Rapid Registration of Multimodal Images Using a Reduced Number of Voxels

Xishi Huang[1,3], Nicholas A Hill[2,3], Jing Ren[3,4], and Terry M. Peters[1,2,3]
[1]Biomedical Engineering Graduate Program, [2]Department of Medical Biophysics,
The University of Western Ontario;
[3]Imaging Research Laboratories, Robarts Research Institute,
P.O. Box 5015, 100 Perth Drive, London, ON, Canada N6A 5K8
[4]Canadian Surgical Technologies & Advanced Robotics (CSTAR), London, Canada

ABSTRACT

Rapid registration of multimodal cardiac images can improve image-guided cardiac surgeries and cardiac disease diagnosis. While mutual information (MI) is arguably the most suitable registration technique, this method is too slow to converge for real time cardiac image registration; moreover, correct registration may not coincide with a global or even local maximum of MI. These limitations become quite evident when registering three-dimensional (3D) ultrasound (US) images and dynamic 3D magnetic resonance (MR) images of the beating heart. To overcome these issues, we present a registration method that uses a reduced number of voxels, while retaining adequate registration accuracy. Prior to registration we preprocess the images such that only the most representative anatomical features are depicted. By selecting samples from preprocessed images, our method dramatically speeds up the registration process, as well as ensuring correct registration. We validated this registration method for registering dynamic US and MR images of the beating heart of a volunteer. Experimental results on *in vivo* cardiac images demonstrate significant improvements in registration speed without compromising registration accuracy. A second validation study was performed registering US and computed tomography (CT) images of a rib cage phantom. Two similarity metrics, MI and normalized cross-correlation (NCC) were used to register the image sets. Experimental results on the rib cage phantom indicate that our method can achieve adequate registration accuracy within 10% of the computation time of conventional registration methods. We believe this method has the potential to facilitate intra-operative image fusion for minimally invasive cardio-thoracic surgical navigation.

Keywords: multimodal image registration, image-guided cardiac procedures, mutual information, normalized cross-correlation, cardiac disease diagnosis

1. INTRODUCTION

Cardiac illness is the leading cause of death by disease in developed countries and many people undergo heart surgery every year. Conventional cardiac procedures require that a sternotomy be performed to provide the surgeon access to the heart. Conventional approaches also require arresting the patient's heart and using a cardiac bypass machine to maintain the patient's oxygenation and blood circulation during the surgery. These two facets of conventional cardiac surgeries are extremely traumatic and while they are necessary to access the site they are not part of the actual cardiac therapy. Despite the invasive nature and associated risk of complications, most cardiac surgeries are still performed with an open chest. Conversely, minimally invasive cardiac therapies significantly decrease patient trauma and morbidity; lower the risk of infection and other complications; decrease patient recovery time and the length of hospital stay, and ultimately reduce health care costs.

Although minimally invasive cardiac surgery greatly benefits both the patient and the health-care system, it poses a difficult challenge to surgeons and engineers. Unlike open chest surgery, which provides surgeons with a direct view of and access to the heart, minimally invasive cardiac therapies offer limited surgical access and restricted surgical vision because the surgeon must rely on endoscopes or other medical imaging devices to provide a view of the heart. Minimally invasive cardiac procedures call for a real time, high quality image guidance method.

Interventional MR or CT offers high quality image guidance, however, a number of implementation issues (including restricted surgical access; increased expense and complexity of procedure; in the case of CT, both the patient and surgeon are exposed to harmful ionizing radiation; compatibility issues of the surgical instruments and equipment

Medical Imaging 2006: Visualization, Image-Guided Procedures, and Display,
edited by Kevin R. Cleary, Robert L. Galloway, Jr., Proc. of SPIE Vol. 6141, 614116, (2006)
1605-7422/06/$15 · doi: 10.1117/12.653593

with the magnetic field of the MR unit), make using these imaging techniques during minimally invasive therapies, difficult. On the other hand, US imaging is safe, comparatively low in cost, simple to use, and avoids radiation and compatibility problems. However, the cardiac images produced are of poor quality and difficult to interpret unless they can be placed into their corresponding anatomical context. By augmenting cardiac images produced by US with MR/CT images the surgeon would have an image guidance tool with the flexibility of US and the quality anatomical information provided by MR/CT. In order to guide cardiac therapies these US and MR/CT images have to be integrated in real time.

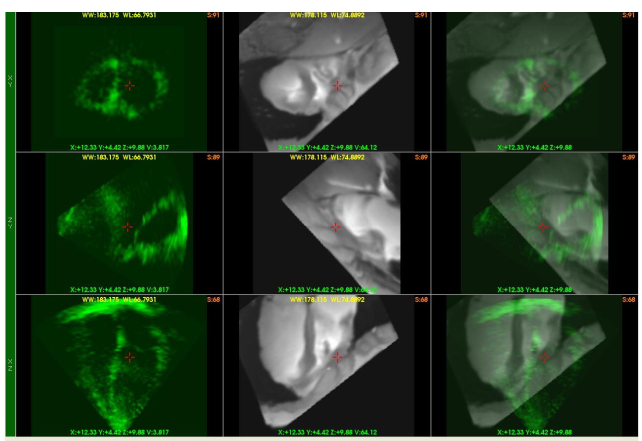
Mutual information (MI) is widely used in multimodal image registration [1]. However, correct registration may not coincide with a global or local maximum of MI [2], [3]. This difficulty is quite apparent when MI is used to register 3D US images to dynamic 3D MR images of the beating heart. Standard MI methods tend to converge to an incorrect registration result mainly due to artifacts and low signal-to-noise ratio (SNR) in the US images. Moreover, most multimodal image registration methods converge too slowly to be used in real time applications. Registration speed depends on the complexity of the images being registered and on the number of voxels being processed.

In order to resolve the problem of incorrect registration and to reduce the registration time, we present a method which registers preprocessed US and MR/CT images using a subset of the original voxels. Our approach requires preprocessing the US and MR/CT images to generate masked images, which depict only the most representative features and filter out speckle, noise and other unnecessary or redundant information prior to selecting samples for registration. The proposed method can be used to register intra-operative US images with previously acquired MR/CT images for surgical guidance during minimally invasive cardiac surgeries or to facilitate easier and more accurate cardiac disease diagnosis by combining relevant diagnostic information from two imaging modalities.

This paper addresses 3D US/CT-MR image registration in two distinct but related situations. We first discuss the registration of 3D US to 3D MR images of the heart, and then 3D US to CT images of the rib cage. This latter technique forms part of the protocol to initially register a pre-operative dataset to the patient.

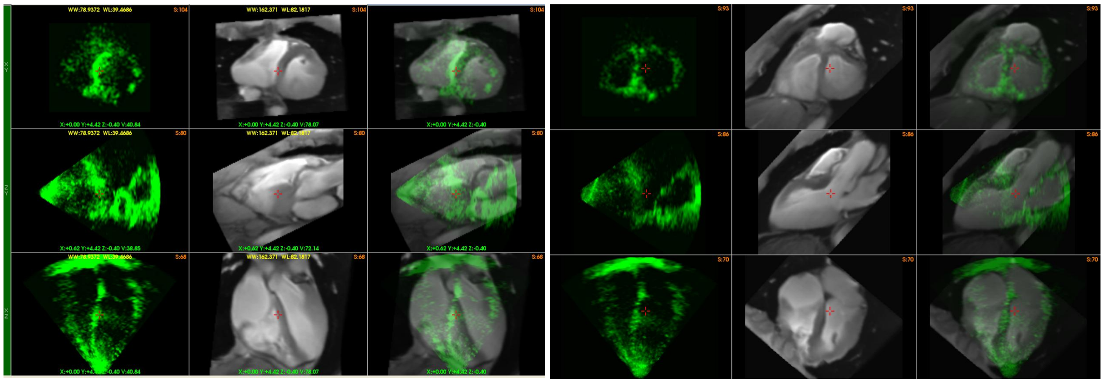
2. BACKGROUND

The discrepancy between correct registration and the global or local maximum of the MI metric [2], [3] is quite evident when MI is used to register 3D US images to dynamic 3D MR images of the beating heart. The MI metric considers both the heart and background information in its calculation and often results in a mismatch (Figure 1). Masking out the background does not necessarily eliminate this problem, as a significant mismatch may still remain due to the presence of artifacts in the US images and the low SNR of US (Figure 2a). The correct registration (Figure 2b), which was confirmed by visual inspection by a cardiologist, does not correspond to any local, let alone a global, minimum (Figure 3). To help ensure correct image registration, we restrict the field of view (FOV) in the US images to depict only the heart and then threshold these images such that only the most representative features of the heart are shown.



(a) (b) (c)

Fig. 1. Multimodal image registration using the MI similarity metric. (a) three orthogonal slices of the original US images (b) three orthogonal slices of the high resolution MR images (c) the image registration results using MI. Note the three panels in each column display the 3D object in three orthogonal views. The top panel depicts the axial view, the middle the sagittal view, and the bottom panel the coronal view. This convention is followed for all similar images in this paper.



(a) (b)

Fig. 2. (a) a large mismatch is still apparent after masking out the background of the US, (b) the correct registration is achieved after thresholding and masking the US images to depict only the most representative features of the heart.

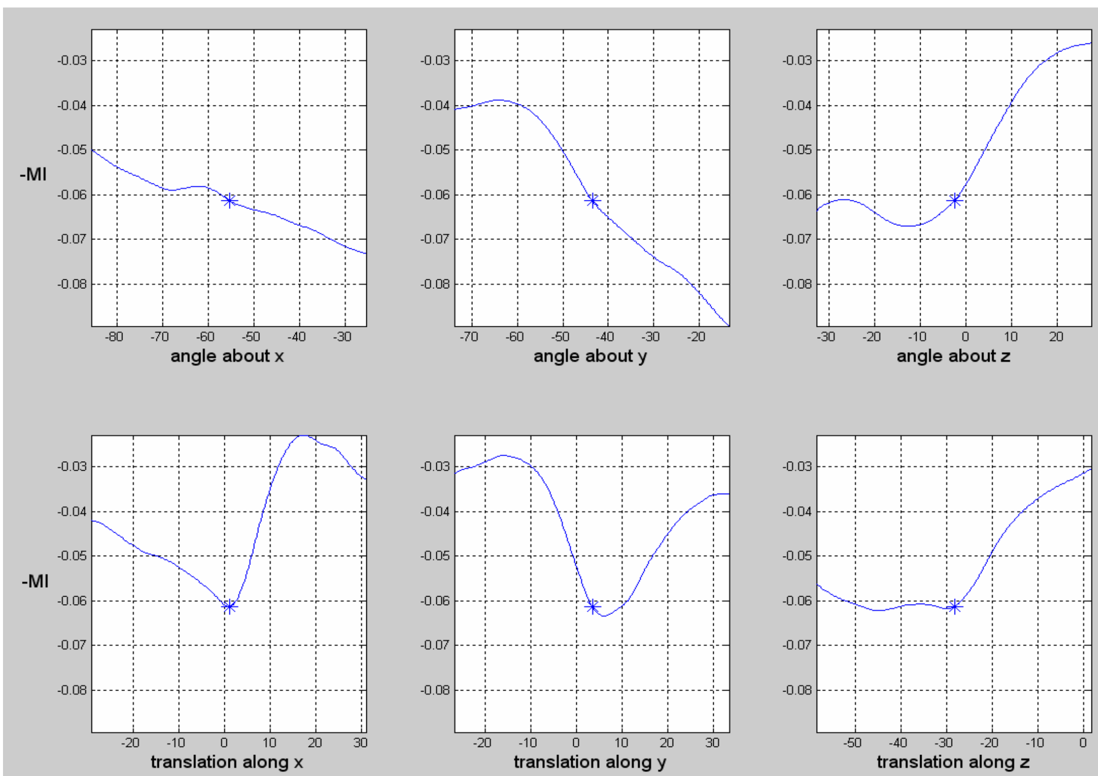


Fig. 3. The MI similarity metric for each rigid-body transformation parameter. The top three plots illustrate the MI curves for the rotational parameters and the bottom three plots illustrate the MI curves for the translational parameters. Correct registration does not correspond to a global or local minimum using MI. In each plot the * indicates the transformation parameter value for correct registration. The * should be located at a minimum, but this is often not the case when registering US and MR images of the beating heart.

The speed of convergence for intensity-based image registration is also a concern. The registration speed of intensity-based methods depends greatly on the number of voxels being registered. The number of voxels registered in MI methods is often selected as 10 - 20% of the total voxels of the images. In this study, the complete 3D US image consists of approximately 4.8 million voxels. The region of interest, the heart, contains roughly 1.2 million voxels. If, for example, we select a sample size of 100 000 voxels, it takes 37.27 seconds to complete an acceptable registration result on typical hardware (Intel P4 2.8GHz).

3. METHODS

3.1 Registration Procedure

In order to overcome the issues of incorrect registration and slow registration speed, we present a registration method that only considers the most representative features of both images. We reduce the number of voxels employed for image registration by selecting them from the most pronounced features in the original images. This registration method is not only more robust than an approach that considers all the voxels, but also greatly accelerates registration process.

The registration procedure consists of three discrete steps:

- **Step 1. Extraction of the most representative features of images**. We preprocess original images to generate masked images containing only the most representative features captured in both US and MR/CT volumes. In order to generate these masked images, speckle and less salient features were removed using an application-specific preprocessing routine (discussed in the following examples).
- **Step 2. Registration sample selection**. Sample voxels are randomly selected from the most pronounced features on the masked images and are used in the registration procedure.
- **Step 3. Registration using reduced samples.** Perform rigid-body registration of US samples to preprocessed MR/CT images using the MI and/or NCC similarity metrics, using a gradient descent method to optimize registration parameters.

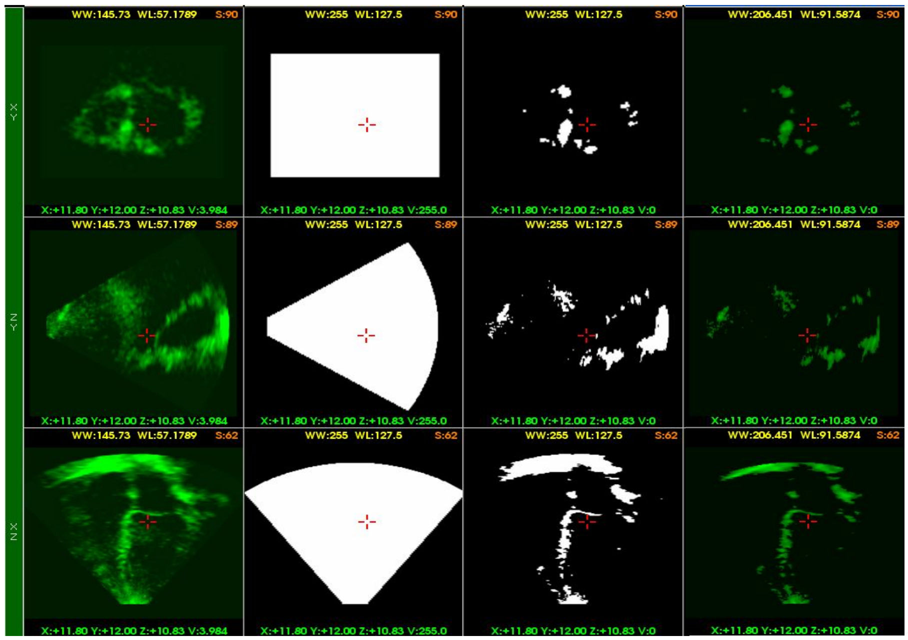
The following examples describe the utilization of this approach to register US-MR images of a volunteer's beating heart and US-CT images of a rib cage phantom.

3.2 Example 1. Registration of 3D US-MR images of the beating heart

In this example, 20 dynamic 3D MR images of one cardiac cycle (with a resolution of 256 x 256 x 75 and a voxel size of 1.48 x 1.48 x 1.50 mm^3) were acquired on a 1.5T GE CVi scanner (GE Medical systems, Milwaukee), and 14 3D US images (with a resolution of 160 x 144 x 208 and a voxel size of 1.24 x 1.26 x 0.80 mm^3) obtained from the same volunteer, were acquired on a Philips SONOS 7500 US machine [4]. Breath-holding was employed during both image acquisitions in an attempt to ensure that the complex motion of the heart was solely a function of myocardium contraction.

Since the US and MR data sets were acquired with different acquisition rates, the US and MR frames were not temporally synchronized. The MR images were therefore temporally interpolated to generate MR volumes corresponding to the phases in the cardiac cycle at which the US volumes were acquired, using the method described in Huang *et al.*, 2005 [4]. The US volume, and the appropriate, temporally interpolated MR volume were viewed in custom software (OCCIViewer) and the two image sets were manually translated and rotated until the volumes were approximately aligned. Our experience has indicated that such preliminary manual registration can be achieved to within ±10 mm and ±5 degrees, well within the capture range of the MI-based registration algorithm (as illustrated in Figure 7).

To overcome the problem of incorrect registration using MI, we extracted the most representative anatomical features and reduced any artifacts and noise in the US images. First, thresholding was applied to the US images such that only the most apparent anatomical features (i.e. cardiac chamber walls) were depicted. Then the FOV in the US images was restricted to highlight the cardiac structures. For convenience, we refer to the thresholded US image with the restricted FOV as *masked US images* (Figure 4). A small sample set (1 000 voxels) was randomly selected from these masked US images, and registered to the MR images using MI.



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

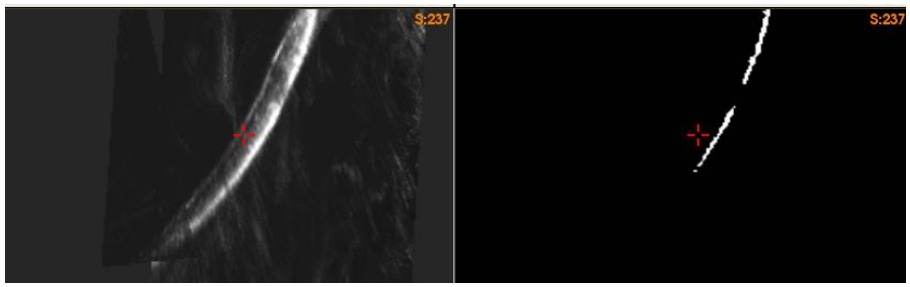
Fig. 4. Generating the masked US images. (a) original US images, (b) the FOV is restricted, (c) thresholding is applied, and (d) the final thresholded US images with the restricted FOV.

3.3 Example 2. Registration of US-CT images of a rib cage phantom

Registration of 3D US and dynamic 3D MR images of the beating heart is a challenging task. The registration result is often dependent on how close to the optimal solution the registration starts. The rib cage provides a frame of reference for thoracic organs, (heart and lungs), therefore registering multimodal images of the rib cage results in a suitable initial position/orientation for beating heart image registration and eliminates the need for an initial manual registration. In this section, we use our registration method to register US and CT images of a rib cage phantom.

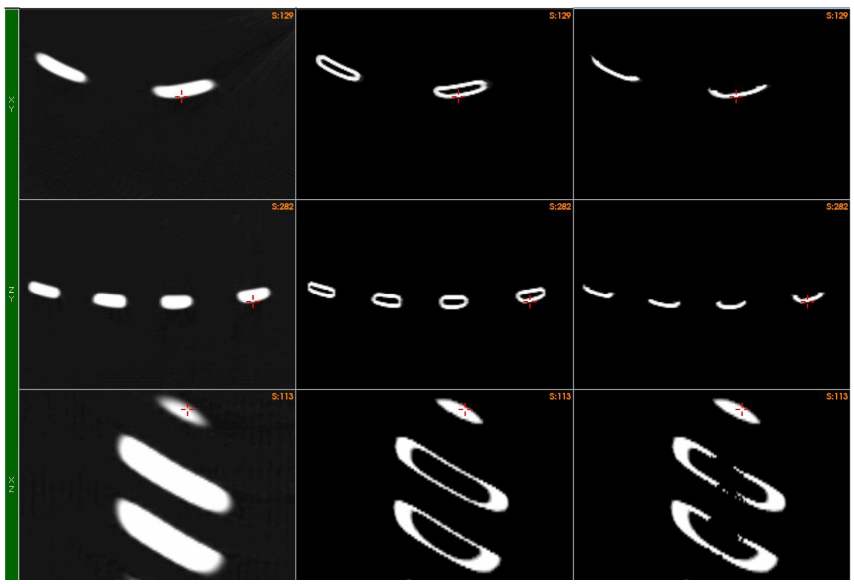
A high-resolution CT data set of the rib cage phantom was acquired using a helical CT imaging technique with an image resolution of 512 x 512 x 321 and a voxel size of 0.78 x 0.78 x 1.25 mm^3. Two-dimensional (2D) US images were acquired using an Aloka SSD 1700 US machine. The US probe was tracked with a Polaris optical tracking system (OTS) and the resulting 2D images were reconstructed into 3D US volume using a freehand reconstruction method [5]. Five 3D US volumes, representing different regions of the rib cage phantom, were acquired and assembled into one composite US 3D volume [6].

The US images were noisy due to speckle. Thresholding the US images reduced the noise and only one boundary corresponding to the top surface of the ribs was depicted (Figure 5b). Applying a masking procedure to the CT images such that only one side of the ribs is shown, (as is the case for the preprocessed US images), improves and speeds up the registration process. A morphological erosion operation with the radius of one voxel was used to extract similar anatomical boundary features in the CT images. A gradient magnitude image was then generated using a directional gradient magnitude filter. If the gradient component along the US scanning direction was positive, we set the output to the gradient magnitude, otherwise we set it to 0. Finally, a thresholding operation was used to isolate the intensity values of the rib cage from the background of the image (Figure 6). After preprocessing the US and CT images, the MI and NCC metrics were maximized using an iterative registration procedure. While the NCC metric is generally only applicable to images from the same modality that have similar appearances, by processing the US and CT images in the manner described we obtain processed US and CT images with similar appearance, and may therefore use NCC for this application.



(a) (b)

Fig. 5. US image preprocessing, (a) original US image, (b) preprocessed US image



(a) (b) (c)

Fig. 6. CT image preprocessing, (a) original CT image, (b) preprocessed CT image with whole boundary, (c) preprocessed CT image with single boundary

4. EXPERIMENTAL RESULTS

4.1 Registration of 3D US and MR images of the beating heart

The MI method was employed to register preprocessed 3D US and 3D MR images of the beating heart with 1 000 sampled voxels. Figures 2b and 7, illustrate that our method not only ensures the correct registration, but also reduces the registration time from 37.27 seconds to 0.89 seconds per frame, compared to our original trials using 100 000 samples. The pre-processing task takes 0.35 seconds.

4.2 Registration of 3D US and CT images of the rib cage phantom

3D US and 3D CT images were registered using both NCC and MI algorithms. The registration results are summarized in Table 1 and Figure 8. In Table 1, the target registration error (TRE) is defined as the root mean square distance between the centers of matched fiducial markers as described in [6]. In the NCC algorithm, the number of voxels used for registration can be reduced from 70 000 to 200, while still maintaining adequate registration accuracy. Figure 9 demonstrates that the NCC similarity metric curves are quite smooth and have similar trends in the neighborhood of correct registration for different sample sizes, justifying the use of a small number of voxels selected from regions representing salient features. Registration time using the NCC and MI methods is approximately linear with the number of samples.

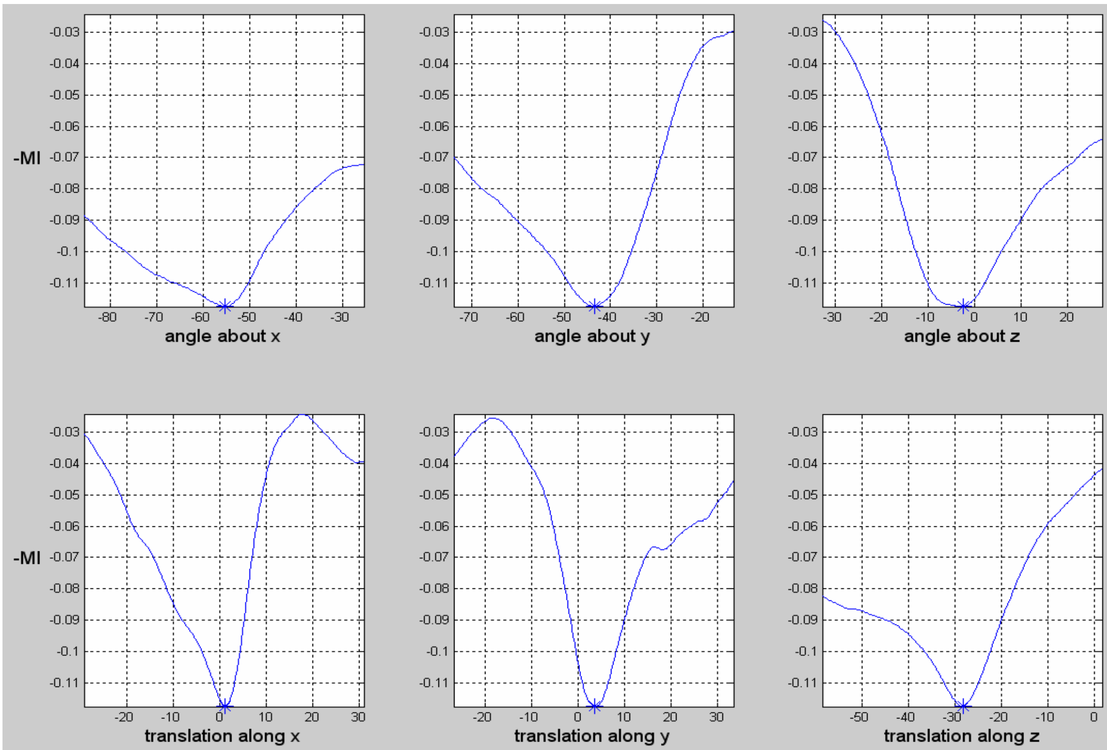
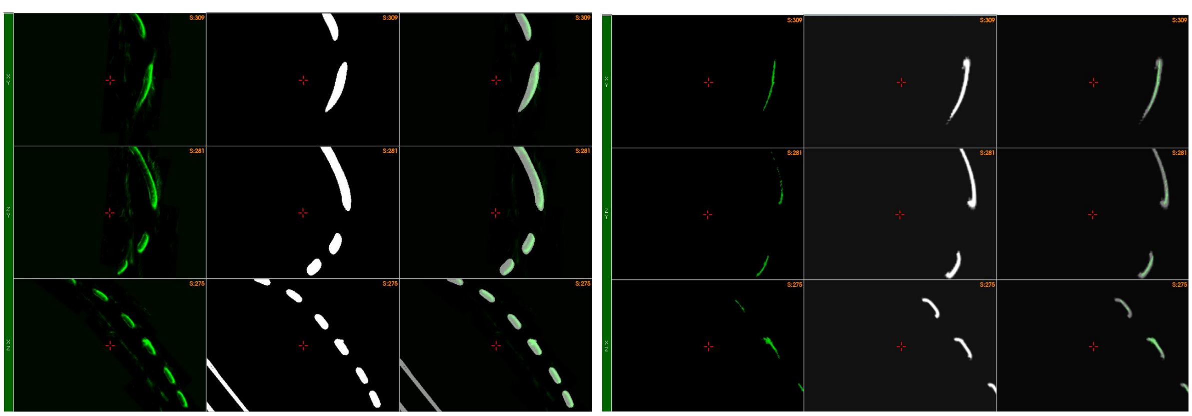
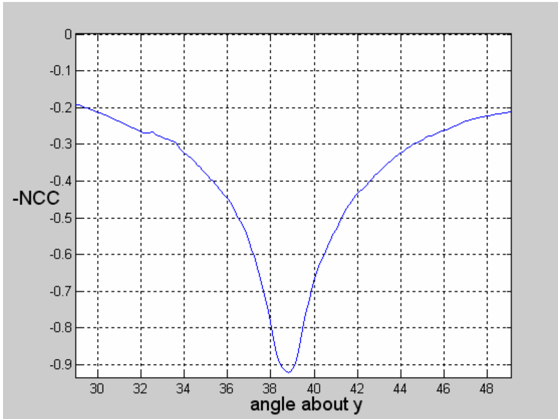


Fig. 7. Correct registration corresponds to a local minimum using the proposed method. In each plot, * denotes the transformation parameter value for correct registration, which is located at a minimum.



(a) (b) (c) (d) (e) (f)

Fig. 8. Registration of preprocessed US and CT images using 200 samples. Each column shows three orthogonal slices of volumes. (a) original US, (b) original CT, (c) overlay of original US & CT, (d) preprocessed US, (e) preprocessed CT, (f) overlay of preprocessed US & CT.



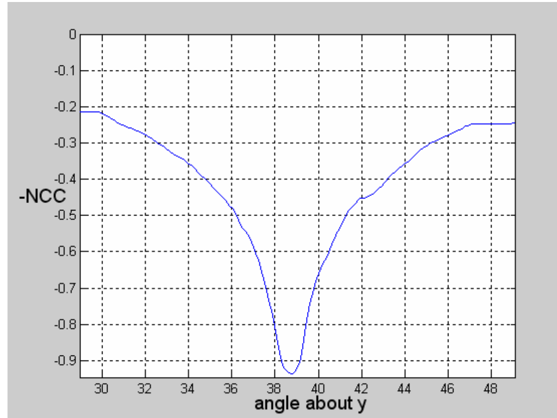


Fig. 9. NCC with single boundary (a) 1 000 samples (b) 200 samples. Different sample sizes display similar trends.

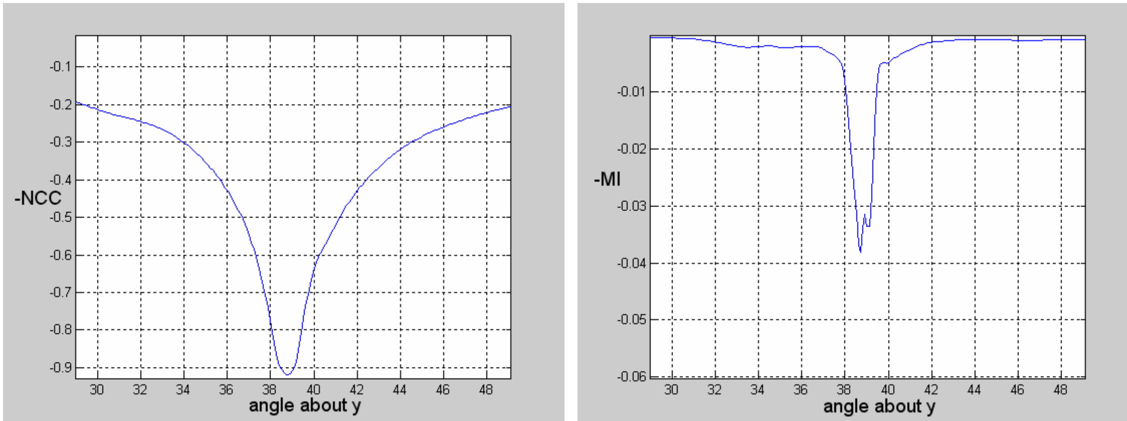
	NCC		MI	
#Samples	**Time (s)**	**TRE (mm)**	**Time (s)**	**TRE (mm)**
200	0.73	1.45	4.41	1.62
500	1.58	1.33	6.49	1.58
1000	2.65	1.34	9.13	1.48
5000	15.04	1.25	34.09	1.46
10000	29.10	1.22	57.56	1.46
50000	158.57	1.22	224.92	1.45
70000	216.36	1.21	301.97	1.45

Table 1. Registration accuracy and time

5. DISCUSSION

Our method for rapidly registering multimodal images requires generating masked images, which display only the most pronounced features captured in both modalities, and then randomly selecting a small number of samples from these masked images for registration. We have validated this method using images from rib cage phantom experiments and images from a volunteer beating heart. Validation results suggest that this technique can effectively eliminate mis-registration problems of standard MI methods, as well as significantly reduce registration times.

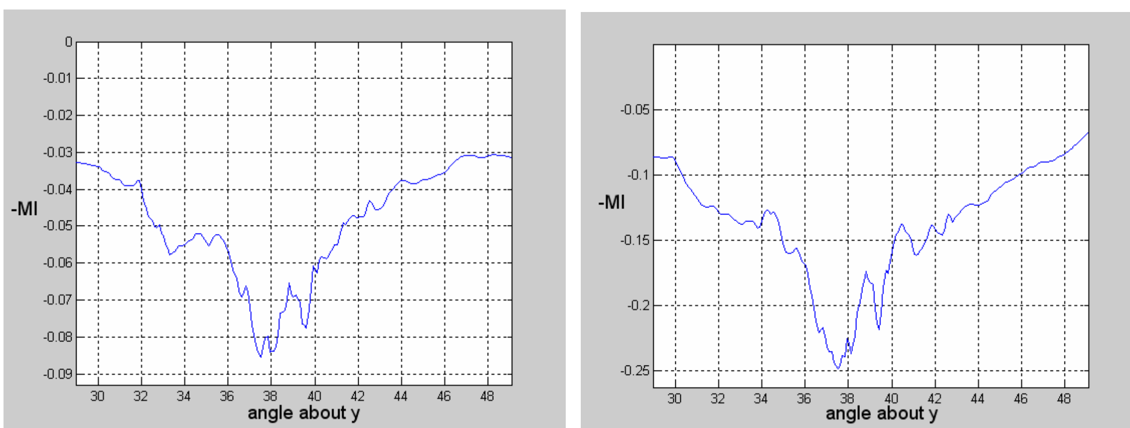
From our registration experiments of the rib cage phantom, we see that the NCC metric produces a better registration result than the MI metric in terms of accuracy, smoothness of the metric function and less local minima (Figures 9, 10 and 11). In the case of the single boundary example of the preprocessed CT image, NCC has no local minima near the global minimum (the correct registration value) (Figures 9 and 10a), while MI has many such minima close to correct registration (Figures 10b and 11). If only 200 voxels are used for image registration, the NCC function is still very smooth (Figure 9b), but the MI function becomes noisy making it difficult to achieve a correct registration (Figure 11b). Although the NCC method performs better than the MI method, NCC can only be employed to register images generated from the same modality or from images of different modalities that can be processed into images having similar characteristics. We note that the TRE results obtained for the US-CT (rib) protocol (which translates into a TRE at the center of the heart of approximately 2.5 mm) demonstrate that this is a potential means of generating an initial registration starting point for an 3D US-3D MRI (which needs to be within about ±10 mm of the final registration result) and therefore would avoid the manual initial registration step. Note however that since we do not have access to the CT data from the same subject for whom the US-MR registration was performed, we were not able to use this result directly in the example described in Section 3.2.



(a) NCC

(b) MI

Fig. 10. (a) NCC with single boundary. There are no local minima close to the global minimum (correct registration) using the NCC metric. (b) MI with single boundary. The MI curve has local minima in the vicinity of correct registration.



(a)

(b)

Fig. 11. The negative MI curve for the single boundary case for (a) 1 000 samples and (b) 200 samples. There are more local minima in the case of small samples. As the number of samples is reduced, the MI curve becomes noisier and the number of local minima increases.

6. CONCLUSIONS

By extracting the most representative features from both imaging modalities, the number of voxels considered for registration is reduced and the registration process can be completed faster without sacrificing registration accuracy. This was demonstrated for two different, yet related scenarios, the registration of US to MR images of the beating heart and the registration of US to CT images of a rib cage phantom. Furthermore, we demonstrated that the NCC metric is more robust than the MI metric in cases where the source and target images have similar characteristics.

REFERENCES

[1] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever, "Mutual-information-based registration of Medical Images: A Survey," IEEE Trans. on Medical Imaging, 22(8):986-1004, 2003.

[2] F. Maes, D. Vandermeulen, and P. Suetens, "Medical image registration using mutual information," Proceedings of IEEE, 91(10):1699-1722, 2003

[3] D.L.G. Hill, P.G. Batchelor, M. Holden, and D.J. Hawkes, "Medical image registration", Phys. Med. Biol., 46:R1-R45, 2001.

[4] X. Huang, N.A. Hill, J. Ren, G. Guiraudon, D. Boughner, and T.M. Peters, "Dynamic 3D ultrasound and MRI image registration of the beating heart," J. Duncan and G. Gerig (Eds.): MICCAI 2005, LNCS 3750:171-178, 2005.

[5] Gobbi DG and Peters TM, "Interactive intra-operative 3D ultrasound reconstruction and visualization," Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002, LNCS 2489:156–163, 2002.

[6] X. Huang, N.A. Hill, and T.M. Peters, "Ultrasound-based technique for intra-thoracic surgical guidance," Proceedings of SPIE, 5744:822-830, 2005.