Medical Image Registration and Statistical Analysis of Its Results

By

Hongshu Wang

School of Information Technologies

The University of Sydney

NSW 2006, Australia

A thesis submitted in fulfilment of the

requirements for the degree of

Master of Science

May 2001

Abstract

Precision of medical image registration is very important in clinical diagnosis and treatments. It is usually assessed by visual inspection or by referring to other methods that require special expertise and extensive experience. In this study, we proposed a novel automatic approach based on statistical theory to estimate confidence intervals of the registration parameters, which allows the precision of registration results to be objectively assessed.

Under the assumption of local linearity, statistical confidence intervals of model fitting (regression) can be used to evaluate registration precision. Monte Carlo simulations using the Hoffman brain phantom with various amounts of displacement, noise and spatial filtering were conducted to evaluate the formula for estimating the confidence intervals in 2D image registrations. Monte Carlo simulation results are consistent with the calculated confidence intervals, and the agreement is applicable to different amounts of translation, angular rotation and spatial smoothing. The estimated parameter values fall within the predicted 90%, 95% and 99% confidence intervals in most cases. The present results indicate that the use of statistical confidence intervals can provide an objective assessment of individual image registration results.

Acknowledgments

I would like to attribute every achievement throughout my study and research to my supervisor, Professor David Dagan Feng, for his inspiration and experienced guidance. I would also like to thank Professor Henry Sung-Cheng Huang of the University of California at Los Angeles (UCLA) for his insightful discussions leading to this cutting-edge research. It is a rare privilege and a great honour for me to work with them.

I would like to express my sincere gratitude to associate Professors Jesse Jin and Doan Hoang of the Basser Department of Computer Science, and Drs Stefan Eberl and Koon-Pong Wong of the Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, not only for their kindly assistance and support, but also for their invaluable advice and fruitful discussions.

I am very grateful to the Department of Pharmacology, UCLA School of Medicine, for providing me excellent research facilities and environment during my stay in 1997. I would also like to thank the staff and students at the Basser Department of Computer Science for providing stable and efficient research environment. The students in BMIT Research Group, in particular, Weidong Cai, Fan Guo, Xiaomei Yu and Cindy Bai, gave me excellent suggestions to improve my study skills and their friendship is one of the motivations toward the success.

Finally, I would like to express my deepest gratitude to all my family members. Without their encouragement and support, I would not have the strength and determination to continue my research during many difficult times.

This work is supported by UPA and ARC grants.

Author's Publications

1. H. S. Wang, F. Guo, D. Feng, and J. Jin, "A signature for content-based image retrieval using a geometrical transform", *Proceedings, 6th ACM International Multimedia Conference*, pp. 229-234, Bristol, UK, Sept 1998.

2. F. Guo, H. S. Wang, D. Feng, and J. Jin, "Measuring similarity of images without a sophisticated segmentation", *Proceedings, Pan-Sydney Area Workshop on Visual Information Processing*, pp. 83-88, Sydney, Nov 27, 1998.

3. H. S. Wang, D. Feng, E. Yeh, and S. C. Huang, "Objective assessment of image registration results using statistical confidence intervals", *Proceedings, 1999 IEEE Nuclear Science Symposium and Medical Imaging Conference*, vol. 3, pp. 1578-1582, Seattle, USA, Oct 27 – 30, 1999.

4. H. S. Wang, D. Feng, and S. C. Huang, "Objective assessment of 3-D medical image registration results using statistical confidence intervals", *Conference Record, 2000 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Lyon, France, Oct 15 – 20, 2000.

5. H. S. Wang, D. Feng, and S. C. Huang, "A statistical method for the assessment of 3-D medical image registration", *Proceedings, Pan-Sydney Area Workshop on Visual Information Processing*, Sydney, Dec 1, 2000. (in press)

6. H. S. Wang, D. Feng, E. Yeh, and S. C. Huang, "Objective assessment of image registration results using statistical confidence intervals", *IEEE Transactions on Nuclear Science*, 48:106-110, 2001.

Table of Contents

1. Introduction

1.1 Medical Image Registration

Medical imaging is a vital part of a large number of applications in clinical setting. The applications do not only limit to diagnosis setting but treatment planning, and evaluation of surgical and radiotherapeutical procedures as well. The medical imaging modalities can be divided into two categories – anatomical modalities and functional modalities. Anatomical imaging modalities depict morphological structures. They include, for example, computed tomography (CT), magnetic resonance (MR) and ultrasound (US). Conversely, functional imaging modalities depict information on the physiological functioning of the underlying anatomy. They include planar scintigraphy, single-photon emission computed tomography (SPECT), and positron emission tomography (PET).

It is of interest to scan the same subject using different imaging modalities for clinical diagnosis. Due to the complementary nature of different imaging modalities, proper integration of the data obtained from different scans is required. The integration process that brings the modalities involved into spatial alignment is referred to as *image registration*. In simple wording, image registration matches an image to another usually of different modalities on the same subject through rigid transformation, for instance, three-dimensional (3D) rotation and translation.

Image registration plays a major role in at least two types of applications - data fusion and data comparison. In the former case, one takes advantage of the availability of multiple instances of supposedly identical data. The registration of these instances

with one another allows the extraction of common features, for example by averaging or by more refined processes (Irani and Peleg, 1991). In the latter case, however, a new problem arises as the registration process tries to align data that are possibly dissimilar. This class of problem leads to robust registration methods using an internal criterion that is not sensitive to outliers (Herbin *et al.*, 1989; Chiang and Sullivan, 1993). After registration, the tasks usually proceed by the detection of dissimilar regions given a statistically significant level of confidence.

A good example that demonstrates the use of registering different modalities is in the area of epilepsy surgery. Patients may undergo various MR and CT for anatomical reference, and fluorodeoxyglucose (FDG) PET studies for functional imaging. Registration of the images obtained from any combination of the above modalities can benefit the surgeon a lot in locating the foci and deciding whether surgical operations can be done or not.

1.2 A Challenging Issue

Most of the medical image registration methods (Maurer and Fitzpatrick, 1993; van den Elsen *et al.*, 1993; Maintz and Viergever, 1998) minimise or maximise values of certain cost functions to achieve the global optimised matching of the images. These functions are usually the sum of squares of the distances between certain homogeneous features in the two image sets. The sum of distances between homogeneous point pairs of the two image sets (Evans *et al.*, 1989), distances between skin surfaces of CT, MR and PET images of the head in the "head-hat" method (Pelizzari *et al.*, 1989), the absolute difference between pixel values of PET

image and pixel values of image simulated by MR image (Lin *et al.*, 1994), and the ratio between pixel values and their means in the same tissue class (Woods *et al.*, 1993, Ardekani *et al.*, 1995) are examples of these cost functions. However, most of these cost functions do not directly reflect the distance between the actual and estimated positions of targets - the target registration error (TRE). Most medical registration applications require accuracy and precision assessment methods to justify their results. Woods *et al.* (1998) used the so-called internal consistency measures to replace 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 cannot be easily achieved. Different methods may not always be comparable to each other under identical criteria. Thus, it is desirable if we could estimate statistical confidence intervals of the registration parameters, which allow the precision of registration results to be objectively assessed.

A novel automatic method to estimate statistical confidence intervals of the resulting registration parameters is proposed in this work. It is based on the assumption of local linearity, which holds when the cost function is near the neighbourhood of the minimum. The method allows the precision of registration results to be objectively assessed for 2D and 3D images. We evaluate its performance by Monte Carlo simulations.

1.3 Organisation of this thesis

This thesis will be organised in the following manner. Chapter 1 (i.e. this chapter) gives a general overview of this work. Chapter 2 presents the general background and literature review of image registration. Chapter 3 presents the theory of the proposed approach for the objective assessment of registration results using statistical confidence interval and its application in 2D image registration. Chapter 4 extends the proposed approach to 3D image registration. Chapter 5 summarises and concludes the work present in this thesis.

2. Image Registration – An Overview

Classification of image registration techniques has been formulated by van den Elsen *et al.* (1993) and a considerably augmented and detailed discussion can be found elsewhere (Maintz and Viergever, 1998). In this chapter, an overview of different image registration methods and their natures are given. Basically, image registration methods can be divided into two classes - image based and non-image-based. In the following sections, we will discuss them in more detail.

2.1 Image-based Methods

Image-based registration can be divided into extrinsic, i.e. based on foreign objects introduced into the imaged space, and intrinsic, i.e. based on the image information as generated by the subjects.

2.1.1 Extrinsic Registration

Extrinsic methods rely on artificial objects attached to the subjects. The objects are designed to be well visible and detectable in all pertinent imaging modalities. The registration of the acquired images is comparatively easy and fast and can usually be automated since the registration parameters can often be computed explicitly without the need for complicated optimisation algorithms. However the major drawback of these methods are the prospective character must be made in the image acquisition and often it involves the use of invasive markers for references. Although non-invasive markers can be used, the registration results are less accurate. A commonly used fiducial object is a stereotactic frame (Lunsford, 1988) that screwed to the subject's outer skull table, a device which has been regarded as the *gold standard* for

registration accuracy.

It is apparent that extrinsic methods cannot include subject-related image information, the nature of the registration transformation is often restricted to rigid, i.e. rotations and translations only. Because of such limitation and various other practical issues, use of extrinsic 3D to 3D methods is largely restrained to brain and orthopedic imaging (Simon *et al.*, 1995; Ellis *et al.*, 1996), although markers can be used in projective imaging of any body area.

2.1.2 Intrinsic Registration

Intrinsic methods only rely on the image content generated from the subject. Registration can be based on a limited set of identified salient points (points / landmarks) on the alignment of segmented binary structures (regional / structural-based), most commonly object surfaces, or directly onto measures computed from the image grey values (voxel property-based).

2.1.2.1 Point / Landmark-based Methods

Using anatomical landmarks such as morphological salient or visible points identified by the users (Evans *et al.*, 1989), or points or curves located by automatic segmentation algorithms. Landmark-based methods usually require subjectively recognized landmark points, and, to achieve accurate registration by high dimensional transforms, a large number of landmark point pairs are desirable. As a result, expertise and experience is the fundamental factor deciding the precision of the registration results. On the other hand, automatic segmentation usually cannot achieve high

precision due to lack of highly reliable segmentation method for medical images, especially for low-resolution and non-anatomical images like PET.

Precision of landmark or point based methods is usually depending on fiducial localization error (FLE), and evaluated by mentally comparison with other methods, or by calculation of average FLE by simulations made on phantoms or previous patients (Fitzpatrick *et al.*, 1998).

2.1.2.2 Regional / Structural-based Methods

These methods register corresponding region pairs by segmenting of structures or regions of human body, most of them are based on anatomical interpretations of the structures. These parts extracted from both image modalities are then used as sole input for the alignment procedures. The segmentation results are usually adjusted by human experts, due to the limitations of segmentation algorithms and poor quality of current CT, MR and PET images. The most famous example of the primitive methods is the "head-hat" method introduced by Pelizari *et al.* (1989), this method is based on the surface extracted from different image modalities of human head.

An obvious drawback of this kind of methods is that the registration accuracy is limited to the accuracy of the segmentation process. Most of the regional/structural-based methods are based on elastically deformation of structures from one image into the same corresponding structures of the second images. The optimization criterion is always locally defined and computed and the deformation is constrained by elastic modeling constraints imposed onto the segmented curve or surface. Deformation

curves often appear as snakes or active contours. Deformation methods are in theory very well suited for inter-subject and atlas registration. They often need a good initial position in order to converge properly, which is generally realized by rigid pre-registration of the images involved. Also, they still suffer the in-accuracy of segmentation methods.

2.1.2.3 Voxel Property-based Methods

Voxel property-based methods using the full image content are the most interesting methods of current research. Theoretically, these are the most flexible of the registration methods, since, unlike all other methods, they do not start by reducing the grey-level image to relatively sparse extracted information, but use all of the available information throughout the registration process. These methods have been pushing to be used in clinical practice due to the increasing clinical call for accurate and retrospective registration.

Principal axes and moment based methods compute image center of gravity and its principal orientations from zeroth and first order moments. Registration is then performed by aligning the center of gravity and the principal orientations (Alpert *et al.*, 1990). Examples of other full content based methods include: cross-correlation of original images or extracted feature images (Junck *et al.*, 1990); Fourier-domain-based cross-correlation (Leclerc and Benchimol, 1987); histogram based methods (Hill, 1993); mutual information (Colligon *et al.*, 1995); zero crossing in difference images (Venot *et al.*, 1983); and absolute or squared intensity differences (Hoh *et al.*, 1993), etc.

The methods having most important influence and having the potential to dominant future research is variance minimization methods. They make use of full image content of both images to be registered, utilize the anatomic information of pixels, and at the mean times, focus on minimization of the difference (variance) between the two image sets. These methods include minimization of variance of intensity ratios (Woods *et al.*, 1993, 1998) and minimization of variance of grey values within segments (Ardekani *et al.*, 1995).

The intensity ratio variance minimization algorithm calculates the ratio of one image to the other on a voxel-by-voxel basis and then iteratively moves the images relative to one another to minimize the variance of this ratio across voxels. The technique is fully automated and is independent of the specific anatomic structures being imaged. It is also retrospective, so that the head position can vary markedly from one image set to another. The alignment technique is based on the idealized assumption that if two images sets are accurately aligned, the value of any voxel in one image set is related to the value of the corresponding voxel in the other image set by a single multiplicative factor. If the image sets are misaligned, the multiplicative factor is no longer constant but varies from voxel to voxel throughout the image. The alignment algorithm systematically moves the two image sets relative to one another until this voxel-to-voxel variation is minimized.

The method using grey value minimization within segments achieves registration by optimizing the segmentation induced on the second image. For example, the head contour is detected on the MR image using a gradient threshold method. The head

region in the MR image is then segmented into a set of connected components using *K*-means clustering algorithm. When the two image sets are registered, the segmentation of the MR image indirectly generates a segmentation of the PET image. The best registration parameters are calculated to be the one that optimizes the segmentation induced on the PET image. This method is an improvement of the method based on intensity ratio variance in the way that it utilizes anatomical information of the image. It has the potential to achieve clinically satisfied quality without user interaction and can be applied to a wide range of registration problems. Further research can be applied to this direction by retrospective segmentation, based on intermediate registration results, and by minimize the variance inside segmented regions.

2.2 Non-image-based Methods

Registration of multimodal images can also be non-image-based. This seems paradoxical but it is possible that if the imaging coordinate systems of the two imaging modalities cross calibrated to each other in a certain extent. This requires the scanners to be brought into the same physical location and the assumption that the patient remains still between the acquisitions. These are prohibitive pre-requisites in nearly all applications but they can be met in certain applications such as ultrasound systems (Hata *et al.*, 1994; Erbe *et al.*, 1996), which can come as hand-held devices that are equipped with a spatial localisation system. The ultrasound systems are easily calibrated and can be used while the subject is immobilised on the CT, MR or operating gantry. This technique of calibrated coordinate systems is also often used in

registering the position of surgical tools mounted on a robot arm to images.

2.3 Optimisation Methods for Registration

The parameters that made up for registration transformation can either be computed directly from the available data, or searched for by finding an optimum of some function defined on the parameter space. In the former case, the manner of computation is completely determined by the paradigm. In the case of searching the parameters in the parameter space, most registration methods are able to formulate the paradigm in a standard mathematical function of the transformation parameters to be optimised. This function attempts to quantify the similarity as dictated by the paradigm between two images based on certain transformation. Such functions are generally less complicated in monomodal registration applications since the similarity is straightforward to be defined. Standard optimisation techniques can be used if the similarity function is well behaved. Popular optimisation techniques include the Powell's method (Powell, 1964), downhill simplex method (Nelder and Mead, 1965), the Marquardt's method (Marquardt, 1963), and simulated annealing (Kirkpatrick *et al.*, 1983). These methods and their computer implementations are well documented in Press *et al.* (1992).

3. Statistical Confidence Intervals – Application to Assessment of 2D Image Registration

In this chapter, a novel approach for the objective assessment of the registration results is presented. Under the assumption of local linearity, statistical confidence intervals of model fitting (regression) is used to evaluate registration precision. Monte Carlo simulations using the Hoffman brain phantom with various amounts of displacement, noise and spatial filtering were conducted to evaluate the formula for estimating the confidence intervals in 2D image registrations.

3.1 Nonlinear Estimation and Confidence Interval

Image registration can be regarded as a nonlinear estimation problem for finding an optimal set of transformation parameters in one set of image (function) that can best fit another set of image (data) in some sense. For least-squares criterion, the cost function to be minimised is given by:

$$\min_{\theta \in \Re^p} f(\theta) = \left\| \mathbf{I}(\theta) - \mathbf{I} \right\|^2 \tag{3.1}$$

where f represents the cost function, $\mathbf{I}$ is the image to be_registered, $\mathbf{I}(\theta)$ is an image function that is used to register to $\mathbf{I}$, θ is a solution vector that yields minimum f, and p is the number of parameters to be estimated. The parameters in θ are dependent on the registration problem. In this study, three parameters including translation in x and y directions and rotation are to be estimated.

For least square estimation methods, the cost function could be assumed to be linear around the neighborhood of the current estimates as the cost function can be expanded into a Taylor series about the current estimates θ_0 as (Draper, 1981):

$$f(\theta) \approx f(\theta_0) + \frac{\partial f}{\partial \theta_0}^T (\theta - \theta_0) \tag{3.2}$$

and the confidence intervals (or regions) of the parameter estimates can be calculated using the following equation (Draper, 1981):

$$(\theta - \theta_0)^2 \cdot \sum(f') \leq \frac{s^2}{n-1} \cdot F(p, n-p, 1-\alpha) \tag{3.3}$$

where F is the F-test value of the corresponding confidence level, s^2 is the residual sum of squares (value of the registration cost function at the location of the estimated parameters), $\sum(f')$ represents the sum of the derivatives of the reference model image with respect to the transformation parameters, n is the number of data points, and α is the level of significance.

3.2 Data and Monte Carlo Simulations

Monte Carlo studies to simulate 2D PET images and subsequent registrations of the simulated images were conducted. The resulting distributions of the estimated transformation parameters were used to assess the consistency of 90%, 95% and 99% confidence intervals with the distributions in parameter space. Two-dimensional gray matter and white matter sinograms of the segmented 2D Hoffman brain phantom (Hoffman *et al.*, 1991) were combined with the gray-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 with various filters (i.e., Hanning, Ramp, Butterworth (fifth order), Ham, Parzen and Shepp-Logan filters) (all with cutoff at the Nyquist frequency) were employed to reconstruct images of size 128×128. Various amounts of spatial displacements (i.e., rotations of 0.3, 0.8, 1.2 and 3.3 degrees, and translations of 0.16, 0.8, 1.6 and 2.4 mm) were introduced. Various levels of Poisson noise (i.e., total counts of $5x10^5$, $1x10^6$ and $2x10^6$) were simulated. A Gaussian smoothing filter with a FWHM of 5 mm was applied to both sets of images before registration. Powell's algorithm (Powell, 1964) was used to optimize the parameter θ in (3.1). This algorithm was used because it has been found to be effective for image registration and is used by many image co-registration programs. Furthermore, the confidence interval results are not expected to be dependent on the particular algorithm used.

3.3 Effective Number of Independent Data Points

Since all the data points involved in the calculation of (3.3) should be statistically independent from each other, and the data points in the images are correlated, the total number of points in the image could not be used directly as n and the effective number of independent data points needs to be estimated.

To determine the effective number of independent data points involved in the estimation of confidence intervals, we first used one Monte Carlo simulation study based on normal condition where there is no discrepancy in gray-to-white matter ratio. A pair of images was tested with different values of n in order to determine the optimum value of n that best fits (3.3) for 95% confidence interval in which we

expected that there are about 5% of data points falling outside of it. The same number n selected according to this simulation result was also found to be consistent for both the 95% and 90% confidence levels. We have further investigated the validity of the selected number n in various simulated conditions in other parts of the study.

3.4 Adjustments of the RSS

In cases of extreme noise conditions and large contrast discrepancies, the residual sum of squares (RSS) consists of two parts: the systematic error and the error due to statistical noise:

$$RSS = RSS_{system} + RSS_{noise} \tag{3.4}$$

The contributing factors to the systematic error include the innate difference between the two images, inappropriate registration method, and precision error of the program. Such errors are independent of the initial displacements and noise. The second part of the residual sum of squares is due to statistical noise. If the systematic error is relatively large compared to the noise term, i.e., for cases with very low noise levels and high gray-to-white ratio discrepancies, the estimated residual sum of squares needs to be adjusted for systematic error.

Since the systematic component in the RSS is much less sensitive to the spatial resolution of the images than the other component in (3.4), it can be estimated by applying smoothing filters to both sets of images with relatively large FWHMs when the parameters are found. By removing the systematic component, the resulting RSS provides the estimation of the noise component in (3.4).

3.5 Results and Discussion

3.5.1 The Number of Independent Data Points

To determine the actual number of independent data points, n, in the calculation of (3.3) for the case of matched image contrast, the estimated value of n is plotted against the average percentage of data points falling outside of the 90% and 95% confidence intervals for every 100 realizations, as shown in Figure 3.1.

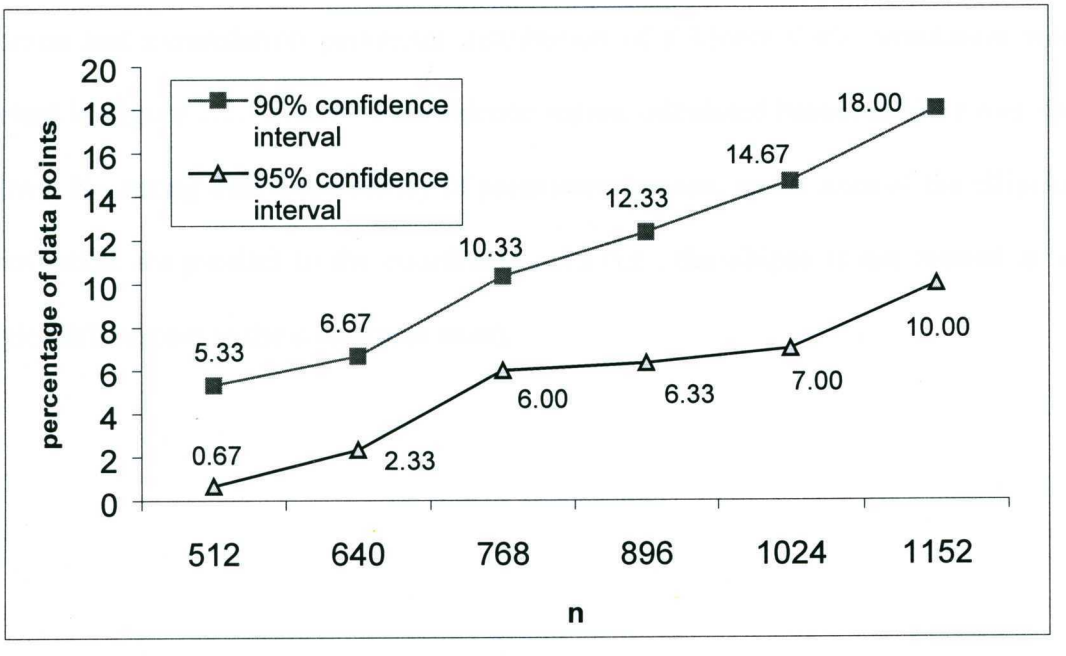


Figure 3.1 Percentage of data falling outside of the 90% and 95% confidence intervals as a function of the estimated number of independent data points in 2D simulations. At n=768, the 90% curve has a value of 10.33 and the 95% curve has a value of 6.0.

It can be seen that when n is chosen as 768, the percentage of data points falling outside of the confidence intervals are 6 and 10.33 for the 95% and 90% confidence intervals, which are close to the expected percentage of data points falling outside of the corresponding confidence intervals. Hence, the effective number n is set to 768 for all of the remaining calculations in this study.

3.5.2 Low Correlation Between Parameters

To illustrate the low correlation between the parameter estimates, the estimated rotation and x-translation parameter distribution of a Monte Carlo simulation were plotted in Figure 3.2. The 95% confidence region calculated based on (3.3) was also shown. No strong inter-dependency of parameters is seen, as the axes of the elliptical distribution are parallel to the coordinate axes (i.e., the ellipse is not rotated at an angle with respect to the coordinate axes).

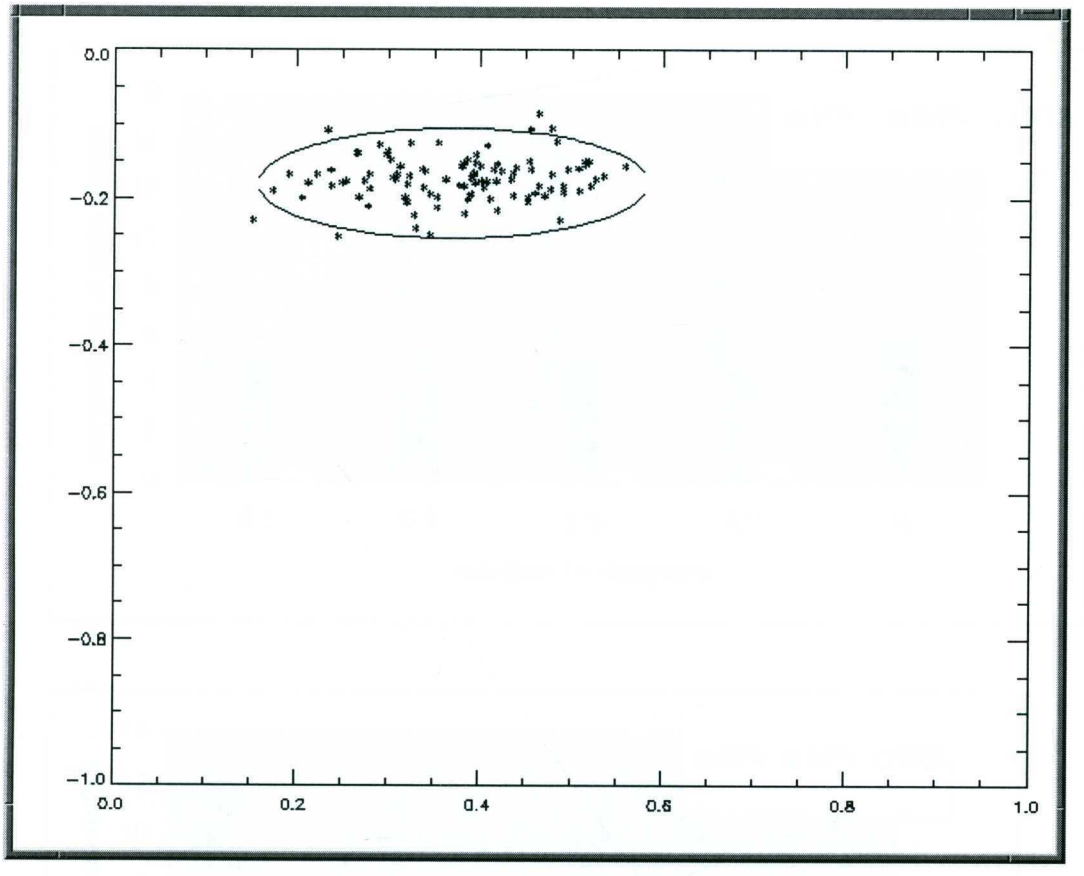
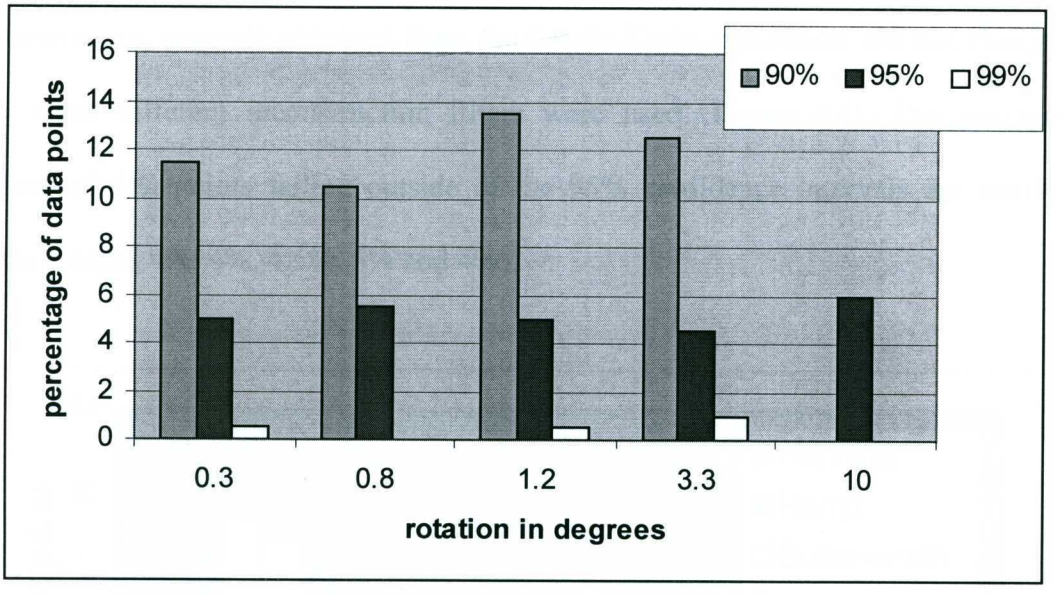


Figure 3.2 Distribution of registration results in the parameter space together with the calculated 95% confidence region. It shows that around 95% of the parameters distribute in the confidence region in normal noise condition (total counts of $5x10^5$). The x- and y-axes represent the amount of translation (in mm) and rotation (in degree), respectively. Approximately 95% of the result points fall in the region and no apparent correlation is found between the two-parameter estimates.

3.5.3 Displacement Invariant

The accuracy of the confidence intervals was found to be invariant to displacements as shown in Figure 3.3. For example, the percentages of data points falling outside of 95% confidence intervals are approximately constant (5%, 5.5%, 5% and 4.3%, for various rotation angles). When larger ranges of spatial displacement and translation (e.g. rotation of 10 degrees and translate of 5 mm) were used, the maximum percent data points falling outside of the 95% confidence intervals was approximately 6%, which is also close to what one would expect.



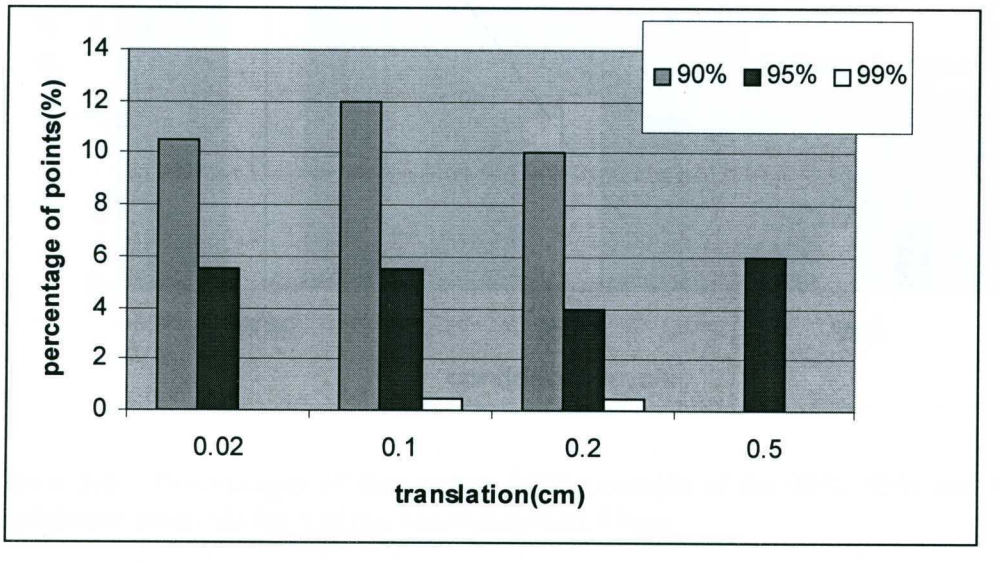


Figure 3.3 Percentages of data points falling outside of the 90%, 95% and 99% confidence intervals for various cases when the two sets of images to be registered have different amounts of rotation or translation between them. (* Only 95% confidence interval data is available for 10-degree rotation and 0.5cm translation)

3.5.4 *Different Reconstruction Filters*

The confidence intervals obtained from the Monte Carlo simulation did not change much when different reconstruction filters were used (Figure 3.4). The average numbers of data points falling outside of the 95% confidence intervals are nearly constant (4.3%, 4%, 4%, 4.5%, 5% and 4%).

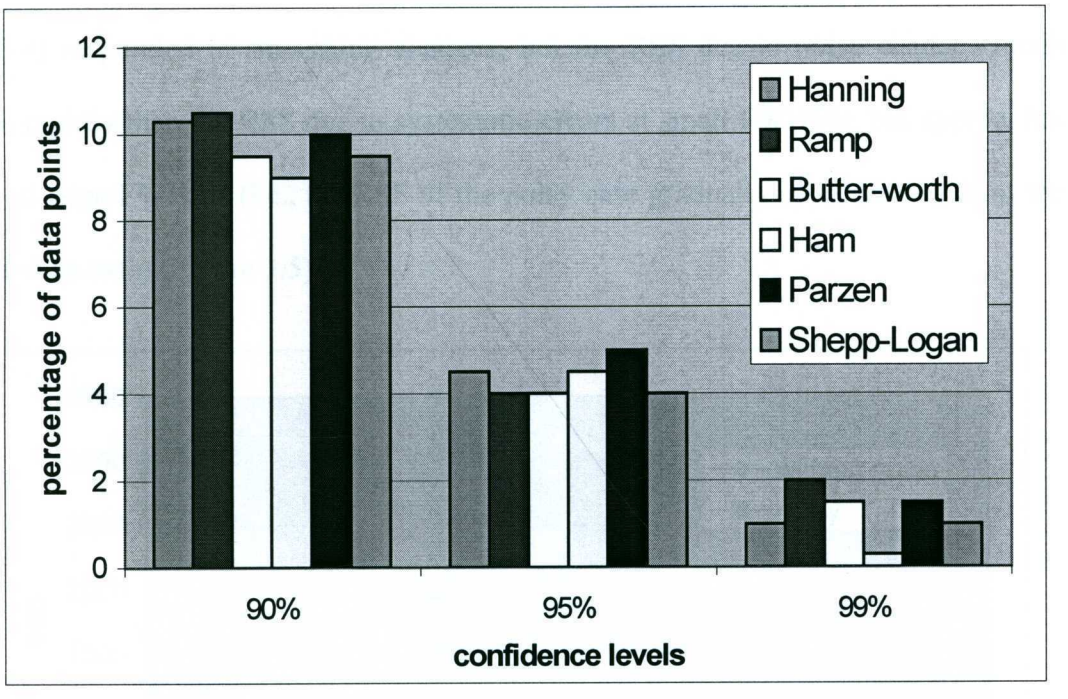


Figure 3.4 Percentages of data points falling outside of the 90%, 95% and 99% confidence intervals for various reconstruction filters.

3.5.5 *RSS Versus Spatial Smoothing*

Gaussian smoothing filters with various FWHM kernels were applied prior to registration. Two noise-free PET image sets with different gray-white contrast in the distributions were reconstructed and registered to study the systematic errors. Noise was then added to these two sets of image data and the same registration process was applied to study the shape of the RSS as a function of spatial smoothing. Both terms in (3.4) responded to smoothing changes, but the RSS due to noise changes much more rapidly than the RSS due to systematic errors at small FWHMs, but approaches zero at large FWHMs (i.e., the RSS of the noisy case gradually approaches that of the noise-free case) (Figure 3.5).

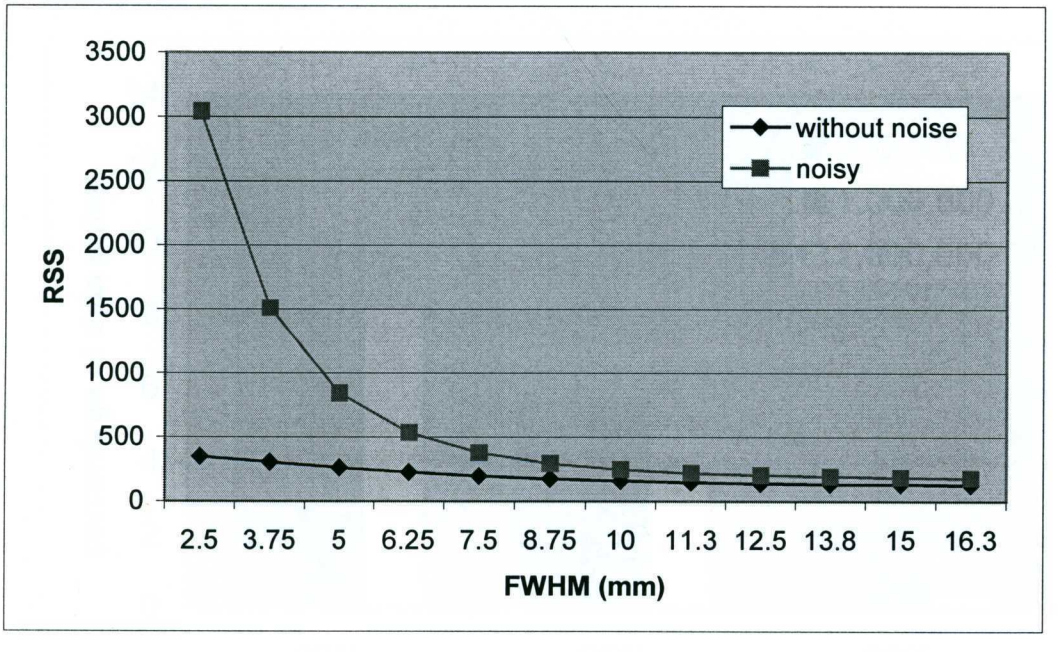


Figure 3.5 Responses of RSS and RSS_{sys} to spatial smoothing for normal PET image contrast condition (total counts of $5x10^5$). The curve of the RSS for the noisy case changes much faster than the curve of the noise-free case at low FWHMs but the two curves are nearly equal at high FWHMs. Similar results are found for various image contrast conditions.

3.5.6 *Effects of Noise Levels and Grey-to-White Ratio Discrepancies*

Without adjustment for the systematic component of the RSS, the numbers of data points falling outside the confidence intervals are not always constant at different noise levels. The numbers of data points falling outside of the calculated confidence increases for higher noise levels. At high noise levels ($5x10^5$ total counts), the average percentage of data points falling outside of the 90% confidence interval is 11.78, while at a low noise level ($2x10^6$ total counts), the average percentage of data points falling outside of the 90% confidence interval is 8.27 (Figure 3.6). This implies that when the noise level is low, the RSS is dominated by systematic errors, making the estimated confidence intervals relatively larger than they should be.

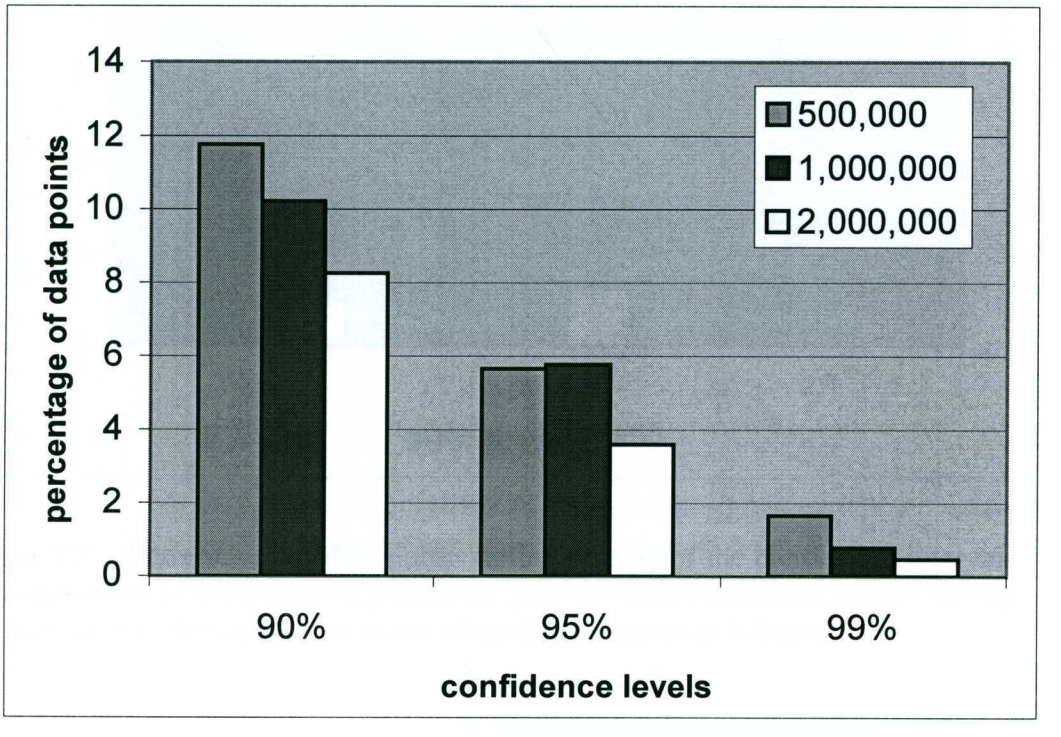


Figure 3.6 Percentages of data points falling outside of the confidence intervals for different levels of noise in the images to be registered. The numbers of outside data points are inversely related to noise levels.

When the ratios of the gray-to-white matters were not the same in the two images, the confidence intervals calculated without adjustment for the systematic component of the RSS did not always agree with simulation results. If discrepancies between gray scale ratios in the two image sets are present, the calculated confidence intervals will increase (Figure 3.7). This implies that when the RSS due to systematic error is large, the estimated confidence intervals could be in error.

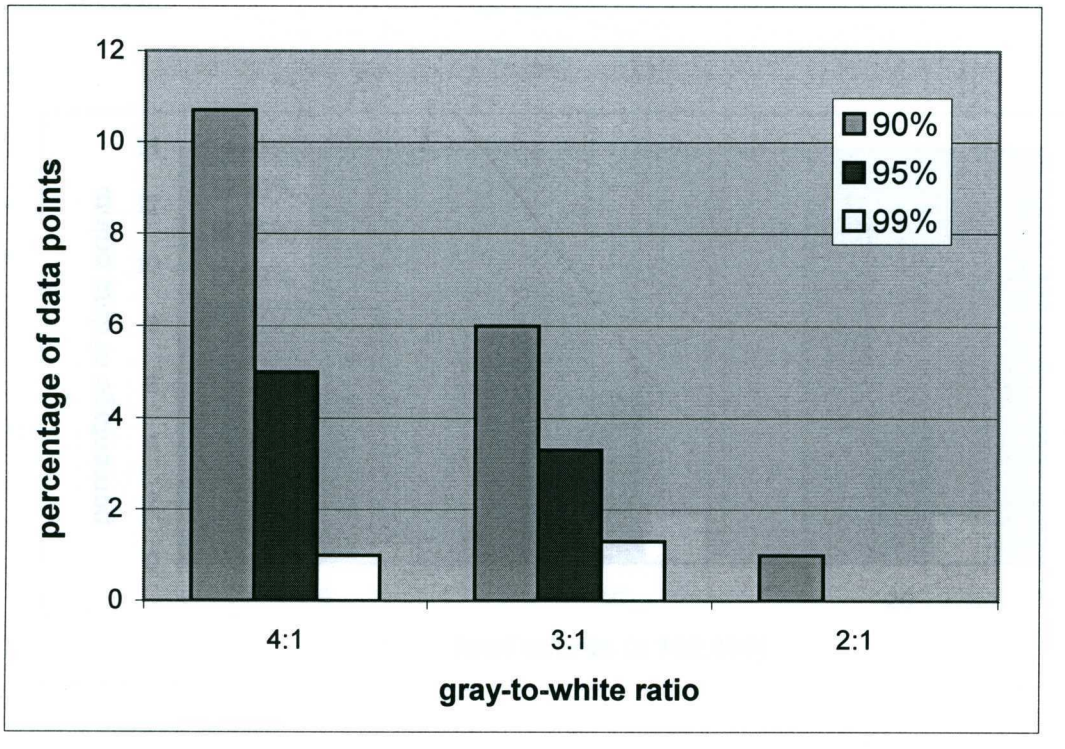


Figure 3.7 Percentages of data points falling outside of the confidence intervals for registration of images having different gray-to-white ratios and images having the ratio of 4:1. The number is lower when the discrepancy is large.

3.5.7 *Effects of Removing the Systematic Errors*

Figure 3.8 shows the percentage of data points falling outside of the confidence intervals for different noise levels after the estimated systematic component of RSS was removed before applying (3.4) as described in Section 3.4. Figure 3.9 shows the results after removing the estimated systematic component of RSS but for different gray-to-white matter ratio discrepancies in the image pairs. It can be seen that the predicted confidence intervals again do not vary much for different image contrast discrepancies.

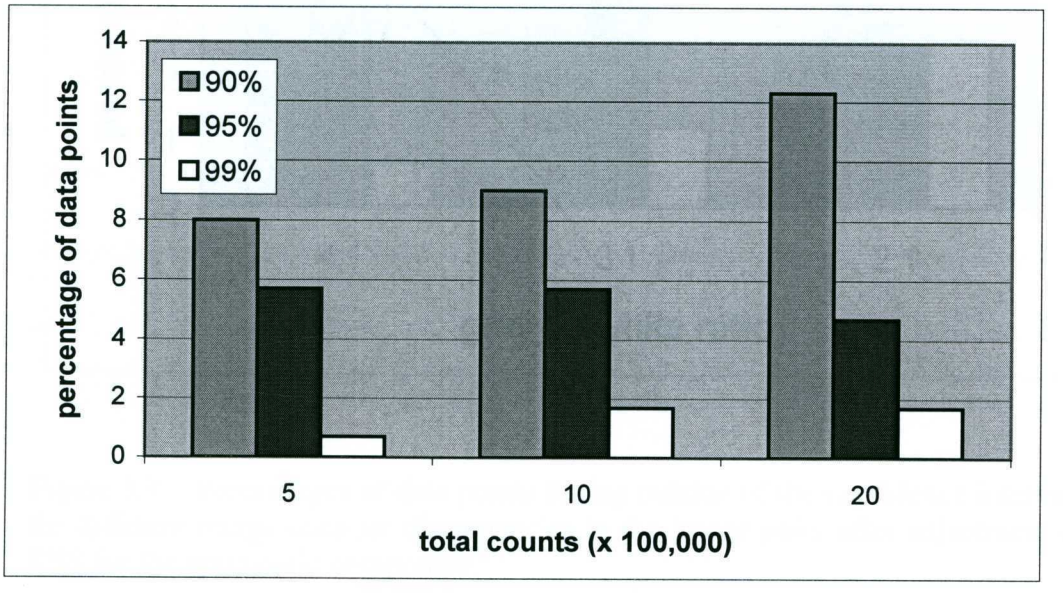


Figure 3.8 Percentages of data points falling outside of the confidence intervals for different noise levels after adjustment of RSS for the systematic component.

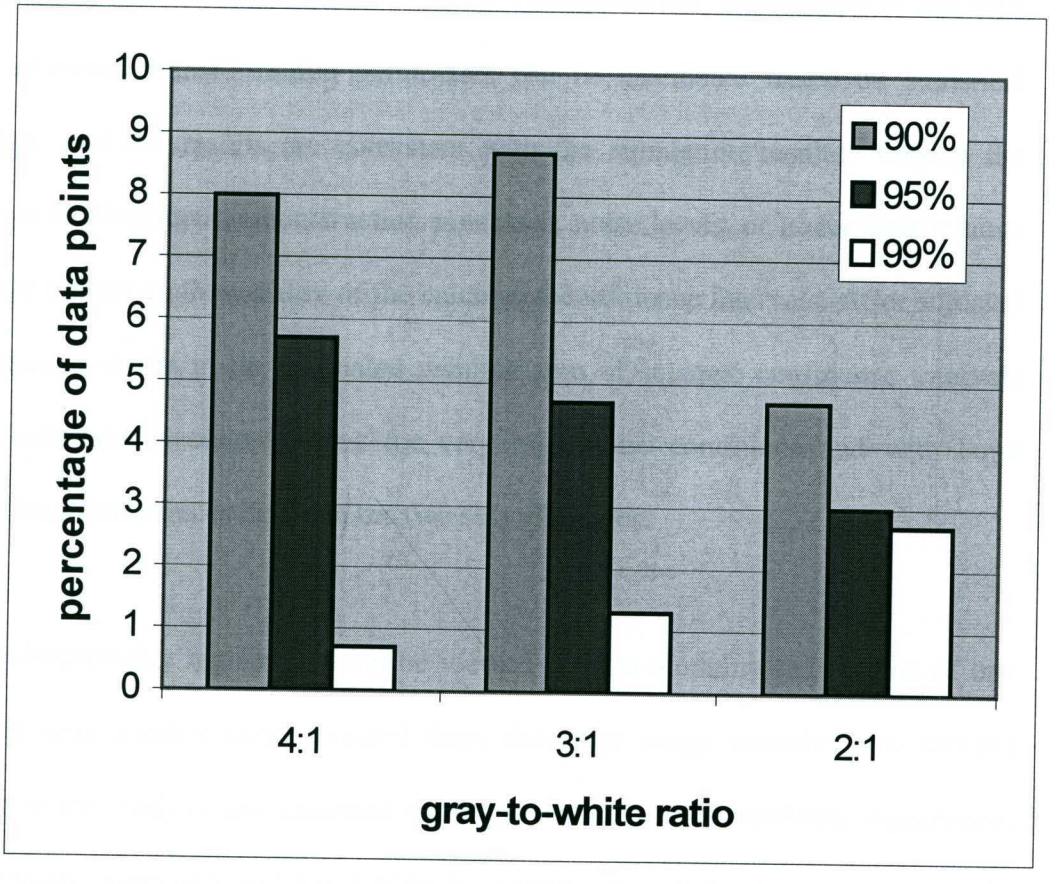


Figure 3.9 Percentages of data points falling outside of the confidence intervals for different image contract discrepancies in the image pairs after adjustment of RSS for the systematic component.

3.6 Summary

We have shown in this chapter that for image co-registration, the confidence intervals of the estimated transformation parameters can be calculated based on statistical regression, and the results are consistent with the simulation results. Varying the amount of displacement, reconstruction processes, noise levels, or tracer distributions have little impact on the validity of the calculated confidence intervals. After adjusted for systematic errors in the estimated residual sum of squares, confidence intervals can be calculated accurately even for very high noise conditions and with large distribution discrepancies between the two sets of images.

Since multi-modality registration can be viewed as intra-modality registration of one image set with another one simulated from the other image modality, the method proposed in the study is also expected to be applicable to multi-modality registration. Hence, visual inspection and validations by experts are not necessary to assess the precision of the registration results. The present results indicate that the use of statistical confidence intervals has a high potential to be used to provide an automatic and objective assessment of individual image registration result.

4. Statistical Confidence Intervals – Application to Assessment of 3D Image Registration

In previous chapter, we proposed a novel automatic method to estimate confidence intervals of the 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. In this work, we extend the proposed automatic approach to 3-D cases and we evaluate its performance by Monte Carlo simulations.

4.1 Theory and Monte Carlo Simulations

The problem of image registration can be regarded as finding a set of transformation parameters that gives an optimal fitting/matching of one set of image to the reference image. For least-squares problem, the confidence intervals or regions can be calculated using the following modified equation:

$$k_1(\theta - \theta_0)^2 \cdot \sum (f') \leq ((s^2 - s_{sys}^2)/(k_2(n-1))) \cdot F(p, n-p, 1-\alpha) \quad (4.1)$$

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 $s^2 \cdot \Sigma(f')$ represents the sum of the derivatives of the reference image to the transformation parameters. θ and θ_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 points available in the reference image. The residual sum of squares (RSS) between two images 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, it can be estimated by applying Gaussian 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. Three-dimensional grey matter and white matter sinograms of the segmented 3D Hoffman brain phantom [Hoffman *et al.* 1991] 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. Filtered back-projection reconstruction algorithm with different filters (i.e., Hanning, Ramp, Butterworth, Ham, Parzen and Shepp-Logan filters) was then employed to reconstruct images of size 128x128x128. Various amounts of spatial displacements were introduced. Various levels of Poisson noise (total counts of $5x10^5$, $1x10^6$ and $2x10^6$ 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 [Powel, 1964] was selected as the optimization procedure.

4.2 Results

To determine the effective number of independent data points in the calculation of Eq.

(4.1), 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 4.7×10^{-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. (4.1). 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.

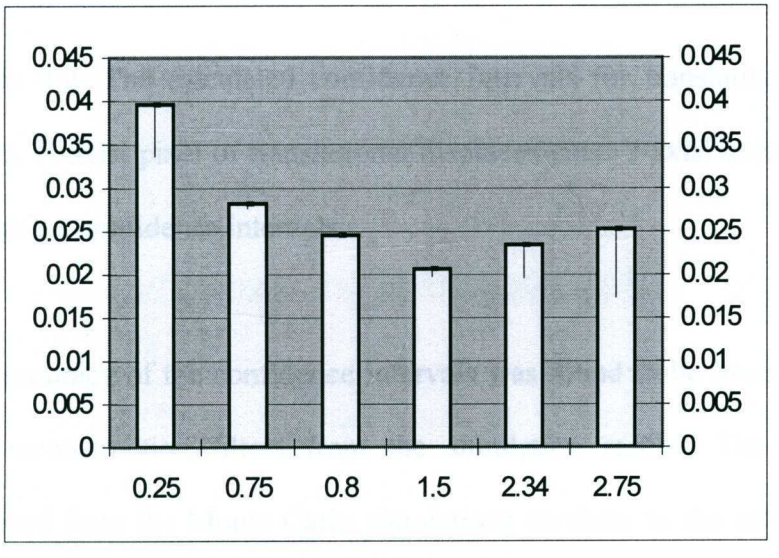


Figure 4.1: The calculated confidence intervals for rotation (in degree) and their errors. X-axis: degree of rotation displacements. Y-axis: mean and variance of the calculated confidence intervals.

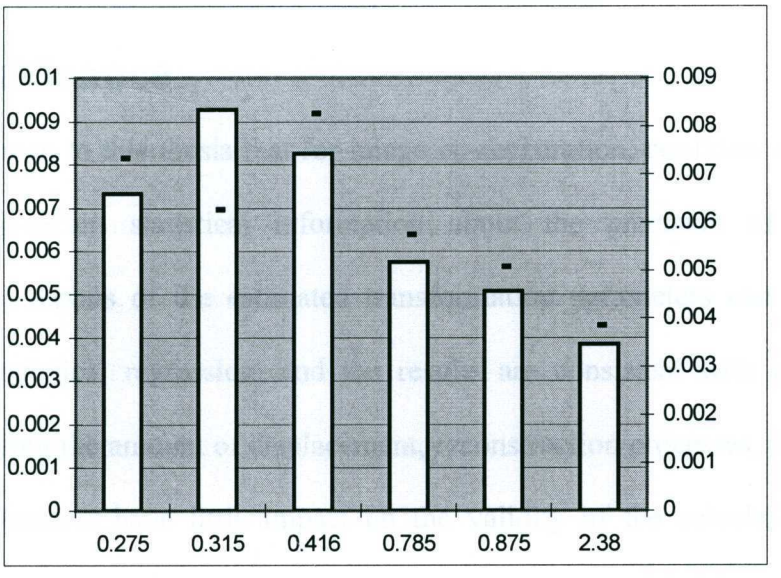


Figure 4.2: The calculated confidence intervals for translation (in pixel) and their errors. X-axis: pixel of translational displacements. Y-axis: mean and variance of the calculated confidence intervals.

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, as shown in Figures 4.1 and 4.2.

5. Conclusions

5.1 Significance

We have shown in this thesis that for image co-registration, confidence intervals can provide significant statistical information about the precision of results. The confidence intervals of the estimated transformation parameters can be calculated based on statistical regression, and the results are consistent with the simulation results. Varying the amount of displacement, reconstruction processes, noise levels, or tracer distributions have little impact on the validity of the calculated confidence intervals. After adjusted for systematic errors in the estimated residual sum of squares, confidence intervals can be calculated accurately even for very high noise conditions and with large distribution discrepancies between the two sets of images.

5.2 Applications to multi-modality registration and future works

Multi-modality image registration can be viewed as the process of finding the optimised transformation parameters by comparing image of one modality and another simulated image of another modality, or minimisation of pixel value variance between the two modality images. Thus, the confidence interval proposed by this thesis has the potential of being applied to multi-modality registration result assessments.

Based on this assumption, future studies can be proposed on new registration method, and the same precision analyse method can be applied the new method. Woods' method minimises variance in histogram domain. The assumption is pixels having the same gray value belong to the same tissue type, and, they will have similar gray

values in another image modality. But pixels having same values in one image modality do not necessarily belong to one same tissue. For example, pixels having gray value of 100 may not all belong to white matter in PET. Brighter pixels in PET may have darker values in MRI. As a consequence, the corresponding pixels in MRI image in the same locations may have various values, depending on the tissue type of those pixels. The variances of these pixels may not be minimised when registered.

So if we can register the two images sets base on local groups of pixels, ie, minimise the variance inside each of these groups, that will overcome Woods problem. There are plenty of methods to do local grouping, one of the simplest is to minimise the variance within each local groups. The precision analyse method proposed in this thesis is based on minimisation of variance, so it can by applied to both Woods method and the proposed new registration method without significant modifications.

References

N. M. Alpert, J. F. Bradshaw, D. Kennedy, and J. A. Correia (1990). "The principal axes transformation - A method for image registration", *J. Nucl. Med.*, 31:1717-1722.

B. A. Ardekani, M. Braun, B. F. Hutton, I. Kanno, and H. Iida (1995). "A fully automatic multimodality image registration algorithm", *J. Comput. Assist. Tomogr.*, 19:615-623.

J. Y. Chiang and B. J. Sullivan (1993). "Coincident bit counting - A new criterion for image registration", *IEEE Trans. Med. Imag.*, 12:30-38.

A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal (1995). "Automated multimodality image registration using information theory" In Bizais, Y. and Barillot, C. (eds), *Information Processing in Medical Imaging*, pp. 263-274, Kluwer Academic Pulishers, Dordrecht.

N. R. Draper (1981). *Applied regression analysis*, 2nd Ed., New York, Wiley.

R. E. Ellis, S. Tijsvug-Larsen, M. Marcacci, D. Caramella, and M. Fadda (1996). "A bicompatible fiducial marker for evaluating the accuracy of CT image registration", In Lemke, H. U., Vannier, M. W., Inamura, K. and Farman, A. G. (eds), *Computer Assisted Radiology*, vol. 1124, *Excerpta Medica-Int. Congress Series*, pp. 693-698, Elsevier, Amsterdam.

H. Erbe, A. Kriete, A. Jodicke, W. Deinsberger, and D. Boker (1996). "3D-ultrasonography and image matching for detection of brain shift during intracranial surgery", In Lemke, H. U., Vannier, M. W., Inamura, K. and Farman, A. G. (eds), *Computer Assisted Radiology*, vol. 1224, *Excerpta Medica-Int. Congress Series*, pp. 225-230, Elsevier, Amsterdam.

A. C. Evans, S. Marrett, L. Collins, and T. M. Peters (1989). "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*, vol. 1092, pp. 264-274, SPIE Press, Bellingham, WA.

J. M. Fitzpatrick, J. B. West, and C. R. Maurer, Jr. (1998). "Predicting error in rigid-body point-based registration", *IEEE Trans. Med. Imag.*, 17:694-702.

N. Hata, M. Suzuki, T. Dohi, H. Iseki, K. Takakura, and D. Hashimoto (1994). "Registration of ultrasound echography for intraoperative use: A newly developed multiproperty method", In Robb, R. A. (ed), *Visualization in Biomedical Computing*, vol. 2359, pp. 251-259, SPIE Press, Bellingham, WA.

M. Herbin, A. Venot, J. Y. Devaux, E. Walter, J. F. Lebruchec, L. Dubertret, and J. C. Roucayrol (1989). "Automated registration of dissimilar images: application to medical imagery", *Comput. Vision Graphics Image Process.*, 47:77-88.

D. L. G. Hill (1993). *Combination of 3D medical images from multiple modalities*, Ph.D. Thesis, University of London.

E. J. Hoffman, P. D. Cutler, T. M. Guerrero, W. M. Digby, and J. C. Mazziotta (1991). "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.*, 11:17-25.

C. K. Hoh, M. Dahlbom, G. Harris, Y. Choi, R. A. Hawkins, M. E. Phelps, and J. Maddahi (1993). "Automated iterative three-dimensional registration of positron emission tomography images", *J. Nucl. Med.*, 34:2009-2018.

M. Irani and S. Peleg (1991). "Improving resolution by image registration", *Comput. Vision Graphics Image Process.*, 53:231-239.

L. Junck, J. G. Moen, G. D. Hutchins, M. B. Brown, and D. E. Kuhl (1990). "Correlation methods for the centering, rotation and alignment of functional brain images", *J. Nucl. Med.*, 31:1220-1226.

S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi (1983). "Optimization by simulated annealing", *Science*, 220:671-680.

V. Leclerc, and C. Benchimol (1987). "Automatic elastic registration of DSA images". In *Computer Assisted Radiology*, pp. 719-723, Springer-Verlag, Berlin.

K. P. Lin, S. C. Huang, *et al.* (1994). "A general technique for interstudy registration of multifunction and multimodality images", *IEEE Trans. Nucl. Sci.*, 41:2850-2855.

L. D. Lunsford (1988). *Modern Stereotactic Neurosurgery*, Martinus Nijhoff, Boston, MA.

J. B. A. Maintz, and M. A. Viergever (1998). "A survey of medical image registration", *Med. Imag. Analy.*, 2:1-36.

D. W. Marquardt (1963). "An algorithm for least-squares estimation of nonlinear parameters", *J. Soc. Ind. Appl. Math.*, 11:431-441.

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

J. A. Nelder and R. Mead (1965). "A simplex method for function minimization", *Comput. J.*, 7:308-313.

C. A. Pelizzari, G. T. Y. Chen, D. R. Spelbring, R. R. Weichselbaum, and C. T. Chen (1989). "Accurate three-dimensional registration of CT, PET and/or MR images of the brain", *J. Comput. Assist. Tomogr.*, 13:20-26.

M. J. D. Powell (1964). "An efficient method for finding the minimum of a function

of several variables without calculating derivatives", *Comput. J.*, 7:155-163.

W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in C. The Art of Scientific Computing*, Cambridge University Press, New York.

D. A. Simon, R. V. O'Toole, M. Blackwell, F. Morgan, A. M. DiGioia, and T. Kanade (1995). "Accuracy validation in image-guided orthopaedic surgery", In *Medical Robotics and Computer Assisted Surgery*, pp. 185-192, Wiley, New York.

P. A. van den Elsen, E. J. D. Pol, and M. A. Viergever (1993). "Medical image matching – a review with classification", *IEEE Eng. Med. Biol.*, 12:26-39.

A. Venot, J. L. Golmard, J. F. Lebruchec, L. Pronzato, E. Walter, G. Frij, and J. C. Roucayrol (1983). "Digital methods for change detection in medical images", in Deconinck, F. (ed.), *Information Processing in Medical Imaging*, pp. 1-16. Nijhoff Publishers, Dordrecht.

R. P. Woods, J. C. Maziotta, and S. R. Cherry (1993). "MRI-PET registration with automated algorithm", *J. Comput. Assist. Tomogr.*, 17:536-546.

R. P. Woods, S. T. Grafton, C. J. Holmes, S. R. Cherry, and J. C. Mazziotta (1998). "Automated image registration: I. General methods and intrasubject, intramodality validation", *J. Comput. Assist. Tomogr.*, 22:139-152.