_{x+} - \hat{I}_{x-}\right)^2 & \left(\hat{I}_{x+} - \hat{I}_{x-}\right)\left(\hat{I}_{y+} - \hat{I}_{y-}\right) \\ \left(\hat{I}_{x+} - \hat{I}_{x-}\right)\left(\hat{I}_{y+} - \hat{I}_{y-}\right) & \left(\hat{I}_{y+} - \hat{I}_{y-}\right)^2 \end{pmatrix} \quad (5)$$

where we omitted the i and j indices to remove notational clutter. We diagnolize (5) as

$$\hat{J}(I) = (v_1 \;\; v_2) \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \end{pmatrix} \quad (6)$$

Here, the eigenvectors and the eigenvalues correspond to the directions of maximum and minimum variations and the strength of these variations, respectively. From starting search point (s_x, s_y), we find the maximum variation direction $(d_x, d_y) = \psi v_1^T, \psi = -1$ or 1. If $|d_x| \geq |d_y|$, then the main component of search direction(MCSD) is x component, else MCSD is y. If we assume that the MCSD is x (the same for y), then the unit increment Δx in MCSD is 1 for the each step of search, the unit increment Δy of another component is $|d_y|/|d_x|$. Along search direction (d_x, d_y), the structure matrix of the k-th sampling point $\left(s_x + \frac{\psi k d_x}{|d_x|}, s_y + \frac{\psi k d_y}{|d_x|}\right)$ is calculated as (5), and the feature value v_f of this sampling point is equal to $|\mu_1 - \mu_2|$, which is derived from (6). Then the elastic spring model [6] is applied to determine the nodal positions. The criteria of model is the potential energy ε_m at node m, which is a combination for the internal spring restoration energy ε_s and the feature energy ε_f, it is

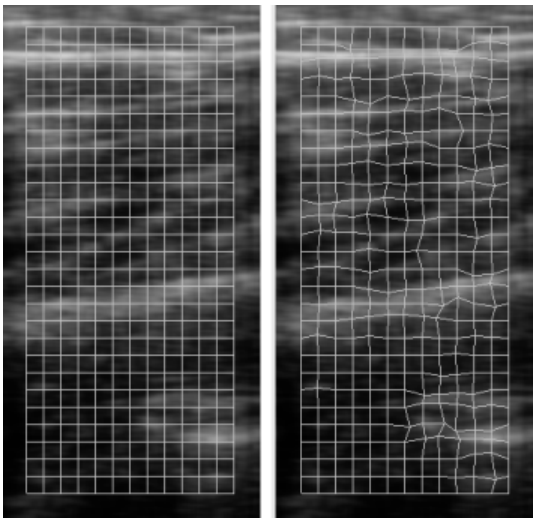

Fig. 1. Initial Mesh and Adapted Mesh

$$\varepsilon_m = \varepsilon_s + \varepsilon_f = \frac{1}{2}k\varepsilon^2 - \frac{1}{2}v_f\|D\|^2 \quad (7)$$

where $\varepsilon^2 = \sum_{i=1}^{4} d_i^2$ at each node is the total sum of the squared length change of springs connected to the node, k is the spring constant, and $\|D\|^2$ is the squared magnitude of the nodal displacement vector. $d_i = l_i - L$ is the length change of rest length L and stretched length l_i.

Therefore, nodal positions are determined where $\varepsilon_{pre} \geq \varepsilon_{cur}$ and $\varepsilon_{cur} \leq \varepsilon_{post}$, ε_{cur} is potential energy at current position. This adaptive mesh method is able to adapt to the actual ultrasound images, and its advantage is low calculation cost. The search is 1D compared with 2D gradient descent method, and more importantly, it does not need to compute an additional feature energy map for the whole image but only the 2×2 structure matrix at certain points. As shown in Fig. 1, we obtain the adapted mesh comparing with the initial mesh on pre-smoothed images by AOS method. After the adaptive mesh is positioned at areas of feature, motion estimation using blocks centered at the individual mesh nodes generates a field of more accurate motion vectors. The details of the Improved Minimum Sum of Absolute Differences (IMSAD) motion estimation are discussed in [2].

IV. Angle Based Correction

Error of motion estimation, local tissue motion, speckle decorrelation, out-of-plane motion and scanning backwards pose a significant challenge in motion estimation. Therefore

the angle based correction method is used to regulate bad motion vectors. The mean vector is the criteria during the correction processing for two reasons. First, from the viewpoint of statistics, the mean vector can express the actual movement between sequence of images excluding the impact of bad motion vectors in general cases. Second, using thresholds in the matching process of IMSAD makes the distribution of valid blocks irregular, so median filtering for block vectors is not suitable here. Moreover, we must follow an assumption that the corrected motion of the previous frame should dominate the motion of the current frame. The angle based correction method has three steps developed in [2]. Here, we expand this method to adapt the overall motion estimation in the next section. The correlation coefficient α, which represents a function of the individual motion vector's deviation angle compared with corrected mean vector, is treated as a weight for future processing.

After the correction processing, the bad motion vectors of blocks are regulated with the relatively correct mean vector, so that the blocks motion estimation is more reliable.

V. Linear Transformation for Recovery of Overall Motion

From the above two sections, the relative positions of the one-to-one corresponding blocks in the previous and the current frame (images 1 and 2) are obtained. We can treat these relative positions as two point patterns and calculate the transformation parameters to represent the overall transformation parameter between images 1 and 2. All that needs to be done then is to transform image 2, place it in the WFOV image buffer, and update the pixels. Previous work in [2] used the conclusion of a well-known problem that is to find the similarity transformation parameters of (R: rotation, t: translation, and c: scaling) in computer vision applications, giving the minimum mean squared error (MSE) of two point patterns, x_i and y_i, $i = 1,2,\ldots,n$ in an n-dimensional space. The MSE is expressed as:

$$e^2(R,t,c) = \frac{1}{n}\sum_{i=1}^{n} \|y_i - (cRx_i + t)\|^2 \tag{8}$$

Using [7], we can get these optimal rigid transformation parameters in the two patterns with a priori known correspondences. However the rigid transformation can not be a good description of the actual situation, for some deformation of tissue and speckle happens in the adjacent images. In contrast to finding optimal rigid transformations, many methods have been tried in computer graphics to simulate deformable objects, and here the method mentioned in [8] is referred to compute optimal linear transformations to give a better description of the deformation. In this approach, the point patterns matching problem with a priori one-to-one relationship can be stated as follows: Given two sets of points x_i and y_i, find the rotation matrix R and the translation vectors t and t_0 which minimize $\sum_i w_i(R(x_i - t_0) + t - y_i)^2$, where the w_i are weights of individual points. In the situation here, the weights are calculated in previous section. The optimal translation vectors turn out to be the center of mass of the initial shape and the center of mass of the actual shape, i.e.

$$t_0 = x_{cm} = \frac{\sum_i w_i x_i}{\sum_i w_i}, \;\; t = y_{cm} = \frac{\sum_i w_i y_i}{\sum_i w_i} \tag{9}$$

In order to find the optimal rotation R, this method defines the relative locations $q_i = x_i - x_{cm}$ and $p_i = y_i - y_{cm}$ of points with respect to their center of weight and relaxes the problem of finding the optimal rotation matrix R to finding the optimal linear transformation A. And the term to be minimized is $\sum_i (Aq_i - p_i)^2$. Setting the derivatives with respect to all coefficients of to zero yields the optimal transformation

$$A = (\sum_i w_i p_i q_i^T)(\sum_i w_i q_i q_i^T)^{-1} \tag{10}$$

which contains the rotational part R.

As we mentioned before, fully rigid motion can not give an accurate description of the actual deformable motion, so we use the linear transformation matrix computed in (10) to extend the range of motion from rigid to linear. Finally, the goal positions can be computed as

$$g_i = A(x_i - x_{cm}) + y_{cm} \tag{11}$$

Expanding (11), we get $g_i = Ax_i + T$, and define $T = y_{cm} - Ax_{cm}$, which represents the translation vector (t_x, t_y). Therefore, the linear transformation method is used to recover the motion between successive images in a sequence, i.e. from images 1 to 2, 2 to 3 and so on. Briefly, if the motion from images 1 to 2 is given by the parameters A, T, and that between images 2 to 3 is given by A', T'. The transformation matrix representing the overall motion from image 1 to 3 is given by:

$$M_{13} = \begin{pmatrix} A'_{11} & A'_{12} & t'_x \\ A'_{21} & A'_{22} & t'_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & t_x \\ A_{21} & A_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix} \tag{12}$$

The overall motion estimation can obtain the accurate estimation of probe motion and the image registration processing is finished.

VI. Polygon Filling for Sequence of Images Compounding

After the motion recovery is done, the WFOV image is combined from a sequence of images with the transformation information. However, the point-to-point compounding method causes float-to-int position mapping errors, which will degrade the quality of compounded image. In order to compound the adjacent images accurately, we use the algorithm of polygon filling [9] to eliminate the float-to-int mapping errors. The main process of this method has the following steps:

1) Finding the filling range of the sampled image in the WFOV buffer. This step places the four corners of image to the corresponding position in the WFOV buffer first, and this is a float-to-int mapping. Then the edge of the polygon is calculated. Because we use a linear deformation model, the filling range is an arbitrary quadrilateral.

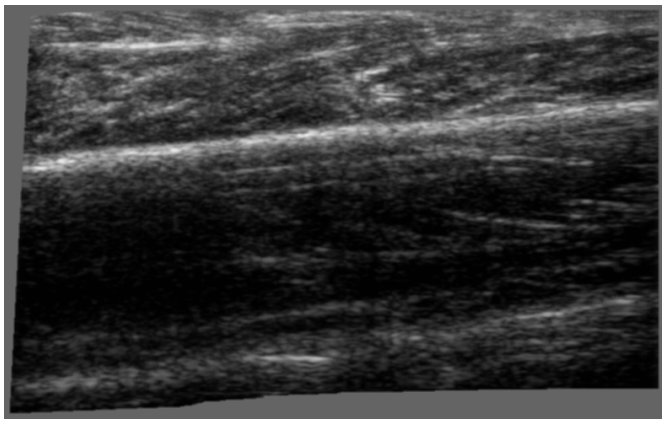

Fig. 2. Compounded femur WFOV image from a 5 MHz linear probe.

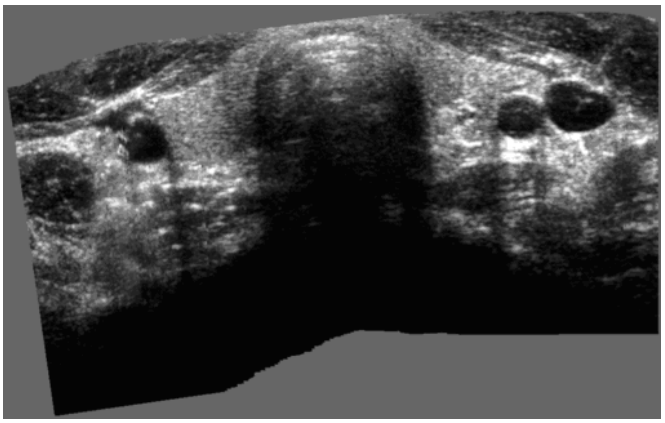

Fig. 3. Compounded neck WFOV image from a 7.5 MHz linear probe

2) Getting the integer coordinate P_i of each pixel in the filling range using the polygon filling algorithm. All edge pixels of the polygon are discarded, for these pixels' positions are not integers in the filling range.
3) Calculating the inverse matrix M' of the current linear transformation matrix M. Then we get the starting coordinate P_i' of P_i, where $P_i' = M'P_i$. So we can calculate the pixel value $\hat{v}_i$ using bilinear interpolation with the original coordinate P_i' in sampled image.

When the process of polygon filling is finished, we get accurate pixel values for each position of integer coordinate in the compounding range, then the persistence method and some other image quality optimization strategies in [2] are used for other processing in the WFOV compounding.

VII. Computer Simulations and Results

To verify our proposed improved WFOV algorithms, we have tested lots of sequences of successive images using *in vivo* data. We used Saset iMago color ultrasound scanner for data acquisition which is the system we designed in our lab. The standard test set contain 300-400 images. The proposed WFOV method is interactive, i.e., process activates when the user begins moving the probe and presents the WFOV image continuously with the moving of the probe. As shown in Fig. 2, we provide the final WFOV image scaned from a 5 MHz linear probe, and this is a femur scanning. We can see that the WFOV image contain the complete information of this tissue, and the image is denoised. In Fig. 3, a neck scanning WFOV image is provided, and this image is obtained from a 7.5 MHz linear probe. We can clearly see the two sides of the thyroid besides the carotids and jugulars in the denoised neck image.

VIII. Conclusion and Future Work

In this paper we present an improved method for wide field of view imaging. Additive Operator Splitting based Gaussion smoothing method is used for pre-processing, and the adaptive mesh method improves the IMSAD searching, so block-based motion estimation is improved. The weighted linear transformation makes the recovery of overall motion be able to adapt both rigid and deformable conditions. Applying the polygon filling algorithm improves the quality of compounded WFOV image. In future work, more robust motion estimation methods and more accurate registration models [10] need to be considered for the situation of speckle decorrelation and tissue deformation. Moreover, some high performance computational resources, such as GPU [11], should be implemented for faster processing in diagnosis.

IX. Acknowledgments

The authors thank Alex Huang and Xiaohui Zuo of Saset Healthcare Inc. for obtaining the ultrasound images.

References

[1] L. Weng, A. P. Tirumalai, C. M. Lowery, L. F. Nock, D. E. Gustafson, P. L. Von Behren, and J. H. Kim, "US extended-Field-of-View imaging technology," *Radiology*, vol. 203, pp. 877-880, Jun. 1997.
[2] C. Tan and D. C. Liu, "Image Registration Based Wide-Field-of-View Method in Ultrasound Imaging", *Second International Conference on Bioinformatics and Biomedical Engineering*, Shanghai, China, 2008.
[3] J. Weickert, B. M. T. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Trans. Image Processing*, vol. 7, no. 3, pp. 398-410, Mar. 1998.
[4] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, Jul. 1990.
[5] K. Z. Abd-Elmoniem, A.-B. M.Youssef, and Y. M. Kadah, "Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion," *IEEE Trans. Biomedical Engineering*, vol. 49, no. 9, pp. 997-1014, Sep. 2002.
[6] F. Yeung, S. F. Levinson, F. Dongshan, and K. J. Parker, "Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh," *IEEE Trans. Medical Imaging*, vol. 17, no. 6, pp. 945-956, Dec. 1998.
[7] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376-380, Apr. 1991.
[8] M. Müller, B. Heidelberger, M. Teschner, and M. Gross, "Meshless deformations based on shape matching," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 471-478, Jul. 2005.
[9] D. F. Rogers, *Procedural Elements for Computer Graphics, 2nd edition*. Chapter 2. New York: McGraw-Hill, 1998.
[10] M. Holden, "A Review of Geometric Transformations for Nonrigid Body Registration," *IEEE Trans. Medical Imaging*, vol. 27, no. 1, pp. 111-128, Jan. 2008
[11] S. Mazaré, R. Pacalet, and J.-L. Dugelay, "Using GPU for fast block-matching," *14th European Signal Processing Conference*, Sep. 2006.