Multi-scale Adaptive Mask 3D Rigid Registration of Ultrasound and CT Images

Zhijun Zhang, Jerome Schmid, Man Kwan Soo,
Bailly Yan, Chung Kwong Yeung
Surgery Department
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
alexzhang@surgery.cuhk.edu.hk

Abstract

3D Rigid intra-operative image registration is an important technique to provide guidance for pre-operative information from different image modalities. Due to the artefacts that cause correspondence ambiguity, accurate registration of US with other image modalities such as computed tomography (CT) is still a challenging problem. We propose a method which considers the registration problem of US and CT images as a multi-scale regional information saliency and local similarity selection process. We design our method as a multi-stage approach in which global and local rigid registrations alternate in each stage. During the local registration, the US image regions with high feature saliency and similarity with CT image will be selected and joined as region masks for the next global registration, other parts will not be considered in order to remove the correspondence ambiguity. The performances of our method are compared with a typical intensity based registration method on phantom and real patient images.

1 Introduction

3D Ultrasound (US) imaging is widely used in image guided surgery due to its non-ionizing effect, low cost and real-time properties. However, US images are spoiled by speckle noise and artefacts [3]. The artefacts regions do not correspond to meaningful anatomical structure information. This information incompleteness phenomenon often makes the comprehension of anatomical structure very difficult. Registration between the US images with a complementary modality, such as computed tomography (CT) or magnetic resonance imaging (MRI), appears as a promising solution to improve US image understanding. By correctly aligning the CT images with the US images, all of the pre-operative information extracted from the CT (e.g. segmentation of organs, major vessels and pathologies detections) can be augmented on the US image [2, 7]. This technique will bring great convenience and improve the efficiency and safety during the surgical practice. Accurate registration between the CT and US images is the critical problem in ultrasound augmented reality.

The multi-modal image registration can be categorized into model based, feature based and voxel based registration methods. Among these methods, voxel based methods use directly the intensity information to match the source and target images, it is not necessary to segment the images or extract models from images. They are very suitable for computer implementation. There are several intensity based multi-modal image registration methods [9, 13, 11, 15] and they work well on registrations between CT and MR images. Maes *et al* [9] proposed mutual information (MI) based registration method, Studholme *et al* [13] proposed an improvement of overlap independent method by using normalized mutual information (NMI). Roche *et al* [11] proposed a method on correlation ratio (CR) and Woods *et al* [15] proposed a method on partition intensity ratio (PIU). In these methods, image intensities are considered as random variables with identical independent distributions. The image similarity functions will measure the dependency or correlation between the two random variables. It can be represented as function of the joint probability density function (PDF) which is evaluated from corresponding intensity pairs. The correspondence between artefacts regions of US image and CT image will add the unreliable intensity pairs into the joint PDF and bias it. The influence of US artefacts in registration has been reported in [5]: the result will not always correspond to the global optimum or even to a local optimum.

A solution is to detect the useful information for registration. By extracting the regions where the images have better correspondence information, the performance of the intensity based methods can be greatly improved. Huang *et al* [5] used a threshold to extract the useful voxels in the US. However, this simple operation does not work properly on US image with complex anatomical structures such as abdomen and brain. Roche *et al* [11] proposed a robust estimation of bivariate function together with a correlation ratio method to suppress the correspondence outliers between MRI and US. The results are usually dependent on the parameters tuning and the Powell's optimization process is time consuming. Penney *et al* [10] extracted a vessel probability density map from the US images and used it to register with the MRI images. This method needs a learning process by using a large amount of US images together with a manually determined threshold for MRI images. Leroy *et al* [8] and Wein *et al* [14] used noise models to detect the artefacts regions. In real application, it is difficult to find a general model for artefacts to achieve good performances. Local features can provide unique and reliable information for registering the images with less trustable information. Stewart *et al* [12] proposed a method to register the retinal images by using local features. The registration starts from the most accurate local feature matching and then propagates with more global feature matching. Salient regions have been used as features for registration recently because of its higher robustness. Huang *et al* [1] has used multiple salient regions for 2D image registrations.

We propose a multi-scale 3D adaptive mask MI based CT to US images registration method. It is an iterative method with several stages. In each stage a global registration and some local block matchings are carried out. Regions with low saliency or low similarity with CT image will be excluded from registration, they usually coincide with the artefact image regions. The global and local registration process will alternate until the whole registration converges.

The rest of the paper is organized as follows. Our novel intensity based registration method is explained in the Section 2. The experiments and datasets are introduced in Section 3. Results and discussion are shown in Section 4. Finally, conclusions are given in Section 5.

2 Methods

2.1 Matching Ambiguity in Intensity Based Registration Methods

Due to the artefacts in the US images, the correspondence between the US and CT images may have ambiguity. MI based method prefers the transform which brings the joint histogram to be more clustered.This problem will happen to all of the information theory based registration method including NMI method. Fig.1 shows a negative MI metric of a 3D phantom US and CT images with different masks of the US image. The images in left column are axial slices of a CT image and the overlaid US and CT images which are already registered. We can see in the right part of the US image, a region with shadow artefacts exists. Images in middle column shows two different mask images of US volume used for registration, the upper one is the fan shape mask and the lower one is the vessel area mask. The figures in right column show negative MI metrics with different US image masks as functions of translational parameters errors in *x* and *y* directions. We can see the the similarity metric with fan mask does not correspond with a global optimum while that with a vessel area mask shows a global minimum in the correct transform parameters.
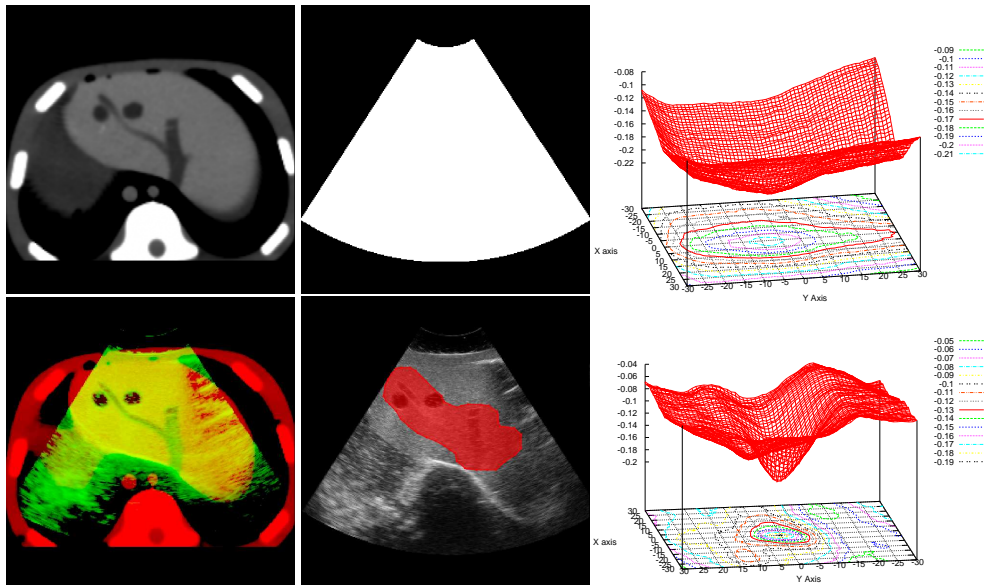

Figure 1: Left column: one axial slice of the CT image and the color overlaid US and CT corresponding axial slice; Middle column: the fan shape mask and the vessel mask of US image; Right column: the negative MI metrics plotted with fan shape mask and the vessel mask against the registration parameters error around the ground truth parameters.

2.2 Combination of Global and Local Registrations

We design our method as a combined global and local rigid registration method. This method combines the advantage of robustness for global registration and accuracy for local registration. We use MI for global and local similarity metric in our implementation
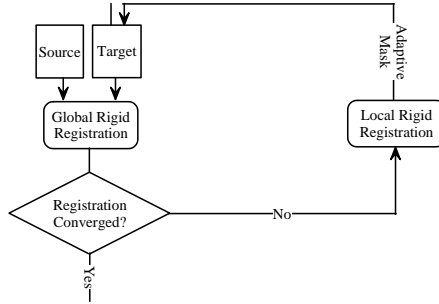
Figure 2: The combination of global and local rigid registrations.

but any multi-modal intensity based similarity metric can be embedded into our registration framework. We begin our registration using global MI registration method with whole US fan mask. We then divide our image into uniform blocks and analyze the local region saliency and their similarities with CT image. Blocks with high local saliency and high similarity with CT will be selected for next global registration. The selected local region will be joined and used as the mask for global registration. The whole registration process will alternate between global and local registrations until the it converges. Fig.2 shows the collaboration of the global and local registration methods.

2.3   3D Region Saliency Calculation

The region with useful information and the artefacts can be extracted by using the concept of region saliency [6]. Saliency is the measurement of an unpredictability of local attributes over a scale. It is proposed for general images and it is also suitable for medical images. The local attributes can be intensity values or colors. The scale can be considered as a sphere with a certain radius. Larger saliency measurement means bigger unpredictability and probability density function magnitude change over scale, so intuitively it means a bigger dissimilarity over scale. Higher saliency regions will be less possible to be the area of artefacts since in these areas when changing the scale, the intensity is usually a reordering and the difference of PDF magnitude is very small.

The region saliency detection consists of three steps as described in [6]. The local saliency will be the product of local entropy $H_D(s, \mathbf{x})$ and local probability density difference $W_D(s, \mathbf{x})$ at the optimum scale:

$$S_D(s_p, \mathbf{x}_0) = H_D(s_p, \mathbf{x}_0) W_D(s_p, \mathbf{x}_0). \tag{1}$$

In continuous case, the entropy will be described by:

$$H_D(s, x_0) = - \int p(i, s, \mathbf{x}_0) log\, p(i, s, \mathbf{x}_0) di, \tag{2}$$

with i the intensity index, $p(i, s, \mathbf{x}_0)$ the Parzen window estimation of the probability density function around $\mathbf{x}_0$ with a scale s. We denote the Parzen window local PDF estimation as in [4]:

$$p(i, \sigma_s, \mathbf{x}_0) = \frac{1}{|\Omega|} \int_{\mathbf{x} \in \Omega} g_\sigma(i - I(\mathbf{x})) g_{\sigma_s}(\mathbf{x} - \mathbf{x}_0) d\mathbf{x}, \tag{3}$$

with g_σ a Gaussian kernel function and $g_{\sigma_s}(\mathbf{x} - \mathbf{x}_0)$ a Gaussian function weighting each of the intensity Gaussian value according to the distance of the points to the region center $\mathbf{x}_0$, Ω is the local image area and $|\Omega|$ is the volume size of these region. Change of σ_s will changes the scale of the local joint entropy, so we directly use σ_s as the scale, that is $\sigma_s = s$.

The optimal local scale s_p is determined by:

$$s_p = \{s : \frac{\partial H_D(s, \mathbf{x}_0)}{\partial s} = 0, \frac{\partial^2 H_D(s, \mathbf{x}_0)}{\partial s^2} < 0\}, \tag{4}$$

and the continuous partial derivative of $H_D(s, \mathbf{x}_0)$ with respect to scale will be:

$$\frac{\partial H_D}{\partial s} = -\int\int (log\, p(i, s, \mathbf{x}_0) + 1) g_\sigma(i - I(\mathbf{x})) g_s(\mathbf{x} - \mathbf{x}_0)(\frac{\mathbf{x} - \mathbf{x}_0}{s})^2 di\, d\mathbf{x}. \tag{5}$$

Then the optimal local scale can be obtained by solving nonlinear equation (4) by substituting $\frac{\partial H_D}{\partial s}$ by (5). The inter-scale saliency measure, $W_D(s, \mathbf{x}_0)$ is defined by:

$$W_D(s, \mathbf{x}_0) = s \int |\frac{\partial}{\partial s} p(i, s, \mathbf{x}_0)| di \tag{6}$$

When the optimal scale of a point is obtained, we can substitute s_p into equation (6) and (1) to obtain the saliency measurement around point $\mathbf{x}_0$.

Instead of evaluate the saliency at each of the voxel, we divide the 3D US image into uniform blocks, then we evaluate the saliency at the block center. We assume the saliency measurement is continuous and smooth, the saliency evaluated at the block center represents the block saliency. This will decrease the computing time distinctly. At each of block center, we obtain the optimum scale for local entropy and the region saliency. We reorder the blocks by their center points saliency from high to low. We choose a portion (we use 70%) of the blocks with high saliency value for local block matching.

2.4 Polar Coordinate Image Processing

Because the beam ray characteristic of US imaging, the image is actually sampled along the ultrasonic beams. The artefacts are also distributed along the beam rays. When we evaluate the entropy of the US image or joint entropy of US and CT images, it is more reasonable to sample the US image along the beam rays. The beam rays form a coordinate space of Polar coordinate space. So instead of measuring the entropy and MI by sampling uniformly in Cartesian coordinate space, we sample the US images in Polar coordinate space. A typical 3D Cartesian and Polar coordinate space for US is shown in fig.3 in the top row. The coordinate transform information can be easily obtained from the geometrical information of the 3D US volume. An example of the same number of sampled points for MI evaluations in the Cartesian and Polar coordinates are shown in the bottom row of fig.3 from left to right. The US image is used as source image and the sample points are randomly chosen in it. We can see the uniform distributions in Cartesian and Polar coordinate space respectively in these two methods. We can see obviously that the sample points are uniformly distributed along the ultrasonic beam rays instead of in the 3D rectangular space.
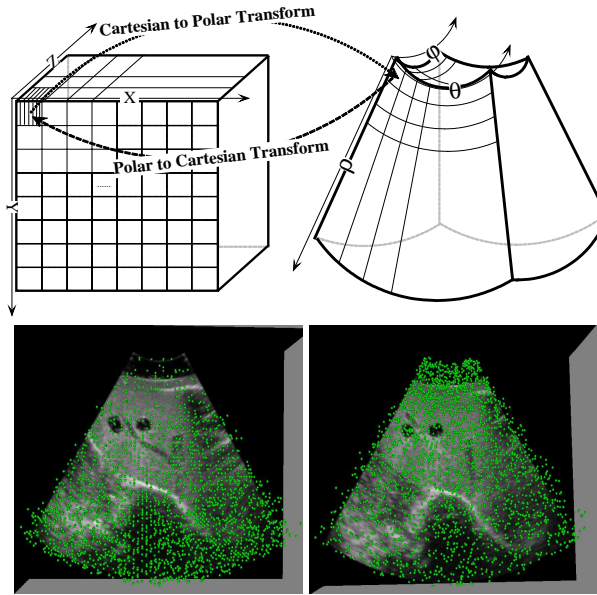
Figure 3: 3D Cartesian coordinate and Polar coordinate spaces. The upper row shows the two coordinates. The bottom row shows the Cartesian coordinate sampling scheme and Polar coordinate scheme in MI calculating. Each sample point is represented by a green dot.

2.5   Registration by Dominant Block Matching

After the blocks with high saliency value are detected, we will use these blocks to locally recover the rigid transform between the US and the CT images. For each of the blocks, an MI based registration is used to acquire the local rigid transform. The region is defined as a box in 3D Polar coordinate. The side length of the region is equal to the size of the optimum scale at the block center. The region with the optimal saliency scale as the radius is used for the registration. The US sub-region is sampled in 3D Polar coordinate and the MI similarity metric is optimized. When all of the local block matchings are finished, we rearrange these blocks by the final MI value. We again take the portion (70%) of the blocks with higher local similarity measurement as useful blocks for next stage. The local rigid transform parameters obtained from these selected blocks will be averaged and it will be used for the initialization of the next global registration.

3   Data and Experiments

3.1   Data Acquisition

We used both *in vitro* and *in vivo* datasets to evaluate our method together with MI method. The former came from a multi-modal abdominal phantom (model 057, CIRS Inc.®), and the latter were the abdominal images of a patient. All 3D US images were taken from a GE Voluson® 730 machine with a 3D transducer of model RAB2-5L. The CT images were taken from a Helical CT machine of GE system®. The images were taken while the

patients held their breath. The CT and US image characteristics for *in vitro* and *in-vivo* experiments including image dimension, voxel size and number of datasets are shown in Table 1.

Experiments	Data	Image Information		
		Dimension	Voxel size(*mm*)	Numbers
in vitro	US phantom	$256 \times 256 \times 256$	$0.915 \times 0.708 \times 0.580$	3
	CT phantom	$256 \times 256 \times 119$	$1.25 \times 1.25 \times 1.25$	1
in vivo	US patient	$256 \times 256 \times 256$	$0.840 \times 0.591 \times 0.640$	3
	CT patient	$512 \times 512 \times 177$	$0.625 \times 0.625 \times 1.25$	1

Table 1: The *in vitro* and *in vivo* US and CT dataset characteristics.

3.2 Experiments

The registration results are evaluated from both visual inspection and quantitative experiments. We evaluate the registration method by starting the registration with random initial parameters for multiple times. For each pair of the images to be registered, a ground truth rigid transform is obtained by using a feature point initialized and intensity based registration software. Several pairs of corresponding points are manually picked by a radiologist and an intensity based method with a manually labeled ROI will refine the initial result. We represent the 3D rigid transform by using six parameters, three for rotations and three for translations. For each datasets, we evaluated these parameters by running 100 registrations, each of which was initialized with an arbitrary transform. Each of the parameters was generated by adding an arbitrary displacement error of parameters to the ground truth parameters. In generating the arbitrary displacement error for parameters, each component of the parameters displacement was chosen within an error range. In our tests, the ranges for the translational and rotational error components were $\pm30mm$, and $\pm0.349rad$ respectively.

4 Results and Discussion

4.1 Accuracy

We list the accuracy test results in Table.2. For phantom and patient registration results, the parameters errors from multiple datasets are averaged. In both the phantom and the patient experiments, the parameter components errors with the ground truth parameters are quite high after MI based registration, while after our proposed method, the component errors with ground truth are much decreased.

4.2 Qualitative Evaluation

Registration of a US and CT phantom image is shown in shown in fig. 4. The first two images in upper row are the CT and US image respectively. The first two images in bottom row are the registration results by using MI registration method and our proposed method. The registration results are shown by one axial slice of color overlaid images. We can see there is small misalignment after the MI registration while after registration by using our method, the resampled US image overlays with the CT image much better. We use three stages in this experiment and figure in upper right shows the adaptively selected

Experiments	Methods	Parameters Errors							
		Rotation(*rad*)				Translation(*mm*)			
		ΔR_x	ΔR_y	ΔR_z	$\|\|\Delta R\|\|$	ΔT_x	ΔT_y	ΔT_z	$\|\|\Delta T\|\|$
US to CT phantom	MI	0.035	0.05	0.033	0.07	1.60	6.52	1.77	6.94
	AMMI	0.033	0.029	0.027	0.06	0.78	0.83	0.92	1.46
US to CT patient	MI	0.05	0.04	0.06	0.08	0.65	5.17	1.33	6.64
	AMMI	0.03	0.04	0.03	0.05	0.49	0.96	1.24	1.65

Table 2: The results of MI and our proposed registration method for the random initial parameters tests. MI: mutual information registration method; AMMI: our proposed adaptive mask mutual information registration method.

mask. We can see the most of shadow region in the right part of the image is excluded. The bottom right figure shows the negative MI metric with adaptive mask plotted against the translational parameters errors around the alignment. The metric function shows an unique and accurate optimum at the matching parameters and the metric function is quite smooth. Registration of a real patient is shown in fig.5. The top row shows the CT and US
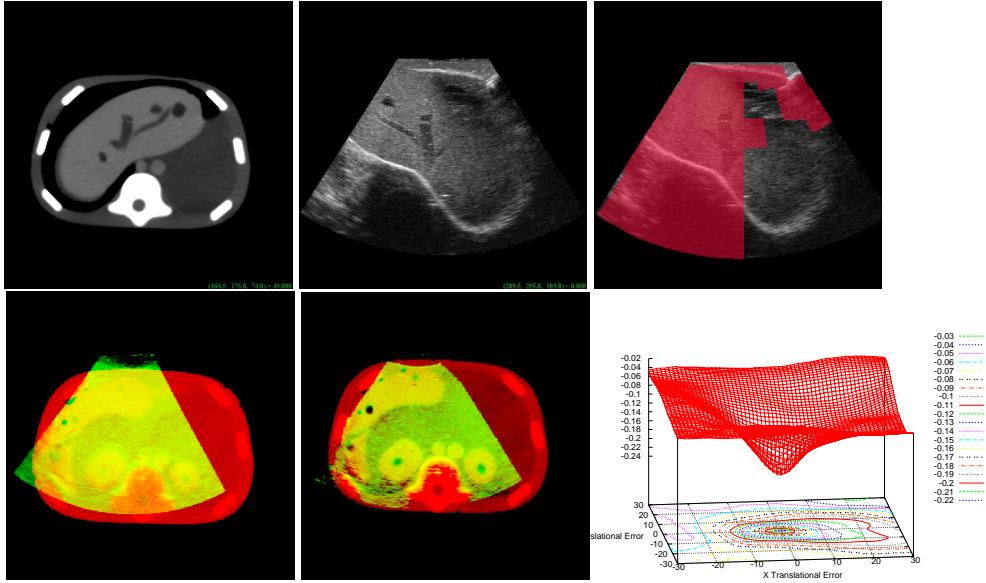

Figure 4: Phantom US to CT registration results by using MI and our method.

images of a patient liver. In the bottom row, from left to right are the registration results by using MI and our method shown by color overlaid . We can see the improvement of alignment near the inferior vena cava.

5   Conclusion

In this paper, we have presented a new rigid 3D US to CT image registration method. We have adaptively selected the regions with high saliency and similarity with CT image
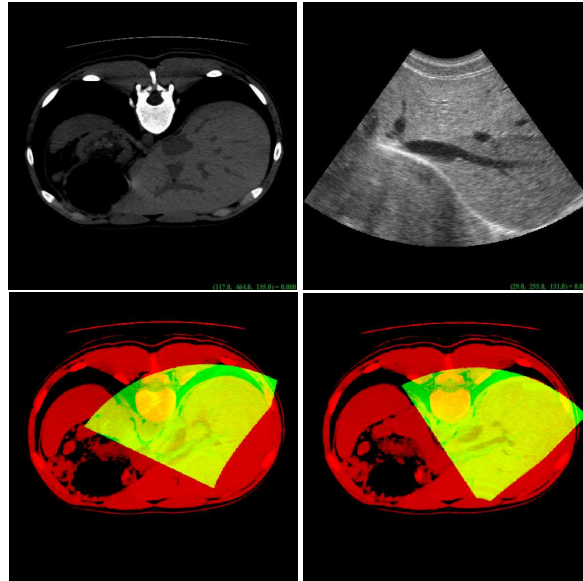
Figure 5: Registration results of US to CT patient liver by using MI and our proposed method.

as useful information for registration. We compared our method with a typical intensity based multi-modal registration method — MI based method, the results of phantom and real patient datasets show the improvement of the accuracy of the registration parameters. This method can be applied to the applications where only partial image exists for registration.

6   Acknowledgement

References

[1] Hybrid image registration based on configural matching of scale-invariant salient region features. In *Second IEEE Workshop on Image and Video Registration*, pages 167–176, 2004.

[2] S.R. Aylward, J. Jomier, J.P. Guyon, and S. Weeks. Intra-operative 3d ultrasound augmentation. In *IEEE International Symposium on Biomedical Imaging*, pages 421–424, 2002.

[3] B. Block. *The Practice of Ultrasound – A Step by Step Guide to Abdominal Scanning*. 2004.

[4] G. Hermosillo and O. Faugeras. Variational methods for multimodal image matching. *International Journal of Computer Vision*, 50(3):329–343, 2001.

[5] X.S. Huang, N.A Hill, J. Ren, and T.M. Peters. Dynamic 3d ultrasound and mr image registration of the beating heart. In *Medical Image Computing and Computer Assisted Intervention , MICCAI 2005*, volume 3750 of *Lecture Notes in Computer Science*, pages 171–178, 2005.

[6] T. Kadir, A. Zissereman, and M. Brady. An affine invariant salient region detector. In *European Conference on Computer Vision - ECCV 2004*, volume 3021 of *Lecture Notes in Computer Science*, pages 228–241. Springer Berlin, 2004.

[7] T. Lange, S. Eulenstein, M. Hunerbein, H. Lamecker, and P.M. Schlag. Augmenting intraoperative 3d ultrasound with preoperative models for navigation in liver surgery. In *Medical Image Computing and Computer Assisted Intervention , MICCAI 2005*, volume 3217 of *Lecture Notes in Computer Science*. Springer, Sep 2004.

[8] A. Leroy, P. Mozer, Y. Payan, and J. Troccaz. Rigid registration of freehand 3d ultrasound and ct-scan kidney images. In *Medical Image Computing and Computer Assisted Intervention , MICCAI 2005*, volume 3216 of *Lecture Notes in Computer Science*, pages 837–844. Springer, Sep 2004.

[9] F. Maes, A. Collignon, and D. Vandermeulen. Multimodality image registration by maximization of mutual information. *IEEE Trans. on Med. Imag.*, 16(2):181–187, 1997.

[10] G.P. Penney, J.M. Blackall, M.S. Hamady, T. Sabharwal, A. Adam, and D.J. Hawkes. Registration of freehand 3d ultrasound and magnetic resonance liver images. *Medical Image Analysis*, 8(1):81–91, 2004.

[11] A. Roche, G. Malandain, N. Ayache, and X. Pennec. Multimodal image registration by maximization of the correlation ratio. *INRIA Research Report*, (3378), 1998.

[12] C.V. Stewart, C.L. Tsai, and B. Roysam. The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Transaction on Medical Imaging*, 22(11):1379–1394, 2003.

[13] C. Studholme, D.J. Hawkes, and D.L.G. Hill. An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, 32(11):71–86, 1999.

[14] W. Wein, R. Barbara, and N. Nassir. Automatic registration and fusion of ultrasound with ct for radiotherapy. In *Medical Image Computing and Computer Assisted Intervention , MICCAI 2005*, Lecture Notes in Computer Science, pages 303–311. Springer, Oct 2005.

[15] R. P. Woods, J. C. Mazziotta, and S. R. Cherry. Mri-pet registration with automated algorithm. *Journal of Computer Assistant Tomography*, 17(4):536–546, 1993.