A Novel Subsampling Method for 3D Multimodality Medical Image Registration Based on Mutual Information

Maryam Zibaeifard[i]; Mohammad Rahmati[ii]

ABSTRACT

Mutual information (MI) is a widely used similarity metric for multimodality image registration. However, it involves an extremely high computational time especially when it is applied to volume images. Moreover, its robustness is affected by existence of local maxima. The multi-resolution pyramid approaches have been proposed to speed up the registration process and increase the accuracy of the result. In this paper, we present a new improved method of sample selection for multi-stage registration based on mutual information. Instead of down-sampling of the whole image as it is done in the pyramid methods, we propose a new technique to find a suitable subset of image samples based on image information content, which results in a better estimate of the optimal transformation. A comparison for MR images indicates that our proposed method yields better registration than subsampling method, especially when subsampling factor is low. The experimental results involving three-dimensional clinical images of CT, MR and PET are presented for rigid registration.

KEYWORDS

Image registration, Mutual information, Simulated annealing.

1. INTRODUCTION

Image registration is the process of spatial alignment of two or more images acquired from different sensors, viewpoints or time intervals [1]. Registration is widely used in medical imaging applications. A common practice of these applications could be found in fusion of multimodality images when patients have to undergo epilepsy surgery [2]. Registration and fusion of MR and PET images will benefit the surgeon. Besides multimodality registration, there exist important application areas in monomodality registration. Examples include alignment of two images given in different times in order to detect silent differences. Over the years, a large number of registration techniques have been developed including point-based methods, deformation models [3] and voxel-based methods. In voxel-based approaches, the optimum registration is determined by iteratively optimizing a similarity measure calculated using gray values of both images. Such methods do not require user interaction but may suffer from high computational cost. When mutual information is selected as the similarity measure, estimating the joint histogram which is a time-consuming process is iteratively required. Another difficulty with this method is the possible existence of local maxima in MI function. If original misregistration is small, the algorithm converges to the global maximum more possibly. For this reason, coarse to fine multi-resolution strategies have been introduced. These approaches increase accuracy of the registration and speed up the optimization algorithm. In multi-resolution methods, the pyramid of images could be obtained using the wavelet transformation [4], averaging method or image subsampling [5]. When using image pyramids, no preference is considered for different regions of images. We propose that more samples should be selected from high entropy areas such as regions of tissue transition. Furthermore, mutual information criterion suffers from lack of spatial information. Using the proposed method, we try to compensate this shortage.

The most common optimizers used for voxel-based

This work was supported by Iran Telecommunication Research Center (ITRC) under Grant T-500-1826.

i Maryam Zibaeifard is with Department of Computer Engineering and Information Technology, Amirkabir University of Technology, Tehran, Iran (email: zibaeifard@ce.aut.ac.ir).

ii Mohammad Rahmati is with Department of Computer Engineering and Information Technology, Amirkabir University of Technology, Tehran, Iran (email: rahmati@aut.ac.ir), Corresponding author.

registration are Powell's routine and Simplex method [6] which both are non-stochastic. Probabilistic optimization routines such as genetic algorithms and simulated annealing are seldom used. The main reason behind unpopularity of stochastic search routines for MI function is the fact that the desired maximum corresponding to the optimal transformation may not be the global maximum of the search space. A limited part of search space, called capture range, leads to the desired solution and probabilistic search methods can move outside the capture range. We propose a multi-stage search method which applies Powell's routine in early stages and a simulated annealing scheme in the final stage when initial range of the search space is limited enough to be safe for any stochastic search routine.

In Sections 2, we first review image registration based on mutual information. Section 3 presents our proposed subsampling method and Section 4 describes the search technique that we applied. In Section 5, the experimental results for registration of CT to MR and PET to MR images are presented and finally Section 6 gives a summary of our method.

2. IMAGE REGISTRATION

The registration problem is to find the optimal transformation T^* which best aligns the images. For reference image I and floating image J, image registration can be defined as follows:

$$T^* = \arg\max_{T} \rho(I, T(J)) \quad (1)$$

where ρ refers to a similarity measure. Therefore, different registration methods can be derived from different similarity measures and different search strategies.

The most common transformations applied to register medical images are rigid and affine. An affine transformation includes transition, rotation, scaling and shearing where it maps parallel lines to parallel lines [7]. Rigid transformation is a special kind of affine transformation when only transitions and rotations are allowed. In rigid transformation, the objects retain their relative shape and size and it is generally used for brain images.

A. Mutual information criterion

Mutual information is an entropy-based measure representing the amount of information that one of the images gives about the other. Consequently, registration is assumed to be gained in the state of maximizing of mutual information [8]. Given reference image I and floating image J, mutual information of the two images can be defined using marginal entropies $H(I)$ and $H(J)$ and joint entropy $H(I, J)$ in overlapping parts of the images:

$$MI(I, J) = H(I) + H(J) - H(I, J). \quad (2)$$

Obviously, *MI* function is sensitive to the size and contents of the overlapping area of the two images. A normalized measure of mutual information which is less sensitive to the overlapping area is defined as [9]:

$$NMI(I, J) = \frac{H(I) + H(J)}{H(I, J)} \quad (3)$$

It can be shown that $0 \le NMI(I, J) \le 2$ because:

$$H(I, J) \ge H(I), H(I, J) \ge H(J) \quad (4)$$

from (4) we find an upper bound for the normalized measure:

$$NMI(I, J) = \frac{H(I) + H(J)}{H(I, J)} \le 2 \quad (5)$$

3. SAMPLE SELECTION METHOD

Optimizing *MI* function requires estimation of the joint histogram iteratively. Given reference image I and floating image J and for each transformation T belonging to the search space, the joint histogram $H(I(s), T(J(S)))$ is computed for overlapping region of the images, where $s \in S$ and S is the set of grid points of the reference image.

To speed up the computation process, a subset of S may be used, where subsampling factor $\alpha = 1/N$ means only one in N voxels is selected for computing the joint histogram. The complexity of computing *MI* function [10] varies linearly with number of selected samples. Consequently, when subsampling factor is small, an estimation of optimal transformation can be obtained relatively fast. The result can be used as initial value in finer level (with higher α).

A. Variant subsampling factors

The idea of the proposed method [11] is based on using higher subsampling factors for the regions which contain more information. The common subsampling method applies a fixed factor α for all areas of an image. In multi-resolution techniques, subsampling is done by averaging or other methods, but again α is the same for all regions, whereas tissue regions deserve more attention than background. Furthermore, edge or tissue transition regions contain important information for adjustment.

A method of incorporating spatial information of edges with mutual information that was introduced in [12] uses gradient vectors of corresponding points but calculation of gradient vectors in each iteration increases the computational cost. By using variant subsampling factors, we can emphasize the role of edge regions in *MI* measure.

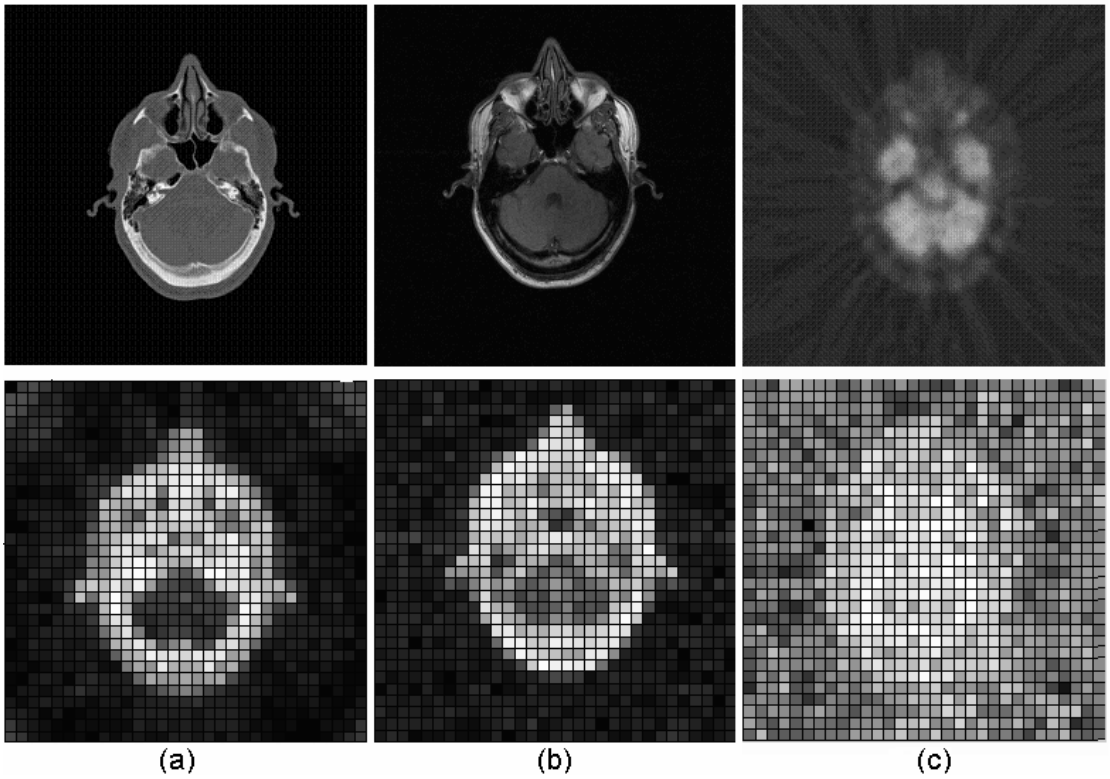

Figure 1: Entropy maps for (a) CT, (b) MR, and (c) PET images. Blocks with higher entropy are shown with higher intensity.

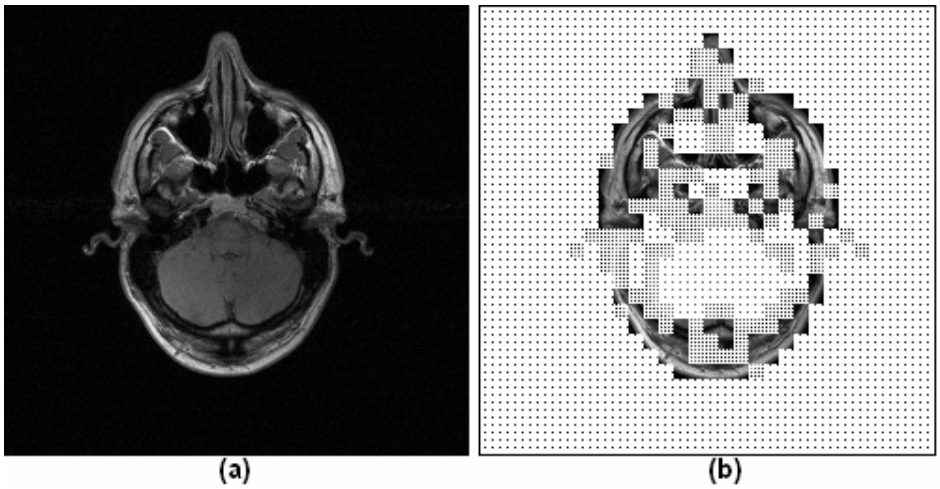

Figure 2: (a) A brain MR image, (b) Selected samples using proposed method where $n = 3$, $c = 4$, $\alpha_1 = 1/16$, $\alpha_2 = 1/4$, and $\alpha_3 = 1/1$.

In order to measure the information content of each partition of an image, entropy is used. In Fig. 1, the entropy maps of MR, CT and PET images which divided into small blocks are shown. The block size depends on the width of edge transition regions in the image. For brain images of Fig. 1 and 2, blocks of 32 × 32 were used. Clearly, for three-dimensional images, the entropy blocks should be three-dimensional. In general, to subsample an image with n variant factors, a series of block sets $A_1,...A_n$ in the reference image are defined as:

$$A_i = \{s \in B_j \mid \tau_{i-1} < H(B_j) \le \tau_i\}, \quad \begin{matrix} \tau_0 = 0 \\ i = 1,..., n \\ j = 1,..., M \end{matrix} \qquad (6)$$

where M is the number of blocks in the image, B_j is the jth block, $H(B_j)$ is entropy of the jth block and τ_i is entropy threshold for block set A_i which can determine

how many blocks belong to the block set. Assuming blocks B_j are sorted ascendingly with respect to their entries and $P(A_i)$ to be the ratio of number of blocks in block set A_i to M, the entropy threshold τ_i can be found as:

$$\tau_i = H(B_l), \qquad l = Round\left(\sum_{m=1}^{i} P(A_m)\, M\right) \tag{7}$$

For each block set A_i, a corresponding subsampling factor α_i is used which holds the condition $\tau_p > \tau_q \rightarrow \alpha_p > \alpha_q$ that means the blocks with higher entropy should be subsampled with higher factors. A simple way is to consider a constant c as coefficient factor:

$$\alpha_{i+1} = c\,\alpha_i, \quad c\,\alpha_i \le 1 \tag{8}$$

Fig. 2 shows selected samples for a brain MR image. Eighty percent of samples belong to the first block set A_1 with subsampling factor α_1, the block set A_2 contains ten percent of samples with corresponding factor α_2 and the rest are members of the third block set A_3 with subsampling factor α_3.

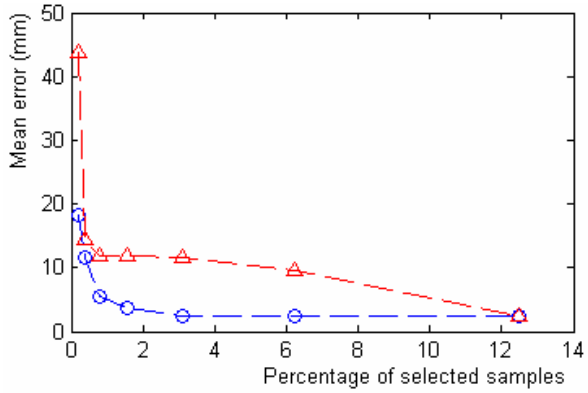

Figure 3: Circles and triangles indicate mean errors for proposed method and uniform subsampling method, respectively.

A comparison of proposed method and uniform subsampling method for MR to MR registration of 10 image pairs under simulated rigid transformations is shown in Fig. 3. The vertical axis shows the mean of registration errors computed in several volumes of interest (VOIs) and the horizontal axis represents the percentage of samples which selected by both methods. It is noted that due the voxel-size of images and the interpolation method, a minimum registration error is unavoidable. The comparison indicates that the proposed method reaches the minimum error at lower percentage of samples and yields a better registration than uniform subsampling method especially when subsampling factor is low. Fig. 4 represents *MI* functions for a transition and a rotation parameter, where zero position corresponds to the optimal solution. It shows that by using our proposed method, the MI curve artifacts and local maxima are reduced (i.e., the curve is smoother) in compare with the uniform subsampling method.

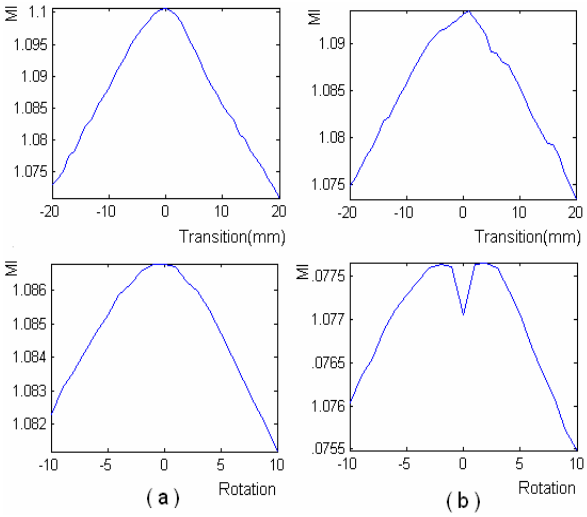

Figure 4: MI function representation under transitional and rotational distortion based on (a) proposed method and (b) uniform subsampling method.

4. SEARCH TECHNIQUE

Local maxima in MI function may result from low resolution, interpolation artifacts [13]-[14] or small overlapping area between the two images. An appropriate search technique is needed to deal with this problem. In non-stochastic optimization routines such as Powell's method, a maximum (or minimum) is found by decreasing the initial range of each parameter iteratively. The methods stop when encountering to a local maximum.

In contrast, stochastic methods are able to get out of local maxima, but for two reasons optimization of MI function through these methods are less popular. The first, stochastic routines usually require more iteration cycles to converge than non-stochastic methods. The more important reason is the possibility of moving out of the capture range and accepting a local or global maximum which is far from the desired solution as the best value. The graph shown in Fig. 5 represents the mutual information of a PET and an MR image as a function of one transition and one rotation parameters. For large values of parameters when overlapping area of the two images is too small, the MI function displays an abrupt behavior.

In the past, the multi-resolution approaches were introduced to improve the registration results. Those methods find a rough estimate of the optimal transformation using low resolution images, which is subsequently used as initial point for the higher resolutions. If the optimizer is trapped into a local

maximum in a coarse stage, it is likely that a local maximum is located at this point for the higher resolution stage too. As a result, the multi-stage method fails to improve the solution.

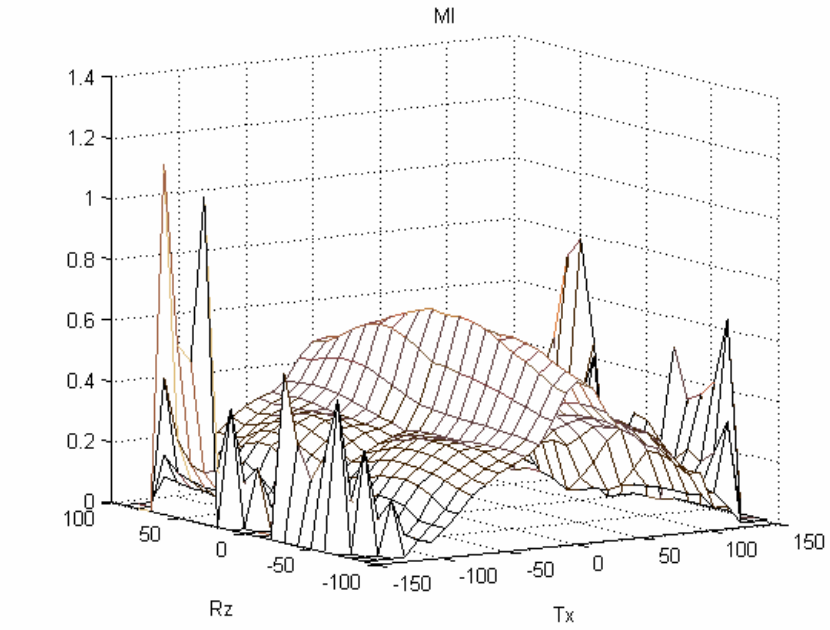

Figure 5: Mutual information of PET/MR as a function of one transition and one rotation parameters.

Because the boundaries of capture range for all parameters can not be determined exactly, it is necessary to reduce the search space before applying the simulated annealing method. Therefore, we propose using Powell's routine or other non-stochastic optimizers in primary stages and a simulated annealing scheme [15] in the last one or two stages.

A. Simulated annealing

The main concept in simulated annealing [16] is based on a phenomenon in thermodynamics which involves heating and controlled cooling of a material to increase the size of its crystals and reduce their defects. The atoms in their initial positions present a state of local minimum of the internal energy. The heat causes the atoms to become unstuck and wander randomly through states of higher energy; the slow cooling gives them more chances of finding configurations with lower internal energy than the initial one.

The Boltzmann's probability distribution describes this phenomenon:

$$prob(E) \approx \exp(E / c_b t) \tag{9}$$

where, E is the energy level, t is the system temperature and c_b is called Boltzmann's constant. The expression (9) describes how the system is able to get out of the local minima. We introduce a simulated annealing scheme based on Metropolis's procedure to optimize the normalized mutual information of images.

B. Random generation function

For optimizing normalized mutual information function $f_{NMI}(X)$, where $X = [x_1, x_2, ..., x_N]$ is a vector of N independent parameters of a spatial transformation, the first step is to define a function which generates random numbers to be used as jump values from current variable vector to a new test vector. Bell-shaped probability functions are preferred for optimizing the mutual information function [17] because the probability of small jumps around the current value is expected to be greater than the probability of large jumps. We applied a Gaussian generation function as follows:

$$G_k^i(d) = \frac{1}{\sqrt{2\pi t_k^i}} \exp\left(\frac{-d^2}{2t_k^i}\right) \tag{10}$$

here, $G_k^i(d)$ with mean value $\mu = 0$ and standard deviation $\sigma = \sqrt{t_k^i}$ is the probability for jump value d and t_k^i is the temperature variable for parameter x_i in iteration k.

In order to cover the initial range by the Gaussian function as in [17], the initial temperature for each parameter is selected as follows:

$$\begin{aligned} R_0^i &= 3\sigma = 3\sqrt{t_0^i} \\ t_0^i &= \left(R_0^i / 3\right)^2 \end{aligned} \qquad i = 1,2,..., N \tag{11}$$

where, R_0^i is the initial range and t_0^i is the initial temperature for parameter x_i. To ensure that the variance of the random generation function (10) never gets too

small, a minimum variance for each parameter is defined to be 1/4 of the initial temperature:

$$\sigma_{\min}^{i\ 2} = t_0^i / 4 . \qquad (12)$$

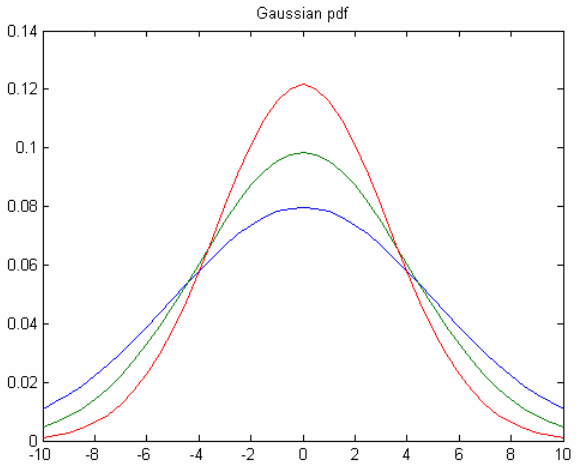

Figure 6: Gaussian generation function for a parameter with initial range $R_0 = 16$ in three consecutive iterations.

C. Cooling schedule

The cooling schedule expresses how the temperature variables are lowered during the iterations [18]. The temperature t_k^i in the k*th* iteration is determined by following:

$$t_k^i = (\lambda)^k t_0^i . \qquad (13)$$

In our proposed algorithm, we used a cooling factor of $\lambda = 0.9$ for all the variables.

D. Acceptance function

We made a trivial modification in Metropolis's procedure to find the maximum, instead of the minimum of energy state. Considering the change in energy level to be:

$$\Delta E = f_{NMI}(X_n) - f_{NMI}(X_c) \qquad (14)$$

where X_n and X_c are the values of new and current variable vectors, respectively, the probability of acceptance of a test value for parameter x_i is defined as:

$$pr_k^i(\Delta E) = \begin{cases} 1 & \Delta E \geq 0 \\ \exp\left(\dfrac{\Delta E}{c_b t_k^i}\right) & \Delta E < 0 \end{cases} \qquad (15)$$

Equation (15) expresses that even the worse values can be accepted with probability $pr_k^i(\Delta E)$. It is desired that the same ΔE made by different variables results in an identical probability of acceptance. In other words, the acceptance function should be independent of initial temperature for each variable. Furthermore, the probability of acceptance in primary iterations should be close to 1, even for $\Delta E < 0$. As mentioned in Section 2, the normalized mutual information of two images is a number between zero and 2 and $\Delta E = -0.1$ corresponds to a large distortion with respect to the current values of parameters. We consider $pr_0^i(-0.1) = 0.99$ for all variables, so the constant c_b is determined as:

$$pr_0^i(-0.1) = \exp\left(\frac{-0.1}{c_b t_o^i}\right) = 0.99 \Rightarrow c_b \approx \frac{10}{t_0^i} \qquad (16)$$

Given (13), (15) and (16), we define an acceptance function that is independent of the temperature variables:

$$pr_k^i(\Delta E) = \exp\left(\frac{\Delta E}{c_b t_k^i}\right) = \exp\left(\frac{\Delta E}{10(\lambda)^k}\right), \quad \Delta E < 0 \qquad (17)$$

The random generation function, the cooling schedule and the acceptance function for a variable with initial range $R_0 = 16$ are shown in Fig. 6, 7, and 8. Fig. 8 shows that the system reaches steady state about iteration 100 that means after iteration 100, a worse jump rarely is accepted. A pseudo code for the simulated annealing algorithm is presented in Fig. 9.

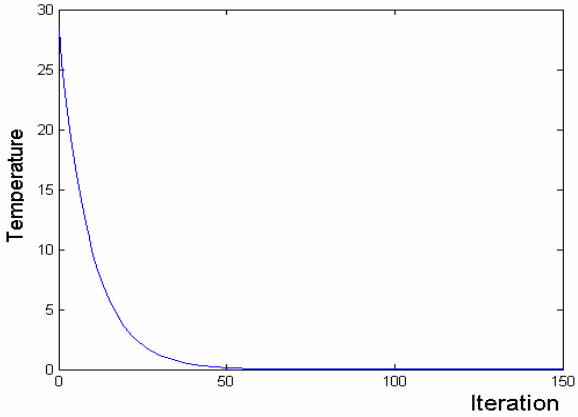

Figure 7: Cooling schedule for a parameter with initial range R_0= *16*.

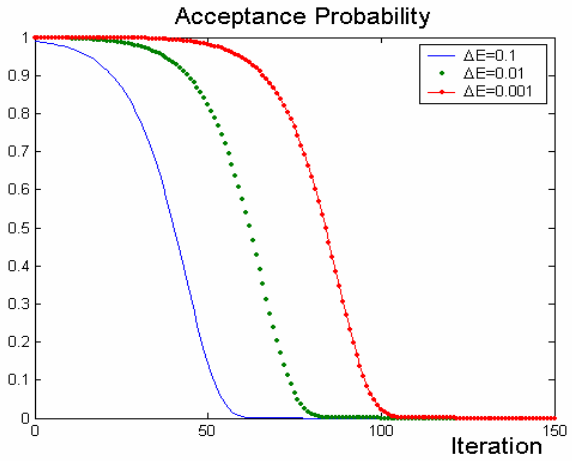

Figure 8: Acceptance function.

```
Initialize:
   Initial temperatures  t_0^i ,      i = 1,...,  N
   Initial ranges  R_0^i ,      i = 1,...,  N
   Cooling factor  λ
   Maximum numbers of iterations  max
   Iteration counter  k ← 0
While  k ≤ max
      For all variables  x_i
          Generate a test value inside the initial range
          If  ΔE ≥ 0
                Accept the test value
          Else
                Accept it with a probability calculated by acceptance function
          EndElse
       EndFor
       k ← k + 1
       For all variables  x_i
          Reduce temperature using the cooling schedule
        EndFor
EndWhile
```

Figure 9: Pseudo code for the simulated annealing algorithm.

5. EXPERIMENTAL RESULTS

The proposed method was applied to register CT to MR and PET to MR volume images of Retrospective Registration Evaluation Project (RREP). Image volumes of three modalities (CT, MR, and PET) were obtained from patients undergoing neurosurgery at Vanderbilt University Medical Center on whom bone-implanted fiducial markers were mounted. These volumes had all traces of the markers removed and were provided via the Internet to project collaborators outside Vanderbilt, who then performed retrospective registrations on the volumes, calculating transformations from CT to MR and/ or from PET to MR. These investigators communicated their transformations again via the Internet to Vanderbilt, where the accuracy of each registration was evaluated. (See [19] for more detail.)

MR images include T1, T2, PD and rectified versions of them. Geometrical distortions are corrected for rectified images. Typical voxel size of the images is (1.25 × 1.25 × 4) for MR images, (0.65 × 0.65 × 4) for CT images and (2.59 × 2.59 × 8) for PET images in mm.

We used a three-stage search approach and reduced the search space when starting a new stage. The optimizer was Powell's routine in the first and second stages and the adaptive simulated annealing in the final level. Tables 1 and 2 describe the proposed subsampling method used in each stage with $n = 2$ (two subsampling factors) and $c = 4$. The joint histograms were estimated using bilinear interpolation.

TABLE 1
PARAMETERS OF THE PROPOSED SUBSAMPLING METHOD IN EACH STAGE FOR CT TO MR PAIRS.

Stage	Joint Histogram bins	$P(A_1)$	$P(A_2)$	α_1	α_2
1	64 × 64	0.8	0.2	1/1024	1/256
2	128 × 128	0.8	0.2	1/256	1/64
3	256 × 256	0.8	0.2	1/64	1/16

TABLE 2
PARAMETERS OF THE PROPOSED SUBSAMPLING METHOD IN EACH STAGE FOR PET TO MR PAIRS.

Stage	Joint Histogram bins	$P(A_1)$	$P(A_2)$	α_1	α_2
1	32 × 32	075	0.25	1/256	1/64
2	64 ×64	0.75	0.25	1/64	1/16
3	256 × 256	0.75	0.25	1/16	1/4

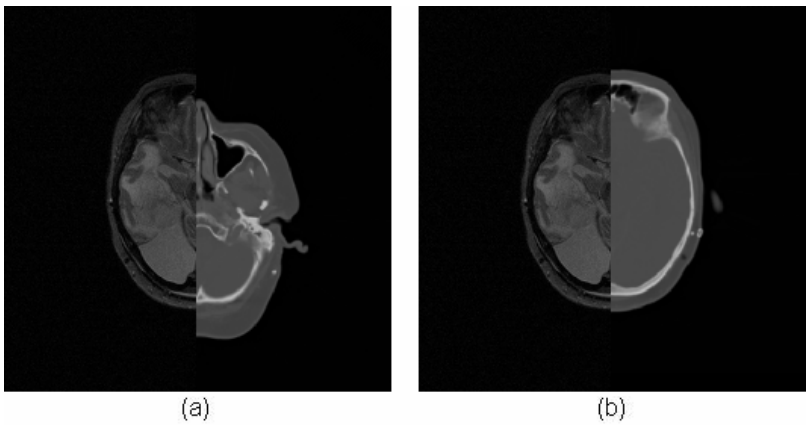

Figure 10: Fused slices of CT/MR images (a) before registration, (b) after registration.

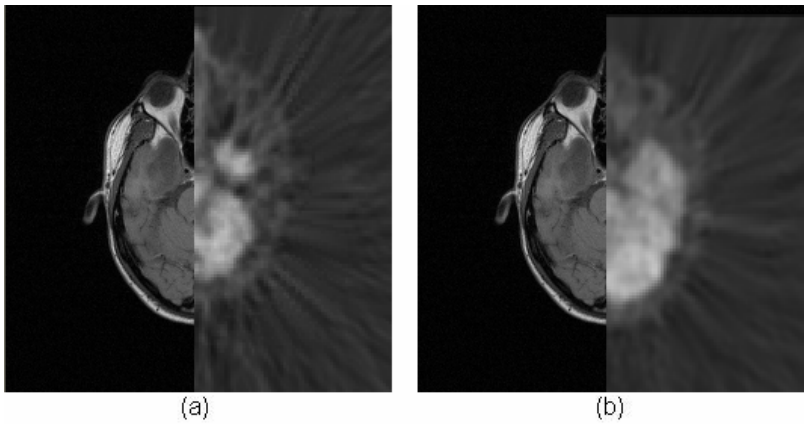

Figure 11: Fused slices of PET/MR images (a) before registration, (b) after registration.

The difference between marker-based gold standard and our registration has been evaluated in 10 VOIs. The results are summarized in Tables 3 and 4. The mean errors are less than the largest voxel size (4 mm for CT and MR and 8 mm for PET), which means that all mean errors represent subvoxel accuracy. Fig. 10 and 11 visualize our registration results.

TABLE 3
REGISTRATION ERRORS (IN MM) WITH RESPECT TO THE GOLD STANDARD FOR CT TO MR PAIRS.

Modality	Number of datasets	Median	Mean	Maximum
CT-T1	7	1.4604	1.5638	3.6254
CT-T1rect	6	0.6144	0.9798	5.0428
CT-T2	7	2.1742	2.1251	3.3467
CT-T2rect	7	0.9596	1.2045	5.3249
CT-PD	7	2.2918	2.4021	5.0634
CT-PDrect	7	0.7153	0.9664	3.4570

TABLE 4
REGISTRATION ERRORS (IN MM) WITH RESPECT TO THE GOLD STANDARD FOR PET TO MR PAIRS.

Modality	Number of datasets	Median	Mean	Maximum
PET-T1	5	3.1852	4.3604	9.7388
PET-T1rect	4	2.6150	2.7462	4.5467
PET-T2	4	3.3800	3.4876	6.5244
PET-T2rect	4	2.9284	3.5646	7.8524
PET-PD	4	3.6630	4.4628	9.0148
PET-PDrect	4	2.9474	4.0269	9.8907

6. CONCLUSION

An improved multi-stage registration method based on mutual information has been introduced. The method employs a new subsampling technique which selects more relevant samples for joint histogram estimation. Because of existence of local maxima, simulated annealing is used

in the final stage to increase the chance of finding the optimal solution.

The method is tested on volume image pairs of CT/MR and PET/MR. The registration errors are evaluated with respect to the marker-based gold standard. The results represent subvoxel accuracy for all mean errors.

7. ACKNOWLEDGEMENTS

The images and gold standard transformations were provided as part of the project, "Retrospective Image Registration Evaluation", National Institutes of Health, Project Number 8R01EB002124-03, Principal Investigator, J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN, USA.

This work was supported by Iran Telecommunication Research Center (ITRC), under grant T-500-1826.

8. REFERENCES

[1] Lisa G. Brown, "A survey of image registration techniques," ACM Computing Surveys, vol. 24, no. 4, pp. 325-376, 1992.

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

[3] V. Noblet, C. Heinrich, F. Heitz, J. P. Armspach, "3-d deformable image registration: a topology preservation scheme based on hierarchical deformation models and interval analysis optimization," IEEE Transactions on Image Processing, vol. 14, no. 5, pp. 553-566, 2005.

[4] Hua mei Chen, Pramod K. Varshney, "A pyramid approach for multimodality image registration based on mutual information," Proceedings of 3rd international conference on information fusion, vol. I, pp. 9-15, 2000.

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

[6] S. A. Teukolsky, B. P. Flannery, W. T. Vetterling, *Numerical recipes in C*, Cambridge University Press, Cambridge, UK, 1992.

[7] R. Shekhar, V. Zagrodsky, "Mutual information-based rigid and nonrigid registration of ultrasound volumes," IEEE Transactions on Medical Imaging, vol. 21, no. 1, pp. 9-22, 2002.

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

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

[10] F. Maes, A. Collignon, "Multimodality image registration by maximization of mutual information," IEEE Transactions on Medical Imaging, vol. 16, no. 2, pp. 187-198, 1997.

[11] Zibaeifard Maryam, Rahmati Mohammad, "An improved multi-stage method for medical image registration based on mutual information," Proceedings of 17th British Machine Vision conference, vol. 2, pp. 479–487, Edinburgh, UK, Sep. 2006.

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

[13] Josien P. W. Pluim, J. B. Antoine Maintz, Max A. Viergever, "Interpolation artifacts in mutual information-based image registration," Computer Vision and Image Understanding. vol. 77, pp. 211–232, 2000.

[14] Jeffrey Tsao, "Interpolation artifacts in multimodality image registration based on maximization of mutual information," IEEE Transactions on Medical Imaging, vol. 22, no. 7, pp. 854-863, 2003.

[15] Zibaeifard Maryam, Rahmati Mohammad, "An adaptive simulated annealing scheme for medical image registration by maximization of mutual information," *Proceedings of 8th International Conference on Signal Processing*, pp. 1157-1160, Guilin, China, Nov. 2006.

[16] K. A. Dowsland, Simulated annealing. In C. Reeves (ed), "Modern Heuristic Techniques for Combinatorial Problems," Wiley, pp. 20–69, New York, 1993.

[17] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, P. P. Van Saarloos, "Registration of stereo and temporal images of the retina," IEEE Transactions on Medical Imaging, vol. 18, no. 5, pp. 404-418, 1999.

[18] Ilias Maglogiannis, Elias P. Zafiropoulos, "Automated medical image registration using the aimulated annealing algorithm," Proceedings of Methods and Applications of Artificial Intelligence, Third Hellenic Conference on AI, pp. 456–465, 2004.

[19] J. West, J. M. Fitzpatrick, "Comparison and evaluation of retrospective intermodality brain image registration techniques," Journal of Computer Assisted Tomography, vol. 21, no. 4, pp. 554-566, 1997.