International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:2, No:1, 2008

Medical Image Registration by Minimizing Divergence Measure Based on Tsallis Entropy

Shaoyan Sun, Liwei Zhang, and Chonghui Guo

***Abstract*—**As the use of registration packages spreads, the number of the aligned image pairs in image databases (either by manual or automatic methods) increases dramatically. These image pairs can serve as a set of training data. Correspondingly, the images that are to be registered serve as testing data.

In this paper, a novel medical image registration method is proposed which is based on the *a priori* knowledge of the expected joint intensity distribution estimated from pre-aligned training images. The goal of the registration is to find the optimal transformation such that the distance between the observed joint intensity distribution obtained from the testing image pair and the expected joint intensity distribution obtained from the corresponding training image pair is minimized. The distance is measured using the divergence measure based on Tsallis entropy. Experimental results show that, compared with the widely-used Shannon mutual information as well as Tsallis mutual information, the proposed method is computationally more efficient without sacrificing registration accuracy.

***Keywords*—**Multimodality images, image registration, Shannon entropy, Tsallis entropy, mutual information, Powell optimization.

I. INTRODUCTION

THE geometric alignment or registration of images is a fundamental task in numerous applications in three-dimensional medical image processing. Medical images from different modalities can provide complementary information about anatomy or functional in a synergistic manner. Since images may be acquired in different poses, considerable effort has been placed on developing methods for image registration. Existing image registration techniques can be broadly classified into two categories: feature-based and intensity-based methods [1,2]. A feature-based method requires the exaction of features common in both images. Obviously, a feature-based method is data dependent. Since different image data may have different features, the feature exaction algorithms adopted in a feature-based image registration algorithm are expected to be different for different registration tasks. In contrast, intensity-based image registration techniques are free from this limitation because they do not deal with the identification of geometrical landmarks.

Recently there has been an active research into the use of intensity-based image registration techniques [3, 4]. The key step of intensity-based image registration is to find a spatial transformation such that a similarity metric (or distance metric) between two or more images taken at different times, from different sensors or from different viewpoints, achieves its maximum (or minimum). The use of image similarity measures, especially those derived from information theory, have been shown to allow fully automated registration in a number of important clinical applications. The Shannon mutual information (Shannon-MI) which is based on the Kullback-Leibler divergence (KLD) has recently received substantial attention [5] for it is a robust and accurate registration method. In addition to the widely-used Shannon-MI, generalized information-theoretic similarity metrics, such as Rényi entropy [6, 7] and Tsallis mutual information (Tsallis-MI) [8, 9, 10], have properties that make them conductive to medical image registration.

On the other hand, as the use of co-registration packages spreads, the number of the aligned images pairs in image databases (either by the manual or automatic methods) increases dramatically. These image pairs can serve as a set of training data, in which the statistical joint intensity properties can be observed and learned in order to acquire useful *a priori* knowledge for future registration tasks. Chung *et al.* [11] and Gan *et al.* [12] proposed to make use of the *a priori* knowledge of the joint intensity distribution expected to obtain as a reference distribution. This expected joint distribution can be estimated from aligned training images. Any two images of the same or different acquisitions are aligned when the distance from their observed joint distribution to the corresponding expected distribution is minimized.

In this paper, we propose to measure the distance using the Tsallis divergence measure (denoted as TDM). Unlike KLD, TDM is non-logarithmic and has an adaptive parameter. The performance of the proposed method has been investigated by applying it to 3D brain image registration problems both on simulated and clinical images. Experimental results show that the proposed TDM registration algorithm can provide an accurate and rapid registration.

Manuscript received January 23, 2007. This research is supported by Natural Science Foundation of China (Grant No. 10571018).

Shaoyan Sun is with the Department of Applied Mathematics, Dalian University of Technology, Dalian, 116024, People's Republic of China (corresponding author, e-mail: sunshaoyan1979@yahoo.com.cn). She is doing research on image registration.

Liwei Zhang is with the Department of Applied Mathematics, Dalian University of Technology, Dalian, 116024, People's Republic of China (e-mail: lwzhang@dlut.edu.cn).

Chonghui Guo is with the Department of Applied Mathematics, Dalian University of Technology, Dalian, 116024, People's Republic of China (e-mail: guochonghui@tsinghua.org.cn).

II. Registration Algorithm

A. The Spatial Coordinate and Transformation

Each image is associated an image coordinate frame with origin positioned in a corner of the image, with x axis along the row direction, the y axis along the column direction, and the z axis along the plane direction [13].

One of the images is selected to be the floating image, f, from which samples s are taken and transformed by the transformation T_θ into the reference image, r. In general, $T_\theta s$ will not coincide with a grid point of image r and interpolation of the reference image is needed to obtain the image intensity value $r(T_\theta s)$. Nearest neighbor interpolation of r is generally insufficient to guarantee subvoxel accuracy, as it is insensitive to translations up to one voxel. Other interpolation methods, such as trilinear interpolation, may introduce new intensity values which are originally not present in the reference image, leading to unpredictable changes in the marginal distribution of the reference image foe small variations of θ. To avoid this problem, F. Maes proposed a partial volume interpolation (PVI) to update the joint histogram for each voxel pairs $(s, T_\theta s)$. Instead of interpolating new intensity values in image r, the contribution of the image intensity $f(s)$ of the sample s of f to the joint histogram distributed over the intensity values of all eight nearest neighbor's of $T_\theta s$ on the grid of r, using the same weights as for trilinear interpolation [12].

In this paper we restrict the transformation T_θ to rigid-body transformation. The rigid-body transformation is a superposition of a 3-D rotation and a 3-D translation and the registration parameter θ is a six-component vector consisting of three rotation angles ϕ_x, ϕ_y, ϕ_z (measured in degrees) and three translation distances t_x, t_y, t_z (measured in millimeters).

Transformation of image coordinates P_f to P_r from the image f to image r is given by

$$V_r \cdot (\mathrm{P}_r - C_r) = R_x(\phi_x) \cdot R_y(\phi_y) \cdot R_z(\phi_z) \cdot V_f \cdot (\mathrm{P}_f - C_f) + t(t_x, t_y . t_z) \qquad (1)$$

with V_f and V_r being 3×3 diagonal matrixes representing the voxel sizes of images f and r, respectively (in millimeters), C_f and C_r the image coordinates of the centers of the images, $R = R_x \cdot R_y \cdot R_z$ the 3×3 rotation matrix, with the matrixes R_x, R_y, and R_z representing rotations around the x, y, and z axis, respectively, and t the translation vector [13].

B. The Observed and Expected Joint Intensity Distributions

Assuming I_f and I_r are the intensity values of two testing images of the same or different acquisitions, where f and r respectively represent the floating and the reference images that are to be registered.

Given a hypothesized transformation T_θ defined by the registration parameter θ, then we can obtain their joint intensity distribution denoted as observed joint intensity distribution $P^{T_\theta}(I_f, I_r)$ [11, 12]. Note that $P^{T_\theta}(I_f, I_r)$ is a function of the transformation T_θ.

The corresponding expected joint distribution $P^*(I_{f^*}, I_{r^*})$ used as the reference distribution can be estimated from aligned training images f^* and r^*. Here images f^* and f should be the same acquisition, so should images r^* and r.

The joint probability distribution of two images is estimated by calculating a normalized joint histogram of the grey values. The marginal distributions are obtained by summing over the rows, resp. the columns, of the joint histogram. We use partial volume interpolation (PVI) to construct the joint histogram since it provides subvoxel accuracy (as opposed to nearest neighbor interpolation) and avoids the introduction of spurious grey values (in contrast to trilinear interpolation). The histogram size we use here is 256×256 which indicates that the images have been pre-rescaled into 256 grey levels.

The goal is to align two testing images of the same or different acquisitions such that the expected distribution and the observed joint intensity distribution are well matched. In other words, the registration algorithm aligns two different images based on the expected outcomes.

C. The Distance Measured by the Tsallis Divergence Measure

In this paper, the distance between two distributions P and Q is measured by the Tsallis divergence measure (TDM), which is proposed by Tsallis [9]:

$$D_{TDM}(P \| Q) = \frac{1}{1-\alpha}\left(1 - \sum_i \frac{p_i^\alpha}{q_i^{\alpha-1}}\right) \qquad (2)$$

with $\alpha \in \Re - \{1\}$, where p_i and q_i denote the probability distributions associated with the distributions P and Q. Here distributions P and Q represent the observed and expected joint distributions $P^{T_\theta}(I_f, I_r)$ and $P^*(I_{f^*}, I_{r^*})$ respectively.

The TDM value, D_{TDM}, tends to zero when the two distributions become equal. So the registration goal is to search the optimal transformation T_{θ^*} which minimizes the distance from the observed distribution to the expected joint distribution, i.e. D_{TDM} . The registration procedure is an iterative process, and is terminated when D_{TDM} becomes sufficiently small.

D. Mutual Information Based Registration Measures

To investigate the performance of the proposed TDM method, we compare it with other two registration method. The first is Shannon mutual information (Shannon-MI) method, which was proposed independently by Collignon *et al.* [14] and Viola *et al.* [15] in 1995. It is a very general and powerful criterion, and has been accepted by most researchers. So we would compare the new method with this classical method. The second method we considere here is Tsallis mutual information (Tsallis-MI), which is based on Tsallis entropy. It was applied into solving registration problems recently.

Shannon mutual information is defined as follows:

$$Shannon-MI = H(I_f) + H(I_r) - H(I_f, I_r) \ , \quad (3)$$

where

$$H(I_f) = -\sum_{i_f} p(i_f) \log p(i_f) \quad (4)$$

is the Shannon entropy.
The expression of Tsallis-MI is as follows:

$$\begin{aligned} Tsallis - MI &= H^{\alpha}(I_f) + H^{\alpha}(I_r) \\ &- (1-\alpha) \times H^{\alpha}(I_f) H^{\alpha}(I_r) - H^{\alpha}(I_f, I_r) \end{aligned} , \quad (5)$$

where

$$H^{\alpha}(I_f) = (1-\alpha)^{-1} \left(\sum_{i_f} p(i_f)^{\alpha} - 1\right), \ \alpha \in \Re - \{1\} \quad (6)$$

is the Tsallis entropy.

The mutual information based methods state that, for two images that are to be registered, the value of their mutual information will be maximal if the images are geometrically aligned.

E. Optimization Method

The images are initially positioned such that their centers coincide and that the corresponding scan axes of both images are aligned and have the same orientation. Powell's method is used to minimize D_{TDM} [16]. This method is a reasonable method between robustness and speed.

Powell optimization requires no derivative information of the objective function. It involves a series of one-dimensional maximizations for each dimension; these maximizations are carried out by Brent's method. Having found an optimum in one direction, the maximization is continued in the next direction, starting from the current position. Once all six parameters have been optimized, the loop is repeated until the

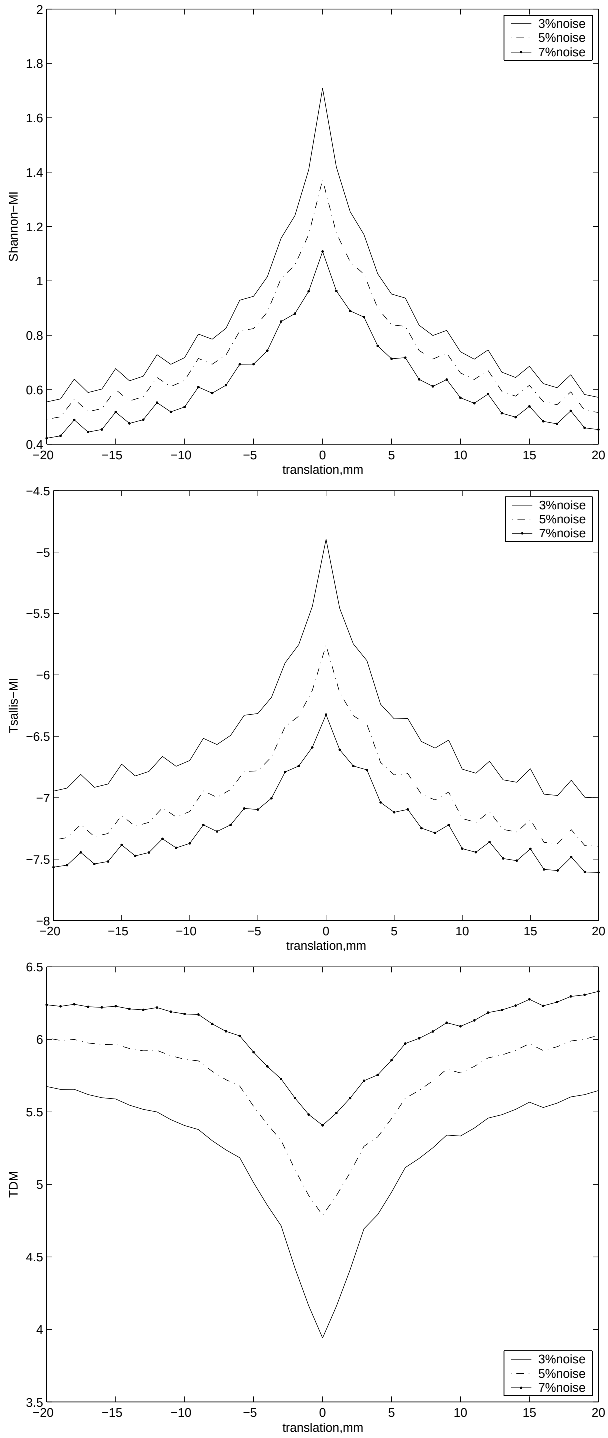

Fig. 1 Examples of three different registration measures as functions of translations along the slice direction (in millimeters) for three testing image pairs

TABLE I
SIMULATED REGISTRATION RESULTS OF THE IMAGE PAIR WITH 3% NOISE LEVEL

Methods	α	Results	Function Evaluations
Shannon-MI	None	10^{-3} ×[-0.1087, -0.1035, 0.0798, -0.0564, -0.1401, 0.0266]	704
Tsallis-MI	0.9	10^{+6} ×[-0.0125, 0.0171, 0.0001, 0.0027, -0.0022, 5.2893]	200 (failed)
Tsallis-MI	0.8	10^{+4}×[0.0768, -4.6148, 0.0005, -0.0000, 0.0001, 0.0768]	194 (failed)
TDM	0.9	10^{-3} ×[0.2429, 0.2267, 0.0424, -0.0981, -0.0079, 0.0226]	640
TDM	0.8	[0.0002, -0.0012, 0.0000, 0.0002, 0.0001, -0.0000]	593
TDM	0.7	[-0.0027, -0.0240, -0.0003, 0.0046, -0.0021, 0.0030]	504
TDM	0.6	[0.0234, -0.0769, 0.0090, 0.0095, 0.0177, 0.0103]	417
TDM	0.5	[0.0833, 0.1380, -0.0165, -0.0187, 0.0371, -0.0003]	417
TDM	0.4	[0.0882, 0.2215, 0.0394, -0.0256, 0.0927, 0.0206]	427
TDM	0.3	[-0.2349, 0.2134, 0.0576, -0.0140, -0.0372, -0.0277]	665
TDM	0.2	[0.2306, -0.3270, -0.0397, 0.1032, 0.1053, 0.1179]	573
TDM	0.1	[21.0586, -14.1472, 6.3378 -5.7688, 5.0797, 19.8442]	237 (failed)

TABLE II
SIMULATED REGISTRATION RESULTS OF THE IMAGE PAIR WITH 5% NOISE LEVEL

Methods	α	Results	Function Evaluations
Shannon-MI	None	10^{-3} ×[0.0966, 0.1908, 0.0744, -0.0598, 0.2922, 0.0586]	705
Tsallis-MI	0.9	10^{+6} ×[7.4277, -0.0253, 0.0036, 0.0151, 0.0000, 0.0250]	317 (failed)
Tsallis-MI	0.8	10^{+3}×[2.7646, -4.4706, -1.6623, -0.0016, -0.0014, 0.0016]	265 (failed)
TDM	0.9	10^{-4} ×[-0.2680, -0.6750, -0.1251, 0.7588, 0.5101, -0.0908]	498
TDM	0.8	[-0.0058, 0.0060, 0.0037, 0.0000, 0.0000, 0.0000]	487
TDM	0.7	[-0.0148, -0.0851, 0.0098, -0.0026, 0.0012, 0.0007]	486
TDM	0.6	[-0.0228, 0.0702, 0.0192, 0.0442, -0.1451, 0.0698]	415
TDM	0.5	[0.0271, 0.2358, 0.0116, 0.0369, -0.1695, 0.1034]	429
TDM	0.4	[-0.0054, 0.0462, -0.2152, 0.1597, -0.3532, 0.3608]	500
TDM	0.3	[0.3640, 0.3407, -0.0679, 0.0613, -0.0100, 0.2758]	496
TDM	0.2	[0.5475, 0.3337, 0.0715, 0.0537, 0.8598, 0.3013]	732
TDM	0.1	19.7359, -9.2095, 5.1806, -4.9970, 6.2699, 19.2690	297 (failed)

improvement achieved in the most recent iteration is within predefined boundaries. The direction matrix is initialized with unit vectors in each of the parameter directions.

An appropriate choice for the order in which the parameters are optimized needs to be specified, as this may influence optimization robustness. For instance, when matching images of the brain, the horizontal translation and the rotation around the vertical axis are more constrained by the shape of the head than the pitching rotation around the left-to-right horizontal axis. Therefore, first aligning the images in the horizontal plane by first optimizing the in-plane parameters $(t_x, t_y . t_z)$ may facilitate the optimization of the out-of-plane parameters (ϕ_x, ϕ_y, ϕ_z) [13].

In this work, the convergence parameters for the Brent and Powell optimization algorithms are set to 10^{-3} and 10^{-4} respectively.

III. EXPERIMENTS AND RESULTS ANALYSIS

A. *Simulated Registration Function*

Four pairs of T1 and T2 image volumes are obtained from the Brain Web Simulated Brain Database [17, 18] (181×217×60 voxels, 1 mm ×1 mm ×3 mm and noise levels were 0%, 3%, 5% and 7%, respectively), in which all the images have already been perfectly aligned and can be used as a testing platform for studying the performance of different objective functions. The image pair with 0% noise level is used as the training data in our experiments, from which the expected joint intensity distribution is estimated; other image pairs are used as the testing data to test the proposed TDM registration algorithm. It should be noted that, to improve the registration speed, these images have been subsampled by a factor of two in each dimension, and image T2 was taken as the floating image in each image registration task.

TABLE III
SIMULATED REGISTRATION RESULTS OF THE IMAGE PAIR WITH 7% NOISE LEVEL

Methods	α	Results	Function Evaluations
Shannon-MI	None	$10^{-3}\times$[-0.1123, 0.1957, 0.0657, 0.0933, -0.0257, 0.0690]	653
Tsallis-MI	0.9	$10^{-3}\times$[-0.0170, 0.2962, 0.0731, 0.1543, -0.1109, 0.0243]	803
Tsallis-MI	0.8	$10^{+4}\times$[2.7646, -4.4706, -1.6623, -0.0016, -0.0014, 0.0016]	308 (failed)
Tsallis-MI	0.7	$10^{+5}\times$[0.0769,-7.2242, 0.0005, -0.0000, -0.0000, -0.0006]	571 (failed)
TDM	0.9	$10^{-4}\times$[-0.4171, -0.4666, 0.2120, 0.4221, 0.4145, -0.3949]	490
TDM	0.8	[0.0207, -0.0694, 0.0000, -0.0005, -0.0000, 0.0297]	479
TDM	0.7	[0.0002, -0.1803, 0.0155, -0.1445, -0.0035, 0.0455]	493
TDM	0.6	[0.1077, -0.3369, 0.0434, -0.1680, 0.0221, 0.1284]	506
TDM	0.5	[0.2177, -0.0909, 0.2355, 0.4110, 0.2731, -0.2918]	742
TDM	0.4	[0.2391, 0.2295, 0.5410, -0.1280, -0.3759, -0.5405]	580
TDM	0.3	[1.1501, 0.0259, 0.5311, 2.0414, -0.5289, 0.3612]	105 (failed)

To visually investigate the performance of the proposed TDM registration method, here we first plot the behavior of the three measures (Shannon-MI, Tsallis-MI and TDM) as functions for translations (measured in millimeters) along the slice direction in Fig. 1, and each measure is plotted as a function on all the three testing image pairs. The value of α in Tsallis entropy used in function's plotting is 0.9.

Observing Fig. 1, it is obvious that the three TDM curves are smoother than the others, which could be expected to benefit the subsequent optimization process of TDM function.

B. Simulated Registration Experiments

To further demonstrate the performance of the TDM method, we conducted a series of simulated registration experiments on the three testing image pairs mentioned above using Shannon-MI, Tsallis-MI and the proposed TDM registration method. With respect to each registration task, the T2 image (floating image) was first transformed with an initial registration parameter θ_0 = [20, -20, 20, 5, -5, 5], prior to starting the optimization process. Since all the images have already been perfectly aligned, complete registration is obtained when the transformation parameter θ was optimized to θ = [0, 0, 0, 0, 0, 0]. The simulated registration results we obtained are shown in Table I, Table II and Table III, respectively including the transformation parameters we obtained and the number of the registration function evaluations during each registration task.

From tables we find that Shannon-MI succeeds registering three image pairs with higher precision, and the proposed TDM method can provide a more accurate registration with faster convergence if the value of α is chosen as 0.9. However, the performance of Tsallis-MI is poor and it is only successful in the registration of the testing image pair with 7% noise level. From three tables, we also can find that the convergence speed of the TDM method becomes faster as the value of α decreases with the exception of Tsallis entropy with α <0.7. Though the registration precision becomes poor as the value of α tends to 0, it still can succeed registering the three image pairs until α =0.3. So α can be chosen by trading convergence speed and registration precision.

C. Clinical Registration Experiments

Furthermore, clinical images obtained from Retrospective Image Registration Evaluation (RIRE) are used to test the proposed TDM method [19]. Registration estimates derived from bone implanted markers in this data (removed from the images prior to experimentation) provide an accurate estimate of rigid registration.

Images CT (512×512×29 voxels, 0.65 mm ×0.65 mm ×4.00 mm) and MR-PD (256×256×26 voxels, 1.25 mm ×1.25 mm ×4.00 mm) from the practice data, for which the gold standard transformation was available to us, are chosen as the training image pair, and images CT (512×512×28 voxels, 0.65 mm ×0.65 mm ×4.00 mm) and MR-PD (256×256×26 voxels, 1.25 mm ×1.25 mm ×4.00 mm) from patient 001 are used as the testing image pair that are to be registered. Note that the gold standard transformation of the testing data is unknown to us.

From the simulated registration results we know that compared with the Shannon-MI and TDM, Tsallis-MI performs poorer when applied into registration problem, so in this subsection, we only compare the performance of the proposed TDM method with α =0.9 and Shannon-MI by using them to register image CT to image MR-PD from patient 001. All translation and rotation parameters were initialized to zero. This is a typical starting estimate for automated registration in clinical use.

After 239 function evaluations, the TDM function converges to the finally transformation θ_1 = [8.1861, -15.5135, -0.2944, 0.1727, 0.3914, -0.6064], and the Shannon-MI declares the convergence when it reaches θ_2 = [8.1381, -15.2125, 0, 0, 0, -0.4571] after 252 evaluations. It is obvious that the difference between θ_1 and θ_2 is a little, and the convergence speed of

ISSN: 2517-9934
Vol:2, No:1, 2008

TDM method is relative faster. This is identical with the conclusions just drawn from simulated registration experiments.

So we can conclude that the proposed TDM is indeed feasible for intra-modality and multi-modality image registration. And compared with the classical Shannon mutual information and Tsallis mutual information, it enjoys faster registration convergence and higher registration precision.

IV. Conclusion

The divergence measure based on the Tsallis entropy, combined with the *a priori* knowledge of the expected joint intensity distribution, has been proposed to enhance the image registration process. The *a priori* knowledge of the expected joint intensity distribution is obtained from the pre-aligned images, which is called training image in this paper. Correspondingly, the images that are to be registered are called testing images. The joint intensity distribution computed from the testing images depends on the transformation from the floating image to the reference one, and this distribution is denoted as observed joint intensity distribution. As the transformation varies, the distance from the observed joint intensity distribution to the expected joint intensity distribution will change constantly. If the distance is enough samll, then we can declare that the two testing images are registered successfully. The distance is measured by a divergence measure based on the Tsallis entropy.

Experimental results show that, compared with other registration methods based on Shannon-MI and Tsallis-MI respectively, the TDM-based registration algorithm could obtain faster convergence without loss of registration precision. Future work will continue investigating our proposed method on large clinical data, and combine spatial information to the proposed TDM function to get further enhancement about the accuracy and robustness of the registration tasks.

Acknowledgment

The clinical images and the standard transformation(s) 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.

References

[1] L. G. Brown, “A survey of image registration techniques,” *ACM Computing Surveys,* vol. 24, no. 4, pp. 325-376, 1992.
[2] H. M. Chen, and P. K. Varshney, “Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation,” *IEEE Trans. Med. Imaging,* vol. 22, no. 9, pp. 1111-1119, 2003.
[3] R. P. Woods, J. C. Mazziotta, and S.R. cherry, “MRI-PET registration with automated algorithm,” *J. Comput. Assisted Tomography*, vol. 17, pp. 536-546, 1993.
[4] D. L. G. Hill, C. Studholme, and D. J. Hawkes, “Voxel similarity measures for automated image registration,” in *Proc. Visualisation in Biomedical computing*, U.S.A., 1994, pp. 205-216.
[5] J. P. W. Pluim, J. B. Antoine Maintz, and M. A. Viergever, “Mutual-information-based registration of medical images: a survey,” *IEEE Trans. Med. Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.
[6] Y. He, A. B. Hamza, and H. Krim, “A generalized divergence measure for robust image registration,” *IEEE Trans. Sig. Processing,* vol. 51, no. 5, pp. 1211–1220, 2003.
[7] M. P. Wachowiak, R. Smolikova, and T. M. Peters, “Multiresolution biomedical image registration using generalized information measures,” in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Montréal, 2003, pp. 846–853.
[8] S. Martin, G. Morison, W. Nailon, and T. Durrani, “ ‘Fast and accurate image registration using Tsallis entropy and simultaneous perturbation stochastic approximation,” *Electron. Lett.,* vol. 40, no. 10, pp.595-597, 2004.
[9] C. Tsallis, “Possible generalization of Boltzmann-Gibbs statistics,” *J. Stat. Phys.*, vol. 52, pp. 479-487, 1988.
[10] N. Cvejic, C N. Cangarajah, and Bull D. R, “Image fusion metric based on mutual information and Tsallis entropy,” *Electron. Lett.,* vol. 42, no. 11, pp.626-627, 2006.
[11] A. C. S. Chung, W. M. Wells III, A. Norbash, and W. E. L. Grimson, “Multi-modal image registration by minimizing Kullback-Leibler distance,” in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*,
[12] R. Gan, J. Wu, A. C. S. Chung, S.C.H. Yu, and W.M. Wells III, “Multiresolution image registration based on Kullback-Leibler distance,” in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Saint-Malo, 2004, pp. 599–606
[13] F. Maes, A. Collignon, D. Vandermeulen, G. Machal, and P. Suetens, “Multimodality image registration by maximization of mutual information, ” *IEEE Trans. Med. Imaging,* vol. 16, no. 2, pp. 187-198, 1997.
[14] A. Collignon, F. Maes, and D. Delaere et al, “Automated multimodality medical image registration using mutual information theory,” in *Proc. 14th International Conference Information Processing in Medical Imaging,* France 1995, 3, pp. 263-274.
[15] P. Viola, and W. M. Wells, “Alignment by maximization of mutual information,” in *Proc. 5th International Conference on Computer Vision,* USA 1995, pp.16-23.
[16] W. H. Press, B. P. Flannery, S.A. Teukolsky, and W.T. *Vetterling, Numerical recipes in C.* Cambridge, U.K.: Cambridge Univ. Press, 1992
[17] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, “Design and construction of a realistic digital brain phantom,” *IEEE Trans. Med. Imaging,* vol. 17, no. 3, pp. 463-468, 1998.
[18] McConnell Brain Image Centre Montreal Neurological Inst, McGill University, Montreal, QC, Canada. [Online]. Available: http://www.bic.mni.mcgill.ca/brainweb/.
[19] Vanderbilt Univ., Nashville, TN. Retrospective Image Registration Evaluation (RIRE). [Online]. Available: http://www.vuse.vanderbilt.edu/~image/registration/.

Shaoyan Sun received her B.S. degree in the Department of Applied Mathematics from QuFu Normal University, Qufu, China in 2002. She received her M. S. degree in the Department of Applied Mathematics from Dalian University of Technology, Dalian, China in 2004. She is currently studying for a Ph. D. degree at the Department of Applied Mathematics of Dalian University of Technology. Her research interests include image registration, pattern recognition, and neuro-informatics.

Liwei Zhang received his B. S., M. S. and Ph. D. degrees in the Department of Applied Mathematics from Dalian University of Technology respectively in 1989, 1992, and 1998. His research interests include nonlinear programming, variational analysis and optimization.

Chonghui Guo received his B. S. degree in the Department of Mathematics from Liaoning University in 1995. He received his Ph. D. degree in Management Science and Engineering from Dalian University of Technology in 2002. His research interests include signal processing, machining learning, bioinformatics and neuro-informatics.