# Objective Assessment of 3-D Medical Image Registration Results Using Statistical Confidence Intervals

Hong-Shu Wang[1], Dagan Feng[1,2] and Sung-Cheng Huang[3]

[1]Basser Department of Computer Science, The University of Sydney, NSW 2006, Australia

[2]Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong

[3]Department of Molecular and Medical Pharmacology, UCLA School of Medicine, Los Angeles, CA, USA

*Abstract*

It is of great interest to provide statistical information of image registration results. We successfully proposed a novel automatic approach based on statistical theory to estimate confidence intervals of the parameters in 2-D registrations. In this paper, the theory has been extended to more sophisticated conditions in the current 3-D phantom study.

Extensive Monte Carlo simulations have been conducted and the results are consistent with the calculated confidence intervals when different amounts of displacement and smoothing are presented. The predicted 95% confidence intervals have less than 20% errors of their sizes in normal noise conditions. By properly removing the systematic errors, the new theory works well even when large amounts of noise and gray value inconsistency are present in the images. The present results indicate that the use of the statistical confidence intervals developed in this paper can provide an objective assessment for 3-D image registration results.

## I. INTRODUCTION

Most of the medical image registration methods [1-3] minimize or maximize values of certain cost functions to achieve the global optimized matching of the images. These functions are usually the sum of squares of the distances between certain homogenous features in the two image sets to be registered. The sum of the distances between homogenous point pairs of the two image sets [4], distances between skin surfaces of CT, MR and PET images of the head in the "head-hat" method [5], the absolute difference between pixel values of PET image and pixel values of image simulated by MR image [6], and the ratio between pixel values and their means in the same tissue class [7, 8] are examples of these cost functions. However, most of these cost functions can not directly reflect the distance between actual and estimated positions of targets, i.e., the target registration error (TRE). Most medical registration applications demand accuracy and precision assessment methods to justify their results. Internal consistency measures [9] were used by Woods et. al. to place limits on registration accuracy for MRI data. Almost all other registration accuracy assessment methods fall into two broad categories: qualitative evaluations by visual inspection and quantitative evaluation by reference to results from a gold standard registration method. The former methods require special expertise and extensive experience, while the latter methods require an extremely accurate gold standard that can not be easily achieved. Different methods may not always be comparable to each other under identical criteria.

We developed a novel automatic method [14] to estimate confidence intervals of the resulting registration parameters and allow the precision of registration results to be objectively assessed for 2-D images. However, most of the registration applications are conducted in 3-D and the computational complexity of 3-D registration is remarkably higher, we extend the automatic approach to 3-D. The performance of the newly developed 3-D method has been rigidly validated by Monte Carlo simulations.

## II. THEORY AND MONTE CARLO SIMULATIONS

We refer the problem of image registration as nonlinear sum of square estimation of the transformation parameters, which results in the optimal fitting of one set of image (function) to the reference image (data). The confidence intervals or regions can be calculated using the following modified equation [10]:

$$\begin{aligned} &k_1(\theta-\theta_0)^2\cdot\sum(f')\le \\ &((s^2-s_{sys}^2)/(k_2(n-1)))\cdot F(p,n-p,1-\alpha) \end{aligned} \qquad (1)$$

where F is a chosen F-test value of the corresponding confidence level, $s^2$ is the residual sum of squares (registration cost function) value at the location of the estimated parameters, $s_{sys}^2$ is the systematic error present in $s^2$. $\sum(f')$ represents the sum of the derivatives of the reference image to the transformation parameters. $\theta$ and $\theta_0$ are the parameters corresponding to the confidence level and the optimal parameters found by the registration procedure, respectively. The parameter $k_1$ reflects the difference in units between translation (pixel) and rotation (degree). $k_2$ represents the portion of independent data points in all of the data point available in the reference image.

The residual sum of squares (RSS) consists of two parts: the systematic error and the error due to statistical noise. Since the systematic component in RSS is much less sensitive to spatial smoothing than the other component in Eq. (2), it can be estimated by applying smoothing filters to both sets of images with relatively larger FWHMs.

Monte Carlo studies to simulate 3D PET images and subsequent registrations of the simulated images were conducted. The resultant distributions of the estimated transformation parameters were used to assess the consistency of the 95% confidence intervals with the distributions in parameter space. 3D grey matter and white matter sinograms



0-7803-6503-8/01/$10.00 2001 IEEE

of the segmented 3D Hoffman brain phantom [11] were combined with the grey-to-white ratios of 2:1, 3:1 and 4:1 before reconstruction to see whether the discrepancies of the ratios in two images can affect the confidence intervals. Then, filtered back-projection reconstruction programs with various filters (i.e., Hanning, Ramp, Butterworth, Ham, Parzen and Shepp-Logan filters) were employed to reconstruct images of size 128x128x128. Various amounts of spatial displacements were introduced. Various levels of Poisson noise (i.e., total counts of $5\times10^5$, $1\times10^6$ and $2\times10^6$ per slice) were simulated. A 3-D Gaussian smoothing filter with a FWHM of 5 mm is applied to both sets of images before registration. The Powell's algorithm [12] was selected as the optimization procedure.

## III. RESULTS AND DISCUSSIONS

To determine the effective number of independent data points in the calculation of Eq. (1), we have drawn curves of the average numbers of data points falling outside of the 95% confidence intervals for every 100 simulations. It shows that $k_2$ should be chosen as $4.7\times10^{-4}$.

To illustrate the low correlation among the parameter estimates, the estimated rotation and translation parameters of 100 simulations were plotted with the 95% confidence region calculated based on Eq. (1). No strong inter-dependency of parameters is seen and the displacement parameter estimates concentrate at the center of the confidence region. The relative independence of the parameter estimates indicates that the confidence intervals could be calculated without considering the covariance terms and thus can be easily extended to higher dimensional parameter spaces.

The accuracy of the confidence intervals was found to be invariant to displacements and reconstruction filters from the simulation results. The confidence intervals obtained from the Monte Carlo simulations conform to the calculated intervals. The calculated confidence intervals need less than 20% adjustment of their values (Figures 1 and 2).

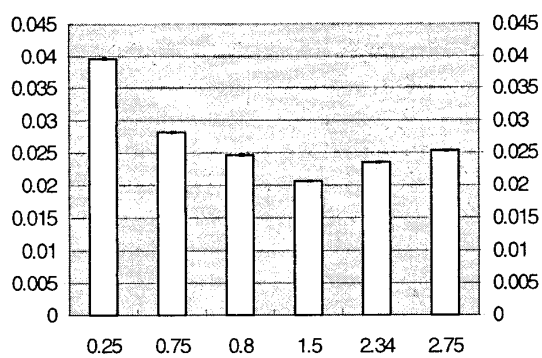


Figure 1: The calculated confidence intervals for rotation (in degree) and their errors.

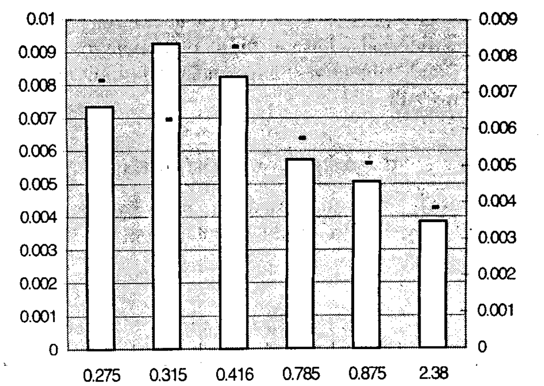


Figure 2: The calculated confidence intervals for translation (in pixel) and their errors.

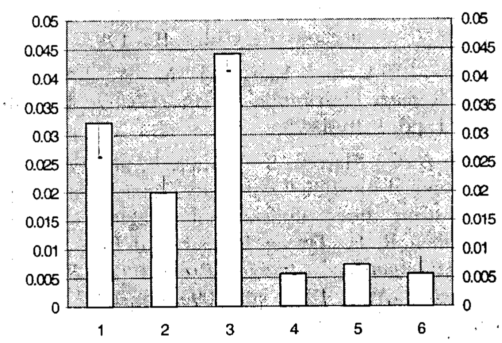


Figure 3: Calculated confidence intervals/parameters and their errors.

3-D Gaussian smoothing filters with various FWHM kernels were applied prior to registration. The RSS kernel curves were drawn. It is found that both terms in Eq. (2) responded to smoothing changes, but the RSS curve due to noise is much sharper at small FWHMs and has a shape gradually approaching that of the systematic RSS curve. When the derivatives of the curve is less than a certain threshold value (0.001 scaled), the corresponding RSS provides the appropriate estimation of the systematic errors.

Figure 3 shows the calculated confidence intervals/parameters and their errors. With adjustments for the systematic component of RSS, the estimated 95% confidence intervals require an average adjustment of only 10%, and about 35% adjustment in extreme cases.

## IV. CONCLUSIONS

The predicted confidence intervals based on statistical regression are consistent with the simulation results. Varying the amount of displacement, reconstruction filters, noise levels, or tracer distributions has little influence on the confidence intervals calculated, which demonstrated the robustness of this method. This method is also expected to be applicable to image registration with higher dimensional parameter space and images obtained from multiple modalities. The present results indicate the use of statistical confidence intervals has a high potential to provide an automatic and objective assessment of individual image registration result.

## V. Acknowledgments

This work was supported in part by the Australian Research Council (ARC), and by DOE contract DE-FC0387-ER60615.

## VI. References

[1] Van den Elsen, P. A., Pol, E.J.D. and Viergever, M. A., "Medical image matching – a review with classification", *IEEE Engng Med. Biol.*, vol. 12, pp. 26-39, 1993.

[2] Maurer, C.R. and Fitzpatrick, J.M., "A review of medical image registration", In Maciunas, R.J. (ed.), *Interactive Imageguided Neurosurgery*, American Association of Neurological Surgeons, Parkridge, IL., 1993, pp. 17-44.

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

[4] Evans, A. C., Marrett, S., Collins, L. and Peters, T. M., "Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems", In Schneider, R. H., Dwyer III, S.J. and Jost, R. G. (eds), *Medical Imaging: Image Processing*, SPIE Press, Bellingham, WA., 1989, vol. 1092, pp. 264-274.

[5] Pelizzari, C. A., Chen, G.T.Y., et. al., "Accurate three-dimensional registration of CT, PET and/or MR images of the brain", *J. Comput. Assist.* Tomogr., vol. 13, pp. 20-26, 1989.

[6] Lin, K.P., Huang, S.C. et. al., "A general technique for interstudy registration of multifunction and multimodality images", *IEEE Trans. Nucl. Sci.*, vol. 41, No.6, pp. 2850-55, December 1994.

[7] Ardekani, B.A., Braun, M., et. al., "A fully automatic multimodality image registration algorithm", *J. Comp. Assist. Tomogr.*, vol. 19(4), pp. 615-623, July/August 1995.

[8] Woods, R.P., Maziotta, J.C. and Cherry, S.R., "MRI-PET registration with automated algorithm" *J. Comp. Assist. Tomogr.*, vol. 17, pp. 536-546, 1993.

[9] Woods, R.P., Grafton S. T., Holmes C. J., Cherry S. R. and Mazziotta J. C., "Automated image registration: I. General methods and intrasubject, intramodality validation", *J. Comp. Assist. Tomogr.*, vol. 22(1), pp. 139-152, January-February 1998.

[10] Draper, N.R., Applied regression analysis, 2nd Ed., New York, Wiley, 1981.

[11] Hoffman, E. J., Cutler P. D., Guerrero T. M., Digby W. M. and Mazziotta, J. C., "Assessment of accuracy of PET utilizing a 3-D phantom to simulate the activity distribution of [$^{18}$F]fluorodeoxyglucose uptake in the human brain", *J. Cereb. Blood Flow Metab.*, vol. 11, pp. 17-25, 1991.

[12] Powell, M. J. D., "An efficient method for finding the minimum of a function of several variables without calculating derivatives", *Comput. J.*, vol. 7, pp. 155-163, 1994.

[13] Wang, H. S., Feng, D., Yeh, H. and Huang, S. C., "Objective Assessment of image registration results using statistical confidence intervals", *Conference Record, IEEE Medical Imaging Conference*, Seattle, Washington, 1999.