Non-rigid Registration of Volumetric Images Using Ranked Order Statistics

Ruwan B. Tennakoon*, *Member, IEEE,* Alireza Bab-Hadiashar, *Senior Member, IEEE,*
Zhenwei Cao, *Member, IEEE,* and Marleen de Bruijne

***Abstract*—Non-rigid image registration techniques using intensity based similarity measures are widely used in medical imaging applications. Due to high computational complexities of these techniques, particularly for volumetric images, finding appropriate registration methods to both reduce the computation burden and increase the registration accuracy has become an intensive area of research. In this paper we propose a fast and accurate non-rigid registration method for intra-modality volumetric images. Our approach exploits the information provided by an order statistics based segmentation method, to find the important regions for registration and use an appropriate sampling scheme to target those areas and reduce the registration computation time. A unique advantage of the proposed method is its ability to identify the point of diminishing returns and stop the registration process. Our experiments on registration of end-inhale to end-exhale lung CT scan pairs, with expert annotated landmarks, show that the new method is both faster and more accurate than the state of the art sampling based techniques, particularly for registration of images with large deformations.**

***Index Terms*—Biomedical Imaging, Non-rigid Registration, Robust Statistics, Sampling, Ranked Order Statistics**

I. Introduction

NON-RIGID image registration is the task of finding a transformation that spatially aligns two images with globally non-uniform differences. The Non-rigid registration problem has received substantial attention in previous years and several methods to solve this problem have been proposed. The non-rigid registration methods are broadly classified into two categories based on their transformation models [1]: Physical models and basis function expansions. Physical models are based on physical phenomena such as viscous fluids or elasticity and are described using partial differential equations [2], [3]. Basis function expansions use parametric models such as B-splines [4], radial basis functions [5] and wavelets [6] to represent the transformation [1]. Among all these models the B-Spline transformation model is commonly used for non-rigid registration largely due to its lower computational complexity compared to physical models. A survey of the existing methods of non-rigid registration has been provided in [1], [7]–[9].

Although the implementation of non-rigid intensity based registration is varied in practice, the intensity based registration framework can generally be segregated into three modules [10]: Transformation space, similarity measure and optimization algorithm, for finding a parametrized transformation by minimizing an intensity based cost function. This is essentially a nonlinear optimization problem: Given the intensity values of two images, namely fixed ($I_f(x)$) and moving ($I_m(x)$) images, the transform parameters vector θ is found by solving the following minimization problem [11]:

$$\hat{\theta} = \arg\min_{\theta} \left[\Psi(\theta)\right] \tag{1}$$

where Ψ is the cost function (or dissimilarity measure) and $\hat{\theta}$ is the optimized vector of the transformation parameters.

To determine the optimal set of parameters that satisfies the above equation, gradient based iterative optimization strategies are commonly used [11]. However, for non-rigid registration of medical images with large number of voxels and transformation parameters, calculation of the cost function gradient ($\nabla_\theta \Psi(\theta)$) takes a significant amount of time [12]. There are two commonly used techniques to reduce the computational complexity of this task. The first technique, based on the stochastic sub-sampling method introduced by Robbins and Monro [13], uses only a subset of voxels to estimate the model parameters [12], [14] while the second technique reduces the number of registration parameters by carrying out registration only on selected regions of the image [15] which affect the registration outcome most.

In this paper we propose a fast and accurate non-rigid registration method for images of the same modality that combines the properties of the above two techniques using a robust statistical approach. Our approach exploits the information available in the difference image, by using an order statistics based segmentation method [16], to find the important regions for registration and use an intricate sampling scheme to target those areas and reduce the registration computation time. Our comparative experiments on registration of end-inhale end-exhale lung CT scan pairs, with expert annotated landmarks, show that the new method is faster and more accurate than the state of the art sampling based techniques [12], [15], particularly for registration of images with large deformations.

The outline of the idea of using the information available

*R. B. Tennakoon is with the Faculty of Engineering & Industrial Sciences, Swinburne University of Technology, Hawthorn, VIC 3122, Australia. (e-mail: RTENNAKOON@swin.edu.au).

A. Bab-Hadiashar is with the School of Aerospace, Mechanical and Manufacturing, RMIT University, Melbourne, VIC 3001, Australia (e-mail: abh@rmit.edu.au).

M. de Bruijne is with the Department of Computer Science, University of Copenhagen, DK-2100, Copenhagen, Denmark. and with the Biomedical Imaging Group, Departments of Radiology and Medical informatics, Erasmus MC - University Medical Center, Rotterdam, The Netherlands.

Z. Cao is with the Faculty of Engineering & Industrial Sciences, Swinburne University of Technology, Hawthorn, VIC 3122, Australia.

in difference images to speed up registration was presented in [17]. The current paper builds upon our earlier work by presenting an extensive set of comparative experimental results and a detailed discussion of the merits of the proposed algorithms. The rest of this paper is organized as follows. Section II provides a brief description of previous research in this area. Section III explains the proposed registration scheme. The experimental set up and the comparative registration results of the proposed methods are presented in section IV. The performance of the proposed methods in comparison with the best available techniques and the advantages and drawbacks of those techniques are discussed in Section V. Section VI concludes the paper.

II. Previous Research

There are two commonly used techniques to reduce the computational cost of parametric non-rigid image registration. Rohde et al. [15] first introduced the idea of carrying out the registration only on the regions that are not so well registered. Their method, called Adaptive Registration (AR), uses a hierarchical framework, where at each level a set of radial basis functions are placed on an irregular grid. The nodes of the irregular grid are the centers of the misregistered regions in the image at a particular level. Using such an irregular grid reduced the number of parameters compared to a fixed grid registration method leading to a reduction in the computational cost. In order to identify the misregistered regions, the AR method uses an additional step in which another set of basis functions are placed on a fixed grid and the gradient magnitudes of the cost function with respect to the parameters of the fixed grid are evaluated. The locations of the remaining grid points after eliminating the ones with low gradients are then considered as centers of the regions of misregistration (nodes of the irregular grid at that level).

A careful analysis of the above method reveals that it is necessary to recalculate the cost function derivatives, involving all the voxels, at the beginning of each registration step to determine misregistered regions. This calculation is computationally expensive hence the overall registration process takes a relatively long time to complete. Our inspiration for the proposed method stems from the fact that the registration time will be significantly reduced if the misregistered regions are identified more efficiently.

The second technique to reduce computational time is to use only a subset of voxels to estimate the parameters on a fixed grid. Klein et al. [14] were the first to advocate the use of uniform sampling for non-rigid medical image registration. The uniform sampling gives equal significance to all the voxels in an image and since the true function in this case is not distributed uniformly, using only a subset of voxels produces a biased estimate of the gradients. To reduce the sampling bias they proposed to renew the set of chosen samples at every iteration of the optimization routine giving equal significance to all the voxels. To improve the registration accuracy Bhagalia et al. [12] introduced the idea of using the importance sampling (IS) technique to select a subset of voxels. Importance sampling is a way of obtaining the properties of the desired distribution using another but related distribution. In their method, they used the properties of the edge magnitude to approximate the probability distribution of the gradients.

In this paper we propose a new method that combines the benefits of both techniques mentioned above. Although our method focuses on misregistered parts in the image, the method used for finding them is completely different to that of the method used by AR. Here, we move away from trying to define the misregistered regions based on cost function gradients and directly use the information carried by the misregistered points to improve the optimization outcome.

We propose a new approach for the identification of misregistered points in which the square intensity differences of fixed and transformed moving image voxels guide the identification process using a rank ordered statistics based robust segmentation technique. In addition to improving the accuracy of registration our method provides a computationally efficient means of stopping the registration process once the improvements become small. In practice it is very important to know precisely when to end the registration process [10]. Our proposed method is described in the following section.

III. The Proposed Method

A. How to Sample Efficiently?

The main idea of our proposed method is to turn the registration focus to the regions of the image which is not so well registered. The challenge is to efficiently identify the regions of misregistration at each step of the registration process. Our contribution here is to devise an efficient method that exploits the information available in the difference image (the image intensity difference between the fixed image and the transformed moving image) to identify the regions of misregistration. We note that the sum of square difference (SSD) cost function is not significantly influenced by the small intensity differences. A very small difference between two corresponding voxels means that the voxels are either almost registered or they are both in a low textured region. As such, the significance of those voxels to the registration process is minimal (those voxels are referred to as group one voxels). On the other hand very high variations are often caused by occasional differences such as misregistered vessels, organ boundaries or gross measurement errors (outliers). Although those are almost always present in every image and could be relevant for registration, their inclusion has the potential to skew the optimization and bias the final results (those are referred to as group three voxels). To visualize how the intensity differences are distributed for a given image, the joint histogram between the fixed image and the transformed moving image during a registration process is shown in Fig. 1. The joint histogram shows that a large number of voxels are clustered around the center line. These are the voxels that correspond to an almost registered group of voxels (group one).

The challenge is to find computationally efficient ways of excluding the group one and three voxels and use only the voxels in the middle group (group two) for registration at every

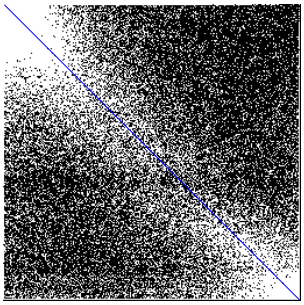


Fig. 1. The joint histogram of the fixed image ($I_f(x)$) and the transformed moving image ($I_m(T(x))$) for case 5 of DIR-dataset at the 10^{th} iteration.

step. This is, to an extent, similar to the multi-structural data segmentation and many different methods to dichotomize data based on the distribution of residuals have been developed. The emphasis of these methods is to group data points that are most similar and the analysis presented in [18] showed that most of these methods have similar performances.Following recommendations of [18], we used the Modified Selective Statistical Estimator (MSSE) [16] to segment voxels. The implementation of MSSE is very straightforward (computationally efficient), the estimator has nice asymptotic properties [19] and the value of its required parameter (i.e. the minimum size of an acceptable group) is known in this case. Having said this, we expect the use of other similar robust estimators to produce comparable results.

In our implementation of MSSE, the members of the first group are found by using the following criterion starting from the median of absolute residuals [16]:

$$|r_{i+1}| < T\sigma_i \tag{2}$$

here i is the index after sorting, $|r_{i+1}|$ is the absolute image intensity difference (residual) at index $i+1$, T is a constant threshold (2.5 is used to include 99% of inliers based on a normal distribution [16]) and $\sigma(i)$ is the standard deviation of sorted residuals up to index i. Fig. 2 shows the classification of different voxels based on their absolute residuals using the MSSE constraint (2).

As we mentioned earlier, the voxels with very large differences have the potential to bias the registration process. As such, the last five percent of the largest residuals are also discarded from the group identified as outliers by the MSSE. Once the voxels that belong to the second group are identified, we modified the sum of squared differences (SSD) cost function to include only voxels that belong to this group at every iteration. Since the group memberships are decided at every iteration, voxels with large differences that have the potential to improve their registration will eventually be included in the second group. The proposed cost function is written as:

$$\Psi(\theta) = \frac{1}{(i_2 - i_1)} \sum_{n=i_1}^{i_2} r_\theta^2(x_i) \tag{3}$$

where i_1 and i_2 are the first and last sorted indices of the squared residuals that belong to the middle group and $r_\theta^2(x_i) = (I_f(x_i) - I_m(T_\theta(x_i)))^2$. $T_\theta(x_i)$ is the B-Spline transformation with parameter θ.

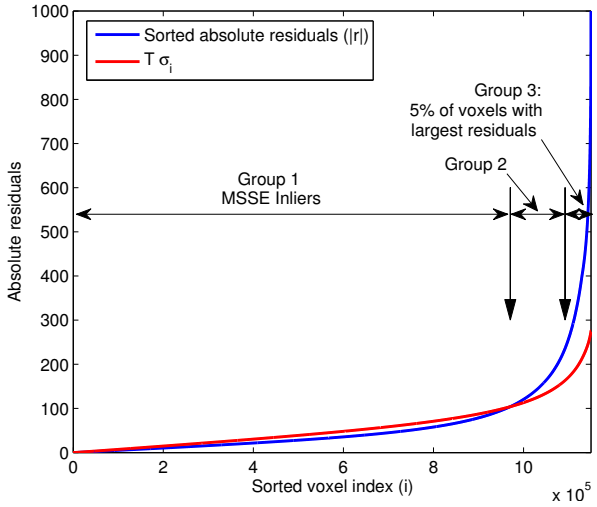


Fig. 2. Classification of different voxels based on their absolute residuals using the MSSE constraint. The arrows show the boundaries between groups in a sample data set.

Having defined the cost function, the next step is to estimate the gradient of the cost function. The estimation of the gradient using all the chosen points would still be computationally expensive. Adopting sampling based stochastic optimization techniques [13] enables us to perform the registration using a computationally cheaper estimate of the cost function gradients. To estimate the gradients using this approach we can either use random sampling or importance sampling techniques [12].We have developed registration algorithms using both techniques and those are named robust random sampling (RRS) and Robust importance Sampling (RIS), respectively. In the importance sampling algorithm [12], the gradient magnitudes of both the fixed ($|\nabla I_f(x)|$) and the transformed moving ($\left|\nabla I_m^\theta(x)\right|$) images are first used to compute the following probability distribution function (PDF):

$$P_s^\theta(i) = \frac{e_i^\theta}{\sum_{j=1}^n e_i^\theta}; i = 1....n \tag{4}$$

where $e_i^\theta = \frac{|\nabla I_f(i)|}{\sum_{j=1}^N |\nabla I_f(j)|} + \frac{\left|\nabla I_m^\theta(i)\right|}{\sum_{j=1}^N |\nabla I_m^\theta(j)|}$

. The inverse cumulative distribution function (CDF) sampling [20] is then used to choose the appropriate voxels for gradient estimation. To explain how the above method works, we first note that when a random variable X has a cumulative distribution function F, then the values of $F(X)$ is uniformly distributed in the range $[0, 1]$. As such, a set of uniformly distributed samples of the calculated probabilities can be used to generate samples of the fixed image that has the desired distribution P.

B. Convergence Criterion

Another important challenge in successful implementation of stochastic optimization schemes is to find an appropriate stopping criterion. Simple criterion used for gradient descent

Algorithm 1 Step-by-Step Algorithm of Proposed Methods

Inputs: Fixed Image, Moving Image, transform Parameters

1: Apply the transform to the moving image and recalculate the intensity values at grid points (Transformed Image).
2: Calculate the absolute differences between the fixed and the transformed images ($|r|$).
3: Sort the $|r|$ values and keep their orders (ascending).
4: Find K, K is the first index of the sorted residuals where $|r_{i+1}| < T\sigma_i$ is not satisfied .
5: Find $K_L = 0.95 \times$ Total Number of Voxels.
6: **if** Sampling Scheme = RRS **then**
7: Randomly Select N voxels starting from the K^{th} position in the sorted array up to the K_L^{th} position as the samples .
8: **else if** Sampling Scheme = RIS **then**
9: Select N voxels starting from the K^{th} position in the sorted array up to the K_L^{th} position using inverse CDF (CDF of the edge magnitudes of the fixed and transformed moving images) sampling, as the samples.
10: **end if**
11: **if** ($C_k \leq$ THRESHOLD) **then** Stop Optimization

algorithms such as the magnitude of the gradient or the change in cost function are inappropriate for stochastic optimization as the gradient approximate is noisy and may not vanish near the optimal solution. To overcome this limitation most practical algorithms simply use a fixed number of iterations [11], [12], [15]. In practice this is not an appropriate solution and the required number of iterations is not known a priori. Fixed number of iterations for all cases would either lead to inadequate accuracy or waste of computation resources.

An important advantage of the proposed algorithm is that the degree of convergence of the registration can be directly measured by looking at the number of voxels identified by the MSSE as inliers (the voxels that are almost registered). As such, the increase of voxels in this group at every iteration is proportional to the increase in the registration accuracy. To take advantage of this characteristic we devised a stopping criterion by developing a normalized measure of this attribute. The measure defined is as follows:

$$C_k = \frac{N_k - N_{k-1}}{N_{k-1}} \times 100\% \tag{5}$$

where N_k denotes the number of voxels in the inlier group at iteration k. The registration is then deemed successful once C_k is less than a specific threshold (0.1% in all of our experiments).

An added advantage of this approach is that the computational cost associated with calculating this criterion is negligible. A complete description of the proposed registration algorithms is provided in Algorithm 1.

IV. Analysis of the Proposed Algorithms

A. Experimental Setup

To compare the performance of the proposed registration methods with the best available techniques, an extensive set of experiments using lung CT images were conducted. The elastix registration toolbox [14] based on the Insight Segmentation and Registration Toolkit (ITK) was used as the basis of the registration framework that was developed to implement both the Importance Sampling and the proposed Robust Sampling methods. The code provided by the first author of [15] was used to obtain the results for the AR method. All experiments were conducted using the HP Z400 workstation with single Intel Xeon W3550 3.06GHz processor.

Two CT datasets were used to estimate the registration accuracy of the two proposed methods and to compare those with the state of the art methods. The first was the CT dataset used by Bhagalia et. al. [12] (referred to as IS-dataset). This dataset consists of 8 pairs (breath-hold) of maximum inhale to maximum exhale CT images with $1.87 \times 1.87 \times 5.0\ mm^3$ voxels [21]. The second data set, referred to as DIR-dataset, was provided by the University of Texas MD Anderson Cancer Center [22]. This dataset consists of 4D-CT images of ten different patients each consisting of a sequence of images taken through a full respiratory cycle. In our experiments we registered the maximum inhale and maximum exhale images in each case. Every case of the 4D-CT dataset has 300 expert identified landmarks and their associations at full inhale and exhale images. In our experiments all images were cropped and segmented to include only the lungs. The segmentation was done automatically using ITK-SNAP [23] software.

B. Implementation

The first step in our registration framework is to implement an appropriate hierarchical scheme to ensure that the registration process is not trapped in a local minimum. To make our results comparable with the ones published in [12], we also implemented a two level Gaussian pyramid scheme [8]. In this scheme, the amount of data in the initial level of the registration process is down-sampled by a factor of two and smoothed using a Gaussian kernel $\mathcal{N}(0,1)$ while in the second level, the original image is used.

For the transformation model the cubic B-Spline described by the following equation [24] was used:

$$T_\theta(x) = x + \sum_{x_k \in N_x} \theta_k \beta^3(\frac{x - x_k}{s}) \tag{6}$$

where θ_k is the k^{th} parameter, $\beta^3(x)$ is the 3rd order B-Spline polynomial given by [25]:

$$\beta^3(x) = \begin{cases} \frac{2}{3} - |x|^2 + \frac{|x|^3}{3} & 0 \leq |x| \leq 1 \\ \frac{(2-|x|)^3}{6} & 1 \leq |x| \leq 2 \\ 0 & 2 \leq |x| \end{cases} \tag{7}$$

and N_x is the set of all control points (x_k) within the compact support of the B-Spline at point x. The control points for each level were organized in a fixed grid with voxel spacing of $16 \times 16 \times 8$ for the first level and $8 \times 8 \times 4$ for the second level. In order to optimize the cost function the stochastic gradient descent described by the following equation was used:

$$\theta_{k+1} = \theta_k + a_k \hat{g}(\theta_k) \tag{8}$$

where a_k is a vector that controls the step size in different directions at every iteration of the algorithm and $\hat{g}(\theta_k)$ is an appropriate estimate of the gradient. For the stochastic optimization solution to converge, the step size a_k should satisfy the following conditions: $a_k \geq 0$; $a_k \to 0 \; as \; k \to \infty$; $\sum_{k=0}^{\infty} a_k = \infty$; $\sum_{k=0}^{\infty} a_k^2 < \infty$. A common way to ensure that the step size satisfies the above criterion is to use the following step size formula [26]:

$$a_k = \frac{a_0}{(A+k)^{\alpha}} \tag{9}$$

where k is the iteration number and a_0 and A are empirically determined constants. However, for problems with large number of parameters, the optimal step size for every parameter may vary widely and therefore a specific step size for each parameter needs to be calculated. In this case, to ensure convergence, the following equation is commonly used to calculate the step size for each parameter [27]:

$$a_k^i = \frac{a_0}{(A+Q_k^i)} \tag{10}$$

where Q_k^i is the number of sign changes in the i^{th} parameter up to the k^{th} iteration.

Another commonly used method to ensure convergence in stochastic approximations is to increase the sample size progressively during the iterations [28]. In our work we have adopted a combination of these two methods as advocated by [12]. Using the combination scheme not only ensures the convergence of our proposed methods but also makes the final results directly comparable with [12]. There are four key parameters to be selected in this optimization scheme: number of iterations, A, a_0 and number of samples in each iteration.

All competing registration methods, in contrast to those proposed here, need a predefined measure to stop the registration process. For these methods the maximum number of iterations at each level of hierarchy was set to 60 and 100, respectively. These represent the best case scenario (in terms of timely registration with the best accuracy) for those methods determined by manual inspection of mean landmark errors in registered test images. In practice, this information is not known for each image and can significantly increase the overall required computation time of those registration methods. To make all the results comparable, in all our experiments involving sampling based methods, we used 4096 and 8192 samples at the sequential registration levels. We observed that the registration accuracy does not change significantly with small variations of the number of samples. The parameters A and a_0 are manually tuned to achieve the overall best registration outcomes. Following [12], the value of A was set to 10 and the mean land mark error for different values of a_0 were calculated. Overall, the $a_0 = 500$ appeared to be the most suitable, this value was kept constant for all the registration methods in order to make the results comparable. Our experiments showed that the overall results is not significantly affected by the small variations of a_0 (increased or decreased by two or three times).

In MSSE implementation the basic quick sort algorithm was used in our experiments and the timing of different runs showed that on average around 15% of the total computation time was spent on the sorting step.

C. Analysis With synthetic images

To demonstrate the effectiveness of the proposed algorithm in identifying the misregistered regions, a simulation study involving a pair of synthetic images (constructed by deforming a lung CT image with known deformations and adding normal noise) was conducted. The second image was designed to exhibit localized motions: changes are confined to two specific regions of the image.

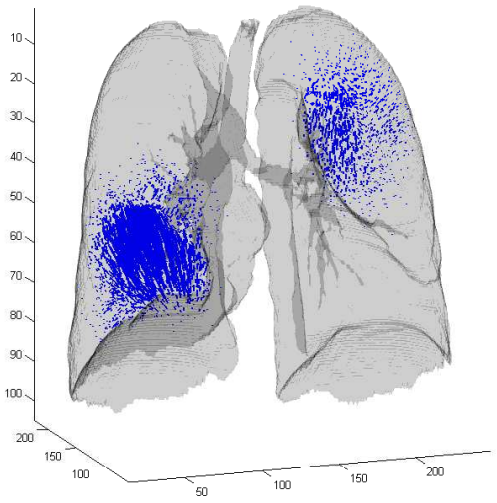
The above images were then registered using the proposed method and the voxels classified as the second group (used in the actual registration process) were separated and plotted in Fig. 3(b) together with the surface rendering of the lung. To show the correspondence between actual motion and identified voxels, the actual motion field together with the surface rendering of the lung image is shown in Fig. 3(a). These images show that the proposed method has been able to correctly identify misregistered voxels. The proposed methods will in turn place more emphasis on these voxels during the registration process.

D. Performance variation with noise

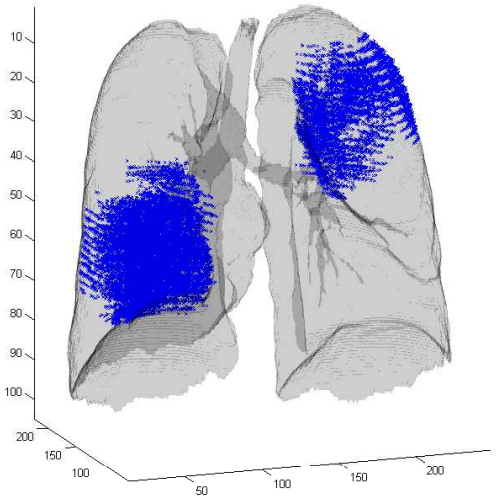
The impact of noise on the proposed registration algorithm has been studied by creating images with known deformations and additive normal noise. In those simulations a lung CT image is first deformed using a predefined B-spline transform that was calculated in a real image experiment using DIR dataset. The intensities of both fixed and deformed images were normalized between zero and one. Zero mean normally distributed noise with standard deviation ranging from 0.0001 to 0.05 was added to both images. It is important to note that since the lung densities cover only around 25% of the entire range of Hounsfield values in the CT images, the maximum noise of 0.05, which is added to the entire range of normalized density values of both images, would represent a substantial amount of actual noise in lung regions. The noisy images were then registered with the proposed algorithms. Finally, registration errors for each image pair were calculated using a set of randomly distributed points and averaged over five repeated experiments. The results of those experiments are shown in Fig. 4. To show the actual amount of added noise the registration errors are plotted in terms of the variance of added noise in Hounsfield units. The figure shows that for relatively large amounts of additive noise (compared to image intensities), the registration error remains unaffected.

E. Experimental Results

The first set of experiments were conducted using the IS-dataset. The registration accuracy of the proposed methods in comparison with the competing methods using the above dataset is shown in Table I. The result shows that the proposed RIS method always achieves better results than the IS and AR. Following recommendations by Demšar [29] on statistical comparisons of classifiers over multiple datasets, the final



(a)



(b)

Fig. 3. The results of the simulation demonstrate the effectiveness of the proposed algorithm in identifying the misregistered regions (a) The predefined motion field together with the surface rendering of the lung. (b) Plot of all voxels classified as the second group in the simulation.

mean registration errors achieved using different algorithms were compared with RIS method using the Wilcoxon rank-sum test [30]. The p-values, after Holms [31] correction, are given in Table III. Those results show that RIS produces significantly better results than the competing methods.

The limitation of the IS-dataset is that it has only six annotated landmarks. To provide more compelling evidence, we also conducted experiments using DIR-dataset [22], which has 300 annotated landmarks. Those landmarks were used to evaluate the registration performance of all methods based on

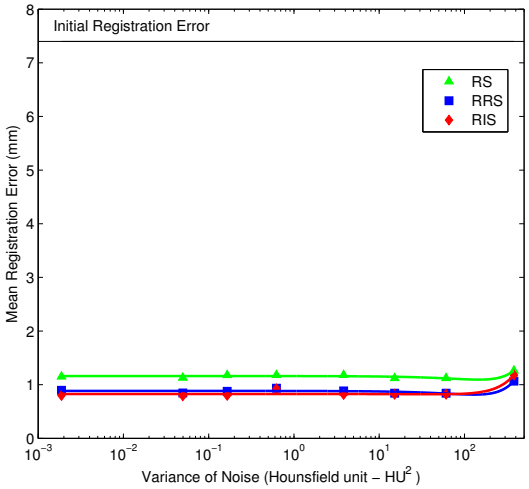


Fig. 4. Performance variation of registration error of the proposed algorithms verses the variance of manually added normal noise. The RS method is also included for the sake of comparison.

their mean landmark errors. The results of those experiments are shown in Table II. The table also includes the standard deviation of the landmark errors and the maximum/minimum of the mean errors of ten repeated experiments. A Wilcoxon based statistical analysis similar to the IS-dataset was performed on these results as well and the results are given in Table III. The results show that the proposed RIS algorithm performs significantly better than the competing algorithms. Table III also shows the results for the Wilcoxon test using all 18 test images, which is in-line with the previous results.

To compare the performance of RIS in comparison with the competing algorithms for each image of the DIR-dataset, a Wilcoxon rank-sum test [30] was used (this test uses the error at each landmark point averaged over ten repeated runs). The p-values after correcting for multiplicity using Holms procedure is given in Table IV. These results show that the proposed RIS method in comparison to all competing methods (AR, RS, IS) produce significantly better results in cases with large deformations (DIR-C4 to C8). It is important to note that the first few cases include very small motions and all methods achieve very good registration results.

TABLE III
AVERAGE RUNNING TIME FOR EACH ALGORITHM AND THE HOLMS CORRECTED P-VALUES OF WILCOXON RANK-SUM TEST THAT COMPARE THE FINAL MEAN REGISTRATION ERROR OF COMPETING METHODS WITH THE RIS METHOD.

	IS-dataset		DIR-dataset		Combined
	Time (s)	p-value	Time (s)	p-value	p-values
AR	348.0	0.03	1,757.0	7.81E-03	**7.00E-04**
RS	55.0	0.03	24.8	7.81E-03	**7.86E-04**
IS	112.0	0.03	100.0	0.03	**1.07E-03**
RRS	43.0	0.04	41.5	0.03	**2.47E-03**
RIS	65.0	-	86.5	-	**-**

To compare the computational complexity of these methods, the average computation time of all competing methods for registration of all images in both datasets were measured and the results are shown in Table III. We note that both RS and

TABLE I
COMPARISON OF MEAN LANDMARK ERROR (EUCLIDEAN DISTANCE) FOR IS DATASET MEASURED USING EXPERTLY IDENTIFIED LANDMARK POINTS.

		IE[1]	AR[2]	RS[3]	IS[4]	RRS[5]	RIS[6]
IS-C1	**Mean**	15.1	4.5	4.39	3.69	3.61	3.39
	Std		2.32	1.28	2.04	2.10	2.14
	Min/Max			4.39/4.86	3.55/3.81	3.37/3.79	2.98/3.88
IS-C2	**Mean**	14.52	6.49	4.53	2.56	2.27	2.51
	Std		9.81	3.64	3.00	2.67	2.38
	Min/Max			4.34/4.65	2.11/3.05	1.59/2.79	2.48/2.58
IS-C3	**Mean**	13.31	4.57	2.74	2.62	2.31	1.77
	Std		3.50	1.47	2.60	1.94	2.33
	Min/Max			2.38/3.30	2.45/2.76	1.90/2.94	1.64/2.03
IS-C4	**Mean**	11.73	1.69	3.21	2.46	2.01	1.73
	Std		0.89	2.70	2.02	1.46	1.10
	Min/Max			2.52/3.96	2.26/2.81	1.69/2.39	1.38/2.00
IS-C5	**Mean**	9.13	2.66	3.46	3.24	2.14	1.62
	Std		2.08	2.30	2.29	1.70	2.02
	Min/Max			3.35/3.60	3.18/3.29	1.77/2.45	1.20/1.82
IS-C6	**Mean**	8.62	3.18	1.16	0.94	0.96	0.86
	Std		2.23	1.30	0.99	1.17	1.19
	Min/Max			1.06/1.20	0.62/1.25	0.62/1.19	0.75/1.06
IS-C7	**Mean**	7.77	1.82	2.43	1.86	1.86	1.39
	Std		1.95	2.10	2.05	2.05	1.61
	Min/Max			1.51/3.18	1.46/2.35	1.46/2.35	0.94/1.77
IS-C8	**Mean**	6.89	3.29	2.34	2.70	2.73	2.32
	Std		2.35	1.93	2.17	2.33	1.96
	Min/Max			2.08/2.79	2.53/2.79	2.48/2.92	2.15/2.47
Overall	**Mean**	10.88	3.53	3.03	2.51	2.24	1.95

[1] Initial Error without any registration [2] Adaptive Registration [3] Random Sampling
[4] Importance Sampling [5] Robust Random Sampling [6] Robust Importance Sampling

TABLE II
REGISTRATION PERFORMANCE FOR IMAGES IN DIR-DATASET MEASURED USING EXPERTLY IDENTIFIED LANDMARK POINTS. EACH ALGORITHM IS REPEATED 10 TIMES TO ACCOUNT FOR THE RANDOMNESS.

		IE	OE[1]	AR	RS	IS	RRS	RIS
DIR-C1	**Mean**	4.01	0.85	0.98	1.11	1.03	1.00	**0.94**
	Std	2.91	1.24	1.51	1.01	1.00	0.99	0.96
	Min/Max				1.06 / 1.14	0.99 / 1.12	0.96 / 1.03	0.90 / 0.99
DIR-C2	**Mean**	4.65	0.7	1.10	1.15	0.97	0.92	**0.88**
	Std	4.09	0.99	1.87	1.23	1.06	0.97	0.97
	Min/Max				1.11 / 1.25	0.90 / 1.00	0.85 / 0.98	0.84 / 0.96
DIR-C3	**Mean**	6.73	0.77	1.26	1.39	1.16	1.08	**1.03**
	Std	4.21	1.01	1.68	1.25	1.15	1.10	1.08
	Min/Max				1.32 / 1.48	1.11 / 1.21	1.04 / 1.12	0.96 / 1.10
DIR-C4	**Mean**	9.42	1.13	1.93	1.84	1.68	1.64	**1.55**
	Std	4.81	1.27	2.24	1.51	1.32	1.33	1.26
	Min/Max				1.77 / 1.89	1.62 / 1.74	1.61 / 1.69	1.48 / 1.64
DIR-C5	**Mean**	7.10	0.92	3.06	2.91	2.38	2.12	**1.70**
	Std	5.14	1.16	4.50	3.03	2.66	2.29	1.68
	Min/Max				2.70 / 3.09	2.26 / 2.56	2.02 / 2.23	1.59 / 1.84
DIR-C6	**Mean**	11.10	0.97	5.27	2.05	1.73	1.65	**1.58**
	Std	6.98	1.38	8.29	1.70	1.26	1.21	1.19
	Min/Max				1.95 / 2.22	1.66 / 1.85	1.54 / 1.70	1.48 / 1.71
DIR-C7	**Mean**	11.59	0.81	4.16	3.31	2.31	2.16	**1.71**
	Std	7.87	1.32	6.03	3.40	2.35	2.08	1.27
	Min/Max				3.03 / 3.55	2.14 / 2.58	1.95 / 2.58	1.63 / 1.81
DIR-C8	**Mean**	15.16	1.03	5.93	3.79	2.42	2.21	**1.76**
	Std	9.11	2.19	7.12	4.54	2.74	2.55	1.74
	Min/Max				3.33 / 4.05	2.27 / 2.62	1.96 / 2.49	1.64 / 1.86
DIR-C9	**Mean**	7.82	0.75	2.85	1.51	1.48	1.47	**1.43**
	Std	3.99	1.09	2.64	1.22	1.23	1.17	1.09
	Min/Max				1.47 / 1.57	1.42 / 1.51	1.41 / 1.54	1.34 / 1.53
DIR-C10	**Mean**	7.63	0.86	3.26	1.89	**1.64**	1.65	1.74
	Std	6.54	1.45	5.40	1.84	1.46	1.55	2.03
	Min/Max				1.79 / 1.93	1.53 / 1.71	1.50 / 1.73	1.54 / 2.13
Overall	**Mean**	8.52	0.88	2.98	2.09	1.68	1.59	1.43

[1] Observer Error: The mean reproducibility of the landmark extraction process which is measured using several independent landmark selection experiments (see [22] for details).

TABLE IV
HOLM'S ADJUSTED P-VALUES FOR THE WILCOXON RANK-SUM TEST CONDUCTED BETWEEN THE LANDMARK ERROR OF RIS AND OTHER COMPETING METHODS.

	AR	RS	IS	RRS
DIR-C1	0.34	1.36E-04	0.04	0.39
DIR-C2	0.07	1.64E-05	0.10	0.14
DIR-C3	0.10	9.54E-09	4.00E-04	0.03
DIR-C4	0.05	1.56E-04	1.37E-03	0.04
DIR-C5	1.46E-05	5.41E-16	7.61E-08	2.43E-04
DIR-C6	8.48E-09	1.09E-07	3.00E-04	0.17
DIR-C7	3.66E-08	7.65E-20	3.35E-08	1.09E-09
DIR-C8	2.47E-19	1.64E-08	1.37E-04	0.01
DIR-C9	3.67E-14	0.50	0.65	0.61
DIR-C10	2.27E-04	0.18	0.73	0.26

IS methods use fixed number of iterations as the stopping criterion and their computation times would heavily depend on the number of required iterations, which is considered an input. The experiments showed that the computation time of our methods are significantly lower than that of the competing methods. The registration time for the AR method is significantly higher than other competing methods. This point is shown in Table III and Fig. 5 (a) where the mean error of all the algorithms (RS, IS , RIS, RRS and AR) are compared against the time taken to achieve that level of accuracy for a sample image (case 1). We also note that the proposed RRS method is more than twice faster than the IS in both datasets.

Fig. 5 (b-d) shows the mean errors of all the sampling based algorithms with the number of iterations taken to achieve that level of accuracy. Those figures show that both proposed methods reduce the mean error faster than RS and IS. We also observed in our experiments that there are less sign changes in the RIS estimated gradient of the cost function compared to RS and IS. This implies that the step size reductions in RIS are slower which contributes to its faster convergence. The graphs also show that the convergence criterion described in section III-B has been successful in stopping the registration process at appropriate points (stable results).

To further analyse the performance of the proposed convergence criterion, we compared the accuracy and computation time of the proposed RIS method with and without the proposed convergence criterion. Here, without convergence criterion refers to RIS registration with a fixed number of iteration (similar to IS and RS registrations). The results are shown in Table V. The p-values (Holms Corrected) of Table V are the results of a Two One-Sided Test (TOST) between the final landmark errors achieved with and without the convergence criterion. These numbers show that the mean landmark errors achieved with and without the proposed convergence criterion are equivalent (within a margin of 5% of the initial landmark error). Calculating the average of the last column of Table V shows that an average time saving of around 44% is achieved by using of the proposed convergence criterion.

To show that there is a linear correlation between the proposed convergence criterion C_k and the registration accuracy throughout the registration process, Pearson's correlation coefficients between the actual registration accuracy measured using mean landmark error and C_k at the end of each iteration for the top level are enumerated in Table VI. The results show

TABLE V
THE IMPROVEMENT OF THE REGISTRATION TIME DUE TO THE USE OF CONVERGENCE CRITERION AND THE RESULTS OF THE TOST (LOW P-VALUES MEANS THE RESULTS ARE STATISTICALLY EQUAL).

	Mean LM Error With CC	**Mean LM Error Without CC**	**TOST p-value**	**Time Without CC**	**Time saved %**
DIR-C1	0.94	0.96	**5.6E-04**	85.3	13.4
DIR-C2	0.88	0.93	**1.3E-03**	123.6	43.9
DIR-C3	1.03	1.11	**2.3E-04**	119.2	64.7
DIR-C4	1.55	1.60	**2.4E-07**	115.4	34.0
DIR-C5	1.70	1.62	**6.7E-04**	124.3	62.1
DIR-C6	1.58	1.66	**5.1E-10**	142.7	64.0
DIR-C7	1.71	1.79	**5.1E-10**	138.7	87.7
DIR-C8	1.88	1.85	**5.3E-14**	153.4	20.9
DIR-C9	1.43	1.45	**9.8E-07**	99.23	40.6
DIR-C10	1.74	1.74	**4.4E-05**	124	12.3

that there is a strong linear correlation between the measure for convergence criterion and the actual registration error.

TABLE VI
THE MEASURE OF CORRELATION BETWEEN THE PROPOSED CONVERGENCE CRITERION AND THE ACTUAL REGISTRATION ERROR MEASURED USING EXPERT IDENTIFIED LANDMARKS. A VALUE IN EXCESS OF 0.5 FOR PEARSON'S CORRELATION INDICATES A STRONG LINEAR CORRELATION. IMAGES FROM THE DIR DATASET IS USED.

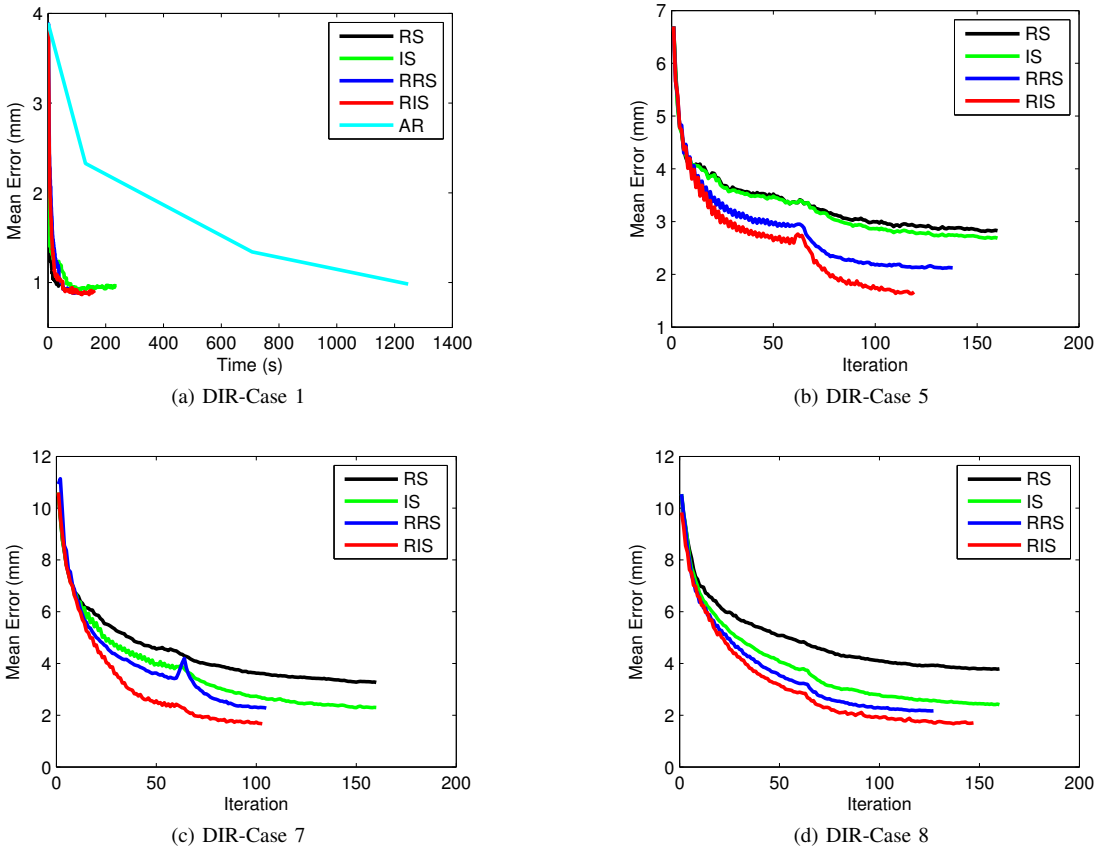
C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
0.93	0.96	0.76	0.86	0.8	0.89	0.7	0.9	0.87	0.82

To visualize the registration quality of the proposed methods the intensity difference image of three slices before and after the RIS registration are shown in Fig. 6. Those slices show that the registration errors have reduced significantly after the registration with the proposed method.

V. DISCUSSION

This paper has presented a new approach to focus the registration on misregistered points identified using a rank ordered statistics based robust segmentation technique. Experiments using lung CT images showed that the proposed method achieved high registration accuracies compared to similar methods. The registration time for the proposed method was also significantly reduced using a new convergence criterion.

It is not surprising to see that the RS method, due to its simplicity, is computationally more efficient than all of the above methods. However, the accuracy of the random sampling method is poor. Our experiments showed that running the RS with more iterations would not improve its registration accuracy. To provide an example, the mean errors of case five using the random sampling algorithm with significantly higher number of iterations (100 and 400 for two registration levels compared to 60 and 100 originally) are shown in Fig. 7. This figure shows that the accuracy of the RS method does not increase by increasing the number of iterations. This is due to the fact that in stochastic optimization, to ensure convergence, the step sizes are incrementally reduced and when the step size is very small, no significant improvement is made from additional iterations. A Two One-Sided Test (TOST) was conducted with the null hypothesis: Mean registration Errors



(a) DIR-Case 1 (b) DIR-Case 5

(c) DIR-Case 7 (d) DIR-Case 8

Fig. 5. (a) Rate of the mean landmark error reduction over time for different registration methods for Case 1. (b-d) Rate of decrease of mean landmark error vs number of required iterations for different methods in cases where registration involves significant deformations.

of RS with low and high number of iterations are not equal (the margin is set to 5% of the error before registration). The p-values after correcting for multiplicity using Holms procedure were also calculated and the results showed that the accuracy of the RS does not improve significantly as the number of iterations increases. The RS algorithm was also tested with three times more samples per iteration. Again, the RS final mean error did not significantly change by increasing the sample size.

To show the effect of using rank order statistics based segmentation during registration, joint histograms of fixed and moving images before and after the application of MSSE are shown in Fig. 8. Those figures show that the segmentation strategy has been successful in removing the high density clusters that appeared in the joint histogram and focusing the optimization on the target voxels.

To test the hypothesis: misregistered regions can be identified using robust segmentation of voxels in the intensity difference image, the median landmark errors were calculated for the landmarks that either belong to inlier (group one) or outliers (Groups two and three). The median is robust to influence of outliers and presents a rigorous performance measure. The median landmark errors, presented in Table VII, show that landmark errors of group one are significantly lower than those of the group two and three.

Registration results for images in the DIR-dataset are reported at www.dir-lab.com and according to these results the best registration is achieved with Least median of square Filtered Compressible flow (LFC) [32] and 4 Dimensional Local Trajectory Modeling (4DLTM) [33]. Comparison of those results shows that the above methods achieved comparable or better registration accuracies to the proposed methods. It should however be noted that the errors of LFC and 4DLTM cannot be directly compared with the errors we computed for RIS and RRS, as those reported in the paper were on 1200 landmarks per image while a subset of 300 landmarks per image was available publicly. It is important to mention that the 4DLTM method uses all the images from the 4D-

TABLE VII
MEDIAN OF THE LANDMARK ERRORS BELONGING TO GROUP ONE AND GROUPS TWO PLUS THREE.

	Median Inlier Error	**Median Outlier Error**
DIR-C1	1.94	2.68
DIR-C2	2.76	4.12
DIR-C3	2.98	3.45
DIR-C4	3.39	5.13
DIR-C5	3.33	5.18
DIR-C6	4.00	13.12
DIR-C7	5.01	12.50
DIR-C8	6.40	7.93
DIR-C9	3.71	5.09
DIR-C10	2.85	4.62

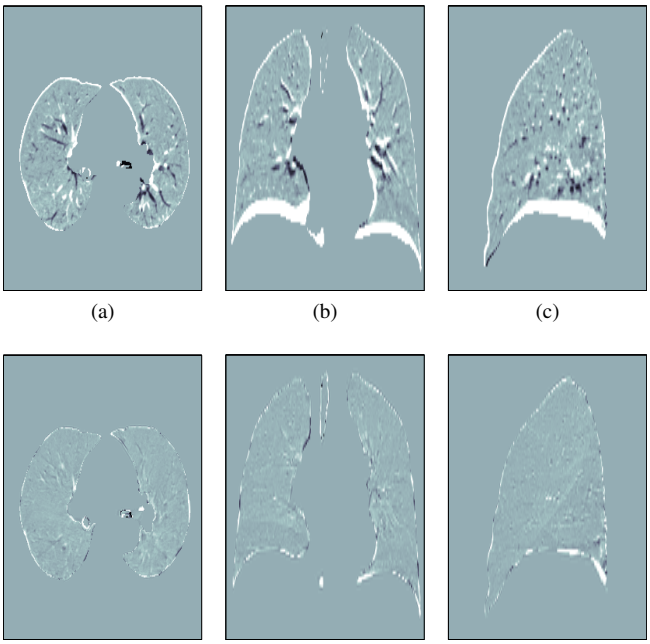


Fig. 6. Axial, Sagittal and Coronal slices of the intensity difference image between fixed and moving images (DIR-Case 5) before registration are shown in (a),(b),(c), respectively. (d),(e) and (f) show the Axial, Sagittal and Coronal slices of intensity differences between fixed and transformed moving images after RIS registration.

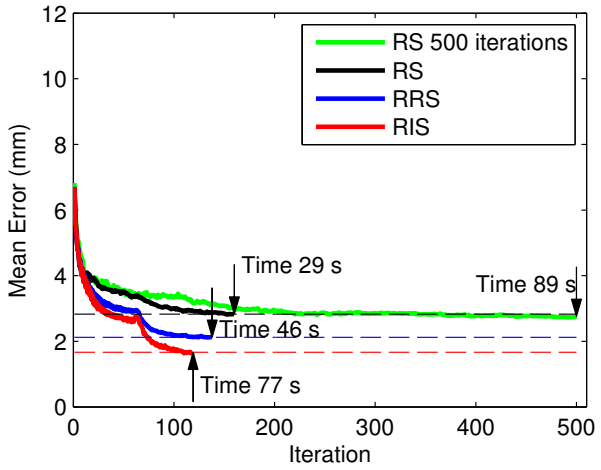


Fig. 7. Mean error of each iteration for algorithms RIS, RRS and RS. The maximum number of iterations for RS has been increased to 100 and 400 in the two registration levels respectively. The image used is DIR-Case 5.

CT dataset whereas the proposed methods only use the two images corresponding to the extremes of the respiratory cycle. The use of intermediate images by the 4DLTM method means that the algorithm has to process significantly more data than methods using only two extreme images. The execution time of the LFC method (reported in [32]) is an indication of the large amount of required computation in this method. As such, the above methods are specifically developed to register lung CT images and use more information than proposed methods making them computationally expensive.

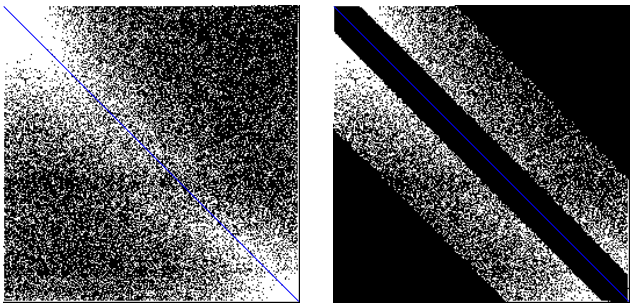


Fig. 8. The joint histogram of the fixed image ($I_f(x)$) and the transformed moving image ($I_m(T(x))$) for case 5 of DIR-dataset at the 10^{th} iteration (a) contains all the voxels. (b) contains only the voxels identified as the middle group by the MSSE.

A drawback of the proposed methods is that, since those methods rely on the intensity difference for sample selection, the change of intensity with a change in inspiration level in case of large deformations can affect the performance of the these methods. Other authors have shown benefit of using a mass preserving intensity model that adjusts intensity according to volume changes [34]. Such a model can be incorporated in the proposed approach and may still improve results over those presented here. Extension to multimodal registration is however less straightforward. In such cases one would need to find appropriate clustering schemes to remove the almost registered data from the images of different modality for the application of the proposed registration methods.

Several recent registration algorithms have been implemented using Graphics processing units (GPUs) to accelerate the computation of those methods [32], [35]–[37]. An algorithm has to be parallelizable in order to take advantage of these computers. In our implementation we use a B-spline transformation to represent the deformation field and the optimization is performed using analytical gradients of the SSD cost function with respect to the B-spline coefficients. In our approach the computation time is reduced by using only a carefully selected subset of voxels in the calculations. Another method to decrease the computation time would be to use a GPU to calculate the gradients. This is because the B-spline coefficients has only a compact support and the memory on the GPU is shared by all the cores, therefore the derivative calculation step can be easily parallelized as illustrated by Rohlfing et al. [38]. An issue for the GPU implementation of this method is that the sorting step of the proposed methods cannot be easily parallelized. Having said this, we have observed that the sorting step on average takes only around 15% of the overall computation time even when the basic "qsort" algorithm was used for sorting.

VI. Conclusion

The paper presented a new approach for non-rigid registration of medical images based on using robust segmentation of squared intensity differences to intelligently guide the sampling part of the stochastic optimization. The performance of the proposed approach in terms of registration accuracy and

computation time was compared with several existing methods via an extensive set of experiments on images with ground truth. The experiments showed that the proposed method substantially improves both the accuracy and the computational complexity of the registration task. More importantly, the method incorporated a computationally efficient means of measuring the quality of registration. Since the registration schemes are all iterative, this measure proved to be highly useful for deciding when to stop the registration process.

ACKNOWLEDGEMENT

The authors would like to thank Martha Coselmon, James Balter, and Marc Kessler for providing the Images for registration and Gustavo Rohde for supplying the code for the AR method. This research was partly supported under Australian Research Council's Discovery Projects funding scheme (DP130102524).

REFERENCES

[1] M. Holden, "A review of geometric transformations for nonrigid body registration," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 1, pp. 111–128, 2008.

[2] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "3d brain mapping using a deformable neuroanatomy," *Physics in Medicine and Biology*, vol. 39, no. 3, p. 609, 1994.

[3] J. P. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[4] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images," *Medical Imaging, IEEE Transactions on*, vol. 18, no. 8, pp. 712–721, 1999.

[5] M. D. Buhmann, *Radial basis functions: theory and implementations*. Cambridge university press, 2003, vol. 12.

[6] Z. Wu, "Compactly supported positive definite radial functions," *Advances in Computational Mathematics*, vol. 4, no. 1, pp. 283–292, 1995.

[7] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[8] H. Lester and S. R. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognition*, vol. 32, no. 1, pp. 129–149, 1999.

[9] K. Murphy et al., "Evaluation of registration methods on thoracic ct: The empire10 challenge," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 11, pp. 1901–1920, 2011.

[10] S. Damas, O. Cordón, and J. Santamaría, "Medical image registration using evolutionary computation: An experimental survey," *Computational Intelligence Magazine, IEEE*, vol. 6, no. 4, pp. 26–42, 2011.

[11] S. Klein, M. Staring, and J. P. W. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines," *Image Processing, IEEE Transactions on*, vol. 16, no. 12, pp. 2879–2890, 2007.

[12] R. Bhagalia, J. A. Fessler, and K. Boklye, "Accelerated nonrigid intensity-based image registration using importance sampling," *Medical Imaging, IEEE Transactions on*, vol. 28, no. 8, pp. 1208–1216, 2009.

[13] H. Robbins and S. Monro, "A stochastic approximation method," *The Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, 1951.

[14] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. Pluim, "elastix: A toolbox for intensity-based medical image registration," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 1, pp. 196–205, 2010.

[15] G. K. Rohde, A. Aldroubi, and B. M. Dawant, "The adaptive bases algorithm for intensity-based nonrigid image registration," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 11, pp. 1470–1479, 2003.

[16] A. Bab-Hadiashar and D. Suter, "Robust segmentation of visual data using ranked unbiased scale estimate," *Robotica*, vol. 17, no. 6, pp. 649–660, 1999.

[17] R. B. Tennakoon, A. Bab-Hadiashar, M. Bruijne, and Z. Cao, "Efficient nonrigid registration using ranked order statistics," in *IEEE 10th International Symposium on Biomedical Imaging*, 2013, pp. 492–495.

[18] R. Hoseinnezhad, A. Bab-Hadiashar, and D. Suter, "Finite sample bias of robust estimators in segmentation of closely spaced structures: A comparative study," *Journal of Mathematical Imaging and Vision*, vol. 37, no. 1, pp. 66–84, 2010.

[19] R. Hoseinnezhad and A. Bab-Hadiashar, "Consistency of robust estimators in multi-structural visual data segmentation," *Pattern Recognition*, vol. 40, no. 12, pp. 3677–3690, 2007.

[20] L. Devroye, *Non-uniform random variate generation*. Springer-Verlag, 1986.

[21] M. M. Coselmon, J. M. Balter, D. L. McShan, and M. L. Kessler, "Mutual information based ct registration of the lung at exhale and inhale breathing states using thin-plate splines," *Medical Physics*, vol. 31, no. 11, pp. 2942–2948, 2004.

[22] R. Castillo and et al., "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Physics in Medicine and Biology*, vol. 54, no. 7, p. 1849, 2009.

[23] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3d active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[24] J. Kybic and M. Unser, "Fast parametric elastic image registration," *Image Processing, IEEE Transactions on*, vol. 12, no. 11, pp. 1427–1442, 2003.

[25] M. Unser, "Splines: a perfect fit for signal and image processing," *Signal Processing Magazine, IEEE*, vol. 16, no. 6, pp. 22–38, 1999.

[26] J. C. Spall, "Implementation of the simultaneous perturbation algorithm for stochastic optimization," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 34, no. 3, pp. 817–823, 1998.

[27] H. Kesten, "Accelerated stochastic approximation," *The Annals of Mathematical Statistics*, vol. 29, no. 1, pp. 41–59, 1958.

[28] P. Dupuis and R. Simha, "On sampling controlled stochastic approximation," *Automatic Control, IEEE Transactions on*, vol. 36, no. 8, pp. 915–924, 1991.

[29] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *The Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.

[30] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

[31] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian journal of statistics*, pp. 65–70, 1979.

[32] E. Castillo, R. Castillo, B. White, J. Rojo, and T. Guerrero, "Least median of squares filtering of locally optimal point matches for compressible flow image registration," *Physics in Medicine and Biology*, vol. 57, no. 15, pp. 4827–4833, 2012.

[33] E. Castillo, R. Castillo, J. Martinez, M. Shenoy, and T. Guerrero, "Four-dimensional deformable image registration using trajectory modeling," *Physics in Medicine and Biology*, vol. 55, no. 1, p. 305, 2010.

[34] V. Gorbunova, J. Sporring, P. Lo, M. Loeve, H. A. Tiddens, M. Nielsen, A. Dirksen, and M. de Bruijne, "Mass preserving image registration for lung ct," *Medical Image Analysis*, vol. 16, no. 4, pp. 786–795, 2012.

[35] X. Gu, H. Pan, Y. Liang, R. Castillo, D. Yang, D. Choi, E. Castillo, A. Majumdar, T. Guerrero, and S. B. Jiang, "Implementation and evaluation of various demons deformable image registration algorithms on a gpu," *Physics in Medicine and Biology*, vol. 55, no. 1, p. 207, 2010.

[36] A. Eklund, P. Dufort, D. Forsberg, and S. M. LaConte, "Medical image processing on the gpu-past, present and future," *Medical Image Analysis*, 2013.

[37] O. Fluck, C. Vetter, W. Wein, A. Kamen, B. Preim, and R. Westermann, "A survey of medical image registration on graphics hardware," *Computer methods and programs in biomedicine*, vol. 104, no. 3, pp. e45–e57, 2011.

[38] T. Rohlfing and C. R. Maurer Jr, "Nonrigid image registration in shared-memory multiprocessor environments with application to brains, breasts, and bees," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 7, no. 1, pp. 16–25, 2003.