3-D DATA HANDLING AND REGISTRATION OF MULTIPLE MODALITY MEDICAL IMAGES

Mohammad Ali Oghabian

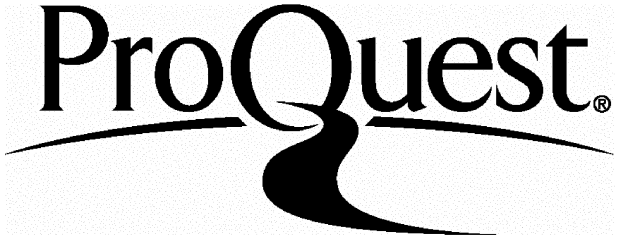
Department of Medical Physics

University College London

Thesis for the Degree of Doctor of Philosophy

University of London

1993



Abstract

The many different clinical imaging modalities used in diagnosis and therapy deliver two different types of information: morphological and functional. Clinical interpretation can be assisted and enhanced by combining such information (e.g. superimposition or fusion). The handling of such data needs to be performed in 3-D.

Various methods for registration developed by other authors are reviewed and compared. Many of these are based on registering external reference markers, and are cumbersome and present significant problems to both patients and operators. Internal markers have also been used, but these may be very difficult to identify. Alternatively, methods based on the external surface of an object have been developed which eliminate some of the problems associated with the other methods. Thus the methods which have been extended, developed, and described here, are based primarily on the fitting of surfaces, as determined from images obtained from the different modalities to be registered.

Annex problems to that of the surface fitting are those of surface detection and display. Some segmentation and surface reconstruction algorithms have been developed to identify the surface to be registered. Surface and volume rendering algorithms have also been implemented to facilitate the display of clinical results.

An iterative surface fitting algorithm has been developed based on the minimization of a least squares distance (LSD) function, using the Powell method and alternative minimization algorithms. These algorithms and the qualities of fit so obtained were intercompared. Some modifications were developed to enhance the speed of convergence, to improve the accuracy, and to enhance the display of results during the process of fitting. A common problem with all such methods was found to be the choice of the starting point (the initial transformation parameters) and the avoidance of local minima which often require manual operator intervention.

The algorithm was modified to apply a global minimization by using a cumulative distance error in a sequentially terminated process in order to speed up the time of evaluating

of each search location. An extension of the algorithm into multi-resolution (scale) space was also implemented. An initial global search is performed at coarse resolution for the 3-D surfaces of both modalities where an appropriate threshold is defined to reject likely mismatch transformations by testing of only a limited subset of surface points. This process is used to define the set of points in the transformation space to be used for the next level of resolution, again with appropriately chosen threshold levels, and continued down to the finest resolution level. All these processes were evaluated using sets of well defined image models. The assessment of this algorithm for 3-D surface registration of data from (3-D) *MRI* with *MRI*, *MRI* with *PET*, *MRI* with *SPECT*, and *MRI* with *CT* data is presented, and clinical examples are illustrated and assessed.

In the current work, the data from multi-modality imaging of two different types phantom (e.g. Hoffman brain phantom, Jaszczak phantom), thirty routinely imaged patients and volunteer subjects, and ten patients with setting external markers on their head were used to assess and verify 3-D registration. The accuracy of the sequential multi-resolution method obtained by the distance values of 4-10 selected reference points on each data set gave an accuracy of 1.44±0.42 mm for MR-MR, 1.82±0.65 for MR-CT, 2.38±0.88 for MR-PET, and 3.17±1.12 for MR-SPECT registration. The cost of this process was determined to be of the order of 200 seconds (on a *Micro-VAX II*), although this is highly dependent on some adjustable parameters of the process (e.g. threshold and the size of the geometrical transformation space) by which the accuracy is aimed.

Table of contents

CHAPTER 3: SURFACE FITTING

CHAPTER 4: SURFACE RECONSTRUCTION

CHAPTER 5: MINIMIZATION METHODS

CHAPTER 6: SEQUENTIAL METHODS

CHAPTER 7: MULTI-RESOLUTION REGISTRATION

CHAPTER 8: EXPERIMENTS, RESULTS AND DISCUSSION

CHAPTER 9: DISPLAY METHODS

CHAPTER 10: CONCLUSIONS

List of illustrations

List of tables

Acknowledgements

I wish to thank Mr. Andrew Todd-Pokropek for his supervision on this project and for his guidance and valuable effort throughout the course of this work, and in writing this thesis. I would also like to thank Dr. Iski Gorden, Alan Connely, and Nino Margetic, whose insights and experiences helped the assessment of the results, and thus contributed to this thesis. My thanks also go to Mrs F Pourheydari who helped me in typing this thesis.

The author would also like to thank all members of Computer unit in Rayne institute and the MRI unit in Great Ormond Street hospital, where the work was mainly carried out.

I wish also to thank all members of X-ray department in Great Ormond Street hospital, MRC unit in Hammersmith hospital, and Radiological Science department in Guys hospital, who supplied data and helped in generation of special data for the verification of the proposed methods.

Lastly, I would like to thank my family, parents, and friends, whose patient and moral support encouraged me to carry on this work.

CHAPTER 1
INTRODUCTION

1.1- General conception

The clinical imaging modalities in diagnostic, therapeutic and surgical fields deliver two major types of information: morphological information which represents the shape, borders and relational position of an object, and functional information which shows the uptake of a specified organ (object's activity). It is sometimes desirable to display the two kinds of information simultaneously in one picture. For example, mapping abnormal (e.g. tumour), or normal structures of an object obtained from different modalities such as *MRI* (Magnetic Resonance Imaging), *CT* (Computerized Tomography), *PET* (Positron Emission Tomography) and *SPECT* (Single Photon Emission Tomography) provides complementary information about an organ. Displaying morphological or functional information obtained at different time instant can also be of interest to be interpreted simultaneously.

Since these images are taken under different patient positioning and imaging parameters, interpretation of them, as reconstructed originally, does not provide an easy and accurate understanding of similarities and differences between them. The problem becomes more crucial where a clinician would like to map accurately volumes of interest from one study to the other, by which some surgical or therapeutical planning may be based.

For each original slice of any study, the corresponding slice can be reconstructed on another, if the relative geometrical transformation parameters between them are found. A process called *registration* is used to correct for the appropriate translation, rotation, and magnification differences between the images. By applying this process the geometrical image differences induced by the scene (patient and detector position) can be discarded and thus a better understanding of the object structural (anatomical or functional) differences can be obtained. The research described in this thesis is an investigation into the problems of registering the images in order to obtain some correlated (matched) images of an object, and displaying them.

The definition of the research problem and general aims of the project are presented in this chapter. Some clinical uses and applications of the process are given in the next section. The objectives of the research are presented in section 1.3. In section 1.4 an overview of the research and the organization of the thesis is outlined. The computer facilities used throughout this project are introduced in section 1.5.

1.2- Clinical application

Both surgical planning and radiation therapy, may require metabolic and physiological activity level as obtained by *PET* or *SPECT* to be correlated with bony or soft tissue anatomy from *CT* or *MRI*. A knowledge of lesion activity as well as the determination of accurate boundaries is an essential factor in volume measurement, internal dose calculation and radiotherapy planning. Functional information associated with particular brain structures (e.g. regions of interest) also enhance the interpretability of medical images. For example correlated diagnostic information can be obtained from quantitative analysis of the nuclear medicine images (*PET* or *SPECT*) and morphological investigation of *MRI* or *CT* images. Moreover, tumour regions which are well visualized in *MRI* images may need to be placed in the context of bone structure derived from *CT* scans.

In many clinical circumstances, the original 2-D slices are not ideal. The 3-D structures of medical objects are not easily interpretable from 2-D slices which usually lack some clinical problems such as size and shape of lesions. Looking at shaded object surfaces in relation to each other in a scene enhances the interpretability of the images within the surrounding 3-D anatomy.

1.3- Research objectives

The concept of image registration and data matching have been considered by a number of researchers (see chapter 2). Interest in registering medical images is extensive and also continuing in clinical applications. If image registration (or data matching) is to be usable in clinical assessments then it should be simple, relatively inexpensive, and do not include any extra constraint for patient during the imaging. If it is also to prove valuable judgments in clinical decisions, it must be accurate, free from subjective interaction error, and reliable in both

quantitative and qualitative measurements.

In recent years effort has been concentrated on improving the techniques which use the routine images taken without any special patient positioning or tedious setting of e.g. external markers on a patient. These techniques which are based on external surfaces of an object are also independent of noise and grey-level changes inside the object. However, the actual problems associated with an accurate and automatic registration of medical images remain unrevealed. There is also no reliable method to handle low resolution and noisy images (e.g. *SPECT* images).

The main objective of the research presented herein is to develop a suitable technique for correlation and display of tomographic brain images. The criteria which define the suitability of a registration algorithm should be determined, first.

A review of all available registration techniques and their limitations was done initially. It is important to assess all sources of errors incorporating these techniques and to overcome problems associated with them. The choice of the most suitable registration algorithm is one important issue to start with. Selecting or developing any registration algorithm, a number of related processes might be required to be explored (e.g. surface reconstruction and minimization). All these processes need to be improved by implementation and development of new approaches.

Developing an accurate and automated technique which is less dependent on manual interaction of an operator is of great interest and importance. It is essential to verify all the processes involved and measure their accuracy and timing response using some well-defined clinical images (models). The clinical utility of the registration and superimposition processes, and clinical values obtained by them should be considered. Finally, evaluating the computational aspects of the algorithm and overcoming programming obstacles which effect the applicability of the process in routine medical imaging and clinical applications are important research objectives.

1.4- Overview of the research

This thesis is organized into ten chapters discussing different related approaches of the

project. An overview of the research, the aims, and clinical applications of the registration and superimposition processes are presented in this chapter.

Chapter 2 addresses several methods which have been introduced in literature for image registration. Most of them are based on setting some internal or external landmarks on a patient. Since this is a significant problem for both patient and technician during imaging, a technique has been considered in the current work which is based on fitting together two models of external surfaces as obtained by the various modalities (surface fitting).

The general surface fitting algorithm as defined by previous workers is outlined in chapter 3. In this chapter, a least square distance function (LSD) is constructed to evaluate the fit between surfaces (as extracted by methods described in chapter 4). Each variable (transformation parameter) of this function is updated using a minimization process in order to minimize the distance between the two surfaces and to find the relative transformation between them. The new approaches to improve the surface fitting technique and overcome the obstacles associated with that are also presented.

Surface reconstruction, as presented in chapter 4, is the first requirement of the surface fitting process. Thus, consideration is initially given to the task of detecting surfaces subjected to noise and other characteristic differences (e.g. grey level difference). Filtering of noisy data is applied as a preprocessing stage and a 3-D binary surface image is created for each data set. New approaches are proposed for extracting surfaces of noisy images (e.g. SPECT brain images).

Different approaches to the minimization process together with the choice of a suitable algorithm are assessed in chapter 5. Using a new global minimization technique is suggested in chapter 6 which constructs a global search (i.e. grid search) through all possible transformation parameters instead of minimization in a downhill trend as implied by most minimization methods (e.g. direction set methods). The proposed technique is a type of sequential process employing a cumulative distance error and a variable threshold to speed the performance of the grid search.

To improve the computational cost, a new algorithm is proposed in chapter 7 which is based on a multi-resolution scale based algorithm. The computational implementation of the process is presented, and the accuracy and cost measurements are assessed.

The experimental results obtained by using sequential and multi-resolution processes are outlined in chapter 8. In this chapter, the imaging parameters and data types used for the verification of the techniques are introduced. Based on these data, the techniques are verified and different factors influencing the performance are presented. A probability analysis is also performed to investigate the effects of different threshold level and various number of sample points at each resolution levels.

The problem of superimposition is addressed in chapter 9. Some of the commonly used shading and display techniques are first outlined and their strengths and weaknesses in displaying clinical images are discussed. Different approaches of overlaying the correlated images and displaying them are then presented.

A summary of the research objectives and the results obtained by this project are given in chapter 10. Suggestion for further works and extension of the research are also presented in this chapter.

1.5- Instrumentation and computer facilities

The workstation on which most of the project was done, consists of two main components. A *Kontron* image processing machine with 16-Mb dynamic frame memory, and a 5-Mb *Micro-VAX 11* which is interfaced with the *Kontron* machine in order to increase the capability of system in terms of programming. The *Kontron* system is not programable and its own routines lack some essential programs such as 3-D data handling and geometric transformations required for registration and display. There is also a link between the *Kontron* and a *VAX* computer which runs the *Siemens MRI* imaging system.

The *Micro-VAX* system has been installed with *Fortran* and *Pascal* compilers. It is possible to call most of the *Kontron* routines via *C*-defined-library subroutines and *Fortran* wrappers. However, the actual registration programs and display routines need to be written in *Fortran* on the *VMS*-based *Micro-VAX* system (version 4.5). All the images are displayed on *Kontron* display system.

CHAPTER 2
REGISTRATION METHODS

2.1- Introduction

In order to interpret and diagnose a clinical problem, different images might be required. These images are usually taken in different geometric environments (i.e. with the patient in a different position, or at a different time, or alternatively using different imaging modalities). These imply geometrical differences such as translation, rotation, and scaling between them. In most clinical cases, it is useful to know exactly some corresponding clinical information between the different types of images taken from one object.

Registration is defined as finding corresponding points or regions in two or more images of a single object or scene. The ability to access these corresponding points (pixels in 2-D or voxels in 3-D) is the most essential factor for data alignment and, in turn, for the generation of a superimposed image. In this respect, any geometric difference should be revealed between two sets of images recorded under different patient positioning or imaging conditions. Moreover, intensity differences, geometric distortions and noise which typically occur in each data set, independently, need to be taken into account. Accurate registration enables potential use of different imaging parameters, time-varying information and alignment of anatomical and functional information in a superimposed image. These add clinical values in most diagnostic and therapeutic applications.

The aim of this chapter is to review the most common registration methods and to outline the grounds based on which the methods used in this project were selected. Different registration methods are addressed in section 2.2 and 2.3. grey level based methods are briefly addressed in section 2.2. Various feature based registration methods are outlined in section 2.3. Section 2.3.1.2 describes the use of external landmarks in the registration of medical images. The use of edge-based (e.g. surface fitting) methods as a potential and widely used method of registering complex medical images are explained in section 2.3.3. Section 2.4 addresses a least-square correlation algorithm for the evaluation of match between the two images based on their corresponding

points. The application of the transformation parameters are also represented as a matrix notation in this section. A least-square algorithm is developed as a multi-parameter minimization problem which can be solved by the use of a non-iterative algorithm. Two non-iterative algorithms are defined in section 2.4.1 and 2.4.2. The definition of iterative algorithms (using a local or global minimization process) as required for some types of registration methods are left for chapter 5. Various factors influencing the selection of a reliable registration algorithm are summarized in section 2.5. A summary of the methods and their properties are given in section 2.6.

2.1.1- Current registration methods

Any registration technique attempts, firstly, to find a number of corresponding pairs of points (control points or pixels) in different images of a single object. Secondly, the geometric differences between the pairs are computed either by minimizing a dissimilarity function e.g. a distance measure, or by maximizing similarities between them. The different types of registration

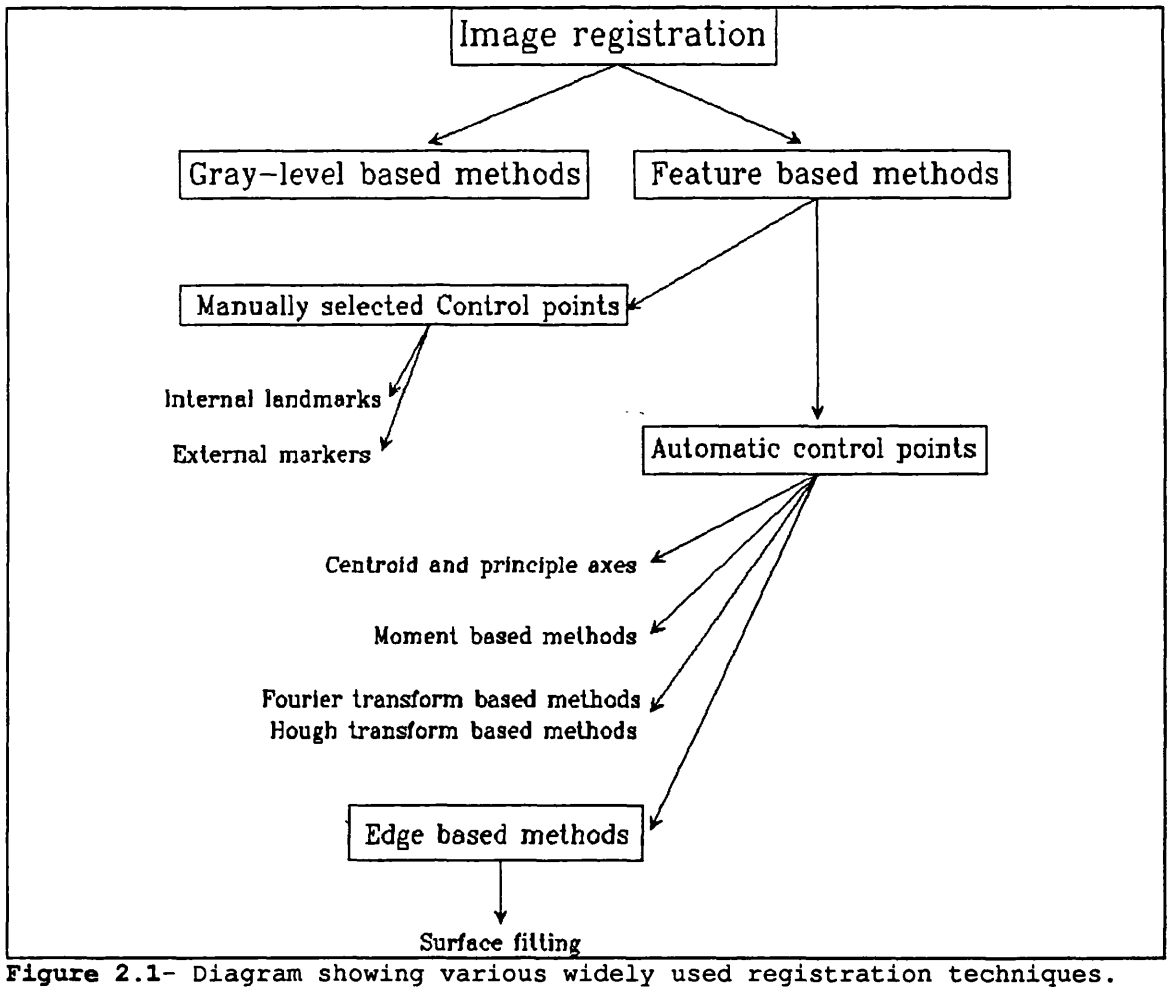


Figure 2.1- Diagram showing various widely used registration techniques.

are based on the type and level of object description (e.g. representation) at which the matching is attempted and thus based on the type of corresponding pairs. The schematic diagram in figure 2.1 shows some of the widely used registration methods. Two general approaches of registration known as *grey-level-based* and *feature-based* matchings are reviewed in the next sections.

2.2- Grey level based matching

Pairs of corresponding points (or regions) between the two images can be found by matching the image values (pixel intensities) using some forms of correlation. This is also known as a *signal-based* registration method. A similarity criterion such as a correlation coefficient, correlation function (Barnea & Silverman 1972, Lillestrand 1972, Pratt 1974, Wong and Hall 1979, Appledorn et al 1980, Ros et al 1984), or the sum of absolute values of the differences (SAVD) (Svedlow et al 1978) can be performed on 2-D grey scale images. The grey level based methods are sensitive to the variation in pixel properties, which may lead to misregistration of complex images due to their intensity variation. Thus, their use is limited in matching of medical images whose intensity fluctuate greatly under different imaging conditions. The problem became severe due to the variable detectability of an object (e.g. tumour, or contrast medium used in radiography) imaged under different imaging modalities.

Recently, a technique which uses intensity value of the voxels, and minimizes the variance of the ratio of intensity (VIR) in two images, was introduced by Wood et al 1992. The technique was originally applied on images from the same modality (PET). It is very sensitive to partial volume effects, resolution (axial resolution) and uneven change in intensity of the two modalities (e.g. tumours which are visible in only one of modalities). Some tissue classification by intensity ranges has been applied (Hill et al 1993) to overcome some of the problem associated with this technique. However the occurrence of local minima (see section 5.1) are not avoidable in this matching process.

2.3- Feature based registration

An alternative method which is more resistant to imaging condition changes, is based on the use of some image features (e.g. contour segments, high curvature points, angles) which characterize the shape, position and arrangement of an object. The features participating in

registration should correspond to the regions in the image which are constant under different imaging conditions. In more complex images taken under very different imaging and geometric conditions, application of a signal based matching process on a close initial alignment obtained by a feature based registration may be required for an acceptable registration system.

These features are more likely to be invariant to imaging changes (e.g. intensity differences), when a higher level of description is derived, precisely. In this respect, registration and matching at the level of object boundaries has its own advantage over signal based methods. However, deriving these object descriptions from some segmentation methods is not trivial and can lead to incorporating error and deficiency.

In the next subsections, various feature-based methods which are widely used for the registration of 2-D or 3-D images are outlined. The features employed in these methods are either obtained manually under the interaction of a human operator, or automatically by a machine operation.

2.3.1- Methods based on Control points

One of the widely used approaches to registering 3-D volume data is based on the use of a (small) set of points defined by internal or external anatomical landmarks on the image. In this respect, the points which have a constant relationship to the patient anatomy in both studies are used.

2.3.1.1- Internal landmarks

Particular internal points such as points having specific location can be obtained manually (Merickel, Carthy 1985, Kessler and Chen 1987). William et al 1978, Singh et al 1979 and Frei et al 1980 have also implemented techniques by which registration is obtained by a relative geometric transformation of some manually selected *internal landmarks* between the two images. In this respect, the approach taken by Singh et al 1979 was to compare the relative position of landmarks by, visually, estimating their relative location, and correcting the estimates by maximizing a cross-correlation coefficient function. This correlation function is applied between a window surrounding the landmark in one image and several locations adjacent to the landmark in the other image. They used these techniques efficiently for registration of 2-D *Nuclear*

medicine images (*myocardial perfusion studies*) taken at different times and under a varied physiological conditions.

These techniques are user dependent and their accuracy is due to the operator interaction in selecting the control points. An alternative control based method which is less dependent on the operator decision for the selection of control points uses *external landmarks*. However, these methods still involves some manual interactions due to placing of the external landmarks on the patient.

2.3.1.2- External landmarks

Fiducial markers as part of a *stereotactic frame* (Kall et al 1985), or some point sources attached to the patient (Price et al 1975, Fleming 1984) are used as *external landmarks* for the registration. The stereotactic system was used originally (Goerss et al 1982 and Kelly 1984) for localization and transposition of points and lesion volumes in *CT*. The technique was augmented, by Kall 1985, for transposition of both *CT* and *MR* data into a common 3-D stereotactic space, thereby registering them. Fox et al 1985, also, applied stereotactic method for anatomic localization of *PET* images. Since the relation between the original tomographic and the stereotactic reference coordinate system is established by this method, the correlation and alignment of the images from different modalities can be easily obtained.

External point sources visible in both studies can be attached to the patient during imaging by both modalities. In this approach, three or more corresponding points between the two studies are identified. Once the set of point-to-point correspondences is established, the transformation differences can be derived between the two sets of points and consequently between the two data sets (Chen and Kessler 1985). Radio-opaque materials (e.g. lead beads) in *CT*, *CuSO4-doped-water* capsules in *MR* and radioactive-filled capsules in *SPECT* are ideal widely-used *fiducial markers*. In general, positioning and fixing the landmarks on a patient is not technically easy and it is inconvenient for both technician and patient.

2.3.2- Automatic control points and features

Some techniques have been developed using *control points* computed from specific common features of both sets of images without any user participation. Using *centroid, major*

and minor axes (Chow and Aggarwal 1977, Mitiche and Aggarwal 1983), or *invariant moments* (Hu 1962, Sadjadi and Hall 1978, Maitra 1979, Abu-Mostafa and Psaltis 1984) are some examples of these methods.

The *control points* are defined to be; **1)** shape-specific picture points with respect to the relevant transformation of the problem, **2)** sufficiently apart to tolerate addition of noise in the location of them, **3)** easy to compute, and **4)** have as much information about the shape as possible. Various methods have been introduced in the literature satisfying these requirements differently. *Centroid registration, moment based methods, Hough transform* and *Fourier transform methods* are examples of the machine based methods as described in the following subsections.

Computer generation of control points is often not suitable for the complex 3-D medical images. These features are highly affected by image artifacts and noise which are very common in medical images. Moreover, the 3-D structures contained in the images obtained by different modalities are not exactly the same, whereas this is an essential requirement of this registration method.

2.3.2.1- Centroid and principal axes

Registration of two planar images (figures) was done as a new approach by Mitiche & Aggarwal 1983, using points that are specific to the shape with respect to transformation of interest (e.g. centroid). The *centroid* is defined as

$$C_x=\Sigma\Sigma\Sigma\ x.f(x,y,z) \,/\, (\Sigma\Sigma\Sigma\ f(x,y,z)) \quad \text{which is } \Sigma x/n \text{ in binary images,}$$
$$C_y=\Sigma\Sigma\Sigma\ y.f(x,y,z) \,/\, (\Sigma\Sigma\Sigma\ f(x,y,z)) \quad \text{which is } \Sigma y/n \text{ in binary images,}$$
$$C_z=\Sigma\Sigma\Sigma\ z.f(x,y,z) \,/\, (\Sigma\Sigma\Sigma\ f(x,y,z)) \quad \text{which is } \Sigma z/n \text{ in binary images,}$$

Equ. 2.1

where f(x,y,z) is the voxel value of a 3-D image and x, y and z are positional information of the voxels. n is the number of voxels covered by the object in each orthogonal direction.

The *unweighed mean point* (centroid) and *radius weighted mean point* used by Mitiche & Aggraval are shape-specific points in spatial domain and describe shapes independently of the coordinate system. These points were used to register two images of a particular object. After

registering images in their method, applying a matching process (e.g. template matching) was also suggested, in order to check how closely the boundaries matched and to search for any further (minor) transformation.

The *principal axes* are defined as the best straight line fitted to a shape (like a Chromosome as defined by Paton 1970)(see section 4.5.1). These axes have also a property of being independent of the coordinate system. The *modified major and minor axes* (two lengths along and across the principal axes), the area, and the angle between the major axes and the vertical or horizontal axes are other descriptors used for registration (Chow and Aggarwal 1977). Points on an object that are furthest or closest from the centroid, the point whose distance from the centroid is equal to the medium of distances of all points on the shape from centroid have also been suggested for the registration problem.

The main problem with matching using the above shape-based registration techniques is that the registration can not be applied when only partial information about shapes are available.

2.3.2.2- Moments based methods

It was suggested by Hu 1962 to represent an image by an infinite set of moments. Denoting the image plane distribution by f(x,y), the moments (of order j, k) can be obtained by

$$M_{jk}= \sum\sum f(x,y).x^j.y^k. \qquad \text{Equ. 2.2}$$

It is believed that only a few of these *moments* are necessary to represent most shape characterizations of an object (Teague 1980). These finite set of *moments* were shown to contain closely the same shape information as the original image (using inverse moment problem). Although lower moments characterize the centroid of the data, the principal axes and the size and orientation of the image, the discrimination properties of a complex patterns would be increased if higher moments are also used.

A set of moments which are invariant to geometric transformation (size change, translation, rotation) are desirable for general pattern recognition and registration. *Invariant moments* can be obtained by normalizing the moments in respect to size, *centroid, principle axes* or other shape specific functions (Maitra 1979, Abu-Mostafa and Psaltis 1984, 1985). As shown

by Hu 1962, the set of moments normalized in size can still characterize a particular pattern (object), where they are themselves independent of the pattern position in visual field and of the pattern size. The representation of a pattern can be converted to the representation of points in a n-dimensional moment space where n is the number of invariant moments with respect to for example the centroid or principle axes.

The main problem of matching images using the moment methods is that, like other shape-specific methods mentioned earlier, registration based on partial information is not possible. Therefore, the process is inefficient for registering 3-D medical images which create image data set not mutually complete.

2.3.2.3- Fourier Transform based methods

The problem of estimating the 3-D motion parameters was presented by Lin et al 1986 in an approach which does not require the establishment of the initial points correspondence between the two images. They attempted to find the motion parameters based on the *Fourier transforms* of the two images. Each image data set is firstly shifted to its own *centroid.* Then the rotation matrix R (see section 2.4.1) is found by correlating the values of the two Fourier transforms which are related by the same rotation matrix as the original images. In principle, a search strategy through all possible rotation matrices should be established to determine the rotation parameters. Since the centroids of the two data sets are also governed by the same rotation matrix and translation vector, the estimated rotation matrix R can be used to determine the translation vector by a centroid correlation function.

Registration of translated and rotated images using *Fourier transforms* were also suggested in a different approach by Castro and Morandi 1984 and 1987. The algorithm was based on the phase correlation technique developed from a purely translational displacement problem introduced by Kuglin and Hines 1975. The Fourier transforms of an image f(x,y) and its translated and rotated replica $f_t(x,y)$ in a 2-D case can be related by

$$F_t(\xi,\eta)=e^{-j2\pi(\xi x0+\eta y0)}\, F(\xi\cos\theta+\eta\sin\theta\ ,\ -\xi\sin\theta+\eta\cos\theta),$$

Equ. 2.3

where x0 and y0 are the translations in x and y directions, and θ is the angle of rotation. The phase (exponential form) in the above correlation is obtained when the actual angle of rotation

found. The angle θ is first determined using a directed search strategy and, then, the translation x and y is determined from the phase correlation (Kuglin and Hines 1975).

Other types of operators such as *Hough* based transformation was used by Yam & Davis 1981, and extended by Davis 1982, to register images automatically. A set of unique control points are chosen (e.g. local maximum of the curvature measures of a binary edge image) in one image. The *Generalized Hough Transform* (*GHT*) operator then attempts to find the matching feature in another image.

The major disadvantage of the Fourier-based methods which restricts their use for the registration is the requirement of two identical image volumes. Therefore, it is very difficult and even in some cases not possible to register medical images whose parts (regions) are not usually identical in different imaging modalities.

2.3.3- Edge based methods

Other frequently used registration methods are based on *edge* information obtained in a preprocessing stage. Only the edge information needs to be registered in this approach. Thus, the edge extraction step is the most critical part of this process.

These methods have been used by a number of workers (Andrus and Campbell, 1975, Svedlow and Gillem, 1978, Medioni and Nevatia, 1984, Bhanu and Faugeras, 1984). Similarity measurement was used by them to evaluate the fit between the images. Venot et al 1984 introduced a different class of similarity measure for registering simple 2-D images. This technique was based on the calculation of the number of *stochastic sign changes* (*SSC*) in the digitized subtraction angiographic images. A precise edge detection technique is required in order to have a reasonable maximum for the similarity measure at match location.

Applying a distance function as a measure of misregistration overcomes the problem associated with the requirement of a precise edge detection technique. The use of distance function enables the definition of the edges as a symbolic representation (e.g. edge segments) as proposed by Bhanu & Faugeras 1984, Price & Reddy 1979, and Medioni & Nevatia 1984. This type of edge definition technique can be applied to noisy edges in order to reduce the error

associated with the edge detection technique.

Surface fitting algorithms have been derived directly from *edge-based* methods in 3-D space. It was implemented by Pelizzari and Chen 1987 who used models of external surfaces obtained from various scans (also see Chen et al, 1988, Pelizzari and Chen, 1989). In this approach, a least squares distance (LSD) function, is constructed to evaluate the fit between the surfaces computed from the two sets of images. The mean value of the LSD is then minimized by a local minimization algorithm (e.g. Powell's gradient convergent method), in order to determine the best transformation between these two data sets.

The edge-based methods overcome some of the difficulties associated with the other registration techniques. Firstly, there is no need for special precautions for subject positioning during imaging. Secondly, the technique works well for registering brain images (using skin or brain surfaces), even when there are some differences in the internal structures of the original images. The technique has a reasonable potential for accuracy, reproducibility, and ease of use which are important factors of data fitting. However, the technique suffers from the presence of multiple local minima. To overcome this, the initial transformation parameters need to be chosen by a human operator, and modified during the minimization process, in order to align the two surfaces, initially, as closely as possible.

An alternative edge-distance based registration method which has been proposed by Barrow et al 1977, and improved by Borgefors 1984 and 1988, is based on distance transform (e.g. Chamfer transform) matching. This technique was used and modified recently by Hill et al 1993 for surface fitting purpose. As for other surface based methods, it needs a good start hypothesis of the transformation to bring the surfaces into correspondence, otherwise it suffers from trapping into local minima.

2.4- Least-square fitting of two images

Least-square fitting is defined as a multi-parameter minimization function which shows the mismatch between the two images. The general *L2 norm* least square function can be expressed as

$$\Sigma^2 = \Sigma\,(\mathbf{P1'}_i - \mathbf{P2}_i)^2, \quad \mathbf{P1'}_i = \mathbf{P1}_i[\mathbf{R}][\mathbf{S}]+\mathbf{T}, \quad \text{for } i=1,\ldots,n \qquad \text{Equ. 2.4}$$

where **P** and **P'** represent the corresponding points in the two registering images. **P1'** represents the points which are geometrically transformed and n is number of registering points. **R, T** and **S** are rotation, translation (shift) and scaling parameters, respectively.

There are two approaches of least square fitting being *iterative* and *non-iterative* methods. In next two sections, non-iterative method are described as applied to a number of control points (e.g. external landmarks) obtained before the registration. The correspondence between the points used in a non-iterative method should be known in advance. An Iterative least square function can be solved by using a local or global minimization process. This type of least square fittings as applied to the general surface fitting algorithm is addressed in chapter 5.

2.4.1- Analytical methods

If the coordinates of a number of corresponding points (four are the minimum required in 3-D cases; see equation 2.5) are known between the two images, the coordinate transformation can be defined analytically. *Analytical techniques* for finding the geometric differences between two sets of data are based on simultaneous equations which define the mapping of coordinates of one study on the other. This coordinate mapping can be expressed by the following equations.

$$\begin{aligned} x2 &= a_1x1+a_2y1+a_3z1+a_4 \\ y2 &= b_1x1+b_2y1+b_3z1+b_4 \\ z2 &= c_1x1+c_2y1+c_3z1+c_4 \end{aligned}$$

Equ. 2.5

Where (x1,y1,z1) are the coordinate of the original image and (x2,y2,z2) are the coordinate of the second image. The a's, b's and c's are parameters which are to be determined for the corresponding image points (e.g. corresponding markers) to be transformed. These parameters contain all translational and scaling information as well as the values of Sine and Cosine corresponding to the rotational differences between the two images. The parameters can be expressed in matrix notation as

$$[x2,y2,z2] = [x1,y1,z1][\mathbf{R}][\mathbf{S}][\mathbf{T}] \qquad \text{Equ. 2.6}$$

where **R, S** and **T** are Rotation, scaling and translation matrices, which can be represented as

$$[x2,y2,z2] = [x1,y1,z1] \begin{bmatrix} R_{11} & R_{12} & R_{13} & 0 \\ R_{21} & R_{22} & R_{23} & 0 \\ R_{31} & R_{32} & R_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ T_x & T_y & T_z & 1 \end{bmatrix}$$

Equ. 2.7

In the above representation, the rotation matrix is the concatenation of three matrices which represent rotation in three orthogonal direction x, y and z. The whole transformation can then be achieved with a single transform matrix, as follows, representing the parameters *a*'s, *b*'s, and *c*'s defined in the above equations.

$$[x2,y2,z2]=[x1,y1,z1] \begin{bmatrix} S_xR_{11} & S_yR_{12} & S_zR_{13} & 0 \\ S_xR_{21} & S_yR_{22} & S_zR_{23} & 0 \\ S_xR_{31} & S_yR_{32} & S_zR_{33} & 0 \\ T_x & T_y & T_z & 1 \end{bmatrix} = [x1,y1,z1] \begin{bmatrix} a_1 & b_1 & c_1 & 0 \\ a_2 & b_2 & c_2 & 0 \\ a_3 & b_3 & c_3 & 0 \\ a_4 & b_4 & c_4 & 0 \end{bmatrix}$$

Equ. 2.8

In the case of four points (markers), twelve equations in the form of the above equations would be generated which can be solved for the parameters.

2.4.2- Non-iterative least-square fitting using SVD

A *non-iterative* algorithm for least square fitting of 3-D images, based on the *Singular Value Decomposition*[1] (*SVD*), was implemented by Arun et al 1987. Two sets of 3-D corresponding points $\mathbf{P1}_i$ and $\mathbf{P2}_i$ (where i=1,2,...,n; and n is the number of points) are used to determine the rotation and translation differences between the two images. The method assume the correspondence between the two sets to be known before registration. By translating the centroid of each data set to the origin of its own coordinate system, all the points are assumed to be corrected in terms of translational shift. The least square function Σ^2 of equation 2.4 is then expressed by

[1] SVD: Any M*N matrix **A** whose rows are greater than or equal its columns, can be decomposed as an M*N column-orthogonal matrix **U**, an N*N diagonal matrix **W** (with positive or zero element) and the transpose of an N*N orthogonal matrix **V**.

$$\Sigma^2 = \Sigma^2 \ |\mathbf{q1}_i\mathbf{R}-\mathbf{q2}_i| \ = \Sigma \ (\mathbf{q1}_i\mathbf{q2}_i + \mathbf{q2}'_i\mathbf{q2}_i - 2\mathbf{q2}'_i\mathbf{Rq1}_i),$$

Equ. 2.9

where $\mathbf{q1}_i=\mathbf{P1}_i-C1$, $\mathbf{q2}_i=\mathbf{P2}_i-C2$ are the points corrected for centroid difference, $C1=1/n(\Sigma\ \mathbf{P1}_i)$ and $C2=1/n(\Sigma\ \mathbf{P2}_i)$ are the centroids of two data sets.

By *SVD* of a 3*3 matrix $H=\Sigma\ (\mathbf{q2}_i\mathbf{q1}'_i)$ (where *t* denotes *matrix transposition*), the rotation matrix **R** can be derived as $\mathbf{R}=\mathbf{VU}^t$ (see Arun 1987), where **V** and **U** are 3*3 orthogonal matrices (see footnote 1). Having found the rotation matrix **R**, the actual translation **T** can be derived by $\mathbf{T}=\mathbf{P2}-\mathbf{P1}[\mathbf{R}]$ (see equ. 2.4).

2.5- Selecting a registration algorithm

The selection of a registration algorithm is based on the type of the images to be registered, the information content in them and various factors influencing the reliability of the process. These factors can be characterized for the registration of brain data sets as follow.

1) The algorithm should be useable with the routine medical images obtained by current imaging modalities. The need to arrange for special imaging conditions may restrict its use in clinical applications.

2) The technical performance of the algorithm should be easy and not involve a practical difficulty and tedious manual intervention. In this respect, the techniques which do not require a special precaution such as patient positioning during the imaging are preferable.

3) The implementation aspect of the algorithm on a computer is very important in terms of feasibility and the amount of processing time required for the solution. Different algorithms involve various complexity of the registration process. The algorithm must be performed at an acceptable speed.

4) The accuracy of the process depends on the quality of the information used for the registration. The information and data obtained by a computer are usually more robust than those handled by a human being. As the involvement of an operator to interact reduced, the efficiency

and the accuracy of the process can also increase. However, the manual interaction is sometimes useful for example when the correctness of a match needs to be controlled and tested by an operator during a visual inspection (see section 3.2.6).

5) The type and amount of data existing in both the original images are two important factors in the choice of an algorithm.

6) The reproducibility of the technique is also very important for the repetition of the process, without any patient inconvenience.

7) There are often grey-level changes and geometric distortions which are not due to misregistration of the two images. The ability to compensate for these differences is important in matching process. In this respect, the methods which avoid the use of internal grey-scale structures are preferred.

After reviewing all registration algorithms and by considering all the above criteria, an edge-based method (surface fitting in 3-D) was selected. In this approach, a distance function was used to evaluate the fit between the two surfaces. The minimization algorithms used so far for surface fitting approach suffer from the presence of multiple local minima. The minimization process could easily fail as a result of these minima.

Attempts were made to overcome the deficiencies of the surface fitting algorithm as defined by previous workers. The need for implementing a new minimization strategy was an objective in the surface fitting approach. A number of modifications were introduced to make the process useable in most of the imaging modalities. Some new approaches were also developed to further increase the feasibility of the surface fitting in terms of accuracy and time (see chapter 3).

2.6- Summary

Having introduced the various common registration algorithms, the applicability of them for use with medical images was described in this chapter. The advantages and disadvantages

of each algorithm were also described in the relevant sections. Among all these methods, an *edge-based* algorithm based on a *least-square-distance* matching was chosen as a technique which meets most of the requirements for the registration of medical images. This algorithm minimises the sum of square-distances between the two surfaces obtained from the original grey scale images. The minimization is applied over sets of geometrical transformation parameters which indicate how one surface is allowed to be transformed in order to obtain a match with the other surface.

The general surface fitting algorithm is outlined in chapter 3, and the problem of surface detection is presented in chapter 4. In chapter 5, different minimization algorithms and their use in the least-square fitting are reviewed.

CHAPTER 3
SURFACE FITTING

3.1- Introduction

A brief review of various registration algorithms was given in chapter 2. Among the different registration techniques introduced in this chapter, *surface fitting* was used to register two sets of brain data obtained under different imaging modalities. The fundamental grounds used to select this registration method were: its ease, applicability in the routine imaging, and power of the algorithm to achieve the result regardless of the internal structure of the object. The surface fitting algorithm was first implemented by Pelizzari and Chen 1987 on brain medical images. The basic concepts of the method remain the same in this project. However, an attempt was initially made to overcome all the obstacles believed to exist in this method. Some new approaches were also suggested mainly with respect to minimization algorithm to facilitate the technique for use in the routine clinical applications.

The surface of an anatomical structure visible in different scans of interest is used to describe the patient-specific geometrical information inside the coordinate system of each imaging modality. These models of surface can be transformed geometrically in a spatial *Euclidean* (Cartesian) coordinate system in order to be brought into a 'fit' position. Since anatomical surfaces (e.g. brain or head) are not symmetric (e.g. sphere or cylinder), a unique transformation is expected based on matching of the two surfaces. It is assumed that two surfaces and the images from which they are derived can be scaled to the same size based on the prior knowledge of data acquisition and imaging parameters, and thus scaling is not part of the fitting procedure. Correction of both data sets for pixel size prevents any geometric inconsistency which might lead to misregistration.

The aim of this chapter is to describe the most commonly used method of registration known as surface fitting. The algorithm and some results obtained by applying various aspects of the process are presented in this chapter. Section 3.2 deals with the algorithm, in general, as implemented by other workers (e.g. Pelizzari and Chen 1987, Chen et al 1988). Various sub-

sections of this section outline the requirements and programming steps of the algorithm. The general mismatch function as will be used throughout this project is defined in this section. This is a distance error measure between the two registering surfaces as expressed in sub-section 3.2.5. The remaining problem is to minimize this function in order to find the best global match between the two data set. A number of modifications was employed, addressed in section 3.3, giving some new approaches to make the algorithm suitable for a routine clinical application. Section 3.4 outlines the types of data used for registration and the characteristics of imaging systems from which the data were acquired. The data types which have been used as known models for assessment and evaluation of the algorithm are presented in section 3.5. The expected accuracy of using different data types are also presented in this section. Section 3.6 presents two types of results: first, the behaviour of the distance function to different parameters (e.g. transformation parameters), and second, the accuracy and cost measurements obtained by applying different minimization algorithms. At last, a summary is given in section 3.7.

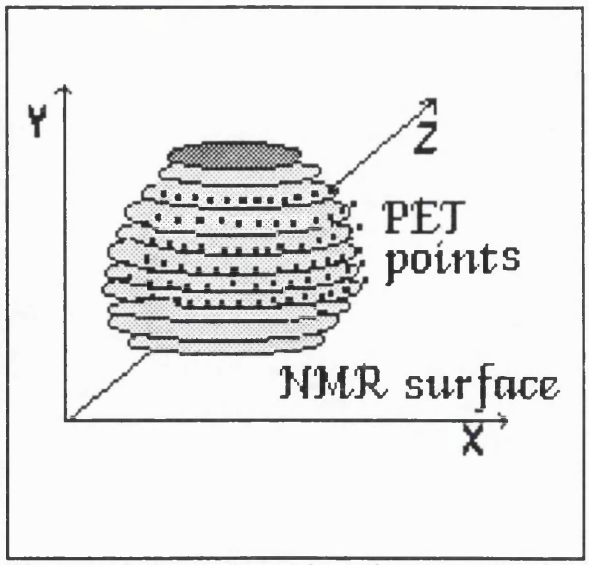


Figure 3.1- Schematic diagram showing the solid model of the objective surface (e.g. NMR) and sample points from the transforming surface (e.g. PET) (fitting model).

3.2- Algorithm

The original 2-D or 3-D grey level image data sets are processed and then geometrically scaled in a preprocessing stage. The external surfaces of both registering objects are generated by application of a surface detection algorithm on the serial slices of the pre-processed data set (see chapter 4). Then, the binary surfaces (being the set of voxels lying on the surface as defined by the surface detection algorithm in 3-D space) are used as primitives to be registered.

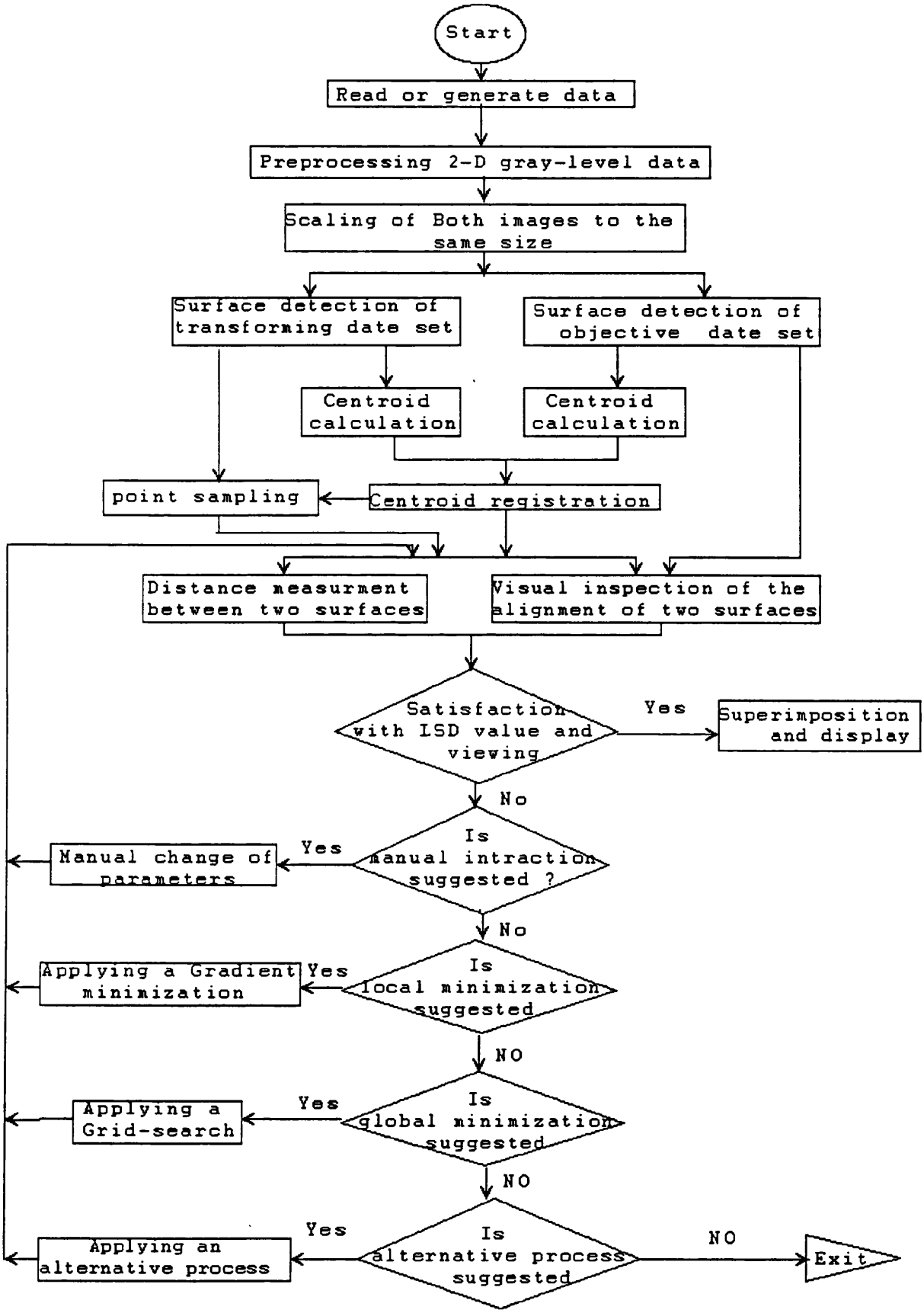


Figure 3.2- Diagram of the surface fitting algorithm showing different steps of the process. LSD denotes the least square distance between the two registering images.

The surface to be fitted is referred to as the objective surface $\mathbf{P}_o$ in the current work. The surface which is transformed at each stage of the fitting process is called the transforming surface $\mathbf{P}_t$. For the objective surface, a 3-D binary surface (solid model) is used for the surface fitting (see figure 3.1). However, in order to save computation effort, only a series of 3-D points selected from the 3-D surface are used for the transforming surface.

The distance between the two surfaces are measured as the mismatch value between them. A Least Square Distance (LSD) function (see section 3.2.5) is defined to evaluate this distance and to minimize the misregistration between the two surfaces.

Figure 3.2 shows a block diagram of the surface fitting algorithm as used in the current work. As shown in this figure, the algorithm consists of several steps: surface detection, centroid and manual registration, point sampling, point correspondence (using an intersection process), distance measurement, visual inspection, minimization and finally evaluation of the mismatch between the two surfaces. The various stage of the fitting process are discussed in the following sections.

3.2.1- Centroid registration

The centroid of each binary surface image is calculated using the information obtained in the surface detection process. The centroids of both data sets are transformed to the origin of the coordinate system of the objective surface (centroid registration). This new coordinate system of the objective surface is considered as a world reference coordinate system. Alignment of the two centroids brings the two surfaces into a close position which makes possible the application of the surface fitting algorithm as described below. This process is shown schematically in figure 3.3.

3.2.2- Manual registration

Manual registration is based on two types of information. These are the knowledge of geometrical information and parameters obtained by the imaging system, and the visual judgement obtained by a viewing screen. In this respect, the knowledge of the coordinate system in which the images are generated, and the approximate position of the starting and ending slices are examples of the first type of these information. Knowing these information are essentially required for a proper registration, especially when the whole 3-D data set are not acquired during

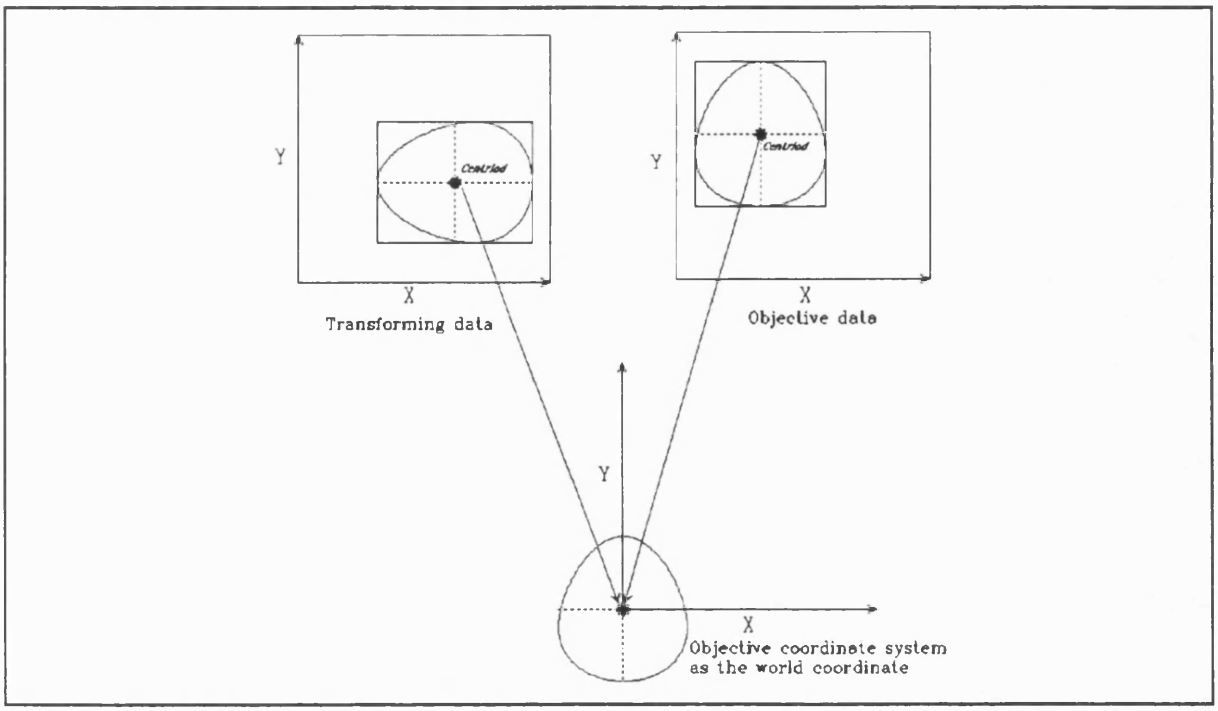


Figure 3.3- Schematic diagram showing the process of centroid registration (i.e. the initial transformation) based on the system known directional (geometric) information.

the imaging. The visual information obtained from a display of overlapping surfaces can also facilitate the manual registration. This visual inspection is discussed in section 3.2.6.

3.2.3- Point Sampling

As stated in section 3.2, only a number of points selected (sampled) from the transforming surface participate in the surface fitting process. The choice of sampling points (the position and number of selected points included in the registration process) should satisfy the assumption that the points selected (i.e. a fraction of all surface points) are representative of all the points on their corresponding surface image. In other word, they should reflect all potential misregistration information about the whole data set. This was attempted, here, by dividing the area of the surface into m regions, each having different distinct locations and where possible, different slopes (surface normals). Those points having the local maximum curvature in each region are then selected as sample points (see figure 3.4). As discussed in section 4.3.4, the curvature can be obtained during the *circularity check* in the surface detection process. The coordinates x,y,z of each selected point are then compared with those of previously selected points in order to reject any point having two or more coordinates similar or close to the

coordinates of the other (neighbouring) points.

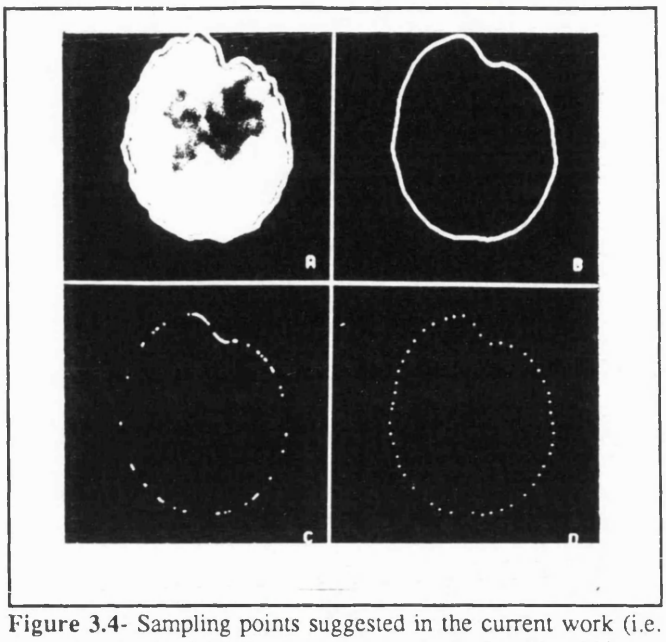


Figure 3.4- Sampling points suggested in the current work (i.e. the points having high local curvatures; shown in image D).

Some special considerations must be taken into account in this sampling process. **(1)** Sampling should be performed on those regions of the transforming surface image for which there are corresponding data on the objective surface. For example the region of eyes are not used for sampling since the *SPECT* data usually does not adequately represent the complexity involved in this region. There are also situations where some non-overlapping regions exist between the two studies. Any attempt to register these regions or to include them in the registration is a misleading in the process. Sampling the points after the initial centroid or manual registration helps to reduce the inconsistency between the registering pairs. **(2)** The number of points and their locations (where they are selected from) are two important factors influencing the sensitivity of the Least Square Distance (LSD) function to an applied geometric transformation. In this respect, selecting the points from high-curvature regions of the head surface yields more specific geometric information (see section 4.3.4) and thus can be expected to yield a process which is more sensitive to change of the transformation parameters. **(3)** As suggested by Nagel and Rosenfeld 1972, the points should be ordered in such a way that any sub-set of them reflect similar misregistration information as the whole set. The characteristics of the sampling points are also revealed in section 6.2.6. The experimental results and discussions about this process are presented in section 8.3.2.1 and 8.3.2.2.

3.2.4- Intersection routine

The distance between the two surfaces is calculated by evaluating the distances between some corresponding points of these surfaces (see figure 3.5a-b). In this respect, for each sample point of the transforming surface, a corresponding point should be determined on the objective

surface. The corresponding point is assumed to be at the intersection point of a ray originated at the centroid, passing through the sample point of the transforming surface, and intersecting with the objective surface. This 3-D intersection can be found using different approaches of a ray tracing algorithm. Since the intersection routine should be applied to all the sampling points, reducing the search required for finding these intersections is one of the main objectives regarding the timing response of the registration process. Two intersection methods used in the current work are based on a line-voxel (ray-voxel) intersection and a line-polygon intersection algorithms.

The first method is applied when the surface is represented by a 3-D binary array (voxels in a 3-D discrete space). The problem of intersection in this approach is similar to the problem of a straight line drawing in raster display system, but in 3-D space. The simplest and quickest approach which is known as an *incremental* method is based on the equation of straight line in 3-D space. In this method, the larger digital differential ($\nabla x=x_2-x_1$, $\nabla y=y_2-y_1$ or $\nabla z=z_2-z_1$, where points (x_1,y_1,z_1) and (x_2,y_2,z_2) are the end points of the line) is selected as a discrete length whose units are considered as incremental units used in the selection of the next point on the straight line. The increment in any other direction is based on the slope (e.g. (x_2-x_1)/length) of line in respect to the first long-length direction.

The Modified Bresenham's Algorithm MBA (Bresenham 1965) was also used which seeks the optimum 3-D discrete location to represent a straight line. Like the incremental method, the MBA algorithm always increments by one unit in the direction of higher length (based on the slope of the line). However, in this algorithm the increment in the other variables (directions) are determined by examining the distance (error) between the actual line location (unlike the latest line location in the incremental method) and the nearest grid location. The MBA algorithm defines a straight line more precisely (i.e. using partial voxels), but it is slower than the simple incremental method. It should be pointed out that due to the initial alignment of surfaces, only a short ray of voxels (e.g. about 30 voxels) in a straight line are required to be searched for the actual intersection with the objective surface. This ray is shown schematically by a solid line in figure 3.5b.

The second type of intersection algorithm is based on a ray tracing algorithm (denoted as RTA) (Roger 1985) which is applied on polygon-defined surfaces (see section 4.4.3). A ray

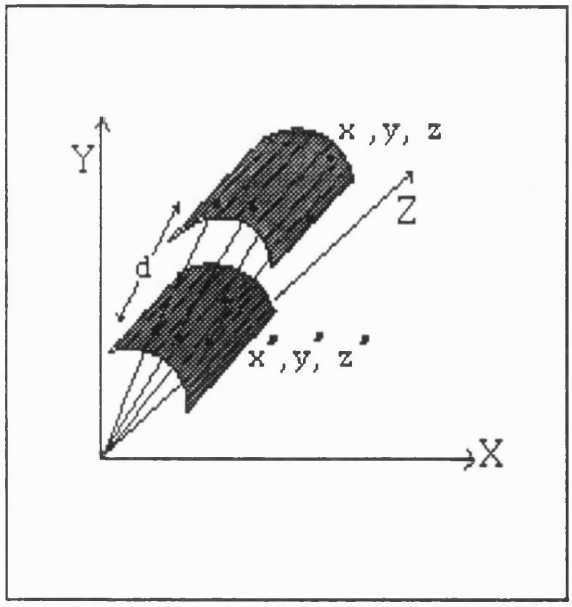


Figure 3.5a- Schematic diagram showing two fitting surfaces. It shows the method by which the correspondence between the transforming and objective points is set.

is traced in an appropriate direction from a sample point to the objective surface and terminated when it intersects a patch of the objective surface. The direction of the ray is defined as the direction of a line passing through two known points, the centroid and the sample point. Accurate implementation of the intersection routine involves a high number of search to detect the target patch between a number surface patches. The distance between the sample point and the intersected patch is then used as the residual to be minimized.

Although, the number of surface elements on a tiled surface is much smaller than that on a voxel-based surface, the intersection process is not substantially faster due to the fact that these tiles (surface elements) must be examined by an expensive process in respect of each ray before the actual intersection is found. To eliminate unnecessary intersections, the intersection of a ray with some bounding volumes (such as bounding sphere or bounding box which contains a large surface patch) of an object is examined, initially. Then, those patches whose bounding volumes fail to be intersected by the ray, are not considered for further examination (e.g. distance measurement) (see figure 3.6). On the other hand, an intersecting ray is further tested against the bounding volumes of the sub-patches obtained by a subdivision of the original patch (Catmull 1974) if it initially intersects the patch bounding volume. Examining a ray for an intersection with a sub-patch (surface element) itself would be only applied if its bounding volume is

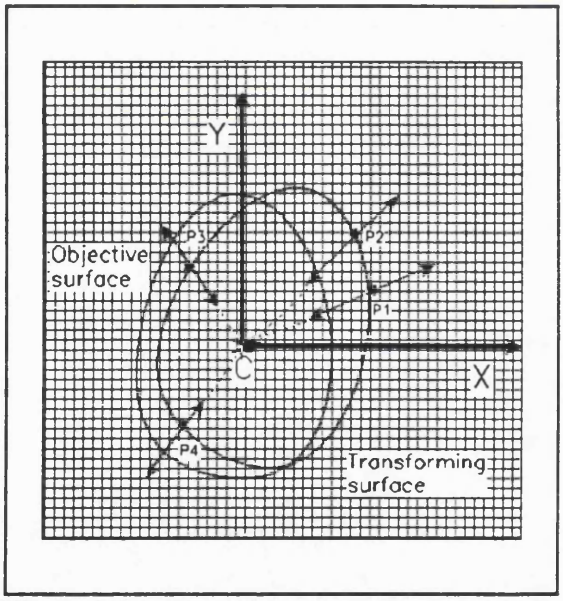


Figure 3.5b- Shows the intersection between the sample points (of transforming surface) and the objective surface in a 2-D cross section (contours) image.

intersected by the ray.

The RTA procedure uses translation and rotation about the coordinate axes to make the ray coincident with a selected axis for example the z axis, thereby to reduce the 3-D intersection problem to an intersection of a ray with a 2-D object. The intersection process is, therefore, defined as a simple intersection of a line with a 2-D plane (Apple 1968, Goldstein and Nagel 1971)

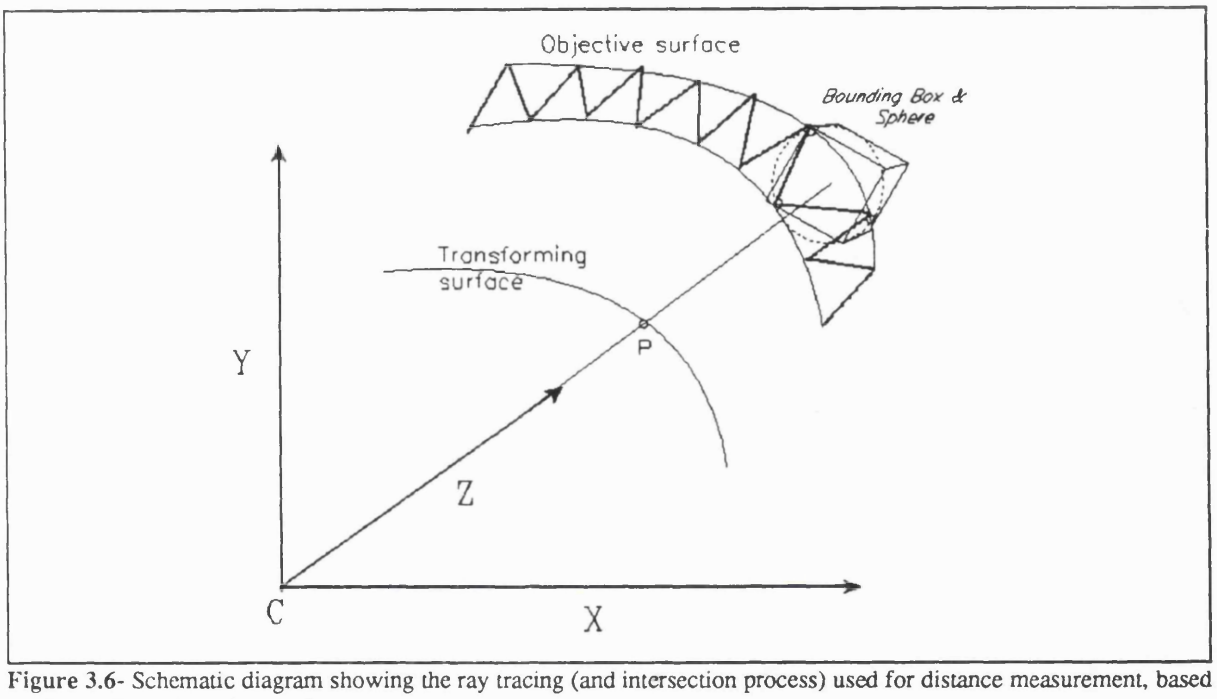


Figure 3.6- Schematic diagram showing the ray tracing (and intersection process) used for distance measurement, based on the surface triangular patches.

3.2.5- Mismatch measurement

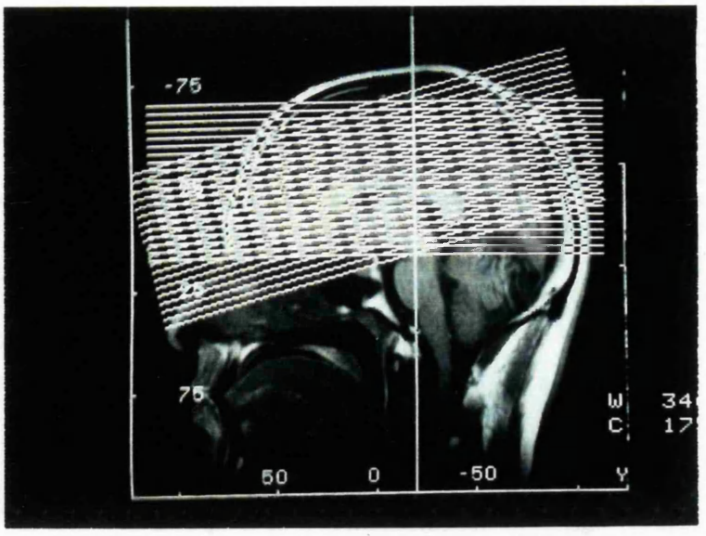
The fit is evaluated by a L2 norm least squares distance function (LSD) minimized in respect of the best transformation parameters. The residual is defined as the mismatch value which is the sum of the distances (or square distances) between the points on the transforming surface (set of sampling points of hat) and the intersection points on the objective surface (head). The distance between the sampling points and the corresponding points on the surface to be fitted (objective surface) are determined by tracing a ray through the 3-D surface data after an initial registration (e.g. centroid and/or manual registration). This distance is computed using the L2 norm distance function for each transformation. The distance function $\Sigma^2 d_i$ which is to be

minimized can be expressed as:

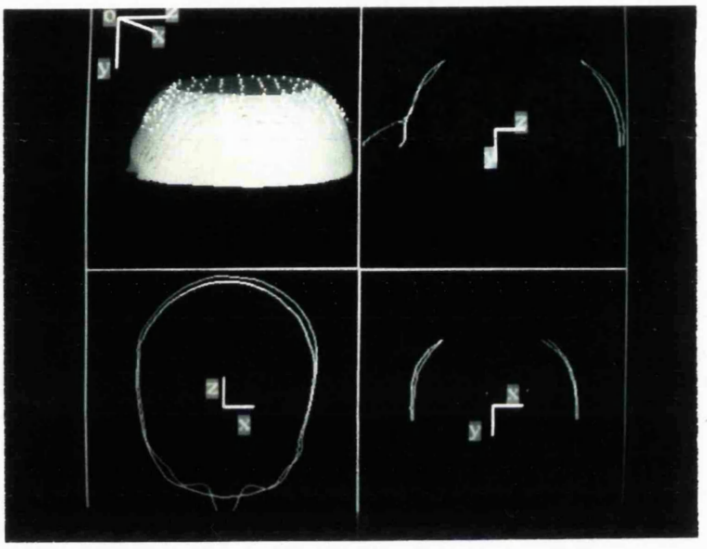
$$\Sigma^2 d_i = \Sigma\ (\mathbf{P}_{ti}[T] - \mathbf{P}_{oi})^2; \quad (i=1,..,n),$$
$$\Sigma^2 d_i = \Sigma\ \{(X'_{ti}, Y'_{ti}, Z'_{ti}) - (X_{oi}, Y_{oi}, Z_{oi})\}^2 \qquad \text{Equ. 3.1}$$

where $\mathbf{P}_t$ and $\mathbf{P}_o$ are vectors representing the corresponding points on each surface: the surface to be transformed and the target (objective) surface, respectively. n is the number of sample points. X'_t, Y'_t and Z'_t are the transformed coordinates of point $\mathbf{P}_t$. $\Sigma^2 d_i$ is the mismatch value (LSD value) which is the sum of the squares of the distances between each individual pair of points. The average value of individual square distance errors is referred to mean square distance (MSD) error and is usually used as the mismatch measure in surface fitting process. The square root of this value is denoted as mean distance error (MDE) as widely used through this thesis. Individual distance errors can be accumulated as sample points are examined and this type of error is known as cumulative distance error (CDE).

(a)



(b)



(c)

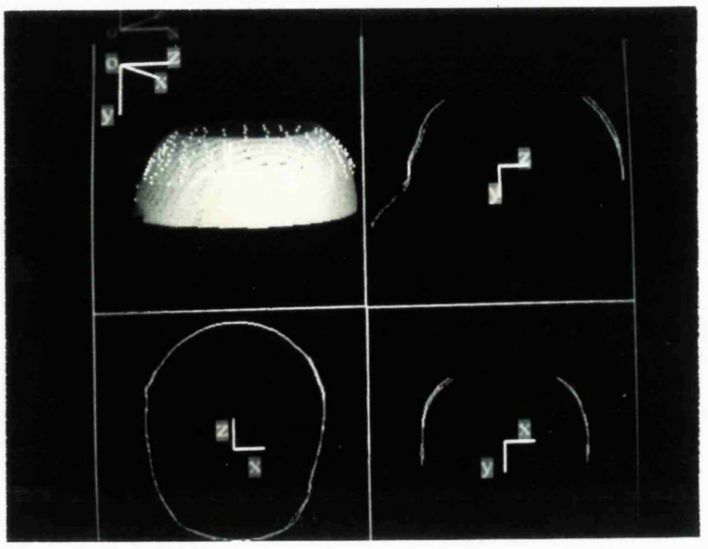


Figure 3.7- A type of visual inspection displayed during the surface fitting process. Top image shows a *MR* sagittal slice. The lines in the top image show the position of planes through the head, before and after transformation (tilt) by the system parameters (see section 8.2; 'data type C'. Middle images show the contours in three orthogonal planes display before registration, bottom images show these contours after registration.

3.2.6- Visual inspection

In the present algorithm, a set of boundary contours were generated which show the positional (geometric) information of both registering surfaces in respect to each other. Figure 3.7 shows a typical example of this display. The coordinate system in which the registration process is applied and the directional information of each surface data set are displayed in these images. The 2-D cross sectional contours of head displayed in three main orthogonal planes provide visual information to judge about the correctness of a fit. The manual registration is also based on these contour displays, which can put the two surfaces in a very close match location. More explanation and images defining this display strategy is outlined in chapter 9.

3.2.7- Minimizing distance function values

The transformation parameters can be determined by a minimization process using either a local search (for example by using the direction set method of Powell, 1964) or a global search strategy (see chapter 4 for both these techniques). Since the correspondence between the points of the two registering images are not exactly known in advance, an iterative manner was used. Initially, the correspondence between the points can only be predicted. Nonetheless, this prediction is improved after each stage of the (minimization) process when the two surfaces come to a close match state. In this respect, the (transforming) surface position is updated after each certain promising transformation and the ray intersection process is applied on the latest aligned surfaces. Routinely, in an iteration of the process, only those points selected on transforming surface are shifted. The objective surface is kept constant during the process.

The six rigid geometric transformations (three shifts in x, y, and z direction, and three rotations around x, y, and z axes) are modified until the residual as defined in the previous section is minimized. Linear scaling is only applied in a limited way to account for minor uncertainties in pixel size or in accuracy of the detected surfaces of a low contrast image (due to possible sub-optimal choice of threshold level; see section 4.3.2). These uncertainties can be detected either by viewing the superimposed (overlaid) contours in three main planes during the visual inspection, or by calculating the variance of individual distance values. In the later approach, scaling is only applied if the variance of the distance errors is low in spite of having a high residual value (i.e. all the distance errors are comparably high).

The choice of minimization method is very important in computer cost (timing and space) and the accuracy of a detected minimum. Most direction set methods (e.g. Powell 1964) suffer from being trapped in multiple local minima. Different methods of minimization used through the current project are discussed in chapter 5. A sequential and multi-resolution method is also introduced in a novel approach in chapters 6 and 7. These transformation parameters are termed the *search location* in the next chapter where different search algorithms are presented to find the best transformation parameters giving the minimum of the distance function.

3.2.8- Data transformation and alignment

Having been registered the two images, the appropriate transformation may be used to transform any structure in either image set (transforming or objective data) into the coordinate of the other. The images can then be fitted to generate a useful superimposed image for clinical interpretation. Contours and volumes of interest (VOIs) can be outlined on 3-D data set (or multiple slices) of one scan and transformed to coordinate system of the other scan and finally superimposed on the corresponding slices. Various methods of display and superimposition of the registered data sets are shown and discussed in chapter 9.

3.3- Modifications to the previous methods and new approaches

Various methods of surface fitting outlined in the literature can be characterized by differences in the type and level of description at which the matching was attempted. Accordingly, some new approaches are used in the current work to facilitate the surface fitting algorithm and reduce its previous deficiencies.

a) The matching performed by most of the previous workers was based on the surface patches delineated between a number of selected points on the contours of 2-D slices. In this work, the fitted surfaces are characterized by voxels in 3-D geometrical space. Since the surfaces are represented at the voxel level, a good accuracy can be expected in presentation of fine structures and surface details.

b) In this work the circularity check (see section 4.3.4) on the detected boundaries is performed.

With respect to that, any hole which is likely to be due to presence of tumour or cold regions in the surface of *ECT* (Emission Computerized Tomography) data, can be interpolated. As discussed in section 4.5, extrapolating over such holes can reform the defective parts of surfaces, and thus facilitate the application of the surface fitting method for registration of *SPECT* data.

c) A new approach, based the shape interpolation (see section 4.4.4), for reconstruction of the surface between the adjacent slices is used. It is believed (Herman 1991) that a better approximation (than the one of trilinear interpolation) is obtained between the contours of thick slices or in the gaps between them. In this method, the geometrical form of the surface is represented by the locations of the (binary) voxels, and grey scale values are not used.

d) Unlike some of previous methods, only rotations and translations are applied as 3-D rigid transformation. In this work, no attempt is made to scale the data during the routine registration process. The scaling is applied only at the preprocessing stage based on the knowledge of imaging parameters. However, if the statistic of the individual distance errors confirms an uncertainty in image size, an attempt is made to scale the registering data set. This uncertainty is due to segmentation errors for example a possible sub-optimal choice of threshold in surface detection process. In this respect, the variance of distance errors are evaluated as well as the distance errors (e.g. MSD value). Special attention is paid, then, to a high MSD value which has a low variance. These are counted as behaviour of two images being registered at different wrong geometrical scaling.

e) In the current method, after each stage in which a promising transformation is detected, the coordinates of the transforming surface are updated. The sample points are then reselected from the transformed surface. This updating and re-sampling strategy minimizes the error introduced due to sub-optimal choice of the point-correspondence (as described in section 3.2.4).

f) Due to the inherent type of surface representation, the method of ray tracing and 3-D intersection used in the current work is based on a line drawing algorithm in 3-D geometrical space. Although this intersection method ought to be slower than those based on intersection of a ray with surface patches (tiles), some efficient modifications, such as tracing of a small line segment around the surface points (instead of a line through 3-D volume data), make the algorithm fast enough to be applicable in the surface fitting process. The efficiency becomes

consistently higher with reducing the size of intersection line across the surfaces.

g) Viewing and visual inspection play a major and important role in registration process. Three orthogonal contours (on the 3 orthogonal planes) of cross sectional images are displayed before starting an automatic registration and also between some stages of the registration. This visualization makes the manual interaction and operator modifications applied during the registration easier. The superimposed 3-D registering points (from hat) and the surface data (head) are also illustrated in one corner of the same contour-viewing screen.

h) The surface fitting technique was augmented in order to preclude undesirable local minima obtained by previous methods, thereby allowing it to be used on *SPECT* imaging. A multi-dimensional symmetric space of geometric transformations is globally searched, in a novel approach, for the global transformation parameters (at the true match location). Using a cumulative distance error instead of a mean distance error provides a sequential process which can be terminated at certain mismatch locations. An alternative type of algorithm which uses a multi-grid (known as multi-resolution) coarse-to-fine search space to speed up the registration process is suggested. Using these two strategies greatly reduces the search effort and makes the global search applicable for registration of the medical images, at a reasonable computing time. These are the main originalities of the project, being set forth in chapters 6 and 7.

i) Finally, display methods of the registered images and superimposition routine were developed which are new in their own approach. The implementation was mostly based on the requirements of clinicians in the hospital where the work was performed. In chapter 9, different display methods as well as some examples of superimposed images are outlined.

3.4- Material and method

Routine clinical imaging studies were used without any special patient positioning during imaging. All the machine parameters and scaling information (e.g. pixel size, field of view, zoom factor) were obtained from each imaging system itself. These parameters were examined, initially, using a well-defined phantom with known size information. Each imaging system has a specific and known coordinate system which defines the image coordinates based on the patient

directional position inside the imaging system. A typical imaging coordinate system is shown in figure 3.8.

Different images can be made available from any imaging modality by applying different imaging parameters (e.g. different slice thicknesses, or different pulse sequences in *MRI*). The following are two main criteria based on which an imaging modality and in turn its images are chosen in the fitting process.

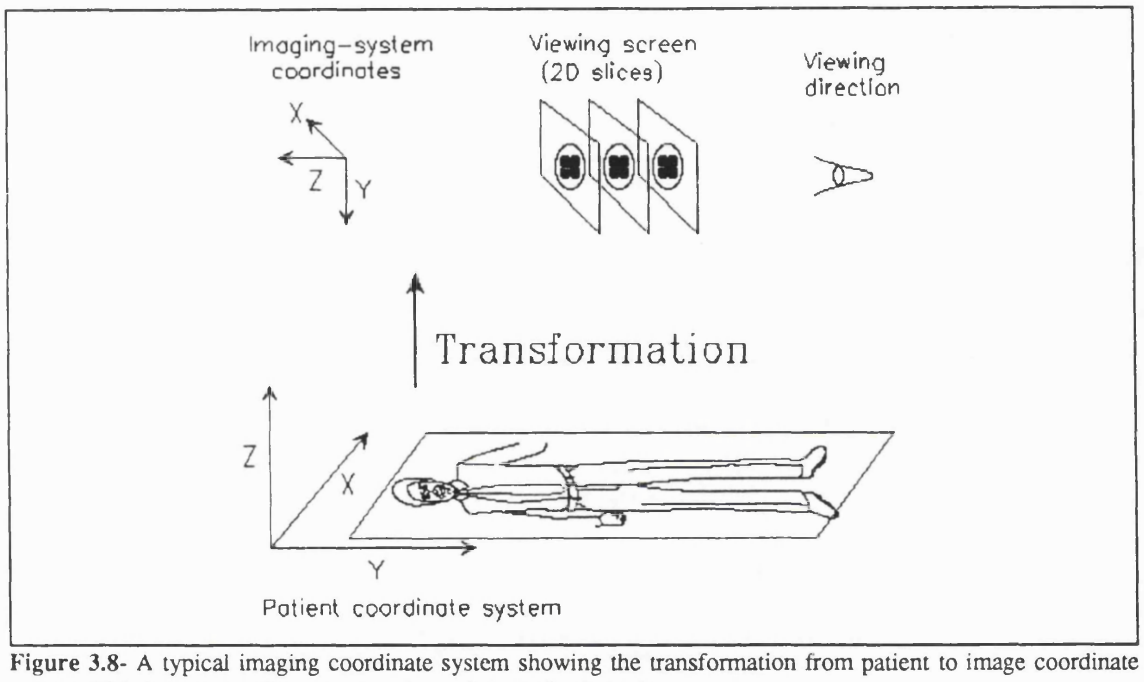


Figure 3.8- A typical imaging coordinate system showing the transformation from patient to image coordinate system. This image coordinate system is used throughout thesis.

1) The 3-D surface model (objective data) is usually taken from the scan having higher resolution, smaller slice thickness, and the one which covers a larger volume of the patient. On the other hand, the transforming points are sampled, only at slice level, from the contours of the image which have coarser resolution and worse slice thickness.

2) Both the external surface of scalp and brain surface might be visible and thus formed by different segmentation techniques. The external surface of the scalp (skin surface) is the most commonly used object for surface fitting process, and is readily outlined on most images such as *CT* and *MR*. Using this external scalp surface is also well suited for automating edge detection methods and thus requires little user interface (see section 4.3.3). In general, an effort is also made to employ skin surfaces instead of brain surfaces where there is a substantial brain surface

deformity (such as a hole) in one of the images.

The technique was used to register the *MR* with *MR*, *PET* with *MR*, *SPECT* (*HMPAO*) with *MR*, and *CT* with *MR* brain images of a number of patients. Skin surfaces were used for registration of *MR* with *MR*, *MR* with *PET* and *MR* with *CT* images. Brain surfaces were usually used for the registration of *MR* with *HMPAO* brain images. The characteristics of the imaging systems and the routine imaging parameters used in the current project are as follow.

MR scans were performed with a 1.5 Tesla *Siemens* system operated by a *VAX* host computer. Both 3-D *FLASH* (*Fast Low Angle Shot*) images and 2-D slices were used and dealt with separately. Different imaging sequences and parameters were used, creating T2-weighted, T1-weighted or/and *STIR* (*Short Tau Inversion Recovery*) images of a pixel size typically in the order of 0.70-1 mm (300 mm field of view, 256*256 matrix size and zoom factor of 1.2-1.6). Acquisition was typically done with TR=4.0 sec and TE=90 msec for T2, and TR=600 msec and TE=15 msec for T1-weighted images, slice thickness of 5-6 mm and a gap of 2-2.5 mm (when STIR sequence is used). Transversal slices were usually obtained starting few millimetres (about 10 mm) from the top of the head, down to the base of skull. *MR* images obtained by a short echo time were used for the detection of skin surfaces. A *STIR* sequence was used to create 2-D fat-suppressed *MR* image which makes the detection of the brain surfaces easier.

PET images were generated from a *PET* scanner having eight rings of *BGO* (bismuth germanate) detectors (512/ring) of width 5.6 mm (transaxial). This is a CTI system operating at MRC cyclotron unit (in Hammersmith hospital, London). Scans were performed after inhalation of $C\ ^{15}O_2$ for the measurement of cerebral blood flow. Three other different studies were routinely done using O, or CO for emissional tomography, and ^{68}Ge (Germanium) for transmissional information which is used for attenuation correction. 2-D transversal slices of 6.75 mm thickness and 4.69 mm pixel size were reconstructed, on a 128*128 matrix, from the vertex to the base of skull. The physical resolution of the machine corresponds to a full width half maximum (FWHM) of 6 mm (Spinks et al 1988). The surface of *PET* data was obtained from the transmission data set which are usually acquired during scanning for attenuation correction purpose. The external surface can also be defined from emission data as the outer envelope of the emission distribution detected under a threshold-based segmentation technique.

SPECT images were created by an *Elscint SP4 SPECT* system using ^{99m}Tc (Technetium) labelled radiopharmaceuticals. The actual resolution of the system is about 7-10 mm at FWHM (at 10 cm in depth). The data were acquired after injection of *HMPAO* (*hexamethylpropylene amine oxime*) and the transverse scans were reconstructed with a slice thickness of the same pixel size. A matrix size of 64*64 gives a pixel size of about 6 mm at a zoom factor of 1, which was used to obtain 2-D slices. External surfaces from the *SPECT* can be defined as the outer envelope of the emission distribution detected by setting of a threshold value. However, decisions on threshold value and the type of threshold selection technique are very important factors in the accuracy of the detected objects (brain surface). Our experiments to achieve the best result are outlinec in section 4.3.5.

Contiguous *CT* transverse scans were performed with a *T60 Toshiba CT* machine using system setting of 240 mm field of view and 5 or 10 mm slice thickness. Using a matrix size 320*320 and zoom factor of 1 gives 2-D slices having a pixel size of 0.75 mm. For viewing purpose, the original high-grey-level data (represented by a 16-bits value) were scaled to a grey level of only 256 (correspond to 8 bits) due to the limitation of our system. *CT* images are characterized by a distinct skin tissue image adjacent to a low noisy background. Therefore, surface detection is usually easy and relatively trivial on the *CT* data.

3.5- Verification

Registration of some well-defined clinical head images was performed to verify the accuracy of the present surface fitting method. Since phantom studies were done and analyzed in detail by Pelizzari and Chen 1989, and our objective was to use more realistic clinical situation, the verification was obtained by various clinical and simulated data. Three different types of well-defined data were used for this assessment.

(1) A set of high definition *MR* data was employed initially, and a known transformation (3-D tilt and shift) was applied. The data were then degraded and incorporated with noise to simulate some *PET* or *SPECT* images. An effort was made to register these simulated data to the original ones and the algorithm response and misfit values were measured.

(2) Different images (scans) of one volunteer patient were also obtained by setting of some arbitrary geometric changes (such as setting combination of slice-selection gradients in different orthogonal directions to rotate the image in different directions) on the system while the object itself did not move or tilt between the different studies. The top image shown in figure 3.7 demonstrates the registering regions of head on an example of this data type.

(3) Finally some external landmark studies were used and the results of the current registration were compared to those of the landmark studies.

The results obtained from these known misregistered simulated data are presented in section 3.6 and also in chapter 8. The limiting factors in accuracy of a match are the large pixel size, the limited resolution, and the large slice thickness. Accordingly, accuracy (as the minimum mismatch value) is expected to be not worse than the larger voxel size of the two registering images. However, interpolation and scaling of the coarser data set to the same size of the high resolution data, decreases voxel size and thus increases the accuracy.

3.6- Results and discussion

The results presented in this section are based on the well known data sets introduced in section 3.5. The geometric and directional information of such data types is known in advance, and thus provides a well-defined data set for the evaluation of the algorithm. The aim of this section is to show that the surface fitting algorithm and the type of mismatch function which is minimized during the process, are efficient for obtaining a proper registration.

3.6.1- Sensitivity of distance function to geometric transformation

In order to assess the behaviour of the least square distance function to the individual transformation parameter, a set of experiments was designed. In these experiments, sets of well-defined registered surface data were used. These surfaces were then misregistered by a set of transformations (-15 to +15 voxels shift or/and -15 to +15 degrees rotation), and the mismatch values (MDE) were calculated.

Technique	MSD Value											
	$Z^{\rightarrow}$	$Z^{\leftrightarrow}$	$X^{\rightarrow}$	$X^{\leftrightarrow}$	$Y^{\rightarrow}$	$Y^{\leftrightarrow}$	$Rx^{\rightarrow}$	$Rx^{\leftrightarrow}$	$Ry^{\rightarrow}$	$Ry^{\leftrightarrow}$	$Rz^{\rightarrow}$	$Rz^{\leftrightarrow}$
Random samples from region *V2*	1.6	2.8	0.65	0.89	1.0	1.81	0.48	0.71	1.3	2.95	0.31	0.39
Random samples from region *V1*	1.8	3.5	0.82	1.49	1.1	2.15	0.55	1.13	1.4	3.45	0.42	0.76
High Curvature points	2.0	2.4	1.45	2.15	1.1	1.85	0.55	0.75	1.6	2.7	0.51	0.65
Low Curvature points	1.2	2.1	0.45	0.65	0.9	1.35	0.40	0.66	0.8	1.95	0.26	0.33

Region *V1* is the whole head surface obtained from the subject scan.
Region *V2* is the head surface excluding face (eyes and below-eyes region).
The shifts or rotations in different directions are expressed by X,Y, Z, Rx, Ry and Rz.
MDE denotes the mean distance error (value).
Superscriptions:
$\rightarrow$ denotes one unit of transformation-parameter change applied at a near match location.
$\leftrightarrow$ denotes one unit of transformation-parameter change applied at a certain mismatch location.

Table 3.1- Showing the effect of different sampling strategies on sensitivity of the distance function to one unit of transformation (in x, y or z) for *MR-MR* registration.

The results of these experiments are presented in figure 3.9, 3.10 and 3.11. For type of object which we are interested in i.e. the head which is almost symmetrically spherical with a radius of about 150 mm, figure 3.9 shows that the distance function is less sensitive to rotation than it is to translation (shift). However, it has a sharp single minimum when two surfaces are truly matched. Figure 3.10 shows the relative response of the rotation and shift at each direction, x, y and z. In general, it is desirable that the change in MDE value becomes high even in case of a minor transformation. As revealed by some of our experiments, the sensitivity of the distance function is influenced by the position, characteristic, and the number of sample points. The results obtained by varying the first two factors are presented in table 3.1. The number of points required is assessed in section 3.6.2 and also in chapter 8. As shown in this table, the change induced by one degree shift or rotation in each direction confirms that the sensitivity is highly dependent on the position and type of sample points. The sampling restricted to the regions of the head apart from eyes and face shows a better performance than the whole head in term of the sensitivity. The high-curvature features (points) are also shown to have a much better sensitivity compared to the low-curvature features.

The graph in figure 3.11 is based on the changing all six transformation parameters by the same value (i.e. a shift or rotation within the limit of -15 to +15 degrees rotation, or voxels

shift). Although a sharp global minimum is visible, like the one obtained by changing the individual parameters, the situation is more complex in the real minimization circumstances, when any combination of these parameters may be applied. This complexity is due to the existence of a large number of transformations between -15 to +15 (i.e. of the order of 31^6 for various combination of 6 parameters) which, in turn, may lead to a large number of local minima. These local minima are shown in figures 3.12a-d for different sizes of transformations in x, y and z direction. Note that, in figures 3.12a-d, the parameters are changed by the stated steps (see the caption on the figures), but by changing one parameter at a time. When all the steps of one parameter are applied, the process is continued by changing the next parameter. As shown in these figures, there is a number of local minima, for example, even when the two surfaces are misregistered by 180 degrees rotation. These results indicate the crucial need for the initial manual registration, based on the knowledge of the system parameters and visual inspection.

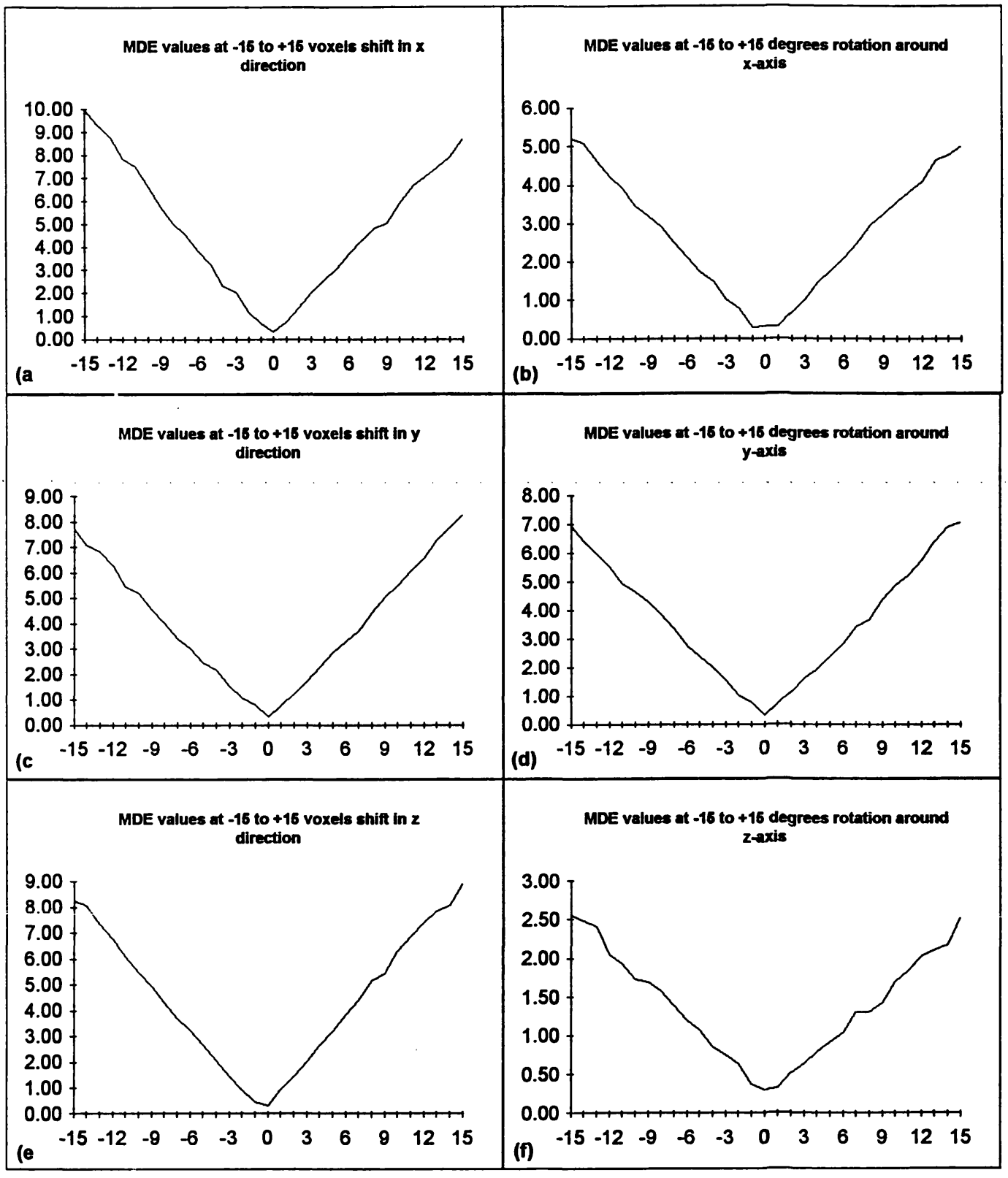


Figure 3.9- Graphs showing the response (sensitivity) of the distance function to each individual transformation parameter. MDE denotes *mean distance error*.

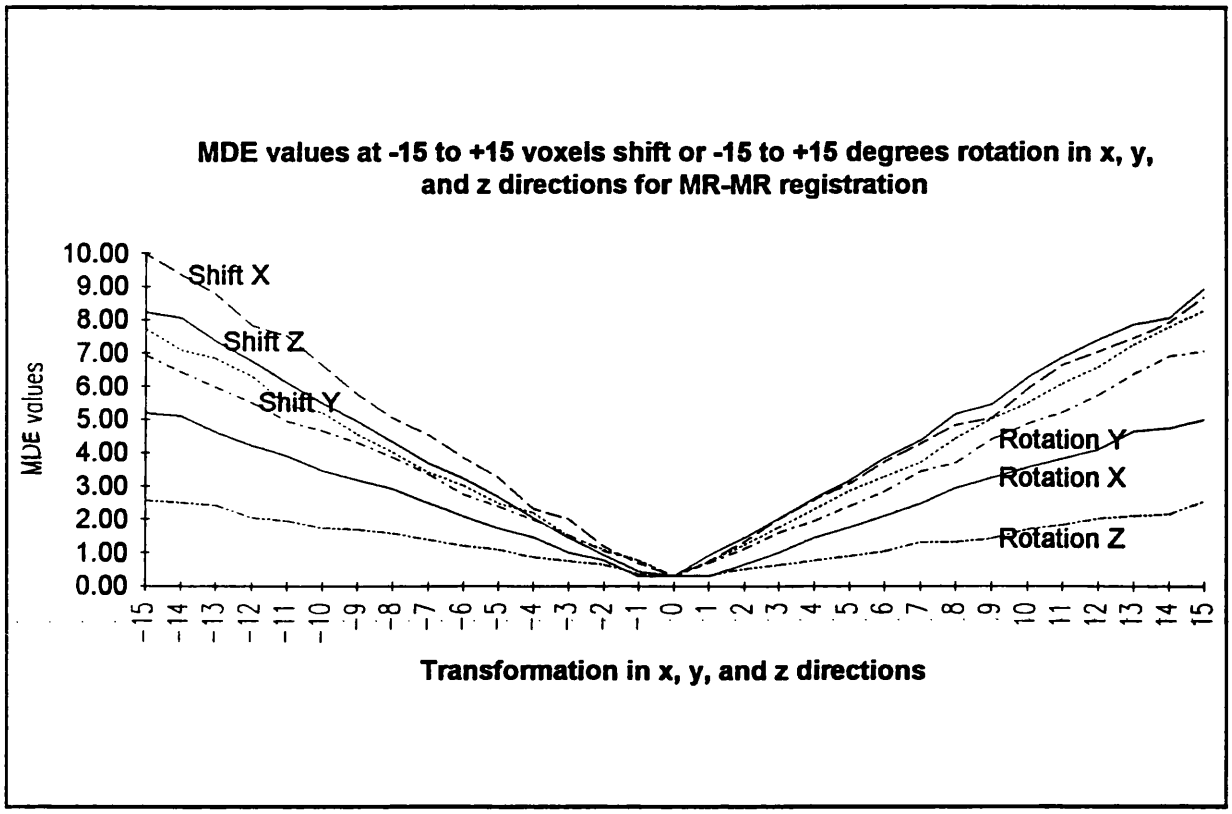


Figure 3.10- The relative response (behaviour) of the distance function (as MDE) to different geometric transformations.

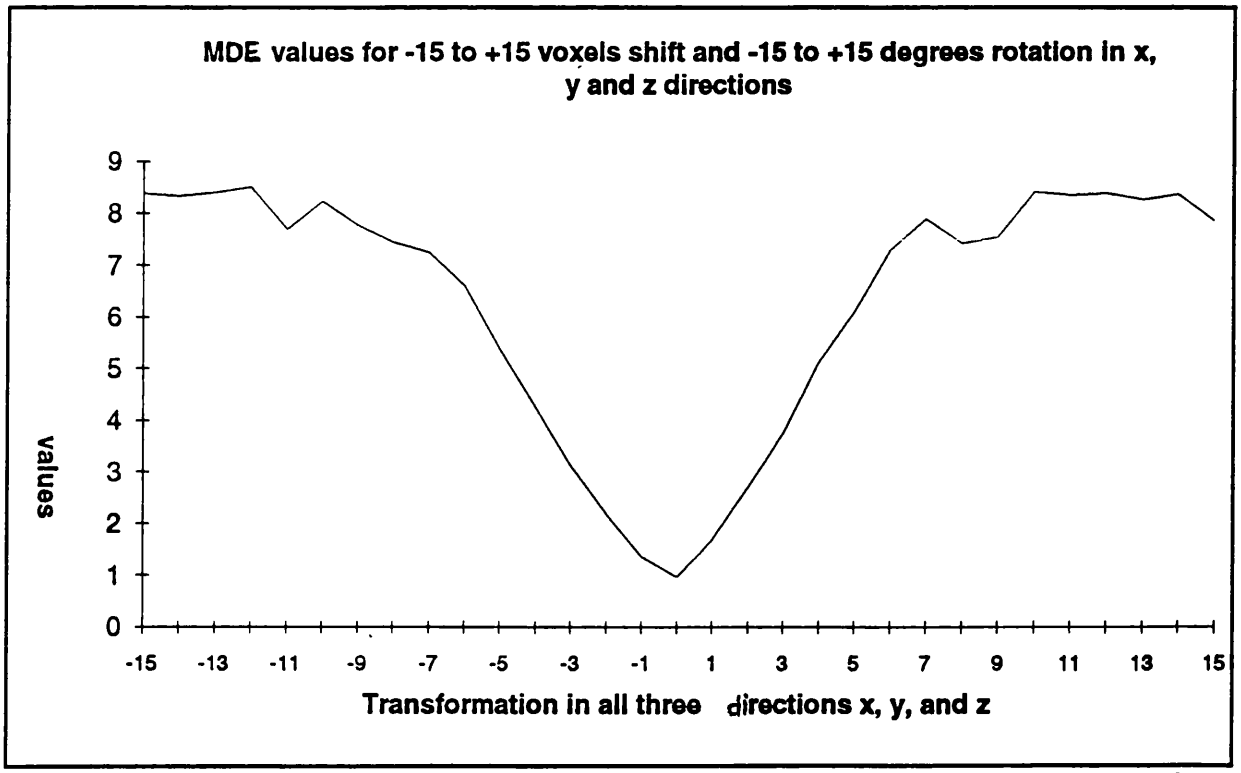


Figure 3.11- The response of MDE to the combination of transformations in all directions (i.e. the same value of shift and rotation applied between -15 to +15).

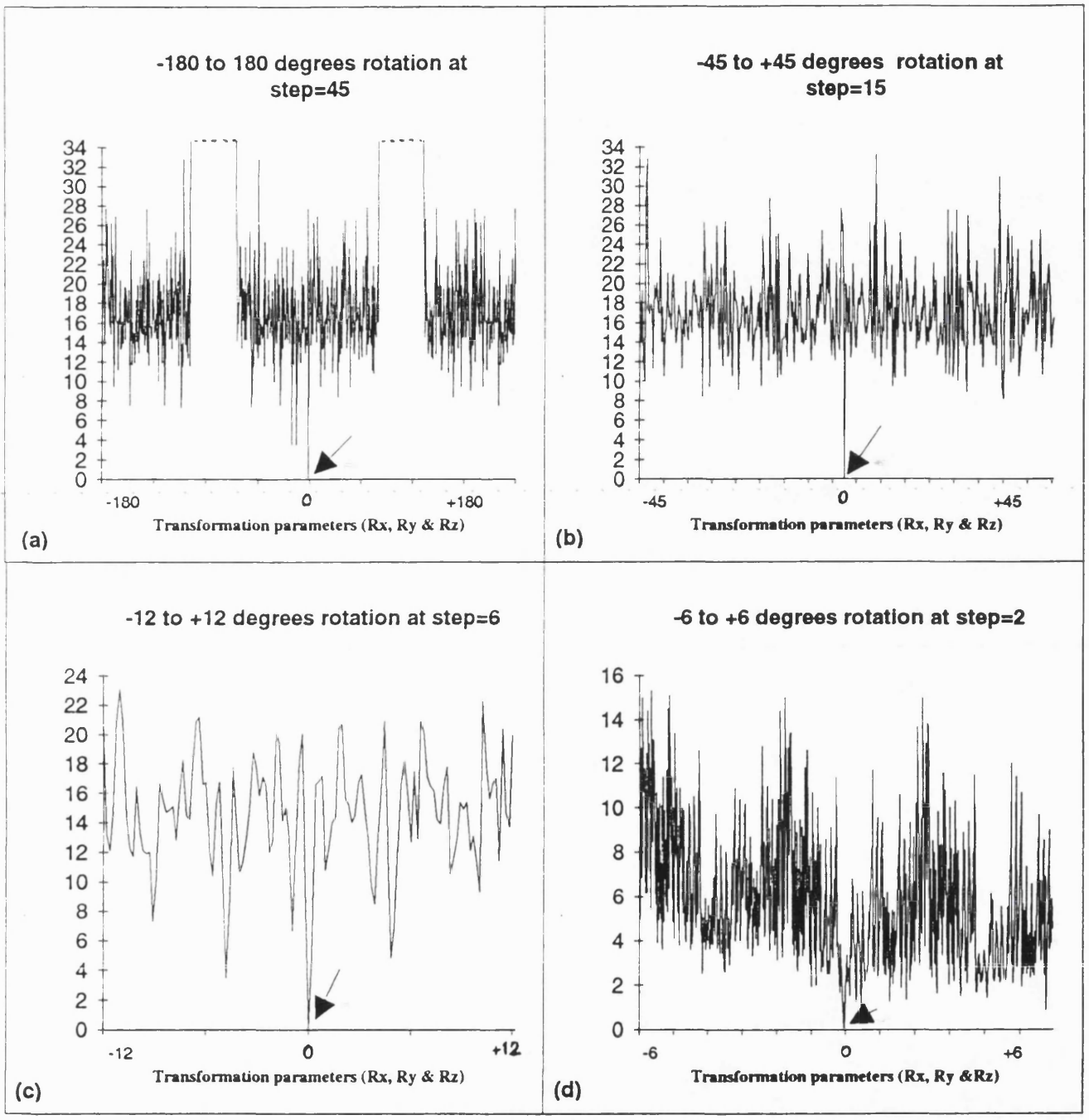


Figure 3.12- The distance values (MDE) at different combinations of the geometric transformation around the true match location. The parameters are changed between the two extremes by the stated steps. In these intervals, one parameter is changed at a time. The next parameter is changed only when all the steps of the previous parameter is completed. The arrows indicate the global minimum (true match location).

3.6.2- Accuracy and cost assessment

The accuracy and cost of applying the surface fitting algorithm are presented in table 3.2. The accuracy can defined as the minimum mismatch value obtained by the registration process (see section 8.3.4). The results presented in this section are based on the first type of data sets introduced in section 3.5. The original MRI data and the simulated PET and SPECT images were arbitrary misregistered in order to generate a number of well-defined misregistered data sets. The results shown in table 3.2 are obtained from two types of local minimization algorithms: the direction set method of *Powell*, and conjugate gradient method of Fletcher-Reeves (e.g. *FRPRMN*)(see section 5.3.1 and 5.3.2). As shown in this table, the process may not converge to a global minimum, due to an early detection of a local minimum. This response depends highly on the relative location (i.e. in relation to the actual required transformation for the true match) from which the minimization starts. As shown in table 3.2, starting the local search at different locations (i.e. different initial transformation) can change the efficiency of finding of the global minimum. The results presented in each group-transformation in this table, include different starting transformations (denoted by subscription IA, IB, or IC) among which some lead to a match, some to a mismatch. As suggested by previous workers, a number of re-initializations can be employed to guess a new initial set of parameters when detecting a local minimum, thus re-starting the process from the new location. This is one of the main disadvantages of the local minimization process since it relies on the manual interaction of the operator. Accordingly, the process is user-dependent, and the convergence is undesirably prolonged.

As shown in table 3.2, the accuracy obtained by *MR*-to-*MR* registration is of the order of 1 voxel (corresponding to a transformation error of 1 voxel shift or 1 degree rotation). This accuracy increases to about 1.5 voxels for *MR*-to-*PET* registration and 2.5 voxels for *MR*-to-*SPECT* registration. These correspond to an error of about 1-2 voxels in shift or 1-2 degrees in rotation for *MR-to-PET* registration, and 2-3 voxels in shift (or 2-3 degrees in rotation) for *MR-to-SPECT* registration. The timing response of the process was shown to be better for the *Powell* method (which is about 300 seconds) than it is for the *FRPRMN*'s (about 360 seconds as averaged). The number of points was also shown to influence the accuracy and cost of the minimization. As the number of points increases, the accuracy improves but the cost expectedly increased.

Table 3.2- Showing the expected accuracy (as the minimum obtained distance error), and cost of the surface fitting using the Powell and FRPRMN minimization methods.

Application			Original TP Differences		Errors in TP after Registration			
Data	Method	Scheme	x,y,z	Rx,Ry,Rz	x,y,z	Rx,Ry,Rz	MDE	Time
MRI With *MRI*	Powell	R_3	-	5,8,10	-	0,1,0	0.88	75
		S_3	5,8,10	-	0,0,1	-	0.2	115
		T_{IA}	5,8,10	5,8,10	2,2,1	-2,-3,-5	3.35	980
		T_{IC}	5,8,10	5,8,10	0,0,2	-1,1,0	1.22	140
		T_{IA}	5,5,5	0,0,0	0,0,0	0,0,0	0.00	70
		T_{IA}	5,8,10	0,0,0	-1,-1,0	0,0,0	0.25	370
		T_{IB}	5,8,10	0,0,0	1,-3,-3	2,2,1	1.80	500
		T_{IA}	0,0,0	5,5,5	1,2,-4	-1,1,1	2.19	150
		T_{IA}	0,0,0	5,8,10	1,0,-4	-3,-3,1	2.67	750
		T_{IB}	0,0,0	5,5,5	0,0,0	0,0,1	0.80	114
		T_{IB}	0,0,0	5,8,10	0,0,1	0,1,1	1.32	180
		T_m	0,0,0	5,5,5	0,0,1	0,1,0	1.22	250
		T_m	5,5,5	5,5,5	0,0,-1	1,-1,0	0.90	230
		T_m	5,8,10	5,8,10	-1,0,0	-1,1,0	0.97	370
		T_{50}	5,8,10	5,8,10	-1,-1,0	0,0,2	1.47	180
		T_{100}	5,8,10	5,8,10	0,0,1	0,1,1	1.00	250
		T_{200}	5,8,10	5,8,10	-1,0,0	-1,1,0	0.97	370
	FRPRMN	T_{IA}	5,8,10	-	0,0,0	-	0.03	220
		T_{IA}	-	5,8,10	-	2,0,0	0.50	200
		T_m	5,5,5	5,5,5	0,-1,0	0,-1,1	0.79	480
		T_{IC}	5,8,10	5,8,10	0,-1,0	0,0,-1	0.66	290
MRI With *PET*	Powell	T_m	5,5,5	5,5,5	0,1,-1	-1,-1,1	1.42	323
		T_m	5,8,10	5,8,10	0,2,0	0,0,-2	1.60	280
	FRPRMN	T_m	5,5,5	5,5,5	0,0,2	1,2 0	1.75	322
MRI With *SPECT*	Powell	T_m	5,5,5	5,5,5	0,1,2	0,0,3	2.45	235
		T_m	5,8,10	5,8,10	1,0,2	1,-1,1	1.95	320
	FRPRMN	T_m	5,5,5	5,5,5	0,0,2	-1,2,0	2.30	280

R_3= Only 3 rotations around x, y and z axes were used. S3= Only 3 shifts were used.
T_{IA}= Initial transformation of 0 was used for shifts x, y, z and rotations Rx,Ry,Rz.
T_{IB}= Initial transformation of -10,-10,-10,0,0,0 was used for x,y,z,Rx,Ry,Rz.
T_{IC}= Initial transformation of 15,15,15,15,15,15 was used for x,y,z,Rx,Ry,Rz.
T_{50}= 50 sample points were used for the registration.
T_{100}= 100, and T_{200}= 200 sample points were used for the registration.
MDE denotes the mean distance error between the two registering surfaces.
T_m= Manual alteration of the transformation was applied at a false local minimum.
(Note that simulated *PET* and *SPECT* data (see section 3.5) were used).

3.7- Summary

Firstly, the general technique of the surface fitting as defined by other workers was described in this chapter. A number of improvements to the general surface fitting algorithm was suggested in section 3.3. One of the key improvements is due to the use of a global minimization algorithm which increase the accuracy of the registration. After defining the surface fitting algorithm, a strategy to achieve the registration in a reasonable duration of time remains essential, and is still unexplored. These aims have led to two novel approaches, sequential and multi-grid registrations, which are discussed in chapters 6 and 7.

The imaging parameters and data types used through this project were also introduced in this chapter. The characteristics of the imaging systems introduced here are conventional in most medical departments. Thus, they can provide some image models for verifying the applicability of the different aspects of the registration process.

The methods used for the verification of the algorithm and the different well defined data sets used for this purpose are also outlined in this chapter. Based on these data, the technique was verified and its results are presented in section 3.6. It has been shown that the type of mismatch function (e.g. LSD) and, in turn, the surface fitting algorithm, are capable to register two clinical data sets by producing a reasonably small mismatch between them. However, a number of local minima may still exist, for which care needs to be taken. The sampling strategy (e.g. position and number of sample points) was also shown to influence the behaviour and sensitivity of the distance function. It was shown that the accuracy depends on the resolution of data sets and thus varies with the types of imaging studies. This accuracy was shown to be of the order of the average pixel size of the two original grey-scale images.

CHAPTER 4
SURFACE RECONSTRUCTION

4.1- Introduction

The main task and major application of the surface reconstruction technique, in this project, is found in recovering the 3-D external envelope of objects used in the registration process (surface fitting). Some reconstruction techniques involve *3-D boundary following* whereas in some others only *slice-level boundary detection* is required. The later needs to be complemented by a surface formation process between the contours of successive slices. The major difference between a number of surface reconstruction methods lies in the method of interpolation applied between the contours of successive slices.

A *stacking up* process is the simplest surface formation method achieved either by *replication* of contours or for example by a more complicated method such as *dynamic elastic* surface interpolation. In the later approach, as suggested by Lin and Liang 1988, a shape specifying function is used to define the intermediate contours. A *3-D boundary (surface) following* algorithm has also been introduced in different approaches; by *tracking the voxel faces* (Herman 1979, Artzy et al 1981) on a binary thresholded 3-D image, or by a *3-D edge detection* schemes (Liu 1977). A triangulation-based technique was suggested by Kepple 1975, for approximating a closed surface between two contours. These techniques have been extensively improved by other workers (Fuchs 1977, Wang & Aggrawal 1985). Alternatively, sectional curves can be represented by *parametric equations* (e.g. *B-spline*), and then the surface is interpolated by *cardinal splines* (Gordon 1969, Wu & Greenber 1977).

Surface reconstruction consists of several stages as discussed in this chapter. In general, these can be defined as surface representation, segmentation, and surface formation. Section 4.2 discusses different representation methods. The criteria for selecting an optimal representation method is outlined in section 4.2.1. Sections 4.2.2 and 4.2.3 give an overview of the various object (surface) representation methods and data structures. Section 4.2.4 outlines the representation methods used in the current project.

Two dimensional segmentation and edge detection methods are addressed in section 4.3. *Grey level thresholding* and *gradient based methods* are described as the basis of edge detection methods in section 4.3.2.1 and 4.3.2.2, respectively. The edge detection algorithm used in most of applications of the current work is outlined in section 4.3.3. In this respect, *curvature* measurement and a 'desirability' test known as the *circularity check* is described in section 4.3.4.

Surface formation needs to be applied where a series of 2-D contours are initially detected from the original cross sectional images. Different surface formation methods are based on the choice of the appropriate representation method to define an object, and image characteristics. Three different interpolation methods are described in section 4.4. These are known as *linear interpolation*, *triangulation*, and *shape based interpolation* as addressed in section 4.4.2, 4.4.3, and 4.4.4, respectively.

Section 4.5 describes the surface reconstruction of the SPECT images which are deformed by some tumours and abnormal structures. The problem of the correspondence between the registering points of the two surfaces in the deformed region, and some techniques to resolve these problems are discussed in this section. A summary of the surface reconstruction methods and their advantages and disadvantages are given in section 4.6.

4.2- SURFACE REPRESENTATION

4.2.1- Introduction

The measurement of the geometrical properties of an object (e.g. size, shape, connectedness) depends on how the object is represented. The problem of finding an effective and practical representation is primary to any 3-D object reconstruction scheme. The following are a number of criteria that must be satisfied in order to give an accurate and reliable representation for use in the matching process.

a) The representation must be computable in a practical situation where a large amount of data exists. These data require a large storage requirement and thus a large number of calculation are involved in handing such data.

b) The representation should easily manipulate the complexity of an object (i.e. medical object) and level of details required for the matching process.

c) The representation should be sensitive to any undesirable geometrical change made to the object. This means that it should reflect any undesirable mismatch between two images of one object taken under different imaging parameters such as position, viewing angle, and zooming factor. It is very important for determining a unique solution of the positioning problem for the registration process that the representation of the object be invariant under certain transformation parameters (rotation and translation).

d) Uniqueness in specifying a single object inside the image data coming from both imaging modalities is essential during the matching process. It prevents matching an object against other undesirable (different) objects.

A 3-D object in a scene can be thought as a function f(x,y,z) of three real variables (x, y, and z position in 3-D space), having a value even for a very small portion of it. However, due to the lack of ideal imaging facilities and computer hardware, only a discrete scene comprising a number of voxels (volume elements) is normally represented. Accordingly, the density assigned to a voxel [x,y,z] in a 3-D array is an estimate of the average value of f(x,y,z). The object of interest, in scene, occupies some of those voxels totally, some of them partially and miss the rest of them (in background). In the current project, these density values, themselves, were not considered to be important because the geometry and position of objects inside the scene were used as primitives for the registration process. This information can be extracted from a binary segmented object or from its edge data, and thus the representation problem is reduced to that of representing a simple binary scene or border information.

4.2.2- Overview of the representation methods

Among the different methods of object representation, the following are frequently used in computer graphics and pattern recognition.

a) *Parametric representation*, in which a shape type and a set of parameter values (mathematical equations) are used to specify an object. *Parametric bicubic patches* are

commonly used for representation of 3-D surfaces. Three mathematical equations (one for each x, y, and z) are used to define the coordinate of points on a curved surface. Each equation has two variables (parameters) defining the position of each point on the bicubic patches. *Hermite bicubic patches*, *Bezier bicubic patches* and *B-spline bicubic patches* are three different approaches to the parametric representation.

b) An object can also be represented by a list of cubical or non-cubical spatial cells (*spatial numeration*) in three dimensional space. These cells represent a discrete 3-D scene containing the object of interest. The scene region, in this representation, is considered to be a rectangular region subdivided by three orthogonal planes into smaller rectangular volume elements (voxels). The object in the scene is defined as a collection of connected voxels and can be represented either by the voxel itself (*volume-based method*), or by a set of voxel faces laying on object surface (*surface-based methods*).

Volume-based methods have been proposed by Meagher 1982 who described the object as a set of voxels in 3-D scene. The object can be represented and recorded as a 3-D array of voxels (Oswald 1983, Barillot 1985), an octree (Jackins 1980), or a sequence (list) of voxels laying on the object surface. Here, a priority order among the voxels is defined as an important approach for data access during registration or visualization.

Surface-based methods were first proposed by Herman and Liu 1977 and evaluated by Artzy and Frieder 1981. They represented an object by its surface, which is the set of all the voxel faces separating voxels inside the object from voxels outside the object in background. They defined connectivity and adjacency criteria for describing objects and their surfaces. The set of all the closed faces obtained in this way are represented and recorded as a sequence of voxel faces.

c) Finally, boundary representation treats the object as an enclosing surface (planes, quadric surface, patches, etc). *Polygon meshes* (or patches) are sets of connected planar surfaces (polygons) which approximately represent a curved surface (Keppel 1975). The disadvantage of this technique in some applications, however, is that it does not satisfy the criterion (b) required to represent the complexity and details of some medical objects.

In practical situations an object (or object surface) can not be represented exactly by all its details. This is due to the fact that only a limited number of points (e.g. integer variables) are defined by any representation. Even mathematical surfaces, in spite of their efficiency, may exactly represent the surface only at a set of points picked out on the object. Therefore, all surface representations are considered to be only an approximation of the surface of interest.

The representation capabilities of most of the above techniques are poor in satisfying all of the criteria outlined earlier especially that of the effectiveness in handling the details of complex medical objects. The *volume-based approach* was adopted, by which the level of information (details) is comparable to that of the original 2-D slices. However, a fine level of details can not be recaptured at the interslice spacing when the thickness or interslice gap is large.

4.2.3- Overview of data structures

The object of interest (e.g. surface) can be described by one of the following common data structures.

a) *3-D binary arrays*: are 3-D arrays of 0's in background and 1's inside an object or on the object surface. Good reliability and data access is achieved by describing the object by 3-D binary arrays, although this has to be paid for by use of a huge unnecessary data volume.

b) *Run length representation*: is slice-by-slice, row-by-row representation which treats the object as a collection of binary segments (Merrill 1973). Only the left and right end-points of all segments which represent a continuous run of object voxels (1-voxels) are recorded as a sequence of integers for further processing. The data accessability of this method is poor and it is not efficient for data registration.

c) *Borders*: is also a slice-by-slice representation in which the borders of object are represented by a sequence of voxels. In 3-D, the positional information of each voxel of the border (x, y and z coordinate information) are stored in a sequential list of elements. In a similar approach directed contours can be used which assign the directional information to the contour.

A variant of this representation deals with the border (e.g. in 2-D) as a sequence of moves originating from an arbitrary starting point and moving to one of the neighbours in some specified direction (Freeman 1974). The moves are recorded as a sequence of numbers (codes), where each number (0-7) correspond to one of the eight neighbours of the current point in 2-D. Restricting the moves to up, down, left and right is referred to a chain code. Diagonal moves can be added to the chain code in a type of structure called crack code.

d) *Octrees*: The 3-D scene is recursively subdivided into eight sub-scenes until each sub-scene is either full (inside object) or empty (out side object). The result is represented as a tree structure containing a number of nodes. The main disadvantage of this representation is that it is highly variant with respect to any transformation in that any small transformation (e.g. shift) may give a very different octree. Thus, it is very complicated to extract positional information with respect to the registration process.

Two important factors which effect the choice of an efficient data structure are the size of physical memory and the accessability required during the manipulation of an object (e.g. surface). The storage (physical memory) available in computers in the medical environment is sometimes smaller than the memory required for manipulation of two 3-D binary images of size 256*256*256. Using a list of surface data (x, y, and z coordinates of surface voxels), on the other hand, in a consecutive fashion does not provide a directly addressable memory where it is necessary to search for a specific voxel in 3-D scene. Some other pointer and block structure are needed in order to generate a number of smaller addressable blocks or lists of data (Frieder 1985).

4.2.4- Representation method

The *3-D binary array* has been used as the main data structure. An auxiliary data structures, a 2-D plane map was constructed to speed up the data checking (search) procedure and to reduce the size of memory allocation. The plane map has two indexes (x and y) identical to two coordinates of the 3-D scene. The value of each point in the plane map indicates whether or not the surface is passed through any depth (z coordinate) of the particular x and y location (coordinate). The value of 0 is assigned, if a surface does not pass through it, otherwise, a non-zero value is assigned to P(x,y), indicating the number and position of the surface points in z

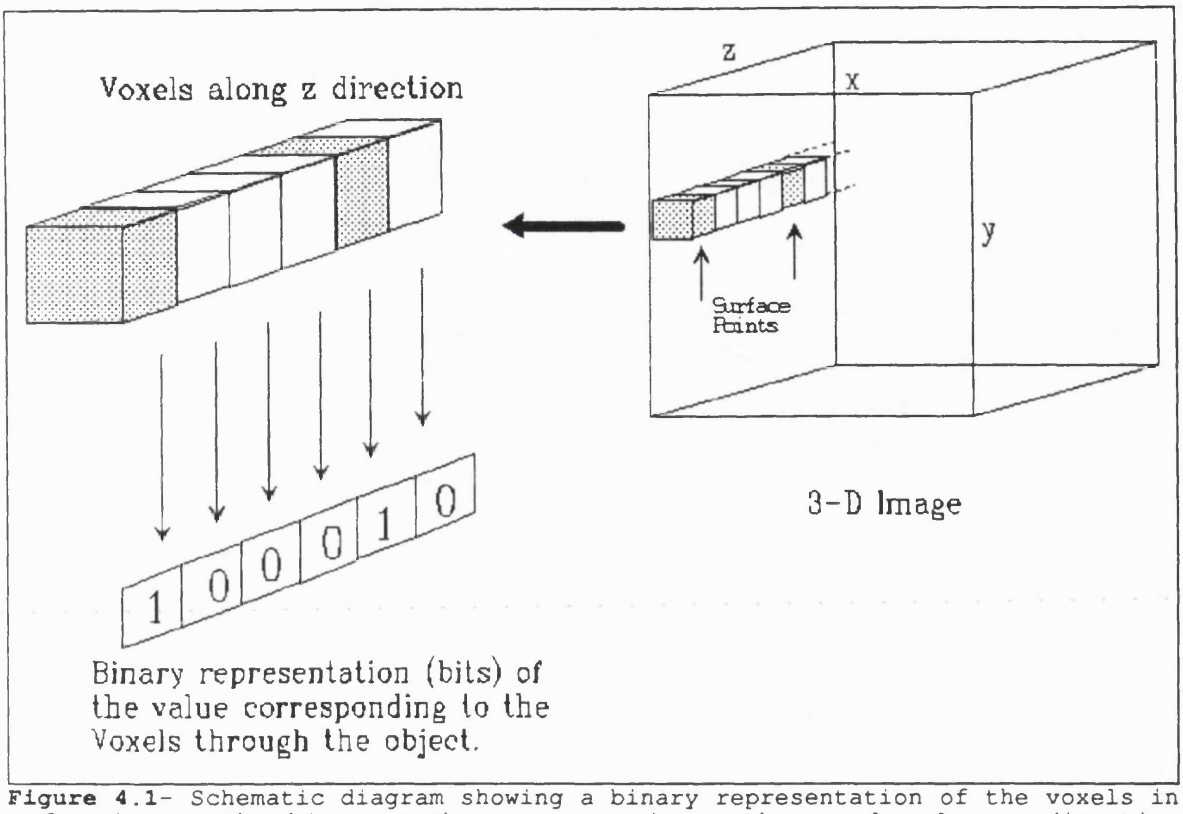


Figure 4.1- Schematic diagram showing a binary representation of the voxels in a 3-D image. The binary number corresponds to the voxels along z-direction (shaded voxels and 1's represent the surface points).

direction along the line from each (x,y) coordinates. In this respect, one bit is allocated in respect of each binary voxel in 3-D array (see figure 4.1). The voxels (e.g. 256 voxels) along z direction of a 3-D array can then be expressed by one or more integer values corresponding to e.g. 256 bits. In order to access (or search) the surface voxels, the integer values can be decoded into a binary number (or processed by a bit manipulation operation) showing the bits which have been set corresponding to the position of surface points.

4.3- 2-D SEGMENTATION

4.3.1- Introduction

Segmentation aims at isolating regions which consist of a contiguous collection of voxels (volume elements of an image) with different characteristic to their surrounding area. Let us take as the characteristic 'grey value'. Therefore, an edge, as a local feature, occurs where this grey value changes as the border between two adjacent extensive regions is crossed. It should be stressed that the grey level change depends, not only on object characteristics such as attenuation coefficient in *CT*, or radioactivity uptake in *Nuclear Medicine*, but may also be affected by *partial volume effect* due to surface slope, and noise in the image.

Different edge detection algorithms have been developed using different specifications and properties of an image. The edge detection technique suggested by Prewitt 1966 specifies grey level as this property and uses threshold to discriminate histogram peaks and valleys. Grey level uniformity was specified as a property by Bric and Fennema 1970. The use of derivative operators for edge detection dates back to 1955 and was proposed by Kovasznay and Joseph. In a similar approach, Ausherman 1972 suggested the maximum derivative of grey levels along one direction as a property. Step-fitting edge detection was used by Hueckel 1971. However, this was believed to be not an easy or suitable technique for boundary detection of complex images. An optimization approach to edge (curve) tracking was introduced by Montanari 1971 and Ballard 1973 who defined a decision function based on prior knowledge of the boundary. This function is minimized or maximized in a iterative method giving the minimum cost paths for the best boundary point selection. In general this method is computationally expensive due to its iterative nature.

In some complex images, using a single local property does not provide enough information to classify the pixels of the image. A more powerful technique needs to be defined for example using several local properties of a pixel itself and/or of its neighbouring pixels. For example, a threshold criterion can be used to determine whether the gradient value is large enough to accept or reject the presence of an edge point.

The boundaries can be defined as an ordered and connected subset of all pixel elements in an object, having some specified properties. These boundaries can be obtained by several stages of more advance segmentation processes. In this approach, the boundary of an object can be extracted by attempting additional segmentation processes based on connectivity (Herman and Liu 1977), size, and shape criteria (e.g. '*circularity*' as defined in section 4.3.4) on a pre-segmented edge image (using threshold or gradient operators).

Some of the above techniques can be applied either in a *sequential* or in a *parallel approach*, whereas some, such as boundary tracking methods, are inherently sequential. When using a sequential approach, the criteria for accepting a boundary point depends on the information obtained from earlier processing of other points. In a parallel approach, the decision made at each point is usually independent of the decisions at other (neighbouring) points. The successive iterations required for adjusting the decision to select an edge point make the parallel segmentation slow and computationally expensive. In contrast to the parallel segmentation, the sequential methods in which the decision at each point is made only once, are much faster. However, the sequential processes may require a further decision-making (i.e. backtracking process) at any detected edge point if an error is detected during the tracking stage. Consequently, it is desirable that a boundary detection technique, as defined by Liu 1977, be a single stage, sequential-type method and use some (voxel) properties to guide the boundary search in a small space.

4.3.2- Edge detection methods

As shown in figure 4.2 an ideal edge has a steplike cross section, which almost never occurs in a normal noisy and coarse resolution image. The lack of perfect resolution in an image affects the size of the blurred region substantially. Partial volume effects, also, produce an uncertainty in the location of object edge. Thus, the important question is where exactly the actual object edge occurs in this blurred region. The decision to use a *threshold-based method* or a *gradient-based operator* to extract some features, such as object edges, is the most important stage of boundary detection. Density thresholding is frequently used in most applications. However, it is believed (Rosenfeld and Kak 1982) that gradient methods lead to a more precise edge detection. In both cases, further edge processing may be needed, in order to validate the detected contours by studying the relation of some neighbouring points and angle between them.

This is the objective of the circularity assessment as described in section 4.3.4.

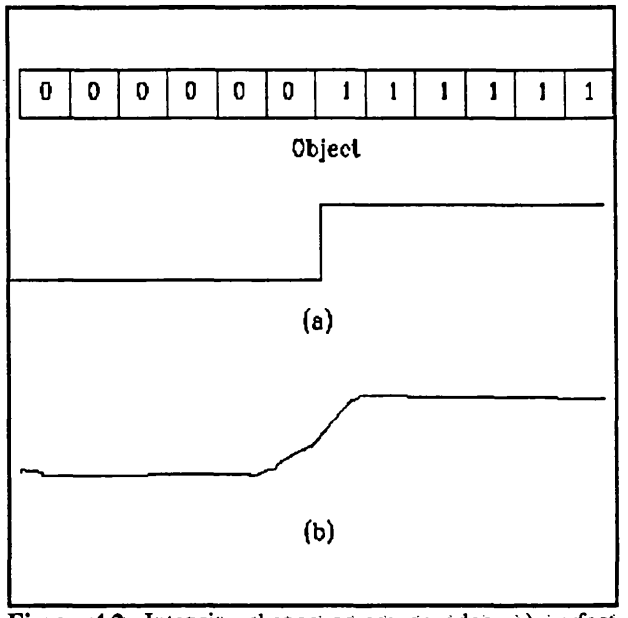


Figure 4.2- Intensity change across an edge; a) perfect step edge, b) Noisy and blurred edge.

4.3.2.1-Grey level thresholding

Thresholding is a widely used technique in image segmentation. It is used where an image has a low contrast. A variety of techniques based on grey level histograms have been proposed for automatic threshold selection (Prewitt 1966). Local thresholding (Katz 1965, Wolfe 1969, Morrin 1974) is applied where uneven illumination (grey level value) exists in different parts of an object. This method depends on some local properties (e.g. edge value [Katz 1965, Morrin 1974], or neighbourhood grey level [Wolfe 1969]) of the point itself, or a window around the point. The grey level information is also used with edge strength information to improve the grey level histogram forming two distinct peaks (regions) for threshold selection (Katz 1965, Weszka 1974).

4.3.2.2- Gradient based methods

Wherever an intensity change occurs (e.g. edge), there should be a corresponding peak in the first directional derivative (Figure 4.3), or equivalently a zero-crossing in the second directional derivative of intensity. Therefore, the task of detecting edges can be reduced to that of finding the peak of first derivative, or the zero-crossing of the second derivative.

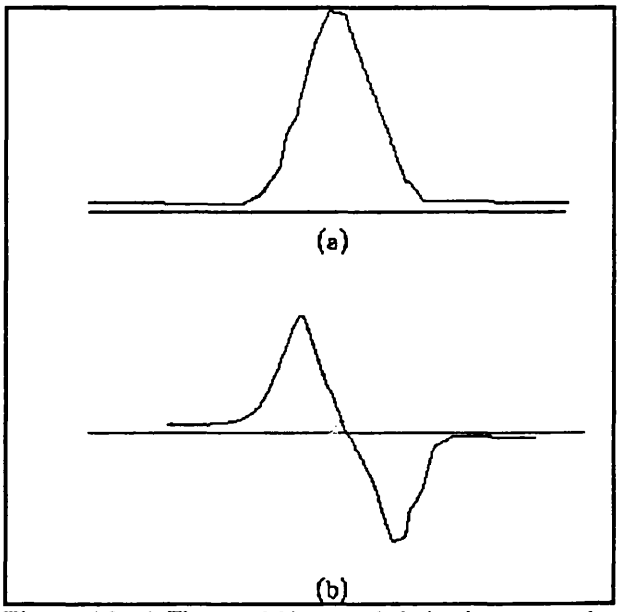


Figure 4.3- a) First, and b) second derivative across the edge shown in figure 4.1.

The gradient is a direction dependent operator. The gradient in any direction θ (measured from the X-axis) is given by

$$\partial f/\partial x' = (\partial f/\partial x)\cos\theta + (\partial f/\partial y)\sin\theta$$

where $\partial f/\partial x$ and $\partial f/\partial y$ are the rate of change of the function (f) in two perpendicular directions (x and y), and x′ expresses that the direction θ is measured from x-axis. The direction of maximum gradient magnitude (edge) is then given by arctangent operation, $\mathrm{Tan}^{-1}[(\partial f/\partial x)/(\partial f/\partial y)]$, and the magnitude itself is defined as $\sqrt{[(\partial f/\partial x)^2+(\partial f/\partial y)^2]}$. This magnitude is used where no directional derivative, as a measure of edge inclination, is necessary. The maximum of the absolute value between two directional gradients (e.g. perpendicular gradient) may also be used to represent the maximum gradient. For a 2-D digital image the following difference operations (e.g. Roberts Δ_+ and Δ_- (Gonzalez & Wintz 1977)) are used to define the gradient (Δf) of image.

$$\Delta_+ f(x,y) = f(x+1,y+1)-f(x,y) \qquad \Delta_- f(x,y) = f(x,y+1)-f(x+1,y)$$
$$\Delta_x f(x,y) = f(x,y)-f(x-1,y) \qquad \Delta_y f(x,y) = f(x,y)-f(x,y-1)$$
$$\Delta_{2x} f(x,y) = f(x+1,y)-f(x-1,y) \qquad \Delta_{2y} f(x,y) = f(x,y+1)-f(x,y-1)$$

The Laplacian filter ($\nabla^2 f$) which is an direction-invariant operator of the second derivative, can also be used as an edge detector. This can be expressed as

$$\nabla^2 f=[f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)] - 4 \times f(x,y),$$

which responds on each side of an edge, where there is a change in the rate of change of grey level (Top and bottom shoulders of a ramp as shown in figure 4.3b. It is useful where no directional information is desirable in the image. However, this type of the simple Laplacian is very noise sensitive and thus it is not normally used for boundary detection.

It is possible to reduce the effects of noise on the response of an operator by smoothing the picture before applying the operator. Using the average at adjacent pixels before taking differences leads to weakening the edge, since the neighbourhoods for the pixels overlap and therefore the differencing will cancel out the common values (Rosenfeld and Kak 1982). Some operators for the difference of averages have been introduced which take into account some adjacent but non-overlapping neighbourhoods. The following convolution filter is one of these operators.

$$\begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

The blur produced by the above operator, can be sharpened by suppressing non-maxima, over a distance less than the size of averaging neighbourhood, in the direction across the edge. Weighted averaging can also be used which gives greater weight to points lying closer to the central pixel (x,y). It gives a reasonable smoothing power reducing the effect of noise on edge detection (e.g. Sobel gradient averaging operator).

$$1/4\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix}, \qquad 1/4\begin{bmatrix}1 & 2 & 1\\ 0 & 0 & 0\\ -1 & -2 & -1\end{bmatrix}$$

4.3.3- Algorithm

Both, grey level *thresholding* and *gradient based methods* are used to detect edges of an object. The process is designed to switch between these two methods based on the characteristics of the image. In general, the same algorithm is used to track a boundary of an object using either types of edge detection.

The threshold methods are used where the edges of an object (e.g. brain) are too blurred and the image contrast is low. In complex medical images with a wide range of object grey levels, it is not possible to have a bimodal histogram from which the threshold be defined. In order to select a global threshold, two blocks having low variability were manually chosen, one from inside the object and from background. The average grey level of each block is used as the grey value of the histogram peak of the corresponding region. The threshold is then defined at the valley between the two peaks by calculating the mean value of them.

The gradient methods are used to detect the edges in a high contrast image such as *MRI*. In order to track a boundary, an edge direction map is used as well as the gradient magnitude (Robinson 1977). A set of compass gradient masks (see figure 4.4) which generalize the Sobel operator is formed by rotating the two masks expressed in the previous section, in eight directions. Note that, only the first four masks are sufficient to obtain both the gradient and the edge direction information. In general, these operators provide the difference of the weighted average which is desirable, as discussed in the previous discussion.

The edge direction map is also used to examine the local connectivity of the edge points.

As shown in figure 4.4, eight principal directions (k=0,...,7) exist in edge direction map tested by the compass masks. As a definition (given by Robinson 1977), if the direction at the centre of the 3*3 grid shown in figure 4.5 is k, then the edges are connected provided the directions of the proceeding and succeeding edge vectors are k-1, k or k+1. This means that the direction of a boundary point is not more than 90 degrees from the direction of the previous or next boundary point. The connected neighbours of pixel **P**(i,j) are defined as

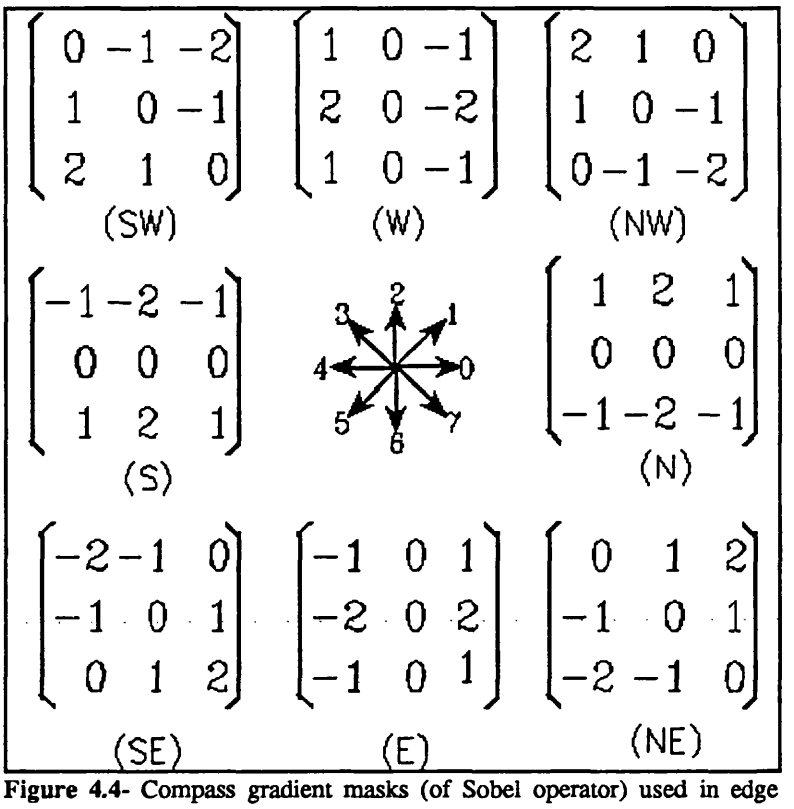


Figure 4.4- Compass gradient masks (of Sobel operator) used in edge detection process. The 8 directions shown in the centre are obtained by the compass masks and used for edge direction map.

$$N_{P(i,j)}=\{P(\ell,m)|\ 0<Max(|i-\ell|,|j-m|)\leq 1,$$

where ℓ and m correspond to (x,y) coordinates of any neighbouring pixel.

The boundary detection algorithm used in the current application was a sequential type process based on three criteria; *edge direction map*, *gradient magnitude*, and *connectivity*. *Curvature* information was also used as a shape based criterion to examined the detected boundaries based on a priori knowledge of the object boundary. The algorithm is outlined in the following sub-sections.

1) An initial boundary point is picked from a manually selected box in a high contrast region around the edge. The point which has the maximum compass gradient value is firstly selected from inside the box, and is placed on a boundary list.

2) The edge direction is defined at the current boundary point based on the compass

gradient masks. For the next point, only the neighbouring points which might satisfy the connectivity criterion are examined (i.e. tracking algorithm). This reduces the number of candidate connected neighbours to three points for each element.

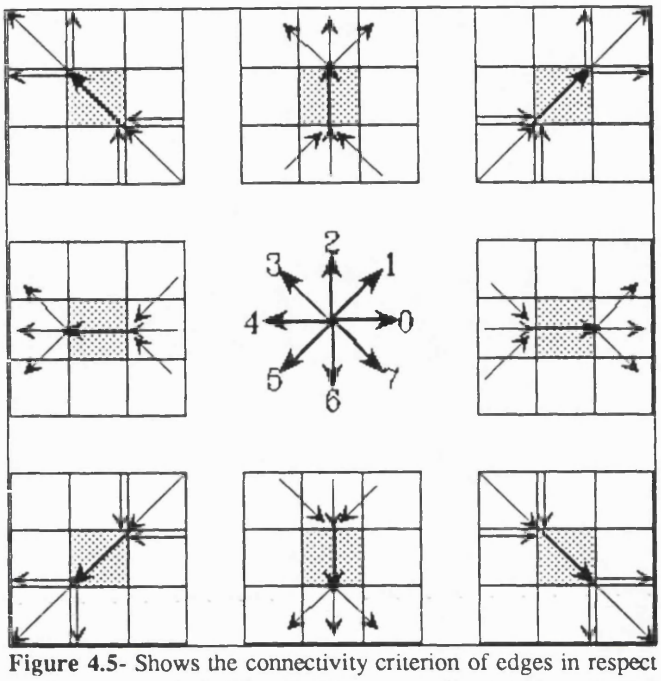


Figure 4.5- Shows the connectivity criterion of edges in respect to the central pixel. The direction map shows eight principal directions of the edge at the central pixel.

3) The point which has the maximum compass gradient value is first selected and compared with a pre-defined threshold. Then, it is considered as a candidate for the boundary if it has a gradient value above the threshold, and has not been previously marked as a bad point. All other neighbouring points on the track are stored in a separate list known as the backtracking list if their gradients are also above the threshold.

The threshold is measured using the mean gradient value of a high contrast region (box) on the object boundary where the tracking starts. This threshold can be updated locally as the boundary of the object tracked sequentially. T is defined as $T=aT_A$, where T_A is the average gradient value inside the box (for a fixed threshold), or along the boundary found so far (for a local threshold). 'a' is a tolerance factor, $a \leq 1$. Empirically, $a=0.4$ was found to be appropriate for all the applications used in this project.

4) The direction of edge at the current point is then determined by the direction of the maximum compass gradient mask, by which the connectivity check is performed. The current point is counted as a boundary element if it satisfies the connectivity criterion. If this criterion fails for the current point, the process is continued, examining the next promising neighbouring point, and going on to step 3. An array comprised of bad paths (which stores both the bad boundary elements and bad directions; i.e. the direction of the boundary ended to the current bad element) is generated and updated accordingly.

5) If the algorithm fails to find the next acceptable boundary point, a backtracking routine is called by which the elements in the backtracking list are examined. The first element from the top of this list is selected and the process continues if the path is not flagged as returning to the previous wrong path or the algorithm does not fail again in the new direction. Otherwise, the previous point in the backtracking list is taken into account and the process repeated.

6) If the initial point or one of its neighbours is detected during the process, the algorithm might find a close boundary. The process is then stopped (termination process) in the current slice. Steps one to six are repeated in the next adjacent slice. If the slice thickness or interslice spacing is not big (i.e. more than 2 millimetre), the selection of initial points can be automated using the connectivity criterion.

4.3.4- Curvature and circularity check

Slope and *curvature* are important contour segmentation criteria used by Freeman 1967 and Rosenfeld 1973, 1975, and reviewed extensively by Pavlidis 1977-1978. These two are the most useful geometrical properties describing the shape of a curve and the relational information among subsets of a single digital curve or segments of two adjacent curves. The slope at a point can be derived from its tangent vector, whereas curvature is defined as the rate of change of this slope. Both, slope and curvature at a point **P** can be defined as the slope and curvature of a line segment on each side of **P** along the curve. Using a line segment enables smoothing of the slope/curvature and thus the noise suppression on a contour is also achieved. The length of the line segment determines the amount of smoothing and should not be bigger than a straight line desired to match the curve segment starting at **p**.

Significant features on a curve are defined as the points where the absolute curvature is high (the slope changes abruptly), or where the sign of the curvature changes. These are defined as *corners* (angles) and *inflections*, respectively. The inflections give useful information about convexity and concavity of a curve, and are also important features of a contour image.

Two methods were implemented for calculation of the curvature. For a one-dimensional curve y=f(x), *curvature* κ is expressed as a function of arc length and in terms of the derivatives at a point x.

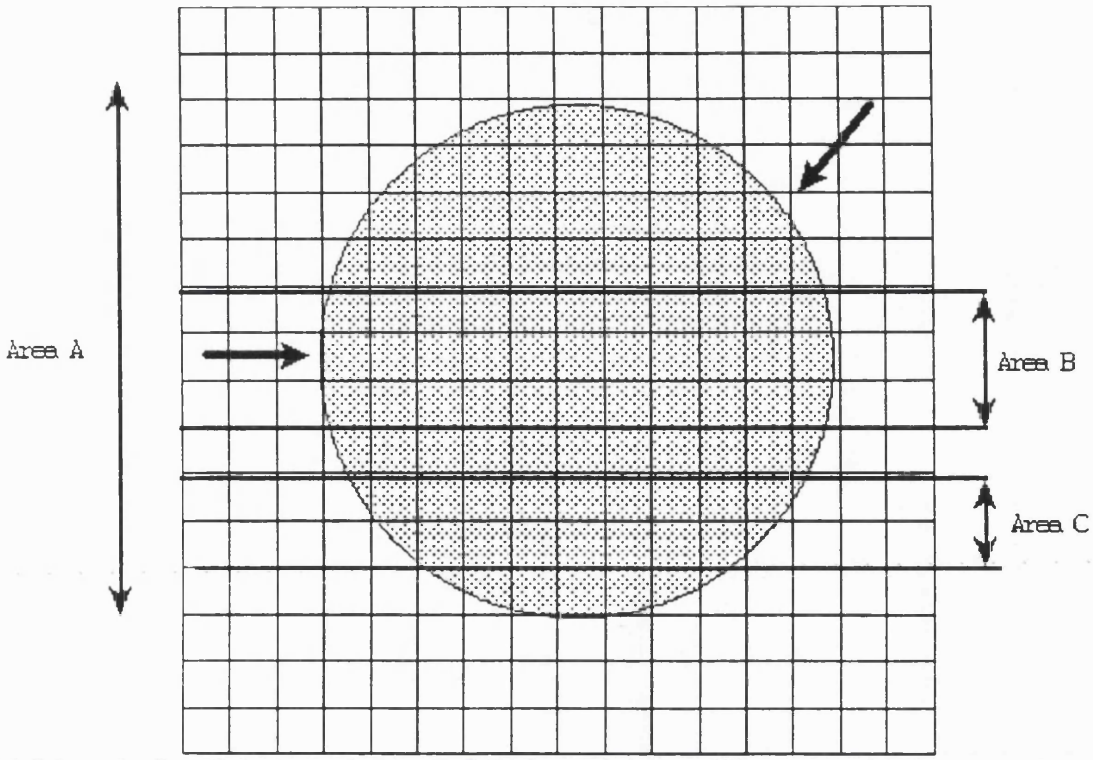


Figure 4.6- Schematic of a cylinder image showing partially occupied pixels at different parts of the boundary, where the slopes are different (see arrows) in respect to the basis coordinate axes.

$$\kappa(x) = f''(x)/(1+(f'(x))^2)^{3/2} = (d^2y/dx^2)/(1+(dy/dx)^2)^{3/2}.$$

where $f'(x)$ and $f''(x)$ are the first and second derivative of the function f(x). The curvature can be defined on a digital curve ($\mathbf{P}_i=(x_i,y_i)$, i=1,...,n) by replacing the above derivatives by differences. In this respect, a smoothed slope is considered to be $(y_{i+k}-y_i)/(x_{i+k}-x_i)$ for some integer $k\geq 1$.

The curvature can also be defined as the cosine of the angle between two k-vectors $\mathbf{a}_{ik}$ and $\mathbf{b}_{ik}$, where

$$\mathbf{a}_{ik} = (x_i-x_{i+k},y_i-y_{i+k})$$
$$\mathbf{b}_{ik} = (x_i-x_{i-k},y_i-y_{i-k}),$$

and the choice of k acts as a smoothing factor. The cosine of the angle between each two curve segments can then be expressed as

$$C_{ik} = (a_{ik} \cdot b_{ik}) / |a_{ik}||b_{ik}|,$$

$$C_{ik} = (\Delta x^+)(\Delta x^-)+(\Delta y^+)(\Delta y^-) \,/\, (\sqrt{((\Delta x^+)^2+(\Delta y^+)^2)}\sqrt{((\Delta x^-)^2+(\Delta y^-)^2)}),$$

where $\Delta x^+ = x_i - x_{i+k}$, $\Delta x^- = x_i - x_{i-k}$, $\Delta y^+ = y_i - y_{i+k}$, and $\Delta y^- = y_i - y_{i-k}$.

Having found the curvature at every point $\mathbf{P}_i$ (i=1,...,n) of a detected boundary image, their values are compared with a pre-set threshold value to check the desirability of the angles and curvatures of the boundary points (see fig 4.13). Since this curvature threshold value corresponds to a well-defined regular, and approximately a circular, shape of the head contour, this process is referred to the *circularity check.*

4.3.5- Experimental results

In order to see the effect of the choice of threshold value on the size of an object determined by an edge detection process, a cylindrical phantom having a known physical size was used in this experiment. This phantom was also used to calibrate and verify the parameters

ble 4.1- Phantom study using threshold and gradient methods to detect edges of a 2-D phantom image (cross section of a cylinder with wn physical size) at different resolution.

Applied Process	Res. *FWHM* (mm)	Dm by *Thrs*	Dm by *Grd*	Area *A* by *Thrs* (Pixels)	Area *B* by *Thrs* (Pixels)	Area *B* by *Grd* (Pixels)	Area *C* by *Thrs* (Pixels)	Area *C* by *Grd* (Pixels)
None	1.0	175	174	24208	3547	3539	3119	3112
Gaussian Filter	2.5	176	174	24352	3559	3542	3135	3115
	5.0	178	176	25196	3580	3541	3159	3120
	6.5	179	176	25620	3597	3546	3172	3117
	9.0	180	176	25852	3600	3547	3187	3120
	11.0	182	176	26324	3615	3544	3197	3118
	13.5	184	177	26810	3621	3534	3213	3116
Gaussian Filter + Guassian Noise	2.5	178	174	24620	3574	3540	3148	3110
	4.0	178	176	25136	3589	3548	3169	3111
	5.0	180	174	25688	3597	3542	3182	3129
	6.5	184	180	26160	3600	3549	3199	3142
	9.0	184	176	26436	3615	3545	3215	3099
	11.0	186	180	27050	3621	3560	3236	3195
	13.5	189	181	28520	3632	3561	3247	3165

Phantom study: Physical diameter= 173mm; Physical area = 23506; Threshold was set at half of image intensity. *Thrs* denotes threshold method, *Grd* denotes gradient method, and Dm denotes diameter of the cylinder. Area *A*= Total area of cylinder; Area *B*= Area of region around horizontal diameter; Area *C*= area of a region where the contour slope is high, as shown in figure 4.6.

(i.e. physical information) of the imaging modalities. The results of edge detection on a 2-D cylindrical phantom image are shown in table 4.1 and 4.2. As shown in the tables, the edge detection process which is based on a threshold criterion is very sensitive to a threshold value. The problem of selecting a proper threshold value is thus critical in determining the size of an object in the image. The maximum gradient across an edge occurs where the edge has its maximum slope, and thus can occur in the middle of a blurred edge, between the two shoulders, regardless of setting a user dependent factor such as threshold value. Therefore, the gradient method provides a more reliable edge detection algorithm.

Threshold	Diameter	Area of *A*	Area of *B*	Area of *C*
40	185	26535	3722	3334
50	183	26285	3686	3287
60	182	25900	3649	3251
70	181	25550	3611	3213
80	180	25280	3590	3178
90	178	24160	3556	3144
100	177	23780	3526	3115
110	*175*	*23425*	*3486*	*3077*
120	174	23080	3456	3048
130	172	22760	3422	3013
140	170	22385	3381	2976
150	169	22070	3346	2935
160	167	21800	3308	2891
170	164	21370	3248	2838
180	159	20780	3208	2770

Table 4.2. The effect of threshold level on the size of a cylinder image (max. pixel value=200) at resolution of 5mm *FWHM*. Areas *A*, *B*, and *C* are defined in figure 4.6.

The results obtained by area *B* and *C* in figure 4.6 confirms that different parts of the object border (e.g. surface) have different uncertainty in edge location (i.e. different partial volumes) with respect to the discrete voxels. The surface regions, which have a higher slope with respect to the orthogonal directions of the image (along which the square discrete pixels are aligned), undergo more partial volume effects. That also implies more uncertainty in the actual position of the surfaces (see figure 4.6). Figure 4.7 shows intensity-profile across the images of the cylindrical phantom, at different resolution and noise levels. As shown in the graphs represented in figure 4.7, edges of the cylinder image are more blurred (i.e. having a higher slope) as the image is degraded (i.e. resolution is reduced). Consequently, the measured size of the cylinder in the degraded image is more sensitive to the threshold value.

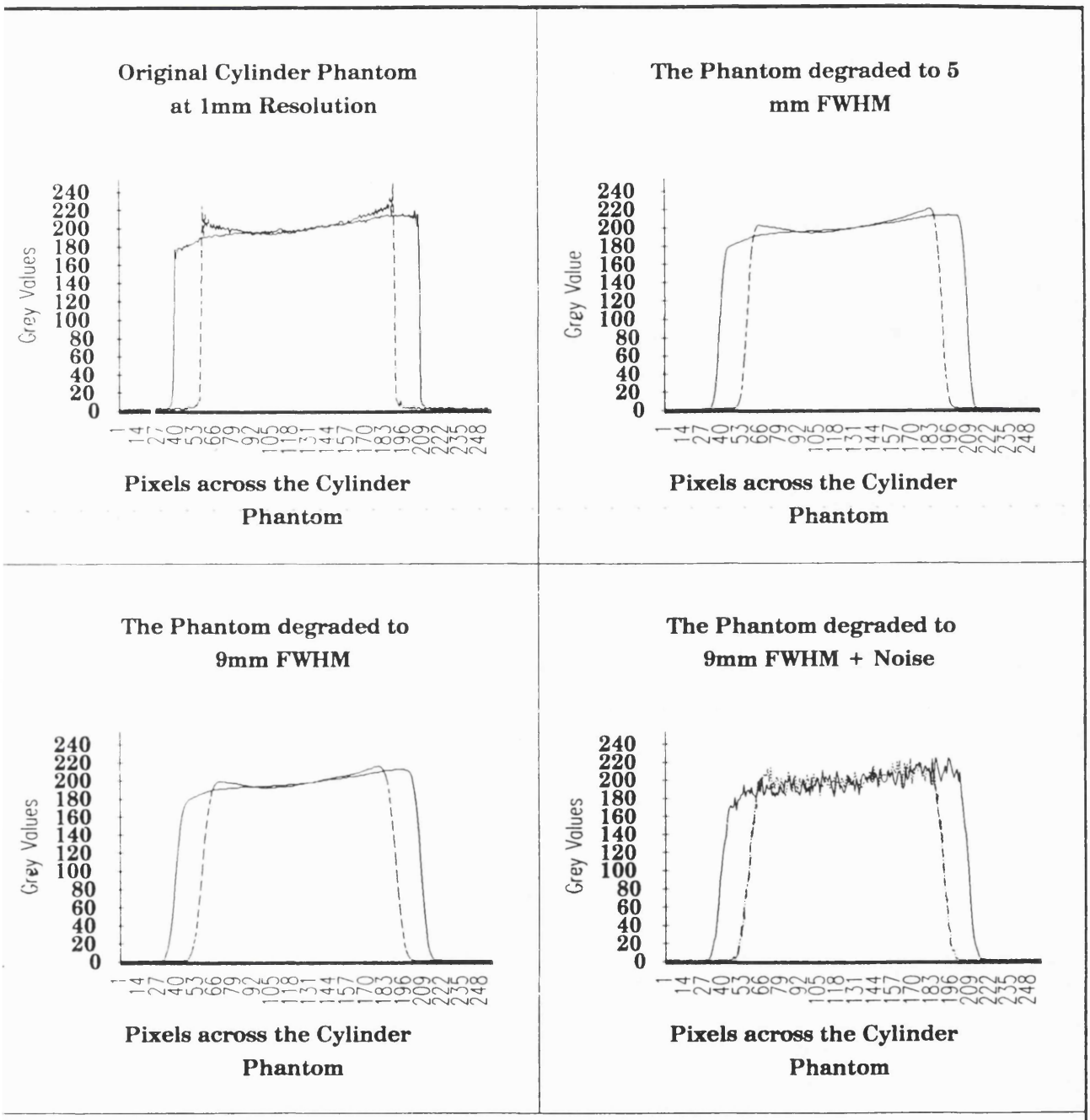


Figure 4.7- Pixel values (intensity-profile) across a 2-D image of a cylinder phantom at different resolution and noise levels. Solid lines show the profile through the cylinder diameter along the x direction (i.e. at area *B* as shown in figure 4.6). Dash lines show the results corresponding to region *C*.

4.4- SURFACE FORMATION

4.4.1- Introduction

The object of interest (the external surface of the head or brain) should have a smooth surface with connected boundaries, a regular shape, and an approximately circular outline. Exceptions are likely in some abnormal clinical situations where a lesion deforms the external surface. In the routine practice of 3-D medical imaging, 2-D cross sectional slices which comprise the original data, are often reconstructed by imaging systems. It is desirable to form a 3-D solid surface from 2-D or 3-D contours of a region of interest for both display and registration purpose.

The surface of a 3-D object can be captured (approximated) by stacking up the contours of the 2-D slices, if the slice thickness and interslice distances between successive slices are small. However, in many routine medical application only a series of 2-D thick slices are reconstructed. the contours of these slices are not sufficient when employing a simple stacking up process to reconstruct a surface, even if the interpolation of the interslice region is performed in advance. In this situation, the empty space between each two adjacent contours should approximately be captured by an appropriate interpolation process (e.g. triangulation).

4.4.2- Linear interpolation

In the current work, a series of intermediate slices were interpolated between each pair of adjacent cross sections using a 3-D grey level interpolation process, when the interslice gap and slice thickness were small (i.e. 2-3 millimetres). In this approach, the 2-D contours were generated using the techniques discussed in section 4.3.3. The 3-D surface was then generated by stacking up these contours.

The smoothness and continuity (connectness) of boundaries between two adjacent contours is hard to maintain by this method and hence it does not provide a true 3-D solid model of the object. However, this method was found to be sufficient for the reconstruction of external head surfaces to be used in the fitting operation, where the original interslice gap is not

significantly large (e.g. more than few millimetres).

4.4.3- Triangulation

If the slices are not closely spaced, the empty space between two adjacent contours are approximated using *triangulation* process. In this respect, an initial attempt is made to find the corresponding points or curve segments between adjacent contours participating in interpolation. Since the number of points **P** comprising (laying on) each contour may not be equal, and it is desirable to keep the size of **P**s small for the computational speed, the contours are segmented as explained below. Then only the breakpoints enclosing these segments participate in surface interpolation. Splitting and merging some of these segments might be necessary in order to initiate one-to-one correspondence between the segments of adjacent contours. In this approach, the problem of surface formation is attempted in three parts; contour segmentation (decomposition), segment matching, and surface interpolation between the curve/line segments, or on the other hand, between the breakpoints which separate the segments.

The triangulation starts from an initial pair of breakpoints on the two adjacent contours and proceeds through all the corresponding breakpoints. A dissimilarity measure can be defined between the two contours. In each step, if the dissimilarity measure between the two contour segments is higher than a predefined threshold, then the segment is split or merged with its neighbouring segment and the dissimilarity measure is recalculated. The breakpoints on adjacent contours are then grouped into triplets so that two of them from one sequence (contour) and one from the other reconstruct a planar surface patch (triangulation) between the two contours. Let one contour be defined by a sequence of m distinct points $\mathbf{P}_1,\mathbf{P}_2,...,\mathbf{P}_m$ and the other contour by $\mathbf{Q}_1,\mathbf{Q}_2,...,\mathbf{Q}_m$. A surface can then be represented by tile $\{\mathbf{P}_i,\mathbf{Q}_i,\mathbf{P}_{i+1}\}$, or tile

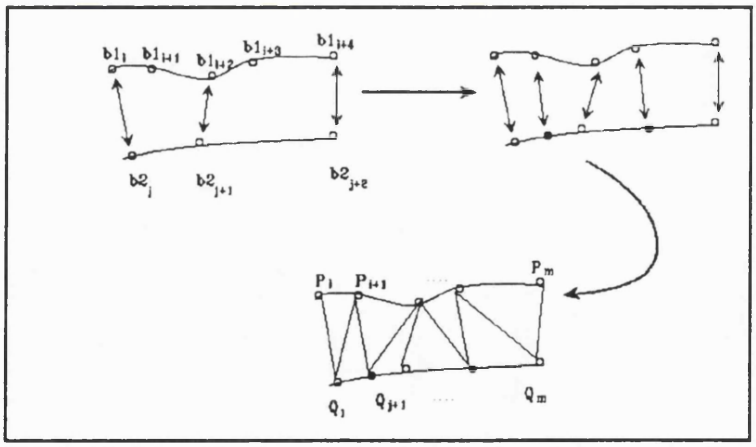


Figure 4.8- Schematic diagrams showing the triangulation process. Top left diagram shows breakpoints, top right: splitting the curve segments, and bottom: the triangular tiles.

$\{\mathbf{Q}_i,\mathbf{P}_i,\mathbf{P}_{i+1}\}$ (see figure 4.8). The problem of correspondence between the pairs of vertices (breakpoints) and selecting the optimal tiles, can be defined in a directed graph strategy, as suggested by Kepple 1975 and Fuchs 1977. They showed that there are a number of acceptable surfaces (tiles) between the vertices. Therefore, a cost path measure (Johnson 1977) was defined to choose an optimal surface from among them. In this respect, the minimum cost (C^{min}) path between a number of paths, was defined to be the minimum area of the associated triangular surface tiles. This minimum cost can be expressed as

$$C^{min}_{ij}=\text{area}(\mathbf{P}_i\mathbf{Q}_j,\mathbf{P}_{i+1}\mathbf{Q}_j).$$

(for i=1,...,n and j=1,...,n; n is number of breakpoints)

4.4.3.1- Breakpoint selection

Having found the curvature of every point $\mathbf{P}_i$ (i=1,...,n) as discussed in section 4.3.4, only a number of these point which have a high curvature (bigger than a threshold) are selected. These *high curvature points* are referred to as ***breakpoints*** (**b** as used in the triangulation) and approximately represent the shape of the original curve, with an accuracy depending on their number. in order to select the breakpoints from all over the contour, the curve is divided into m equal size curve segments, and one point having maximum curvature is selected from each segment. These local curvature maxima provide a better overall shape representation than the global maxima (values greater than some global threshold) whose distribution about the whole curve could be highly biased.

4.4.4- Shape based interpolation

A *shape-based interpolation* can be used to estimate the location of the intermediate voxels between the slices. As proposed by Raya and Udupa 1990, the binary segmented slice images are used and converted into grey level images, in which the grey values represent the distance from the boundaries (edges) of the object. In this definition, positive values represent the pixels inside the object and negatives represent those which are outside. The intermediate slices are then obtained by interpolating the distance-representing grey level slices. Finally, thresholding of the intermediate slices (e.g. at level of zero) creates the segmented object in a previously non-existent intermediate region.

The importance of the *shape-based interpolation* is due to the distance function used in converting the original binary segmented images. Various distance functions (called chamfer distances) were proposed by Borgefors 1984 and 1986, and revised by Herman 1991. The method proposed by Herman involves several successive stages which approximates the Euclidean distance more closely than the previous methods. The initialization proposed by him involved the replacement of the values of all the pixels outside the object by a very large negative value and inside by a large positive value. Those pixels inside the object sharing an edge with the outside, are assigned +15 and those standing outside sharing an edge with the inside are replaced by -15. In next stage, two templates (called near-optimal 3*3 or 5*5), as shown in figure 4.9, are used as masks placed on the image and moved in two processes, one row-by-row from top to bottom with a left-to-right ordering and one from the bottom to top with a right-to-left ordering within the rows. The pixel which is covered by the centre of the template is updated depending on the sign of its content, if it does not have the value +15 or -15. For each image pixel covered by a nonempty template pixel, the contents of the two values (image-pixel value and corresponding template-pixel value) are added if the image pixel value is positive and are subtracted if it is negative. In this approach, the current pixel to be updated is replaced by the smallest value of the computed sums where the original pixel value is positive, and is replaced by the largest of the differences if it is negative.

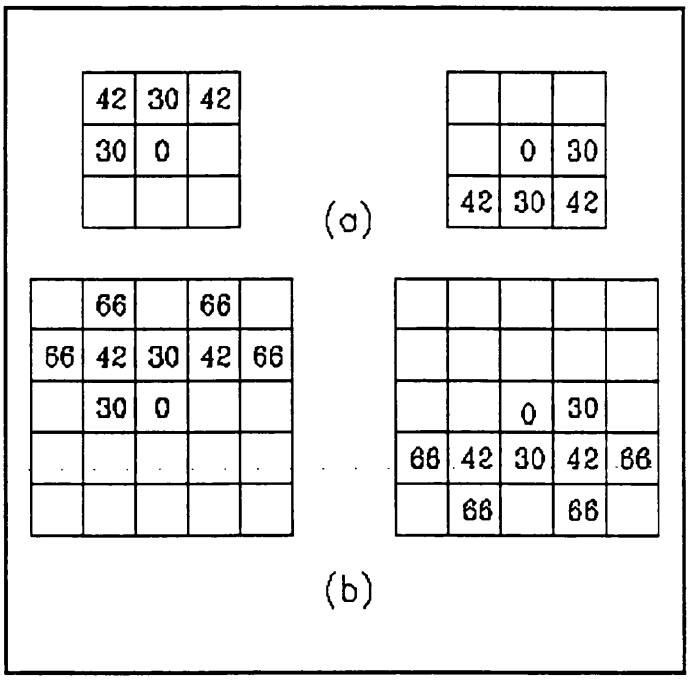


Figure 4.9- a) 3*3 and **b)** 5*5 templates used for shape-based interpolation (Herman 1991). Left and right templates are used for the first and second pass, respectively.

4.5- Surface reconstruction of special cases: Elimination of defects from surfaces

Figure 4.10 shows a typical *SPECT* image of a brain from a *HMPAO* radiotracer study. The irregularity of the brain edges is not unexpected in a normal brain structure. Presence of a

low intensity region (e.g. cold region or hole) is also possible in the *SPECT* brain image due to an abnormal brain tissue (e.g. tumour). However, normal, or even abnormal, structure of a brain might not alter the smoothness of the outer envelope of the brain image or its surface. In general, in order to create two nearly identical surfaces for the registration process, a smooth surface (i.e. with regular edges) must be reconstructed from the *SPECT* data.

As mentioned in section 4.3.4, a circularity check is performed on the contours obtained from the head. The main objective of the process is to detect any deformities, such as holes (cold regions) on the surface which could occur due to, for example, the presence of a tumour in the brain. The angular measurements from adjacent line segments provide an approximation of the change of the slope along the boundary. Comparing the rate of change of the slope with a constant threshold value (i.e. expected curvature used as a predefined threshold) indicates any possible undesirable deformity as well as the start and end points of this lesion. The images shown in figure 4.11a illustrate some *SPECT* images with a severe abnormality of the brain surface. The original contours detected by the contour tracking algorithm, and the points having undesirable curvatures, are shown in figures 4.13a and 4.13b (e.g. points having κ)-0.90 as shown in image D of these figures). During the circularity check, all points with undesirably high curvature are displayed to the operator. The operator can then manually alter and eliminate the unwanted points. The desirability of a curvature value in a deformed region is then automatically decided by setting a curvature threshold value. Visual inspection of displayed contours also plays an essential role in confirming or rejecting the automatic decision.

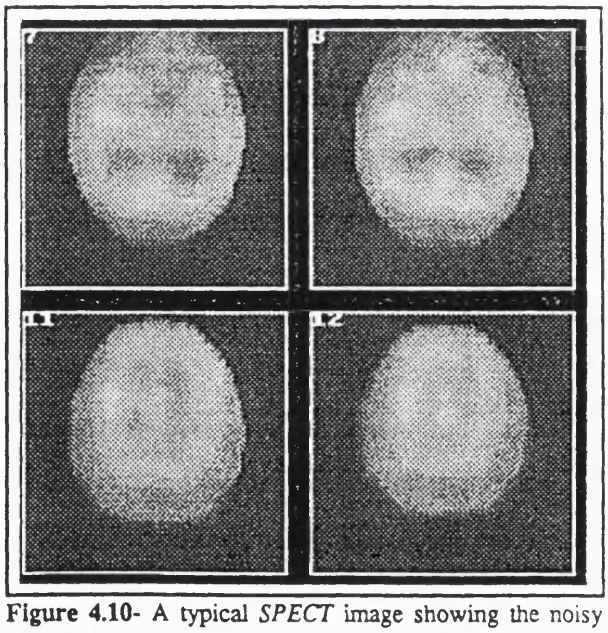


Figure 4.10- A typical *SPECT* image showing the noisy and blurred brain edges.

The technique used to reconstruct (reform) the deformed region depends on the size of the lesion. The application of the *median filtering* on the original 2-D grey-scale slices can 'reform' most of the holes and deformities, due to its different behaviour on long, low curvature, edges and corners (edges with sharp and high curvature). The maximum usable size of the filter is restricted

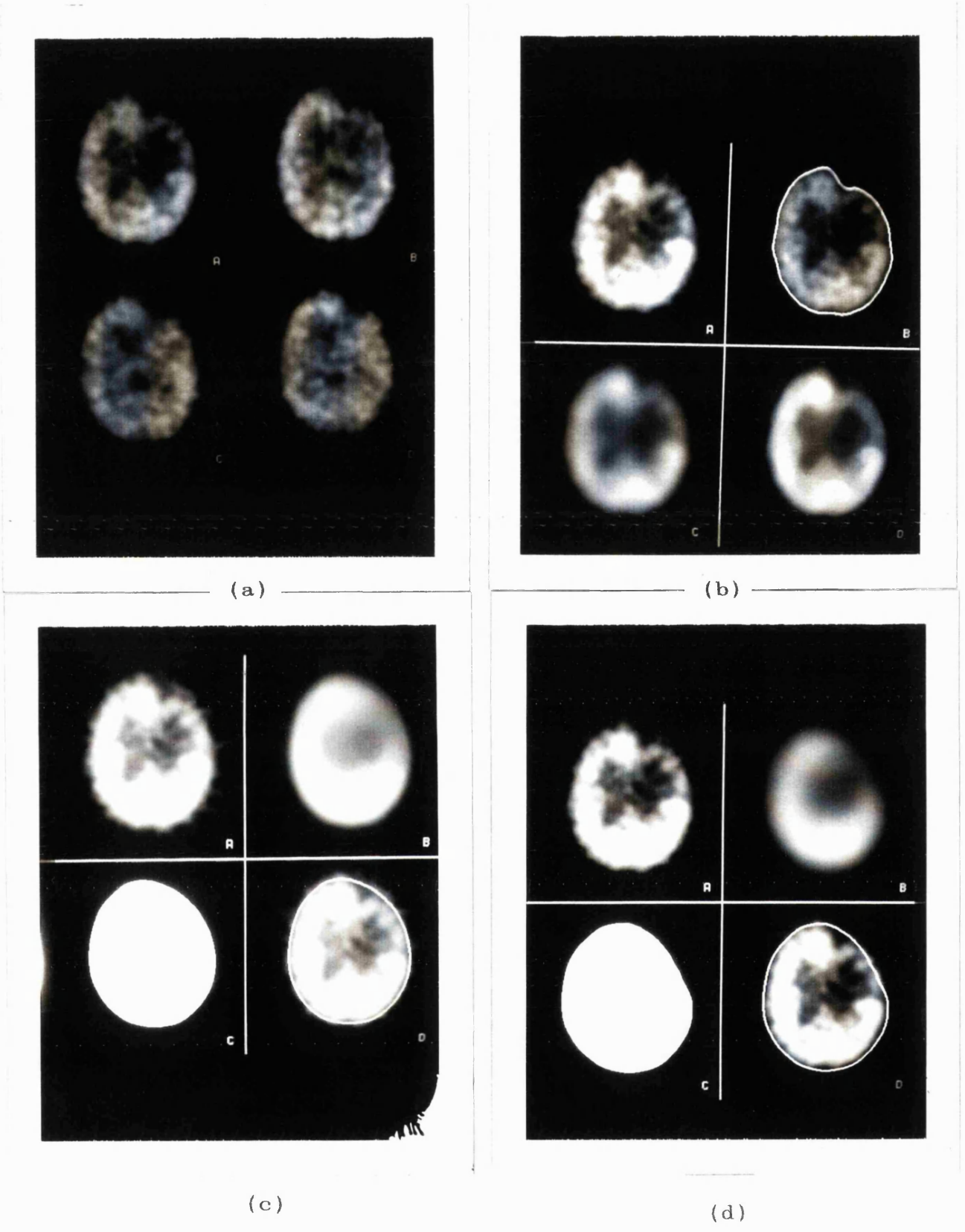


Figure 4.11- a) Four *SPECT* image slices having severe deformity on right frontal lobe. **b)** Shows the effects of a lowpass filtering (bottom left) and median filtering (bottom right) of size 21*21 on the top left 2-D SPECT image. The contour superimposed on top right image was obtained from the lowpass filtered image. **c)** Shows how median filtering (size 51*51) reforms the shape of brain and fills the deformity on a brain surface (B is the filtered image, C is the binary thresholded image of B, and D shows the superimposition of the contour of B on the original image A). **d)** Shows the effect of a lowpass filtering at the same size of the median filter (i.e. 51*51). The characteristics of the images shown in the four corners of figure **d** correspond to the image shown in figure **c**.

[3]by the size (i.e. area) of the normal regions, which occupy half of the filter size, when this filter is placed on the hole region. The effects of median filter of different size on both normal edges and abnormal gap (hole) regions are shown in figures 4.11b, 4.11c, and 4.12b. As shown in figures 4.11c and 4.12b, even a filter size of 51*51 can be used without any significant alteration of the normal edges. However, the use of the filter is limited only to a deformed region whose size is relatively small compared to the size of the normal region occupied inside the filter. In this respect, the size (area) of the normal region surrounding the abnormal lesion (hole), occupied inside the filter, should not be smaller than half of the filter size (area).

The results shown in figure 4.11d, demonstrate the effect of applying a lowpass filter of the same size as that of the median filter (figure 4.11c). For the segmentation of the image shown in figure 4.11d, half of the object intensity value was used. Figure 4.12a shows the effect

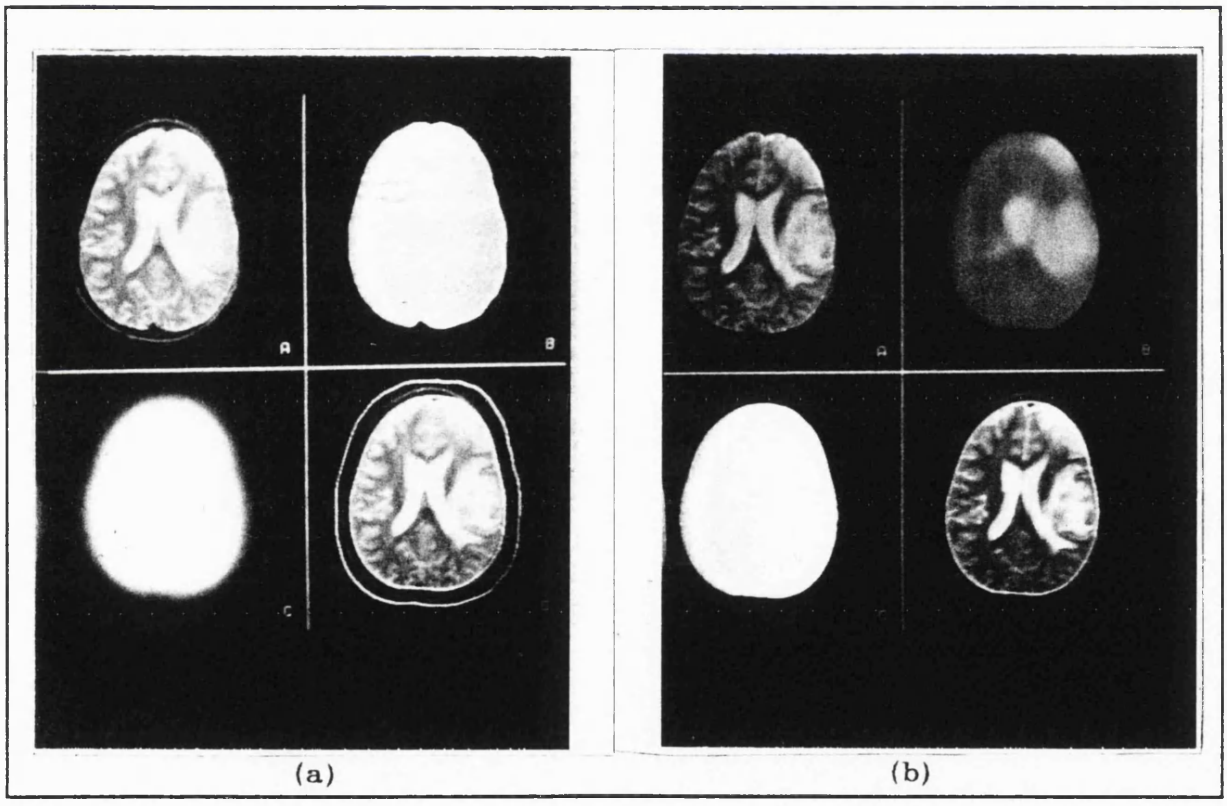


Figure 4.12- Shows the effects of lowpass filtering and median filtering on *MRI* and binary thresholded data. **a)** Top left image A shows the original *MRI* image, top right: binary thresholded image of A, bottom left: lowpass filtered (size 31*31) image of B, bottom right: The contours obtained at two different threshold levels (i.e. at half of the object intensity value, and at a smaller value), and superimposed on the original image. **b)** Top left image A shows the original *MRI* image, Top right corner (B) shows a median filtered image (size 51*51) of A. Bottom left is the binary image of B. Bottom right corner shows the superimposition of the contour obtained from B on the original image.

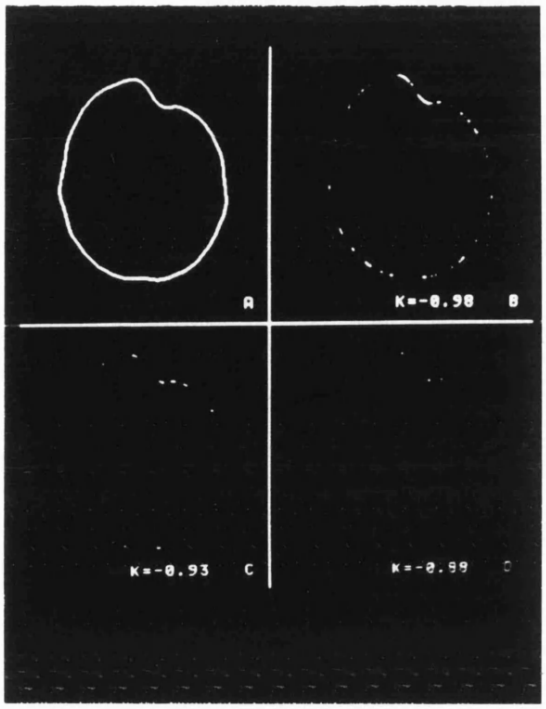
of threshold level in the detection of an object with its actual (correct) size. The critical problem is to set an appropriate threshold value in order to obtain the actual size of the object. Moreover, due to the similarity of the gap blurred region with the normal edge intensity (which is also blurred), it is not possible to precisely discriminate between these two regions by setting a global threshold on the image.

The superiority of median filtering over lowpass filtering to reform a hole or deformed brain region is also shown in figure 4.12. As shown in figure 4.12a, setting a sub-optimal threshold value on a lowpass filtered *MRI* image can enlarge the size of a detected object (e.g. brain). However, a median filtering, as shown in figure 4.12b, does not alter the object size substantially and can preserve the edges of a normal brain image.

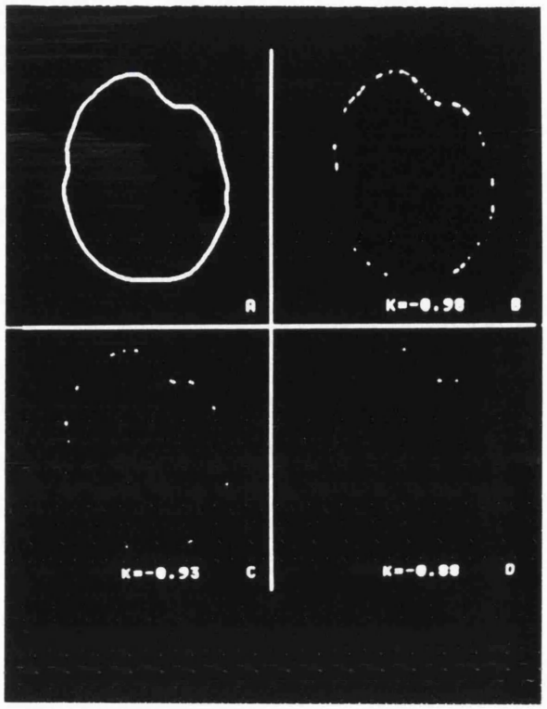
An alternative technique for the surface reformation is used by *reflection* of the normal part of contour from one side on the abnormal part of the other side, thereby ignoring the original detected deformed region. This technique uses the symmetry of the right side of the head to the left side of the head. The defective parts of the contour are identified by checking the contour circularity as discussed in section 4.3.4. The defects in the contours can also be identified or modified by visual inspection and possible manual interaction. Two principal axes (see section 4.5.1) are used as a base-line (reference) on the contour image; one dividing the 2-D head image (contour or binary segmented image) into equal left and right sections, and one splitting it into anterior and posterior regions. All the reflections are performed with respect to the principal axes. The result of applying this technique is illustrated in figures 4.13c-d. As shown in the figures, the principal axes are determined based on the normal structures of the image.

4.5.1- Principal axes

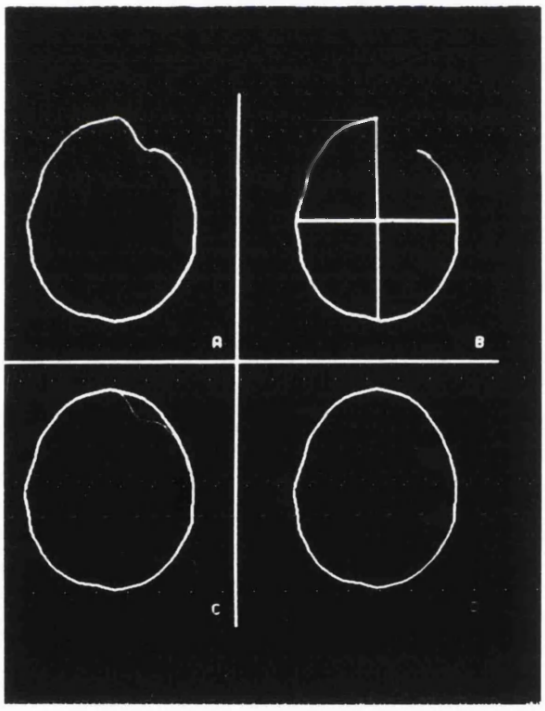
A closed curve such as the external contour of the head can be approximated by the connections between points on a number of straight lines passing through the centroid. In a given x-y plane, a straight line **L** can be represented as $\mathbf{L}(x,y)=ax+by+c=0$, with a slope of $\tan\alpha=-a/b$, where α is the angle between the line and the axis x. Let $\mathbf{P}_i(x_i,y_i)$ represents the points on a contour, the distance from each individual point $\mathbf{P}_i$ to a given line **L** can be defined as



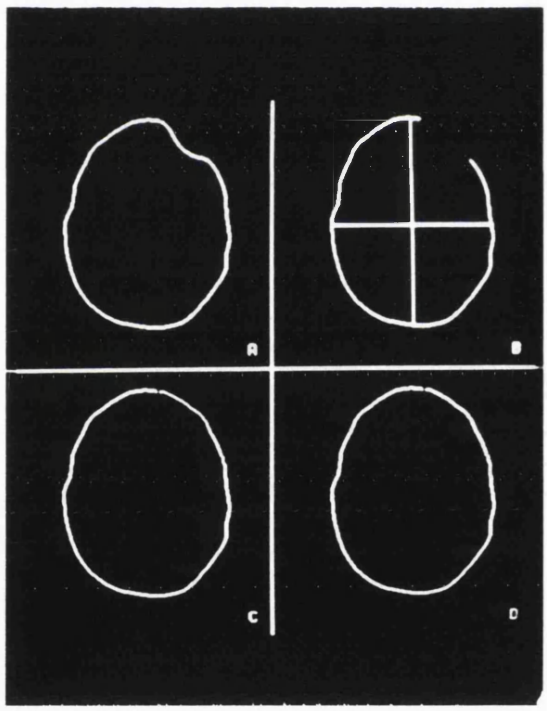
(a)



(b)



(c)



(d)

Figure 4.13- Top figures (**a** and **b**) show the points on the brain contours having different high curvature values (κ is the threshold curvature as defined in section 4.3.4). The points shown in frame D of figures **a** and **b** are the undesirable high curvature points (κ)-0.90). Bottom figures **c** and **d** show the reflection process used to reconstruct the defective brain region. Two principal axes are shown in the top right corners (B) of these figures. The bottom right images D are the corrected (reformed) contours using the reflection process.

$$d_i=(|ax_i+by_i+c|)/\sqrt{(a^2+b^2)}$$

Then the total distance of the contour from such a line is given by

$$D=\sum d_i, \text{ (for } i=1,...,n; \text{ n=number of points on the contour).}$$

For a binary image, the *principal axes* can be defined as two orthogonal straight lines through the centroid, for which the value of D is minimized (see figure 4.13c-d). Since during imaging process the head is routinely directed along the two orthogonal axes of the coordinate system, the longitudinal and transversal directions (on a transaxial head image) can be thought of as the intersection of sagittal and coronal planes with the transaxial slice, respectively. Consequently, the (transaxial) cross section of head image has a symmetric shape (i.e. left side and right side). In this respect, the principal axes are usually aligned along the orthogonal image coordinate axes. This cause the symmetricity of the head cross section with respect to coordinate axes as well. Some geometric transformations such as shifting, scaling, and even rotating the head around the longitudinal direction of the body do not change the symmetricity of the head image. The orientation of principal axes (i.e. axes of symmetry) can also be obtained, in moment based approach, using the maximum and minimum of the *moment of Inertia.*

The symmetry of the transaxial head cross section is an important factor for estimating the principal axes of an abnormal head image slice. The *principal axes* (*major and minor axes*) can be obtained by examining the direction of the principal axes in an normal image slice above or/and below the deformative slice. The centroid is then computed as the central point, on the principal axes, between the borders of the head contour. This process is performed automatically, but may be interactively controlled under the visual assessment.

4.6- Summary

A number of surface reconstruction methods, which are widely used by other workers, were studied in this chapter. All these surface representations are regarded as methods which only approximate the object (surface) of interest. In this respect, it is assumed that a number of points chosen from the whole structure, represent the object with all its required details.

However, even these discrete points themselves do not represent the exact characteristics and position of the actual surface details, due to resolution and partial volume effects, as well as the size of pixels in original grey-scale images.

The surface information can be extracted from a binary segmented object or from its edge data, and thus the representation problem is reduced to one of representing a simple binary scene or border information.

The representation capabilities of most of the representation techniques are poor with respect to satisfying all of the criteria outlined in section 4.2.1. However, the *volume-based approach* using a *3-D binary array* can maintain the level of information (details) comparable to that of the original 2-D slices, and provides easily accessible surface data.

Due to the nature of original grey-scale data in most medical applications, a *slice-level boundary detection* is required for surface reconstruction. A *gradient based 2-D edge(3^]^] *'3n* process was described and suggested for the use in most high contrast images. Alternatively, a *threshold based method* is recommended for the segmentation of objects and thus for edge detection. These techniques need a surface formation process to be employed between the contours of successive slices. Various interpolation methods have also been outlined, among them, shape based methods were found to be superior, and used extensively throughout this project.

Although, most of the abnormal structures of the brain do not alter the shape of the outer envelope of the brain image (surface), some abnormalities can deform the surface extensively. However, this may be a major problem in the registration, since two nearly identical surfaces are required for surface fitting process. A verification known as the *circularity check* for the shape of the detected head contours was defined based on the *curvature measurement.* Any unacceptable deformity (or hole) existing in the brain surface can be detected by the circularity check and 'reformed' by a type of interpolation process. Two techniques were suggested to 'reform' the abnormal regions and holes on the surfaces: one based on median filtering and another on contour reflection.

CHAPTER 5
MINIMIZATION METHODS

5.1- Introduction

The least square function as defined in section 3.2.5 depends on six independent variables (i.e. transformation parameters). The task of minimization is to find the value of those variables giving the minimum of a given function. Starting any minimization task, the location of the global minimum is an unknown inside a relatively large search space. Various methods introduced in the literature require different number of times for the function value or its derivatives to be calculated. In terms of computational cost, it is desirable to calculate a given function as few times as possible during the minimization process.

A minimum of a function can be either the lowest point in a finite neighbourhood (local), or the global lowest point. In general the global minimization methods are too slow to be practically useable, especially when the number of variables are high. Both types of algorithm can be either constrained where some priori limitations are applied on the value of the independent variables, or unconstrained.

Down hill local minimization methods may be divided into two approaches, the ones which search in transformation space in a multidimensional manner, and alternatively those which reduce the problem of multidimension to a series of one-dimensional line minimizations. A multidimensional minimization such as the downhill simplex method (Nelder and Mead 1965) is not very efficient in terms of the number of function evaluations that it requires, as well as the fact that it shares with other down hill minimization methods, the problem of trapping into local minima, for which a multi-start strategy is required. The line minimization based methods converge much faster to a minimum and thus starting the minimization from different locations is computationally more accessible.

Global minimization methods try to search the transformation space globally, either as a random process (e.g. stochastic minimization) or in some predefined order. Although stochastic

optimization, such as simulated annealing and genetic algorithms (Metropolis 1953, Mandava and Fitzpatrick 1989) go uphill as well as downhill and may avoid local minima, their computation is too expensive to be applicable for the minimization of functions (of a small number of variables) which can converge by an alternative method. However, their role is important in solving combinatorial problems where a very large discrete configuration space exists (e.g. the order of circuit elements in a complex integrated circuits).

The alternative global minimization strategy is the grid search which can be reduced in a multiresolution based approach as introduced in chapter 7.

The aim of this chapter is to describe some of the common minimization algorithms, and to define the methods which are mostly used in the current work based on their characteristics. The concept of search strategy in a minimization algorithm is addressed in section 5.2. Section 5.3 constitutes a selection of the widely used algorithms for unconstrained minimization. In this respect, the direction set method of Powell, conjugate gradient convergence method of Fletcher-Reeves, and gradient method of Newton-Raphson are outlined in section 5.3.1, 5.3.2 and 5.3.3, respectively. Section 5.4 presents linear search methods which are known as line minimization. The concept of constrained minimization and its use in a general minimization problem is presented in section 5.5. In section 5.6, the choice of the algorithm and the characteristics of a suitable method is outlined. A brief summary of the methods is given in section 5.7.

5.2- Search methods for minimization

A fundamental problem in the implementation of a minimization algorithm is to design a search strategy by which a point $\mathbf{x}_{k+1}$ is determined, given the points $\mathbf{x}_1,...,\mathbf{x}_k$ and the function values $f(\mathbf{x}_1),...,f(\mathbf{x}_k)$ and/or the derivative of the function. The direct search methods refer to methods which rely only on evaluating $f(\mathbf{x})$ at a sequence of points and comparing their function values, in order to obtain the minimum. The various approaches of these methods are as follow.

The simplest direct search method is known as '*grid search*' in which a rectangular grid of points is used to evaluate $f(\mathbf{x})$ at each of them in turn. The major disadvantage of this scheme is due to a large number of unwanted function values calculated in the search area. To overcome

this problem, a variable grid spacing can be designed instead of a constant step size. In a large grid spacing search region, the function is evaluated at an initial guess and a neighbourhood surrounding that point for the best function value. The minimum point is then considered as the new base point. In this approach, the grid spacing is reduced for further evaluation, but only if there is no substantial difference in function values within the neighbourhood.

Other methods have been implemented using different techniques for direction setting or different step lengths. The problem of minimizing a function of several variables can also be reduced to a linear (one-dimensional) search as defined in the next sections.

5.3- Direction set methods

Two problems should be considered in any multi-dimensional minimization procedure. Firstly, how a direction **d** is chosen, and secondly, what step is taken in that direction away from a current point to the next. In the direction set methods, multi-dimensional minimization is reduced to a number of one-dimensional minimization (*line- minimization*)(see section 5.4). A sequence of directions are constructed, and the function is minimized in each direction, separately.

Generally, it is desirable that a movement along a new direction $\mathbf{d}_{k+1}$ does not spoil the minimization so far obtained along the direction $\mathbf{d}_k$. This is the concept of mutual conjugacy being desirable in any direction set method. This criterion may be regarded as the major part of a minimization algorithm.

Various multi-dimensional minimization algorithms differ in the way that they define and update the next search direction. A diverse class of them are based on the fact that $f(\mathbf{x})$ (function value at point $\mathbf{x}$) increases, or decreases in direction $\mathbf{d}$ according to the sign of the directional derivative $\nabla f(\mathbf{x})$ (i.e. gradient methods). Although minimization along the basis vectors can be attempted, minimizing the function only in these directions is inefficient when many tiny steps are required in a long, twisted and narrow valley. Using gradient information, on the other hand, is the most powerful means of setting the direction $\mathbf{d}$ when the number of variables (directions) are large (Powell 1964). The methods of finding a direction in which to search for the minimum,

are divided into two main categories: the methods that do not require the evaluation of the derivatives, and those based on the directional derivatives. Although, both can depend on the properties of conjugate directions, different complexities and thus accuracies and timings can be obtained by them. Both types of method have been revised and improved several times by different workers since their initial implementations. However, in the following sections only the latest known version of these algorithms are presented.

5.3.1- Powell method

The *Powell* method is a type of gradient method which does not require the evaluation of derivatives. A quadratic function f(**x**) can be approximated by its Taylor series as

$$f(\mathbf{x}) \approx c - \mathbf{b}.\mathbf{x} + (1/2)\mathbf{x}.\mathbf{A}.\mathbf{x}, \qquad \text{Equ. 5.1}$$

where vector **b** and matrix **A** are the first and second partial derivatives of the function. The gradient of such function can be expressed by

$$\nabla f = \mathbf{A}.\mathbf{x} - \mathbf{b}. \qquad \text{Equ. 5.2}$$

Moving along some directions ($\mathbf{d}_{k+1}$) towards the minimum implies a change δ in gradient

$$\delta(\nabla f) = \mathbf{A}.(\delta \mathbf{x}). \qquad \text{Equ. 5.3}$$

As defined by e.g. Powell 1964, the gradient must stay perpendicular to the previous direction $\mathbf{d}_k$, while moving along a new direction $\mathbf{d}_{k+1}$, in order to satisfy a conjugate condition for the two directions $\mathbf{d}_k$ and $\mathbf{d}_{k+1}$. From the above equation this is just

$$\mathbf{d}_k.\delta(\nabla f) = \mathbf{d}_k.\mathbf{A}.\mathbf{d}_{k+1} = 0. \qquad \text{Equ. 5.4}$$

As stated by Powell 1964, a set of n (number of dimensions) linearly independent conjugate directions (or n line-minimization) will find the minimum of a quadratic function. However, in a real situation the function is not exactly quadratic and thus the process needs to be iterated using a number of such the direction sets. Powell's quadratically convergent method is a gradient method which does not require the evaluation of the derivatives, while it depends

on the properties of the conjugate directions. He used the *quadratic interpolation* method to find the minimum of the function f(**x**) along a direction **d** (for a set of n directions).

The algorithm finds a quadratic function (e.g. $Y(\lambda)$) which takes the same value as $f(\mathbf{x}_k+\lambda\mathbf{d})$ for three current values of λ (step), where $\mathbf{x}_k$ is the current point and **d** is a given search direction along the line $\mathbf{x}=\mathbf{x}_k+\lambda\mathbf{d}$. Having found the value of λ at the minimum (i.e. λ_m) of the quadratic function $Y(\lambda)$, one of the old three values of λ is replaced by λ_m. The process is repeated until the desired accuracy is obtained in the value of λ_m in each direction.

The directions $\mathbf{d}_1, \mathbf{d}_2, ..., \mathbf{d}_n$ are defined by the algorithm outlined in the following steps. The minimization in these directions are left for section 5.4.

a) Give an initial guess as starting point $\mathbf{x}_0$ in search space.

b) The n directions in the first iteration are initialized with the direction of the basis vectors

c) For k= 1,2,...,n calculate step-size λ_k by a line-minimization algorithm (see section 5.4), so that $f(\mathbf{x}_{k-1}+\lambda_k\mathbf{d}_k)$ is a minimum and define $\mathbf{x}_k=\mathbf{x}_{k-1}+\lambda_k\mathbf{d}_k$.

d) Find one of the old directions in which the function made its largest decrease ($\mathbf{d}_m$) and define $\Delta f=f(\mathbf{x}_{m-1})-f(\mathbf{x}_m)$, where $1\leq m\leq n$.

e) Calculate the function values $f_0=f(\mathbf{x}_0)$, $f_n=f(\mathbf{x}_n)$ and $f_d=f(2\mathbf{x}_n-\mathbf{x}_0)$ at a point somewhere further along the proposed new direction.

f) If $f_d\geq f_0$ and/or $(f_0-2f_n+f_d)(f_0-f_n-\Delta f)^2 \geq (1/2)\Delta f(f_0-f_d)^2$, repeat the above procedure (steps a to e) using the old directions $\mathbf{d}_1,\mathbf{d}_2,...,\mathbf{d}_n$ for the next iteration, and $\mathbf{x}_n$ for the next $\mathbf{x}_0$, otherwise do next step.

g) Minimize the function $f(\mathbf{x}_n+\lambda\mathbf{d})$ in a new direction $\mathbf{d}=\mathbf{x}_n-\mathbf{x}_0$, set $\mathbf{x}_n+\lambda\mathbf{d}$ as starting point ($\mathbf{x}_0$) for next iteration, and replace the old direction $\mathbf{d}_m$ by **d**, where the rest of the old directions would remain unchanged for the next iteration.

It was also shown by Powell that the way in which **d** is defined ensures the conjugacy criterion of all the directions after proceeding n iterations. As well as being simple, the method was claimed to be more efficient than those methods which are based on the evaluation of the function derivatives. However, it shares the problem of dropping into local minima with other direction set methods.

5.3.2- Fletcher-Reeves method

The *Fletcher-Reeves* method (e.g. *FRPRMN* method) is a type of minimization algorithm which involves the calculation of the first derivative of the function. A sequence of mutually *conjugate directions* were constructed by Fletcher and Reeves 1964, using the derivative of a function approximated by a quadratic form. The method makes each $\mathbf{g}_{k+1}$ (gradient of function) orthogonal to its immediate predecessor, and each $\mathbf{d}_{k+1}$ conjugate to its predecessor.

$$\mathbf{g}_{k+1}.\mathbf{g}_k=0 \qquad \mathbf{d}_{k+1}.\mathbf{G}.\mathbf{d}_k=0 \qquad \text{Equ. 5.5}$$

where **G** is the *Hessian matrix* (which is positive and definite) and $\mathbf{g}_k$ the *gradient vector* of $f(\mathbf{x})$. For a quadratic function $f(\mathbf{x})$, The minimum along the direction $\mathbf{d}_k$ can be found at the point

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda^*_k\mathbf{d}_k, \qquad \text{Equ. 5.6}$$

when moving from a current minimum at $\mathbf{x}_k$. The value λ^*_k is the step size required to minimize $f(\mathbf{x}_k+\lambda_k\mathbf{d}_k)$. The gradient at such a minimum is obtained for a quadratic function as

$$\mathbf{g}_{k+1} = \mathbf{g}_k + \lambda^*_k\mathbf{G}.\mathbf{d}_k \quad (k=1,2,...,n) \qquad \text{Equ. 5.7}$$

λ^* is chosen by a line-minimization algorithm (see section 5.4) to take the function to the minimum along the line where the new gradient $\mathbf{g}_{k+1}$ is orthogonal to the direction $\mathbf{d}_k$, satisfying

$$-\nabla f.\mathbf{d}_k = \mathbf{g}_{k+1}\mathbf{d}_k = 0. \qquad \text{Equ. 5.8}$$

The direction of downhill gradient (as in the steepest descent method $\mathbf{d}_1=-\mathbf{g}_1$) is used initially. Subsequent conjugate directions are chosen by the following direction sequence satisfying the mutually conjugacy of the directions.

$$\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \gamma\mathbf{d}_k \quad (k=1,2,...,n), \qquad \text{Equ. 5.9}$$

where the coefficient γ_k can be determined by the condition of the conjugacy ($\mathbf{d}_{k+1}\mathbf{G}\mathbf{d}_k=0$) and equations 5.7 and 5.9 which give

$$\gamma_k = (g_{k+1})^2 / (g_k)^2. \qquad \text{Equ. 5.10}$$

λ_k can also be determined from these equations (e.g. equ. 5.5 and 5.7) as

$$\lambda_k = \mathbf{g_k d_k} / \mathbf{d_k G d_k}. \qquad \text{Equ. 5.11}$$

Fletcher and Reeves suggested that the direction of search should revert periodically (e.g. in directions $\mathbf{d}_1$, $\mathbf{d}_{n+1}$, $\mathbf{d}_{2n+1}$,...) to the direction of downhill gradient when a non-quadratic function is minimized. Having found the conjugate directions, the method uses the line-minimization algorithm, like the Powell method, to find the minimum of the function along each direction. However, unlike the Powell Method the directional derivative of f(**x**) is used for line-minimization algorithm.

5.3.3- Newton type methods

The basic idea of the *Newton-Raphson* method is based on the approximation of the function f(**x**) by a quadratic function (see equ. 5.1) and taking into account the minimum of the quadratic. The derivative of the quadratic function at the point of minimum can be expressed as

$$\mathbf{G_k}.(\mathbf{x_m}-\mathbf{x_k}) - \mathbf{g_k} = 0, \qquad \text{Equ. 5.12}$$

The minimum $\mathbf{x}_m$ can then be obtained as the next point $\mathbf{x}_{k+1}$ in the search direction by

$$\mathbf{x_{k+1}} = \mathbf{x_k} + \lambda^*_k \mathbf{G_k^{-1} g_k}, \qquad \text{Equ. 5.13}$$

where $-\mathbf{G_k^{-1} g_k}$ is the direction of the linear search, and λ^*_k takes the function to the minimum along this direction. In order to avoid the difficulty of evaluating matrix inversion, a new class of gradient method known as *quasi-Newton* (or *variable metric*) method is used (Broyden 1967). In this method, the matrix $\mathbf{G_k^{-1}}$ is replaced by a positive definite symmetric matrix $\mathbf{H_k}$ (e.g. $\mathbf{H_1}$ initially is as unit matrix I) which is updated in each iteration. The resulting equation is then

$$\mathbf{x_{k+1}} = \mathbf{x_k} + \lambda^*_k \mathbf{H_k g_k}. \qquad \text{Equ. 5.14}$$

The algorithms require calculation of the function derivatives, thus can be very time consuming.

5.4- Line-Minimization

It is possible to localize a minimum between two one-dimensional points (*bracketing*), easily, if the multi-dimensional minimization is reduced to a one-dimensional problem (*line minimization*). The *bracketing process* starts with an initial guess for two points **a** and **b** on the search space, which give two distinct function values f_a and f_b. Then a third point **c** is selected in such a way to satisfy the following relation between the three points.

$$f_b<f_a \quad \text{and} \quad f_b<f_c, \qquad \text{Equ. 5.15}$$

where **a<b<c** and these three points bracket a minimum of the function (see for example points 1, 2, and 3 in figure 5.1).

The step size introduced in the process proceeds downhill and is obtained either by using a constant factor or as the result of a *quadratic* or *parabolic extrapolation.* A constant step size is used where the function is not properly quadratic near to the minimum. In this respect, the optimal bracketing interval **a<b<c** has its middle point **b** a fractional distance (see section 5.4.1) between the two ends. The *parabolic extrapolation* method tries to find a new extrapolated point on a quadratic function (parabola) fitted on the original points (see section 5.4.2). In either of these methods, one of the old points is then discarded in the respect of the new point included in the bracket. This process is repeated until it yields its new point in a uphill trend.

Given an optimal triplet of points, the best minimum can be found by stepping downhill into the triplet. Like the bracketing algorithm, the minimization stage uses one of the above methods; constant step size, or the step size based on the result of a parabolic interpolation.

5.4.1- Golden section search

The step length can be defined based on the *Fibonacci* sequence ($F_n=F_{n-1}+F_{n-2}$). This is the concept of a technique known as *Golden Section search* which is used for minimization of a unimodal function. The Golden ratio to define the interval (step length) between the evaluating points is obtained by the *Fibonacci* relation

$$\lim_{n\to\infty} F_n/F_{n+1}=(\sqrt{5}-1)/2=0.618034 \qquad \text{Equ. 5.16}$$

by which for example a point **b** divides a line segment **ac** with the relation **ab/bc = ac/ab.**

The constant step size defined in this technique yields a fractional distance of 0.61803 from one end and 0.38197 from other end of the bracketing interval. The new point obtained by this method is symmetric to one of the initial points, and therefor three resulting points have always the same fractional relation as the three initial points have.

Similar to the bracketing routine, one of the old points is discarded and replaced by the new point in the triplet, if its value is greater than that of the new point. Consequently, this technique provides each new triplet interval just 0.61803 times the size of the preceding interval.

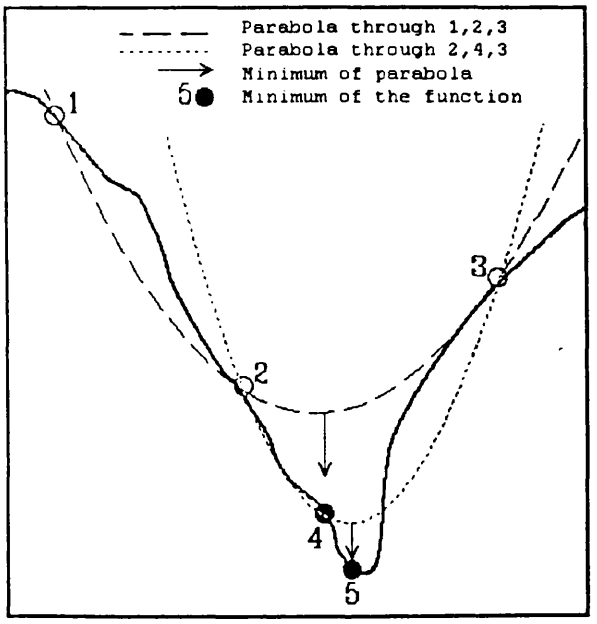


Figure 5.1- Minimization using parabolic interpolation. Function is evaluated at parabola's minimum, 4, which replaces point 3. The minimum of the second parabola nearly corresponds to the minimum of the function.

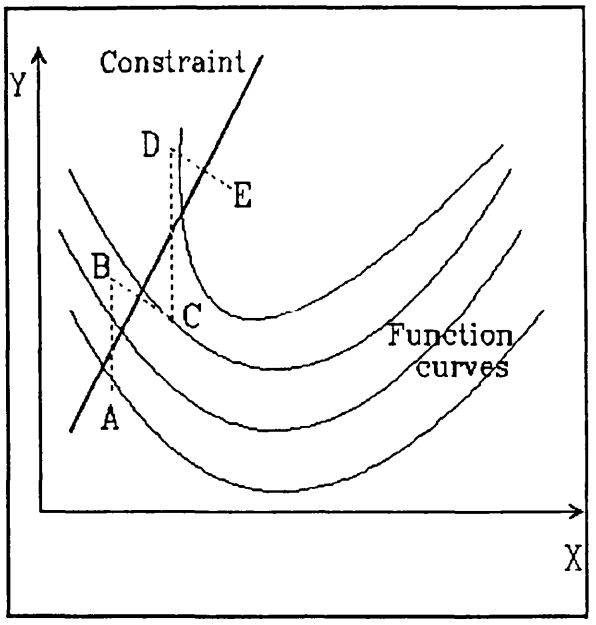


Figure 5.2- 2-D constraint minimization. The thick line shows the constraint boundary. AB and CD represent direction in which the function is minimized. BC and DE correct the process when the constraint is violated.

5.4.2- Quadratic interpolation method

A linear minimization method was defined by Powell 1964, in which a quadratic function is interpolated through three interval points (known as triplet) and its minimum is considered as a new point in the triplet. The process is iterated to obtain the desirable minimum.

The technique has been revised and used by Brent 1973 for functions which are parabolic near to the minimum. In this approach a parabola is fitted through the three points of the bracketing triplet, and its minimum is considered as a new point which is near to the actual function minimum. This procedure also needs to be repeated few times in which discarding one of the old points and adding a new parabolic minimum points, until a function minimum is achieved (see figure 5.1).

5.5- Constrained minimization

The general aim of a *constrained minimization* is to reduce the minimization problem to a limited search area and to eliminate any non-feasible point from that search area. All the methods for dealing with unconstrained minimization can be extended to allow for the presence of constraints. Generally, the procedure is to transform a given constrained problem into an unconstrained problem and then solve that problem. In a common method known as Hemstitching (Robert and Lyvers 1961), the constraints are ignored until one or more of them is violated. The current point is then returned to the feasible region in for example a direction orthogonal to the constraint (see figure 5.2).

5.6- Selecting an algorithm

The choice of minimization method is very important in computer cost (timing and space) and accuracy of the true global minimum of a function. The followings are some of the criteria and factors influencing the choice of the minimization algorithm.

1) Non-iterative methods are much faster than the iterative methods. However, the use of a non-iterative method is restricted due to the nature of some registration problems.

2) In general, any effort to use techniques which require evaluation of the derivatives of the function, should be robust enough to compensate for the additional calculation of the derivatives.

3) Storage requirement should be considered with respect to those methods which require only of order n (number of dimension), and those that require n^2 storage.

4) The methods which reduce the multidimensional problem to a number of one-dimensional problems (i.e. methods requiring only a line minimization algorithm) are potentially of greater interest, because these methods seem to provide an easy and fast minimization algorithm (e.g. direction set methods such as Powell's method).

In the surface fitting algorithm discussed in chapter 3, an iterative method should be used since the actual correspondence between the points are not known in advance. In this respect, each set of corresponding points needs to be updated in respect of the latest promising surface position obtained by applying a set of transformation parameters. In considering the above factors, direction set methods were mostly used throughout this project. Two *direction set* methods, *Powell* and *FRPRMN*, both reduce the multi-dimensional minimization problem to a series of simple one dimensional tasks. *Newton type* methods need higher precision arithmetic and thus are more time consuming.

5.7- Summary

Various methods for solving the minimization problems were described in this chapter. These methods are based on repetitive steps (iterative approaches). The simplest type of algorithm was defined as *direct search* (e.g. grid search) method which uses function evaluations and comparisons, together with a strategy for determining the next search point.

Golden section search and *quadratic interpolation* methods were addressed in a linear search strategy obtained by the conjugate direction methods. Gradient methods were used to generate a sequence of linear searches along successive directions of a multi-dimensional function. Some of these methods, e.g. the *Powell* method, the *Fletcher-Reeves* method, and the *Newton type* methods were presented which are the most efficient general purpose minimization techniques available at the present time. However, the use of the Flecher-Reeves algorithm and Newton type methods which require high-precision arithmetic and thus can be very time-consuming due to calculation of derivatives, is limited in most applications.

In this chapter the properties of a suitable minimization algorithm were also outlined and the best algorithm satisfying most of these characteristics was defined.

CHAPTER 6
SEQUENTIAL METHODS

6.1- INTRODUCTION

The shape of the human head is not strictly symmetric and there should be a unique transformation which could be expected to match exactly two images of the head. However, a number of transformations exist which produce incorrect matches having local minima in the distance function. This is due to the partially regularity and symmetricity of the head shape. The problem implies the need to check all the possible promising transformations which might be a candidate for obtaining the global true match. On the other hand, evaluating the function at an insufficient number of search locations (transformation parameters) might fail to find correct match location, that is, to give proper transformation parameters for registration. As discussed in section 3.2.5, a mean distance error (MDE) is required by the surface fitting process introduced so far, due to the nature of the minimization process (because the MDE value which is essential for setting the next search direction, is obtained after evaluating all the sample points). Accordingly, all the n extracted features (corresponding points) have to be evaluated at each transformation, prior to any decision on the correctness of the fit. Considering the two above requirements would result in a very slow convergence algorithm due to evaluating all points at all search locations. On the other hand, not to do this is the cause of the major disadvantage of the algorithm, that is, the detection of local minima (wrong minimization).

However, the accuracy which can be obtained by evaluating all the sample points is required only for those relatively few locations (transformations) near the match location. Hence, at a vast majority of locations, it would be a considerable waste of time to perform high-accuracy calculation using all sample points. This is the basic idea behind the introduction of a new technique for the surface fitting algorithm. In the minimization process applied in this new technique, all the possible transformation parameters are searched in order to find the best parameters. This process is known as a global search and is based on the grid search explained in section 5.2. Assuming that the scaling parameters of the images are known by the system parameters itself, the search space has six dimensions, three shifts and three rotations, to be

searched during the minimization. The way that the error is computed and subsequently the decision is made about integrity of any transformation, is a sequential process. Accordingly, this new technique is known as the *sequential method.*

A class of sequential similarity detection algorithms was introduced by Barnea and Silverman 1972. In their method, a translational registration was implemented by a template based matching process. The method was extended and a more complicated sequential matching algorithm was introduced by Ramapriyan 1976, in which a template was partitioned and translational shift was applied in different size-level. The method was then assessed and used by other workers (Rosenfeld and Vanderbrug 1977, Vanderbrug and Rosenfeld 1977, Wong and Hall 1978, Wong 1978, Wong and Hall 1979). However, all the modifications made so far to the algorithm, only concern the applicability of the process to 2-D images (specifically optical and radar images) undergoing translational registration.

The concept of the sequential method and its algorithm are presented in this chapter (see also IPMI'91; Oghabian & Todd-Pokropek 1991). In section 6.2, the theoretical discussion and properties of the method are explained. The algorithm is also outlined in this section. The properties (e.g. size) of search space (transformation parameters) used in the sequential process are outlined in section 6.2.1. The sequentiality of the cumulative distance error and its properties are described in section 6.2.2. Different sources of error and their contribution to the expected distance error are outlined in section 6.2.3. Section 6.2.4 and 6.2.5 will describe two types of threshold methods (i.e. constant and variable) used in this technique. The expression for the threshold sequence which depends on the number of sample points involved in the registration process, is presented in section 6.2.5. Some properties of these sample points are presented in section 6.2.6. Computational aspect of the process and cost expectation are presented in section 6.3. Section 6.3.1 provides a quantitative analysis of the search locations and sample points based on which the cost expectation was made. Finally, a brief summary is given in section 6.4.

6.1.1- Terminology

Matching process is used synonymously with *'registration process'* to find the best transformation parameters by which the two registering surfaces are matched. The term *'location'* refers to search location and that is a set of transformation parameters (i.e. 3 shifts and 3

rotations in x, y, and z direction) at which the matching is evaluated. *Search window* is used to denote the whole proposed search locations over which the best transformation is searched. Two analogous terms, *match location* and *registration point*, refer to the best transformation parameters yielding the minimum distance error between two surfaces. *Mismatch location* in this concept refers to any location not giving a match. The basic variables and expressions used in this chapter are listed below. A type of subscription may be used in each case which specifies that expression with a particular type of aspect (feature) as defined in the relevant text.

W	search window
P	sample points; that is the sampled points of the transforming surface
V	tracing voxels; the voxels traced in 3-D space for detecting an intersection of **P** with the objective surface
e	individual distance error
$\bar{e}$	Mean distance error (MDE)
σ	standard deviation (S.D) of the individual distance errors
E	cumulative distance error
r	MDE error at true match location
R	cumulative distance error at true match location
g	number of standard deviations (S.D.) from mean (r)
T	threshold used in the sequential methods
C	computational cost
n	number of entities (denotes number of sample points if it is used without any subscript).

6.2- Algorithm

The new method implemented for registering 3-D medical images, is based on the evaluation of the cumulative distance errors (instead of the mean distance errors). Since building up the cumulative error is a sequential process, a test can be made as each point is added to see if a threshold has been exceeded. An adaptive termination of the function evaluation and abandonment of its corresponding transformation in the processing location, is the second characteristic of the method. The method reduces the redundancy (e.g. time consumption) of the previous methods by performing a sequential process which may be terminated before all points are tested for a particular transformation. These enable methods to be implemented which are

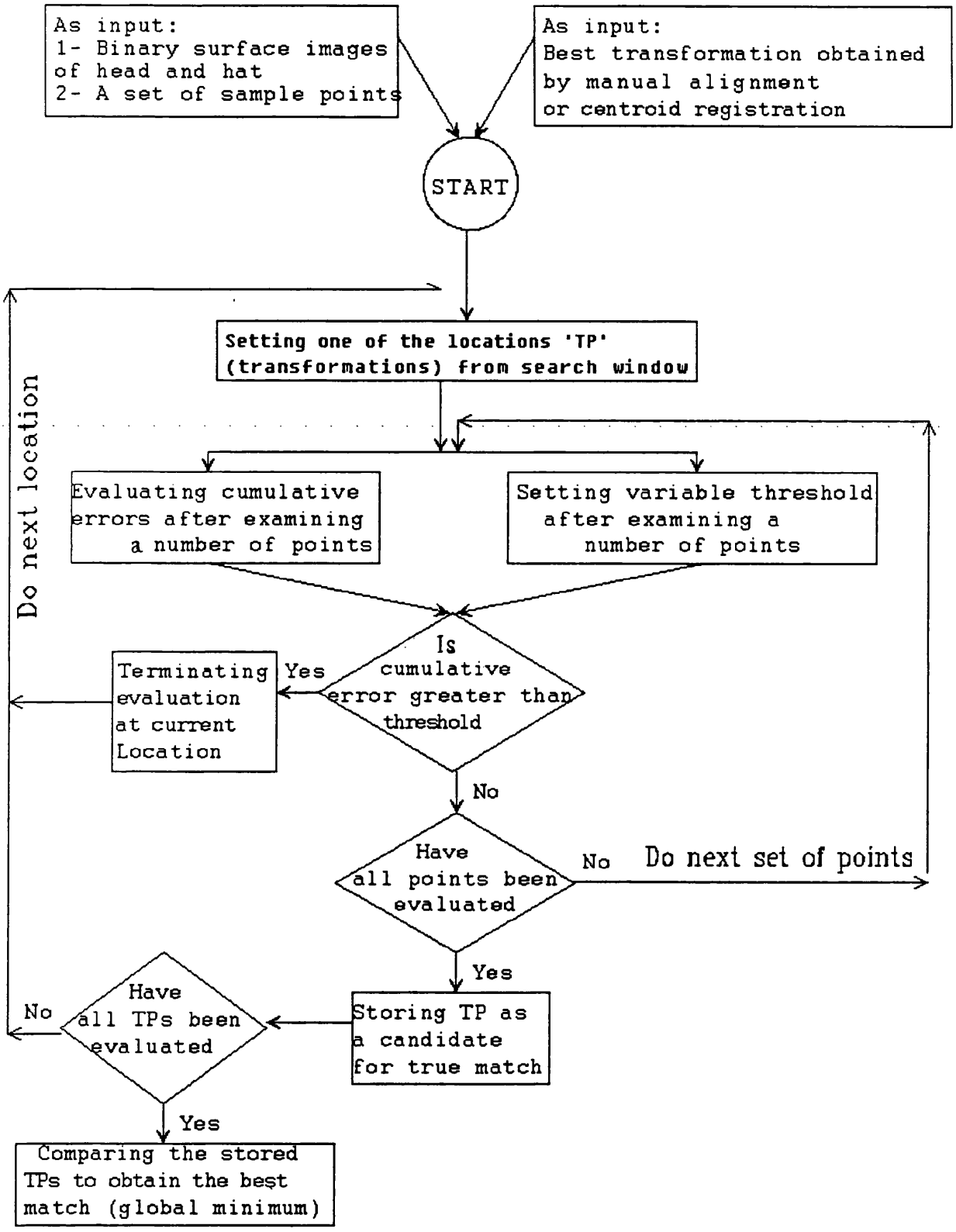


Figure 6.1- Diagram outlining the general algorithm of the sequential method used for registration. (TP denote transformation parameter).

independent of the evaluation of the mean square distance function for determining the next direction in which to search (for obtaining the best transformation parameters as discussed in section 3.2.7).

As stated, checking all the possible transformation parameters and terminating a sequential process when the cumulative distance error exceeds a threshold is the main feature of the new technique. The termination is actually controlled by a preset variable threshold depending on the number of points. This threshold specifies the sufficiency of the number of sample points participating in the achievement of the cumulative distance error. However, there will be a trade-off between the possible misregistration error and the number of points for which the process is allowed to continue before the termination.

The diagram shown in figure 6.1 summarizes the sequential algorithm used for registration of surface images. The input to the algorithm are the surface data obtained by a surface detection algorithm and a set of transformation parameters obtained by either the manual interaction process or any other method of registration. Different steps and requirements of the sequential algorithm are outlined in the next sub-section.

6.2.1- search space

A global search is performed to find the best transformation parameters giving the global minimum for the mean square distance function. Due to the preliminary alignment (e.g. manual registration or/and centroid registration) of the two data sets, all the possible transformations (locations) are not required to be examined. Thus, a search space (search window) is designed which consists of only those likely transformations which can be candidates for the global minimum.

The decision on the size of the search window depends on the accuracy of the transformation parameters estimated in initial stage when the registering surfaces are primarily aligned. The decision also depends on the sensitivity of the distance function to six various transformation parameters (3 shifts and 3 rotations). As shown in figures 3.9 and 3.10, a misalignment of 5 to 10 voxels shift in x, y or z direction yields a mismatch value (distance error) of 3 to 7 voxels which is high enough to be judged by the viewing inspection. Centroid

registration can also align the surfaces within this accuracy. The rotation, as shown in these figures, causes a smaller change in mismatch value, and thus it is more difficult to be assessed by viewing judgement. In this respect, a minimum rotation of 10 to 15 degrees (i.e. mismatch value of 2 to 7 voxels) shown to be adequate for a visually sensitive process. According to these results a maximum translational shift of 10 voxels and rotation of 15 degrees (in x,y and z direction) around the best transformation (obtained by the manual registration) seems to be generally adequate for setting the search window in the surface fitting process.

6.2.2- Cumulative distance error

As stated in section 3.2.5, the error measure is based on the L2 norm function

$$e_i = (P_{ti} - P_{oi})^2 \quad \text{for } i=1,2,..,n$$

where n is number of surface points which are going to be tested from the transforming image, P_t and P_o are vectors specifying the position of the corresponding surface points in 3-D space (for transforming and objective images, respectively).

The cumulative error E_m after m measurements (i.e. the number of points from the samples tested until the current working stage of the process) is

$$E_m = \sum e_i, \quad \text{where } i=1,...,m \text{ and } m \leq n \qquad \text{Equ. (6.1)}$$

As shown by Barnea & Silverman 1972, the density distribution of error measures of type L1 or L2 ($|P_t - P_o|$ or $(P_t - P_o)^2$) can be represented by an exponential function. However, this exponential distribution can be approximated by a normal distribution by invoking the central limit theorem (Papoulis 1965) for large value of n assuming that the individual error terms are statistically independent. Taking into account the above theorem the cumulative error has a normal distribution with mean $\bar{E}_m$ and standard deviation σ_m (Vanderbrug and Rosenfeld 1977). These can be expressed as

$$\bar{E}_m = m\bar{e} \qquad \text{Equ. (6.2)}$$

$$\sigma_m^2 = m\sigma^2, \quad \sigma_m = \sqrt{m}\sigma = \sqrt{m}\bar{e} \qquad \text{Equ. (6.3)}$$

where $\bar{e}$ and σ are the mean and standard deviation (S.D.) of the error distribution of the individual distance errors.

6.2.3- Expected errors

The misregistration error due to surface images misalignment contributes the major part of the distance error. However, there is an expected distance error even at the point of registration (where two surfaces are truly matched) which is due to the surface extraction error affected by image noise, artifact, distortion (in the case of *MRI*), and transformation noise. This is an unavoidable error generated during the registration process due to any factor other than the misregistration error. Let the expected cumulative error be R_m which is the sum of the error terms (square distance errors) for m points in the location of the true match.

$$R_m = mr \quad , \quad r = \bar{E}_n/n, \quad \text{for } m \leq n \qquad \text{Equ. (6.4)}$$

where r is the amplitude of the average individual error measurement and $\bar{E}_n$ is the mean cumulative error for total number of sample points (n), at the true match location.

The probability density distribution of the cumulative error E_m can be expressed by the following well known normal distribution equation as

$$P_E = \frac{1}{\sqrt{2\pi}(\sqrt{m}r)} \int_{-\infty}^{g} exp\left[\frac{-(E_m - R_m)^2}{2(\sqrt{m}r)^2}\right] dE_m = 1 - \frac{1}{\sqrt{2\pi}} \int_{g}^{\infty} exp\left[\frac{-(E_m)^2}{2}\right] dE_m$$

Equ. (6.5)

where g is the number of deviation (S.D.) from mean (expected error value) and a threshold (T_m) is expressed in terms of g. In this respect, g is a function of the probability P_E.

$$g = (T_m - R_m)/(\sqrt{m}r) \qquad \text{Equ. (6.6)}$$

6.2.4- Constant threshold algorithm

The efficiency of a sequential search depends on the proper choice of threshold which causes the termination of the process at some earlier stage when a mismatch location is

examined. In a constant threshold algorithm, the test is made against a preset constant threshold T as the error for the corresponding pairs is accumulated. The value of T is very important with respect to accuracy and the computational cost of the process. A high accuracy may be achieved only by an associated increase in computation.

6.2.5- Variable-threshold algorithm

Optimally, the threshold value should minimize the likelihood that a search location other than the true registration point (match location) remains below the threshold, while it maximizes this criterion for the true match location itself. It should also keep the number of operations reasonably small. As an initial assumption, it would be reasonable for the threshold to have a shape approximately similar to that of error growth curve (see fig. 6.2). Moreover, in order to prevent the elimination of a promising search location influenced by only few number of noisy and high-distance points, the initial value of T should be high enough.

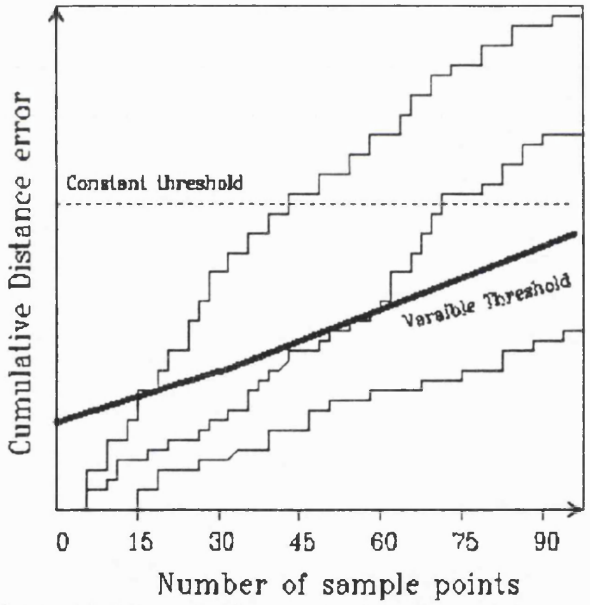


Figure 6.2- Schematic diagram showing error growth curves at true match location and two mismatch locations. The suggested threshold sequence is indicated.

There are three factors influencing the choice of threshold value required at each stage of the evaluation process. **1)** The knowledge of the expected (mean) error at the true registration point (match location) is essential for setting the threshold sequence. **2)** The shape of the error growth curve (cumulative error) is very important in setting the threshold. **3)** As shown in equation 6.5, the threshold is a function of the probability P_E of the match based on the density distribution of the cumulative error E_n. A n-dependent variable threshold (n= number of evaluated points) may be used as a threshold sequence to bound the error growth curve (at match location) from above with distances depending on g. The replacement of the constant threshold by this variable threshold can improve the computation effort while the performance does not decrease.

The threshold T for stopping the evaluation of the distance function can be derived from the value of the expected cumulative error R_m. Since this threshold should take account of possible variation in R_m in order to secure the detection of the true match location, a criterion based on a number of standard deviations (g as defined in equation 6.5) away from the expected error (R_m) is used. The desired threshold T_m as a function of m the number of points tested can then be expressed by the following equation:

$$T_m = R_m + g\sigma_m = R_m + g\sigma\sqrt{m} = R_m + gr\sqrt{m}, \qquad \text{Equ. (6.7)}$$

where r is the amplitude of the average distance errors at the true match location (considered as the expected error), σ is the S.D. of the error distribution of the individual errors, and R_m is the sum of the error terms (expected cumulative error) after evaluating m points of the registering surface data set. R_m is obtained empirically for any particular type of data set based on a priori knowledge of the distance errors at an approximate match location on a well-defined image pair (models).

The initial values in the threshold sequence should be high enough to allow reasonable confidence for retaining all uncertain locations (i.e. whether they are match or mismatch locations) under the threshold. The conditional factor ζ is used which is true when m is small.

$$T_m = R_m + (g+\zeta)r\sqrt{m}, \qquad \text{Equ. (6.8)}$$

where $\zeta>0$ for small m in the sequence (e.g. if $m<20$), and $\zeta=0$ when m is high. In most applications, increasing the threshold value (g) by one or two extra error standard deviations (i.e. $\zeta=1$ or $\zeta=2$) is suggested.

6.2.6- Sample points

Since the evaluation of distance errors is sequential, the order in which the points participate in the evaluation is important. They should be ordered such that, firstly, evaluation in such an order causes the termination decision earlier, and secondly, a small number of selected points mimic the entire response of all surface points (e.g. they should selected from all parts of the data set and not from regions grouped together). The suggested ordering rules for sample

points used in sequential process are shown (for a 2-D image) in the schematic diagram of figure 6.3. As shown in this figure, even a small number of points (i.e. any subset of points) belongs to different parts around the object (head contour) and thus holds an overall positional information. The position and number of sample points, as discussed in section 3.2.3, are also two important factors influencing the accuracy and speed of the registration process (see section 8.3.2 and 8.3.2.1 for experimental results).

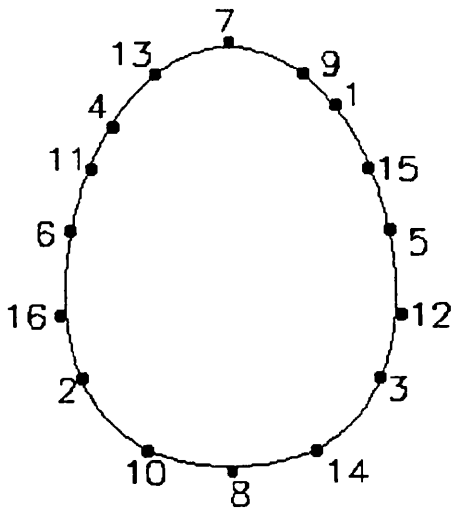


Figure 6.3- Diagram indicating the order of the first few sample points used in the sequential method.

The points are sampled in a three-stage process. Firstly, the local-high-curvature points (i.e. breakpoints) are selected as discussed in section 3.2.3. In the second stage, the required number of sample points are selected from these breakpoints sampled uniformly around the whole head. Finally, the points are ordered and grouped into sets of points (e.g. groups of 6 points) as suggested for the termination decision strategy in the sequential process. In this approach, the points in any group are also distributed equally around the whole head. Taking into account the same ordering strategy for sampling in a 3-D space (e.g. sampling in adjacent slices), the above process is performed on the global breakpoints from all the relevant slices.

6.3- Computation aspects of the sequential method

The expected computational cost for applying the sequential method depends on the number of search locations (n_W), the number of sample points (n_P), and the threshold sequence which restrict the number of points evaluated in the process. The cost also depends on the number of operations required for distance error measurement. As discussed for the surface fitting algorithm, due to the nature of intersection process, evaluation of some number of voxels V are required along the intersection line in a 3-D space (see figure 6.5). Accordingly, the number of operations involved for evaluation of each sample point can be obtained from the number of voxels (n_V) along the intersection line and the number of operations associated with each voxel (n_{Opr}). The diagram shown in figure 6.4 summarises these operations based on the number of the major and essential computational requirements. The cost can be expressed as

$$Cost=n_W * n_P * n_V * n_{Opr} * C_{Opr} \qquad \text{Equ. (6.9)}$$

where C_{opr} is the cost of each operation (i.e. usually expressed as the equivalent integer add operation: e.i.a). Note that the cost of point sampling, threshold sequence calculation, and search window selection are not included in all the calculations and cost predictions expressed in this chapter. This is due to the fact that these operations are required to be done only once in the sequential process, and thus their costs are small in respect to the rest of process. The cost of calculating the transformation matrix (which is required in respect of each rotation) is also negligible and thus is not included. The five parameters expressed in equation 6.9 are discussed in the next few paragraphs.

1) Due to the centroid registration and the initial manual registration, there is no need to apply a shift equivalent to the size of image, or 360 degree rotations in each direction. The number of locations n_W should be of an order sufficient to ensure the detection of the maximum possible distance between the two surfaces without being corrected by the centroid and manual registration (see section 6.2.1). The search space comprises regions (window) W_S referring to shifts, and W_R referring to rotations around each direction (x, y and z) which is applied for each shift step. Based on the experiments performed throughout this project (as discussed in section 6.2.1), it was found that a maximum shift of 16 voxels and maximum rotation of 15 degrees (in each direction x, y and z) are sufficient in most images initially registered by the centroid and manual registrations. In order to save computational cost, the rotation can be applied with step equal to 2. Considering shifts of from -8 voxels to +8 voxels, and rotations of -10 degrees to +10 degrees in the x, y and z directions, the number of search locations can be obtained by

$$n_{W_S}= (8\times2+1)^3, \quad n_{W_R}=(10\times2/2+1)^3,$$
$$n_W = n_{W_S} \times n_{W_R} = 6.54\times10^6. \qquad \text{Equ. (6.10)}$$

1- Do degradation (geometrical scaling) of the surfaces.
2- Select the initial level in which the global minimization is applied.
3- Select the sample points in each resolution level.
4- Calculate the (variable) threshold sequence (as a look-up table).
5- select the search window required in the initial level.

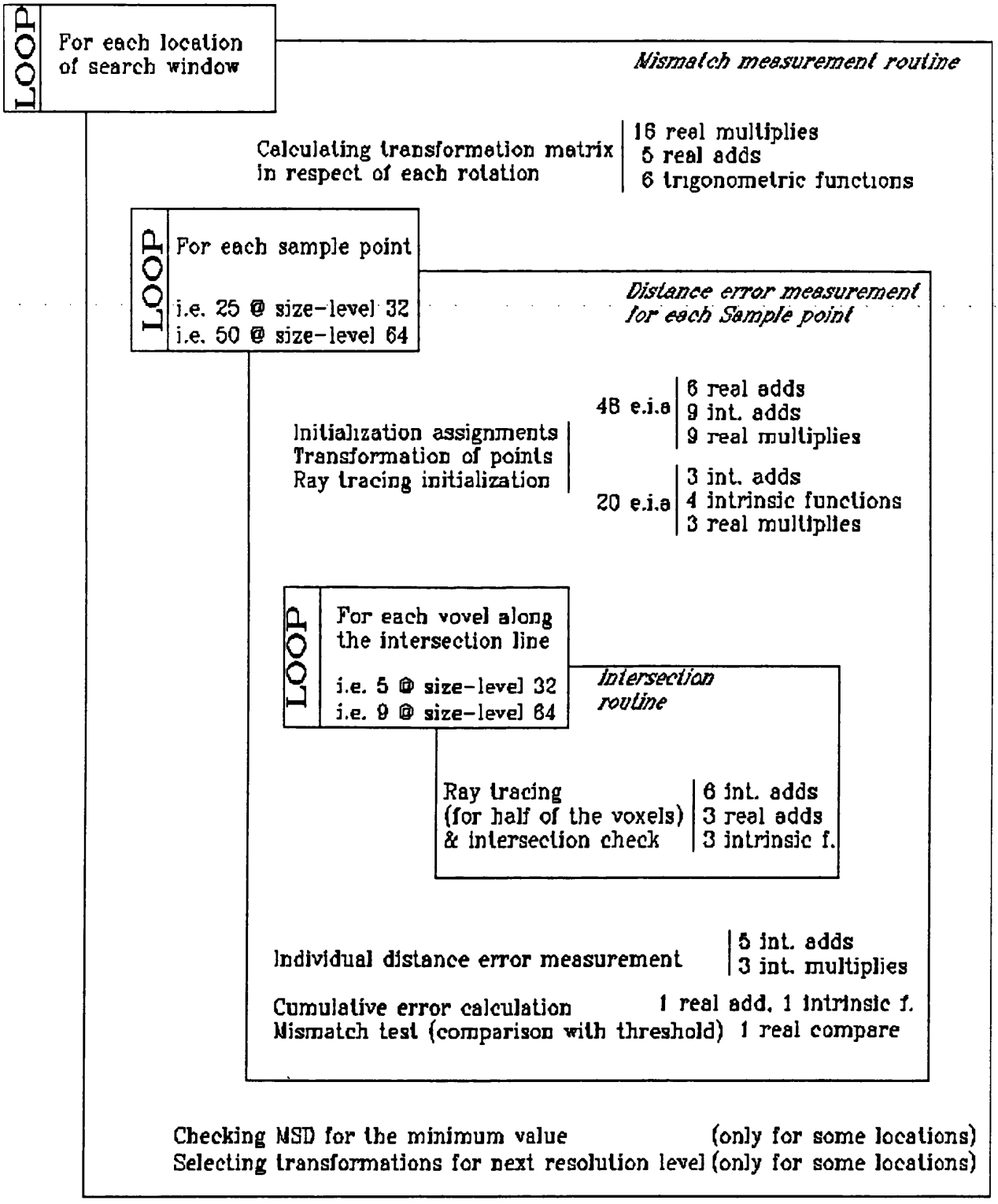


Figure 6.4- Algorithmic-based diagram representing different steps of the sequential registration process and number or required operation. Intrinsic functions are functions used in Fortran, such as Sin().

2) The number of sample points n_p is significant in determining the cost of the process. This is the major factor whose contribution is controlled by the threshold sequence which thus regulates the speed and accuracy of the process. This threshold causes the number of evaluated points in 3-D space to be ideally much smaller than the expected number, due to the automatic termination of the point evaluation at certain undesirable transformations.

3) As stated earlier, due to the application of the initial centroid and manual registration, the two registering surfaces are initially not expected to be far away from each other. Therefore, only a small number of voxels along an tracing ray originated from the transforming surface needs to be examined for the intersection with the objective surface (see figure 6.5). Allowing the maximum of 16 voxels shift in a high resolution image is based on the assumption (derived from the experiments) that the maximum initial distance expected between the two surfaces is not more than about 16 voxels in each direction. Accordingly, no more than 16 voxels (**V**) need to be examined around each sampling point, in order to detect the actual intersection with the objective surface (head). However, in the actual intersection process, the number of examined voxels is usually less than the expected value because of an earlier cessation of the ray tracing when the intersection occurs. Since, the tracing ray is originated from a point in 3-D space corresponding to the position of the sample point and then extended in two sides by 16 voxels (see figure 6.5), the number of examined voxels is

$$n_V = (2 \times 16) + 1 = 33. \qquad \text{Equ. (6.11)}$$

4) As it can be seen in figure 6.4, the number of operations results from contributions from: **a)** the calculation of the transformation matrix at each search location, **b)** transforming each surface sample point to the new coordinate system, **c)** defining a 3-D ray tracing process (intersection line) for each sample points, and finally, **d)** comparing these points with the objective surface, thereby updating the distance error function if the ray intersects

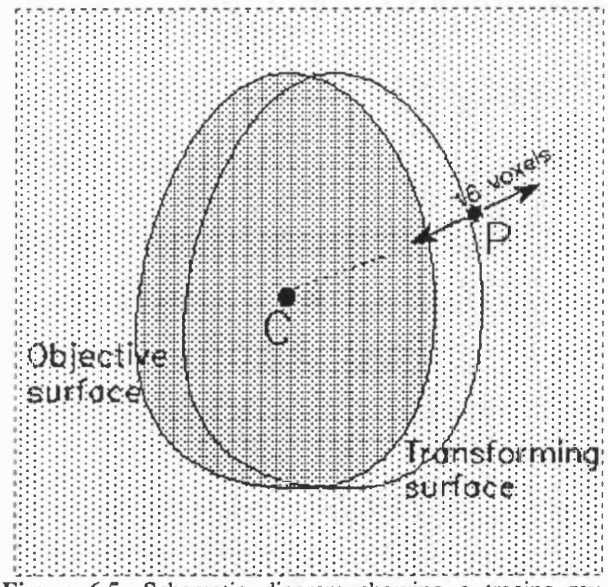


Figure 6.5- Schematic diagram showing a tracing ray originated from sample points to obtain the intersection with objective surface.

with the objective surface. In order to enable the comparison between the number of operations required by using different stages of the process, all the operations are expressed by a relative number based on the *equivalent integer add operation.* The actual timing response of different operations obtained by a μVAX II, and the time ratio relating these operations with integer add operations are illustrated in table 6.1. The number of operations required for evaluating each part of the program (i.e. in respect of each search location W, sample point **P**, and ray tracing voxel **V**) during the registration process is as follows.

1)	1 integer compares / W	≡ 1.5 e.i.a./ W
2)	17 integer adds / Sample point '**P**'	≡ 17 e.i.a./ **P**
3)	7 real adds / **P**	≡ 14 e.i.a./ **P**
4)	1 real compare / **P**	≡ 4 e.i.a./ **P**
5)	5 intrinsic functions (e.g. Sin, Int)/ **P**	≡ 10 e.i.a./ **P**
6)	3 integer multiplies / **P**	≡ 6 e.i.a./ **P**
7)	12 real multiplies / **P**	≡ 36 e.i.a./ **P**
8)	3 integer adds/ ray tracing voxel '**V**'	≡ 3 e.i.a./ **V**
9)	1 real adds / **V**	≡ 2 e.i.a./ **V**
10)	1 integer compare / **V**	≡ 1 e.i.a./ **V**
11)	2 intrinsic functions / **V**	≡ 4 e.i.a./ **V**

where e.i.a. denotes *equivalent integer add operation.* The number of assignments are not considered in these operations. Note that these number of operations do not take into account any improvement (i.e. termination of the evaluation) due to the use of the variable threshold. From this information, the total number of operations can be expressed as

$$n_{Tot\_opr} = (n_W \times 1.5) + n_W \times n_P \times 87 + n_W \times n_P \times n_V \times 10 \quad \text{e.i.a.}$$

$$n_{Tot\_opr} = n_W\,[n_P\,(87+10n_V)] \quad \text{e.i.a.} \qquad \text{Equ. (6.12)}$$

The first part of above expression is eliminated due to its smallness in relation to the other parts. From expressions 6.10, 6.11 and 6.12, the total number of operations at the highest resolution level (L0) is

$$n_{Tot\_opr} = 6.54 \times 10^6 \times (417 n_P) \quad \text{e.i.a.}$$

Due to the improvement obtained by using threshold sequence, it is not possible to specify exactly n_P for each location in the search window and in turn to define n_{Tot_opr}. However,

n_p can be predicted based on some behaviourial knowledge of the distance function at different search locations. These are outlined in the next section.

5) The duration of time taken by an integer add operation (i.e. C_{opr}) is shown in table 6.1. The cost of applying sequential method is obtained by the product of n_{Tot_opr} by the time required for one integer add operation.

Operation	time $s/10^6$	e.i.a.
Integer add	**2.26**	**1**
real add	**4.9**	**2**
integer multiply	**4.9**	**2**
real multiply	**5.9**	**3**
integer compare	**3.5**	**1.5**
real compare	**8.3**	**4**
Trigonometric function	**4.95**	**2**
Intrinsic function e.g. Sin()	**4.95**	**2**

'^R '^]

Table 6.1- Timing response of the system corresponding to different operations (for μVax II, based on seconds per million operation). e.i.a. denotes equivalent integer add operation.

6.3.1- Decomposition of the search window and prediction of n_p

n_p varies with threshold in a way that depends upon the degree of misregistration between the two surfaces. The threshold is ideally set to permit an early termination of the evaluation at a certain misregistration window point, that is, the termination at a search location which give cumulative error greater than the threshold T_m (where $T_m=rm+gr\sqrt{m}=rm(1+g/\sqrt{m})$; see equ. 6.7), when only a fraction of the total number of sample points are examined.

Both, the threshold and mismatch (distance error) values are expressed in terms of the number of voxels. Assuming that the misregistration error is the only source of the distance error, the threshold, which is expressed in voxels not in absolute value, is the maximum acceptable distance error between the two surfaces. Assuming that any unit of the transformation (in one or all directions) moves the two surfaces away (i.e. produces a distance error) with the same magnitude, a neighbouring search window ($\dot{W}_{nei}$) can be defined around the true match location which have its corresponding distance errors less than the threshold. Consequently, in respect of each unit of the threshold, the window $\dot{W}_{nei}$ is expected to be enlarged by about one voxel in each direction corresponding to one voxel shift or/and one to two degrees rotation (because the distance errors are less sensitive to rotation) around the true global match location.

In the example shown in figure 6.6 (central shaded 3×3 square), this window consists of 729 locations corresponding to 1 voxel shift and 1 degree rotation in x, y and z directions (obtained by $(\text{mean}\times 2+1)^{(6)}$, where the threshold is equal to 'mean' and mean=1 voxel).

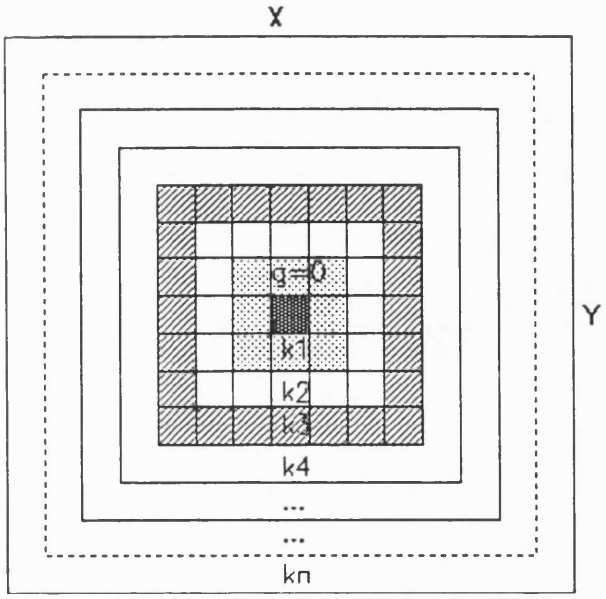


Figure 6.6- The neighbouring locations around the true match (central square) for two pixels being matched together. Each region K corresponds to the same number of transformations around the true match location.

Only the locations of the search window W_{nei} have distance errors comparable with the threshold, and thus require more evaluation to be tested for the true match location. Therefore, all the sample points are expected to be examined at every location of this search window.

From the equation 6.7, the number of neighbouring locations $n_{W_{nei}}$ in which approximately all the sample points are likely to be evaluated can be obtained by

$$n_{W_{nei}} = (r(1+g/\sqrt{m})\times 2 +1)^6 \qquad \text{Equ. (6.13)}$$

Other locations (denoted as search window W_{off_nei}) which produce sufficiently high distance errors would elevate the threshold at some earlier stages of the evaluation and thus only a few number of points are examined. However, the initial number of points should be high enough at any location, in order to establish conclusively that a certain misregistration has occurred. The number of locations in which the points are Partially evaluated can be obtained by

$$n_{W_{off\_nei}} = n_W - n_{W_{nei}} \qquad \text{Equ. (6.14)}$$

As shown in figure 6.6, the window W_{off_nei} can be divided into different sub-windows (W_k) based on the expected distance error which they yield. The number of sample points allowed to be evaluated in the locations of a sub-window W_k (where k is the transformation

parameters, e.g. shift, to build that particular sub-window) depends on the threshold T (i.e. mean and value of g) and on the transformation k. Since the distance error contributed by each sample point increases as the applied transformation increases, a smaller number of points are allowed to be evaluated in these sub-windows than in the window W_{nei}. This means a smaller number of sample points are needed to generate a distance error bigger than the proposed threshold value. The number of points ($n_{P@W_k}$) for a sub-window W_k can be obtained by

$$n_{P@W_k} = n_P \times T/k \qquad \text{Equ. (6.15)}$$

where $_{P@W_k}$ refers to the sample points examined at window location W_k. Assuming that any transformation (e.g. shift) makes the mean square distance between the two surfaces with the same magnitude, the number of locations n_{W_k} of a sub-window W_k can be obtained by

$$n_{W_k} = (k\times2+1)^6 - ((k-1)\times2+1)^6, \qquad \text{Equ. (6.16)}$$

where $k \geq r(1+g/\sqrt{m})$ and r is the mean of the distance between the two surfaces. From the definition of the neighbouring voxels, it can be shown that

$$n_{W_k} = n_{W_{nei}} \text{ if } k=r(1+g/\sqrt{m}) \quad \text{i.e. } k=T$$

The total number of window locations can then be expressed as

$$n_W = \sum_{k=T}^{maxTR} [n_{W_k}] \qquad \text{Equ. (6.17)}$$

where maxTR is the maximum transformation (e.g. shift) applied in the search window W (e.g. maxTR=10, if the maximum transformation is about 10 voxels). From equations 6.12 and 6.17, the total number of operations (n_{Tot_opr}) can be obtained by

$$n_{Tot\_opr} = \sum_{k=T}^{maxTR} (n_{W_k}) (417\times P_n \times n_P) \quad \text{e.i.a.} \qquad \text{Equ. (6.18)}$$

where coefficient P_n denotes the probability of the number of sample points involved in registration when a mismatch location is examined. Since the behaviour of the distance function is not clearly known at a mismatch location, the coefficient P_n should be found empirically. However, an estimated value for P_n can be obtained by considering the parameters effecting the

probability of the matched points (e.g. mean and S.D.).

$$P_n = T/k = \text{mean}(1+g/\sqrt{m})/k. \qquad \text{Equ. (6.19)}$$

In this sense, n_P can be replaced by $n_{P@W_k}$ from the equations 6.15 and 6.19. The total number of operations are then expressed as

$$n_{Tot\_opr} = \sum_{k=T}^{maxTR} (n_{W_k}) (417 n_{P@W_k}) \quad \text{e.i.a.} \qquad \text{Equ. (6.17)}$$

It is apparent that there is some statistical noise and misregistration fluctuation that affect the number of evaluated sample points yet unexplored for determining the exact number of operation and thus the computing time.

6.4- Summary

This chapter introduced a new approach for matching the 3-D surface data obtained from medical images. This new approach was known as the sequential method in which a cumulative distance error was used instead of the mean distance error for the evaluation of the mismatch. The different sources of error contributing the cumulative distance errors were also outlined. Using this type of the error measure was shown to enable the application of a variable threshold which causes the termination of a mismatch location very efficiently. A formulation for this threshold sequence was also presented. It was shown that the application of the sequential process on 3-D surface data reduces the computational cost while it does not effectively change the performance of the process.

The computational aspects of the process and cost prediction were discussed in section 6.3. Various factors influencing the cost were also studied. The expected time required for the algorithm to decide that there is not a match at a given location was shown to depend on the threshold, sampling strategy and ordering of the sample points.

CHAPTER 7
MULTI-RESOLUTION REGISTRATION

7.1- Introduction

To improve the computational cost, the sequential method of registration was modified and implemented in a multi-resolution space (i.e. at different matrix sizes) (see IPMI'91 for initial results; Oghabian and Todd-Pokropek 1991). The proposed technique incorporates a hierarchical search for obtaining the best transformation (match location) which match the two surface images. Although the technique is applied on binary images of different grid (matrix) sizes, it is called *multi-resolution* throughout this project due to the fact that scaling an image to a bigger pixel size yields efficiently a coarser resolution image. In this process, there are some common features with the multi-resolution template matching introduced by Barnea and Silverman 1972, Wong and Hall 1978, and Wong 1978.

The fitting is started in low resolution surface images, and stepped to a higher resolution when a desired fit passes an appropriate threshold. During the search at each resolution level, the sequential testing and adaptive termination (see section 6.2.5) of a search location is also performed, in order to minimize the amount of computation, further. The whole process is summarized in the diagram shown in figure 7.1.

The aim of this chapter is to describe the proposed *multi-resolution* approach and discuss the computational aspect of the process. The general algorithm outlining different parts of the program is described in section 7.2. All the multi-resolution aspects of the surface creation and surface fitting process are presented in the subsections of 7.2. A fast method for creating low resolution surface images is suggested in section 7.2.1. The number of search locations and the basic strategy behind the selection of these locations are presented in section 7.2.2. The error analysis is given in section 7.2.3. The advantage of using a variable threshold is also addressed. Section 7.3 presents the expected computational cost of the multi-resolution algorithm and the cost saving gained by this process. A summary of the method is presented at the end.

7.1.1- Terminology

Most of the variables and terms expressed in this chapter have been defined in chapter 6. Here a superscript L is attached to denote the resolution level (as defined in section 7.2). Level L, in this concept, refers to the degradation (scaling) level applied to create a smaller low-resolution surface image. The size-level is referred to the size of image (in any direction) corresponding to the resolution level. The values assigned to L and the definition of some other variables are as follow.

L can be 0, 1, 2 or 3, denoted as L0, L1, L2 and L3, for image size of 256^3, 128^3, 64^3 and 32^3, respectively. The size 256^3 in this respect is the size of the highest resolution image.

$\check{L}$ denotes the maximum level to which the image is degraded (i.e. the level in which the global grid search is applied)

P the probability of a match (true or false) at below detection threshold.

7.2- Algorithm

The main features of the multi-resolution process are shown in the diagram of figure 7.1. The inputs to the multi-resolution image registration system are two sets of binary surfaces obtained by a surface detection technique, and a set of transformation parameters. A grid search (see section 5.3) is applied on surface images at different resolution levels, by which the computational cost is substantially reduced. Although the purpose of this technique is to apply a grid search including all the possible transformations (each having 3 shifts and 3 rotations in x, y, and z) for minimization process, it is unnecessary to check those transformations which are substantially too far from the expected match location. Therefore, further reduction of computation is also achieved by restricting the search to a smaller window around the best match location obtained by centroid and/or interactive manual registration.

A sequence of lower resolution surface data sets of a smaller sizes (e.g. 128*128, 64*64, 32*32) are generated from two original surface data of size 256*256. The level to which the image size is reduced depends on the value of mismatch and "*effective distance*" error between the two registering surfaces. The effective distance means the average distance of only those

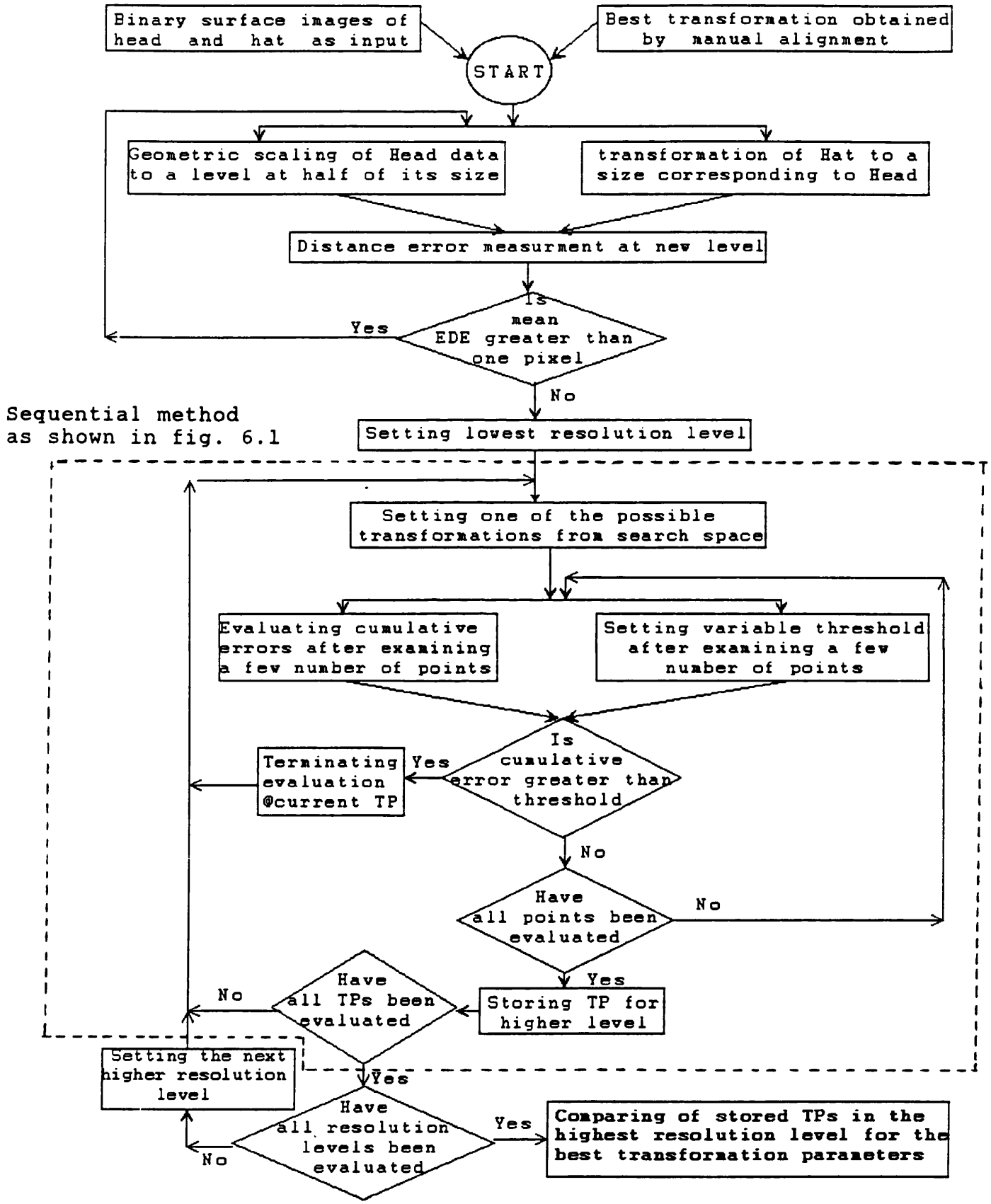


Figure 7.1- The main features of the multi-resolution process. EDE denotes '*effective distance error* (see section 7.2) and TP=transformation parameter.

points on the surface which are not in match or close to match position with respect to their corresponding points. No attempt was made to reduce the image dimension to a size which gives an effective distance error of less than one voxel. This is due to the fact that the surfaces are defined in terms of whole voxels. Distance measures are therefore expressed as an integer number of voxels, which vary as a function of level L. Thus, the effective distance error between two surfaces reduces at the same rate as the image size is reduced.

7.2.1- Surface degradation methods

Two procedures were used to create low resolution surface images. The first process involves two-dimensional low-pass filtering of the original grey-scale image slices at each level and then sampling the filtered data at half of the image sample rate of the higher resolution level. The surface detection algorithm is then applied on each scaled data set at different levels. The major disadvantage of this procedure is that all the time consuming surface detection steps should be employed on the complex grey- scale images at all resolution levels.

An alternative criterion for creating a surface in a lower resolution image (having a bigger voxel size) can be defined which is based on the binary surface data set obtained at the highest resolution level. This method was mainly used in the current project. In this method, placing a surface point at a position (voxel) in a low resolution image depends on the location and number of neighbourhood surface points (for example eight adjacent voxels) in its corresponding higher resolution image. In the simplest sense, a surface point in the small matrix size of a low resolution surface image is set if at least two of its neighbourhood in the corresponding higher resolution image are on the surface. This criterion is explained in an algorithmic manner in the following paragraph.

Let a voxel f_{L+1} in a low resolution level (i.e. *degradation level* L+1) be set as a surface point if any two of the eight adjacent voxels in the higher resolution image (i.e. degradation level L) are on the surface.

$$f_{L+1}(i,j,k)=1, \text{ if } f_L(2i+\alpha,2j+\beta,2k+\gamma)=1$$

for at least two neighbouring voxels (where $\alpha,\beta,\gamma=0,1$),

otherwise; $f_{L+1}(i,j,k)=0$, Equ. 7.1

where α, β and γ are scalar values 0 or 1 in order to define the eight neighbouring voxels for each surface point (2i,2j,2k) of the lower degradation level L (for i,j,k between 1 and n; where n is the size of the 3-D binary surface data set at higher degradation level L+1). The schematic diagram of figure 7.2 shows this process in a two dimensional binary image.

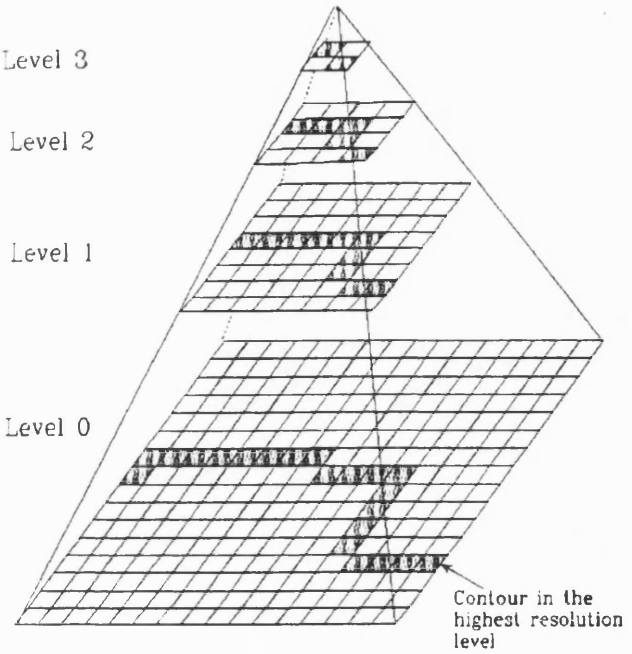


Figure 7.2. Schematic diagram showing the degradation of a 2-D binary image from level L0 (size 256^3) to level L3 (size 32^3).

Both the transforming and objective surfaces are scaled (degraded) to the same level by the preceding algorithm. The sampling points of the transforming surface are then reselected in each new level. Alternatively, the original sample points (at the size-level 256) can be scaled to a level corresponding to the current matrix size of the objective surface. However, in the later method, more errors are incorporated in obtaining the actual locations of the sample points due to possible loss of consistency with the other points (e.g. neighbouring points) of the transforming surface.

7.2.2- Search Locations

All the transformations (locations) in a proposed search window are applied to the surface data sets of lower resolution level. In each location, the distance errors are measured and cumulated sequentially until they exceed an adaptive (variable) threshold. The goodness of any transformation is judged by comparing its corresponding distance error with the threshold. In order to overcome the effect of errors incorporated during the degradation process, a number of transformations which yield a mismatch value less than the predefined threshold, are considered in each level as candidates for acceptable matches in the higher resolution levels. In other words, the method uses some of the best transformation parameters (most promising search location) found at a given level L+1, to pass them as search locations to the next lower degradation level L for further investigation.

7.2.3- Error concept

As well as the registration errors introduced in section 6.2.3, there are other sources of error associated with the multi-resolution algorithm. These errors are mainly associated with the surfaces by losing some detail due to the nature of scaling (*degradation*) algorithm. This is known as scaling or degradation noise throughout this thesis. As discussed in section 7.2.1, any voxel in a lower level is formed from, and thus correspond to, eight voxels in the previous higher resolution level. Accordingly, any angled or oblique edge occurred in these eight voxels may be smoothed out (see fig. 7.2) and is thus counted as a straight line segment after the degradation. Nevertheless, there might be some noise on the surface due to surface detection process, which will smooth out as well. A low resolution image may also be less affected by distortion (in terms of voxels) than an image of high resolution would.

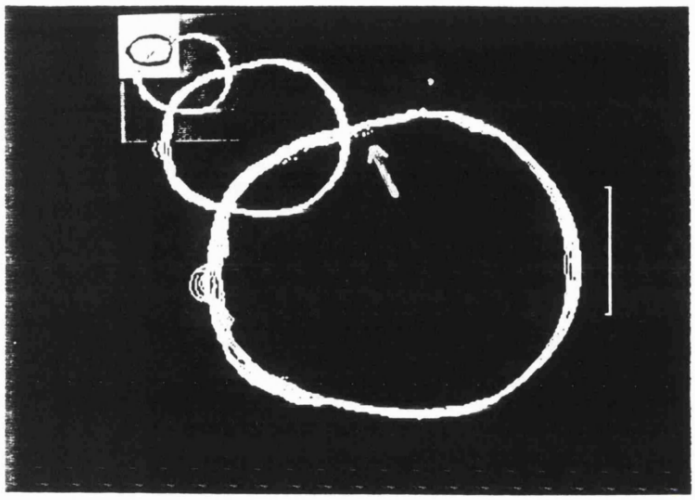


Figure 7.3- 2-D cross section of head surface in the multi-resolution study. These correspond to levels L0 (size 256^3), L1, L2, and L3.

7.2.3.1- Scaling noise

In general, errors introduced in the creation of low resolution images depend on the degradation (scaling) method and thus the amount of detail and geometrical information lost in the process. The contour images extracted from the skin surface at four subsequent resolution levels are shown in figure 7.3 As illustrated in this figure more detailed information is contained in the contour of 256*256 image than 32*32. On the other hand, the lowest level image has a simpler shape which corresponds to the oval shape of the head contour (in axial cross section). This effect can sometimes be regarded as an useful phenomenon causing an easier and quicker convergence of the registration process. However, some noise on the contour of the high resolution image (indicated by a white arrow on fig. 7.3) are eliminated in the contour of the lower resolution image.

7.2.4- Variable-threshold selection

The *variable threshold* T introduced in chapter 6 (equ. 6.8) can be modified to account for the termination of the function evaluation at different resolution levels L.

$$T_m^L = R_m^L + (g+\zeta)\sigma = R_m^L + (g+\zeta)r^L\sqrt{m} \qquad \text{Equ. 7.2}$$

This threshold sequence guides the search from a low resolution to a next higher resolution level. The threshold should be low enough to select only the most promising transformations at each level for testing in the next level. However, it is desirable that this threshold does not eliminate those locations that are likely to be important as a potential true match location at a higher resolution level. The crucial and difficult problem in the selection of an optimal threshold is the choice of the expected (residual) mean error R_m^L which exists between the two registering surfaces at each resolution level (see section 6.2.3). In many cases it is not easy to do this in advance, since different images may be processed with various noise levels. However, the threshold can be set based on the probability analysis (e.g. density error distributions) for two well-defined surface data sets (see chapter 8 for the results).

The expected error in a subsequent low resolution level is expected to be half that of its parent level. However, another factor should be considered in this concept. Since the surface is represented in terms of whole voxels, the individual distance errors are measured to voxel accuracy. Hence, scaling the image to a lower resolution level, forces some of the distance errors to a rounded number of voxels. This forces some of the errors to zero at lower resolution levels when successive scaling have been applied. Accordingly, the coarser the surfaces are made, the lower the standard deviation of the distance error is expected. Therefore, in order to predict the threshold at any level from the expected errors obtained in the previous level (e.g. higher resolution level), a coefficient (τ) should be taken into account for such an unwanted error suppression mechanism. In this concept, τ counts for those errors which are inherently bigger than the threshold, but have been scaled (rounded) to a value of less than threshold. The threshold sequence in a level L (obtained from the highest level L0) becomes as

$$T_m^L = (1-\tau)(R_m^{L0} + (g+\zeta)r^{L0}\sqrt{m})/2^{(L-L0)} \qquad \text{Equ. 7.3}$$

where L0 denotes the highest resolution level. τ is zero for the highest resolution level and varies between 0 to 1 depending on the fraction of points matched due to the degradation process. Since the fractional error resulting from the use of the whole voxel during the degradation process is unknown, their actual contribution to the expected error of a degraded level is not clear. This effect in addition to the other sources of error mentioned in section 7.2.3, causes an uncertainty in the prediction of τ. The prediction of the mean (expected) error in the lower resolution levels are therefore obtained, empirically in each level, by using a well-defined scaled (degraded) data set.

7.3- Computational cost

As mentioned in the previous chapter, the cost of a sequential method is primarily a function of the number of search locations, the number of sample points, and the number of arithmetic operations involved in evaluating of each point. The basic approach in assessing the computational cost associated with the sequential concept of the multi-resolution technique can be found in section 6.3. All the formulations and discussions outlined in section 6.3 and 6.3.1 are applicable to this approach of sequential multi-resolution, except that of incorporating an extra factor (coefficient) L defining the resolution level. The superscript L, in this sense, is used to denote the degradation or scaling level between the processing surface and the original high resolution surface (i.e. L=0 for image size of 256^3, L=1 for 128^3, L=2 for 64^3, and L=3 for 32^3). In this section, the computational aspects of the method at different resolution levels are discussed. Unless otherwise stated, the terminology and variables used in this section are based on those defined in section 6.1.1 and 6.3.

Since at a lower level of resolution both the size of data sets and the number of possible search locations decrease, the improvement (evaluation) of distance function terminates by giving an appropriate outcome for each transformation in a very short period of time. This outcome is either a rejection, or an acceptance of that particular transformation as a potential candidate location for the global minimum. Consequently, the cost efficiency of the registration process substantially increases as the size of the two registering images decreases. The number of sample points required is also reduced at coarser resolution which implies a higher convergence speed for the algorithm. These effects are the origin of some cost differences with the sequential

process as discussed in the following paragraphs.

1) Although the rotation is consistent at all the resolution levels, a bigger step-length is required to produce a noticeable change in position of the surfaces. Shifting at each level (in terms of voxels) is equivalent to its twice that shift in a next higher resolution level. Accordingly, if -8 to +8 voxels shift is required in size-level 256, this is equivalent to -1 to +1 voxels shift in size-level 32. The search window $W^{\check{L}}_{S}$ is the initial shift window applied in the lowest resolution level (e.g. -1 to +1 voxels shifts in each direction x, y and z at the size-level 32) and $W^{\check{L}}_{R}$ is the corresponding rotation window (i.e. -10 to +10 degree rotation applied with step-length 2 in each direction), where the superscript $\check{L}$ denotes the initial level in which the whole transformations is applied (e.g. size-level 32). Note that the step-length is an estimate of the smallest step in transformation space that changes the position of the transforming surface (obtained by step≈ArcTan(1/d); for a shift of one voxel, where d in the distance between the surface and centre of rotation). The total number of transformations initially used in level L3 is

$$n_W^{L3} = n_{WS}^{L3} \times n_{WR}^{L3} = 3^{(3)}+11^{(3)} = 3.6\times10_4. \qquad \text{Equ. 7.4}$$

Although this window size is much smaller than that of the proposed search window at the original resolution level, it covers a wider range of transformations. W in the initial level (i.e. $W^{\check{L}}$) is usually fixed, but it varies at other levels depending on threshold values. Optimally, the size of window $W^{\check{L}-1}$ in the next higher resolution level is much smaller than $W^{\check{L}}$.

2) Based on the window size prediction outlined in section 6.3.1 (Equ. 6.13, 6.16), the size of sub-window W_{nei} which requires all the sample points to be processed is reduced as the resolution (matrix size) is lowered. The fraction of points processed in sub-window W_{off_nei} (i.e. all sub-windows W_k of the shift or rotation space) is also reduced. These are due to the decrease in the value of 'mean' of the distance between the two surfaces in voxels, which gives a lower threshold value (Note that the size of voxels as units of mismatch is relatively bigger at lower resolution levels).

3) Since the size of the registering images is reduced as the resolution level reduces, the optimal number of sampling points n_P^L required for registration is also expected to reduce. The expected number of required sample points at different resolution levels depends on the ratio of size reduction between them. Based on the values selected for L (i.e. L is 0, 1, 2 or 3 for image

size of 266^3, 128^3, 64^3 and 32^3, respectively), the size ratio between the images at different resolution levels and the highest resolution level (L=0), is a function of $2^{(L)}$. Therefore, the number of sample points n_p^L at level L can be obtained by

$$n_p^{L0}/2^{(L)}. \qquad \text{Equ. 7.6}$$

4) The change in the number of voxels n_v^L required to be tested for an intersection in the ray tracing process at different levels, are also equivalent to the ratio of the size-level reduction. For example 2 voxels at size-level 32 corresponds to 16 voxels at size-level 256 (i.e. size reduction is 256/32=8). The general equation which specify the number of required voxels is

$$n_v^L = n_v^{L0}/2^{(L)}. \qquad \text{Equ. 7.6}$$

The number of operations in level L can be obtained from equation 6.18, but with some modification to take into account the influence of the resolution level. The cost of the degradation process and the selection of initial level are ignored in all the assessments, since this needs to be done only once. Thus

$$n_{Tot\_opr}^{L} = \sum_{k=T}^{MaxTR} [n_{Wk}^{L}] \; [P_n \times n_p^{L0}/2^{(L)} \; (87+10n_v^{L0}/2^{(L)})] \quad \text{e.i.a.}$$

Equ. 7.7

The expected cost of the whole multi-resolution process can then be obtained by

$$Cost = (n_{Tot\_opr}^{L} \times C_{Opr}) + \sum_{L=L-1,0}[(P \times n_W^{L+1}) \times P_n \times n_p^{L}(87+10n_v^{L})].$$

Equ. 7.8

where P is probability of a below-threshold mismatch at the lower resolution level L+1. The first part of the equation indicates the cost of applying the coarsest resolution level, and the second part shows the cost corresponding to the rest of levels. There is a trade-off between the probability P as a function of threshold, and the accuracy of the registration. Finding a good accuracy for which the cost is minimum is an optimization problem which can be solved empirically. The saving of computation effort resulting from the process is shown in table 7.1.

7.4- summary

This chapter has presented a novel approach applied to the registration of 3-D surface

Ľ	Size	n_W	n_P (×87)	n_V (×10)	n_{Tot_opr}eia	Time
0	256	3.4×10^6	100	33	1.4×10^{11}	4726 min
1	128	505×10^3	50	17	6.48×10^9	216 min
2	64	86.6×10^3	30	9	4.59×10^8	15 min
3	32	18.7×10^3	20	5	5.12×10^7	102 sec

Table 7.1- Number of entities (e.g. operations) in the multi-resolution technique when it starts at different resolution level Ľ (Ľ denotes the initial resolution level). The reduction in number of operations and time is obvious. n_W, n_P and n_V refer to the number of search locations, sample points and tracing voxels, respectively. n_W is based on a suggested window size (i.e. ±8 voxels shift and ±10, ±6, ±8 degrees rotation with step 2 around x, y, and z direction). n_P and n_V are based on the program and algorithmic diagram shown in figure 5.4. The value in parenthesis show the number of operations for each entity. n_{Tot_opr} denotes the total number of operations based on the equivalent integer add (e.i.a)($n_{Tot\_opr}=n_W\times(n_P\times(87+10n_V)$; see equ. 6.12). Note that no improvement was included resulting from the use of a variable threshold.

data from medical images. The approach employs a grid search on two low resolution (small image size) surfaces, and applies only the most promising transformation parameters on a corresponding bigger (high resolution) image. Therefore, the performance computed on the basis of some few voxels matched in a coarse resolution image is used for the next resolution level, and in turn for the original high resolution image.

A degradation technique has been introduced which uses only the 3-D positional information of the surfaces. The errors incorporated in the surface degradation technique and their effects on the registration process was also studied in this chapter. Other sources of error were outlined with a tendency of a noise suppression effect of the low resolution images.

In section 7.2.4, an expression was determined for the definition of a threshold sequence used in multi-resolution sequential methods. Various parameters influencing this variable threshold was discussed.

The expected computational cost and the factors influencing this cost were explored in section 7.3. The cost prediction of the sequential multi-resolution process was compared with that of the pure sequential method. A time saving ratio of about 1/2800 was obtained when using a multi-resolution method on a grid-based search strategy (in contrast to a global search at the highest resolution level).

CHAPTER 8
EXPERIMENTS, RESULTS AND DISCUSSION

8.1- Introduction

As discussed in chapters 6 and 7, the registration process manages to move and rotate a number of sample points (i.e. a template) from one registering surface across the other to find the best match between them (minimization). The process of searching for the best transformation (location) can be thought as a grid search which locates the best match between two images. The ability to find the correct match is influenced by different types of parameters such as number of locations, number of sample points and the threshold level used in the sequential decision process. Noisy data, the structure of distance function and the deficiency in the registration process itself may distort, broad or mislead the minimization point.

The aims of the experiments and the results outlined in this chapter are to show the fidelity of the proposed *multi-resolution* and *sequential* methods used for registration of medical images. A set of experiments were designed to evaluate the sequential multi-resolution techniques used for registering different types of data set, and to compare practical results with those of the theoretical discussion. Two types of phantom, the Hoffman brain (10 cm thick) and the Jaszczak phantom were initially used in these experiments. The data from thirty routinely imaged patients, and ten patients with setting external markers were also used in the current assessment. Moreover, five sets of brain data from volunteer subjects were used in order to generate some high definition, known data sets. By transforming all the above data sets to a number of known locations (i.e. arbitrary transformation parameters), a number of well defined data sets were generated for verification of the proposed algorithms.

This chapter is organized in two main sections; experimental methods and results. First, in section 8.2, the experimental methods are presented. This introduces the methods applied to evaluate the multi-resolution and sequential process. The section also explains the data used for this assessment. The experimental results from the application of these methods to different data types are presented and discussed in section 8.3. The problem of noise involving in surface

detection, geometrical scaling and registration itself contributes registration errors which are analyzed in this section. The behaviour of the least square distance function to different transformation parameters is also explored. Finally, the issue of accuracy and cost calculation and the analysis of the results are discussed in sections 8.3.4 and 8.3.5. A summary of the results and discussion can be found at the end of this chapter.

8.2- Experimental methods

All the experiments were performed on sets of known arbitrary-transformed data obtained by the following procedures.

Data type A: The images selected for the experimental computations consisted of a set of 2-D adjacent *T2-weighted MR* slices. To simulate the registering data for this assessment, the set of original grey-scale *MR* data were misregistered arbitrarily by different known transformation parameters (e.g. 5, 8 and 10 voxels, shift, and 5, 8,10 degrees, rotation, with respect to x, y, and z direction). These data sets were not truncated after misregistration. The surface of these data sets were formed and stored for surface fitting process and thus for the evaluation of various aspects of the registration algorithm.

Data type B: The original 2-D *MR* slices were degraded before applying the arbitrary misregistration, in order to simulate *PET* and *SPECT* data in a known geometrical environment. This degradation was performed by applying a Gaussian filter to the original *MR* data and then adding Gaussian noise to them.

Data type C: To provide a representative set of data for the real images taken by an imaging system and thus to simulate a more practical set of misregistered images, an alternative set of data was generated. The *MR* data were arbitrarily misregistered by setting different physical parameters (e.g. rates of gradients) of the imaging system itself (i.e. setting different combinations of x-y-z gradient amplitude of the *MR* slice-selection parameters which generate the slices in different directions).

Data type D: External markers visualized in both *MR* and *SPECT* were used. The images taken

with these markers were registered by checking the marker position in both studies. Knowing the geometric information between the two original data sets, *MR* and *SPECT*, the multi-resolution process was then applied and verified.

In the first set of experiments, an attempt was made to register the generated (well-defined) images with the original *MR* slices using the sequential multi-resolution algorithm. Different parameters (e.g. resolution and threshold) were set and tested in these experiments. In order to verify the fidelity and the performance characteristics of the sequential and multi-resolution methods, alternative sets of experiments were designed, in which all the mismatch measures (*MSD* or *MDE* values) were examined in a search window around the optimal transformation (true known location). These transformations were applied following the registration of the two arbitrary-misregistered data sets. For obtaining the performance of the process at other low resolution levels, the original high resolution surfaces were degraded (i.e. geometrically scaled) before the registration. Then the corresponding geometrical transformation (e.g. a shift of 1 voxel on image size of 32^3 in respect of 8 voxels shift on image size of 256^3) was applied at each level to register the two surfaces, originally, before assessing the search window.

When checking all the transformations in a search window, the minimization process recognizes some locations whose *MDE* values are less than a pre-defined threshold (or their cumulative errors are less than the variable threshold). Thus if only the known mismatch locations are checked, then any match location recognized by the process must be a false match. The probability that such an observed occurrence of the distance error being less than threshold T is called *False Alarm Probability* (*FAP*). The window search designed for *FAP* constitutes all the transformations defined in the proposed search window (i.e. $n_W^{L0}=(8\times2+1)^6$ locations corresponding to -8 to 8 voxels, shift, and -8 to +8 degrees, rotation, in x,y and z directions; see section 6.3) except the actual true match location itself and except those neighbouring locations which can be assumed as match locations (e.g. $W^{L0}=(1\times2+1)^6$ locations corresponding to a shift of -1 to +1 voxel, and a rotation of -1 to +1 degree, around the true match location; see region k1 in figure 6.6). In order to evaluate the most uncertain false match locations, a smaller window size was considered for evaluating *FAP* in most experiments. This was done by constituting only the locations which need a minor extra transformation (e.g. ±2 to ±3 voxels, shift, or ±2 to ±3 degrees, rotation; see region k2 and k3 in figure 6.6) in order to be counted as match location.

This helps to established the response of the registration process to various unacceptable transformations which are uncertain. An equivalent transformation range was applied in other resolution levels.

Similar sets of experiments were designed using the above simulated data sets, but at a number of locations which can be assumed as being acceptable matches (defined as true match locations). These locations can be defined as the neighbouring transformations of the true match, where they just pass the test for being as a true match location (i.e. they are very close to the true match location; e.g. region k1 as shown in figure 6.6). In these experiments, the images were also registered in advance using the known transformation parameters. Then the distance errors (*MDE* values) at true match location and its neighbourhood (see figure 6.6 for a schematic diagram of these neighbours) were examined and compared to a preset threshold. The probability of the occurrence of the error being less than the threshold is, then, counted as *True Match Probability* (*TMP*). Ideally, the suggested match and mismatch locations need a small perturbation to reverse their behaviour to being as mismatch or match location, respectively.

In order to be able to use the central limit theorem, and thereby to have a normal error distribution, the cumulative error (being the sum of a number of individual distance errors; e.g. 6 to 10 individual errors as confirmed empirically) was used. The errors were then compared to the corresponding variable threshold (i.e. the threshold corresponding to the number sample points which contribute to the error). The results of applying these experiments are outlined in the next section.

8.3- Results and Discussion

The various parameters (e.g. threshold, number of sample points, scaling noise, etc) which characterise and effect the performance of the registration process are studied in this section. The sensitivity and behaviour of the least square distance function to different transformation parameter is also revealed in this section. As explained in the previous section, in order to obtain a precise behaviour of the distance function to various transformations, a number of acceptable match locations are used for the evaluation of the true match location. The false alarm probability, on the other hand, is obtained by a number of known mismatch locations.

This section is organized as follows. The error density distribution of the distance errors between the two registering surfaces at both match and mismatch positions is presented in section 8.3.1. The probability of the below-threshold match at different resolution levels using various threshold values are examined and presented in section 8.3.2. This section consists of a number of subsections outlining various aspects of the probability analysis. The probability of false matches are demonstrated in section 8.3.2.1. To assess the number of points matched in a particular location (either match or mismatch location), the probability of matching points are outlined in subsection 8.3.2.2. The probabilities and then performance of the sequential process is studied and presented in subsection 8.3.2.3. Next, the response of distance error measures, at differen: resolution levels, to various locations near the minimum are studied and presented in section 8.3.3. The accuracy measurement and cost calculation are addressed in section 8.3.4 and 8.3.5, respectively.

8.3.1- Error density distribution

The error density distributions of the registering MR-simulated data sets described in the previous section are shown in figures 8.1a-d, and their statistics are presented in table 8.1. The error measures in these results are based on a large number of surface points (e.g. not less than 300) obtained at a number of match locations (i.e. when two surfaces have been registered), or at a number of mismatch locations. The graphs in these figures show a normal density distribution (conformed by e.g. the *Kolmogorov-Smirnov* test (Papoulis 1965)) which is due to using the accumulated sum of the individual distance errors. The two graphs (peaks) presented in each figure correspond to error densities at a number of match and mismatch locations, respectively. The common region between the two graphs belongs to some uncertain (ambiguous) match or mismatch locations.

8.3.2- Probability analysis

The probabilistic analyses of the multi-resolution matching given in this section emphases the proper choice of the number of sample points and investigates the effect of different threshold levels. Unless otherwise stated, the results of this section are based on the data and methods introduced in section 8.2 for obtaining true match probability (*TMP*) and false match probability (*FAP*).

Statistics	Level L0	Level L1	Level L2	Level L3
Sample size	3158	728	512	188
Average	1.64	1.01	0.797	0.417
Median	1.64	1.02	0.83	0.39
Mode	1.67	1	0.94	0.32
Variance	0.18	0.05	0.04	0.02
Standard	0.425	0.22	0.20	0.14
deviation	0.0076	0.0083	0.0089	0.0102
Standard error	0.66	0.42	0.3	0.17
Minimum	2.84	1.66	1.29	0.73
Maximum	2.18	1.24	0.99	0.56
Range				

The results based on the MDE values at a number of match locations.

Statistics	Level L0	Level L1	Level L2	Level L3
Sample size	39500	12100	2150	950
Average	3.35	2.61	1.42	0.61
Median	3.39	2.62	1.42	0.6
Mode	2.31	2.82	1.4	0.5
Variance	0.65	0.25	0.045	0.026
Standard	0.806	0.50	0.211	0.160
deviation	0.0278	0.0103	0.0068	0.0049
Standard error	1.46	0.99	0.82	0.18
Minimum	5.18	3.87	1.89	1
Maximum	3.72	2.88	1.07	0.82
Range				

The results based on the MDE values at a number of mismatch locations.

Table 8.1- Statistical information of the error density distribution of the *MR-MR* registration at different resolution levels.

Figure 8.2 shows how the *TMP* increases when using a higher threshold level. The curves in figure 8.2a show that, apart from the lowest resolution level (L3), the probability of match increases as the resolution decreases. This is due to the fact that a simpler shape and inherently smaller misregistration error (see section 7.2.3.1) is expected at a lower resolution level. The occurrence of a significant scaling error in level L3 (i.e. size 32^3) causes a relatively lower *TMP* for this level. The results shown in two sets of graphs in figures 8.2-c and 8.2-d indicate the performance of various number of registering points (5 to 200) at levels L0 and L3. In general a higher probability of the true match is obtained when the number of sample points is high.

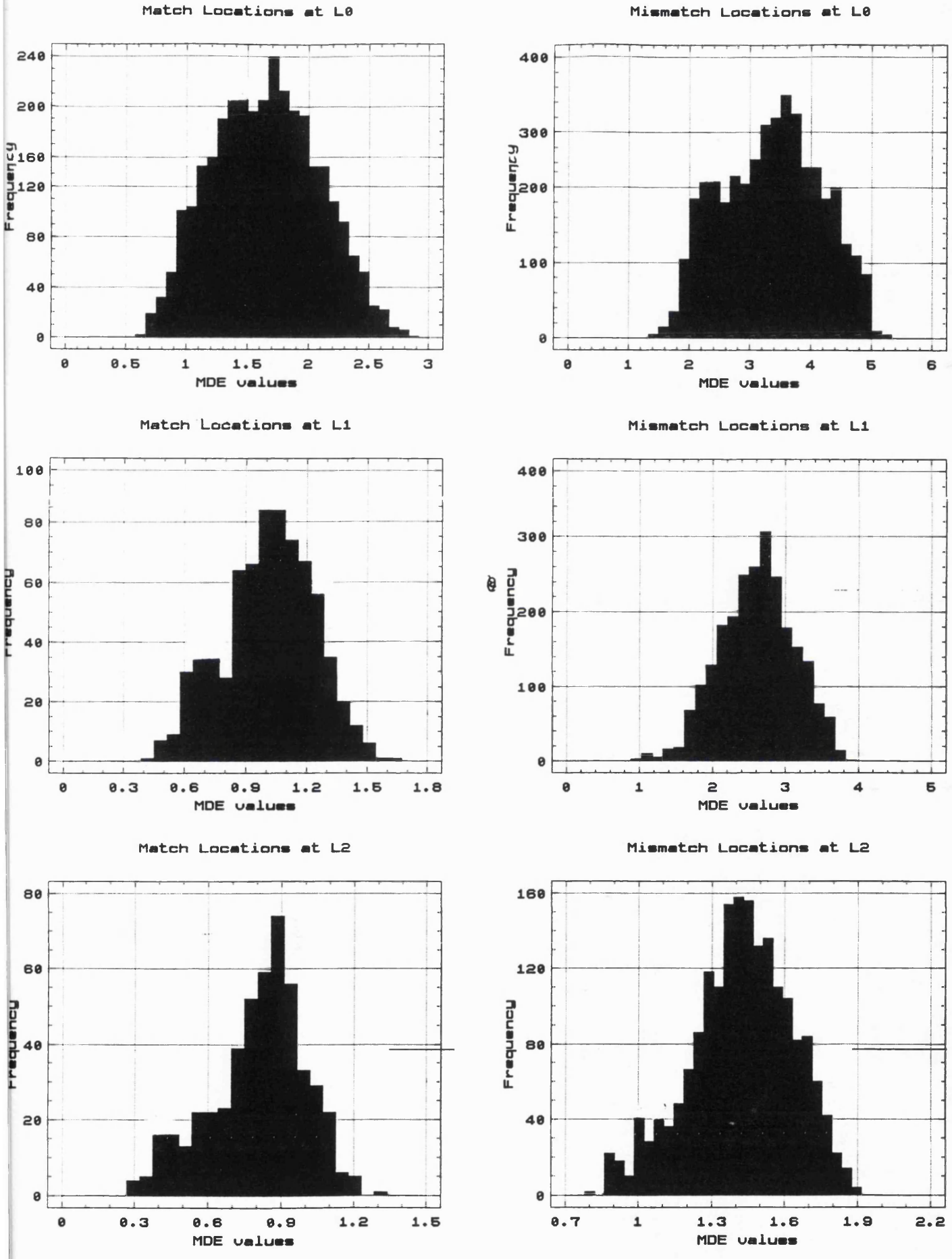


Figure 8.1- a-d) Error density distribution at a number of mismatch (i.e. false match) locations, and at a number of match locations (see section 8.2) for *MR* to *MR* (brain image) matching. These results are based on the sum of 10 individual errors.

The graphs in figures 8.3a-d show the response of TMP to the various size of **P** (number of points). As shown in these figures, the probability of the match increases with number of the points in all the resolution levels except the level L3 which has a roughly similar response (i.e. shows a slight increase in *TMP*) to all different number of registering points. In general, the effect of the size groups of the sample points at different resolution levels is determined by the overall effects of various types of noise (e.g. surface detection noise, scaling noise, and misregistration noise) incorporated in the registration process. However, as will be shown in figure 8.12, in spite of these errors resulting from incorporation of noise, both images being registered are consistent at all resolution levels. That is demonstrated (see table 8.2) by the registration of two well defined image models. Eventually, their alignment lead to a transformation parameter which is very close to the optimal transformation in the highest resolution level, and their mismatch value is of the order of one voxel precision.

8.3.2.1- False alarm probability

The false alarm probability (*FAP*) examined by 5×10^5 mismatch locations is shown in figures 8.4-8.6. These transformations belong to a window size of 9 rotational and translational shifts in each direction, obtained by W^L-W^L_{nei} (see section 8.2). The graphs shown in figures 8.4 and 8.5 present false alarm probabilities versus threshold value at different resolution levels. The *FAP* results in respect to different number of sample points are presented in figure 8.6. Finally, figures 8.7a-b show the *FAP* versus threshold values, corresponding to the registration of *MR* and *SPECT* brain data sets. In this respect, figures 8.7c-d show these results versus the number of sample points. The performance characteristic of matching process at different resolution levels and for various number of sample points are illustrated by *Receiver Operating Characteristic* curves (*ROC*) in figures 8.8 and 8.9, respectively. The results of these figures and the discussion are outlined in the following paragraphs.

I: As shown in figures 8.4 and 8.5, the false alarm probability increases as the threshold increases. This threshold can lead to a reasonable probability, even at the lowest resolution level, if it is set properly. The performance characteristics of the *MR* to *MR* matching at various threshold values shown in figures 8.2 and 8.4, indicates how the threshold plays a critical role in detecting a true match location or, on the other hand, in rejecting a false match location. Although, a lower threshold level is preferable in order to obtain a lower *FAP*, it is not

recommended in lower resolution levels due to lowering *TMP* which might cause the false rejection of a likely match location. When processing in a low resolution level, a higher *FAP* might be acceptable in the favour of a higher *TMP* in order to guarantee the inclusion of the true match location inside the locations passed to a higher resolution level. The fidelity of the registration process is due to the fact that all these uncertain locations which are in the vicinity of the true match location would be tested subsequently in a higher resolution level, where their *FAP* are significantly low.

II: The effect of various types of noise on the probability of match, discussed in the previous section can also be indicated in the graphs of figure 8.5. As the images get coarser, the expected distance error between the two surfaces tends toward zero, so that for any threshold 'g', the *FAP* increases. However, the result obtained from the graph of the image at level L1 (size 128^3) shows how a proper image scaling method can improve the performance of the registration process. In spite of the error introduced during scaling process, the *FAP* at this level is smaller than that at the higher resolution level (i.e. size 256^3). This confirms the noise suppression effect of the scaling (degradation) process. Nevertheless, the level L3 (size 32^3) has undergone a few levels (stages) of the scaling process and thus was incorporated with a higher level of noise. In fact the noise due to scaling degradation, that is shape change, increases, while noise on the surface, resulting from edge detection decreases. The inconsistency in the surface shape (shape deformity) due to this noise phenomenon causes a relatively high *FAP* at the level L3.

III: It can be seen in figure 8.6 that the *FAP* increases as the number of sample points decreases. The effect is due to the existence of a better shape and geometrical information of a surface when a higher number of points is used to define the surface. The *FAP* of the level L3 is much higher than other levels. This can be due to a smaller misregistration noise (error; see section 7.2.3) obtained at a lower resolution surface. Moreover, as illustrated in this figure, a lower resolution level is less sensitive to the number of sample points than a higher level especially when the number of these points are small. In higher resolution levels, the *FAP* decreases substantially with the number of points where the number of examined points are below a certain level (e.g. 20). As shown in this figure, no such significant change occurs in the *FAP* of the level L3. This shows that a smaller number of sample points can be used to define a surface in a lower resolution image. The results obtained at level L1 (i.e. size-level 128) by using a small number of points demonstrates the fact that even some fewer number of points selected from a

higher resolution image (i.e. having bigger size) can carry enough geometric information with less noise about the whole 3-D surface.

IV: As shown in figure 8.7, in general the same results as for *MR-MR* registration was obtained for *MR-SPECT* data registration, except with a more distinct response of *FAP* to the number of points. In general all types of noise (surface detection, misregistration and scaling noises) are expected to increase in *MR-SPECT* registration. These imply a higher expected error in the true match location and thus imply a higher threshold value. Therefore, a higher *FAP* is obtained when registering *MR-SPECT* data. The smaller difference shown between the *FAP* of levels L0 and L1 than that illustrated by *MR-MR* registration, confirms that the degradation error causes more changes in the slope of the *FAP* curve at a lower resolution level. A sharper response in the *FAP* of *MR-SPECT* registration than *MR-MR* is shown with respect to the increase in the number of sample points. This can also be explained due to the greater influence of noise and therefore less geometrical and shape information when a lower number of points are employed.

V: As shown in figure 8.8 the performance characteristic of the matching process changes as the resolution level varies. Note that the curves with higher probability of true match and lower probability of false match belong to the higher resolution levels. Nevertheless, the performance at all levels is reasonably high, allowing the registration of the two subject surfaces, if a good threshold level is employed. The *ROC* curves shown in figure 8.9 are based on the probability measures obtained by using four different size groups of sample points at different threshold values. The superiority of using a larger number of points is evident from these curves.

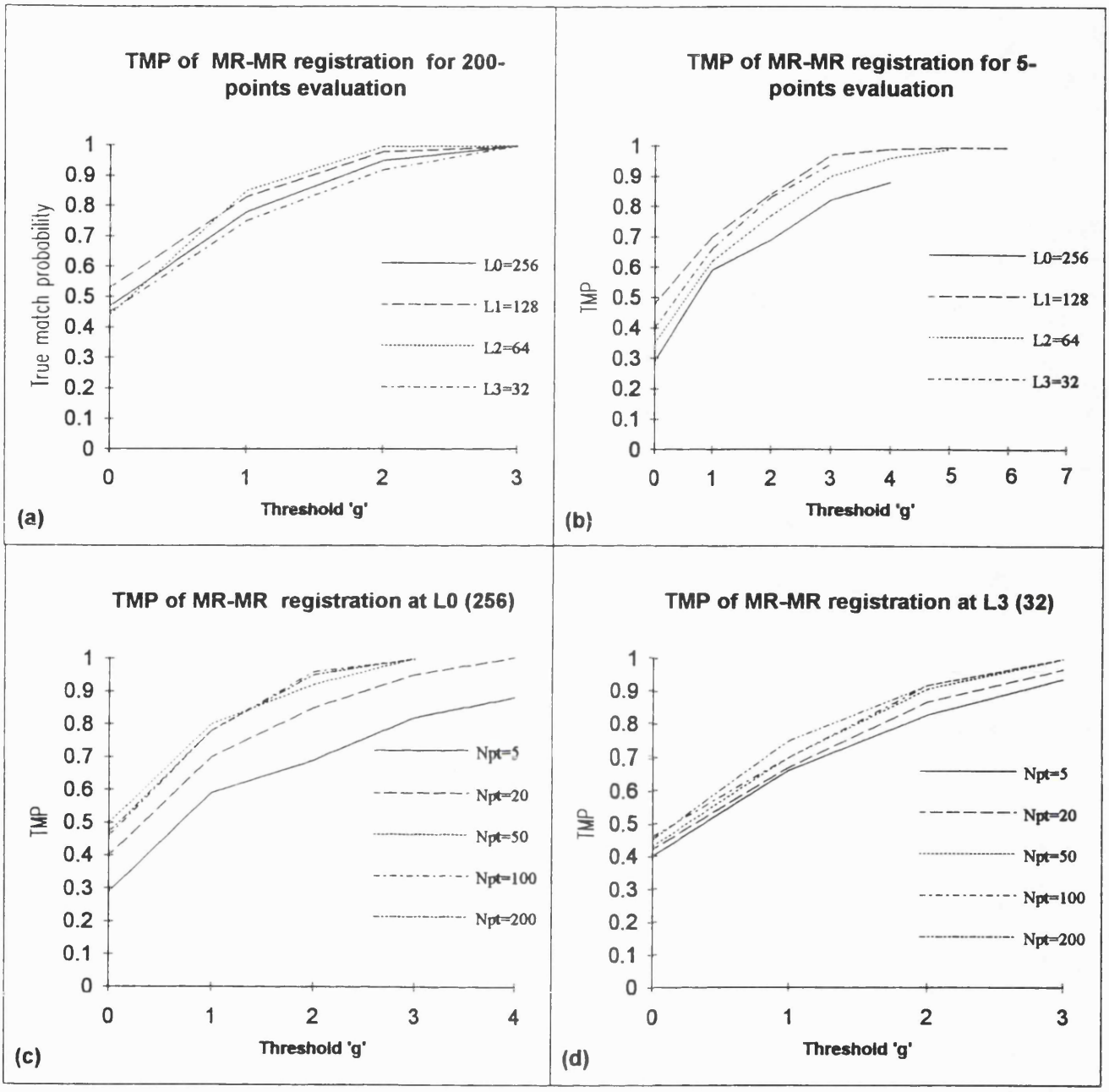


Figure 8.2- The probability of the true match (TMP) as a function of threshold (in terms of g; the number of S.Ds) for the *MR* to *MR* registration. Level L0 to L3 belong to image size of 256*256*256 to 32*32*32, respectively, (denoted by L0(256) to L3(32)). NPT is the number of sample points. **a)** Registration using 200 sample points; **b)** using 5 sample points; **c)** Registration at level L0 (i.e. image size of 256^3); **d)** at level L3 (i.e. image size of 32^3).

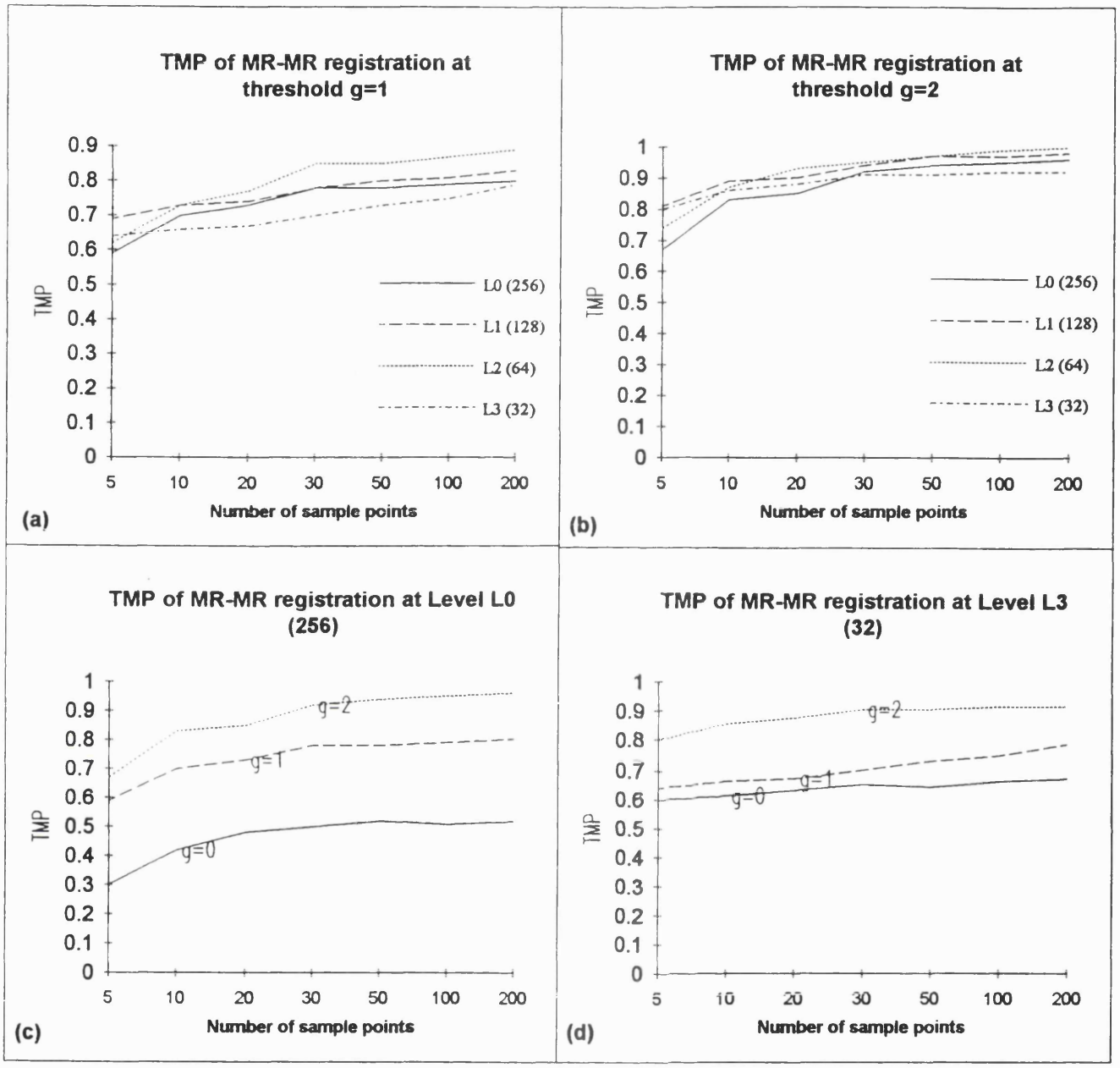


Figure 8.3- The probability of the true match (TMP) versus the number of sample points for the *MR*-to-*MR* brain data registration. 'g' defines the threshold as number of deviation from the mean (expected) error. **a)** At threshold g=1; **b)** At g=2; **c)** At level L0 (i.e. image size of 256^3); **d)** At level L3 (i.e. image size of 32^3).

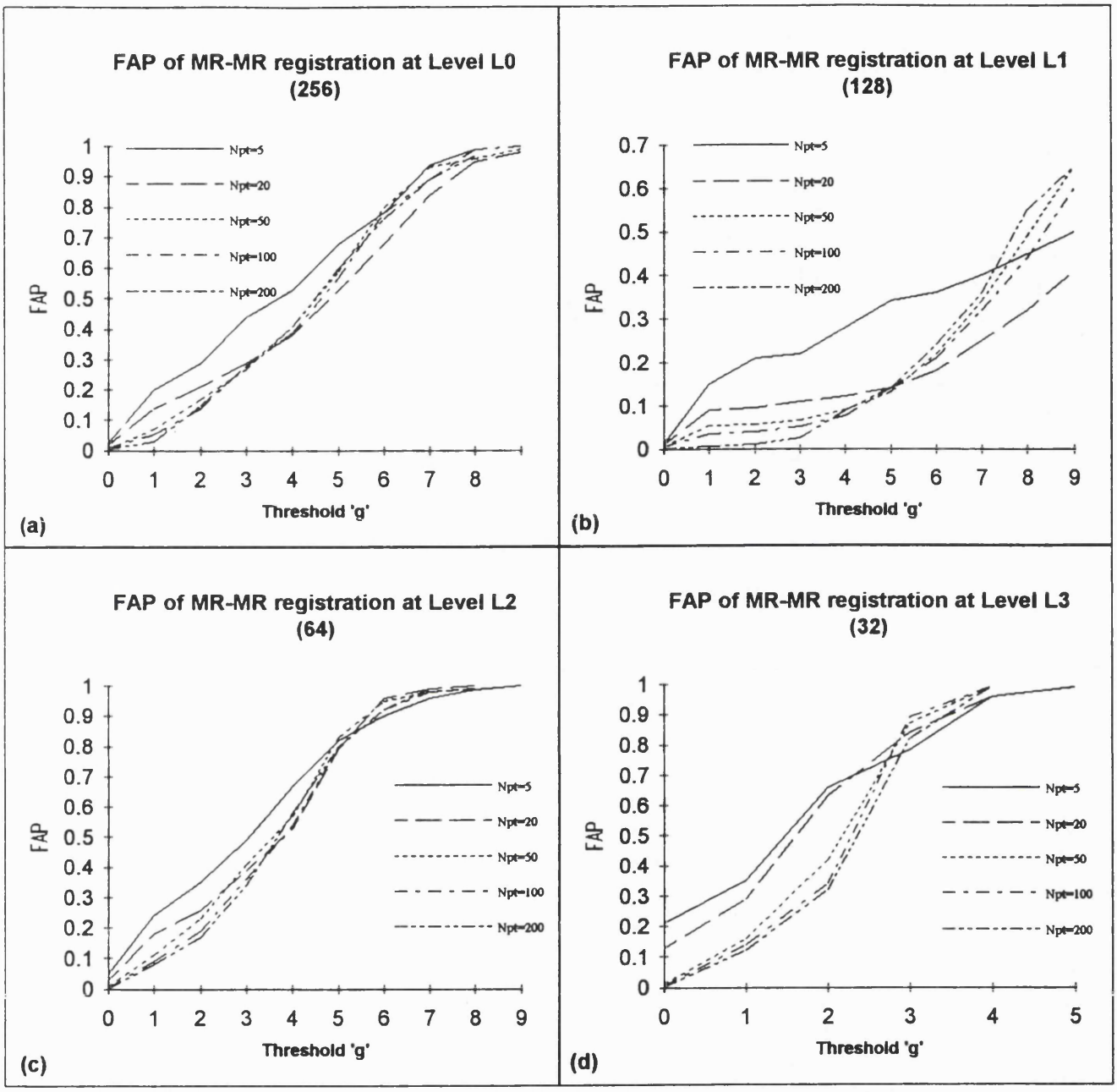


Figure 8.4- False alarm probability (FAP) of the matching process versus threshold 'g' (number of S.D. from the mean error) for different number of points (NPT) at; **a)** Level L0 (denoted as L0(256); i.e. image size if 256^3); **b)** Level L1; **c)** Level L2; **d)** Level L3 (i.e. image size of 32^3)

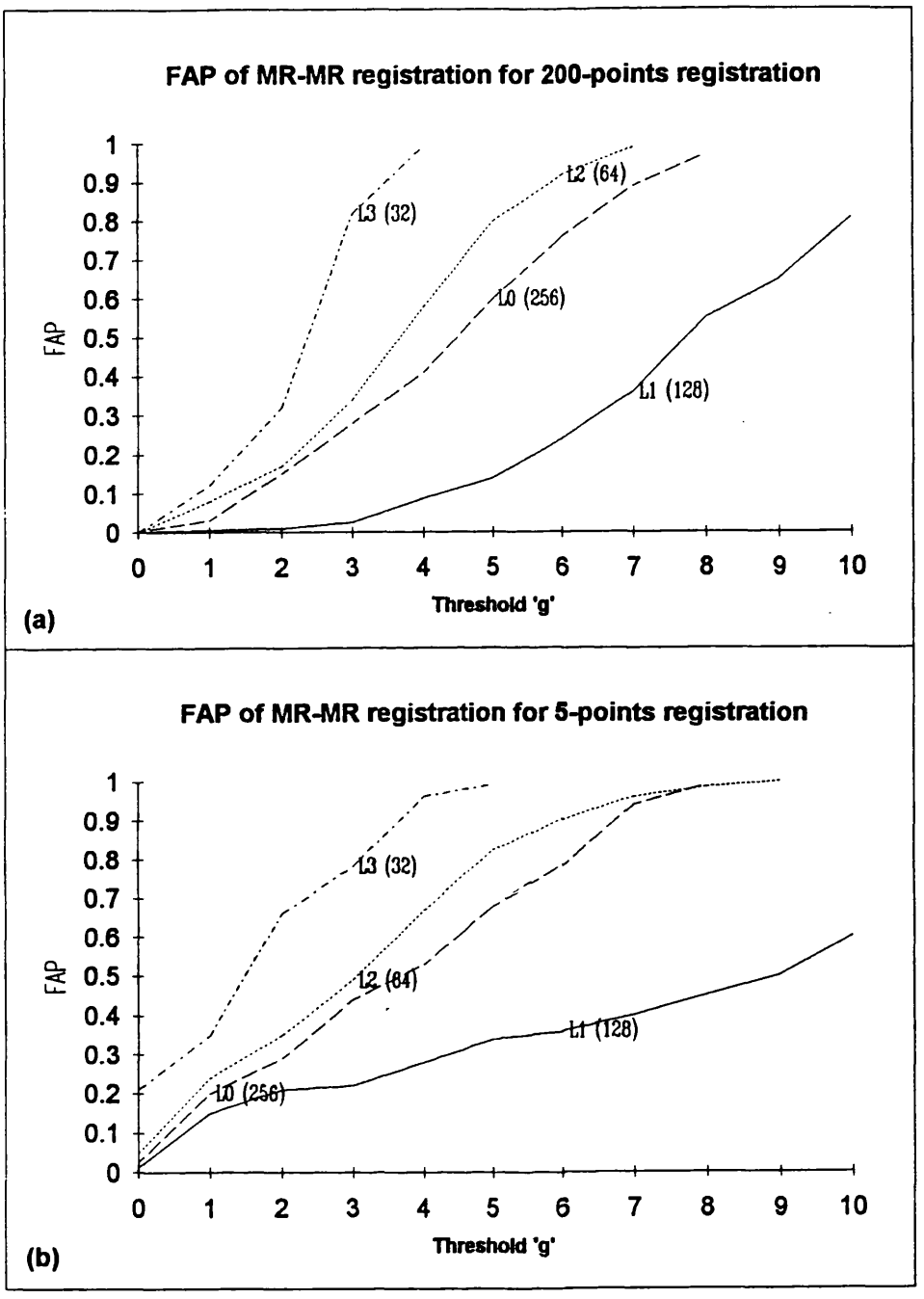


Figure 8.5- False alarm probability (FAP) versus threshold at different resolution levels. **a)** Registration using 200 sample-points; **b)** using 5 sample-points.

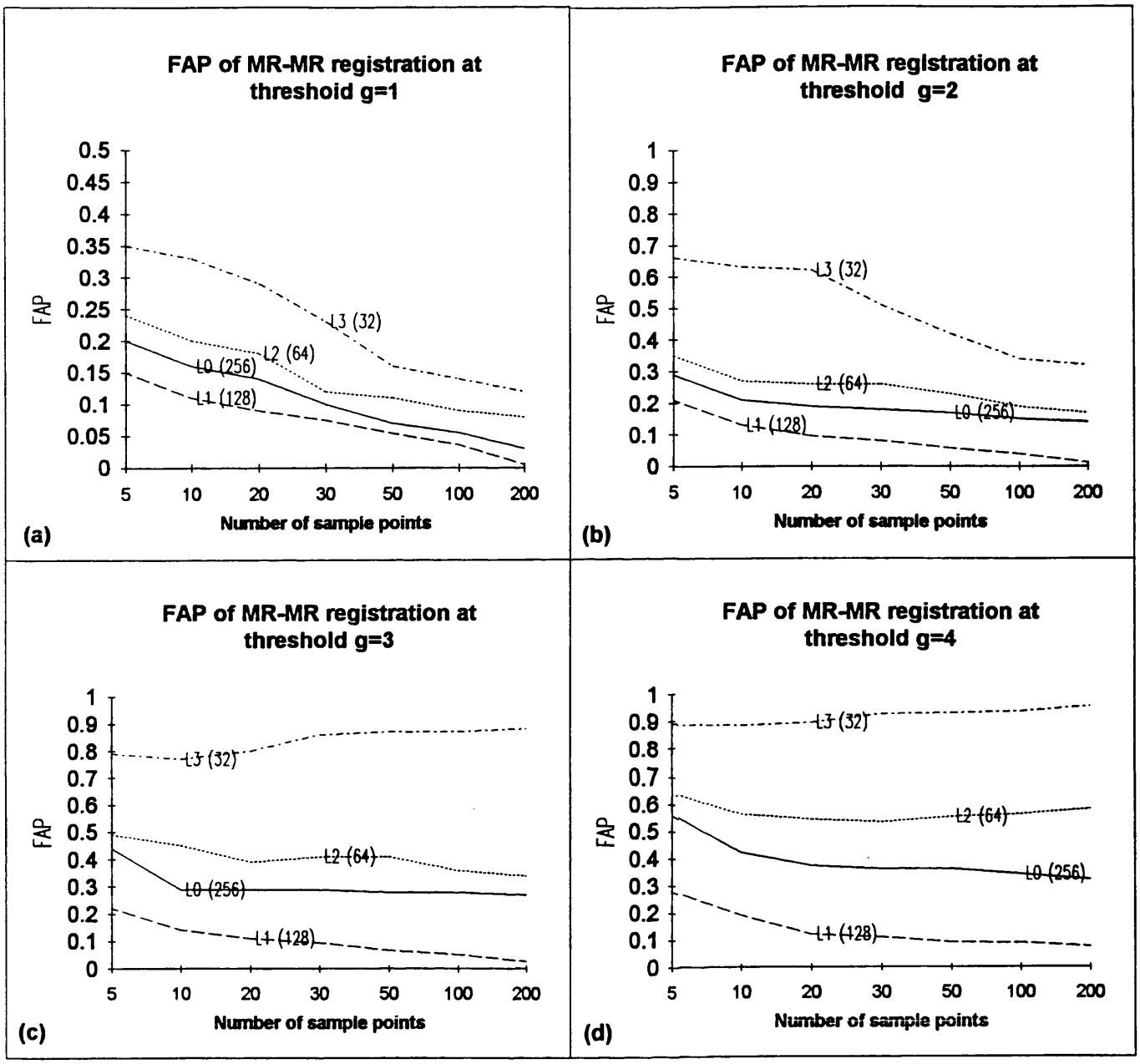


Figure 8.6- False alarm probability (FAP) of the matching process versus number of sample points at levels L0 to L3 (size 256^3 to 32^3, respectively). 'g' is number of S.Ds (standard deviation) from the mean. **a)** Registration at threshold (g=1 S.D.); **b)** at g=2 S.Ds; **c)** at g=3 S.Ds; **d)** at 4 S.Ds.

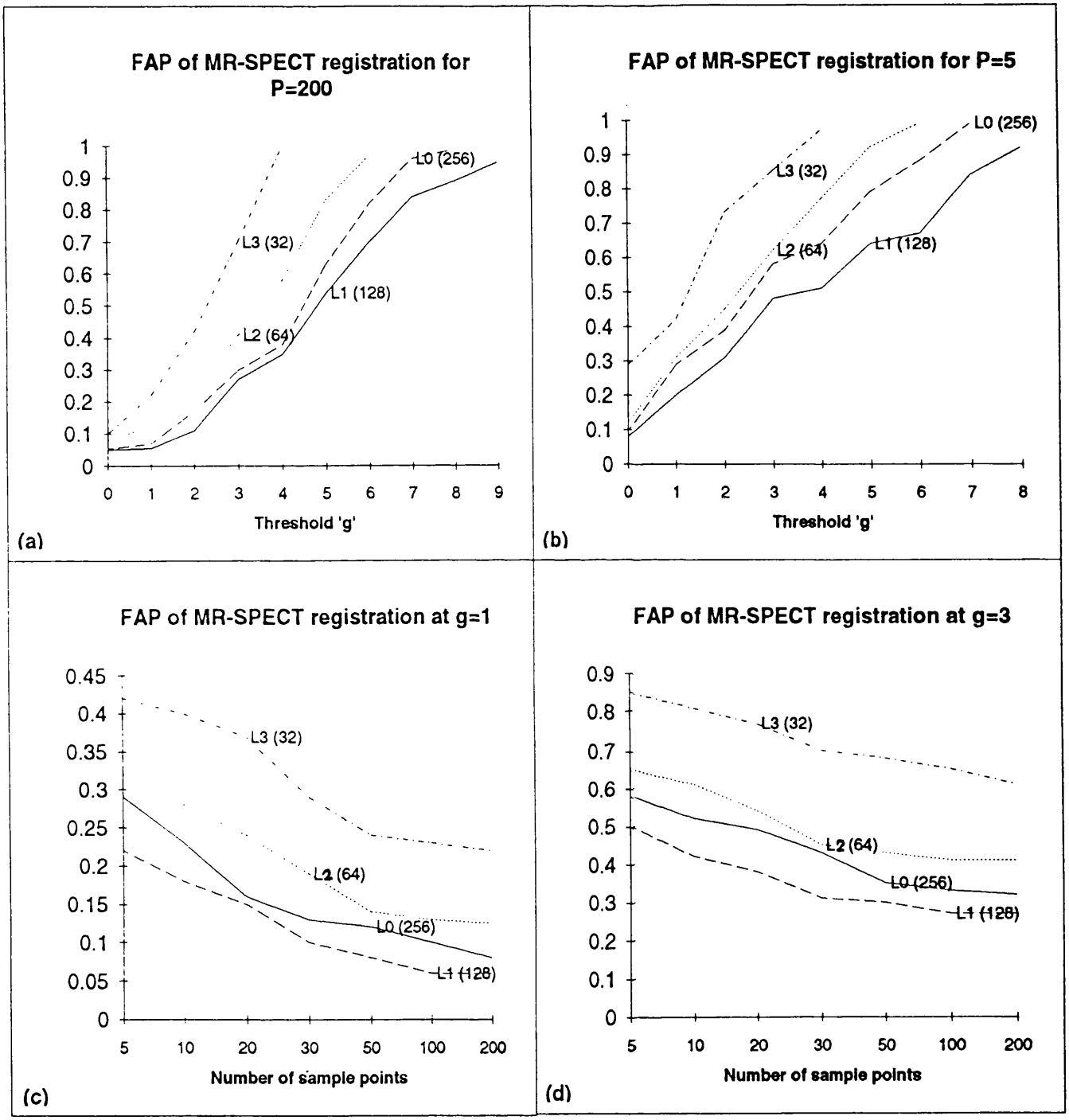


Figure 8.7- False alarm probability (FAP) for *MR-to-SPECT* brain image registration at levels L0 to L3 (i.e. size 256^3 to 32^3, respectively); **a-b)** versus threshold; **c-d)** versus number of points. **a)** Registration of 200 sample points; **b)** Registration of 5 sample points; **c)** Setting threshold level as g=1 S.D. (standard deviation); **d)** threshold level as g=3 S.Ds.

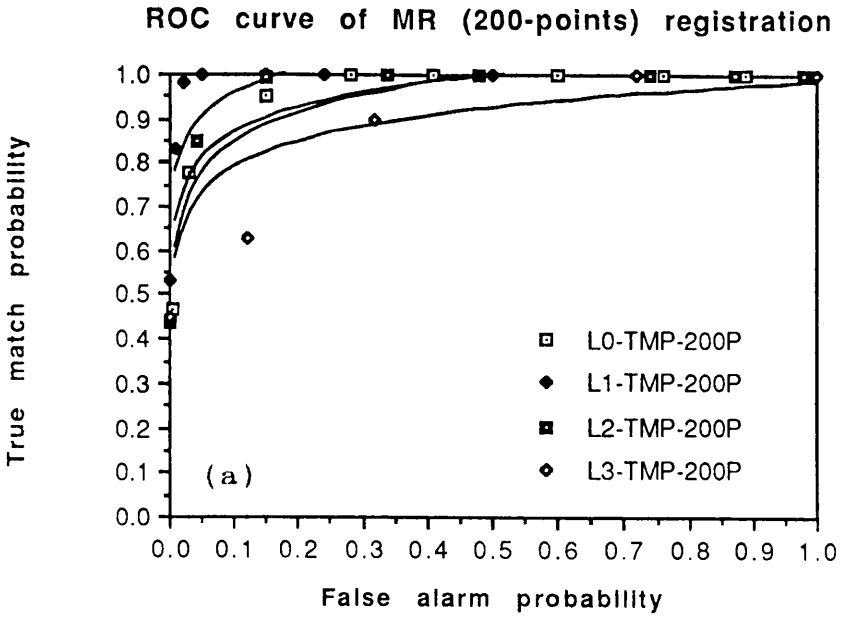


Figure 8.8- Receiver operating characteristic (*ROC* curve) for the registration at different resolution levels, using 200 sample points. L0 to L3 denote image sizes of 256^3 to 32^3, respectively. The curves show that the resolution level L1 is superior to other levels.

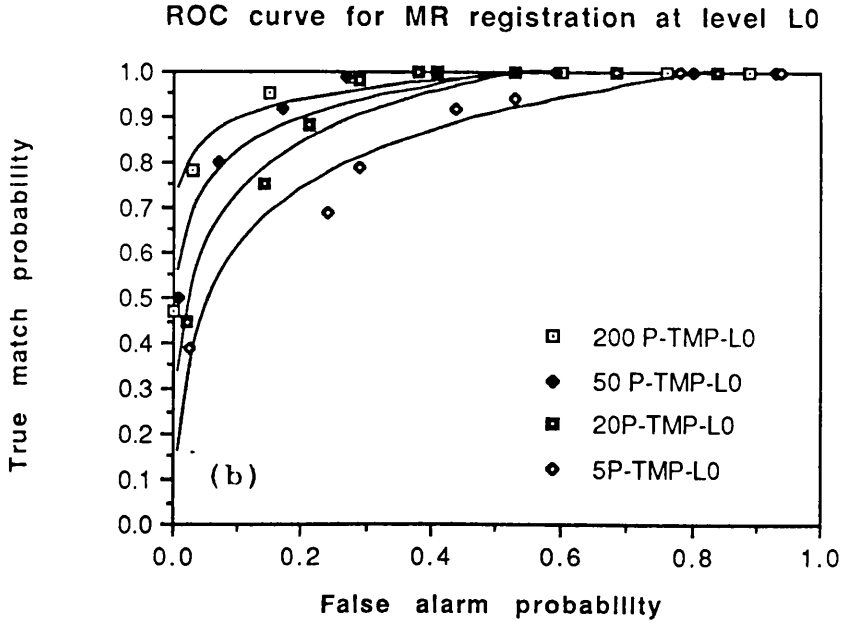


Figure 8.9- Receiver operating characteristic (*ROC*) curves for the registration of different number of sample points (i.e. 200, 50, 20 and 5 points). The curves indicate an improved performance with greater number of points. (Note that some near match locations were used for the TMP in this evaluation; also a number of mismatch locations for FAP).

8.3.2.2- Probability of matching points

In order to assess the role of sample points in building up the cumulative error and compare them at different resolution levels, an alternative set of experiments was performed. The probability was defined as the fraction of number of sample points matched (i.e. have a distance error smaller than a pre-defined threshold) with their corresponding points at some specified locations. Few neighbouring locations adjacent to the true match location were considered for *TMP*. On the other hand, the *FAP* was measured at a few certain mismatch locations.

As shown in figures 8.10a-b, nearly the same response as in the previous experiments is demonstrated by these graphs. Level L3 (size level 32) has the highest number of sample points matched at each location and this leads to a higher probability (i.e. both the *TMP* and *FAP*) than the other levels. The relation between the probabilities at different resolution levels confirms that the distance error is reduced at a lower resolution level due to the alignment (or closeness) of some more registering sample points.

The graphs shown in figure 8.10c illustrates a sharper change in *FAP* of the low-**P** regions of all resolution levels than those with a higher number of sample points. A smaller change is obtained in *FAP*, when the number of points stands above 50 at all resolution levels. The performance obtained by these graphs confirms the preference of some higher number of points in the evaluation of a mismatch value (i.e. the distance between two surfaces).

8.3.2.3- Performance of sequential methods

The method used in this experiment is based on the following transformation set. A range of transformation parameters (e.g. a shift of 0 and 8 voxels, and a rotation of 3, 5 and 7 degrees) were initially applied to the original set of *MR* data (i.e. data types A and B; see section 8.2). The surfaces were detected at each location, separately. For *TMP*, the detected surfaces, from all these locations, were then transformed to the known match location and its neibourhood, and the probability of match was calculated by the methods described in section 8.2. By using this strategy, the effects of the geometrical transformation and interpolation errors are also investigated. The similar procedure was used for calculating *FAP*, except that the surface was not transformed back to the original match location, but was further transformed to -2 to +2

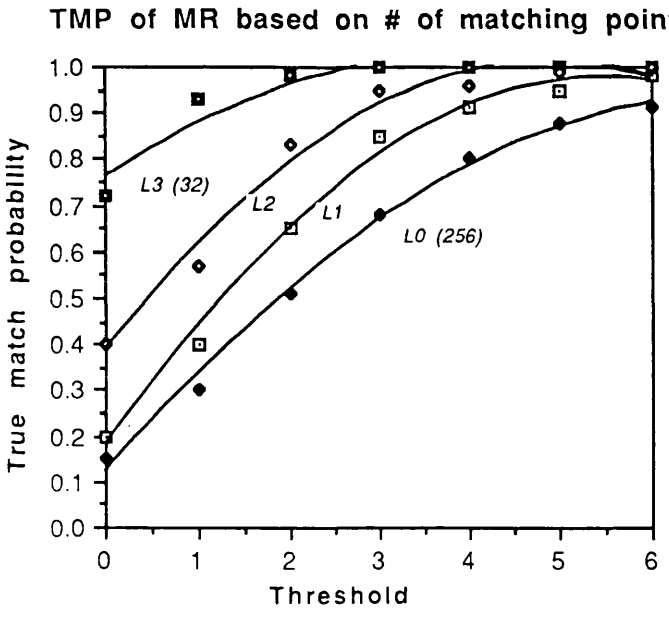
degrees off from the true match location. This procedure leaves the two surfaces at a number of neighbouring mismatch locations (denoted as '*background locations*') which can be counted as true mismatch or false matches depending on the threshold level. These 'background locations' are mismatches in the vicinity of the true match. The decision on the acceptance or, on the other hand, on the rejection of any location (as being a match location) is then made by taking into account a certain number of sample points (e.g. 10, 20, 50, 80, 120 and 150 as done in this experiment). The high resolution surface at each location (after the desirable initial transformation) have the degradation method applied in order to obtain the surface models at all other resolution levels. Then, the corresponding transformation at each level is employed to put the two surfaces at a match (for calculating *TMP*) or a mismatch (for calculating *FAP*) location.

Figure 8.11 shows the probability of match versus the number of sample points where the locations in the background (i.e. outside the acceptable match region) are examined. This figure shows a sharp decrease in *FAP* as the number of sample points increses, where the number of points is small (e.g. less than 20 for level L3 (size 32^3) and less than 50 for all other levels). The *FAP* has a relatively flatter response for higher number of the registering points.

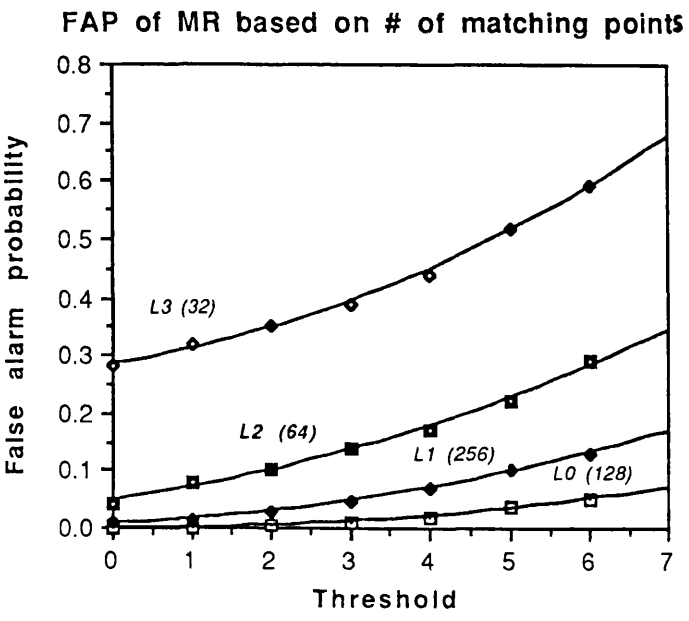
When a sequential process is used, a higher probability of match (i.e. the true or false match probability) is expected at an earlier stage of the process when the number of points are smaller. Accordingly, the fewer the number of points involved in the registration, the higher the probability of below-threshold residuals (i.e. cumulative distance errors) occurs. On the other hand, more (rejecting) examinations are applied at each search location as the number of points comprising in the evaluation increases. Setting a higher threshold level allows any location to be tested further with higher number of points. A high threshold level is not beneficial due to the increase in *FAP* and also in terms of computation effort required by the process. However, when a multi-resolution method is used, all these probable false match locations are evaluated more accurately in a higher resolution level.

8.3.2.3.1- Error growth curve

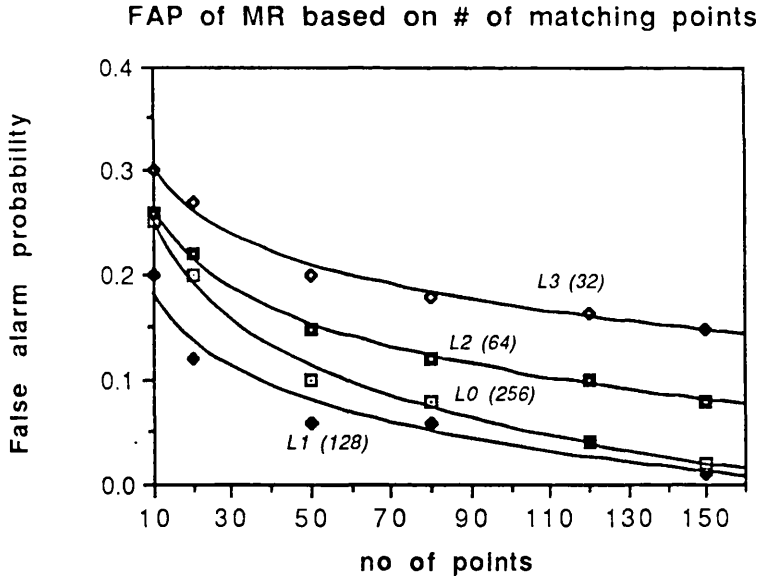
The shape of the error growth curves at match and different mismatch locations is very important for the prediction of error in the sequential process and setting the threshold level used in this decision algorithm. In this respect, the error is incremented as each of the sample points is compared, and the search location is rejected if the mismatch value is bigger than a threshold.



(a)



(b)



(c)

Figure 8.10- The probability of the matching points used in the registration of (*MR*-to-*MR*) brain images. **a)** TMP (i.e. The probability of the points matched to their pair at a match location) versus different thresholds ; **b)** FAP (i.e. The probability of the points matched to their pair at a mismatch location) versus different thresholds; **c)** FAP versus different number of sample points. All the probabilities were obtained at the different resolution levels.

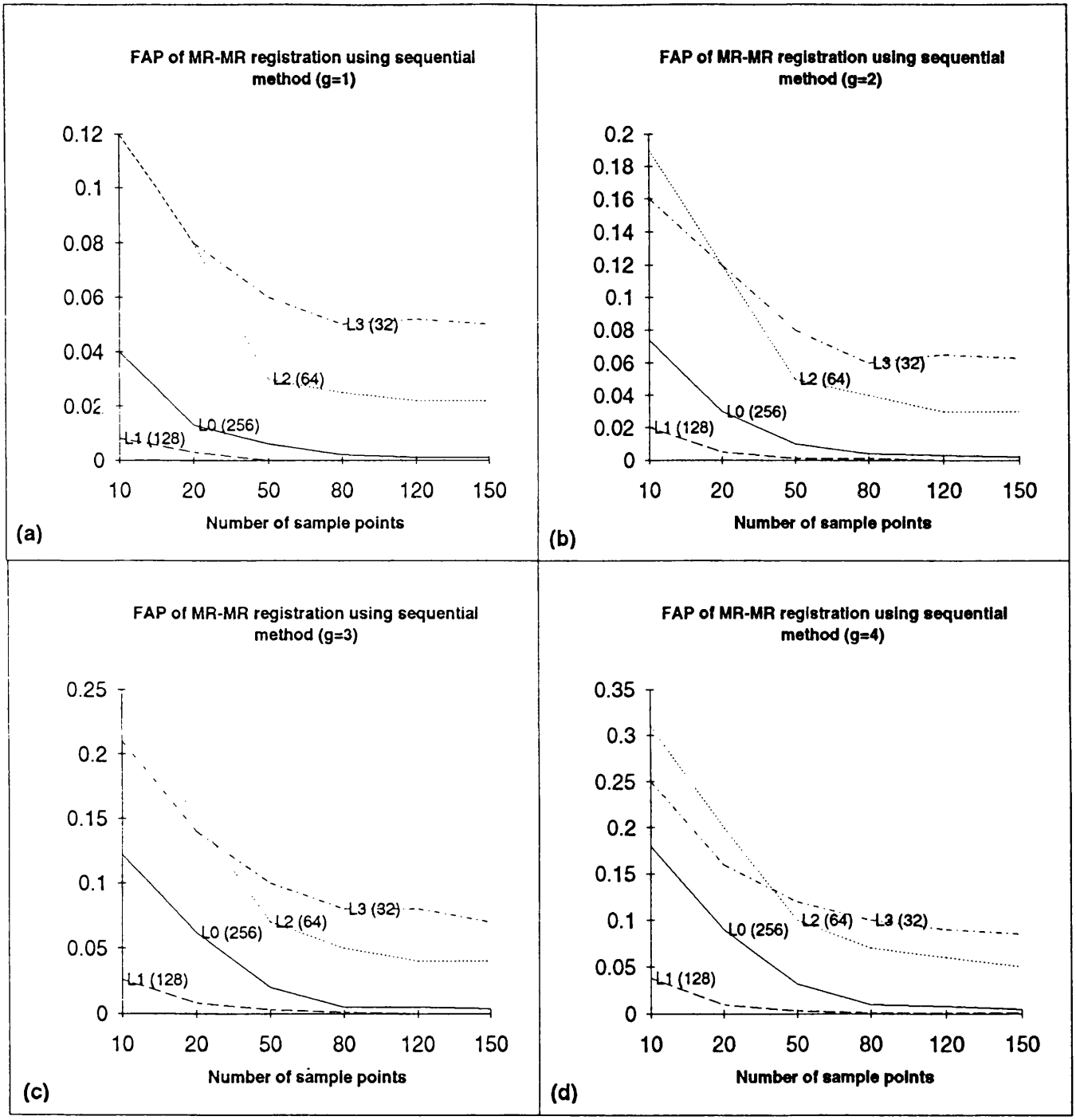


Figure 8.11- FAP of the *MR*-to-*MR* registration based on the sequential decision process, versus different number of sample points. **a-d)** At threshold g=1 to g=4 (g is the number of S.Ds from the mean, used in the threshold sequence), respectively.

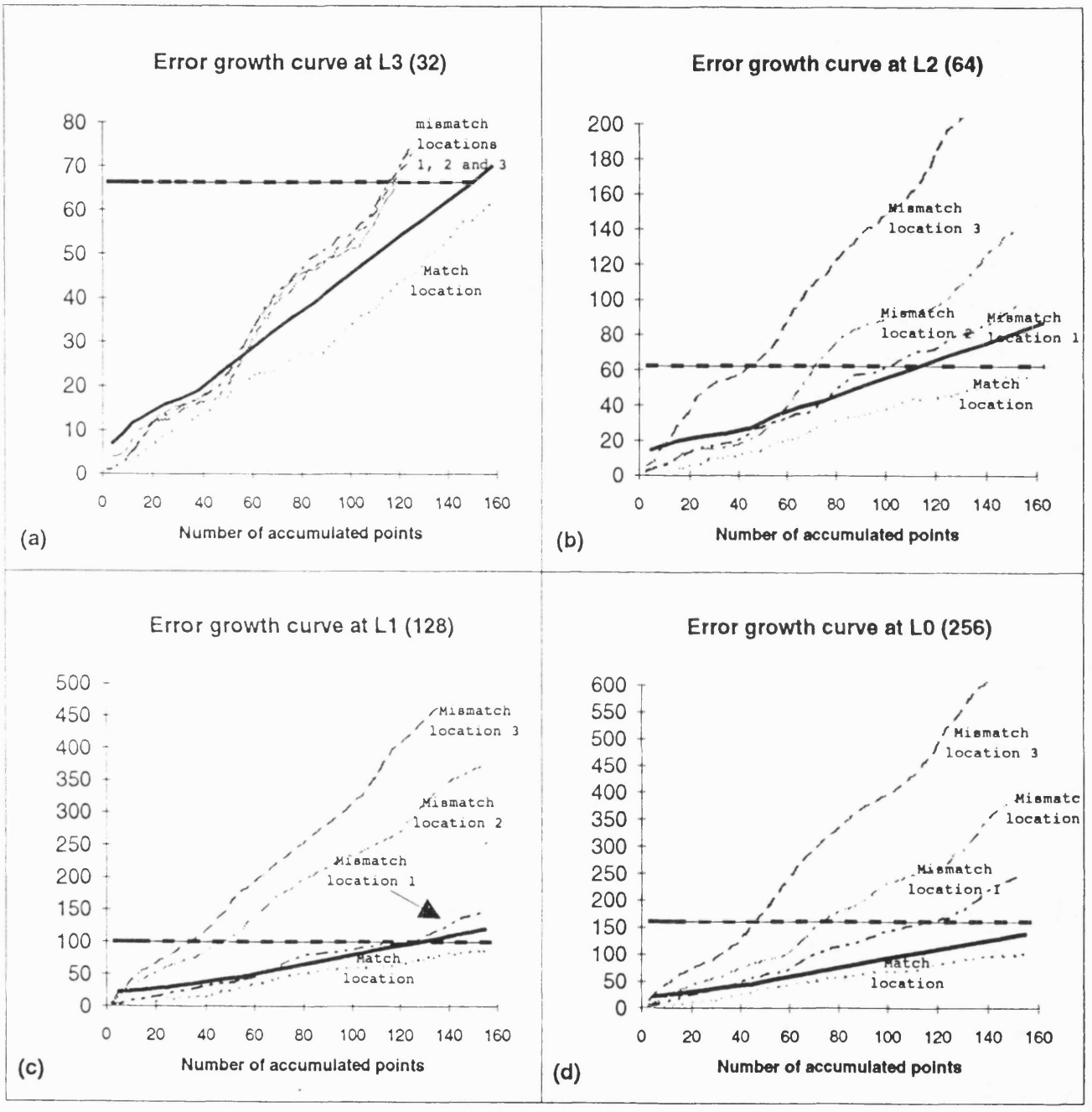


Figure 8.12- Error growth curves at different resolution levels. the curves show the cumulative error values at four different locations (i.e. 4 different transformation settings). The thick plotted lines show the pre-determined variable threshold values for g=2 S.Ds. The superiority of the variable threshold over the constant threshold level is indicated from these graphs. **a-d)** show the results at level L3 (size 32^3) to L0 (size 256^3), respectively.

The cumulative errors in the true match location and three mismatch locations are shown in figure 8.12. The results shown in this figure are based on the cumulative error distance composed of at least the sum of 6 individual error measures. This cumulative concept allows the summation over a larger number of points (i.e. more individual errors) as the error is accumulated. The expected contribution of some points to the cumulative error measure are sufficiently high in some locations and provides a rapid rejection of them based on the variable threshold (as defined in equation 6.8). As shown in figure 8.12b, in the level L2, the mismatch location 1 is rejected after testing 78 points, whereas the match location is not rejected at any point, and thus accepted as a candidate of the true match location for further evaluation. The mismatch locations 2 and 3 are also rejected after testing 54 and 12 points, respectively. The superiority of the variable threshold over a constant threshold level is also shown in this figure. The thick line graph in this figure shows the predetermined threshold value for g=2 (2 standard deviations from the mean) using the equation 7.2. The dashed line in this figure belongs to a constant threshold value at the same value of g.

8.3.3- Sensitivity of the distance function to the transformation

In order to assess the behaviour and sensitivity of the distance function at each resolution level to different size of transformation parameters, a further set of experiments was designed. Two original high resolution registering surfaces from *MR* data were arbitrarily misregistered (e.g. data type A and B), from a match position, by different values of transformation parameters (e.g. 2, 4, 8, and 16 voxels shift or/and degrees rotation). The mean distance error (*MDE*) was then measured at each level after registering them by the multi-resolution process. As shown in table 8.2, the distance function is not sensitive to a small geometric change. The results shown in this table define this small change as being any change less than one voxel when scaled to the lowest resolution level. For example, a shift of 2 voxels at an image size of 256^3 corresponds to a shift of 1 voxel at size 128^3 and 0.5 voxel at 64^3. The multi-resolution technique does not recognize this change at size-level 64 or 32 where the mismatch value is smaller than 1 voxel. The problem with the multi-resolution process can be arised from the fact that the outcome of any low resolution level is used for a higher level. As shown in this table, when a small change is applied to the original high resolution image, it might not be recognized in a coarser image. Thus, when a global search starts for registering two coarse images being originally close to match, any uncertainty in the transformation can be transferred to a higher resolution level.

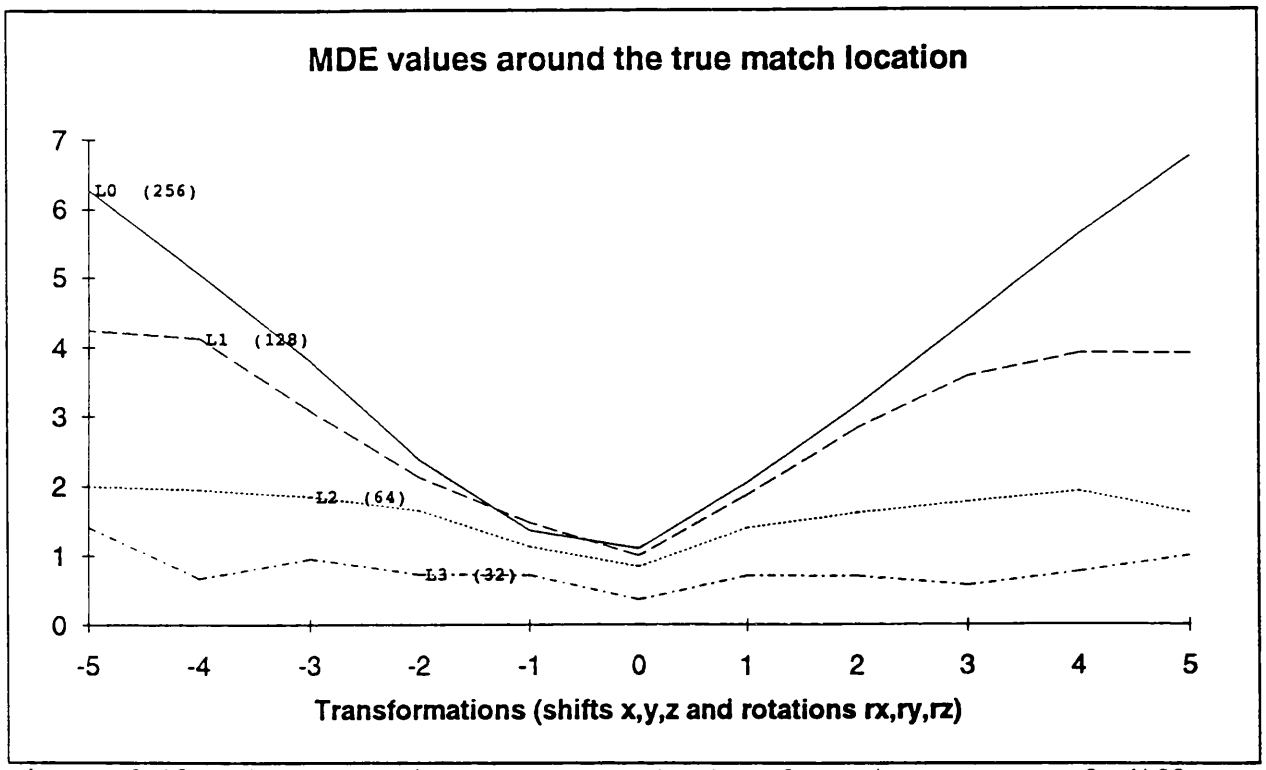


Figure 8.13- The Mean Distance Error (MDE) values in respect of different transformations at different resolution levels for MR data. Different sensitivities are indicated by different slopes of the curves.

The graphs shown in figure 8.13 represent the different sensitivities of the distance function to various transformations around the true match location at different resolution levels. As shown in this figure, the sharper change in *MDE* value of the higher resolution levels confirms that these levels are more sensitive to transformations than the lower levels. It also shows that the range of transformation applied in level L3 was not high enough to provide a reasonable change in *MDE* values in respect of these transformations.

As described in section 7.2, to overcome this problem, the algorithm was designed to recognize the original distance between the two registering surfaces, thereby to decide on the starting level of minimization. Secondly, several of transformations around the minimum found in each low resolution level are used as candidates passed to a higher level.

8.3.4- Accuracy measurement

Convergence of the match to a minimum solution should be possible by either of minimization methods introduced so far. However, the solution is not always correct due to presence of local minima and the contribution of different source of errors. The experiments

show that even at the true global minimum there is a residual e.g. some non zero root mean square distance values (MDE), which may vary due to the repeatition of the registration on the same type of data set. All these create uncertainty in the goodness of a match and, in turn, in the estimation of the match location.

Geometric precision (Accuracy) which shows how good (Accurate) is the estimated location of a match, depends on the distance between the two images (i.e. distance between well-defined reference points) and the standard deviation of these distances. It was shown by Förstner (1982) that apart from the influence of noise (i.e. in the form of correlation value or, on the other hand, signal to noise ratio), the accuracy depends on the sharpness (gradient of the image) of the correlating images. In this sense, high frequency signals (not influence by noise) can lead to better precision.

In order to assess the accuracy of a matching process, some well defined images are required having some well known reference points on them. Measuring accuracy is easy on phantom images, but the results are not exactly relevant to actual clinical images. In the current work the data from multimodality imaging of two different types of phantom, the Hoffman brain phantom, and the Jaszczak phantom were used to assess 2-D and 3-D registration, respectively. The MDE values and standard deviation of 6-10 manually selected reference points on each data set gave an accuracy of 1.44 0.42 mm for MR-MR, 1.82 0.65 for MR-CT, 2.38 0.88 for MR-PET, and 3.17 1.12 for MR-SPECT registration. Note that all these measurement are subject to error in locating the position of reference points.

The accuracy of real clinical data sets was also obtained by setting four external markers on the head in both studies (i.e. MR and SPECT) and the results from 10 such data sets were assessed. This multiple measurement of the registration value shows also whether the residual values obtained by the surface fitting process are due to misregistration (displacement) error, or due to other type of error. The accuracy obtained in registration of MR and SPECT data using multi-resolution surface fitting was 3.47 1.17 mm. In order to have a more accessible method for evaluating the results of multi-resolution in a range of misregistration states, the following data set were used as described in the next paragraph.

The most accessible matching accuracy tests were performed through the use of arbitrarily

misregistered data, degraded to different resolution and noise levels by using Gaussian Filters and Gaussian noise operators. All the methods of measuring accuracy rely on the use of known reference points and are subject to error in locating these points in both images. When using a surface fitting process, all the points contributing to the residual values can be thought as reference points showing the distance between two images (i.e. the accuracy), subject to the condition that the location of the estimated match is within an acceptable displacement (e.g. 1 voxel shift and 1 degree rotation) from the known location. However, when these residual values are used to predict the true expected value (employed in the threshold sequence) for assessment of other un-known misregistered images, the goodness (accuracy) of them should be confirmed. This was done by obtaining the expected error by the use of a group of measurements from different images of the same type, registered after various displacements and tilts (e.g. more than 200). The mean distance error of each image pair and the S.D. of these mean values were, then, calculated to obtain the true expected error. The S.D. of the mean, in this respect, depends on the number of sample points.

The reliability (e.g. reproducibility) of all the experiments were also tested as an alternative measure of the precision and stability of the registration method. An arbitrary misregistration was applied on different images of the same types, before attempting to register each pair of images by the registration process. Their mean value and S.D. were then calculated in order to verify the reliability of the process.

As discussed above, the accuracy of a registration technique can be defined as the average distance error between the two surfaces at the match location. In this respect, the match is obtained by the registration of some pairs of well defined surface images obtained by applying a number of arbitrary transformations. The misalignment allowed (i.e. the decision on a match) between the two surfaces at match location is defined to be less than a certain value (e.g. a shift of 1 voxel and rotation of 1 degree in each direction).

As discussed in chapter 7 (e.g. see section 7.2.4), a higher threshold value at each resolution level allows more locations (transformations) to be passed and thus tested in the next higher resolution level. The results of applying the sequential multi-resolution method are shown in table 8.3. This table shows the number of locations passed to level L0 (i.e. the highest resolution level). This number decreases as the threshold level decreases and thus a faster process

is obtained. However, the accuracy is reduced at a low threshold level due to the elimination of some locations which may be likely candidates for obtaining the true match location.

One can see from the results at table 8.3 that a geometric mismatch of 8 or 10 voxels produces a better accuracy than the mismatch of 5 voxels. This is due to the lose of the partial volume information of the 3-D voxels during the degradation process (see section 7.2.1 and 7.2.4). In spite of this deficiency, it was shown that the algorithm is capable of finding a close match position for both sets of *MRI* and *SPECT* data. The accuracy obtained in this experiment, was 0.81 voxel (about 1 mm) for *MR-MR* and 1.78 voxels for *MR-SPECT* registration.

In the table 8.3, the results of applying both the *multi-resolution* and the *Powell* methods are shown. The *Powell method* employs, a posteriori, the best results obtained by a fast *multi-resolution method* (i.e. with a low threshold setting). As illustrated in this table, an accurate registration is obtained in a reasonably fast manner using both these techniques, together, in the registration process. All likely local minima are also eliminated due to the vicinity of the starting transformation point to the actual minimum point when using a local minimization algorithm.

8.3.5- Cost assessment

Table 8.4 summarizes different parameters involved in the multi-resolution method. Practically, the process might start at different resolution levels depending on the initial mismatch between the two surfaces. The search window size defined in this table includes only the locations (transformations) tested at the initial level. All other parameters shown in the table belong to the initial level. Since only a fraction of the locations are passed to a higher level, the costs of these levels are trivial in respect to that of the initial level.

Three search windows composed of different numbers of locations were suggested for most of these experiments (i.e. about -3 to +3 voxels, shift, and -13 to +13 degrees, rotation, in all x, y and z directions at the coarsest level, L3). Table 8.4 shows the number of operations calculated for the algorithm outlined in the diagram of figure 6.4. The costs (timings) shown in this table result from using the optimal parameters found either by the previous experiments or by the initial knowledge of the surface geometric relations. For example, only few voxels between the two surfaces are required to be traced for an intersection, where the suggested initial

manual interaction can closely align them. Moreover, the cost can further be improved due to the fact that in a sequential process all the sample points are not required to be tested at each location. Accordingly, a faster process is expected when a sequential process is used.

Table 8.5 shows the experimental results of using such a sequential process. The threshold used in this process controls the number of sample points allowed when evaluating each location, which thereby regulates the speed of the registration. As shown in this table, the threshold used for the registration was reliable and permitted a satisfactory speed, in spite of the fact that the process started in level L2 (i.e. image size 64^3). Nevertheless, a better speed could have been achieved at the expense of some loss of accuracy in the process.

8.4- Summary

In this chapter, the application of the *sequential* and *multi-resolution* algorithms developed in chapter 6 and 7 were tested on some well defined clinical data. Some sets of image models were generated to assess the process quantitatively and to predict the expected performance and cost of the algorithm on other data sets. The parameter requirements of the routine imaging processes can be also predicted by the results of these image models.

The probability analyses set forth in this chapter investigated the importance and the effects of different threshold level and various number of sample points at each resolution level. The results of these probabilistic analyses confirmed that a threshold level (g) of 2 or 3 (i.e. number of S.D. of the individual distance error measures at acceptable match locations) is adequate for detecting the true match location while keeping the *FAP* sufficiently low. The feasibility of the registration obtained under different number of sample points demonstrated that using a higher number of points (e.g. 200 in the original finest resolution level; i.e. image size of 256^3) provides a better matching performance than a smaller number of points. The results at other resolution levels confirmed that the required number of points corresponds to nearly half of that of their original image (e.g. 100, 50 and 25 points for size-levels 128, 64 and 32, respectively). However, a smaller number of points (e.g. half of the above number) was shown to be adequate, specially at higher resolution levels, when sufficient care is taken during sampling of points from the whole surface data.

In order to have a reliable match probability, a distance error measured at a mismatch location should correspond to a sufficient number of sample points which are not matched to those points on the other surface. Then, the problem of the existence of some high-distance corresponding pairs (noisy points) contributing to an error measure (*MDE* error or cumulative error) is avoided. A distance error measure at a match location, on the other hand, requires a sufficient number of points matched to their pairs. The experiments outlined in the section '*probability of matching points*', confirms the importance of the choice of the number and position of the matching points used in both match and mismatch locations.

The shape of the *error growth curves* at different mismatch and match locations and the suggested threshold sequence at each level show the advantage of the *sequential process* and the superiority of a *variable threshold* over a *constant threshold* level.

The *sensitivity of the distance function*, at each level, to different geometric transformations indicated the need to start the multi-resolution process at the lowest resolution level where some transformation difference between the two surfaces still exists.

After discussing the problems associated with the registration algorithm in this chapter, a set of experiments were designed with respect to determine the *matching accuracy*. As indicated in the results (see table 8.3), an expected accuracy of about 1, 1.5 and 2 voxels (of 1 mm size) can be achieved for the registration of *MR*, *PET* and *SPECT*, respectively.

The *computational cost* was determined with respect to certain search windows, template sizes (number of sample points) and other parameters required by the process (e.g. threshold). The cost of a multi-resolution process was determined to be in the order of 200 seconds for the actual search process, assuming other parts of the program to be brief enough to ignore. However, the results of applying a sequential multi-resolution process showed that the algorithm is faster even for the whole registration process (i.e. less than 240 seconds on a μVax II). Finally, the cost of about 180 seconds as obtained by applying both the multi-resolution and Powell methods is very satisfactory.

The registration results obtained by using the conventional *MR* and *SPECT* images show that the algorithm as developed is useable in routine imaging studies.

		Transformation x,y,z Rx,Ry,Rz	MDE	np	Transformation x,y,z Rx,Ry,Rz	MDE	np
Original TP mismatch @ L0		2,0,-2 0,0,0	1.93	135	0,0,0 2,-4,0	1.44	146
Obtained TP at match location	L3	1,1,-1 0,0,0	0.07	21	0,0,0 4,8,-6	0.26	17
	L2	1,1,1 0,0,0	0.48	42	1,1,1 0,6,6	0.31	41
	L1	0,0,0 0,0,0	0.67	87	0,0,0, 2,-4,0	0.36	87
	L0	1,1,1 0,0,0	2.39	152	0,0,0 2,-4,0	0.85	152
Original TP mismatch @ L0		4,0,-4 0,0,0	3.51	148	0,0,0 4,-8,0	2.58	149
Obtained TP at match location	L3	-1,-1,0 0,0,0	0.12	17	0,0,0 8,-10,2	0.18	18
	L2	2,0,0 0,0,0	0.50	40	1,1,1 0,2,4	0.30	40
	L1	2,2,-4 0,0,0	1.16	82	1,1,1 2,-8,4	0.42	79
	L0	6,-4,-4 0,0,0	2.32	151	0,0,0 4,-8,0	0.58	153
Original TP mismatch @ L0		8,0,-8 0,0,0	7.07	151	0,0,0 8,-16,0	4.67	127
Obtained TP at match location	L3	2,1,-1 0,0,0	0.11	22	0,0,0 6,-14,2	0.12	19
	L2	2,1,-2 0,0,0	0.42	42	1,1,1 2,-18,2	0.21	42
	L1	4,0,-4 0,0,0	0.25	94	0,0,0 8,-16,2	0.37	83
	L0	8,0,-8 0,0,0	0.56	176	0,0,0 8,-16,0	0.65	140
Original TP mismatch @ L0		16,0,-16 0,0,0	9.60	84			
Obtained TP at match location	L3	1,-1,-2 0,0,0	0.20	5			
	L2	4,1,-4 0,0,0	0.42	42			
	L1	8,0,-8 0,0,0	0.30	94			
	L0	16,0,-16 0,0,0	0.55	176			
Original TP mismatch @ L0		0,0,0 0,0,0	0.25	172			
Obtained TP at match location	L3	1,0,0 0,0,0	0.05	19			
	L2	1,1,0 0,0,0	0.16	43			
	L1	0,0,0 0,0,0	0.21	89			
	L0	0,0,0 0,0,0	0.25	172			

Table 8.2- The results of applying multi-resolution process on different arbitrary misregistered *MR* brain data sets. TP denotes transformation parameters, MDE= mean distance error, n_p= Number of sample points participating in the evaluation (measurement) of distance errors. L0 is the highest resolution level (image size is 256^3). L1 to L3 correspond to image sizes of 128^3 to 32^3, respectively.

Data	Method	Original TP mismatch x,y,z Rx,Ry,Rz	Errors in TP after Registration x,y,z Rx,Ry,Rz	MDE	T at Levels L3/L2/L1	n_W^{L0}	Tm Sec
MRI With *MRI*	Multi-res.	5,8,10 5,8,10	-1,0,0 0,0,0	0.81	2/3/3	368	234
		5,8,10 5,8,10	-1,0,0 0,1,0	1.08	1/3/3	320	195
		5,8,10 5,8,10	-1,0,1 0,1,0	1.25	1/2/2	240	170
		5,8,10 5,8,10	-1,0,0 2,0,0	1.65	1/1/1	176	144
		5,5,5 5,5,5	-1,0,0 0,0,0	0.86	2/2/2	240	433
		5,5,5 5,5,5	1,0,1 0,1,0	1.35	2/1/2	235	419
		5,5,5 5,5,5	-1,1,0 2,0,1	1.82	2/1/1	176	405
		5,5,5 5,5,5	0,-1,-2 2,2,1	2.05	1/1/1	16	180
	Multi-res. & Powell	5,8,10 5,8,10	-1,0,0 0,0,0	0.81	1/1/1	176	172
		5,5,5 5,5,5	0,-1,0 0,0,1	0.83	1/1/1	16	220
MRI With *ECT*	Multi-res.	5,8,10 5,8,10	-1,0,1 2,0,-2	2.28	2/3/3	312	320
		5,5,5 5,5,5	-1,2,0 2,1,-1	2.76	2/2/2	290	395
	Multi-res. & Powell	5,8,10 5,8,10	-1,0,0 1,0,-1	1.78	1/1/1	155	195
		5,5,5 5,5,5	-1,1,0 1,1,0	1.95	1/1/1	37	233
MRI with *PET*	Multi-res. & Powell	5,8,10 5,8,10	-1,0,0 1,0,0	1.6	2/1/1	295	285
		5,5,5 5,5,5	1,0,1 0,1,0	1.45	2/1/1	208	215

TP : Geometric transformation parameters.
T: Threshold; i.e. 'g' as number of standard deviation from mean.
MDE: Mean distance error (here, after registration at the best match).
L3, L2, L1: are resolution levels for image size of 32^3, 64^3 and 128^3.
n_W^{L0}: Number of locations examined at highest resolution level L0.
Tm : Duration of time taken by the process.
ECT : Emission computer tomography (e.g. *HMPAO* brain study)

Table 8.3- Accuracy measurement of the Multi-resolution algorithms for both transformation parameters and residual value (i.e. the minimum obtained distance error, denoted as MDE). *MR* and *SPECT* data (i.e. data type A and B; see section 8.2) were used in this evaluation. (Note that the time is CPU time of *Macro VAX II*).

Ľ	W	n_W	n_P	n_V	$n_{Opr@P}$	$n_{Tot\text{-}opr}$	Tm_{Exp} (Sec)
L3	W1	18711	20	5	137	$5.1*10^7$	102
	W2	57915	20	5	137	$1.58*10^8$	317
	W3	375375	20	5	137	$1.02*10^9$	2057
L2	W4	31185	30	9	177	$1.65*10^8$	331
	W5	225225	30	9	177	$1.19*10^9$	2391
	W6	1156155	30	9	177	$6.14*10^9$	$12*10^3$

Ľ: Initial resolution level for multi-resolution process (e.g. L3 for image size of 32^3 and L2 for size 64^3).
W: Window used for searching transformation parameters.
W1 is 3*3*3 voxels shift and 11*7*9 degrees rotation in x,y and z.
W2 is 3*5*3 voxels shift and 13*9*11 degrees rotation in x,y and z.
W3 is 5*7*5 voxels shift and 15*11*13 degrees rotation in x,y and z.
W4 is 3*5*3 voxels shift and 11*7*9 degrees rotation in x,y and z.
W5 is 5*7*5 voxels shift and 13*9*11 degrees rotation in x,y and z.
W6 is 7*11*7 voxels shift and 15*11*13 degrees rotation in x,y and z.
n_W: Number of locations (transformation parameters) at W.
n_P: Number of points examined at each search location.
n_V: Number of voxels in 3-D space, examined for each point P.
$n_{Opr@P}$: Number of operations (as equivalent integer adds) for each point P (obtained by $87+10n_V$ as expressed in $n_W \times (n_P \times (87+10n_V)$; see equ. 6.12).
$n_{Tot\text{-}opr}$: Total number of operations performed at all points and locations.
Tm_{Exp}: Cost as expected duration of time last by whole process.

Table 8.4- Expected cost calculation based on the sequential and multi-resolution algorithms outlined in diagrams 6.4 and 7.1. Number of sample points and locations (transformation parameters) are derived from the results and discussion set forth in section 8.3.2.2 and 8.3.2.3. All the operations are expressed by the number integer adds which takes about 2 sec/million on a μVAX II machine. Note that no improvement, due to using a variable threshold, was considered in these calculations.

L	T	W	n_W	$n_{p@fit}$	n_{Tot_P}	n_{Tot_V}	n_{Tot_opr}	Tm Sec
L2	2	W4	31185	29	$6.3*10^5$	$3.6*10^6$	$9*10^7$	181
L1	1		3375	47	$9.6*10^4$	$6.75*10^5$	$1.51*10^7$	30
L0	3		1650	95	$6*10^4$	$5.4*10^5$	$5.76*10^6$	11

L: Resolution level (i.e. L2 for size 64^3, L1 for size 128^3 and L0 for size 256^3).
T: Threshold (expressed as g; number of standard deviation from mean).
W: Window used for searching transformation parameters.
n_W: Number of locations (transformation parameters) at W.
$n_{p@fit}$: Number of points examined at match location.
n_{Tot_P}: Total number of points examined at search locations (W).
n_{Tot_V}: Total number of voxels in 3-D space, examined by ray tracing.
n_{Tot_opr}: Total number of equivalent integer add operations in the process.
Tm: Cost as the duration of time last by the process.

Table 8.5- Experimental results showing cost measures and the timing response of the sequential multi-resolution method applied on *MR* to *MR* brain image registration. Note that the process started at the level L2 (i.e. image size of 64^3).

CHAPTER 9
DISPLAY METHODS

9.1- Introduction

Current medical imaging modalities produce 3-D data in a series of adjacent slices which are routinely viewed as a sequence of 2-D grey-scale images. Development of 3-D surface display methods plays an important role in 3-D understanding of anatomical relationships which are required for radiotherapy treatment planning, surgical purpose and diagnosis. Although, a 3-D surface display provides desirable relational information, it does not provide certain types of important information obtained by grey-level changes in 2-D slices. Combining the two kinds of display strategies by either a side-by-side display method or by a superimposition process (i.e. 2-D grey-scale slices combined with 3-D shaded surface images) can be used in many medical applications.

Generally, 3-D rendering techniques (display of visible parts of an object on viewing screen) can be categorised into two approaches. In the first approach, a primarily stage of surface definition is required (Surface rendering methods). This is based on surface positional information regardless of the intensity values assigned to each object volume element. Secondly, a class of techniques are defined, based on voxel intensity, which do not require surface definition in advance to display a surface (volume-rendering methods). Different rendering techniques are due to the various types of object representation and surface definition which were discussed in chapter 4.

The realistic display of a 3-D object on a 2-D surface display system requires depth cues to provide the illusion of the third dimension. Hidden surface removal and shading are essentially required to simulate the surface characteristics, position and orientation of the surface with respect to a viewer. In this chapter various approaches of displaying 2-D slices and 3-D shaded surfaces are discussed with a special interest in superimposition of the registered images from different modalities. Different methods of shading and illusion of 3-D surfaces are also outlined with the aim of obtaining a general and effective 3-D environment for the perception of such data. Various display approaches for the registered and superimposed data are also presented.

Sections 9.2 and 9.3 present an overview of the existing display and shading methods with emphasis on the superiority of some techniques. In a similar respect, volume rendering methods are introduced as more promising display techniques in section 9.4. Various strategies of superimposition and displaying the registered images from different modalities together with the considerations which should be taken in the implementation of these techniques are addressed in section 9.5. Section 9.6 outlines 2-D and 3-D display and superimposition methods used in the current work. The methods of overlaying and combining the correlated images in order to identify the similarities or differences between these images with the aim of obtaining a better clinical interpretation, are also presented in this section. A few of the patients among the total of thirty cases investigated during this project are studied in section 9.7 and their registered images are demonstrated. A summary of this chapter is given in section 9.8.

9.2- Overview of surface rendering techniques

Surface rendering techniques can be divided into two broad categories, *surface-based display* techniques and *binary voxel-based* techniques. Both approaches require binary segmentation of grey-scale data. Geometrical information (e.g. position, depth and distance which are defined as range data) of the surface elements are used to associate depth cues, light intensity and reflection, luminosity, transparency, and shading with the surfaces, and to display them.

In the first category, geometric surface primitives (e.g. closed polygons) can approximate the surface of an object using a number of sample points on 2-D contours extracted from each grey-scale slices. This is usually implemented by a *triangulation* algorithm (Keppel 1975) which reconstructs a polyhedral surface in terms of the planer triangular faces.

In the second approach to surface rendering techniques, a *binary voxel representation* is used to define the surface points, and project them onto a display plane. As discussed in chapter 4, voxels on the surface of a binary or grey scale object can be detected by a 2-D or 3-D *surface tracking* algorithm. The binary surface voxels are then used for shading and display. Either the whole voxels or their faces can be used to represent the surfaces and display them. These methods are outlined in the next two paragraphs.

Voxel faces were used based on an adjacency and connectivity relation (Herman and Liu 1977 & 1979). The voxels are preliminary discriminated in terms of their depths (distance from viewing projection plane) and the information stored in an array called *z-buffer*. For each pixel of the projection plane, the voxels which are closer to the viewer participate in shading and display, and those which are further are removed from the assessment (*hidden surface removal*) (see figure 9.1). For each visible voxel, the visible faces are further specified to define the display parameters and thus surface shading.

A variety of techniques have been used to display visible surfaces (i.e. voxels on the surface) in a binary voxel-based approach. The object volume can be scanned in a *back-to-front* (*BTF*) order (Frieder and Gordon 1985), *front-to-back* (*FTB*) order (Meagher 1982), or along some introduced rays through the volume (*ray tracing*). In *ray tracing* (Tuy H.K., Tuy L.T 1984), the volume is scanned along some introduced rays in a discrete manner. In this display method the work progresses from the screen toward the object, with the assumption that each point on the screen serves as a light source emitting light in parallel rays perpendicular to the screen. Having found the first point of the object intersected by the ray, the appropriate screen point is painted according to the information associated with that intersection point. To examine the object from all possible views, the screen is allowed to move around the object. These methods access all voxels of the scene at display time and the object volume elements are projected on screen as they are encountered. Once a region of the screen has been painted, the voxels projecting on it can be ignored if they are farther than the previous processed voxel from the projection plane (hidden surface removal).

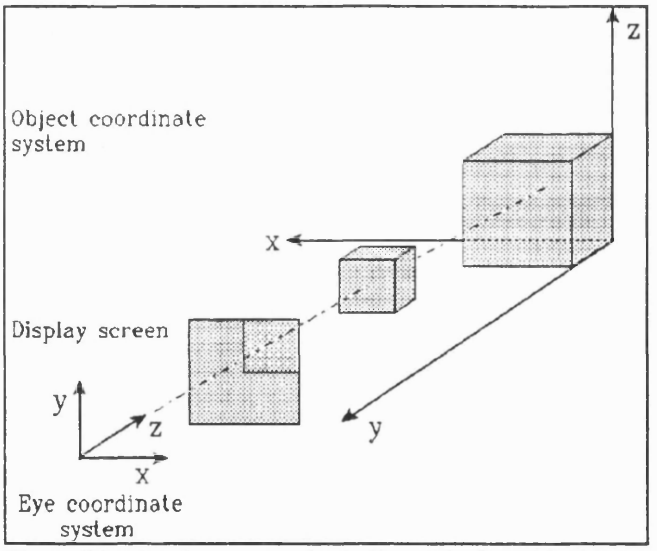


Figure 9.1- showing a general z-buffer method used for hidden surface removal.

Höhne et al 1987 & 1988 and Tiede et al 1988 proposed a display method in which both surface and grey-level data are presented to the viewer simultaneously. In their method, the grey

values in the neighbourhood of a surface voxel are used to measure the relative volume of adjacent tissue types within the voxel. This indicates the direction of the object surface passing through the voxel. An improved 3-D visualization of multiple objects can be obtained by combining the context of grey-scale information (e.g. cut plane, windows, etc) as well as the clipped 3-D shaded surfaces. The grey-scale variation behind a surface can also be visualized by this interactive technique using a clipped surface or transparent projection.

9.3- Overview of shading methods

Shading is the final part of visualization process and provides the viewer with understanding of the surface characteristic, position and orientation (e.g. object normal). Using some information (e.g. orientation and slopes) associated with a 3-D object leads to a vast varieties of shading methods known as *object-space shading*. *Image-space shading*, alternatively, uses only information available in a 2-D image after applying coordinate transformation and hidden surface removal. The different shading methods discussed in the literatures (Foley 1980, Herman and Udupa 1981, Gordon and Reynolds 1985, Chen et al 1985, Bright and Laflin 1986, Höhne and Bernstein 1986) estimate and use the surface orientation in different ways. The general formula for calculating shading values (intensity I_P of point **P**) used by most of the methods is

$$I_P = ((I_{max}-I_a)/D)\ (D-d)\ \mathrm{Cos}(\theta)+I_a \qquad \text{Equ. 9.1}$$

where d is the distance of point **P** from the viewer (i.e. projecting plane or display screen), θ is the angle between the direction of incident light (or emitting light from a surface) and the object normal at **P**, and $\mathrm{Cos}(\theta)$ specifies the characteristic of a diffuse reflecting surface (see figure 9.2). It is assumed that the light source is located at infinity and light rays pass through the projecting plane (display screen). I_{max} is the maximum grey level assigned to the surface (e.g. 255), I_a is background light (intensity) and D is the depth of the object (the distance between the nearest to furthest object point from the viewer). Below, some of the widely used methods based on the above general formula are given.

9.3.1- Object-space shading

Object-space shading methods use the object surface information before projection on a 2-D image plane. Extra data (e.g. face orientation, neighbour face normal, and vertex normal) are required to be stored with the object surface representation. Some of the commonly used object space shading methods are outlined in the next few subsections.

Distance-only shading; is the simplest method of shading which ignores the surface orientation. In this method, the shading is calculated with respect to the centre of each visible voxel (or voxel face as defined by Herman 1979) by setting Cos(θ) to the constant 1 in the above formula. The resulting surface images are too smooth to give true 3-D feeling. Moreover, edges and small structural features are not noticeable by this method. However, *distance-only shading* has extensive use in medical imaging due to the ease of its implementation.

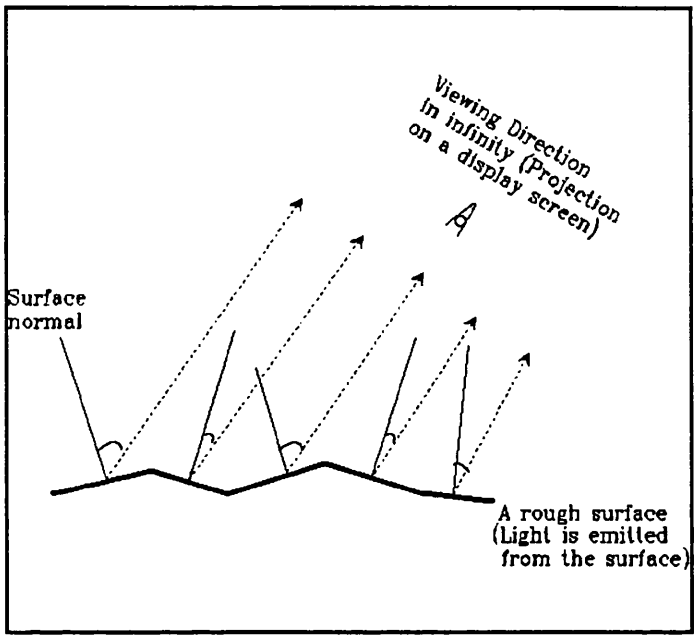


Figure 9.2- Showing the direction of emitted lights (i.e. viewing direction) from a diffuse reflecting surface.

Constant shading; benefits the estimation of surface normal vector for calculation of shading value. It is mainly used for surfaces approximated by planar polygons. A constant intensity is assigned to all visible points on a face (planar facet) of a boundary surface based on the facet normal vector. This normal can be express as

$$\text{Cos}(\theta)=\mathbf{N}\cdot\mathbf{L}. \qquad \text{Equ. 9.2}$$

where **N** is the unit vector normal to the surface (at each planar facet) and **L** is unit vector along the incident light (note that only diffuse reflection is considered where the surface scatters light equally in all directions). **L** can be simplified by assuming an orthogonal incident light in perpendicular direction to projection plane (parallel to z axis), and thus the vector **L** is [0 0 1].

Expression $\mathbf{N}\cdot\mathbf{L}$ is dot product of the two normalized (unit) vector. The angular dependency due to using $\text{Cos}(\theta)$ may create sharp intensity difference for two adjacent faces. The effects of distance and angular dependency can be weighted by

$$(\text{Cos}(\theta)/n)^p \equiv ((\mathbf{N}\cdot\mathbf{L})/n)^p \qquad \text{Equ. 9.3}$$

where n and p are two empirically determined parameters (e.g. n=2, p=0.6 as found experimentally). However, the angular dependency is too high to produce a useful smooth looking image in binary voxel based methods where only six possible orientations exist (6 voxel faces).

Gouraud & Phong shading; by which a smoother surface and better impression of a curved surface can be obtained (Gouraud 1971). The normal vector at each vertex (where polygons or voxels meet) is calculated and an intensity is firstly assigned to the vertices. The intensity assigned to each point of the surface is obtained by interpolation between the adjacent bounding vertices. The possible discontinuity, produced by the *Gouraud* method, between the adjacent polygons can be resolved by the *Phong* method (Phong 1975). He interpolated the normal vector component itself across the polygon surface by first finding the *normal vector* of each edge joining the vertices and then interpolating for the normal vector of each point of the polygon. In *binary voxel based display* methods (e.g. cuberille representation), the normal vector of each vertex is calculated and stored when each voxel face is processed.

Contextual shading; is a visually similar method to Phong shading, introduced by Herman and Udupa 1981 and Chen et al 1985. The angular dependency problem mentioned in constant shading arises if the angle is small for a particular face, while it is not so small for an adjacent face (a face which share an edge with it). The problem can be overcome by using relative orientations of the adjacent faces. Herman & Udupa 1981 suggested the *contextual shading* method which was revised by Chen et al 1985. They assign a shade to points on a voxel face depending on the orientation (normal vector) of the face itself and four adjacent faces.

The major disadvantage of all *object space shading* methods is the necessity to recalculate the normal vector components each time the object is modified. This is very time consuming when the normals of neighbouring faces are required.

9.3.2- Image-space shading

Image-based shading uses a pre-defined image obtained by coordinate transformation, hidden-surface removal and viewer-to-object distance (z-buffer) measurements. No extra data need to be stored with the surface representation and the method is very easy to be implemented. Two well known image-space shading methods are discussed in the following subsections.

Gradient shading; in which the input is an image having distance (depth or z-buffer) information as its pixel values. The representation of such an image is in the form of I=z(x,y), where z is the distance information at any point (x,y). The surface normal (or the angle θ) at each point can be estimated by the gradient vector ∇z at that point. As proposed by Gordon and Reynolds 1985, the normal can be obtained as the vector

$$\nabla z=(\partial z/\partial x,\partial z/\partial y,1), \qquad \text{Equ. 9.4}$$

where the derivatives $\partial z/\partial x$ and $\partial z/\partial y$ can be estimated from the forward difference δ_f and backward difference δ_b, These differences can be expressed, in respect of the x direction, as

$$\delta_f = z_{i+1,j}-z_{i,j}, \qquad \delta_b = z_{i,j}-z_{i-1,j}, \qquad \text{Equ. 9.5}$$

where $z_{i,j}=z(x=i,y=j)$ for i=1,2,...,N and j=1,2,...,N (N*N is the size of image). The weighted average of the forward and backward differences has been suggested (Gordon & Reynolds 1985) for each derivative such as

$$\partial z/\partial x=(w_f\delta_f + w_b\delta_b)/(w_f+w_b), \qquad \text{Equ. 9.6}$$

where w_f and w_b are positive weights which depend on $|\delta_f|$ and $|\delta_b|$, respectively. Small differences are given large weights and vice versa. ($\partial z/\partial y$ is obtained by a similar expression).

Grey-level gradient shading; can represent the fine features (e.g. inclinations) of surfaces which may be lost in the previous shading methods due to the coarse quantization of the surface angles. This occurs in segmentation process even where the voxel structure is used to represent the object boundary. The technique proposed by Höhne and Bernstein 1986 tries to make use

of the partial volume effect especially when two adjacent objects are at different distinct grey level ranges. Since the grey value of a voxel represents the mean density of all the structures within that voxel, the relative change of grey value from one voxel to its neighbour can be considered to represent the surface angle (normal) within the voxel. The intensity (shade) I(i,j) assigned to each pixel of image plane was computed by Höhne 1986 as

$$I(i,j)=A\times\text{Cos}[(f(i,j+1,k)-f(i,j-1,k))/B] \qquad \text{Equ. 9.7}$$

where f(i,j,k(i,j)) is the grey value at the depth k(i,j). The cosine function is used to achieve a diffuse reflection. *A* and *B* are scale-factors depending on the grey-level range of the original images (*A*=255 and *B*=100 for *CT* images, as suggested by Höhne 1986).

9.4- Volume rendering techniques

To overcome the incorrect classification of the small and fine features, which arise from the techniques based on a binary classification decision, the surfaces can be displayed using the original volume data. A variant of voxel based techniques has been developed by Drebin et al 1988 and Levoy 1989, in which the surface is allowed to pass partially through a voxel. Each voxel of a 3-D object is assigned a *colour* (i.e. shading value) and a *partial opacity* (i.e. estimated occupancy fractions for each of the set of materials presented in the voxel). Images are obtained by projecting all the voxels along a line to a similar point on the picture plane and blending together the resulting coloured semi-transparent voxels. This is done by tracing parallel rays (from the viewer) into a data volume in a desired direction.

Using one of the *object-based shading* method (e.g. *Phong*), the colour values are assigned to the voxels based on the grey-scale values of them and their neighbourhood. The depth cuing component of the *shading model* (see equations 9.1 and 9.3) should be scaled to allow the contribution of all the voxels along a ray. Two constants k_1 and k_2 used in the following equation were defined by the previous researchers (e.g. Levoy 1989) for a linear approximation of the depth-cuing.

$$\mathbf{C(i)} = [1/(k_1+k_2\mathbf{d(i)})][\mathbf{C_s N(i).L_s}+\mathbf{C_a}], \qquad \text{Equ. 9.8}$$

where the colour $C(i)$ is indicated for each voxel indexed by a vector $i=(i,j,k)$ $(i,j,k=1,...,N$ and $N*N$ is the size of the projection image), $d(i)$ is perpendicular distance from voxel i to the observer, C_s is the colour of the light source (i.e. the maximum gray level in a medical image approach), and C_a is the colour of ambient light source. L_s is normalized vector in direction of light source (e.g. the perpendicular direction to viewing screen). Surface normal $N(i)$ at voxel i is obtained by gradient vector $\nabla f(i)$ as

$$\nabla f(i)/|\nabla f(i)| \quad \text{for vector } i=(i,j,k). \qquad \text{Equ. 9.9}$$

The gradient $\nabla f(i)$ of grey-scale image $f(i,j,k)$ can be approximated using

$$\begin{aligned}\nabla f(i)=\nabla f(i,j,k)=[&1/2(f(i+1,j,k)-f(i-1,j,k)),\\ &1/2(f(i,j+1,k)f(i,j-1,k)),\\ &1/2(f(i,j,k+1)f(i,j,k-1))]\end{aligned}$$

Equ. 9.10

An opacity $\alpha(i)$ is assigned to each voxel (having value $f(i)$) in a procedure called '*classification*' based on the tissue types falling within any voxel. The important assumptions are that each type of tissue touches tissues of at most two other types and if the types be ordered by their values, then each type touches only the types which are adjacent to it in the ordering. The classification process begins with assigning an opacity α_v^n to voxels having a selected value f_v (correspond to a particular tissue type) and an opacity α_v^{n+1} to voxel f_v^{n+1} (for $n=1,...,N$; $N\geq 1$ is the number of tissue types). The intermediate voxel values are then converted to intermediate opacities by construction a piecewise linear mapping process (Levoy 1989). This mapping (for the case of two tissue) can be obtained by expression

$$\begin{aligned}\alpha(i)=|\nabla f(i)|\ \{&\alpha_v^{n+1}[(f(i)-f_v^n)/(f_v^{n+1}-f_v^n)] +\\ &\alpha_v^n[(f_v^{n+1}-f(i))/(f_v^{n+1}-f_v^n)]\}\\ &\text{if} \qquad f_v^n\leq f(i)\leq f_v^{n+1}\\ &\text{otherwise} \quad \alpha(i)=0\end{aligned}$$

Equ. 9.11

where the scaling by the gradient $\nabla f(i)$ is used to suppress the opacity of the tissues which are interior and enhance the opacity of their bounding surfaces.

Composing the colour and opacity components at each location of projection image plane

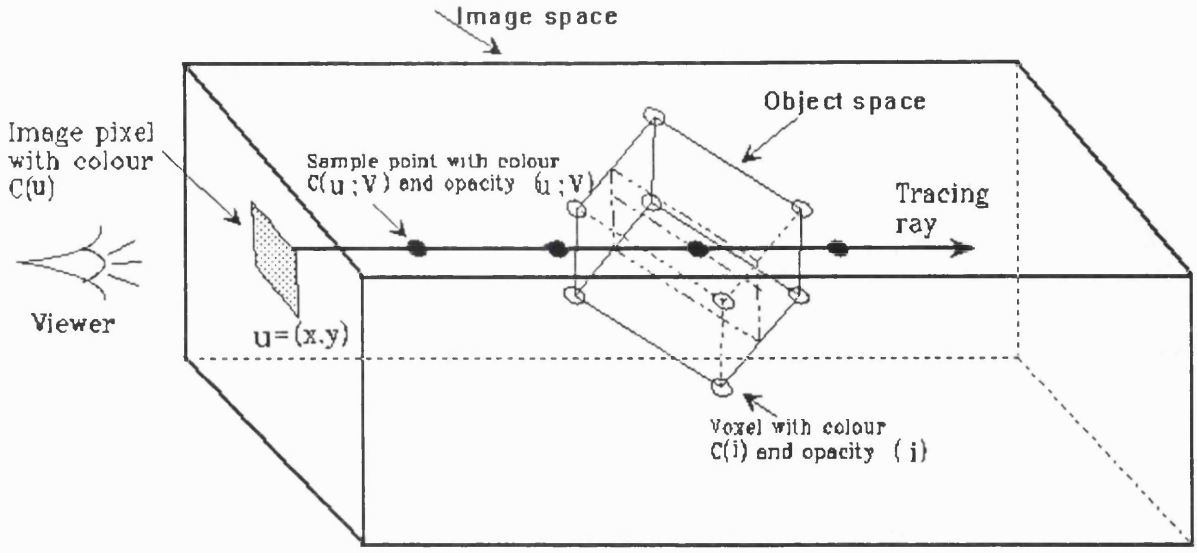


Figure 9.3- Showing the general concept of *volume rendering methods* used to display an object. Both the object space and the image space are represented.

is obtained in a front to back manner by processing the voxels along the ray **u** (see figure 9.3). The colour $\mathbf{C}(\mathbf{u};\mathbf{V})$ and opacity $\alpha(\mathbf{u};\mathbf{V})$ of ray **u** after processing sample voxel **V** inside the data volume is obtained by

$$\hat{\mathbf{C}}(\mathbf{u};\mathbf{V}) = \hat{\mathbf{C}}(\mathbf{u};\mathbf{V}) + \hat{\mathbf{C}}(\mathbf{u};\mathbf{V}^{\circ})(1-\alpha(\mathbf{u};\mathbf{V})), \text{ and}$$
$$\alpha(\mathbf{u};\mathbf{V})=\alpha(\mathbf{u};\mathbf{V})+\alpha(\mathbf{u};\mathbf{V}^{\circ})(1-\alpha(\mathbf{u};\mathbf{V})),$$

Equ. 9.12

where $\mathbf{V}^{\circ}$ is the sample voxel in process, $\hat{\mathbf{C}}(\mathbf{u};\mathbf{V})=\mathbf{C}(\mathbf{u};\mathbf{V})\alpha(\mathbf{u};\mathbf{V})$, and $\hat{\mathbf{C}}(\mathbf{u};\mathbf{V}^{\circ})=\mathbf{C}(\mathbf{u};\mathbf{V}^{\circ})\alpha(\mathbf{u};\mathbf{V}^{\circ})$. After processing all the voxel samples along a ray **u** (see figure 9.3), the colour $\mathbf{C}(\mathbf{u})$ (intensity of the pixel **P** on image plane) is obtained by expression

$$\mathbf{C}(\mathbf{u})=\hat{\mathbf{C}}(\mathbf{u};\mathbf{V})/\alpha(\mathbf{u};\mathbf{V}) \qquad \text{Equ. 9.13}$$

9.5- Superimposition & display of registered images

Comparison of images performed by different imaging modalities or taken at different time instant is due to a two-stage process consisting *registration* and *display*. Registration as discussed in chapter 3, provides only a set of transformation parameters which relates the data

in the coordinate system of one study to the data in other study. Employing the relevant information and displaying the aligned or superimposed images, in the second stage, allows for spotting similarities and differences in clinical features between the two studies and thus improves diagnosis. This benefit is due to the capability to reslice the image data from one study along the planes of the other study. In this section the different approaches of displaying the relevant information are outlined.

The original 2-D images (from different modalities) are collected under different imaging parameters and system characteristics. Superimposition and even side-by-side display of the images should be done in a similar matrix size, zooming, and scaling environment, and nearly identical grey level range. The images may also require geometric and grey-scale transformation (before visualisation) based on the characteristics of the display systems.

In order to display the relevant information between the two studies in a clinically useful form (such as aligned planes), the following requirements are essential. Firstly, three dimensional grey level data are required in order to allow access to the relevant features of both studies. *Image processing* is also necessary to improve the visual quality of the original and resliced grey-scale images. The desirable orientation of images should be decided by an operator (viewer) based on the characteristics and limitations of the display system. The decision on the type of display method (e.g. 2-D or 3-D), data alignment (e.g. side-by-side display or superimposition), and the type of shading used in the visualization are also very important in clinical applications. These requirements are discussed in the next few sub-sections.

1) The routine medical imaging (*MRI*, *SPECT*, *CT*) provides adjacent (e.g. contiguous) 2-D grey-scale slices of a few millimetres thickness (e.g. 6 to 10). In some cases (e.g. *MRI*), a gap is introduced between two adjacent slices. In order to access all the feature in one study corresponding to the other, 3-D information is required. *Grey level interpolation* is the common key to obtained the grey level of the structures within the gap region. For thick slices which contain fine features, some uncertainty might be presented in the definition of the features in the interpolated region between the two slices. The problem is much worse when interpolation is required between routine *MRI* slices having non cubical voxels. The artifacts shown in figure 9.4 are due to this interpolation deficiency. For this reason, it is desirable to keep the original slices of a non-cubical image intact, and to resample through the images having cubical voxels

(reslicing process). The problem can also be overcome by using a *shape-based interpolation* (see section 4.4.4) when binary images need to be interpolated (i.e. for 3-D external surface of an object).

2) The original images formed by medical imaging systems are influenced by the system's physical attributes and thus contain various image characteristics such as noise level, contrast and brightness. A digitally processed image may also occupy a grey-scale range which is different from the range of the other corresponding images (e.g. registered images). The images should be processed to improve their visual appearance and thus enhance their clinical interpretation. In this respect, different methods of image enhancement (e.g. digital image filtering) might be used. However, there is no unique effort to improve the quality of an image with regard to different observers. Moreover, different applications might imply different types of image enhancement.

3) As mentioned above, a set of transformation parameters is the result of a registration process which is used as an input to the desired image reformation algorithm and display system. This coordinate transformation is given in reference to either a world coordinate system (reference coordinate system recognized by both modalities), or the coordinate system of one of the modalities (e.g. the one which produces the objective data). The aim of the superimposition process is usually to display both images of an object in one common coordinate system. However, different viewing directions might be required for clinical assessment. The diagram shown in figure 9.3 demonstrates the relation of two different coordinate systems, for example object coordinate and image coordinate systems (screen coordinate). In order to calculate the position of an image point (being originally on any coordinate system; e.g. the world system) on a display screen, the point must be transformed from the world coordinate system $[x_w, y_w, z_w]$ into the eye or screen coordinate system $[x_s, y_s, z_s]$.

$$[x_s, y_s, z_s]=[x_w, y_w, z_w][\mathbf{TR}] \qquad \text{Equ. 9.14}$$

The transformation matrix **[TR]** may be built up from several rotations and translations, determined by a desired viewing direction. Three viewing rotational angles θ, ϕ, ψ are defined to rotate the world coordinate system around x, y and z axis of screen coordinate system, respectively, using the matrix equations discussed in section 2.4.1. For displaying 2-D slices in

an orthogonal viewing direction, only a -90 or +90 degrees rotation is required to provide the desirable direction. However, it is sometimes of concern in visualization to represent three dimensional coordinates of a 3-D shaded object at some oblique angle to the viewer so that a better 3-D feeling and perception be obtained. The images shown in the top left corner of figures 9.4c, 9.8, and 9.9 were obtained under 20 degrees rotation around z axis. The translation of an image inside the screen coordinate system (e.g. positioning to the coordinate centre) may also be required.

4) In terms of display types, the most straightforward method is to display the 2-D images from different modalities side-by-side on the screen. The original grey-scale image of one modality (i.e. destination image) can be displayed together with the resampled (resliced) slice through the data set of the other study (see e.g. figures 9.6, 9.7). The decision of the choice of destination image is influenced by the clinical use and the type of data sets. In this respect, attempts are made to reslice through the data with cubic voxels and those having a better 3-D resolution.

9.6- Methods

For a 3-D display, the external surfaces represented as discrete binary voxels were used and shaded. The *distance-only shading*, was used to display the fitting of two surfaces during the registration process. The surface normals were also defined in an *image-based approach*, to shade the 3-D surfaces and display them more realistically, during the superimposition process. To improve visualization, the *volume rendering process* was also used for displaying the skin surface and the surface of some internal structures of the head (e.g. brain, tumour, ventricles). As discussed in section 9.4, this process does not need binary segmentation of images, and displays surfaces directly from the sampled scalar field of a 3-D volume. However, its use is limited due to its sensitivity to the presence of artifact and the critical need for a proper opacity assignment (to the tissues) which alter the quality of the image, substantially.

The *3-D shaded display* of external surfaces (skin, brain or ventricles) as shown in figure 9.4b, helps for understanding 3-D environment during the surface fitting process. 3-D shaded surfaces can also be used to show the 3-D relation of various displayed features (e.g. tumour and

ventricles) in the superimposition process. The lines passing through the data on a displayed surface (see top left images of figures 9.6 and 9.7) show the position of destination slice on the 3-D data volume.

In order to produce the 2-D registered slices, the reslicing was performed by a *trilinear grey-scale interpolation* process. The coordinates of the destination slice are transformed by the registration parameters in order to obtain the coordinates of the resliced image. In this respect, the 3-D coordinates of four corner points of a destination sliced image can be used to define the coordinates of four points of the new resliced image plane. The data can then be sampled on the plane defined by these points.

Typical *side-by-side display* of the registered slices (images) is shown in figure 9.5 and in the lower half of figures 9.6-9.10. Three techniques were used to facilitate assessment of the relative structures (information) between the two side-by-side images.

1) The *regions of interest* (*ROI*; e.g. tumour) were selected manually and their contours were superimposed on the 2-D grey-scale images of the other study. The contours of some specific structures (e.g. brain ventricles in *MRI*) were also superimposed on the 2-D images. The top right images of figures 9.6-9.10, demonstrate this visualization technique.

2) As shown in the bottom lower half of figures 9.6 and 9.7a, a *linked cursor* can be used which moves by a mouse digitizer orthogonally on the screen by an operator and shows the corresponding structures between the two images.

3) The *ROI* from one modality can alternatively be mapped on the context of the other image and displayed in different colour coding (scale).

Due to the difference in the grey level ranges of the registered images, *contrast enhancement* was used during display process. Interactive linear image scaling was widely employed to process all the images displayed simultaneously on a screen. Filtering of images (e.g. *median filtering*) was performed where noise degraded the image quality. Images might need to be blurred (i.e. by a low-pass filter) to obtain smoother edges and facilitate the segmentation and contour detection used for the superimposition. Using *colour coding* and colour

scaling by the use of different pre-defined look-up tables allows enhancement of the visualisation and improves diagnostic aspects. All the image processing techniques are performed under the user interaction during the display process while the original grey-scale images are viewed by the observer (e.g. clinician).

9.7- Subjects and results

The registration process was performed on the images of thirty different patients. Most of the patient were children of age less than sixteen years old. The cases were selected from the patients whose images showed a demonstrated abnormality in one or both modalities for which registration was performed. The technique was applied for correlating *MR* with *MR*, *PET* with *MR*, *CT* with *MR*, and *SPECT* with *MR* images. All the scans on each patient were done within a maximum period of two months. In this section a few cases which mostly have a major abnormality are demonstrated.

The images shown in figure 9.4 are taken from a 28-years-old male volunteer, with no abnormality in the *MR* images. As described in section 8.2 (data type C), two sets of images were obtained by setting different system parameters to obtain misregistered images for experimental assessments (see figure 9.4a). Figure 9.4b shows a typical display system corresponding to the images of the subject patient, visualized during the registration process. The registered *T1-weighted MR* images displayed side-by-side in figure 9.4c indicate, visually, that a good correlation has been achieved.

The *MR-SPECT* correlation shown in figures 9.5 was also obtained by the performance of the proposed fitting process on axial *MR* slices and *HMPAO SPECT* images. Four external landmarks were also used (see figure 9.5a) to verify the surface fitting process by the results obtained by the landmarks registration. Both the experimental results and the visual judgement of the displayed correlated images confirmed the reliability of the process.

As illustrated in figure 9.6 metabolic activity level (*O_2-16 FDG; fluoro-deoxy-D-glucose*) as imaged by *PET* is shown, correlated with soft tissue anatomy from axial *T2-weighted MRI* scans. In figure 9.6a, the correlated *MR* slice (bottom right corner) was obtained by resampling

the original grey scale *MR* data oriented along the scan planes of the *PET* study. The resampled *PET* slices shown in figure 9.6b (top and bottom right corners) were taken along an original scan plane of the *MR* data. As demonstrated in these images, the ventricles in the *PET* scan nearly coincide with the ventricle areas in *MR* scan. The contour of the anatomy revealed by the *MR* image is seen in the context of the functional *PET* image by this type of combined display. The linked cursor also demonstrates the correlated structures in the two types of images.

Figures 9.7 show the *MR-SPECT* correlated image of a 7-month old female suffering from seizures. Axial double echo *STIR* sequences were performed to obtain *MR* images which were correlated with *HMPAO SPECT* data. The *MR* data shows high signal in the white matter alongside the body of the right lateral ventricle. There is also a diffuse area of the increased signal in the right temporo-parietal lobe involving the grey matter. As shown in these figures, the low uptake area in the *SPECT* images corresponds to the abnormal area of *MR*.

The images shown in figure 9.8 were obtained from a 4-year male patient with an extensive abnormality of left frontal and left fronto-parietal regions. The axial *T2-weighed MR* images show extensive abnormalities in white and grey matter of these regions, as well as the involvement of the posterior limb of the left internal capsule, with a further lesion in the right striatum and parietal occipital region. As shown in the correlated *SPECT* image, multiple abnormal signal areas are observed, consistent with ischaemic lesions. This suggests that the nature of the multiple infarcts are in the vascular distribution of the brain. In these images, the contour of some lesion regions which are well visualized in *MR* images has been placed in the context of *SPECT* scans acquired under uptake of *HMPAO*.

The *CT-MR* correlation shown in figure 9.9 belongs to the same patient whose images displayed in figure 9.8. Overlaying the features (e.g. contours) of *MR* on the *CT* images, as shown in the top right corner of this figure, demonstrates the relationship between the skull and detailed brain structures in a slice image. The bone and soft tissue morphology demonstrated by the superimposition of *MR* and *CT* images provides useful positional information for surgical purposes.

The results obtained by the registration of three modalities (*MR*, *CT* and *SPECT*) are

presented in figure 9.10. As shown in this figure, three correlated images of the same patient can be displayed on one viewing screen together with a superimposition image showing the correlation between these images more accurately.

9.8- Summary

The problem of image superimposition (overlaying and combining the correlated images) and display were presented and discussed in this chapter. Some of the commonly used display techniques were outlined and their superiorities in displaying clinical images were discussed. These techniques involve two different image shading categories known as object space and image space shading methods.

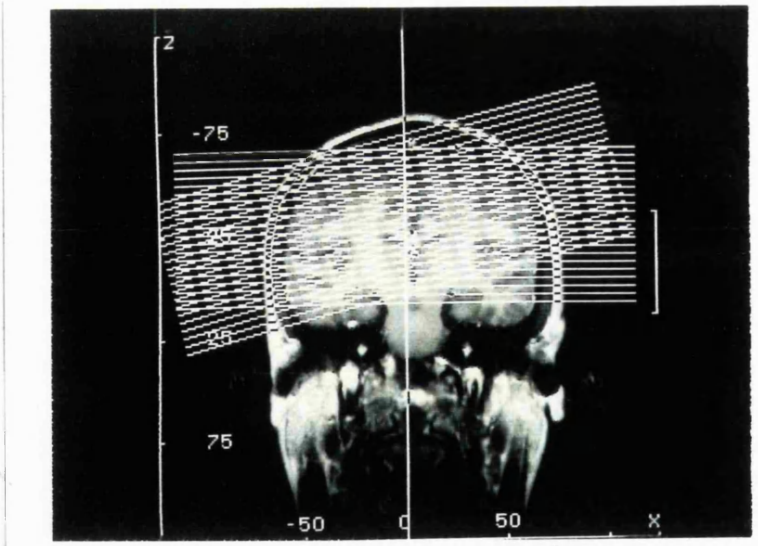
The objects (e.g. 3-D surfaces) displayed in the current work were based on *voxel-based representation methods*. The surface normals were defined in an *image-based approach*, so as to shade the 3-D surfaces and display them realistically during superimposition process. To improve visualization, the *volume rendering process* was also used for displaying the skin surface and the surface of some internal structures of the head (e.g. brain, tumour, ventricles).

Four methods for combining and displaying the correlated (registered) images were employed as described in section 9.6. The correlated 2-D slices from one study were created by resampling their original grey scale data oriented along the scan planes of the other study. The *side-by-side display* of these correlated images was then employed to provide corresponding clinical information between the two studies. The *contours of regions of interest* (e.g. ventricles or tumours) were superimposed on 2-D grey-scale images of the other study. A *linked cursor* was used, moved manually by a mouse digitizer on the screen, showing the corresponding structures between the two images. A *colour-coded display* was also suggested in which different colour ranges (e.g. look-up tables) are assigned to the structures (e.g. tissues) of each image. Superimposition of these structures then has a distinct display of the two overlaying images.

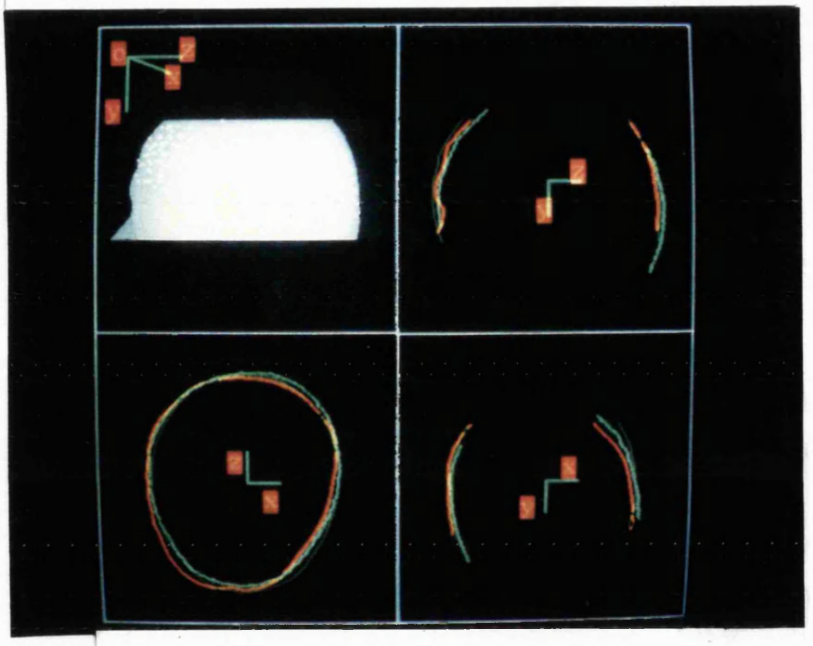
Due to the difference in grey level ranges of the registered images, different image processing techniques (contrast enhancement) were applied during display process. An interactive image processing scheme was employed to display an image desirably and improve the clinical

interpretation.

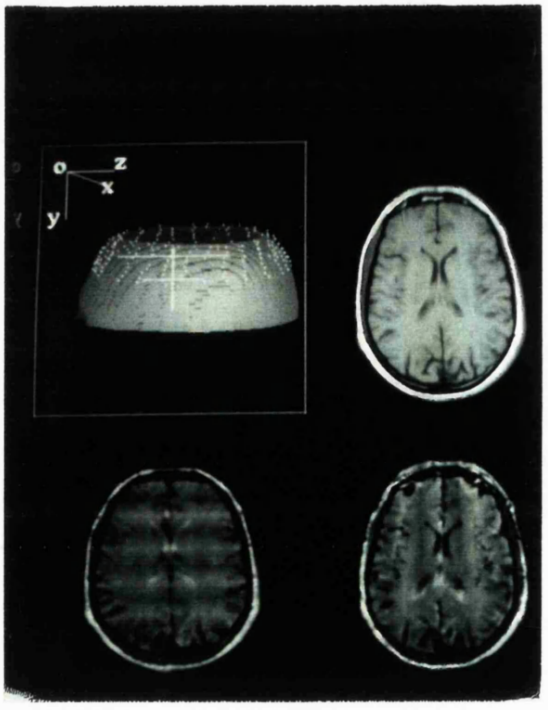
Thirty patients who had been routinely imaged in a clinical department were studied by the proposed registration processes. A few cases among them having abnormal brain regions were discussed and their images demonstrated in section 9.7. The fidelity of the registration process and data alignment was judged by the clinicians and radiologists in the department where the work carried out.



(a)

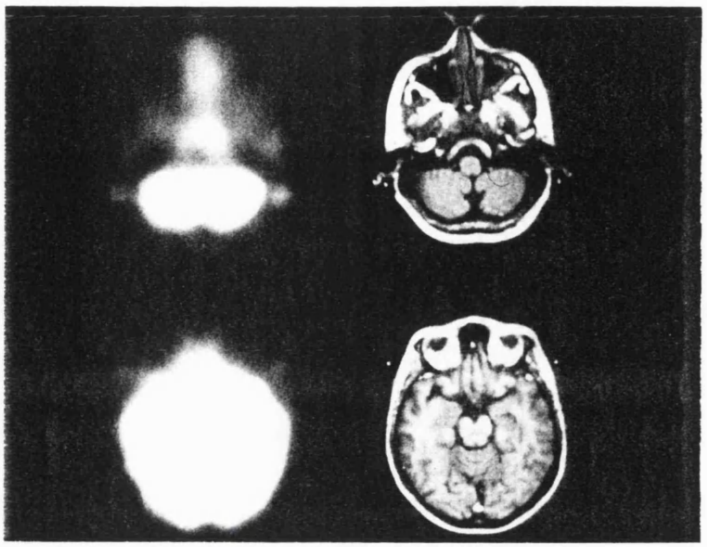


(b)

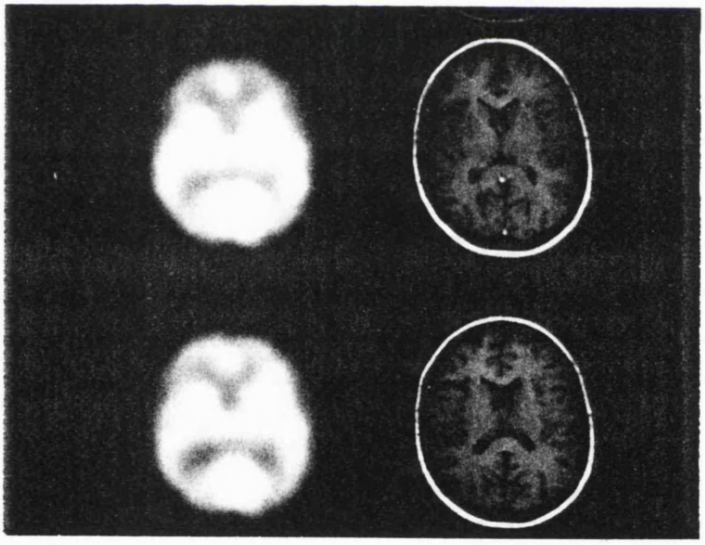


(c)

Figure 9.4- Shows *MR-MR* registration of a volunteer subject. Two well-defined misregistered images (data type C; see section 8.2) were used. **a)** The original known misregistered (i.e. rotated around the z direction) MR images. **b)** A type of the viewing screen (the surface contours at 3 main orthogonal directions). **c)** The registered images displayed as a 2-D side-by-side method. Note that the artifacts shown on the two bottom images are due to the reslicing through *MRI* data set.

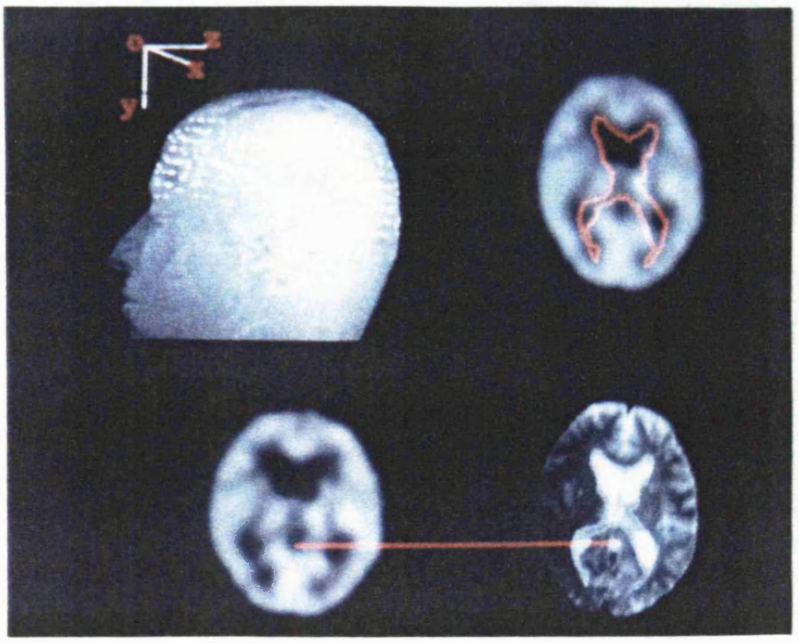


(a)

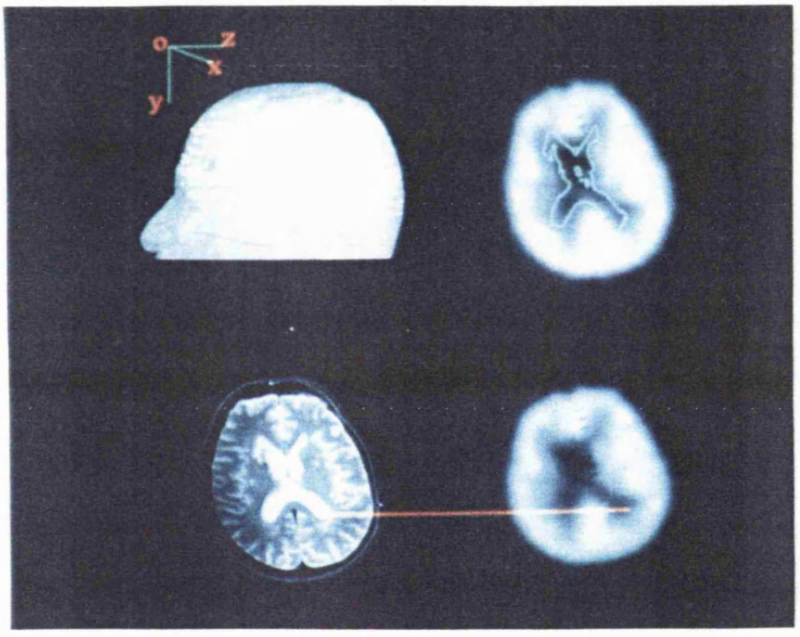


(b)

Figure 9.5- Shows *MR-SPECT* registered images of a patient brain data set. **a)** shows four landmarks (two frontal and two on mastoid process) used to be registered independently from the surface fitting, in order to verify the proposed method. **b)** shows two registered image planes displayed side-by-side.

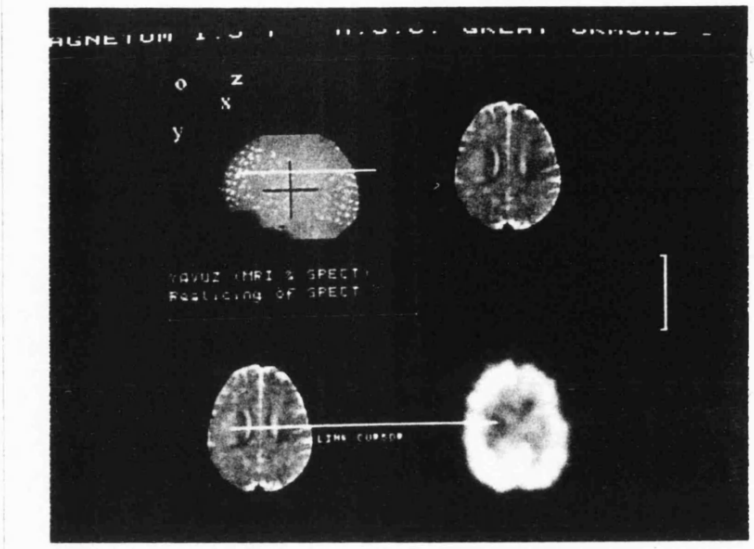


(a)

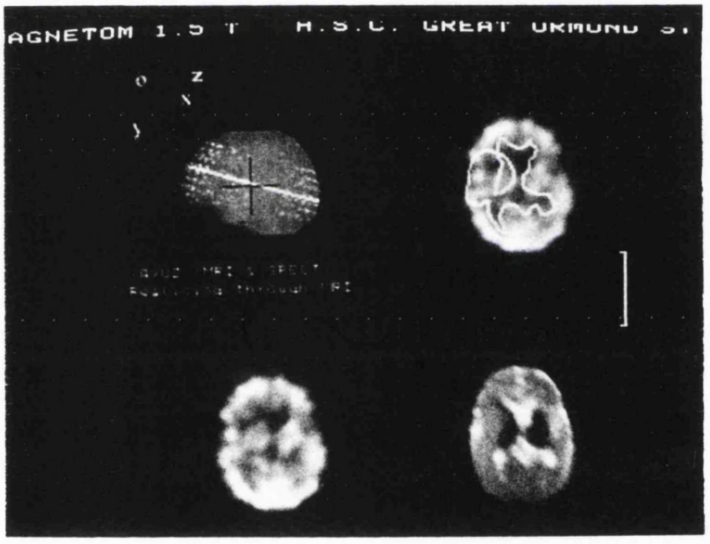


(b)

Figure 9.6- Shows *MR-PET* correlated images (see section 9.7 for the explanation). Both the contour overlay (top right images) and linked cursor (bottom images) are shown on a side-by-side image display. **a)** The reslicing was employed on *MRI* data set from which the contours of brain ventricles were obtained. **b)** The reslicing was performed on *SPECT* data set, but the superimposed contours were obtained from the *MRI* slice (the top right image) .



(a)



(b)

Figure 9.7- Shows *MR-SPECT* registration of a 7-month old female. **a)** Shows the reslicing through *SPECT* data set (bottom right image) with a linked cursor display. The top right image shows the brain surface and sample points used for the registration as well as a solid line showing the position of the displayed image plane. **b)** Shows the reslicing through *MRI* data set (bottom right image) with a contour superimposition of the registered *MRI* brain ventricles on the original *SPECT* image slice (top right image).

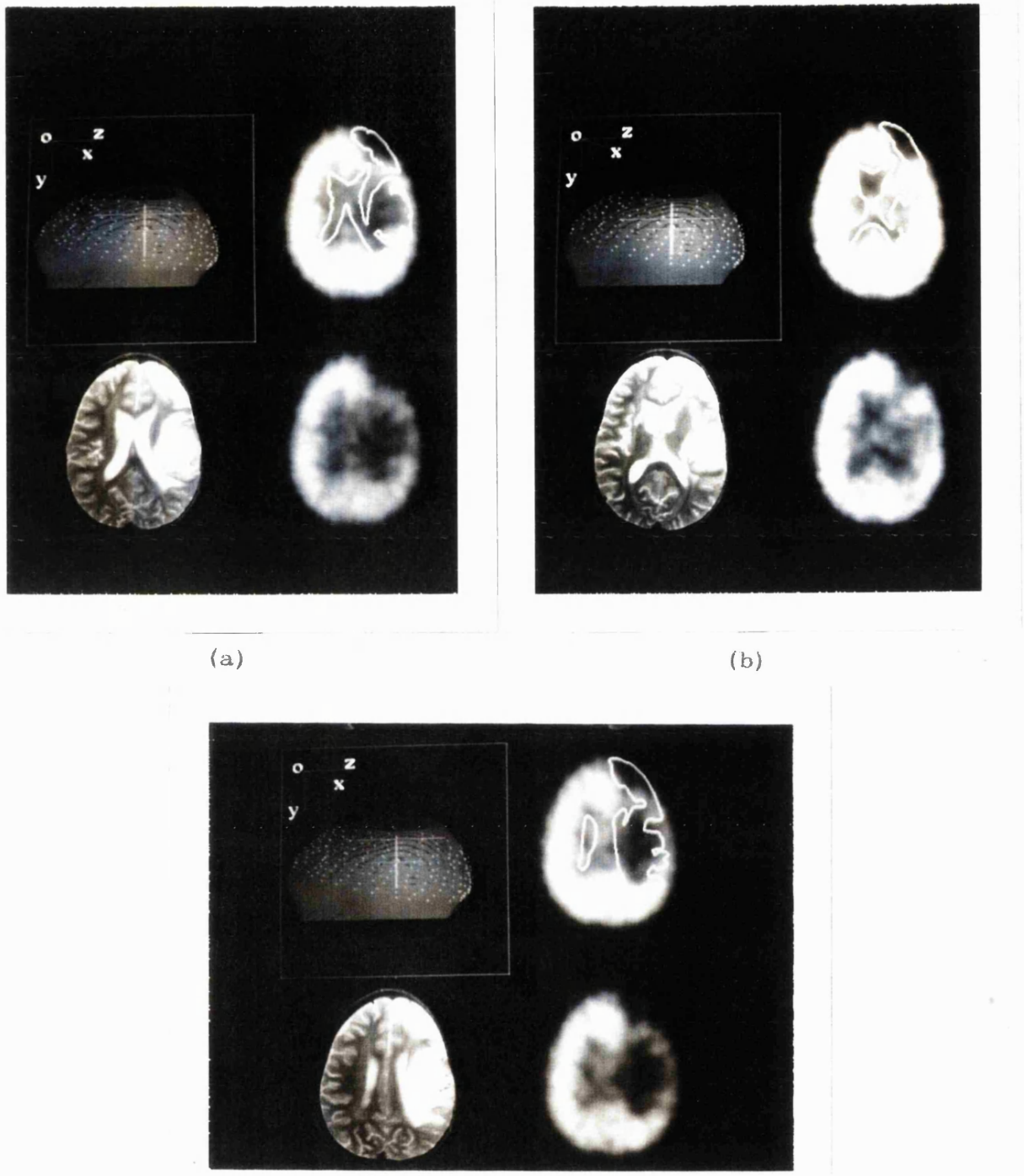
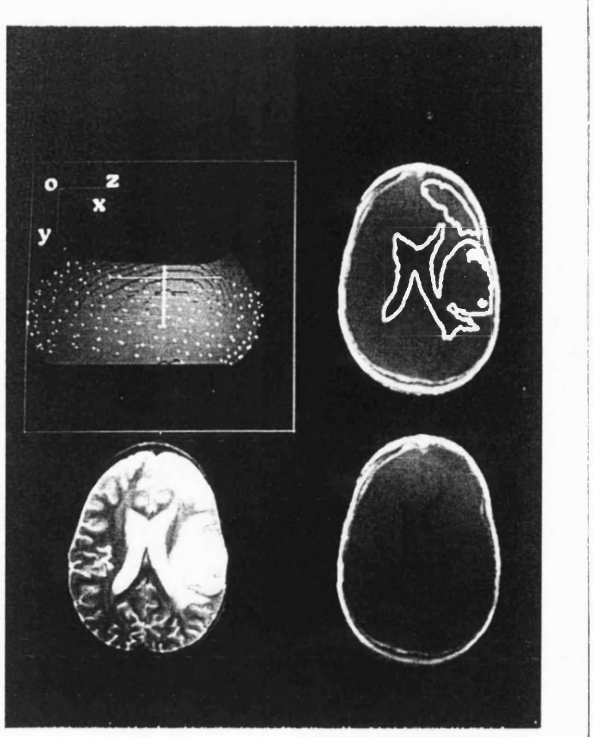
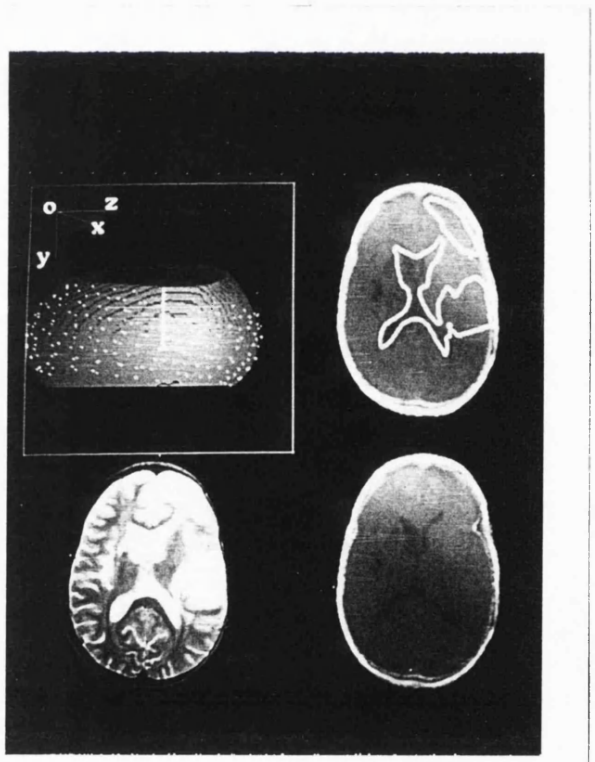


Figure 9.8- Shows *MR-SPECT* registered images of a patient having a big lesion in frontal and parietal lobes. In all images, the reslicing were applied on *SPECT* data set. The original *MR* slice images are presented.



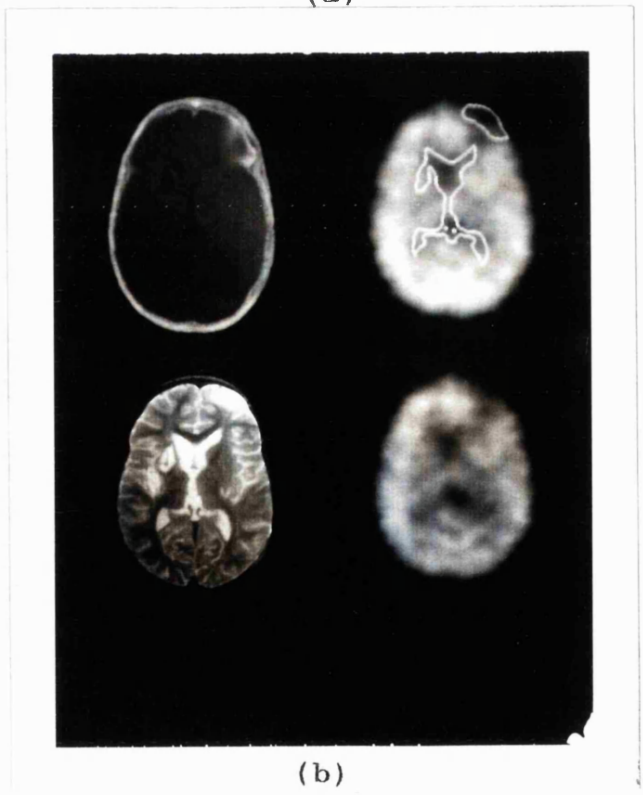
(a)



(b)

Figure 9.9- Shows *CT-MR* registration of a patient brain data set. The side-by-side display is shown in the bottom half, and the superimposed contour-image display in the top right corner.

(a)



(b)

Figure 9.10- Shows *CT-MR-SPECT* registration obtained by the proposed technique (Multi-resolution surface fitting). Three registered images (*CT*, *MR* and *SPECT*) are shown as a side-by-side display, and as a superimposed contour-image display of *MRI-SPECT* data set (the top right corner).

CHAPTER 10
CONCLUSIONS

10.1- Research objectives

The main objective of the work presented in this thesis is to introduce a suitable technique for correlation of tomographic brain images, integrating and combining them by a proper display method. Several attempts have been made to develop algorithmic and computational techniques which allow a reasonable confidence (in terms of accuracy and timing) in the correlation of brain images.

The preliminary objective was to review and verify all available registration techniques and assess those techniques which are suitable for correlating the complex medical images. A method for registration of multiple medical images which does not require any extra information obtained from, for example, stereotactic frames, external markers, or even internal anatomical landmarks, was selected. This is an edge-based method known as surface fitting which discards all the internal structures of the object (e.g. brain).

Due to the nature of the selected registration algorithm, a number of related subjects such as surface reconstruction and minimization have been explored. The obstacles and potential limitations of the surface fitting process were studied. The strengths and drawbacks of this method and its relevant processes were also assessed. Initially, the surface fitting algorithm was modified and revised in order to overcome the obstacles and problems associated with this matching process. There are several new approaches and modifications of the previously existing algorithms, recommended throughout this project (see section 3.3). Overcoming the problems associated with the deformations of the external surface of the brain (e.g. presence of holes on the brain surface) which may occur during imaging or surface reconstruction, is an example of this objective. Other modifications include; 1) using a more reliable and accurate surface representation (e.g. by the use of 'circularity check' and shape-based interpolation, 2) using a better distance measurement between the two surfaces (e.g. by the use of a 3-D ray tracing method and the variance of distance errors), and 3) implementing a closer alignment between

the registering surfaces, both as an initial process and during the stages of the registration process (e.g. by viewing inspection, manual registration, and frequently updating the position of the registering surfaces and re-sampling the sample points after each promising stage of the matching process, i.e. minimizing transformation). In general, most of modifications in the previous algorithms were implemented in the surface reconstruction approach and some in distance-function evaluation and registration.

Two new matching strategies known as sequential and multi-resolution (multi-grid) techniques were introduced as the key improvements of the fitting process, which are the main achievements of the project. In the first approach, a variable threshold sequence was implemented to reduce the search effort of the global transformation space and, therefore, to make the process usable in routine medical image application. In the second approach, a multi-resolution based technique was introduced in order to reduce the size of the registering images, the size of transformation (search) space, and the size of other algorithm's entities (e.g. sample points, ray tracing), thereby to reduce the computational cost required to converge the process in a global search strategy.

All the above processes were verified using different types of data set. Their accuracy and computational performance were measured using some well-known image models and well-defined clinical images (e.g. phantom images, well-known misregistered images, images of the brain with external markers). Analysis of the computational aspects of the algorithms and programming difficulties to make the process usable in routine clinical applications were also investigated.

10.2- Summary of the research

Different scans of brain contain complementary information which may be useful when correlated and displayed simultaneously. Correlation of structural and functional images is essential in quantitative image analysis and understanding of the origin of some abnormalities. Using structural information provided by morphological images is important in data analysis of functional images. Many brain functions are related to structures of very small size, for which

data correlation is required in assessment of the physiological information of these regions.

The research presented in this thesis is an investigation into the problems of the registration and display of brain images obtained by different imaging modalities. Following the introduction of the research objectives, the research topics and organization of this thesis were presented in chapter 1. Some clinical useful application of the registration and superimposition were also defined. This chapter also demonstrated the imaging and computer systems and facilities employed for developing the work presented in this thesis.

The various widely used registration algorithms were introduced in chapter 2. The applicability of these methods to medical images was also described in this chapter. The advantages and disadvantages of each algorithm were described. A registration process based on the reliability criteria was selected and introduced in this chapter. In this approach, the *edge-based algorithms*, which are based on a *least-square-distance matching*, were suggested for registering of brain images. These algorithm minimise the sum of square-distances between the two surfaces. The minimization is applied over a set of six geometrical transformation parameters (3 shifts and 3 rotations) which indicate how one surface should be transformed in order to match with the other surface.

The general technique of the *surface fitting*, as defined by the other workers, was described in chapter 3. A number of improvements to the general surface fitting algorithm were also suggested in this chapter. One of the suggested key improvements was the use of a *global minimization* algorithm which increases the accuracy of the registration. This algorithm has led to two novel approaches: *sequential* and *multi-grid* registrations, as discussed in two separate chapters (6 and 7).

In chapter 3, the imaging parameters and data types used through this project were also introduced. Most of the data introduced in this chapter was used as well-defined image models for verifying the applicability of the different aspects of the registration process. Based on this data the proposed technique was verified and its results were presented in a separate section. In general, a number of local minima were shown to exist in the distance function between the two surfaces. The sampling strategy (e.g. position and number of sample points) was also shown to influence the behaviour and sensitivity of the distance function. It was shown that the accuracy

depends on the resolution of the data set and thus varies with the type of imaging study. The accuracy was shown to be of the order of half the pixel size of the original grey-scale image.

A survey of surface reconstruction methods, as an essential element of the surface fitting, was presented in chapter 4. In this respect, surface representation, segmentation (2-D edge detection), and surface formation were studied. Potential techniques were selected following the introduction of suitability criteria. The representation capabilities of most of the representation techniques were defined to be poor in satisfying all of these criteria. However a voxel-based approach was implemented using a 3-D binary array which can maintain the level of details comparable to that of original 2-D slices. Due to the nature of original grey-scale data in most medical applications, a *slice-level boundary detection* mainly based on a gradient operator was suggested. For the *surface formation* process, various interpolation methods were outlined; among them the *shape based* methods were found to be superior and therefore used extensively throughout this project.

Chapter 4 also outlined the new approaches for handling the abnormal regions and deformities occurring in some nuclear medicine brain images. A verification known as the circularity check, for the shape of the detected head contours was defined, based on the curvature measurement. Any unacceptable deformity or hole in the brain surface can be detected by the *circularity check* and then reformed by a type of interpolation process. Two techniques were suggested, based on median filtering and contour reflection, to correct for (*'reform'*) these abnormal regions.

Chapter 5 described various methods for solving the minimization problems. The iterative approaches were indicated as an essential process for surface fitting algorithms. Gradient methods were used to generate a sequence of linear searches along successive directions of a multi-dimensional function. *Golden section search* and *quadratic interpolation* methods were addressed in a linear search strategy as two powerful methods for quadratic functions. The *Powell* method, the F*letcher-Reeves* method and *Newton type* methods were also described as the most efficient multi-dimensional minimization techniques available at the present time.
In this chapter the properties of a suitable minimization algorithm were also outlined, and the Powell method was selected as the best algorithm satisfying most of these characteristics.

In chapter 6, the application of the *sequential process* for 3-D surface registration was described, in a new approach. A reduction in computational cost with respect to the global grid search was obtained while the quality of the result was not degraded. The *cumulative error* was suggested for use instead of the mean distance error for evaluating the mismatch between two surfaces. The different sources of errors contributing to the cumulative distance error were outlined. Using this type of error analysis, the application of a variable threshold was implemented. It produced a very efficient method for termination of mismatch locations. An expression for the calculation of the threshold sequence was also developed.

In chapter 6, the computational implementation of the sequential process was presented as well as the cost prediction. Various factors influencing the computational cost were also studied. The expected time required for the algorithm to decide that there is not a match at a given location was shown to depend on the threshold, *sampling strategy*, and ordering of the sample points.

In chapter 7, a novel approach for a modified fast registration algorithm was presented. The approach, known as *multi-resolution*, employs a *variable grid search* on two low resolution (small image size) binary surfaces. It applies only the most promising transformations on a corresponding bigger (high resolution) image. In this approach, results computed on the basis of some few voxels matched in a coarse resolution image are used for the next resolution level, and in due course for the original high resolution image.

A *degradation technique* was introduced and the error incorporated in the surfaces created by this technique was studied. In addition to that, other sources of error were also studied. In chapter 7, an expression for the definition of the threshold sequence used in multi-resolution sequential methods was also derived. Various parameters influencing this variable threshold were discussed. The expected computational cost and the factors influencing this cost were also explored. The analysis shown in this chapter confirmed a potential reduction of the computational cost by using a *multi-resolution sequential* method at a grid-based search strategy instead of a pure sequential method.

The application of the sequential and multi-resolution algorithms, developed in chapters 6 and 7, was tested on some well defined clinical data, and its results was presented in chapter

8. A number of well defined clinical image data sets, selected and modified to assess the above mentioned algorithms, were described in chapter 8. The technical aspects of the routine imaging systems producing these clinical data were also addressed. The advantage of a sequential process and the superiority of the variable as opposed to the constant threshold level, was also demonstrated in the experiments set forth in this chapter.

The probability analysis presented in chapter 8 investigated the importance of the effects of different threshold levels and various number of sample points used at each resolution levels. The results of this probabilistic analysis confirmed that a threshold level (g) of 2 to 3 (S.Ds of the individual distance error measures at match location) is required for detecting the true match location while keeping the *FAP* (*false alarm probability*) sufficiently low. It was also shown that using a higher number of points (e.g. 200 in the original finest resolution level; i.e. image size of 256^3) provides a better matching performance than a smaller number of points. The number of points required at any resolution level was shown to correspond to nearly half of that required in the next higher resolution level. The number and fraction of sample points which are matched in a mismatch location, and on the other hand, the number of points mismatched in a match location, are two important factors indicating the reliability of the matching process.

Following the specification of the problems associated with the registration algorithm, a set of experiments was designed to determine the matching accuracy. An expected accuracy of about 1, 1.5 and 2 voxels (whose size is in the order of 1mm) for *MR*, *PET*, and *SPECT* registration was achieved, using the arbitrary misregistered data sets. Using some selected reference points (e.g. external markers or internal landmarks) to verify the proposed algorithm, the accuracy of 1.44 ∓ 0.42 mm, 1.82 ∓ 0.65, 2.38 ∓ 0.88 and 3.17 ∓ 1.12 was obtained for MR-MR, MR-CT, MR-PET and MR-SPECT registration, respectively. The computational cost was also determined with respect to certain search windows, template sizes (number of sample points) and other parameters as required for the process. The cost of the multi-resolution process was shown to be about 200 seconds (on the *VAX II*) for the actual search process. The results obtained by the conventional *MR* and *SPECT* images show that the algorithm can be used for the routine imaging studies.

In chapter 9, the problem of *superimposition* of resulting images was defined, as overlaying and combining the correlated images and displaying them to an observer. Some of

displaying clinical images were discussed. The display technique employed in the current work was a type of surface rendering technique based on voxels in a 3-D space. The *surface normals* used in shading method were defined in an *image-based approach.* The *volume rendering* process was also used for displaying the skin surface and the surface of some internal structures of the head in order to improve visualization.

Four types of combining and displaying the correlated (registered) images were employed, as described in chapter 9. The correlated 2-D slices were displayed *side-by-side* illustrating the corresponding clinical information between the two studies. The contours of regions of interest (e.g. ventricles or tumours) were *superimposed* on the 2-D grey-scale images of the other study. A *linked cursor* was used, moved manually by a digitizer on the screen and showed the corresponding structures between the two images. A *colour-coded* display in which different colour ranges are assigned to the structures of each image was also suggested. Due to the difference in grey level ranges of the registered images, different interactive image processing tasks were applied during display process.

The images of thirty patients were studied using the proposed registration processes. The fidelity of the registration process and data alignment was judged by the clinicians and radiologists in the department where the work was carried out. The potential work of the research outlined in this thesis provides a suitable algorithm for registering medical images, integrating the correlated images, and displaying them.

10.3- Suggestions for further work

The problem of data alignment and superimposition is of great interest in the current available imaging modalities which are incapable of producing an image having both functional and morphological information. However, even if alignment was obtainable, the problem associated with correlating images at different times remains unexplored. For these two simple reasons, this issue requires further investigation in the future. There are a number of suggestions for further work to extend the research presented in this thesis.

The surface fitting process can be extended in order to be applied to body imaging as

well. This involves an improvement of the segmentation process by which noisier images can be processed. For example, internal structures of the chest and abdomen such as heart and liver are adjoined to other soft tissue anatomies and are difficult to extract from their neighbourhood.

The timing response of the global grid search is not very favourable even in the sequential and multi-resolution algorithms. The search window size, which is the main factor influencing the speed, should be modified to include only the most promising search locations. A further suggestion is to generate a location matrix (grid) as a look-up table. In this implementation, all search window elements around the true match location of two well-defined image models should be checked. Then only a single location from each number of neighbouring locations (e.g. 27 locations), having the maximum distance error would be selected, and stored in the location matrix. Using the locations selected by this process, one could avoid inclusion of the local minima in the registration process.

Another area which requires more consideration is the degradation process used to create the low resolution surface images. Shape analysis process should be taken into account, thereby enabling different approaches for binary surface scaling to be investigated. Other surface representation techniques, not used in the surface fitting process, could be employed as aternative procedures. Defining the surfaces by parametric equations could be a potential technique for this task

3-D integrated display of internal objects within a 3-D image and interactive image editing are of great interest. In this respect, further work should be done to improve the 3-D display of the correlated images from different modalities such as MR angiographic images and CT bone images, using an interactive strategy.

The final recommendation for further research is to investigate the display requirements and routine clinical applications (and utilities) of the registration process, to enable extraction of useful clinical values. With respect to that, an easily operated clinically viable package, for use by clinicians, applicable in routine medical imaging environment, needs to be produced.

References

Abu-Mostafa Y S, Psaltis D, (1984) Recognition aspects of moment invariants. IEEE Trans. Pattern Analy. Mach. Intel., Vol 6, 698-706.

Abu-Mostafa Y S, Psaltis D, (1985), Image normalization by complex moments, IEEE Trans. Pattern Analy. Mach. Intel., Vol 7, No 1, 46-55.

Andrus J and Campbell C, (1975), Digital image registration method using boundary maps: IEEE Trans. on Compu., Vol 19, 935-940

Apple A, (1968), Some techniques for shading machine renderings of solids, AFIPS 1968 Spring Joint Compu. Conf., 37-45.

Appledorn C and Oppenheim B E, (1980), An automated method for the alignment of image pairs, J. Nucl. Med., Vol 21, 165-167.

Artzy E, Frieder G, Herman G T, (1981), The theory, design, implementation and evaluation of a 3-D surface detection algorithm, Comp. graph. Imag. Proc., Vol 15, 1-24

Arun K S, Huang T S, (1987), Least-Squares fitting of two 3-D point sets, IEEE Trans. on Pattern Analy. Mach. Intel., Vol 9, No 5, 98-700.

Ausherman D A , Dwyer S J, (1972), Extraction of connected edges from radiographs, IEEE trans. Compu. Vol C21, 753-758.

Ballard D, Sklansky J, (1973), Tumour detection in radiographs, compu. Biomed. Res., Vol 6, 299-321.

Barillot G, Gibaud B, (1985), 3-D Representation of Anatomic structures from CT examination, Proce. SPIE, Cannes, France, Vol 602, Dec, 307-314.

Barnea D, Silverman H, (1972), A class of algorithm for fast digital image registration. IEEE Trans. on Comp., Vol C-21, No 2, 179-186.

Barrow H G, Tenenbaum J M, Bolles R C, Wolf H C , (1977), Parametric correspondence and chamfer matching: Two new techniques for image matching, Proc. 5th Int. Conf. on Artificial Intel., Cambridge, MA, pp659-663.

Bhanu D and Faugeras O, (1984), Shape matching of two dimensional objects, IEEE Trans. Pattern Anal. Mach. Intel., Vol 6, 137-155.

Borgefors G, (1984), Distance transformations in arbitrary dimensions, Compu. Vision Graph. Imag. Proc., Vol 27, 321-345.

Borgefors G, (1984), An improved version of the chamfer matching algorithm, 7th Int. Conf on Patt. Recog., Montreal, P.Q., Canada, pp1175-1177.

Borgefors G, (1986), Distance transformations in digital images. Compu. Vision Graph. Imag. proc., Vol 34, 344-371

Borgefors G, (1988), Hierarchical chamfer matching: A Parametric edge matching algorithm, IEEE Trans. on Patt. Anal. and Machine Intel., Vol 10, No 6, pp849-865.

Brent R P, (1973), Algorithm for minimization without derivatives, (Englewood Cliffs N J), Prentice-Hall, Chapter 5.

Bresenham J E, (1965), Algorithm for computer control of a digital plotter, IBM System J., Vol 4, 25-30.

Bric C R , Fennema C L, (1970), Scene analysis using regions. J. Artificial intel., Vol 1, 205-226.

Bright S, Laflin S, (1986), Shading of solid voxel models, Compu. Graph. Forum, Vol 5, No 2, 131-137.

Broyden C G, (1967), Quasi-Newton methods and their application to function minimization, Mathematics of Computation, Vol 21, 368-381.

Castro E D, Morandi C, (1984), Tracking di immagini in movimento rototraslatorio mediante trasformate di Fourier, Trans. Accademia delle Scienze dell, Istituto di Bologna, Apr 1984.

Castro E D, Morandi C, (1987), Registration of translated and rotated images using finite Fourier transforms, IEEE Trans. Pattern Analy. Mach. Intel., Vol 9, No 5, 700-703.

Catmull E, (1974), A subdivision algorithm for computer display of curved surfaces, PhD Thesis, University of Utah.

Catmull E, (1975), Computer display of curved surfaces, Proce. IEEE Conf. Compu. Graph. Pattern Recog. Data Struc., May, p 11.

Chen G T Y, Kessler M L, and Pitluck S, (1985), Structural transfer in three dimensional medical imaging studies, Proce. of the National Compu. Graph. Association Meeting, 171-177.

Chen L S, Herman G T Y, Reyolds R A, Udupa J K, (1985), Surface shading in the cuberille environment, IEEE Compu. Graph. & Appl., 33-43.

Chen C-Tu, Pelizzari C A and Chen G T Y, (1988), Correlating Functional Nuclear Medicine Images with structural CT or MR Images, Proce. of 1988 meeting of Asian Society of Nuclear Medicine.

Chow W K and Aggarwal J K, (1977), Computer analysis of planar curvilinear moving images, IEEE Trans. Compu. Vol 26, 179-185.

Davis L S, (1982), Hierarchical generalized Hough transforms and line-segment based generalized Hough transforms, Pattern Recog., Vol 15, No 4, 277-285.

Drebin R A, Carpenter L, Hanrahan P, (1988), Volume rendering, Compu. Graph., Vol 22, No 4, 65-74.

Fleming J S, (1984), A technique for motion correction in Dynamic scintigraphy, Europ. J. of Nucl. Medicine, Vol 9, 397-402.

Fletcher R, Reeves C, (1964), Function minimization by conjugate gradients. The Compu. Journal, Vol 7, 149-154.

Foley J D, (1980), An improved illumination model for shaded display, Communications of the ACM, Vol 23, No 6, 343-349

Förstner W, (1982), On the geometric precision of digital correlation, International Archives of Photogrammetry, Vol 24, No III, P176-189

Fox P T, Perlmutter J S, and Raichle M E, (1985), A stereotactic method of anatomical localization for positron emission tomography, J. Compu. Assist. Tomog., Vol 9, 141-153.

Freeman H, (1967), On the classification of line-drawing data, in models for the perception of speech and visual form, W. Wathen Dunn, Ed, Cambridge Mass: MIT press. 408-412.

Freeman H, (1974), Computer processing of line-drawing images, Compu. Surveys, Vol 6, 57-97.

Frei W, Singh M, and Shibata T, (1980), Digital image change detection, Optical Eng., Vol 19, 331-338.

Frieder G, Herman G, (1985), Large software problems for small computers. An example from medical imaging, IEEE software, Sep. 37-47.

Frieder G, Gordon D, Reynolds R A, (1985), Back-to-Front display of voxel-Based objects : IEEE Compu. Graph. & Appl., Jan, 52-60.

Fuchs H, Kedem Z M, (1977), Optimal surface reconstruction from planar contours, communications of the ACM, Vol 20, No 10, 693-702.

Goerss S and kelly P J, (1982), A computed tomographic stereotactic adaptation system, Neurosurgery, Vol 10, No 3, 375-379.

Goldstein R A and Nagel R, (1971), 3-D visual simulation, Simulation, Jan, 25-31.

Gonzalez R C and Wintz P, (1977), Digital image processing, Press, Publisher: Addison-Wesley.

Gordon W J, (1969), Spline-blended surface interpolation through curve network, J. Math. Vol 18, No 10, 931-952.

Gordon D, Reynolds R A, (1985), image space shading of three-dimensional object, Compu. Vision Graph. Imag. proc., Vol 29, 361-379.

Gouraud H, (1971), Continuous shading of curved surfaces, IEEE Trans. Computer, Vol C-20, 623-628.

Herman G T Y, Liu H K, (1977), Display of Three-Dimensional Information in Computed Tomography: J. of Compu. Assis. Tomog. Vol 1, No 1, 155-160.

Herman G T Y, Liu H K, (1979), Three-Dimensional Display of Human Organs from Computed Tomograms, Compu. Graph. Imag. Proc., Vol 9, 1-21.

Herman G T Y, Udupa J K, (1981), Display of three dimensional discrete surface, Proc. SPIE, Vol 283, 90-97.

Herman G T Y, (1988), From 2D to 3D Representation, NATO ASI Series, Vol F39, (Mathematics and computer science in Medical Imaging), Edited by M.A. Viergever, A.E. Todd-Pokropek; Springer-Verlag Berlin, Heidelberg, 197-219

Herman G T Y, Bucholtz C A, (1991), Shape-based interpolation using a chamfer distance, 12th Inter. conf. on Information Proc. in Med. Imag., July, Wye, UK, 314-325.

Hill D L G, Hawkes D J, Harrison N A, Ruff C F, (1993), A strategy for automated multimodality image registration incorporating anatomical knowledge and imager characteristics, Information Processing in Medical Imaging, 13th Int. Conf. IPMI'93, Flagstaff, Arizona, USA, June 14-18, Eds: Barrett H H, Gmitro A F, pp182-196.

Höhne K H, Bernstein R, (1986), Shading 3D-images from CT using grey-level gradients, IEEE Trans. on Med. imag., Vol MI-5, NO 1, 45-47.

Höhne K H, Delapaz R L, (1987), Combined surface display and reformatting for the 3-D analysis of tomographic data, Investigative Radiology, Vol 22, 658-664.

Höhne K H, Bomens M, Tiede U, (1988), Display of multiple 3-D-objects using generalized voxel-model, SPIE, Med. Imag. Vol 2.

Hu M K, (1962), Visual pattern recognition by moment invariants: IRE Trans. on Information Theory, Vol 8, 179-187

Hueckel M F, (1971), An operator which locates edges in digitized pictures, J. of ACM, Vol 18, 113-125.

Jackins C L, Tanimoto S L, (1980), Octree and their use in representing three-dimensional objects, Compu. Graph. Imag. Proc, Vol 14, 249-270.

Johnson D B, (1977), Efficient algorithms for shortest paths in sparse networks, J. of the Association for compu. mach., Vol 24, No 1, 1-13.

Kall B A, Kelly P J, (1985). Cross registration of points and lesion volumes from MR and CT, IEEE 7th Annual Conf. of the Engineering in medicine and Biology Society, 937-942.

Katz Y H, (1965), Pattern recognition of meteorological satellite cloud photography, Proc. Third Symposium on Remote Sensing of Environment, Inst. of Science and Technology, Univ of Michigan, Feb, 173-214.

Kelly P J, Kall B A, and Goerss S, (1984), Transposition of volumetric information derived from Computed Tomography scanning into Stereotactic space, Surg. Neurol., Vol 21, 465-471.

Keppel E, (1975), Approximating complex surfaces by triangulation of contour lines, IBM J. Res. Develop., Jan, 2-10.

Kessler M L, Chen G T Y, (1987), Techniques and application of image correlation in radiotherapy treatment planning, Front. Radiat. Therapy and Oncology, Vol 21, 25-32.

Kuglin C D, Hines D C, (1975), The phase correlation image alignment method, Proc. Inter. Conf. Cybernetics, Sep., 163-165.

Levoy M, (1988), Display of surfaces from volume data, IEEE Compu. Graph. and Appl., Vol 8, No 3, 29-37.

Lillestrand R L, (1972), Techniques for change detection, IEEE Trans. Compu., Vol C-21, No 7, July, 654-659.

Lin Zse-Cherng, Lee H, Huang T S, (1986), Motion estimation from 3-D points sets with and without correspondences; Part I, A frequency-domain algorithm for determining motion of a rigid object from range data without correspondences, Proce. IEEE Conf. Compu. Vision and Pattern recog. Miami Beach; Fl, June, 194-198.

Lin W C, Liang C C, (1988), Dynamic elastic interpolation for 3-D object reconstruction from serial cross-sectional images, IEEE Trans. Med. imag., Vol 3, No 3, 225-232.

Liu H, (1977), Two and three-dimensional boundary detection, Compu., Graph. imag. proc. Vol 6, 123-134.

Maitra S, (1979). Moment invariants: Proce. IEEE. Vol 67, 697-699

Mandava V R, Fitzpatrick J M, Pickens D R, (1989), Adaptive search space scaling in digital image registration, IEEE Trans. on Med. Imaging, Vol 8, No 3, P251-262.

Meagher D J, (1982), Geometric modelling using octree encoding, Compu. Graph. Imag. Proc., Vol 19, 129-147.

Medioni S, Nevatia R, (1984), Matching images using linear features, IEEE Trans. Pattern Analy. Mach. Intel., Vol 6, 675-685.

Merickel M, Mc Carthy M, (1985). Registration of contours for 3D reconstruction: IEEE 7th Annual Conf. of the Engineering in Medicine and Biology Society, 616-620.

Merrill R D, (1973), Representation of contours and regions for efficient computer search, Communications of the ACM, Vol 16, 69-82.

Metropolis N, Rosenbluth A, Resenbluth M, Teller A, and Teller E, (1953), J of Chem. Phys, Vol 21, P1087.

Mitiche A, Aggarwal J K, (1983). Contour registration by shape-specific points for shape matching. Compu. Vision Graph. Imag. Proc. Vol 22, 396-408.

Montanari U, (1971), On the optimum detection of curves in noisy pictures, ACM, Vol 14, 335-345

Morrin T H, (1974), A Black-white representation of a gray-scale picture, IEEE Trans. Compu., Vol 23, 184-186.

Nagel R N, Rosenfeld A, (1972), Ordered search techniques in template matching, Proce. IEEE, Vol 66, 242-244.

Nelder J A, Mead R, (1965), Computer Journal, Vol 7, P308.

Oghabian M A, Todd-Pokropek A (1991), Registration of Brain images by a multi-resolution sequential method, Information Processing In Medical Imaging; 12th International Conference, IPMI'91, Wye, UK, (Eds: Colchester A C F, Hawkes D J), July, 165-174.

Oswald H, (1983), Imaging of discrete 3-D scenes, SPIE, April 19-22, Geneva, 'Application of digital image processing', Vol 397, 310-315.

Paton K, (1970), Conic sections in chromosome analysis, Pattern Recog., Vol 2, 39-51.

Papoulis A, (1965), Probability, Random variables, and Stochastic processes, (Press), New York, McGraw-Hill.

Pavlidis T, (1974), Segmentation of plane curves, IEEE Trans. Compu. C-23, 860-870.

Pavlidis T, (1978), A review of algorithms for shape analysis, Compu. Graph. & Imag. Proc., Vol 7, 243-258.

Pelizzari C A, Chen G T Y, (1987). Registration of multiple diagnostic imaging scans using surface fitting, Proce. of the 9th ICCR, 437-440.

Pelizzari C A , Chen G T Y, (1989). Accurate Three-Dimensional registration of CT, PET and NMR Images of the brain, J. Compu. Assis. Tomog., Vol 13, 20-26.

Phong B T, (1975), Illumination for computer generated pictures, communications of the ACM, Vol 18, 311-317.

Powell M J D, (1964), An efficient method of finding the minimum of a function of several variable without calculating derivatives, The Compu. J., Vol 7, 155-162.

Powell M J D, (1965), A method for minimizing a sum of squares of nonlinear functions without calculating derivatives, The Compu. J., Vol 7, 303-307.

Pratt W K, (1974), Correlation techniques of image registration, IEEE Trans. Aerospace and Electronic Systems, Vol 10, 353-358.

Prewitt J M S, Mendelsohn M L, (1966), The analysis of cell images, Annals NY Academic Science, 1035-1053.

Price K, Reddy R, (1979), Matching segments of image: IEEE Trans. Pattern Analy. Mach. Intel. Vol 1, 110-116.

Price R R, Lindstrom D P, Hillis S, Friesinger G C, Brill A B, (1975), Analytical techniques for image superposition, Proce. fifth symposium on sharing of compu. progr. and technology in Nucl. Med., 241-250.

Ramapriyan H K, (1976), Multilevel approach to sequential detection of pictorial features, IEEE Trans. Compu., Vol 25, No 1, 66-78.

Raya S P, Udupa J k, (1990), Shape-based interpolation of multidimensional objects, IEEE Trans. Med. Imag., Vol TMI-9, 32-42.

Ros D, Juvells I, (1984), A simplified method for the automatic alignment of images affected by random noise, Optica Acta, Vol 31, No 10, 1151-1159.

Robert S M, Lyvers H I, (1961), The gradient method in process control, Indust. and Engineering Chem., Vol 53, 877-882.

Robinson G S, (1977), Edge detection by compass gradient masks, compu. Graph. imag. proc., No 6, 492-501.

Roger D F, (1985), Procedural elements for computer Graphics (Press).

Rosenfeld A, Johnston E, (1973), Angle detection on digital curves, IEEE Trans. Compu., Vol 22, 875-878.

Rosenfeld A, Weszka J, (1975), An improved method of angle detection on digital curves, IEEE Trans. Compu. Vol 24, 940-941.

Rosenfeld A, Vanderbrug G J, (1977). Coarse-Fine template matching, IEEE Trans. on System Man and Cybernetic, Feb., 104-107.

Rosenfeld A, Kak A C, (1982), Digital Picture Processing, Academic Press, Vol 2, Chapter 10.

Sadjadi A F, Hall E L, (1978), Numerical computations of moment invariants for scene analysis, Proc. IEEE Conf. Pattern Recog. Imag. Proc., Chicago, May, 181-187.

Singh M, Frei W, Shibata T and Huth G C, (1979), A digital technique for accurate change detection in Nuclear Medicine images with application to Myocardial perfusion studies using Thallium-201, IEEE Trans. Nucl. Science, Vol 26, 565-575.

Spinks T J, Jones T, Gilardi M C, and Heather J D, (1988), Physical performance of the latest generation of commercial positron scanner, IEEE Trans. on Nucl. Science, Vol 35, No 1, Feb.

Svedlow M, Gillem C D, (1978). Image registration: Similarity measure and preprocessing method comparisons. IEEE Trans. Aerospace and Electronic Systems, Vol 14, 141-149.

Teague M R, (1980), Image analysis via the general theory of Moments, J. Optical Society of America, Vol 70, No 8, Aug., 920-930.

Tiede U, Riemer M, Bomans M, and Höhne K H, (1988), Display techniques for 3-D-tomographic volume data, Proce. NCGA.

Tuy H K, Tuy L T, (1984), Direct 2-D display of 3-D objects, IEEE Compu. Graph. & Appl., Vol 4, No 10, 29-33.

Vanderbrug G J, Rosenfeld A, (1977), Two-stage template matching, IEEE Trans. on Compu. Vol C-26, No 4, Apr.

Venot A, Lebruchec J F, and Roucayrola J C, (1984), A new class of similarity measures for robust image registration, Compu. Vision Graph. Imag. Proc., Vol 28, 176-184.

Venot A, Leclerc V, (1984), Automated correlation of patient motion and gray values prior to

subtraction in digitized angiography, IEEE Trans. Med. Imag., Vol 3, 179-186.

Wang Y F, Aggrawal J K, (1985), Construction of surface representation from 3-D volumetric scene description, proce. IEEE compu. society conf., Sanfrancisco: Compu. Vision and Pattern Recog., Jun, 130-135.

Weszka J S, Nagel R N, Rosenfeld A, (1974), A threshold selection technique, IEEE Trans., Compu., Vol 23, 1322-1326.

Williams D L, Ritchie J L, Hamilton G W, (1978), Implementation of a digital image superposition algorithm for Radionuclide images: An assessment of its accuracy and reproducibility, J. Nucl. Med., Vol 19, No 3, 316-319.

Wolfe R N, (1969), A dynamic thresholding scheme for quantization of scanned images, Proce. of Automatic Pattern Recog., National Security Industrial Association, Washington D.C, May, 143-162.

Wong R Y, Hall E, (1978), Sequential hierarchical scene matching, IEEE Trans. Compu., Vol C-27, No 4, Apr.

Wong R Y, (1978), Sequential scene matching using edge features, IEEE Trans. Aerospace and Electronic Systems, Vol AES-14, No 1, Jan.

Wong R Y, Hall E L, (1979). Performance comparison of scene matching techniques. IEEE Trans. on Pattern. Analy. & Mach. Intel. Jul, Vol 1, No 3, 325-330.

Wood R P, Cherry S R, Mazziotta J C, (1992), Rapid automated algorithm for aligning and reslicing PET images, J. of Compu. Assis. Tomog., Vol 16, No 4, pp620-633.

Wu S C, Abel J F, Greenber D P, (1977), An interactive computer graphics approach to surface representation, Communications of the ACM, Vol 20, No 10, 703-712.

Yam S, Davis L S, (1981), Images registration using generalized Hough Transform, IEEE Conf. on Pattern Recog. and Imag. Proc., Dallas/TX, 526-533.