

**ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ**  
**Τμήμα Ηλεκτρονικών Μηχανικών & Μηχανικών Υπολογιστών**

**Μεταπτυχιακή διατριβή με θέμα:**

***‘A comparative study of landmark-based image registration techniques’***

**Εξεταστική Επιτροπή:**

**Καθ. Μιχαλ. Ζέρβακης  
Καθ. Νικ. Σιδηρόπουλος  
Αναπλ. Καθ. Πετράκης Ευρ.**

**Μαρεντάκης Ιωάννης  
Δεκέμβριος 2005**

## Contents

1.	Introduction.....	8
2.	Content-Based Image Retrieval (CBIR) in Medical Applications .....	12
2.1	Introduction.....	12
2.2	Background and Related work.....	12
2.3	Introduction to IRMA project.....	13
2.3.1	The 7 steps of processing.....	13
2.3.2	Results of the IRMA approach .....	15
2.4	Introduction to the NLM/CEB project.....	17
2.4.1	System Overview .....	18
2.4.2	Results of NLM – CEB project.....	20
2.5	The need for content-based medical image retrieval .....	21
3.	Medical Image Registration Methods.....	23
3.1	Introduction.....	23
3.2	Criteria used for classification of registration methods .....	23
3.2.1	Dimensionality .....	24
3.2.2	Nature of registration basis .....	24
3.2.3	Nature of transformation.....	26
3.2.4	Domain of the transformation.....	27
3.2.5	Interaction .....	28
3.2.6	Optimization procedure .....	28
3.2.7	Modalities involved in the registration .....	30
3.2.8	Subject.....	31
3.3	Overview of Medical Image Registration Methods.....	31
3.11	Why choose ICP and ASM algorithms.....	33
4.	Norm-based approaches in orientation problems .....	34
4.1	Models and Curve fitting .....	34
4.2	Norm – Based solutions.....	36
4.3	Least Squares approach.....	37
4.4	Linear Programming for least absolute error approaches.....	39
4.5	Investigation of Assumptions .....	40
5.	Active Shape Models (ASM) – ‘Smart Snakes’ .....	44
5.1	Introduction and Related work.....	44
5.2	The Point Distribution Model .....	45
5.3	Labeling the Training Set.....	46
5.4	Capturing the statistics of a set of aligned shapes.....	46
5.5	Aligning the training set.....	47
5.6	Examples using Active Shape Models.....	49
5.6.1	Finding resistors with an ASM .....	49
5.6.2	Finding hands with an ASM .....	50
4.7	How it works.....	51
5.8	Results and figures of the ASM algorithm .....	52
6.	The Iterative Closest Point Algorithm (ICP) .....	56
6.1	Introduction and related work.....	56
5.2	Quaternion.....	58

6.2.1 Description.....	58
6.2.2 Practical uses of quaternions.....	59
6.3 The ICP algorithm.....	60
6.4 Convergence Theorem for the ICP algorithm.....	63
6.5 Initial states for Global Matching .....	64
6.6 How it works.....	65
6.7 Results and figures of the ICP algorithm.....	67
7. Nelder-Mead algorithm.....	70
7.1 Introduction.....	70
7.2 The Nelder – Mead algorithm.....	70
7.3 Statement of the algorithm.....	70
7.4 Results and Figures of the Nelder-Mead algorithm.....	75
8. Experiments .....	77
8.1 Implementation of ICP, ASM and Nelder-Mead algorithms with 20 shapes represented by 15 landmarks each. ....	77
8.2 Error Curves.....	84
9. Conclusion and further work .....	92
10. References.....	93

## List of Figures

Figure 2.1 IRMA processing steps (resulting in seven semantic layers (blue)).....	15
Figure 2.2 Web based interface for manual reference categorization (top) and logging of code modifications within the central database (bottom) .....	17
Figure 2.3 CBIR system.....	19
Figure 2.4 CBIR system.....	20
Figure 3.1 Examples of 2-D transformations.....	27
Figure 3.2 Directional optimization along versors vs conjugate directions.....	29
Figure 4.1 X- histograms ICP .....	41
Figure 4.2 Y histograms ICP .....	41
Figure 4.3 X-histograms ASM.....	42
Figure 4.4 Y-histograms ASM.....	42
Figure 5.1 32 point model of the boundary of a resistor.....	45
Figure 5.2 Resistors .....	49
Figure 5.3 Hands.....	51
Figure 5.4 Mouse Pos .....	52
Figure 5.5 ('L04531' image).....	53
Figure 5.6 ASM implementation .....	54
Figure 5.7 X-ray digital image where the landmarks are not fixed properly.....	55
Figure 5.8 The alignment when one of the shapes has random fixed points. ....	55
Figure 6.1 Quaternions.....	58
Figure 6.2 ICP 'The Normal Case' .....	67
Figure 6.3 The 'Random Case' .....	68
Figure 7.1 The triangle BGW, midpoint M and reflected point R for the Nelder-Mead method.....	72
Figure 7.2 The triangle BGW, point R and extended point E.....	72
Figure 7.3 The construction point $C_1$ or $C_2$ for Nelder-Mead method.....	73
Figure 7.4 Shrinking the triangle towards B .....	74
Figure 7.5 'The normal case' .....	75
Figure 7.6 'The random numbering case' .....	76
Figure 7.7 The 'not fixed properly case' .....	76
Figure 8.1 MousePos 'The Normal Case'.....	77
Figure 8.2 X-Ray Image .....	78
Figure 8.3 The 'Normal Case' .....	79
Figure 8.4 Mouse Pos The Not Fixed properly Case.....	80
Figure 8.5 The Not Fixed properly Case .....	81
Figure 8.6 MousePos Different Numbering Case.....	82
Figure 8.7 The Different Numbering Case .....	83
Figure 8.8 Square Error.....	84
Figure 8.9 Std deviation per landmark.....	85
Figure 8.10 X landmarks.....	86
Figure 8.11 Y landmarks.....	86
Figure 8.12 ALL Landmarks .....	87
Figure 8.13 Points under T.....	88
Figure 8.14 Focusing in a part of Fig 8.13.....	88

Figure 8.15 Mean Error for points under T.....	89
Figure 8.16 Focusing in a part of Fig 8.15.....	89
Figure 8.17 Std Error for points under T .....	90
Figure 8.18 Focusing in a part of Fig. 8.17.....	90

## List of Tables

Table 2.1 IRMA 's approach are obvious very good.....	15
Table 2.2 Error rates of different methods on the subset of 1.617 images from the IRMA reference database.....	15
Table 5.1 Eigenvalues of the covariance matrix derived from a set of resistor shapes. ....	50
Table 7.1 Logical decisions for the Nelder-Mead algorithm.....	74

## **ABSTRACT**

The purpose of this thesis is to examine research in medical image retrieval. In particular we are interested in medical X-Ray images. Such research is useful because automation is a desired property in modern medical centres since it can enhance their efficiency and improve the quality of service offered to the public. A general presentation of existing Content – Based image Retrieval (CBIR) systems is provided, however, the thesis focuses on the landmark based image alignment problem. Two error functions, are compared, for solving the alignment problem, the least squares and the least absolute error functions. Two solutions to the least square problem are compared to the solution of the least absolute value problem. Least Squares was solved using a quaternion-based solution and a theoretical approach as proposed in the Iterative Closest Point (ICP) and Active Shape Models (ASM) approaches. After the investigation of the Gaussian distributions, for both algorithms, it was found that the assumptions of the least formulation are not valid so we led to another solution of the error function that was the absolute error approach, and in terms of this Nelder – Mead algorithm was implemented. The two approaches were compared using three criterion and least square found better solution overall but the other one reacts better in specific points of the shapes. Both error functions performed well providing reasonably good alignment.

## 1. Introduction

The purpose of this thesis is to examine the development of a system to help in the analysis of medical X-Ray images. Such a system can be applied to modern medical centres and help doctors and hospital staff in their everyday tasks that are related to medical image processing and especially X-Ray analysis. Automation is a desired property in modern medical centres since it can enhance their efficiency and improve the quality of service offered to the public.

The construction of a full scale medical image analysis system is a challenging task. Such a system should provide storage or access to storage of medical images and efficient retrieval methods on a patient or image content basis. This should ideally be done in an on demand context, since access to such information is usually time critical. Content based access is important since it enhances the ability of doctors and researchers to access information that may improve their understanding of the field of study, through the provision of relevant cases to the ones they study in the everyday work. It is also the case that more efficient retrieval than based on keywords can be performed since retrieval is done on shape basis and can return objects that for example have the same deficiencies. For example, a doctor may require a bone leg with a certain deficiency in the upper part. Not all bone leg images are thus relevant and a refinement base don the particular deficiency has to be made. Due to the aforementioned reasons, Content Based Image Retrieval (CBIR) has been an active research field.

A presentation of the structure of such systems is provided in chapter two. Two of the most prominent such systems are presented. These are the IRMA system and the NLM system. In principle, the construction of such a system relies on a database of classified and indexed images, which is later used for retrieval. Such a database is the NHANES II X-Ray database which is also presented in this thesis and has been used to retrieve images in order to test the ideas that are presented in the text. The purpose of CBIR systems is to classify and index medical images to improve retrieval time and accuracy. Also they play an active role in the content-based retrieval since each such query has to be done in the same terms as the ones used for the classification process and is thus applied to each new image.

Indexing is usually performed in a number of sub steps. These commonly involve segmentation, feature extraction, feature vector organization and classification. Segmentation is the process where each individual element if the medical image is identified. Then, feature extraction on each individual part is performed. Using feature extraction a number of global usually numeric features are extracted from the relevant image part. These features are used to form a feature vector that maps each image on a point in a usually multi-dimensional feature space. This point is the identity of the image in this space. Similar images are likely to exist somewhere in the vicinity of this point. In this sense such a process is basic to classification which is the last step of the storage process. After all these steps are performed, classification provides terms that group images in a meaningful way that assists retrieval. Retrieval can then performed either on lectical terms (for example find arm X-Rays) or on a similarity term, e.g. find images that contain bone structures similar to an input bone structure. In the first case, the query runs through the already stored image classification attributes and returns the matching ones. In an image-content based query the image has first to be analyzed in the same sense as



the already stored images, and then images with similar properties are looked up in the database.

This thesis is focusing in the first step of the indexing process which is also useful in the retrieval process, which is segmentation. Image segmentation is usually performed using edge detection followed by an image alignment. Edge-detection works by scanning the image for edges and results in identifying objects in the image by linking edges. Objects of interest have then to be aligned to a 'model' object that exists in the database so that feature analysis can be performed. If the objects are not aligned properly, feature extraction and subsequently the later steps are likely to fail. The reason for this is that parts of the human body that are represented in medical images appear in different absolute positions and with varying orientations in the medical image. Thus, if not aligned, the feature extraction algorithm that works using image coordinates of shapes will produce noisy and uncertain feature vectors. By aligning the images before feature extraction, the feature space becomes robust and correctly represents the image space. For the purpose of this thesis the Kirsch algorithm is used for edge detection. It is found to be useful and to correctly segment the medical image in the relevant parts. The thesis is focusing more on the alignment problem. In particular, two algorithms for alignment based on landmarks are used to align an image to a 'model' or simply another image. These algorithms require landmarks to be present on the image. Landmarks mark individual points of the image and can be placed either manually from experienced personnel or automatically on the part of the image under investigation. They should definitely be present in points that reveal structural or shape properties of the object to be aligned. Landmark selection is also particularly important since the feature extraction algorithm is also working using the previously defined landmarks and thus requires carefully selected landmarks so that the model shapes will be closer to the reality. However, the landmark based approach is error prone in the sense that human body parts from different people are inherently different and anomalies caused by age or are present from birth in some people do not appear in others. In this sense, it is usual that images have different number of landmarks, and also that landmarks are not positioned in the same order. Thus, a landmark based alignment algorithm should be invariant to the number of landmarks and to the order by which the landmarks were added on the image. Alignment algorithms work by minimizing the distance between the two shapes that have to be aligned using a gradient based iterative process. These shapes are usually a 'model' or mean shape that has been constructed from the database and a new shape to be added in the database or to be used in a query. A simple classification example is to compare the error of the alignment algorithm for a number of model shapes and the new shape. Through this process a rough filtering of the shape candidates can be performed.

Alignment is performed by minimizing a certain error function. We compare two error functions the least squares one and the least absolute value one. Two alignment algorithms that are presented in this thesis and work using least squares minimization of the over determined problem are the Iterative Closest Point (ICP) and the Active Shape Models (ACM) algorithm. They are implemented, tested and compared and the results are presented in the text. In addition, the Nelder Mead non-linear optimization algorithm is used to minimize the least absolute value error function. The algorithm performs comparably to the ICP and ACM algorithms.

In chapter 4 of this thesis, we present a survey of norm - based orientation techniques, along with a theoretical preview of the existing solutions of the over determined problem. Least squares and Linear Programming is briefly discussed in this chapter. After these comes the investigation of the Gaussian Assumptions in the least square approached algorithms. Distributions histograms show the need for another treatment of the error function which is the absolute error approach.

In chapter 5, the ASM (Active Shape Models) algorithm is presented and implemented. Results showing the performance of this algorithm are also presented. In general, this algorithm is found to work properly, however there are some serious deficiencies present. These are related to the fact that no provision is made for shapes with differences in numbering and number of landmarks. In this sense, new landmarks have to be added when missing using interpolation and numbering has to be manually checked. This is a cumbersome process given the variability of human body parts and thus the algorithm is not so efficient. Thus, we equip this algorithm with a closeness metric to match points between shapes in order demonstrate a proper solution. The ASM paper is interesting also because it presents a method for generating model objects of the shapes based on the eigenvectors and eigenvalues of the covariance matrix as this is defined by the landmarks of a large number of images.

In the sixth chapter of this thesis the ICP is presented and implemented. The algorithm is tested in situations where shapes have different number of landmarks and also in cases where landmarks have different numberings. It is found to work very well in both cases. The reason for this, is the fact that the algorithm compares one point to another based on a distance metric. In this sense points are assigned to other points in each algorithm step, based on this metric and possible differences in numbering are eliminated using a proper transformation of the object to be aligned. Also this closeness metric works well in the case where the model and the shape to be aligned have a different number of landmarks. This is due to the fact that the algorithm matches as points of the shape to be aligned to the model point so that distance of only the relevant points is used and ignore the points that are not mapped are ignored. The closest point matching idea thus is proved to be very useful in the context of landmark based image alignment. The ICP algorithm is also found to be quite efficient since a few iterations are usually enough for convergence.

In chapter seven the least absolute value error function is presented and solved using the Nelder-Mead algorithm and an implementation is provided. The algorithm attempts to minimize a scalar-valued nonlinear function of 4 real variables using only function values. The Nelder – Mead algorithm falls in the general class of ‘direct search’ methods, which maintain at each step a simplex, a triangle in two dimensions. The algorithm is tested in similar situations as the others algorithms. It provides with satisfactory results in all the previous cases.

Furthermore, in chapter eight a number of experiments and tests is performed for the three algorithms. Three criteria were used in terms of this comparison. All algorithms were tested in three cases and aligned 20 shapes to a ‘model shape’ Also mean square error and absolute error were used for the comparison of the algorithms. Standard deviation per landmark and for the error was measured. Overall least square approach minimize the error for all the points in stead of absolute error approach which is forcing

as many points as possible to minimum error. A group of diagrams were mined and the results were presented.

Finally, chapter nine and ten present the conclusion and the bibliography on which this thesis was based.

In conclusion, this thesis is focusing in the alignment problem. We find that algorithms that can provide solution to this problem exist and we compare three of them. Least squares solution is optimal in case that the residuals obey a Gaussian distribution. If this is not the case, the assumptions of the least squares formulation are not valid, leading to suboptimal formulations of the problem. In such cases other norm formulations could lead to more robust results. In our problem we investigated the distribution of residuals from the least square formulation and found them to deviate from the Gaussian distribution. For this reason we also tested the least absolute error formulation and compared results.

## 2. Content-Based Image Retrieval (CBIR) in Medical Applications

### 2.1 Introduction

Content-based image retrieval (CBIR) aims at describing the complex object information of digital images by non-textual features, which are applicable for efficient query processing. Color, texture, and shape are used within the first CBIR approaches such as the query by image content (QBIC) system[1]. Besides the raw data layer representing the initial images, QBIC models a feature layer on which the retrieval operates. For the content based retrieval engine (CORE), objects and spatial relationships are described by “concepts” within the interpretation layer[2]. In Blobworld, this layer is referred to as scheme layer. It is build from ellipsoids (“blobs”) representing local image regions of uniform color or texture on an abstract level of interpretation [3]. **However, two or three semantic layers are insufficient to model medical knowledge form image retrieval and results are poor when common CBIR algorithms are applied to medical images.**

Content-Based Image Retrieval refers to the retrieval of images that are indexed by features derived from the image pixels. CBIR features may include texture, boundaries of objects, geometric relationships among objects, grayscale or color histograms, etc. CBIR is different from conventional image retrieval systems in two ways:

- (i) ***the methods used to index*** (a conventional system has a human indexer enter text that describes image contents)
- (ii) ***in the methods used to retrieve*** (conventionally, retrieval of images is by relational database queries on the text used to index the image)

### 2.2 Background and Related work

The ever-increasing volume of medical images, the economic impracticality of manually indexing these images, and the inadequacy of human language alone to describe image contents that are visually recognizable and medically significant, such as shape and geometry, color, texture of objects within images, all provide impetus for research and development toward practical Content-Based Image Retrieval (CBIR) systems that could become a standard offering of the medical library of the future.

Some approaches for content-based image retrieval have been published which are specially designed to support medical tasks. KORN et al[4] describe a system for fast and effective retrieval tumor shapes in mammogram x-rays. Another approach is ASSERT's, that operates only on high resolution computed tomography of the

lung[5]. Here, a physician delineates the region bearing a pathology and marks a set of anatomical landmarks when the image is entered into the database. ASSERT has extremely high data entry costs, which prohibit its application for clinical routine. LONG et al. Access a large collection of 17.000 spine radiographs by means of shape analysis, where biomedical categories such as “anterior osteophytes present/not present” are distinguished automatically[6]. The data entry costs are low, but queries are limited to pre-defined categories. CHU et al. present a knowledge-based image retrieval system with spatial and temporal constructs [7]. Brain lesions are extracted automatically within three-dimensional data sets from computed tomography and magnetic resonance imaging. Their representation model consists of an additional knowledge-based layer within the semantic model. This layer provides a mechanism for accessing and processing spatial, evolutionary and temporal queries. But, these concepts for medical image retrieval are task-specific, i.e. limited to a distinct modality, organ, or diagnostic study and hence, usually not direct transferable to other medical applications. Another two CBIR systems X-rays spine images are:

1. The IRMA project: A Content-Based Image Retrieval in Medical Applications (<http://irma-project.org>)
2. NLM/CEB- A prototype Content-based Image Retrieval System for spine X-rays (<http://nlm.nih.gov>)

Analytically these two projects work as follows:

## **2.3 Introduction to IRMA project**

IRMA project is a project that has been stated in the Department of Medical Informatics and the Department of Diagnostic Radiology of Technology (RWTH) in Aachen at Germany. The objective of this work was to be developed a general structure for semantic image analysis that is suitable for content-based image retrieval in medical applications and architecture for its efficient implementation.

Compared to standard CBIR systems, at least three additional semantic levels of abstraction are needed to be coped with the complex medical knowledge that is to be handled by a general system for content-based image retrieval in medical applications. A low-level of medical knowledge is determined by the imaging modality including technical parameters, the orientation of the patient position with respect to the imaging system, the body region examined, and the functional system under investigation. Based on prototype images, a mid-level of knowledge is described by regions of interest (ROIs) within the images, and a high level is obtained from information regarding the spatial or temporal relationships or relevant objects.

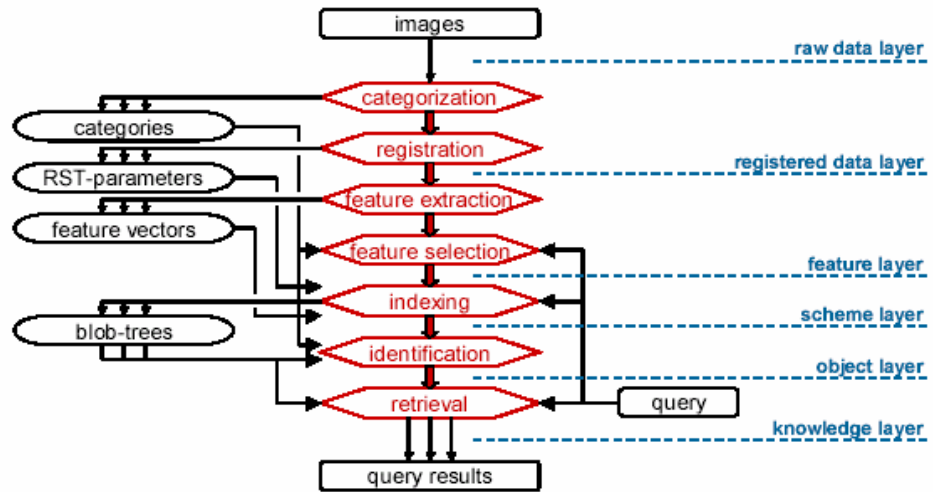
### **2.3.1 The 7 steps of processing**

- Consequently, IRMA splits the retrieval process into seven consecutive steps(Fig. 2.1).
- The categorization step aims at determining for each image entry the imaging modality and its orientation as well as the examined body region and functional

system. So, a detailed hierarchical coding scheme was developed[8], which exceeds the complexity of existing tags of the digital imaging and communications in medicine (DICOM) standard. Automatic categorization is based on a reference of 10.000 images selected from clinical routine and manually classified by experienced radiologists. The automatic categorization of query-by-example images is performed by combining DICOM header information and global image features describing the entire image.

- Registration in geometry (rotation, translation, scaling) and contrast generates a set of transformation parameters that is stored for the corresponding image in each of its likely categories. Registration is based on prototypes which are manually defined for each category, and further incorporate medical expert knowledge into the IRMA systems. The transformation is not performed at this step of processing; instead the generated parameters are utilized at higher layers of abstraction.
- The feature extraction step derives local image descriptions, i.e. a feature value (or a set of values) is obtained for each pixel. These can be category-free or category specific, such as the application of an active shape model (ASM) that uses a priori-knowledge derived from the respective category.
- Decoupling feature selection from feature extraction allows integrating both image category and querying context into the abstraction process.
- Indexing provides an abstraction of the previously generated and selected image features, resulting in a compact image description. According to the selected feature set, this is done via clustering of similar image parts into regions represented by their second area moment description as ellipses. The hierarchical indexing enables the processing of ROIs, which are marked by the user when issuing a query. The ROI is not determined a-priori, the incorporation of medical mid-level knowledge becomes possible.

Only, the retrieval step requires online computations while all other steps can be performed automatically in batch mode at entry time of an image into the database.



**Figure 2.1** IRMA processing steps (resulting in seven semantic layers (blue))

### 2.3.2 Results of the IRMA approach

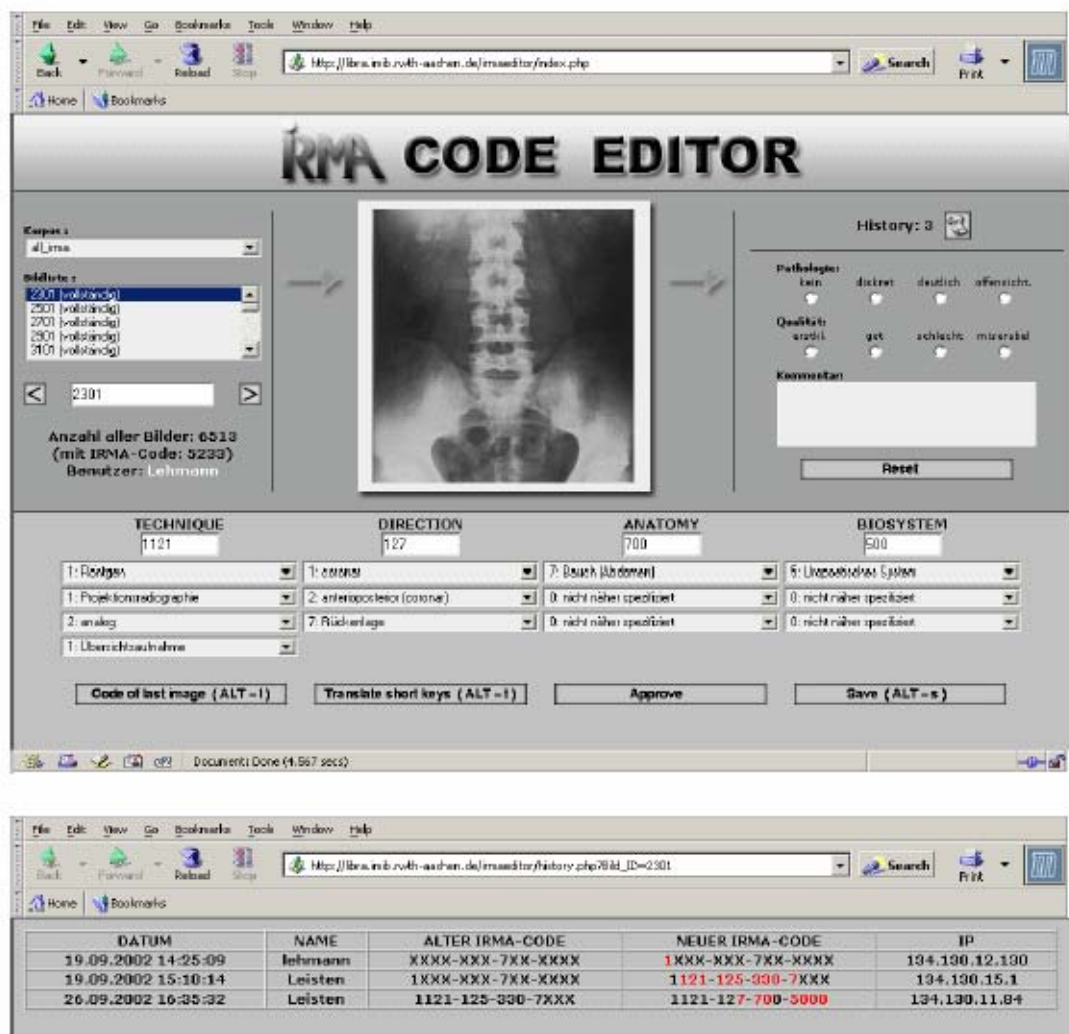
Although manual labeling of the IRMA reference database is still in progress, the system was already used for processing basic queries, i.e. queries regarding the category of images. Based on a subset of 1.617 images from six body regions (abdomen, limbs, chest, breast, skull, and spine) acquired with various modalities in several orientations, a statistical classifier based on the tangent distance performed with 8% error rate [9]. Table 2.2.1 shows results of other methods on this data. The results of IRMA 's approach are obvious very good.

Method	Error Rate
co-occurrence matrices	29.0 %
squared images, 1-nearest-neighbor in Euclidian distance	18.1 %
local representation [28]	9.6 %
statistical approach	8.0 %

**Table 2.2** Error rates of different methods on the subset of 1.617 images from the IRMA reference database.

The IRMA reference database holds 3,879 images that have been labeled according to the IRMA code (Table 2). Since the frequency of imaged body regions reflects the clinical situation, chest radiographs occur most often. In addition, the sub-region classes are distributed irregularly. Classes with less than ten examples were not considered for leaving-one-out experiments. A downscaled image representation of 16 x 16 pixels results in 256 features that were used for a 5-nearest-neighbor correlation-based classifier. Regardless of modality, orientation and biological system, automatic classification of the anatomic region is possible with error rates of 11.6% and 15.8% using leaving-one-out for eight body regions and for 26 relevant sub-regions, respectively. Taking into account the large number of categories and the high variability within the categories due to the various imaging modalities, the obtained results are excellent. Also the IRMA concept pursues all likely categories for later processing steps: When the three or five most likely sub-regions are considered for each sample, the error rate for 26 sub-regions drops to 8.4% and 5.6%, respectively. For these experiments, the 15 and 25 nearest neighbors were analyzed. The error rates can be further decreased by combining the correlation-based classifier with a statistical classifier. If each classifier contributes the 8 and 13 nearest neighbors, error rates yield 6.9% and 4.8%, respectively. Preliminary results were also obtained for the web-based interfaces. The IRMA code editor is used for manual labeling of reference images (Figure 2.2). The IRMA code can be edited either directly by typing in the code or by selecting the entries from selection boxes. Based on prior selections, the sub-codes offered are adopted properly. All changes are recorded in a history protocol stored in the central database for easy error recovery (Fig. 2.2).





**Figure 2.2** Web based interface for manual reference categorization (top) and logging of code modifications within the central database (bottom)

## 2.4 Introduction to the NLM/CEB project

Two of the projects that have been created in “Lister Hill National Center for Biomedical Communications, National Library of Medicine”, are:

- (1) NHANES II
- (2) Content-Based Image Retrieval (CBIR) system

NHANES II is a database that contains demographic, anthropometric, adult health questionnaire and physical examination data for 20,322 NHANES II survey participants. Also contains 17,000 cervical and lumbar spine x-ray images. All these x-ray images have been made publicly available through an FTP archive that is published on the CEB Website. To view these images in full spatial and grayscale resolution, CEB has developed a Java image viewer that is publicly available from the same site, 550 of these images have been converted to the standard TIFF 8-bit form and made publicly available also.

CBIR system is a technique applied to the NHANES II x-ray images, and eventually to other biomedical images. Specific objectives include:

- Conduct R&D in the steps needed for content-based indexing, shape segmentation, feature extraction, feature vector organization and classification
- Conduct R&D into techniques for image-based retrieval
- Develop the algorithms needed to implement both indexing and retrieval
- Design and develop a next-generation CBIR system (incorporating these algorithms) to serve as a platform for evaluating techniques that index and retrieve the NHANES II images in an effective manner
- In the long term, extend CBIR techniques developed for the NHANES II images to other biomedical images[11]

#### 2.4.1 System Overview

CBIR comprises both indexing and retrieval, defined as follows:

- **Indexing** - the computer-assisted data reduction of images into mathematical features. Indexing may be subdivided into the steps of
  - **Segmentation** – the computer-assisted determination of the boundaries of the objects of interest
  - **Feature extraction** – the reduction of the boundaries that result from the segmentation process into mathematical “feature vectors” that capture the shape properties of interest
  - **Feature vector organization** – the organization of the feature vectors in the database into a structure optimized for searching efficiency
  - **Classification** – the computer-assisted labeling of segmented objects into categories of interest.
- **Retrieval** - the user interaction to retrieve desired images from the database, this comprises
  - **User query formulation** – the method used to specify the query, the advanced methods of interest are query by image example and query by sketch applied via different search paradigms, such as target-search, category search, and open ended browsing

- **User query feature extraction** – the reduction of the user's exemplary image or sketch into a mathematical feature vector compatible with the feature vectors stored in the database

- **Query search space strategy** – the search strategy used to efficiently search the database of feature vectors for those near to the user query feature vector; usually, a search strategy is defined by the choice of organization of the feature vectors

- **Similarity matching method** – the method used to compare input and stored image features and to measure similarity between them.

The segmentation work focus in particular on Active Shape Modeling (ASM). The ASM formulation that was followed is that described by Cootes and Taylor[15] of the University of Manchester.

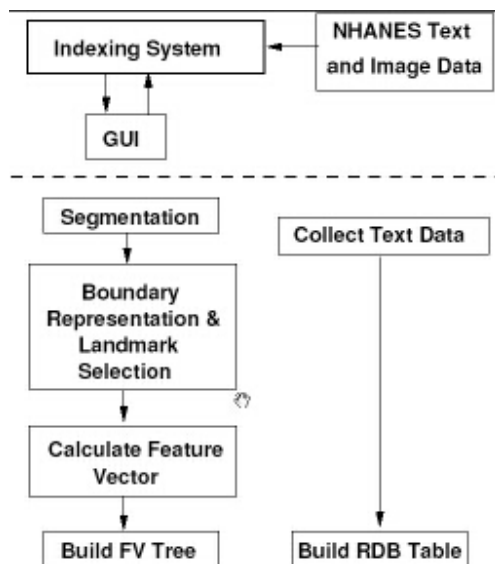


Figure 2.3 CBIR system

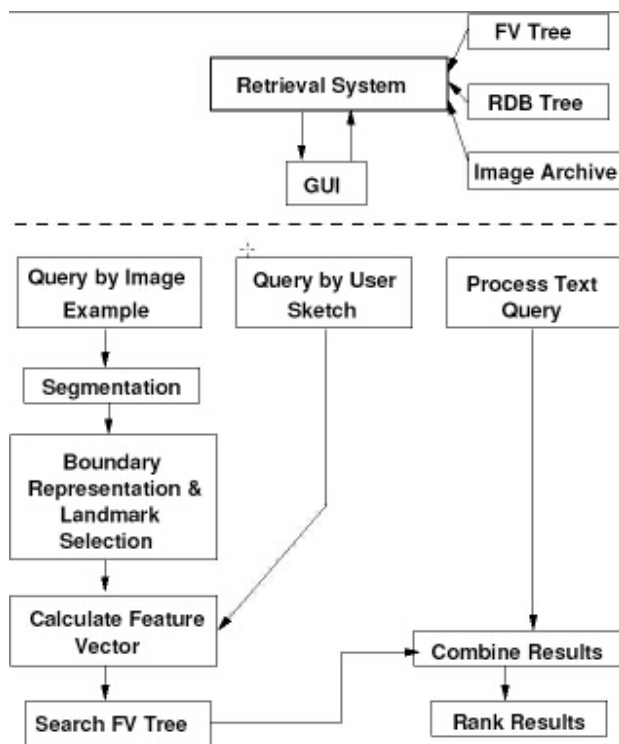


Figure 2.4 CBIR system

In Figures 2.3 and 2.4 are shown the Block Diagram of the CBIR system: Indexing system and Retrieval system, respectively

## 2.4.2 Results of NLM – CEB project

The results acquired from the implementation of the NLM – CEB project can be grouped as follows for each part of the retrieval process:

1. Segmentation: Fully automated segmentation - feature extraction from medical images is a very challenging problem with no method directly applicable as found in the literature. Its quality is affected by three important factors: viz., technique, image quality, and image size - resolution. Hence, computer assisted segmentation is usually a more promising approach. Significant problems were faced with poor image quality in the digitized spine x-ray images where segmentation methods often confuse tissue and vertebra boundaries. Also techniques that have been known to work well for smaller images often do not scale well when presented with higher resolution images. Poor segmentation results were observed on application of fairly robust color and texture segmentation techniques
2. Feature Representation: The dimensionality of the represented feature is highly correlated with the quality and inversely with efficiency of retrieval. Such a system must ensure that the representation technique faithfully captures the image content/semantics. This must be achieved because of the performance evaluation on retrieval quality of

vertebrae with anterior osteophytes. The method selected for representing vertebra boundaries was Fourier descriptors following initial reduction in dimension using polygon approximation. It was found that while the method did fairly well (70% retrieval precision) on complete shape queries, the users were unable to focus on specific regions of the pathology. Intuitive modifications to the query shape were not comparably represented resulting in low quality retrieval. This led the developers of the system to pursue research into partial or incomplete shape matching. In a recent evaluation of retrieval performance improvement through the use of partial shape matching, it was observed that for pathological shapes the retrieval relevance improved by a factor of 2.2. That is, the average retrieval performance improved from 3.815 relevant matches with a standard deviation of 1.66 in the top 10 retrieved shapes to on average 8.125 relevant matches with a standard deviation of 0.25.

3. Feature organization: The initial evaluation of hierarchical clustering showed that there is a significant performance improvement over linear searches. In a set of experiments on a 1298 shape database, on average only 176 leaf nodes needed to be accessed for a 13 nearest neighbor (most similar shapes) query. In scaling the range to 40 nearest neighbors, it needed to access an average of only 315.25 nodes. This is a significant improvement over 1298 accesses required in a linear search. It is, however, optimized for only one feature. Additionally, it uses shape descriptions with a fixed number of boundary points since it requires the similarity measure to be a metric. The Procrustes distance that was used as a similarity measure is a metric, but operates only on boundaries with fixed number of points. Fixed number of points in low dimension can affect image feature detail and consequently adversely affect retrieval quality.

## **2.5 The need for content-based medical image retrieval**

There are several reasons why there is need for additional, alternative image retrieval methods apart from the steadily growing rate of image production.

The goals of medical information systems have often been defined to deliver the needed information at the right time, the right place to the right persons in order to improve the quality and efficiency of care processes. Such a goal will most likely need more than a query by patient name, series ID or study ID for images. For the clinical decision-making process it can be beneficial or even important to find other images of the same modality, the same anatomic region of the same disease. Although part of this information is normally contained in the DICOM headers and many imaging devices are DICOM-compliant at this time, there are still some problems. DICOM headers have proven to contain a fairly high rate of errors, for example for the field anatomical region rates of 16% have been reported[10]

Clinical decision support techniques such as case-based reasoning or evidence-based medicine can even produce a stronger need to retrieve images that can be valuable for supporting certain diagnoses.

It needs to be stated that the purely visual image queries as they are executed in the computer vision domain will most likely not be able to ever replace text-based methods as there will always be queries for all images of a certain patient, but they have the potential to be a very good complement to text-based on their characteristics. Still, the problems and advantages of the technology have to be stressed to obtain acceptance and use of visual and text-based access methods up to their full potential.

Besides diagnostics, teaching and research especially are expected to improve through the use of visual access methods as visually interesting images can be chosen and can actually be found in the existing large repositories. The inclusion of visual features into medical studies is another interesting point for several medical research domains. Visual features do not only allow the retrieval of cases with visual similarity but different diagnoses. In teaching, it can help lecturers as well as students to browse educational image repositories and visually inspect the results found. This can be the case for navigating in imaging atlases [<http://www.loni.ucla.edu/MAP/index.html>]. It can also be used to cross-correlate visual and textual features of the images.

## **2.6 Relation of CBIR systems with Image Registration**

As it was mentioned above 7 steps of processing are required for the image retrieval of a CBIR system. After the image categorization, which is the first one, comes the image registration. Registration requires, in geometry, rotation, translation and scaling also generates a set of transformed parameters that is stored for the corresponding image in each of its likely categories. Image registration is needed for the feature extraction of the image. In the next chapter we are discussing the existing medical image registration methods

### 3. Medical Image Registration Methods

#### 3.1 Introduction

Information gained from two images acquired in the clinical track of events is usually of a complementary nature thus proper integration of useful data obtained from the separate images is often desired. A first step in this integration process is to bring the modalities involved into spatial alignment, a procedure referred to as registration. After registration, a fusion step is required for the integrated display of the data involved.

#### 3.2 Criteria used for classification of registration methods

The classification of registration methods used here is based on the criteria formulated by van den Elsen, Pol & and Viergever [132]. Nine basic criteria are used, which can each be subdivided again. The nine criteria and primary subdivisions are given in the following table.

<b>I. Dimensionality</b>				
<b>II. Nature of registration basis</b>	1.Extrinsic	2. Intrinsic	3. Non-Image based	
<b>III. Nature of transformation</b>	1. Rigid	2. Affine	3. Projective	4. Curved
<b>IV. Domain of transformation</b>				
<b>V. Interaction</b>				
<b>VI. Optimization procedure</b>				
<b>VII. Modalities involved</b>	1. Monomodal	2. Multimodal	3. Modality to modal	4. Patient to modality
<b>VIII. Subject</b>	1. Intrasubject	2. Intersubject	3. Atlas	
<b>IX. Object</b>				

A registration procedure can always be decomposed into three major parts: the problem statement, the registration paradigm, and the optimization procedure. The problem statement, the choice of paradigm and the optimization procedure together

provide a unique classification according to the nine criteria mentioned. Although criteria are heavily intertwined and have many cross-influences, it can be said that the problem statement determines the classification according to criteria ‘Modalities involved’, ‘Subject’, and ‘Object’, and has a direct bearing on the criteria ‘Dimensionality’ and ‘Nature of Transformation’. The paradigm influences the criteria ‘Nature of Registration Basis’, ‘Nature of Transformation’, ‘Domain of Transformation’, and ‘Interaction’ most directly, while the optimization procedure influences criterion ‘Interaction’ and controls ‘Optimization Procedure’.

### **3.2.1 Dimensionality**

The main division here is whether all dimensions are spatial, or that time is an added dimension. In either case, the problem can be further categorized depending on the number of spatial dimensions involved. Most current papers focus on the 3D/3D registration of two images (no time involved). 3D/3D registration normally applies to the registration of two tomographic datasets, or the registration of a single tomographic image to any spatially defined information, e.g., a vector obtained from EEG data. 2D/2D registration may apply to separate slices from tomographic data, or intrinsically 2D images like portal images. Time series of images are acquired for various reasons, such as monitoring of bone growth in children (long time interval), monitoring of tumor growth (medium interval), post-operative monitoring of healing (short interval), or observing the passing of an injected bolus through a vessel tree (ultra-short interval). If two images need to be compared, registration will be necessary except in some instances of ultra-short time series, where the patient does not leave the scanner between the acquisition of two images. The same observations as for spatial-only registrations apply.

### **3.2.2 Nature of registration basis**

Image based registration can be divided into extrinsic, i.e., based on foreign objects introduced into the imaged space, and intrinsic methods, i.e., based on the image information as generated by the patient. Extrinsic methods rely on artificial objects attached to the patient, objects which are designed to be well visible and accurately detectable in all of the pertinent modalities. As such, the registration of the acquired images is comparatively easy, fast, can usually be automated, and, since the registration parameters can often be computed explicitly, has no need for complex optimization algorithms. The main drawbacks of extrinsic registration are the prospective character, i.e., provisions must be made in the pre-acquisition phase, and the often invasive character of the marker objects. Non-invasive markers can be used, but as a rule are less accurate. A commonly used fiducial object is a ‘stereotactic’ frame [25 - 32] screwed rigidly to the patient’s outer skull table, a device which until recently provided the best “gold standard” for registration accuracy. Such frames are used for localization and



guidance purposes in neurosurgery. Sometimes other invasive objects are used, such as screw-mounted markers [33 - 41], but usually non-invasive marking devices are reverted to. Most popular amongst these are markers glued to the skin [42 - 53], but larger devices that can be fitted snugly to the patient, like individualized foam moulds, head holder frames, and dental adapters have also been used, although they are little reported on in recent literature [54 - 58, 42]. The nature of the registration transformation is often restricted to be rigid (translations rotations only).

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

Landmarks can be anatomical, i.e., salient and accurately locatable points of the morphology of the visible anatomy, usually identified interactively by the user[42, 43, 46, 58 - 87, 29, 48 -51, 31], or geometrical, i.e., points at the locus of the optimum of some geometric property, e.g., local curvature extrema, corners, etc, generally localized in an automatic fashion [88 - 97]. Anatomical landmarks are also often used in combination with an entirely different registration basis [58, 42, 46, 72, 76, 78, 81, 82, 83] methods that rely on optimization of a parameter space that is not quasiconvex are prone to sometimes get stuck in local optima, possibly resulting in a large mismatch For the optimization of the latter measure the Iterative closest point (ICP) algorithm [98] (implemented in this thesis) and derived methods are popular. Its popularity can be accredited to its versatility –it can be used for point sets, and implicitly and explicitly defined curves, surfaces and volumes–, computational speed, and ease of implementation. The Procrustean optimum can sometimes be computed, using e.g., Arun’s method [99], but is more commonly searched for using general optimization techniques. Yet other methods perform landmark registration by testing a number of likely transformation hypotheses, which can, e.g., be formulated by aligning three randomly picked points from each point set involved.

Segmentation based registration methods can be rigid model-based [100 - 119] where anatomically the same structures (mostly surfaces) are extracted from both images to be registered, and used as sole input for the alignment procedure. They can also be deformable model based [120 - 130], where an extracted structure (also mostly surfaces, and curves) from one image is elastically deformed to fit the second image. The rigid model based approaches are probably the most popular methods currently in clinical use.

The voxel property based registration methods stand from the other intrinsic methods by the fact that operate directly on the image grey values, without data reduction by the user or segmentation Despite its drawbacks, principal axes methods are widely used in registration problems that require no high accuracy, because of the automatic and very fast nature of its use, the easy implementation. Theoretically, these are the most flexible of registration methods, since they –unlike all other methods mentioned– do not with reducing the grey valued image to relatively sparse extracted information, but use all of the available information throughout the registration process.

It seems paradoxical that registration of multimodal images can be non-image based, but it is possible if the imaging coordinate systems of the two scanners involved are somehow calibrated to each other. This usually necessitates the scanners to be brought in to the same physical location, and the assumption that the patient remains motionless

between both acquisitions. These are prohibitive prerequisites in nearly all applications, but they can be sufficiently met in applications involving the use of ultrasound. Since ultrasound systems can come as hand-held devices that are equipped with a spatial (optical) localization system, they are easily calibrated, and can be used while the patient is immobilized on the CT, MR or operating gantry. The technique of calibrated coordinate systems is also often used in registering the position of surgical tools mounted on a robot arm to images.

### 3.2.3 Nature of transformation

An image coordinate transformation is called **rigid**, when only translations and rotations are allowed. If the transformation maps parallel lines onto parallel lines it is called **affine**. If it maps lines onto lines, it is called **projective**. Finally, if it maps lines onto curves, it is called **curved** or **elastic**. Each type of transformation contains as special cases the ones described before it, e.g., the rigid transformation is a special kind of affine transformation. A composition of more than one transformation can be categorized as a single transformation of the most complex type in the composition, e.g., a composition of a projective and an affine transformation is a projective transformation, and a composition of rigid transformations is again a rigid transformation. A rigid or affine 3D transformation can be described using a single constant matrix (a) equation:  $y_i = a_{ij}x_j$  where x and y are the old and new coordinate vectors. In the rigid case, this equation is constrained as:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ 1 \end{pmatrix} = \left( \begin{array}{ccc|c} & & & t \\ r & & & \\ \hline 0 & 0 & 0 & 1 \end{array} \right) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix},$$

where t is an arbitrary translation vector, and r is a 3 x 3 rotation matrix defined by:

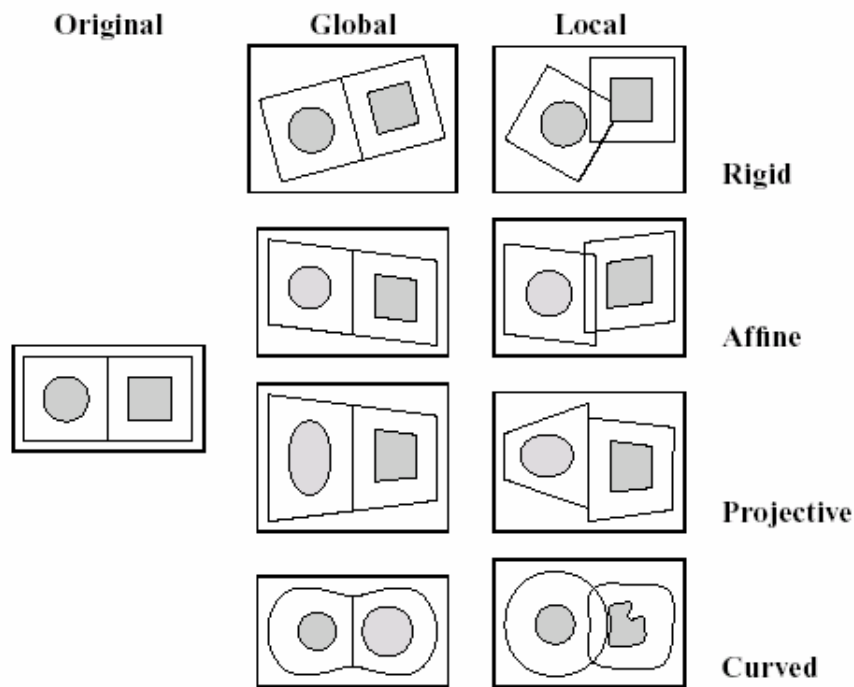
$$r_{kl} = r_{ij}^{(1)} r_{jk}^{(2)} r_{kl}^{(3)}, \quad r^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_1 & -\sin \alpha_1 \\ 0 & \sin \alpha_1 & \cos \alpha_1 \end{pmatrix},$$

$$r^{(2)} = \begin{pmatrix} \cos \alpha_2 & 0 & -\sin \alpha_2 \\ 0 & 1 & 0 \\ \sin \alpha_2 & 0 & \cos \alpha_2 \end{pmatrix}, \quad r^{(3)} = \begin{pmatrix} \cos \alpha_3 & -\sin \alpha_3 & 0 \\ \sin \alpha_3 & \cos \alpha_3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

i.e.,  $r^{(i)}$  rotates the image around axis  $i$  by an angle  $a_i$ . In the affine case,  $r$  is unrestricted. In the projective case, it can only be used a constant matrix representation if employing homogeneous coordinates:  $y_i = u_i / u_1 - u_i = a_{ij}x_j$ , where  $a$  is an arbitrary 4x4 constant matrix. Curved transformations cannot in general be represented using constant matrices. Most applications represent curved transformations in terms of a local vector displacement (disparity) field:  $y_i = x_i + t_i(x)$ , or as polynomial transformations in terms of the old coordinates.

### 3.2.4 Domain of the transformation

A transformation is called **global** if it applies to the entire image, and **local** if subsections of the image each have their own transformations defined. In **Figure 3.1** we can see examples of all transformation types mentioned.



**Figure 3.1** Examples of 2-D transformations

### 3.2.5 Interaction

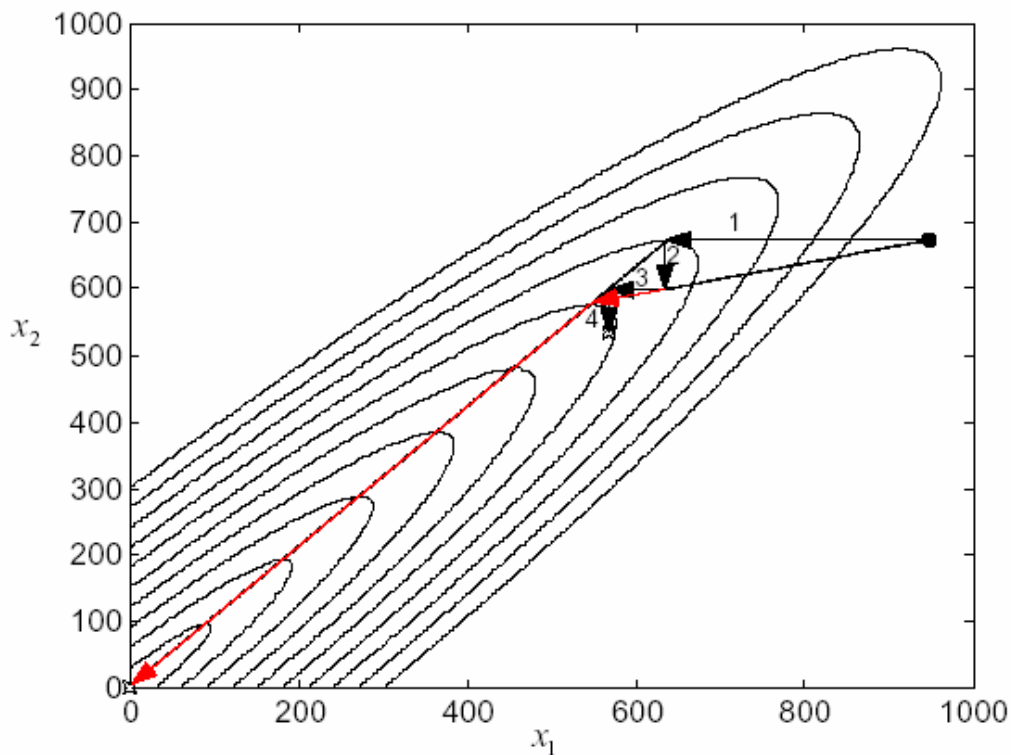
Concerning registration algorithms, three levels of interaction can be recognized. **Automatic**, where the user only supplies the algorithm with the image data and possibly information on the image acquisition. **Interactive**, where the user does the registration himself, assisted by software supplying a visual or numerical impression of the current transformation. **Semi-automatic**, where the interaction required can be of two different natures: the user needs to initialize the algorithm, e.g., by segmenting the data, or steer the algorithm.

Extrinsic methods are often automated, since the marker objects are designed to be well visible and detectable in the images involved. The users are required to roughly point out the marker region, or supply a seed point located in the marker (semi-automatic). Of the intrinsic methods, the anatomical landmark and segmentation based methods are commonly semi-automatic and the geometrical landmark and voxel property based methods are usually automated. Fully interactive methods are reported on very little in the recent literature.

### 3.2.6 Optimization procedure

The parameters that make up the registration transformation can either be computed directly, i.e., determined from the available data, or searched for, i.e., determined by finding an optimum of some function defined on the parameter space. In the former case the use of computation methods is restricted almost completely to applications relying on very sparse information, e.g., small point sets. In the case of searching optimization methods, most registration methods are able to formulate the paradigm in a standard mathematical function of the transformation parameters to be optimized. This function attempts to quantify the similarity as dictated by the paradigm between two images given a certain transformation. Popular techniques are Powell's method [156] the Downhill Simplex method [157]. Frequent additions are multi-resolution (e.g., pyramid) and multiscale approaches to speed up convergence, to reduce the number of transformations to be examined and to avoid local minima. Some registration methods employ non-standard optimization methods that are designed specifically for the similarity function at hand, such as the ICP algorithm (which is implemented in this thesis and is created for rigid model based registration). In optimization procedure the general task of all the methods is to minimize the error function. In ICP algorithm the aim is to minimize the Euclidean distance. Usually this distance is used. In each of the pre-mentioned methods a different algorithm is created and used to success this task.

A very popular algorithm for optimization is the Powell algorithm. It follows the idea of subsequent directional minimizations in order to find optimization problem solution. Once the starting point is chosen, there is always the dilemma how to generate directions for the line search subroutine. Iterating through a set of versors is mathematically correct as they span the optimization domain but can turn out to be ineffective if the objective function forms narrow curving valleys. The improvement would be to choose the next direction so that while optimizing along it the gradient stays perpendicular to the current direction. Such a pair of directions is called conjugate.



**Figure 3.2** Directional optimization along versors vs conjugate directions

Directional optimization along versors vs. conjugate vectors is shown on **figure 3.2**. The first two steps (black arrows) are done in the direction of versors. The third and fourth can be done also along versors (black arrows) but it is much better to perform it using a conjugate vector (arrows nearby) that is created utilizing experience from the last two steps. The steps taken along conjugate vectors lead immediately to the function minimum. Usually to make a new conjugate direction it is needed to have Hessian matrix (or its approximation) of the function being optimized. Powell suggested a routine in which to make next direction one utilizes solely the data from the last  $N$  line searches. The routine preserves algorithm's quadratic convergence rate but its drawback is that directions tend to be linearly dependent. This can be omitted in several ways: one of them is to give up the direction set periodically and to start over with a set of versors.

### 3.2.7 Modalities involved in the registration

Four classes of registration tasks can be transferred based on the modalities that are involved. In monomodal applications, the images to be registered belong to the same modality, as opposed to multimodal registration tasks, where the images to be registered stem from two different modalities. In modality to model and patient to modality registration only one image is involved and the other “modality” is either a model or the patient himself. It is used the term “modality” in a loose sense, not only applying to acquired images, but also to mathematical models of anatomy or physiology, and even to the patient himself. Such inclusions are necessary to properly type-cast the four categories according to the actual registration task to be solved. The classification is closed if only the actual coordinate systems that need to be related are considered, i.e., the coordinate systems referring to the actual modalities named in the problem statement.

For example:

- For diagnostic purposes, two myocardial SPECT images are acquired of the patient, under rest and stress conditions.

Their registration is a monomodal application.

- To relate an area of dysfunction to anatomy, a PET image is registered to an MR image. This is a multimodal application.
- To register an MR to a PET image, a PET image is first simulated from the MR image, and the real and simulated PET images are registered. This is still a multimodal application.
- An example of modality to model is the registration of an MR brain image to a mathematically defined compartmental

model of gross brain structures.

- In radiotherapy treatment, the patient can be positioned with the aid of registration of in-position X-ray simulator images to a pre-treatment anatomical image. Although the registration task is performed using only the images acquired, the actual task of patient positioning is clearly an example of patient to modality registration.

In the categories ‘modality to model’ and the ‘patient to modality’ as it is explained before the term ‘modality’ is not applying to images but to mathematical models of anatomy or physiology or to the patient himself. In the first case the model is the image and the ‘modality’ is registered to the model. In the second case the model is existed from the experience of the past and with this knowledge the patient is positioned creating the ‘modality’.

### 3.2.8 Subject

When all of the images involved in a registration task are acquired of a single patient, it is referred as **intrasubject registration**. If the registration is accomplished using two images of different patients (or a patient and a model), this is referred to as **intersubject registration**. If one image is acquired from a single patient, and the other image is constructed from an image information database obtained using imaging of many subjects, this is named as **atlas registration**. Intrasubject registration is by far the most common of the three, used in almost any type of diagnostic and interventional procedure. Intersubject and atlas registration appear mostly in 3D/3D MR or CT brain image applications. The nature of the registration transformation is mostly curved; these applications are always intrinsic, either segmentation based or voxel property based, using the full image content. Some applications use rigid transforms, but their application is limited. Others use anatomical landmarks for a deformation basis of a curved transformation; unfortunately such applications often require the transformation in large image areas to be interpolated from the nearest landmark transformations, which may prove unreliable. The use of intersubject and atlas matching can notably be found in the areas of gathering statistics on the size and shape of specific structures, finding anomalous structures, and transferring segmentations from one image to another.

### 3.2.9 Object

The list in with the modalities (in p.40) is not theoretically complete, but composed of those imaging areas encountered in recent literature. It is noteworthy that the majority of papers concerns global head registration, even if the registration method used could possibly be used in other image areas.

## 3.3 Overview of Medical Image Registration Methods

There is a definite shift in research from extrinsic to intrinsic methods, although clinically used methods are often still extrinsic. Of the intrinsic methods, the surface based methods appear most frequently, closely followed by “full image content” voxel property based methods. Instances of the latter type are slowly setting the standard for registration accuracy, a place formerly reserved for frame and invasive fiducial based registrations. The application of full image content voxel property based methods is however still largely limited in the extensive application field of intra-operative registration and radiotherapy treatment related registration (both requiring patient to modality registration). Especially in the area of intra-operative registration, surface based

methods are dominant, and voxel based methods almost absent. The reasons may be clear: it is relatively easy to obtain a surface from the patient, either using laser scanning, probes, 2D imagery, etc. , while obtaining reliable image information for voxel property based methods is more difficult: intra-operative imaging may not even be part of the normal surgical routine. If it is, images are usually 2D, and if 3D, of a relative poor quality given common equipment and acquisition sequence constraints in the operating theater. Moreover, surface based methods are, on the average, still faster than voxel property based methods. However, a problem with surface based methods is that they cannot cope with shift of relevant anatomy relative to the surface used in the registration, which may be severely restraining to intra-operative application. This problem may be solved using voxel based methods, but given the current state of affairs considering Registration methods, surgical protocol, and intra-operative imaging, this will not be done in the very near future. In the case of radiotherapy treatment related registration (patient positioning, and patient position verification), the future will certainly include more of voxel based methods: imaging (X-ray simulator images and portal images) is already part of the common clinical treatment routine; radiotherapy relies almost exclusively on imaging for (tumor) localization, unlike surgery, where the visual impression is still the most important cue. It is not unlikely that this will change soon for a number of surgical applications, given the current trend of less and less invasive surgery that requires making use of advanced imaging techniques. Furthermore, given a computed transformation, many applications do not require complex visualization techniques, but can be adequately handled using subtraction techniques. Multimodal applications cannot be discussed in general terms, the applications are simply too diverse. It is tempting, but incorrect, to say registration results are some what more satisfying in methods involving scintigraphic imaging, perhaps because the relatively blurry nature of the images allows for a slightly larger displacement. In, e.g., CT to MR registration, a displacement of a pixel can sometimes be obvious to the naked eye, and to obtain an accuracy in this order of magnitude, we cannot avoid to investigate precision at the acquisition level, (e.g., the distortions induced by field inhomogeneity in MR images), which are of the same order of magnitude. However, the resolution of the images should not be used to formulate a clinically relevant level of accuracy: it is very well possible that a SPECT to MR registration requires a higher accuracy than some instance of CT to MR registration, even though it is likely that the smaller error is more easily assessed by the naked eye in the latter case. The actual level of accuracy needed is in many applications still an unknown, and cannot accurately be quantified, even by the clinicians involved. Intra-operative registration and methods on patient positioning in radiotherapy are in clinical use with apparent good results at a number of sites. On the diagnostic use of registration (modality to modality), much less information can be found. We suspect that, bearing in mind the possible clinical potential of diagnostic registration, it is actually used very little. The reasons for this are, probably, in essence of a logistic nature: unlike in the intra-operative scene (where all imaging and operations take place in the same room), in many multimodal diagnostic settings images are acquired at different places, – often even at different departments–, by different people, at different times, often transferred to different media, and frequently evaluated by different specialist diagnosticians. Besides these logistic reasons, it is also often unclear how a registration can optimally be used in the diagnostic process. It has already been pointed out that much



research can still be done in this area. Many methods can still be considered barred from meaningful clinical application by the fact that they are as yet improperly validated. Although the proper verification methods are known in most cases, and coarsely laid out in the previous section, for most applications the painstaking work of conducting the many experiments involved is only now starting.

### **3.11 Why choose ICP and ASM algorithms**

In this thesis three methods were selected: ICP, ASM and Nelder - Mead algorithms in terms of the alignment process. ASM and ICP algorithms treat the problem in the least squares sense unlike to Nelder – Mead which uses the least absolute solution. These methods do not use any segmentation but are rather based on landmarks. We selected these algorithms in order to test the potential of landmark based approaches to registration, the effects of outliers to the least squares problem and the benefits of a robust formulation steaming from a least absolute error approach. The landmarks that were used for the representation of the shapes were fixed manually with the use of a special designed program called MousePos with programming language C++. The user can easy design the points at the corners of the vertebrae so that minimum loss of data for the followed registration.

## 4. Norm-based approaches in orientation problems

### 4.1 Models and Curve fitting

If all the parameters appear linearly and there are more observations than basis functions, we have a linear least squares problem. The design matrix  $X$  is  $m$ -by- $n$  with  $m > n$ . We want to solve:

$$X\beta \approx y \quad (i)$$

This system is over determined- there are more equations than unknowns. So we cannot expect to solve the system exactly. The Orientation problem minimizes the mean squared error between two matched points sets under rigid-body transformations. In more words, a free point set  $A$  is matched to a fixed point set  $B$  such that a matching function  $\mu: A \longrightarrow B$  exists. An error function

$$E = \sum_{a \in A} \|\mu(a) - R(a) + t\|^2$$

is minimized over all rotations  $R(\cdot)$  and translations  $t$ .

Faugeras and Hebert represent the rotation as a unit quaternion and discovered the first solution in 1983. Horn rediscovered this technique around 1985 and published in 1987 in the most referenced solution to this problem. An alternate solution which represents the rotation as a rotation matrix and uses a singular value decomposition to find the optimal rotation was discovered by Arun in 1986, although a similar computation for computing the angle between subspaces was in Bjorck and Golub. Also representing rotation as a rotation matrix but solving the optimal rotation with the eigenvalue decomposition was discovered independently by Shwarz and Sharir in 1987 and in 1988. A set of degenerate cases was corrected by Umeyama in 1991. A fourth and more obscure solution technique represents rotation and translations together as dual number quaternions and was discovered by Walker. An overview of techniques is presented in the following.

A common approach to the solution of linear equation as in (i) is curve fitting. Let  $t$  be the independent variable and let  $y(t)$  denote an unknown function of  $t$  that we want to approximate. Assume there are  $m$  observations, i.e. values of  $y$  measured at specified values of  $t$ .

$$y_i = y(t_i), i = 1, \dots, m$$

The idea is to model  $y(t)$  by a linear combination of  $n$  basis functions,

$$y(t) \approx \beta_1 \phi_1(t) + \dots + \beta_n \phi_n(t)$$

The design matrix  $X$  is a rectangular matrix of order m-by-n with elements

$$x_{ij} = \phi_j(t_i)$$

The design matrix usually has more rows than columns. In matrix-vector notation, the model is

$$y \approx X\beta$$

The symbol  $\approx$  means "is approximately equal to."

The basis functions  $\phi_j(t)$  can be nonlinear functions of  $t$ , but the unknown parameters,  $\beta_j$ , appear in the model linearly. The system of linear equations

$$X\beta \approx y$$

is over determined if there are more equations than unknowns.

The basis functions might also involve some nonlinear parameters  $\alpha_1, \dots, \alpha_p$ . The problem is separable if it involves both linear and nonlinear parameters.

$$y(t) \approx \beta_1 \phi_1(t, a) + \dots + \beta_n \phi_n(t, a)$$

The elements of the design matrix depend upon both  $t$  and  $a$ .

$$x_{i,j} = \phi_j(t_i, a)$$

Some common models are:

**Straight line:** If the model is also linear in  $t$ , it is a straight line.

$$y(t) \approx \beta_1 t + \beta_2$$

**Polynomials:** The coefficients  $\beta_j$  appear linearly.

$$\phi_j(t) = t^{n-j}, j = 1, \dots, n$$

$$y(t) \approx \beta_1 t^{n-1} + \dots + \beta_{n-1} t + \beta_n$$

**Rational functions:** The coefficients in the numerator appear linearly; the coefficients in the denominator appear nonlinearly.

$$\phi_j(t) = \frac{t^{n-j}}{\alpha_1 t^{n-1} + \dots + \alpha_{n-1} t + \alpha_n}$$

$$y(t) \approx \frac{\beta_1 t^{n-1} + \dots + \beta_{n-1} t + \beta_n}{\alpha_1 t^{n-1} + \dots + \alpha_{n-1} t + \alpha_n}$$

**Exponentials:** The decay rates,  $\lambda_j$ , appear nonlinearly.

$$\phi_j(t) = e^{-\lambda_j t}$$

$$y(t) \approx \beta_1 e^{-\lambda_1 t} + \dots + \beta_n e^{-\lambda_n t}$$

**Log-linear:** If there is only one exponential, taking logs makes the model linear, but changes the fit criterion.

$$y(t) \approx K e^{\lambda t} \quad \text{with } \beta_1 = \lambda_1, \beta_2 = \log K$$

$$\log y \approx \beta_1 t + \beta_2$$

**Gaussians:** The means and variances appear nonlinearly.

$$\phi_j(t) = e^{-\left(\frac{t-\mu_j}{\sigma_j}\right)^2}$$

$$y(t) \approx \beta_1 e^{-\left(\frac{t-\mu_1}{\sigma_1}\right)^2} + \dots + \beta_n e^{-\left(\frac{t-\mu_n}{\sigma_n}\right)^2}$$

## 4.2 Norm – Based solutions

The **residuals** are the differences between the observations and the model,

$$r_i = y_i - \sum_1^n \beta_j \phi_j(t_i, a), \quad i=1, \dots, m$$

Or, in matrix-vector notation,

$$r = y - X(a)\beta$$

We want to find the  $a$ 's and  $\beta$ 's that make the **residuals** as small as possible. There are several possibilities:

- **Interpolation:** If the number of parameters is equal to the number of observations, we might be able to make the residuals zero. For linear problems, this will mean that  $m = n$  and that the design matrix  $X$  is square. If  $X$  is nonsingular, the  $\beta$ 's are the solution to a square system of linear equations.

$$\beta = X \setminus y$$

- **Least squares:** Minimize the sum of the squares of the residuals.

$$\|r\|^2 = \sum_1^m r_i^2$$

- **Weighted least squares:** If some observations are more important or more accurate than others, then we might associate different weights,  $w_j$ , with different observations and minimize

$$\|r\|_w^2 = \sum_1^m w_i r_i^2$$

For example, if the error in the  $i$ th observation is approximately  $e_i$ , then we choose:  $w_i = 1/e_i$ .

- **One-norm:** Minimize the sum of the absolute values of the residuals.

$$\|r\|_1 = \sum_1^m |r_i|$$

This problem can be reformulated as a **linear programming** problem, but it is computationally more difficult than least squares. The resulting parameters are less sensitive to the presence of spurious data points or **outliers**.

- **Infinity-norm:** Minimize the largest residual.

$$\|r\|_\infty = \max_i |r_i|$$

This is also known as a Chebyshev fit and can be reformulated as a **linear programming** problem. Chebyshev fits are frequently used in the design of digital filters and in the development of approximations for use in mathematical function libraries.

### 4.3 Least Squares approach

$$\min_{\beta} \|X\beta - y\|^2$$

A theoretical approach to solving the over determined system begins by multiplying both sides by  $X^T$ . This reduces the system to a square, n-by-n system known as the normal equations:

$$X^T X \beta = X^T y$$

If there are thousands of observations and only a few parameters matrix  $X$  is quite large, but the matrix  $X^T X$  is small. So,  $y$  is projected onto the space spanned by the columns of  $X$ . Continuing with this theoretical approach, if the basis functions are independent, then  $X^T X$  is nonsingular and

$$\beta = (X^T X)^{-1} X^T y$$

However there are several undesirable aspects to this theoretical approach. **Using a matrix inverse to solve a system of equations is more work and less accurate than solving the system by Gaussian elimination.** But. More importantly, the normal equations are always more badly conditioned than the original over determined system. In fact, the condition number is squared.

$$k(X^T X) = k(X)^2$$

With finite precision computation, the normal equations can actually become singular, and  $(X^T X)^{-1}$  nonexistent, even though the columns of  $X$  are independent.

Here is an extreme example:

$$X = \begin{pmatrix} 1 & 1 \\ \delta & 0 \\ 0 & \delta \end{pmatrix}$$

If  $\delta$  is small, but nonzero, the two columns of  $X$  are nearly parallel, but are still linearly independent. The normal equations make the situation worse.

$$X^T X = \begin{pmatrix} 1 + \delta^2 & 1 \\ 1 & 1 + \delta^2 \end{pmatrix}$$

If  $|\delta| < 10^{-8}$ , the matrix  $X^T X$  computed with double-precision floating-point arithmetic is exactly singular and the inverse required in the classic textbook formula does not exist.

One popular solution of the over-determined least squares solution is the pseudoinverse. The definition of the pseudoinverse makes use of the Frobenius norm of a matrix

$$\|A\|_F = \left( \sum_i \sum_j a_{i,j}^2 \right)^{1/2}$$

The Moore-Penrose pseudoinverse generalizes and extends the usual matrix inverse. The pseudoinverse is denoted by a dagger superscript,

$$Z = X^\dagger$$

If  $X$  is square and non-singular, then the inverse are the same,

$$X^\dagger = X^{-1}$$

If  $X$  is  $m$ -by- $n$  with  $m > n$  and  $X$  has full rank, then its pseudoinverse is the matrix involved in the normal equations,

$$X^\dagger = (X^T X)^{-1} X^T$$

The pseudoinverse has some, but not all, of the properties of the ordinary inverse.  $X^\dagger$  is a left inverse because

$$X^\dagger X = (X^T X)^{-1} X^T X = I$$

is the  $n$ -by- $n$  identity. But  $X^\dagger$  is not a right inverse because

$$X X^\dagger = X (X^T X)^{-1} X^T$$

only has rank  $n$  and so cannot be the  $m$ -by- $m$  identity.

The pseudoinverse does get as close to a right inverse as possible in the sense that out of all the matrices  $Z$  that minimize

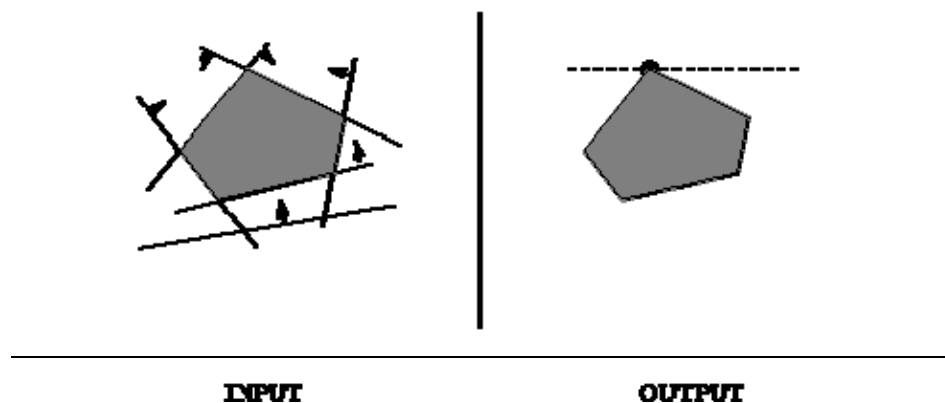
$$\|XZ - I\|_F$$

$Z = X^\dagger$  also minimizes

$$\|Z\|_F$$

It turns out these minimization properties also define a unique pseudoinverse even if  $X$  is rank deficient.

#### 4.4 Linear Programming for least absolute error approaches



Linear programming is the most important problem in mathematical optimization and operations research. Applications include:

- **Resource allocation** - We seek to invest a given amount of money so as to maximize our return. Our possible options, payoffs, and expenses can usually be expressed as a system of linear inequalities, such that we seek to maximize our possible profit given the various constraints. Very large linear programming problems are routinely solved by airlines and other corporations.
- **Approximating the solution of inconsistent equations** - A set of  $m$  linear equations on  $n$  variables  $x_i$ ,  $1 \leq i \leq n$ , is overdetermined if  $m > n$ . Such overdetermined systems are often *inconsistent*, meaning that no assignment of variables simultaneously solves all the equations. To find the variable assignment that best fits the equations, we can replace each variable  $x_i$  by  $x_i + e_i$  and solve the new system as a linear program, minimizing the sum of the error terms.
- **Graph algorithms** - Many of the standard graph problems described in this book, such as shortest paths, bipartite matching, and network flow, can all be solved as special cases of linear programming. Most of the rest, including traveling

salesman, set cover, and knapsack, can be solved using integer linear programming.

The standard algorithm for linear programming is called the **simplex method**. Each constraint in a linear programming problem acts like a knife that carves away a region from the space of possible solutions. We seek the point within the remaining region that maximizes (or minimizes)  $f(X)$ . By appropriately rotating the solution space, the optimal point can always be made to be the highest point in the region. Since the region (simplex) formed by the intersection of a set of linear constraints is convex, we can find the highest point by starting from any vertex of the region and walking to a higher neighboring vertex. When there is no higher neighbor, we are at the highest point. While the basic simplex algorithm is not too difficult to program, there is a considerable art to producing an efficient implementation capable of solving large linear programs. For example, large programs tend to be sparse (meaning that most inequalities use few variables), so sophisticated data structures must be used. There are issues of numerical stability and robustness, as well as which neighbor we should walk to next (so called pivoting rules). Finally, there exist sophisticated **interior-point** methods, which cut through the interior of the simplex instead of walking along the outside, that beat simplex in many applications.

#### 4.5 Investigation of Assumptions

Least squares solution is optimal in case that the residuals obey a Gaussian distribution. If this is not the case, the assumptions of the least squares formulation are not valid, leading to suboptimal formulations of the problem. In such cases other norm formulations could lead to more robust results. In our problem we investigated the distribution of residuals from the least square formulation and found them to deviate from the Gaussian distribution. For this reason we also tested the least absolute error formulation and compared results. The algorithm that is selected and implemented, in terms of this kind of formulation, is the Nelder – Mead algorithm and is presented in chapter 7.

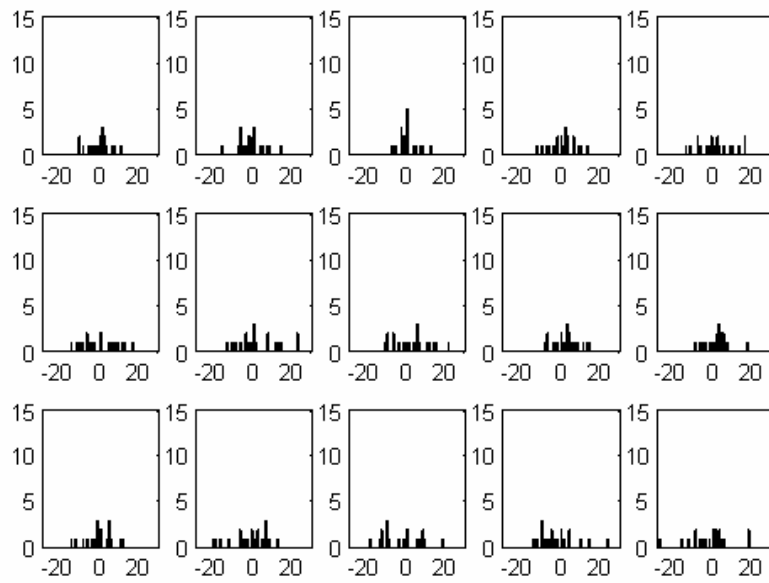
The following figures present the histograms of residuals in the x and y directions, respectively. These distributions appear more as uniform than as normal. In essence, the least squares approach forces the error at individual points to a large area around the optimal (zero).

Three cases were, firstly, investigated for the comparison of the alignment of the three algorithms. These are:

1. 'Normal Case' – This is the case that all the landmarks are well fixed around the boundaries of the vertebrae with the numbering to begin from the up-left corner.
2. 'Random Case' – This is the case that the numbering of the landmarks begins from other part of the vertebrae than the previous (up-left corner).
3. 'Improper Case' – This is the case that the landmarks are not well fixed around the boundaries of the vertebrae.

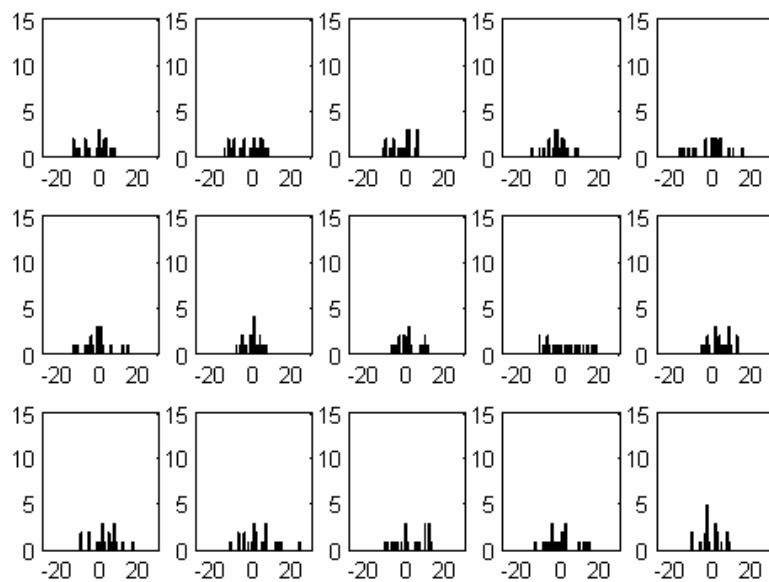


**X histograms - ICP**



**Figure 4.1 X- histograms ICP**

**Y histograms - ICP**



**Figure 4.2 Y histograms ICP**

X histograms - ACM

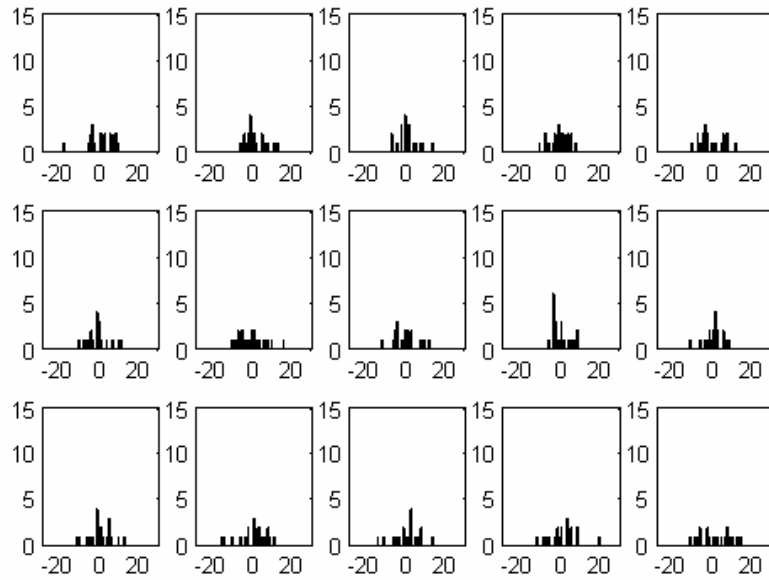


Figure 4.3 X-histograms ASM

Y histograms - ACM

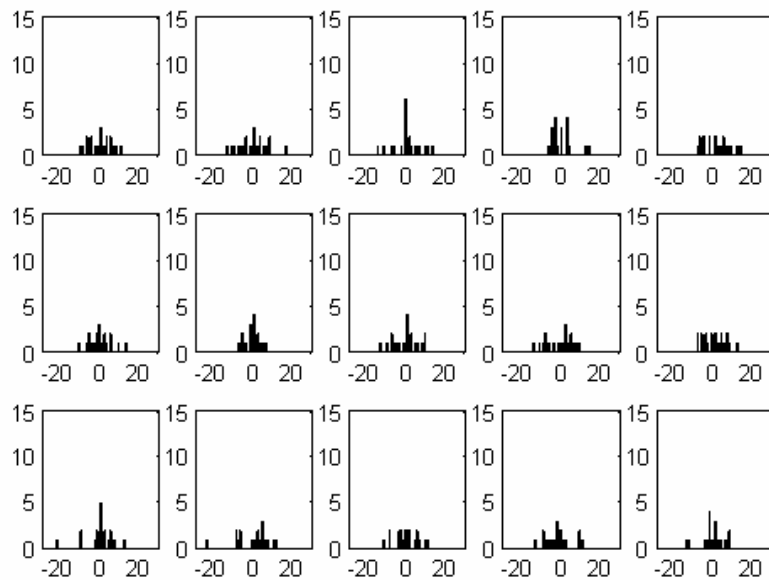


Figure 4.4 Y-histograms ASM

- The residuals are not obeying a Gaussian distribution, so we lead in another way of solving the error function that is discussed above. This is the least absolute error formulation. The algorithm that is selected and implemented, in terms of this kind of formulation, is the Nelder – Mead algorithm and is presented in chapter 7, the results are compared.

## 5. Active Shape Models (ASM) – ‘Smart Snakes’

### 5.1 Introduction and Related work

Accurate and precise segmentation of anatomical structures is an essential component in many biomedical image analysis procedures. Delineation of bone structures, in digital X-ray images, is a prerequisite in many orthopedic examinations. Since manual delineation of these anatomical structures is very tedious and time consuming, fast and accurate computer-aided segmentation methods are required.

Traditional low-level segmentation techniques, relying only on pixel intensity homogeneity or discontinuity, often generate incorrect object delineations, especially in digital X-ray images with typically low contrast between overlapping structures. High-level delineation techniques are additional a priori shape information to constrain the search for the object boundary. In contrast to many other computer vision application areas where the objects to be recognized can be described using simple models with a small number of parameters, anatomical structures in medical images are complex shaped objects with inter-individual shape variability. Deformable models such as popular “Active Contour approaches (snakes)” of Kass et al. [133], are capable of modeling the significant shape complexity by continuously deformable curve. These models, however, are free to take almost any arbitrary smooth shape, without being constrained on the overall shape.

**A single anatomical structure over different individuals is characterized by a generic shape with small deviations around it, which can be learned from a training set of images of such objects.** This a priori knowledge can be used to limit the deformations in order to obtain acceptable shapes. [134,135]

Yuille et al. [describe hand built models consisting of various geometric parts designed to represent image features; they also describe methods for adjusting their models to best fit an image. Unfortunately both the models and the optimization techniques have to be individually tailored for each application.

Staib and Duncan [136] use a Fourier shape model, representing a closed boundary as a sum of trigonometric functions of various frequencies. They also used a form of iterative energy minimization technique to fit a model to an image. However, using trigonometric functions does not always provide an appropriate basis for capturing shape variability, and is limited to closed boundaries.

Lowe[137] describes a technique for fitting projections of three-dimensional parameterized models to two dimensional images by iteratively minimizing the distance between lines in the projected model and those in the image.

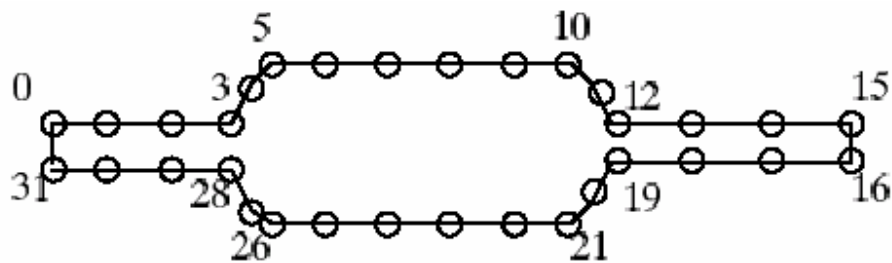
The method that is proposed in [138],[139] and implemented in this thesis, is a method of building flexible models by representing the objects as sets of labeled points and examining the statistics of their co-ordinates over a number of training shapes- point Distribution Models (PDMs). In [138] it is described an iterative optimization scheme for PDMs allowing initial estimates of the pose, scale and shape of an object in an image to be refined. The linear nature of the model leads to simple mathematics allowing rapid execution. Because the models can accurately represent the models of shape variation of

a class of objects they are compact and prevent ‘implausible’ shapes from occurring. Since PDMs can represent a wide variety of objects the same modeling and refinement framework can be applied in many different applications. Given an estimate of the position, orientation, scale and shape parameters of an example in an image, adjustments to the parameters can be calculated which give a better fit to the image. Suggested movements are calculated at each model point, giving the displacement required to get to a better location. These movements are transformed to suggested adjustments of the parameters, giving a better overall fit of the model instance to the data. By applying limits to the ranges of the parameters it can be ensured that the shape of the instance remains similar to the original training examples. Enforcing these limits applies global shape constraints, allowing only certain deformations to occur. Because the models attempt to deform to better fit the data, but only in ways which are consistent with the shapes found in the training set that is called ‘Active Shape Models’ or ‘Smart Snakes’.

## 5.2 The Point Distribution Model

The Point Distribution Model (PDM) is a way of representing a class of shapes using a flexible model of the position of labeled points placed on examples of the class. The points can represent the boundary or significant internal locations of an object (Figure 5).

In [138], [139] the algorithm is applied to resistors (Figure 5).



**Figure 5.1** 32 point model of the boundary of a resistor

### 5.3 Labeling the Training Set

In order to model a shape, it is represented by a set of points. For the vertebrae it have been chosen to place landmarks around the boundary, as shown in Figure 6. The labeling of the points is important, each labeled point represents its boundary. For instance in the vertebrae model points 0 and 45 always represent the ends of the vertebrae, points 0,14,25,37 represents the angles of the vertebrae. The method works by modeling how different labeled points tend to move together as the shape varies. If the labeling is incorrect, with a particular landmark placed at different sites on each training shape, the method will fail to capture shape variability.

MousePos is a program especially designed for the manually placing of landmarks around the vertebrae in the Digital X-Rays images and is designed with the use of Microsoft C++ program language. It is used in both implementations of ASM and ICP algorithms.

### 5.4 Capturing the statistics of a set of aligned shapes

If a set of aligned shapes is available the mean shape and variability can be found. The mean shape is calculate as follows:

$$\bar{x} = \frac{1}{N_s} \sum_{i=1}^{N_s} x_i$$

The modes of variation, the ways in which the points of the shape tend to move together, can be found by applying principal component analysis to the deviations from the mean:

$$dx_i = x_i - \bar{x}$$

The 2n x 2n covariance matrix S can be calculated:

$$S = \frac{1}{N_x} \sum_{i=1}^{N_x} dx_i dx_i^T$$

The modes of variation of the points of the shape are described by the unit eigenvectors of S,  $p_i$  (i=1 to 2n) such that

$$Sp_i = \lambda_i p_i$$

(where  $\lambda_i$  is the i'th eigenvalue of S,  $\lambda_i \geq \lambda_{i+1}$  ,  $p_i^T p_i = 1$ )

The eigenvectors of the covariance matrix corresponding to the largest eigenvalues describe the most significant modes of variation in the variables used to derive the covariance matrix, and the proportion of the total variance explained by each eigenvector is equal to the corresponding eigenvalue.

Any shape in the training set can be approximated using the mean shape and a weighted sum of these deviations obtained from the first t modes

$$x = \bar{x} + Pb$$

where  $P = (p_1 p_2 \dots p_t)$  is the matrix of the first  $t$  eigenvectors

$b = (b_1 b_2 \dots b_t)^T$  is a vector of weights for each eigenvector

the eigenvectors are orthogonal,  $P^T P = I$  so

$$b = P^T (x - \bar{x})$$

## 5.5 Aligning the training set

The modeling works by examining the statistics of the co-ordinates of the labeled points over the training set. In order to be able to compare equivalent points from different shapes, they must be aligned in the same way with respect to a set of axes. If they are not, they can not be compared like to like and any statistics derived would be meaningless. The required alignment is achieved by scaling, rotating and translating the training shapes so that they correspond as closely as possible. The aim is to minimize a weighted sum of squares of distances between equivalent points on different shapes. This is a form of Generalized Procrustes Analysis [140].

We will first consider aligning a pair of shapes. Let  $x_i$  be a vector describing the  $n$  points of the  $i^{th}$  shape in the set;

$$x_i = (x_{i0}, y_{i0}, x_{i1}, y_{i1}, \dots, x_{ik}, y_{ik}, \dots, x_{in-1}, y_{in-1})^T$$

Let  $M_j[x_j]$  be a rotation by  $\theta_j$  and a scaling by  $s_j$ . Given two similar shapes  $x_i$  and  $x_j$  we can choose  $\theta_j$ ,  $s_j$  and a translation  $(t_x, t_y)_j$  mapping  $x_i$  onto  $M_j[x_j]$  so as to minimize the weighted sum

$$E_j = (x_i - M_j(x_j))^T W (x_i - M_j(x_j)) \quad (1)$$

where

$$M_j \begin{pmatrix} x_{jk} \\ y_{jk} \end{pmatrix} = \begin{pmatrix} (s_j \cos \theta) x_{jk} - (s_j \sin \theta) y_{jk} + t_{jx} \\ (s_j \sin \theta) x_{jk} + (s_j \cos \theta) y_{jk} + t_{jy} \end{pmatrix} \quad (2)$$

For 2 shapes (1) is becoming

$$E_j = (x_1 - M(x_2))^T W (x_1 - M(x_2))$$

We can write:

$$a_x = s \cos \theta \text{ and } a_y = s \sin \theta$$

then the least squares approach leads to a set of four linear equations:

Where

$$X_i = \sum_{k=0}^{n-1} w_k x_{jk} \quad Y_i = \sum_{k=0}^{n-1} w_k y_{jk}$$

$$Z = \sum_{k=0}^{n-1} w_k (x_{2k}^2 + y_{2k}^2) \quad W = \sum_{k=0}^{n-1} w_k$$

$$C_1 = \sum_{k=0}^{n-1} w_k (x_{1k} x_{2k} + y_{1k} y_{2k})$$

$$C_{21} = \sum_{k=0}^{n-1} w_k (y_{1k} x_{2k} + x_{1k} y_{2k})$$

The weights can be chosen to give more significance to these points which tend to be most ‘stable’ over the set-the ones which move about least with respect to the other points in a shape. The weighted sum that is used is defined as follows: let  $R_{kl}$  be the distance over the set of shapes; it can be chosen a weight  $w_k$  for the  $k^{th}$  point using

$$w_k = \left( \sum_{l=0}^{n-1} V_{R_{kl}} \right)^{-1}$$

If a point tends to move around a lot with respect to other points in the shape, the sum of variances will be large, and a low weight will be given. If, however, a point tends to remain fixed with respect to the others, the sum of variances will be small, a large weight will be given and matching such points in different shapes will be a priority.

In order to align all the shapes in a set it is used the algorithm as follows:

- 1) Rotate, scale and translate each of the shapes in the set to align to the first shape
  - Repeat
  - 2) Calculate the mean of the transformed shapes
  - 3) Either
    - a) Adjust the mean to a default scale, orientation and origin
    - b) Rotate, scale and translate the mean to align to the first shape
  - 4) Rotate, scale and translate each of the shapes again to match to the adjusted mean.
- Until convergence.

Stage 3 inside the iteration loop is required to renormalize the mean. Without this the algorithm will not convergence- there are in effect  $4(N_s - 1)$  constraints on  $4N_s$  variables  $\theta, s, t_x, t_y$  for each shape- the mean will shrink, rotate or slide off to identify. Constraints on the pose and scale of the mean allow the equations to have a unique solution. Either the mean is scaled, rotated and translated so it matches the first shape, or an arbitrary default setting can be used, such as choosing an origin at its centre of gravity, an orientation so that a particular part of the shape is at the top and a scale so that the distance between two points is one unit.

The convergence condition can be tested by examining the average difference between the transformations required to align each shape to the recalculated mean and its identity transformation. The method converges to the same result independent of which shape is aligned to in the first stage.



And  $W$  is a diagonal matrix of weights for each point

## 5.6 Examples using Active Shape Models

### 5.6.1 Finding resistors with an ASM

T. F Cootes and C. J. Taylor have used this technique successfully to find resistors. They have constructed a Point Distribution Model of a resistor representing its boundary using 32 points (Figure 5). Figure 7 shows an image of part of a printed circuit board with the resistor boundary model superimposed as it iterates towards the boundary of a component in the image. They interpolate an additional 32 points, one between each pair of model points around the boundary, and calculate adjustments to each point by finding the strongest edge along profiles 20 pixels long centered at each point. They use a model with 5 degrees of freedom. Each iteration takes about 0.025 seconds on a Sun Sparc Workstation.

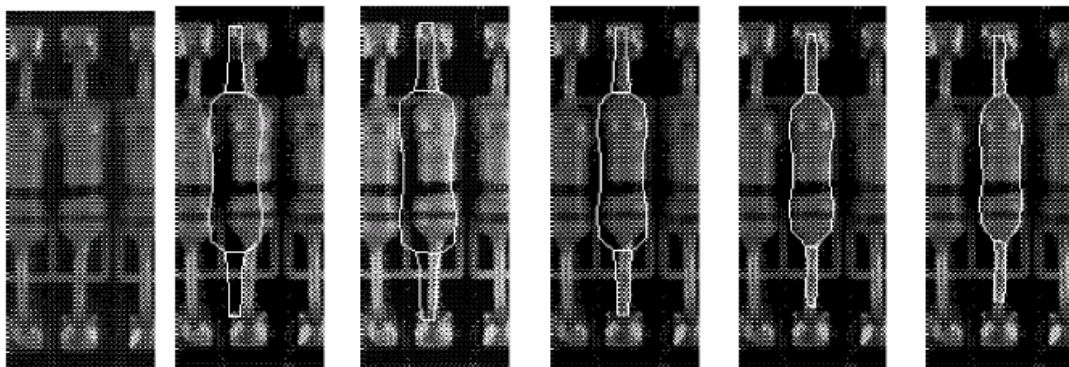


Figure 5.2 Resistors

(a) (b) (c) (d) (e) (f)  
Figure 5.2: Section of printed circuit board with resistor model superimposed, showing its initial position and its location after 30, 60, 90 and 120 iterations.

- (a) Original image
- (b) Initial Position
- (c) After 30 iterations
- (d) After 60 iterations
- (e) After 90 iterations

(f) After 120 iterations

The ends of the wires are not found correctly since they are not well not defined.

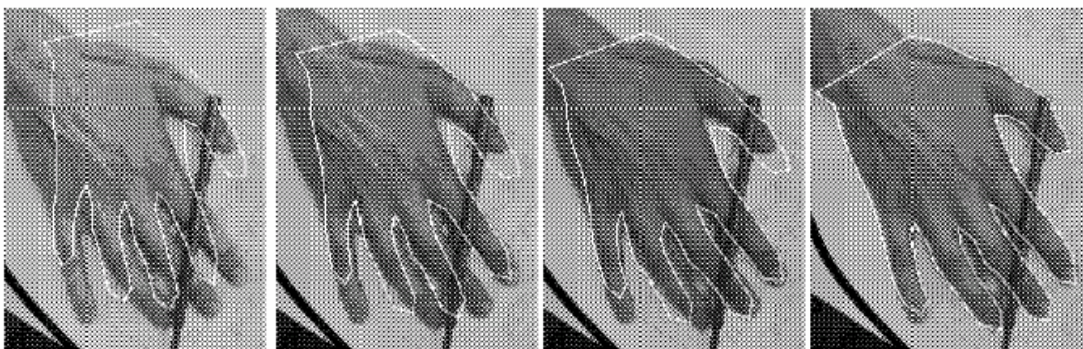
The resistor shapes were aligned, arranging the mean shape to be horizontal and scaling so the average distance of each point of the mean from its center of gravity is one unit. The most significant eigenvalues of the covariance matrix derived are shown in the Table 2 above:

Eigenvalue	$\lambda_i$	$\frac{\lambda_i}{\lambda_T} \times 100\%$	$\sqrt{\lambda_i}$
$\lambda_1$	0.207	66%	0.46
$\lambda_2$	0.026	8%	0.16
$\lambda_3$	0.017	5%	0.13
$\lambda_4$	0.013	4%	0.11
$\lambda_5$	0.010	3%	0.10
$\lambda_6$	0.008	3%	0.09

**Table 5.1**Eigenvalues of the covariance matrix derived from a set of resistor shapes.

### 5.6.2 Finding hands with an ASM

Another Point Distribution Model is constructed of a hand representing the boundary using 72 points.



(a) (b) (c) (d)

**Figure 5.3** Hands

Figure 4.3 : Image of a hand with hand model superimposed, showing its initial position and its location after 100, 200 and 350 iterations.

- (a) Initial Position
- (b) 100 iterations
- (c) 200 iterations
- (d) 350 iterations

Figure 5.3 shows an image of a hand and an example of the model iterating towards it. The adjustments are calculated to each point by finding the strongest edge on a profile 35 pixels long centered on the point. The shape model has 8 degrees of freedom, and each iteration takes about 0.03 seconds on a Sun Sparc Workstation. The result demonstrates that the method can deal with limited occlusion.

As in the previous example the method works reliably, given a reasonable starting approximation. The example shows that the method is tolerant to quite serious errors in the starting approximation, though this depends on the amount of clutter in the image.

#### 4.7 How it works

The ASM algorithm is implemented with a main program which name is solve\_alignment and a sub program which name is solve\_system in MATLAB 7.1.

- In the first 21 lines the main program reads the landmarks of the shapes. The points can be 45 or user defined.
- After this plots the shapes
- Then plots the shapes as they are at the beginning.
- In the next step calls the sub\_program: solve\_system to compute
- The solve\_system finds the solution of the C matrix. This matrix is a 4x4 matrix and it can be solved mathematically with the Cramer law.
- We return to the main program and calculate the mean
- Then we align the mean to first shape
- Align shapes to the mean
- Finally, plots the results.

**The ASM works well with N landmarks fixed for every shape. When one shape has less points from the others the program can not be executed.**

## 5.8 Results and figures of the ASM algorithm

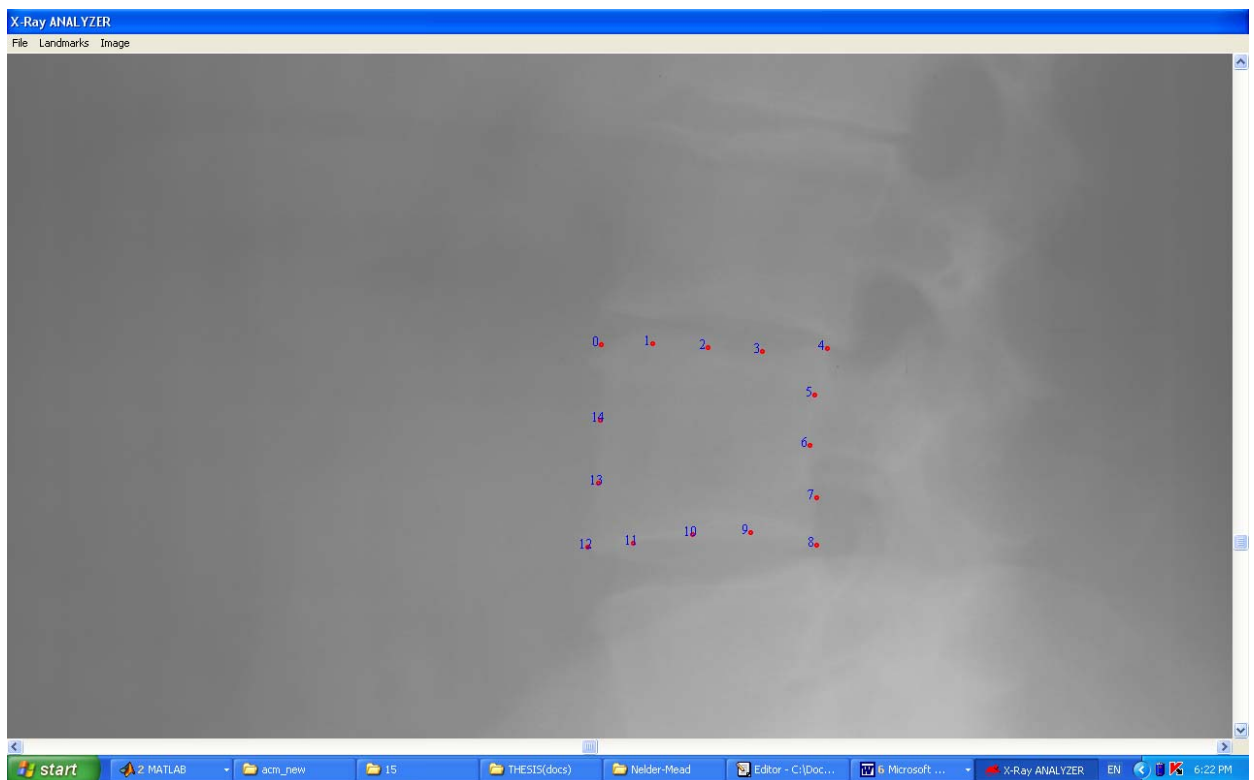
The ASM algorithm is implemented in MATLAB 7.1 with the use of MousePos. MousePos is a especially designed program in C programming language for the manually fixing of the landmarks in the X-ray digital images. These images are from the NHANES II database. These images are available in the website:

- In **Figure 5.4** we can see one X-ray digital image from the NHANES II database. The image is of the lumbar part of the human body.

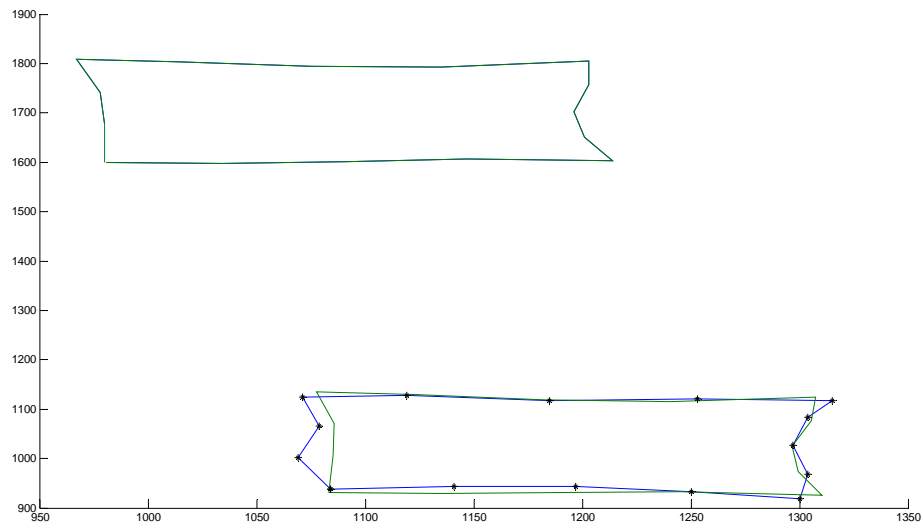


**Figure 5.4** Mouse Pos

- In **Figure 5.5** we can see a screen of the MousePos program with 15 manually fixed landmarks at the boundary of the lumbar:



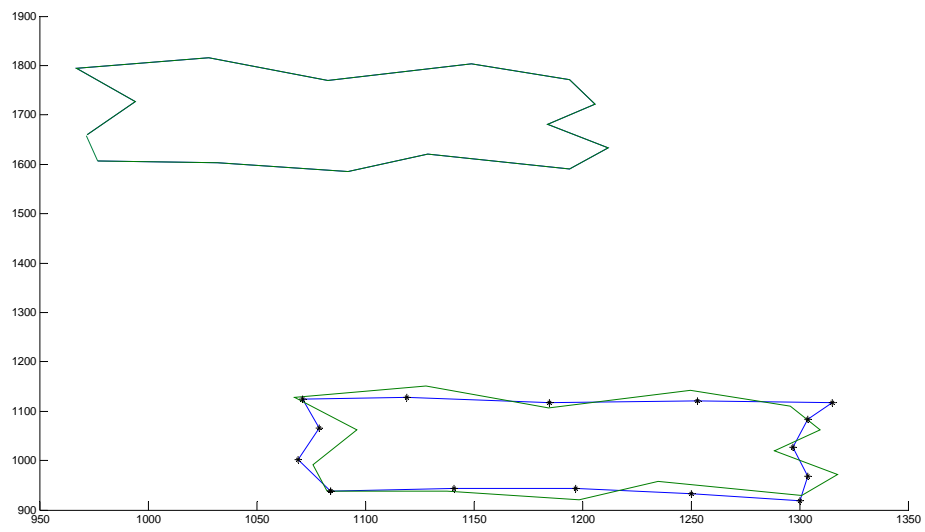
**Figure 5.5** ('L04531' image)



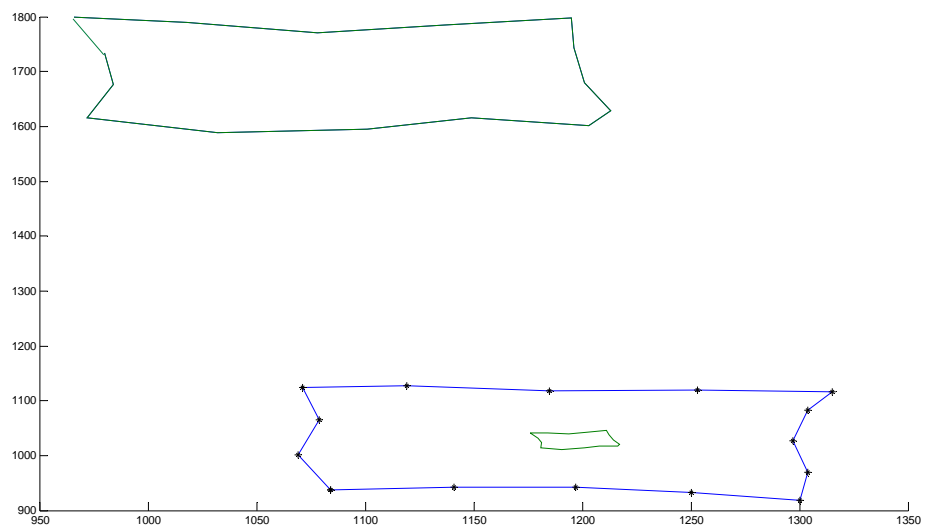
**Figure 5.6** ASM implementation

- In **Figure 5.6** we present the implementation of ASM algorithm, for the alignment of two shapes with 15 landmarks normal fixed, as in **Figure 5.5**, ('L04531' and 'L12966' images), the blue shape is the model, the green is the shape before and after the alignment in all ASM figures:

- In these figures we can see, that in all the cases the corresponding points have been located with accuracy. Even in the case that the shape is not well fixed the algorithm execute the alignment and in the third case where the 'to be aligned' shape was smaller than the 'model shape' ASM seems to align with precision. A disadvantage of this algorithm is that it fails to execute when the one the shape 'to be aligned' has less landmarks from the 'model shape'.



**Figure 5.7** X-ray digital image where the landmarks are not fixed properly



**Figure 5.8** The alignment when one of the shapes has random fixed points.

## 6. The Iterative Closest Point Algorithm (ICP)

### 6.1 Introduction and related work

Little work has been published in the area of registration (pose estimation, alignment, motion estimation) of 3-D free-form shapes. Most of the existing literature addressing global shape matching or registration have been addressed limited classes of shapes, namely,

- 1) polyhedral models,
- 2) piecewise- (super) quadric models [141], [142],
- 3) point sets with known correspondence.

Historically, free-form shape matching using 3-D data was done by Faugeras and his group at INRIA [143], where they demonstrated effective matching with Renault auto part (steering knuckle) in the early 1980's. This work popularized the use of quaternions for least squares registration of corresponding 3-D sets in the computer vision community. The alternative use of the singular value decomposition (SVD) algorithm [144], [145], [146] was not as widely known in this time frame. The primary limitation of this work was that it relied on the probable existence of reasonably large planar regions within a free-form shape.

Schwartz and Sharir [147] developed a solution to the free-form space curve matching problem without feature extraction in late 1985. They used a nonquaternion approach to computing the least squares rotation matrix. The method works well with reasonable quality curve data but has difficulty with very noisy curves because the method uses arclength sampling of the curves corresponding point sets.

Haralick et al [148] addressed the 3-D point-set pose estimation problem using robust methods combined with the least squares SVD registration approach, which provided a robust statistical alternative to the least squares quaternion or SVD point set matching. This algorithm is able to handle statistical outliers and could theoretically be substituted for our quaternion-based algorithm as long as the determinant of the orthogonal matrix is strictly a positive one.

Singular value decomposition (SVD) is a very powerful set of techniques dealing with sets of equations or matrices that are either singular or numerically very close to singular. SVD allows one to diagnose the problems in a given matrix and provides numerical answer as well. Any  $m \times n$  matrix  $a$  ( $m \geq n$ ) can be written as the product of an  $m \times n$  column-orthogonal matrix  $u$ , an  $n \times n$  diagonal matrix with positive or zero elements, and the transpose of an  $n \times n$  orthogonal matrix  $v$ :

where

$$a = u \cdot w \cdot v^T$$



$$W = \begin{bmatrix} w_1 & 0 & . & 0 & 0 \\ 0 & w_1 & . & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & . & w_{n-1} & 0 \\ 0 & 0 & . & 0 & w_n \end{bmatrix}$$

and

$$\begin{aligned} \hat{u} \cdot \hat{u} - \hat{v} \cdot \hat{v} &= 1 \\ \hat{v} \cdot \hat{v} &= 1 \\ w_1, w_2, \dots, w_{n-1}, w_n &\geq 0 \end{aligned}$$

The diagonal elements of matrix  $w$  are the singular values of matrix  $a$  and non-negative numbers.

Horn [149] derived an alternative formulation of Faugera's method [143] of least squares quaternion matching that uses the maximum eigenvalue of a 4x4 matrix instead of the maximum eigenvalue.

Szeliski [150] also describes a method for estimating motion sparse range data without correspondence between the points and without feature extraction. His primary goal was to create a method for estimating the motion of the observer between two range images frames on the same terrain. Given the set of the points from one, he applies a smoothness assumption to create a smoothing spline approximation of the points. Then, a conventional steepest descent algorithm is used to rotate and translate the second data set so that it minimizes the sum of the covariance-weighted  $z$  differences between the points and the surface. His approach is based on a regular  $xy$ -grid structure, and true 3-D point-to-surface distances are not computed. The steepest-descent approach is a slower alternative to reaching the local minima than in ICP algorithm described below.

Horn and Harris[151] also addressed the problem of estimating the exact rigid-body estimation of the observer given sequentially digitized range image frames of the same terrain. They describe a range rate constraint equation and an elevation rate constraint equation. The result is a noniterative least squares method that provides a six-degree-of-freedom motion estimate as long as the motion between frames of data is relatively small. This method is much quicker than the one proposed by Szeliski, but is not clear that this method generalizes to arbitrary rotations and translations of a shape.

Kamgar-Parsi et al. [152] also describe a method for the registration of multiple overlapping range images without distinctive feature extraction. This method works very well using the level sets of 2.5-D range data but is essentially restricted to the three degrees of freedom in the plane since the work was addressed toward piecing together terrain map data.

Li [153] addressed free-form surface matching with arbitrary rotations and translations. His method forms an attributed relational graph of fundamental surface regions for data and model shapes and then performs graph matching using an inexact approach that allows for variability in attributes as well as in graph adjacency relationships. This approach relies on extraction of derivative-based quantities.

The work of Gilbert and Foo [154] and Gilbert et al. [155] is related in that it addresses the computation of distance between two object shapes. Such methods could be the basis for similar shape matching techniques, like ICP.

## 5.2 Quaternion

Quaternion algebra is one possible way to represent 3 dimensional orientation, or other rotational quantity, associated with a solid 3D object.

### 6.2.1 Description

There are a number of ways to think about complex numbers:

- As a mathematical quantity and algebra which, if we do certain operations, allows us to calculate the results of rotations in 3 dimensions.
- It is similar to axis-angle except that real part is equal to  $\cos(\text{angle}/2)$  and the complex part is made up of the axis vector times  $\sin(\text{angle}/2)$ .
- Quaternion numbers are an extension of complex numbers with

Quaternions have 4 dimensions, one real dimension and 3 imaginary dimensions. Each of these imaginary dimensions has a unit value of the square root of -1, but they are different square roots of -1 all mutually perpendicular to each other, known as  $i$ ,  $j$  and  $k$ . So a quaternion can be represented as follows:

$$a + i b + j c + k d$$

It may seem strange that there are 3 square roots of -1, but we are working in 4 dimensions so there are at least 3 ways to get round from +1 to -1. Here is an attempt to draw 4 dimensions in 2 dimensions:

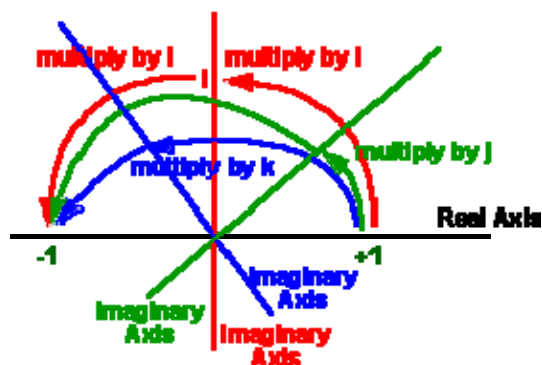


Figure 6.1 Quaternions

### 6.2.2 Practical uses of quaternions.

We can use quaternions to represent 3D rotations and calculate the result of applying one rotation and then applying a subsequent rotation. To calculate the result of rotation represented by  $q$  followed by  $p$  we have to apply the product  $q * p$ . As well as concatenating rotations we can also use quaternions to interpolate between rotations.

Quaternion multiplication and orthogonal matrix multiplication can both be used to represent rotation. If a quaternion is represented by  $qw + i qx + j qy + k qz$ , then the equivalent matrix, to represent the same rotation, is:

$1 - 2*qy^2 - 2*qz^2$	$2*qx*qy - 2*qz*qw$	$2*qx*qz + 2*qy*qw$
$2*qx*qy + 2*qz*qw$	$1 - 2*qx^2 - 2*qz^2$	$2*qy*qz - 2*qx*qw$
$2*qx*qz - 2*qy*qw$	$2*qy*qz + 2*qx*qw$	$1 - 2*qx^2 - 2*qy^2$

This assumes that the quaternion is normalized ( $sqw + sqx + sqy + sqz = 1$ ), if not it should be normalized before doing the conversion.

- The proof of this result follows:

To transform point  $P2$  to point  $P1$  we use

$$P2 = q * P1 * q'$$

we want to find the matrix  $[M]$  that will do the same transform.

$$P2 = [M] P1$$

So assume  $P1 = (0, x, y, z)$  then multiplying out gives:

$$q * P1 * q' = (-qx*x - qy*y - qz*z + i(qw*x + qy*z - qz*y) + j(qw*y - qx*z + qz*x) + k(qw*z + qx*y - qy*x)) * (qw - iqx - jqy - kqz)$$

$$\begin{aligned} &= (-qx*x - qy*y - qz*z)*qw - (qw*x + qy*z - qz*y)*-qx - (qw*y - qx*z + qz*x)*-qy - \\ & (qw*z + qx*y - qy*x)*-qz \\ &+ i((qw*x + qy*z - qz*y)*qw + (-qx*x - qy*y - qz*z)*-qx + (qw*y - qx*z + qz*x)*-qz - \\ & (qw*z + qx*y - qy*x)*-qy) \\ &+ j((-qx*x - qy*y - qz*z)*-qy - (qw*x + qy*z - qz*y)*-qz + (qw*y - qx*z + qz*x)*qw + \\ & (qw*z + qx*y - qy*x)*-qx) \\ &+ k((-qx*x - qy*y - qz*z)*-qz + (qw*x + qy*z - qz*y)*-qy - (qw*y - qx*z + qz*x)*-qx \\ &+ (qw*z + qx*y - qy*x)*qw) \end{aligned}$$

$$\begin{aligned}
&= -qx*qw*x - qy*qw*y - qz*qw*z + qw*qx*x + qy*qx*z - qz*qx*y + qw*qy*y - qx* \\
&qy*z + qz*qy*x + qw*qz*z + qx*qz*y - qy*qz*x \\
&+ i ( qw*qw*x + qy*qw*z - qz*qw*y + qx*qx*x + qy*qx*y + qz*qx*z - qw*qz*y + \\
&qx*qz*z - qz*qz*x + qw*qy*z + qx*qy*y - qy*qy*x \\
&+ j ( qx*qy*x + qy*qy*y + qz*qy*z + qw*qz*x + qy*qz*z - qz*qz*y + qw*qw*y - \\
&qx*qw*z + qz*qw*x - qw*qx*z - qx*qx*y + qy*qx*x ) \\
&+ k ( qx*qz*x + qy*qz*y + qz*qz*z - qw*qy*x - qy*qy*z + qz*qy*y + qw*qx*y - \\
&qx*qx*z + qz*qx*x + qw*qw*z + qx*qw*y - qy*qw*x )
\end{aligned}$$

grouping the x,y and z terms and putting them in a matrix gives:

$qw*qw + qx*qx - qz*qz - qy*qy$	$- qz*qw + qy*qx - qw*qz + qx*qy$	$qy*qw + qz*qx + qx*qz + qw*qy$
$qx*qy + qw*qz + qz*qw + qy*qx$	$qy*qy - qz*qz + qw*qw - qx*qx$	$qz*qy + qy*qz - qx*qw - qw*qx$
$qx*qz - qw*qy + qz*qx - qy*qw$	$qy*qz + qz*qy + qw*qx + qx*qw$	$qz*qz - qy*qy - qx*qx + qw*qw$

since  $qw*qw + qx*qx + qz*qz + qy*qy = 1$  this gives

$1 - 2*qz*qz - 2*qy*qy$	$- 2*qz*qw + 2*qy*qx$	$2*qy*qw + 2*qz*qx$
$2*qx*qy + 2*qw*qz$	$1 - 2*qz*qz - 2*qx*qx$	$2*qz*qy - 2*qx*qw$
$2*qx*qz - 2*qw*qy$	$2*qy*qz + 2*qw*qx$	$1 - 2*qy*qy - 2*qx*qx$

### 6.3 The ICP algorithm

- The Iterative Closest Point (ICP) algorithm proposes a solution to the registration Problem: Given a “model” two-dimensional (2-D) shape and a “data” (2-D) shape, estimate the optimal translation and rotation that register the two shapes by minimizing the mean square distance between them.

- The ICP algorithm is a quaternion-based algorithm. The unit quaternion is a four vector

$\vec{q}_R = [q_0 q_1 q_2 q_3]^T$ , where  $q_0 \geq 0$  and  $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$ . The 3x3 rotation matrix generated by a unit quaternion is:

$$R = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix} \quad (1)$$

- If P is a measured data point set to be aligned with a model set X, where  $N_X = N_P$  and where each point of  $P = \{\vec{p}_i\}$  corresponds to the point  $X = \{\vec{x}_i\}$  with the same index. The mean square function to be minimized is:

$$f(\vec{q}) = \frac{1}{N_P} \sum_{i=1}^{N_P} \left\| \vec{x}_i - R(\vec{q}_R) \vec{p}_i - \vec{q}_T \right\|^2 \quad (2)$$

- The “center of mass”  $\vec{\mu}_P$  of the point set P and the “center of mass”  $\vec{\mu}_X$  for the X point set are given by:

$$\vec{\mu}_P = \frac{1}{N_P} \sum_{i=1}^{N_P} \vec{p}_i \quad (3) \text{ and } \vec{\mu}_X = \frac{1}{N_X} \sum_{i=1}^{N_X} \vec{x}_i \quad (4)$$

- The cross-covariance matrix  $\Sigma_{PX}$  of the sets P and X is given by:

$$\Sigma_{PX} = \frac{1}{N_P} \sum_{i=1}^{N_P} [(\vec{p}_i - \vec{\mu}_P)(\vec{x}_i - \vec{\mu}_X)^T] = \frac{1}{N_P} \sum_{i=1}^{N_P} [\vec{p}_i \vec{x}_i^T] - \vec{\mu}_P \vec{\mu}_X^T \quad (5)$$

- The components of the anti-symmetric matrix  $A_{ij} = (\Sigma_{PX} - \Sigma_{PX}^T)_{ij}$  are used to form the column vector  $\Delta = [A_{23} \quad A_{31} \quad A_{12}]^T$  (6). With this vector we can form the symmetric 4x4 matrix  $Q(\Sigma_{PX})$ .

$$Q(\Sigma_{PX}) = \begin{bmatrix} tr(\Sigma_{PX}) & \Delta^T \\ \Delta & \Sigma_{PX} + \Sigma_{PX}^T - tr(\Sigma_{PX})I_3 \end{bmatrix} \quad (7)$$

where:  $tr(\Sigma_{PX})$  is the trace of the cross-covariance matrix of the sets P and X and  $I_3$  is the 3x3 identity matrix.

- The optimal rotation is selected as the unit eigenvector

$$\vec{q} = [q_0 \quad q_1 \quad q_2 \quad q_3]^t \quad (8)$$

Corresponding to the maximum eigenvalue of the matrix  $Q(\Sigma_{px})$ .

- The optimal translation vector is given by

$$\vec{q}_T = \vec{\mu}_x - R(\vec{q}_R) \vec{\mu}_p \quad (9)$$

- The least squares quaternion operation is  $O(N_p)$  and is denoted as

$$(\vec{q}, d_{ms}) = Q(P, X) \quad (10)$$

where  $d_{ms}$  is the mean square point matching error. The notation  $\vec{q}(P)$  is used to denote the point set P after transformation by the registration vector  $\vec{q}$ .

- In the description of the algorithm, a data shape P is moved (registered, positioned) to be in best alignment with a model shape X. The data and the model shape may be represented in any of the allowable forms. For our purposes, the data shape must be decomposed into a point set if it is not already in point set form. The number of points in the data shape can be denoted as  $N_p$ . Let  $N_x$  be the number of points, line segments or triangles involved in the model shape. The distance metric d between an individual data point  $\vec{p}$  and a model shape X can be denoted:

$$d(\vec{p}, X) = \min_{\vec{x} \in X} \|\vec{x} - \vec{p}\|$$

The closest point in X that yields the minimum distance is denoted  $\vec{y}$  such that

$d(\vec{p}, \vec{y}) = d(\vec{p}, X)$  where  $\vec{y} \in X$ . Let Y denote the resulting set of closest points and let C be the closest point operator:

$$Y = C(P, X)$$

Given the resultant corresponding point set Y, the least squares registration is computed as described above:

$$(\vec{q}, d) = Q(P, Y)$$

- The positions of the data shape point set are then updated via

$$P = \vec{q}(P)$$

- The ICP algorithm can be stated:

- The point set P with  $N_p$  points from the data shape and the model shape X with  $N_x$ , are given.

- The iteration is initialized by setting  $P_0 = P$ ,  $\vec{q}_0 = [1, 0, 0, 0, 0, 0, 0]^T$  and  $k=0$ . The registration vectors are defined relative to the initial data set  $P_0$  so that the final registration represents the complete transformation. Here are the four steps that are applied until convergence within a tolerance  $\tau$ .
  1. Compute the closest points:  $Y_k = C(P_k, X)$
  2. Compute the registration:  $(\vec{q}_k, d_k) = Q(P_0, Y_k)$
  3. Apply the registration:  $P_{k+1} = \vec{q}_k(P_0)$
  4. Terminate the iteration when the change in mean-square error falls below a preset threshold  $\tau > 0$  specifying the desired precision of the registration:  $d_k - d_{k-1} < \tau$ .
- The ICP algorithm guarantees that a local minimum of a mean square objective function is found. This minimum depends on the initial registration and may not be the global one. The registration vector corresponds to the smallest of local minima is selected.
- The idea of the convergence theorem that rules the ICP algorithm is: 1) the least squares registration generically reduces the average distance between corresponding points during each iteration, whereas 2) the closest point determination generically reduces the distance for each point individually.

#### 6.4 Convergence Theorem for the ICP algorithm

A convergence theorem for the ICP algorithm can be stated and proved. The key ideas are that:

- 1) Least squares registration generically reduces the average distance between corresponding points during each iteration, whereas
- 2) The closest point determination generically reduces the distance for each point individually.

This individual distance reduction also reduces the average distance because the average of a set of smaller positive numbers is smaller.

**Theorem:** ‘The iterative closest point algorithm always converges monotonically to a local minimum with respect to the mean-square distance objective function.’

## 6.5 Initial states for Global Matching

Dealing with the initial state problem we can adopt the following definition of the first two moments of the distribution of geometry in P and X :

$$\vec{\mu}_p = E\left[\vec{p} \mid \vec{p} \in P\right], \vec{\mu}_x = E\left[\vec{r} \mid \vec{r} \in X\right], \vec{\Sigma}_p = E\left[(\vec{p} - \vec{\mu}_p)(\vec{p} - \vec{\mu}_p)^t \mid \vec{p} \in P\right], \text{ and}$$

$$\vec{\Sigma}_x = E\left[(\vec{r} - \vec{\mu}_x)(\vec{r} - \vec{\mu}_x)^t \mid \vec{r} \in X\right]$$

Where  $E[.]$  represents the sample expectation (averaging) operator. **If the point data set P covers a significant portion of the model shape X such that the condition**

$$a_1 \sqrt{\text{tr}(\vec{\Sigma}_x)} \leq \sqrt{\text{tr}(\vec{\Sigma}_p)} \leq \sqrt{\text{tr}(\vec{\Sigma}_x)}$$

Holds for a sufficiently large factor,  $a_1 = 1/\sqrt{2} \approx 0.71$  then it is found that it is generally not necessary to use multiple initial translations states, as long as enough rotation states are used. This factor  $a_1$  is the allowable occlusion percentage for global matching. The exact value of  $a_1$  could be computed for any given class of object shapes via exhaustive testing if that is desired.

There are two reasonable options for the initial translation state:

1) Apply the ICP algorithm directly to the point set P using multiple rotation states about its center of mass  $\vec{\mu}_p$ , or

2) transform it first so that the centers of mass of P and X coincide, and then apply ICP.

In our implementation it is applied the second method.

It is found no differences in the final registration results between 1) not translating and 2) pre-translating the point set when using an adequately large set of initial rotations. In fact, any translation state suffices because the ICP algorithm is very insensitive to the initial translation state when used for global matching. Also it is observed that pre-translating the data point set's center of mass generally saves a few iterations (e.g. 2 to 4) out of the usual 20 or so total..

A further simplification in the global shape matching algorithm can be accomplished for most generic shapes, where principal moments demonstrate some level of distinctness. Let  $p_x \geq p_y \geq p_z$  be the square roots of the eigenvalues of  $\vec{\Sigma}_p$ , and let  $r_x \geq r_y \geq r_z$ . If the following sets of conditions hold:

$$\begin{aligned} p_y &\leq a_2 p_x & p_z &\leq a_2 p_y \\ r_y &\leq a_2 r_x & r_z &\leq a_2 r_y \end{aligned}$$



For a specified value of  $a_2$ , e.g.,  $a_2 = 1/\sqrt{2} \approx 0.71$ , one can reliably match the basic shape structure of data and model using only the eigenvectors of the matrices  $\Sigma_p$  and  $\Sigma_x$ . Again, the exact value of  $a_2$  could be computed for any given set of objects and any given level of sensor noise via exhaustive testing if needed. In this case of eigenvalue distinctness, the identity transformation and the 180 degrees rotations about the eigenvector axes corresponding to  $r_x, r_x$  and  $r_z$  provide a very good set of only four initial rotations that yield the correct global minimum for a wide class of model shapes.

If two of the three eigenvalues are approximately equal but significant different from the third for both data and model shapes, the number of initial states need only be expanded for rotations about the nonambiguous axis, thereby reducing the total number of initial rotations states.

If neither of the above cases for global matching hold true (i.e.  $p_x \approx p_y \approx p_z$  and  $r_x \approx r_y \approx r_z$ ) then it must be used a fine sampling of quaternion states that cover the entire surface of the northern hemisphere of the unit 4 sphere uniformly.

## 6.6 How it works

- The ICP algorithm program was implemented in MATLAB 7.1 and analytically works as follows:

1. Recognizes the shapes from the selected 15 landmarks(Fig.1). The landmarks are fixed manually by the MousePos.
2. Computes the mean square for each shape, as it is described in equation (2)
3. Finds the “center of mass” for the point set P and for the point set X, by equation (3),(4).
4. Changes the dimensions of the shapes from 45x2 to 45x3, by adding a third column of nulls so that the equation for quaternion can be used (The Quaternion method can be used only for 3-D shapes, eq.(10)).
5. After finished aligning and adding 3-D dimension, it finds the TR and QT, who are the rotation matrix and the translation vector respectively
6. Then it computes the initial error
  - a. The registration error (reg\_error) is the error between the shape P and Y(Y matrix is described later).
  - b. The y\_error is the error between shape X1 and shape P.
7. After this, it finds the closest points of X1 (model) to shape we want to align (P).

This can be done from the equation:  $d(\vec{p}, X) = \min_{x \in X} \|\vec{x} - \vec{p}\|$ . The closest point is

the point that the distance is minimum. So we compute the minimum distances. This point is added to matrix (Y). This work id done iteretavily, so that Y matrix

- has 45x3 dimensions, as X1 and P (Y may have same points, if they are the closest points).
- Y now contains the closest points of the model. This is done by ICP\_ONLY.
8. ICP\_ONLY is a sub-program and works as follows:
    - Computes rotation and translation matrices to align P to X1 (model).
      - a. First calculates the cross-covariance of X1 and P, from (5).
      - b. Secondly, finds A, delta and Q, from equations (6),(7).
      - c. Thirdly, finds the eigenvectors and the eigenvalues. The eigenvalue that corresponds to the maximum eigenvalue is selected as the optimal rotation, from eq. (8)
    - Then computes the Rotation matrix, from equation (1), and the translation vector, from eq.(9).
  9. Then, in the main program (ICP1teliko), we update the rotation and translation matrix from (ICP\_ONLY). This is done so finally to have one rotation and translation matrix that can be applied to the initial shape. P is the shape that changes after each iteration.
  10. Then, applies rotation an translation matrices and plots the results (figures...)
  11. Finally, computes the registration error.

## 6.7 Equivalence between ICP and ASM

After the statements of ASM and ICP algorithms we can prove the equivalence of the equations that are used for the measure of the distances between the points of the ‘model shape’ and ‘the shape to be aligned’. The problem that we deal with in this thesis can be defined as follows: We have two shapes or more that are represented with N landmarks (points) each one. Let them be  $x_1$  and  $x_2$ , our purpose being to align  $x_2$  to  $x_1$ . We assume shape  $x_1$  to be the model shape. If the above procedure can be accomplished in a satisfactory way, we can use it to align any shape to an ideal model shape  $x_1$ . Mathematically the problem can be treated as follows: We want to minimize the distance between the two shapes. To define the between shape distance we use the sum of the ‘Euclidean distances’ of the landmark points of the two shapes. By minimizing the between shape distance we align the two shapes. The Euclidean distance between two points  $P_1, P_2$  in two dimensions is given by the following formula:

$$d(P_2, P_1) = \sqrt{(P_{2x} - P_{1x})^2 + (P_{2y} - P_{1y})^2} \quad (1)$$

The two algorithms that are implemented in this thesis, so far, for the achievement of the alignment of the N shapes are the ASM and the ICP algorithms. Each of the algorithms is dealing with the problem in a different way; however both are trying to minimize the Euclidean distance between the two shapes.

- In the ASM algorithm the equation to be minimized is:

$$E = (x_1 - M(x_2))^T W(x_1 - M(x_2)) \quad (2)$$

In (2)  $x_1$  is the model and  $M(x_2)$  is  $x_2$  transformed in case to be aligned to  $x_1$  with the use of the rotation and the translation. If there are N landmarks we have N equations. If we

ignore  $W$  then equation (2) effectively represents the sum of all the distances of the each landmark from the respective landmark of the model shape, and is using the Euclidean distance definition as defined in (1).

- In the ICP algorithm the equation to be minimized is:

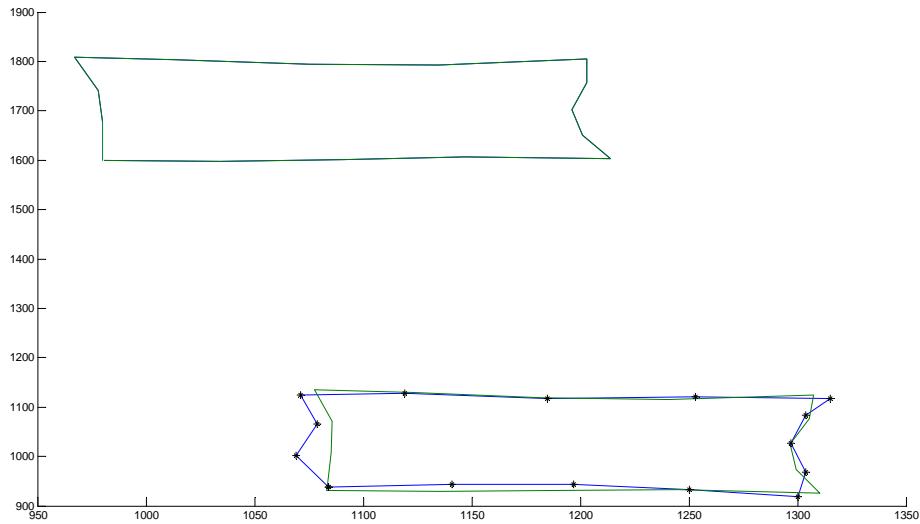
$$f(\vec{q}) = \frac{1}{N_p} \sum_{i=1}^{N_p} \left\| \vec{x}_i - R(\vec{q}_R) \vec{p}_i - \vec{q}_T \right\|^2 \quad (3)$$

$\vec{x}_i$  is the model and  $R(\vec{q}_R) \vec{p}_i - \vec{q}_T$  is the transformed shape. This is obvious that the distance is the Euclidean distance. So we can reach the statement that the two models (2)

## 6.7 Results and figures of the ICP algorithm

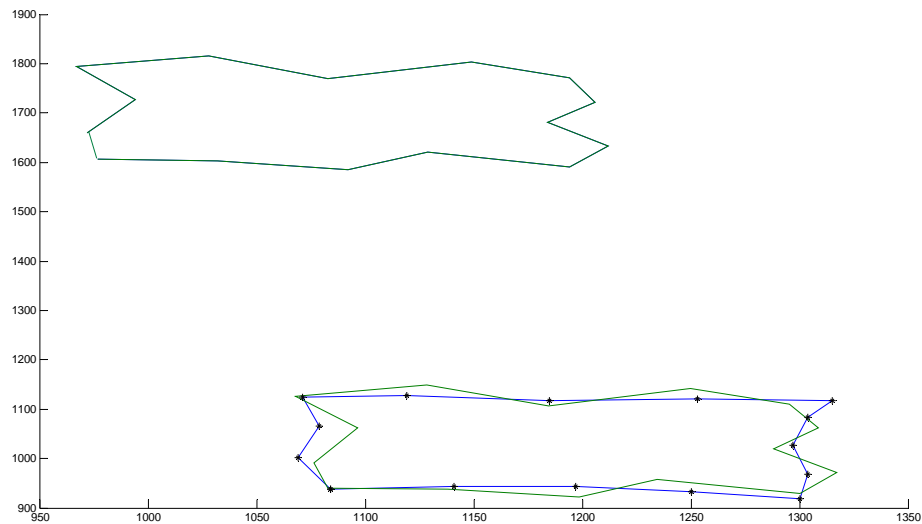
The ICP algorithm is implemented in MATLAB 7.1 with the use of MousePos, which is a especially designed program in C++ programming language for the manually fixing of the landmarks in the X-ray digital images. These images are from the NHANES II database. These images are available in the website:

<http://archive.nlm.nih.gov/proj/ftp/ftp.php>



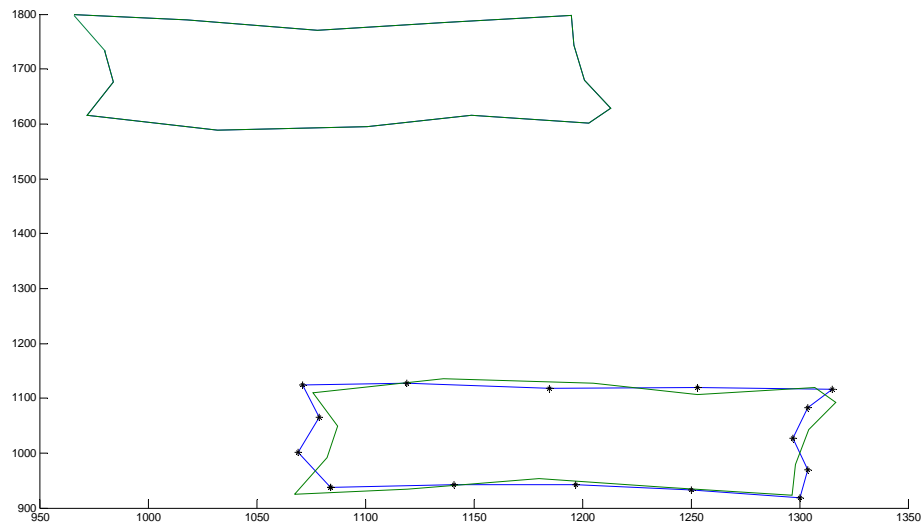
**Figure 6.2 ICP ‘The Normal Case’**

- With color blue is the model shape, with green the shape before and after the alignment (in all ICP figures).
- In **Figure 6.2** it is shown the implementation of the ICP program, with the use of the ICP\_ONLY subprogram, for the alignment of two shapes with 15 landmarks normal fixed, as in Figure 5.6 (of ASM's figures), ('L04531' and 'L12966' images):



**Figure 6.3** The 'Random Case'

- In **Figure 6.3** it is shown the implementation of the ICP program when the one shape has random fixed landmarks



**Figure 6.4** The ‘Improper Case’

- In **Figure 6.4** is the implementation of the ICP algorithm when the points are not fixed properly.

- In the previous figures we show the ICP’s alignment. In all the investigated cases the algorithm worked very efficient. If we focus to the model’s shape points and to the ‘to be aligned’ shape points, in all the cases the distances are very small. A very strong advantage of ICP algorithm, is that, because of the Closest Point matrix, this algorithm can execute the alignment independent of the number of the landmarks that were used for the representation of the shape ‘to be aligned’. This can be done because the Closest point matrix applies to its matrix the closest points to each point of the shape ‘to be aligned’.

## 7. Nelder-Mead algorithm

### 7.1 Introduction

The Nelder – Mead algorithm attempts to minimize a scalar-valued nonlinear function of  $n$  real variables using only function values, without any derivative information (explicit or implicit). The Nelder-Mead method falls in the general class of direct search methods. A large subclass of direct search methods, including the Nelder – Mead method, maintain at each step a *n*-simplex, a geometric figure in  $n$  dimensions of nonzero volume that is the convex hull of  $n+1$  vertices. Each iteration of a simplex direct search method begins with a simplex, specified by its  $n+1$  vertices and the associated function values.

A simplex method for finding a local minimum of a function of several variables has been devised by Nelder and Mead. For two variables, a simplex is a triangle, and the method is a pattern search that compares function values at the three vertices of a triangle. The worst vertex, where  $f(x, y)$  is largest, is rejected and replaced with a new vertex. A new triangle is formed and the search is continued. The process generates a sequence of triangles (which might have different shapes), for which the function values at the vertices get smaller and smaller. The size of the triangles is reduced and the coordinates of the minimum point are found. The algorithm is stated using the term **simplex** (a generalized triangle in  $N$  dimensions) and will find the minimum of a function of  $N$  variables. It is effective and computationally compact.

### 7.2 The Nelder – Mead algorithm

The Nelder-Mead algorithm was proposed as a method for minimizing a real-valued function  $f(\mathbf{x})$  for  $\mathbf{x} \in \mathcal{R}^n$ . Four scalar parameters must be specified to define a complete Nelder-Mead method: coefficients of reflection ( $p$ ), expansion ( $x$ ), contraction ( $\gamma$ ), and shrinkage( $\sigma$ ). According to the original Nelder-Mead paper, these parameters should satisfy:

$$(a) \ p > 0, \ x > 1, \ x > p, \ 0 < \gamma < 1, \ \text{and} \ 0 < \sigma < 1$$

$$(b) \ p=1, \ x=2, \ \gamma=1/2, \ \text{and} \ \sigma=1/2$$

The general conditions (a) are assumed for the one-dimensional case and the standard case (b) in the two-dimensional analysis.

### 7.3 Statement of the algorithm

At the beginning of the  $k$ th iteration,  $k \geq 0$  a simplex  $\Delta_k$  is given, along with its vertices, each of which is a point in  $\mathcal{R}^n$ . It is always assumed that iteration  $k$  begins by ordering and labeling these vertices as:  $x_1^{(k)}, \dots, x_{n+1}^{(k)}$  such that:

$$f_1^{(k)} \leq f_2^{(k)} \leq \dots \leq f_{n+1}^{(k)}$$

where  $f_i^{(k)}$  denotes  $f(x_i^{(k)})$ . The  $k$ th iteration generates a set of  $n+1$  vertices that define a different simplex for the next iteration, so that  $\Delta_{k+1} \neq \Delta_k$ . Because we seek to minimize  $f$ , we refer to  $x_1^{(k)}$  as the best point or vertex, to  $x_{n+1}^{(k)}$  as the worst point, and to  $x_n^{(k)}$  as the next-worst point. Similarly, we refer to  $f_{n+1}^{(k)}$  as the worst function value, and so on.

### - Initial triangle BWG

Let  $f(x, y)$  be the function that is to be minimized. To start, we are given three vertices of a triangle:  $V_k = (x_k, y_k)$ ,  $k = 1, 2, 3$ . The function  $f(x, y)$  is then evaluated at each of the three points:  $z_k = f(x_k, y_k)$  for  $k = 1, 2, 3$ . The subscripts are then reordered so that  $z_1 \leq z_2 \leq z_3$ . The next notation is used:  $B = (x_1, y_1)$ ,  $G = (x_2, y_2)$ , and  $W = (x_3, y_3)$ .  $B$  is the best vertex,  $G$  is good (next to worst or next to best), and  $W$  is the worst vertex.

### - Midpoint of the good side

The construction process uses the midpoint of the line segment joining  $B$  and  $G$ . It is found by averaging the coordinates:

$$M = \frac{B + G}{2} = \left( \frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right)$$

### - Reflection Using the Point R

The function decreases as we move along the side of the triangle from  $W$  to  $B$ , and it decreases as we move along the side from  $W$  to  $G$ . Hence it is feasible that  $f(x, y)$  takes on smaller values at points that lie away from  $W$  on the opposite side of the line between  $B$  and  $G$ . We choose a test point  $R$  that is obtained by “reflecting” the triangle through the side  $BG$ . To determine  $R$ , we first find the midpoint  $M$  of the side  $BG$ . Then draw the line segment from  $W$  to  $M$  and call its length  $d$ . This last segment is extended a distance  $d$  through  $M$  to locate the point  $R$  (Figure 7.1). The vector formula for  $R$  is

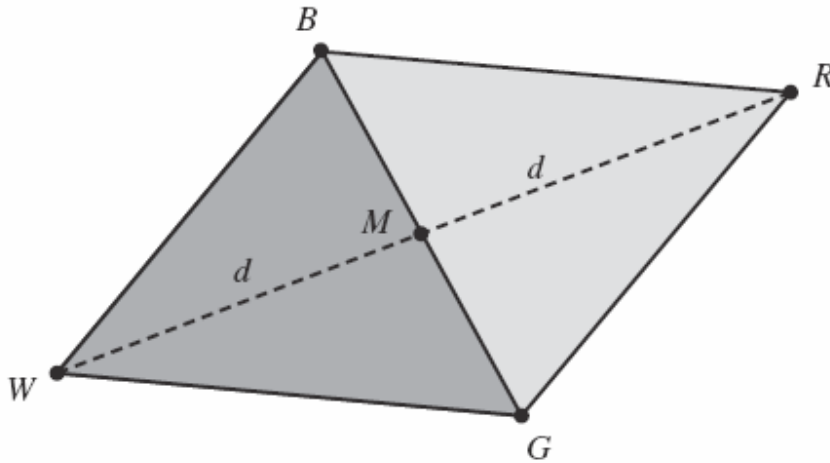
$$R = M + (M - W) = 2M - W$$

### - Expansion Using the Point E

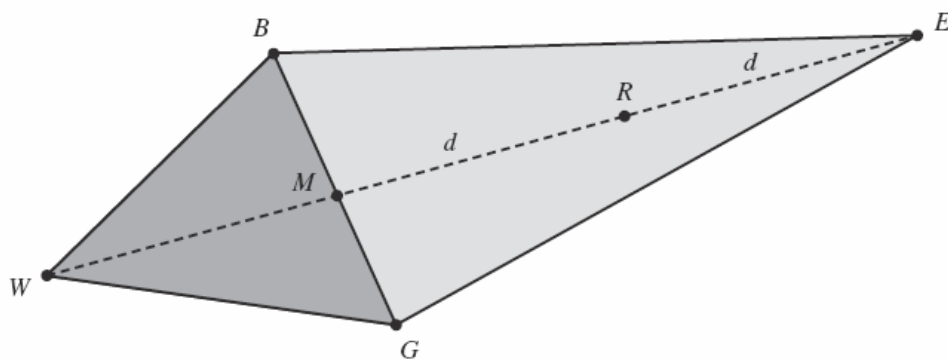
If the function value at  $R$  is smaller than the function value at  $W$ , then we have moved in the correct direction toward the minimum. Perhaps the minimum is just a bit farther than the point  $R$ . So we extend the line segment through  $M$  and  $R$  to the point  $E$ . This forms an expanded triangle  $BGE$ . The point  $E$  is found by moving an additional distance  $d$  along the line joining  $M$  and  $R$  (Figure 7.2). If the function value at  $E$  is less than the

function value at  $R$ , then we have found a better vertex than  $R$ . The vector formula for  $E$  is

$$E = R + (R - M) = 2R - M$$



**Figure 7.1** The triangle BGW and midpoint  $M$  and reflected point  $R$  for the Nelder-Mead method



**Figure 7.2** The triangle BGW and point  $R$  and extended point  $E$

#### - Contraction Using the Point $C$

If the function values at  $R$  and  $W$  are the same, another point must be tested. Perhaps the function is smaller at  $M$ , but we cannot replace  $W$  with  $M$  because we must have a triangle. Consider the two midpoints  $C_1$  and  $C_2$  of the line segments  $\overline{WM}$  and  $\overline{MR}$ , respectively (see Figure 7.3). The point with the smaller function value is called  $C$ , and



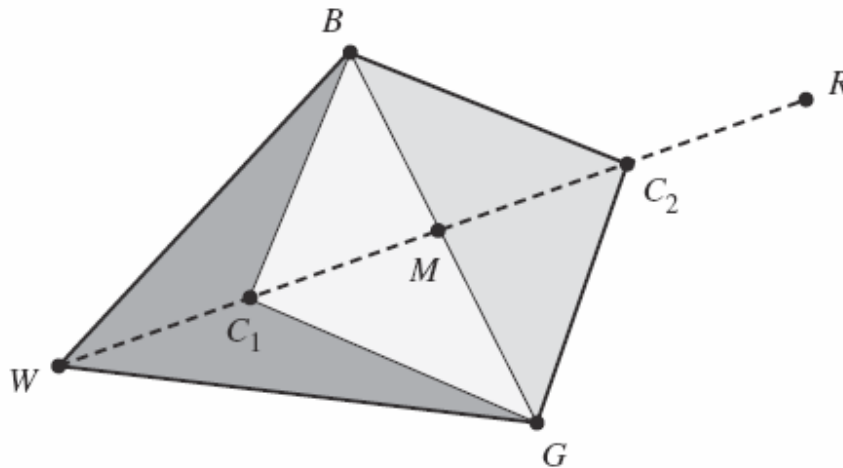
the new triangle is **BGC**. Note. The choice between  $C_1$  and  $C_2$  might seem inappropriate for the two-dimensional case, but it is important in higher dimensions.

**- Shrink toward B**

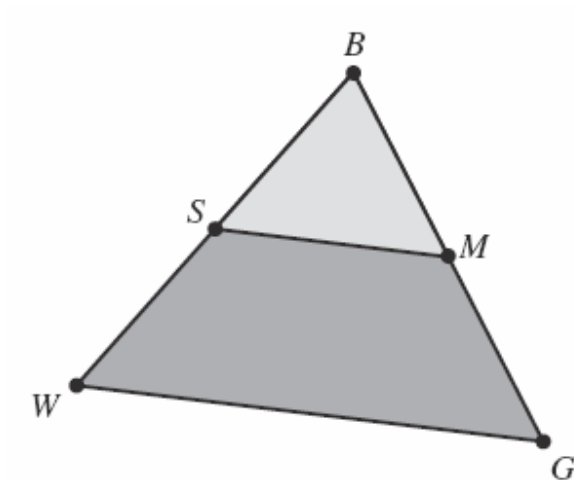
If the function value at **C** is not less than the value at **W**, the points **G** and **W** must be shrunk toward **B** (see Figure 7.4). The point **G** is replaced with **M**, and **W** is replaced with **S**, which is the midpoint of the line segment joining **B** with **W**.

**- Logical Decisions for Each Step**

A computationally efficient algorithm should perform function evaluations only if needed. In each step, a new vertex is found, which replaces **W**. As soon as it is found, further investigation is not needed, and the iteration step is completed. The logical details for two-dimensional cases are explained in table 7.1.



**Figure 7.3** The construction point  $C_1$  or  $C_2$  for Nelder-Mead method.



**Figure 7.4** Shrinking the triangle towards B.

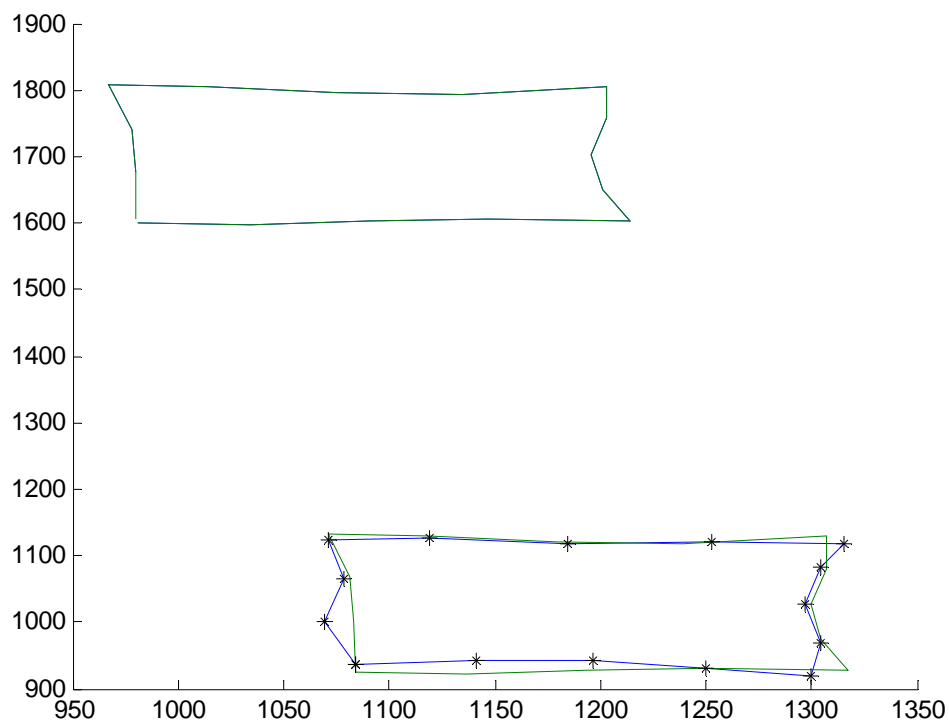
<hr/> IF $f(R) < f(G)$ , THEN Perform Case (i) {either reflect or extend} ELSE Perform Case (ii) {either contract or shrink} <hr/>	
BEGIN {Case (i).} IF $f(B) < f(R)$ THEN replace W with R ELSE Compute E and $f(E)$ IF $f(E) < f(B)$ THEN replace W with E ELSE replace W with R ENDIF ENDIF END {Case (i).}	BEGIN {Case (ii).} IF $f(R) < f(W)$ THEN replace W with R Compute $C = (W + M)/2$ or $C = (M + R)/2$ and $f(C)$ IF $f(C) < f(W)$ THEN replace W with C ELSE Compute S and $f(S)$ replace W with S replace G with M ENDIF END {Case (ii).}

---

**Table 7.1** Logical decisions for the Nelder-Mead algorithm.

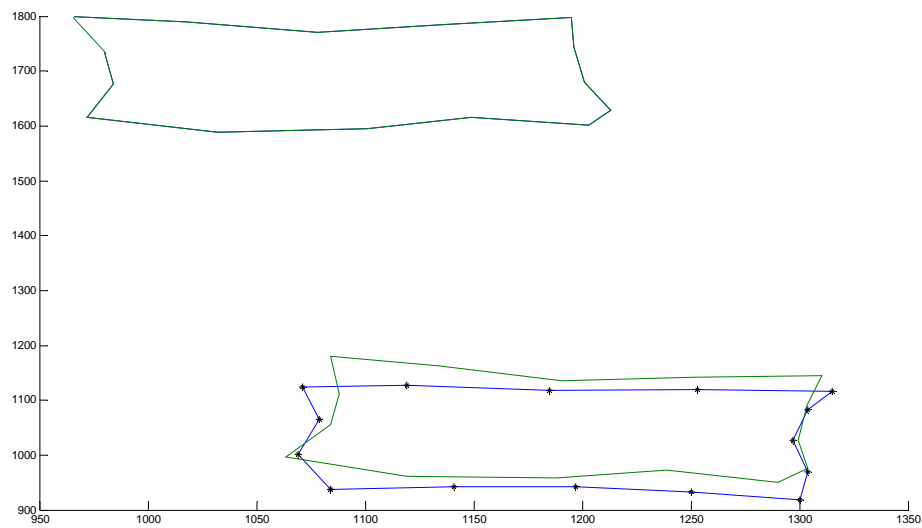
#### 7.4 Results and Figures of the Nelder-Mead algorithm

Nelder Mead algorithm is implemented with the use of Matlab 7.1. In Figure 7.5 is the 'normal case', in Figure 7.6 is the 'random numbering of the points' case, in Figure 7.7 is 'Improper case. If the one of the two shapes has fewer points than the other the algorithm fails to execute the alignment.

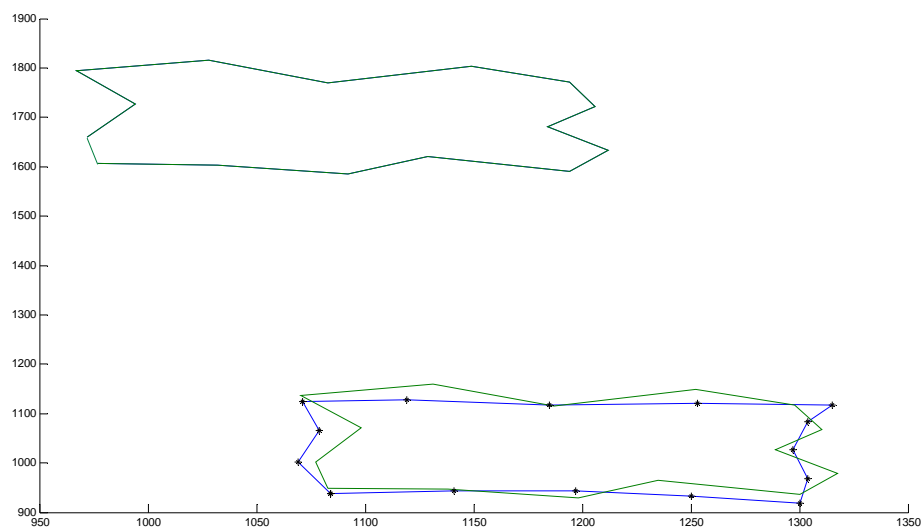


**Figure 7.5** 'The normal case'

- With blue color is the model shape, with green the shape before and after the alignment.



**Figure 7.6** ‘The random numbering case’



**Figure 7.7** The ‘Improper case’

- Overall, the algorithm in the first and the third cases worked well. But in the ‘random numbering case’ it seems that the distances between are models to be bigger than in the other two algorithms. The same disadvantage with the ASM algorithm also appears here; when one shape has different number of landmarks the algorithms cannot align the two shapes.

## 8. Experiments

### 8.1 Implementation of ICP, ASM and Nelder-Mead algorithms with 20 shapes represented by 15 landmarks each.

Three criteria are used for the comparison of the algorithms. In all of them 19 shapes are aligned to the ‘model shape’ which is the ‘mean shape’. In the first one we have 3 cases as we have seen in the previous chapters. 15 points are fixed with the use of MousePos program to represent the shapes. The original digital X-ray vertebrae image (Figure 8.2) is extracted from the NHANNES II database (<http://archive.nlm.nih.gov/proj/ftp/ftp.php>).

In this section 20 shapes represented from 15 landmarks aligned to a model-shape. As it is mentioned before ICP algorithm aligns the shapes to the ‘Closest point shape’, ASM to the ‘mean shape’, and Nelder-Mead to the ‘first shape’ of the 20. Here the three algorithms were aligned the 20 shapes to the ‘mean shape’, in case to be compared.

- The first case is if the landmarks are fixed properly, this means that landmarks are fixed exactly at the boundary of the vertebrae and with right to left numbering. (Figure 8.1). The two algorithms worked very well, ICP and ASM works with more accuracy. All the alignments are in Figure 8.3.

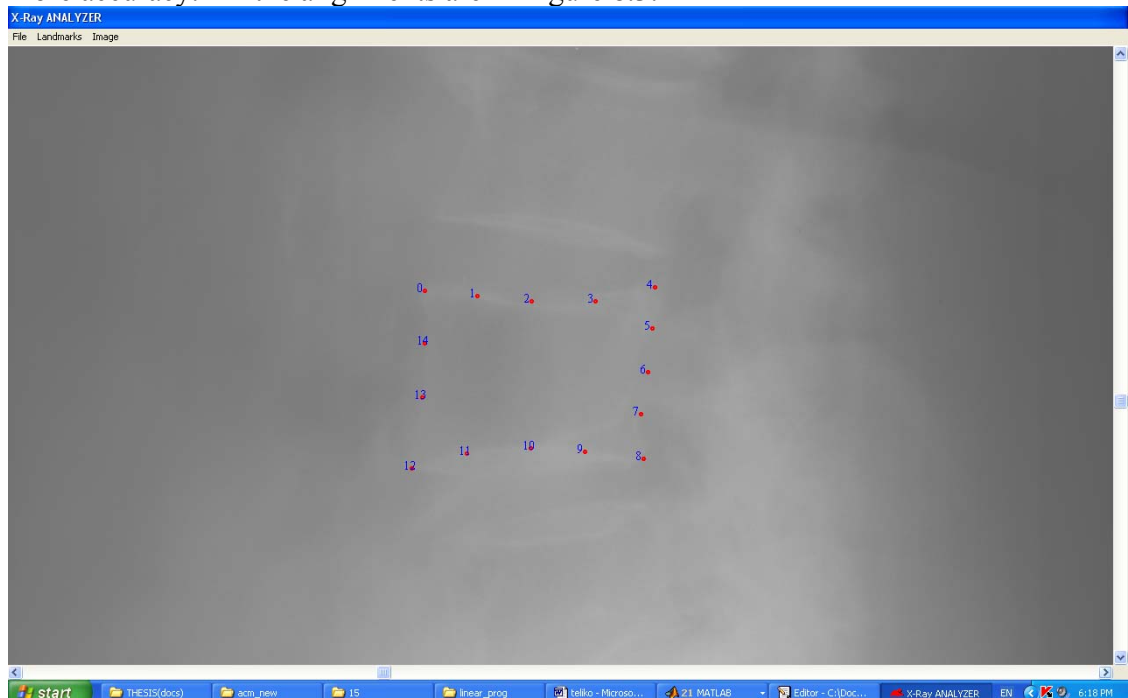
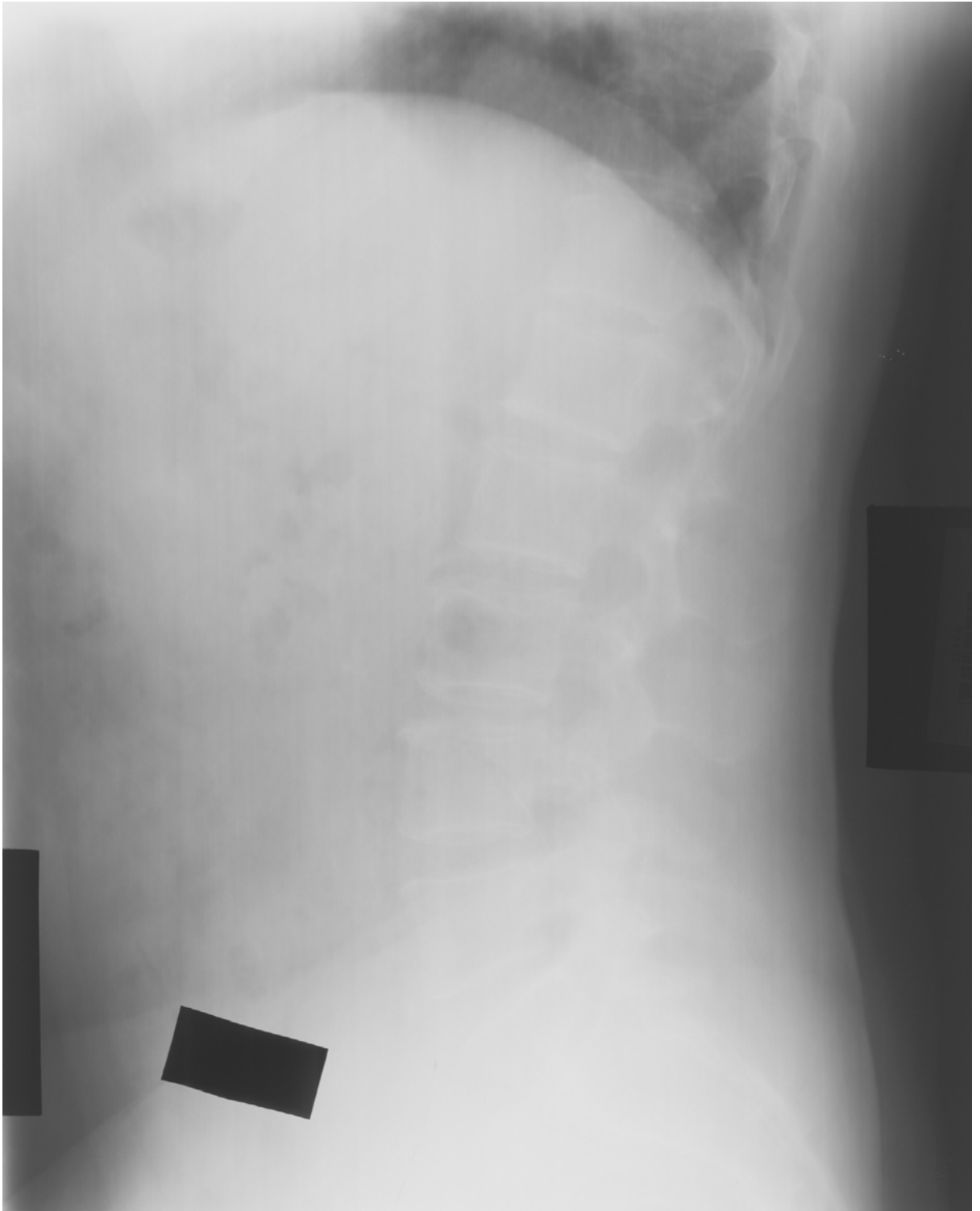
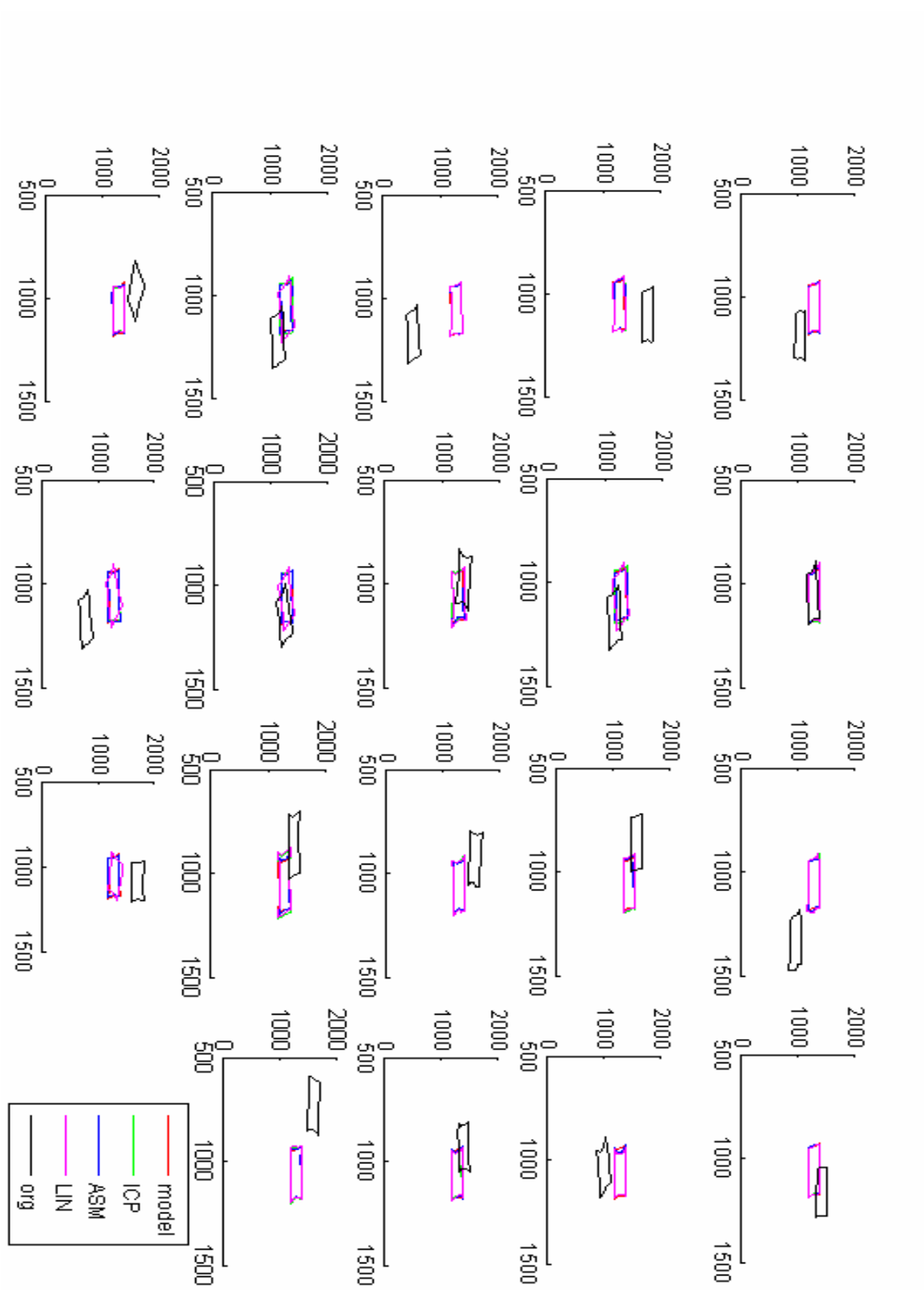


Figure 8.1 MousePos ‘The Normal Case’



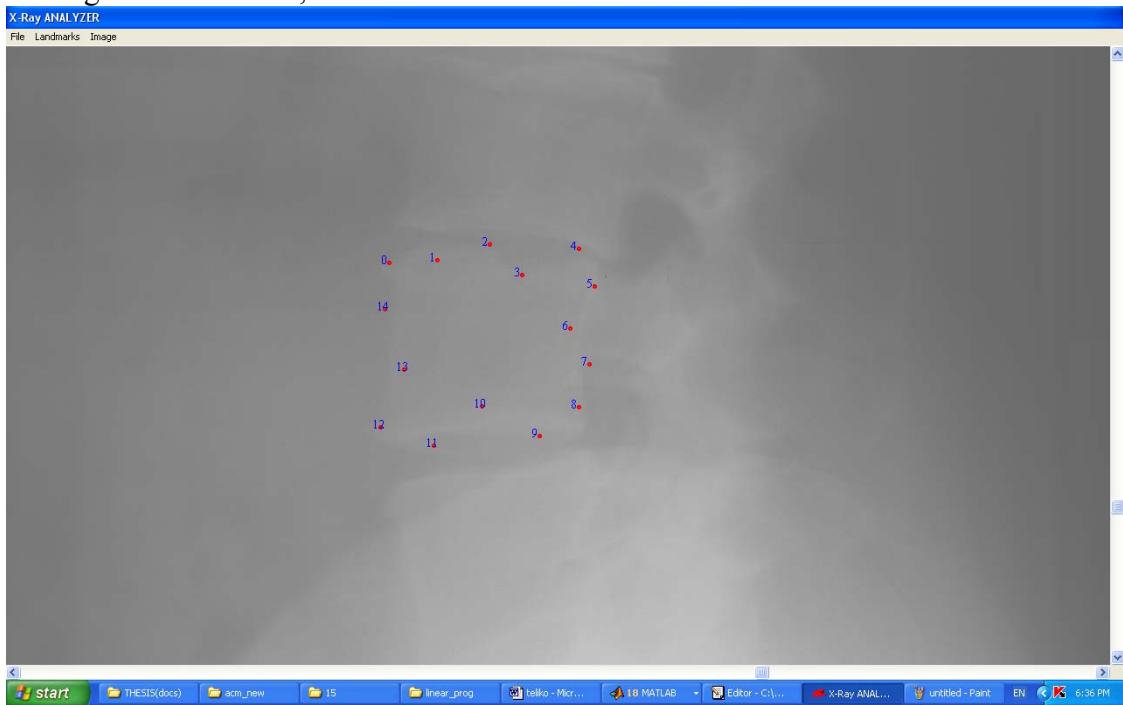
**Figure 8.2** X-Ray Image



**Figure 8.3** The ‘Normal Case’

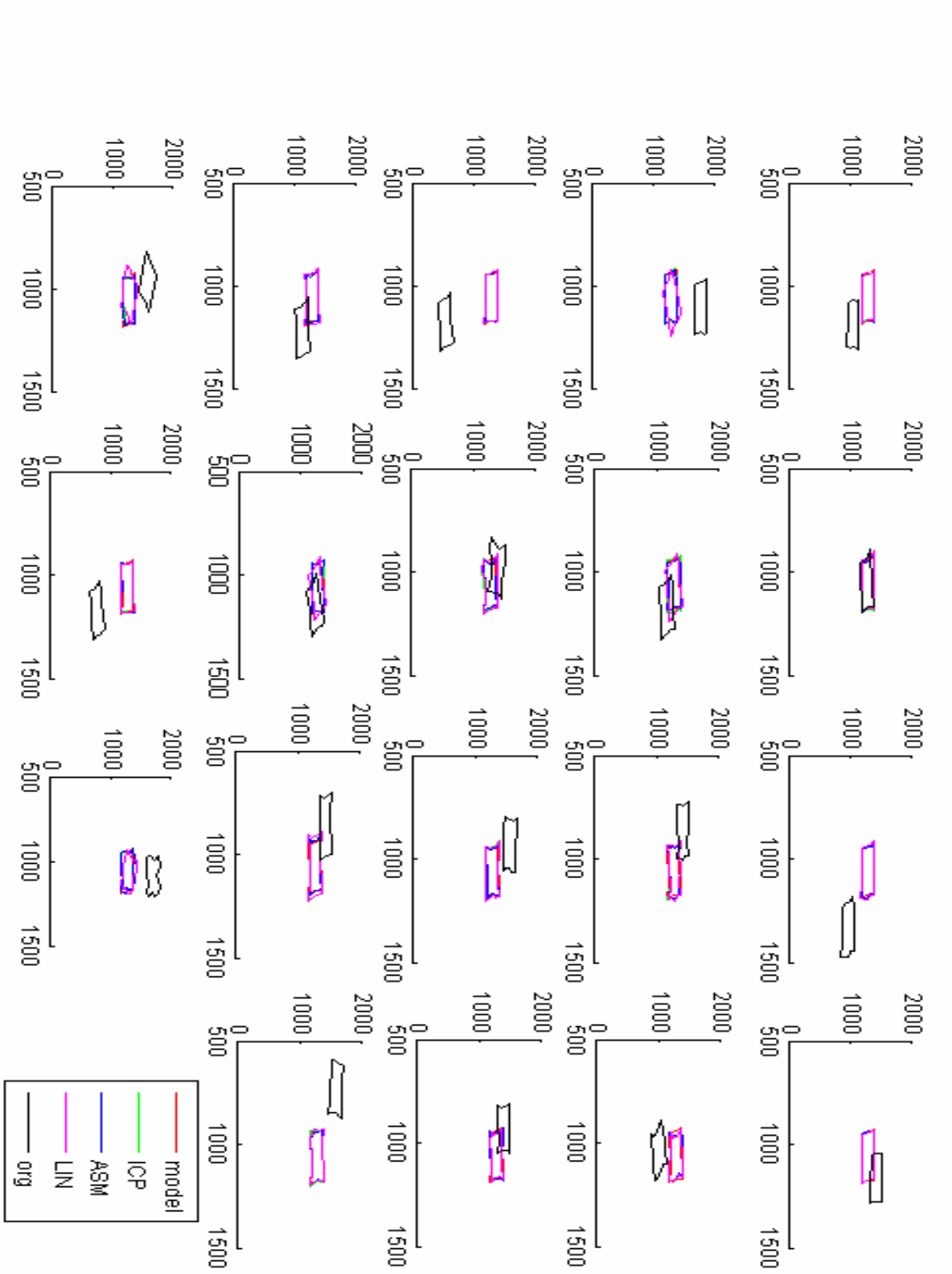
- With black is the shape before the alignment (original), with red is the ‘model shape’, with blue is ASM, with cyan LIN and with green the ICP alignment.

- The second case is when the landmarks are not fixed properly (not properly means that the landmarks are not fixed exactly at the boundary of the shape). Again, here the three algorithms behave very well (ICP has more precision). Figure 8.4 shows how the landmarks can look like and in figure 8.5 the alignments of ASM, ICP and Nelder-Mead.



**Figure 8.4** MousePos ‘The Improper Case’



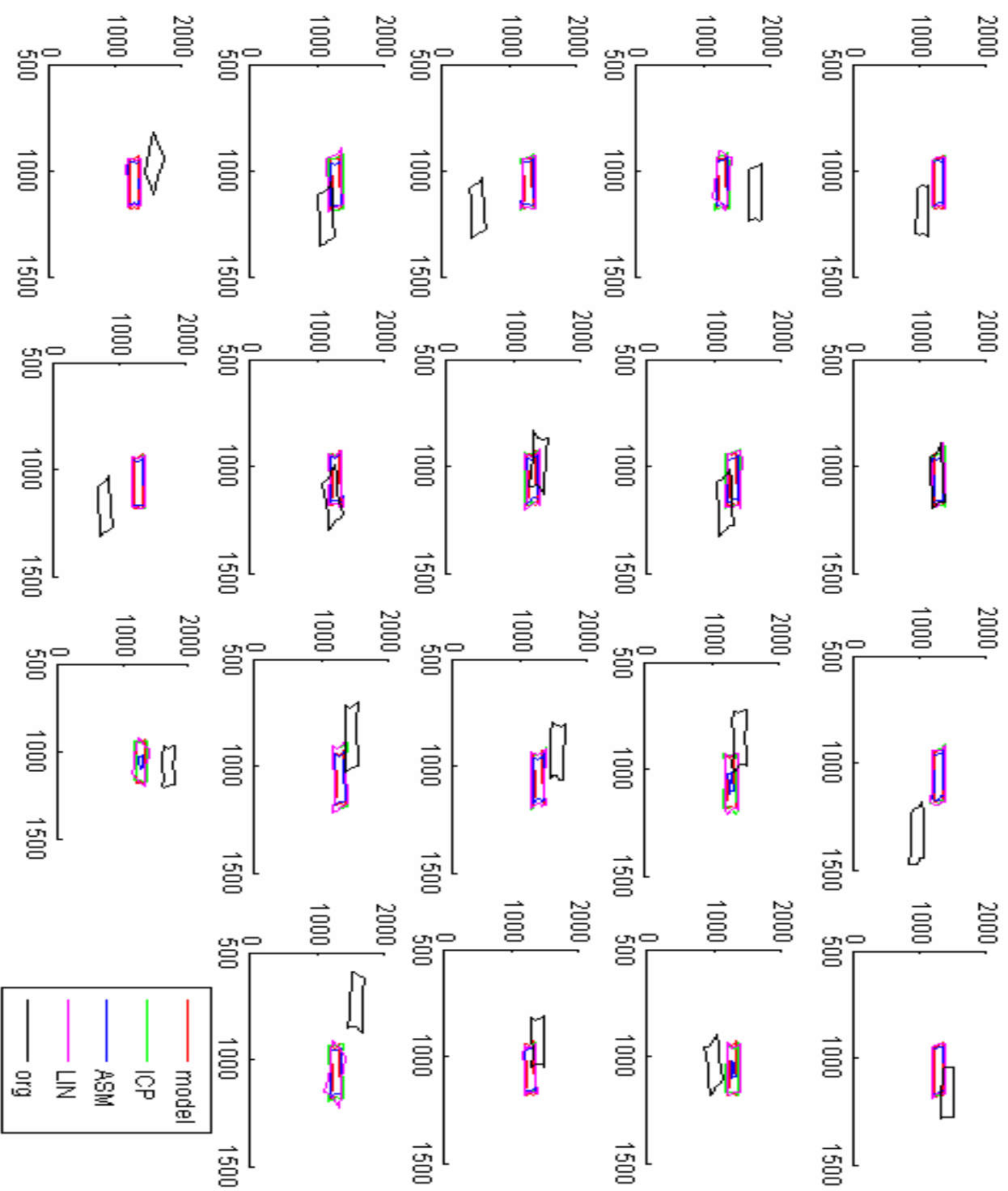


**Figure 8.5** The ‘Improper case’

- The third case considers if the numbering of the points begins not from the left corner of the vertebrae but from a random part as shown in Figure 8.6. In Figure 8.7 are shown the registration of the algorithms.



**Figure 8.6** Mouse Pos – ‘Different Numbering Case’



**Figure 8.7** ‘The Different Numbering Case’

- Overall algorithms work with a lot of accuracy in all the 3 cases that are investigated in the first section.

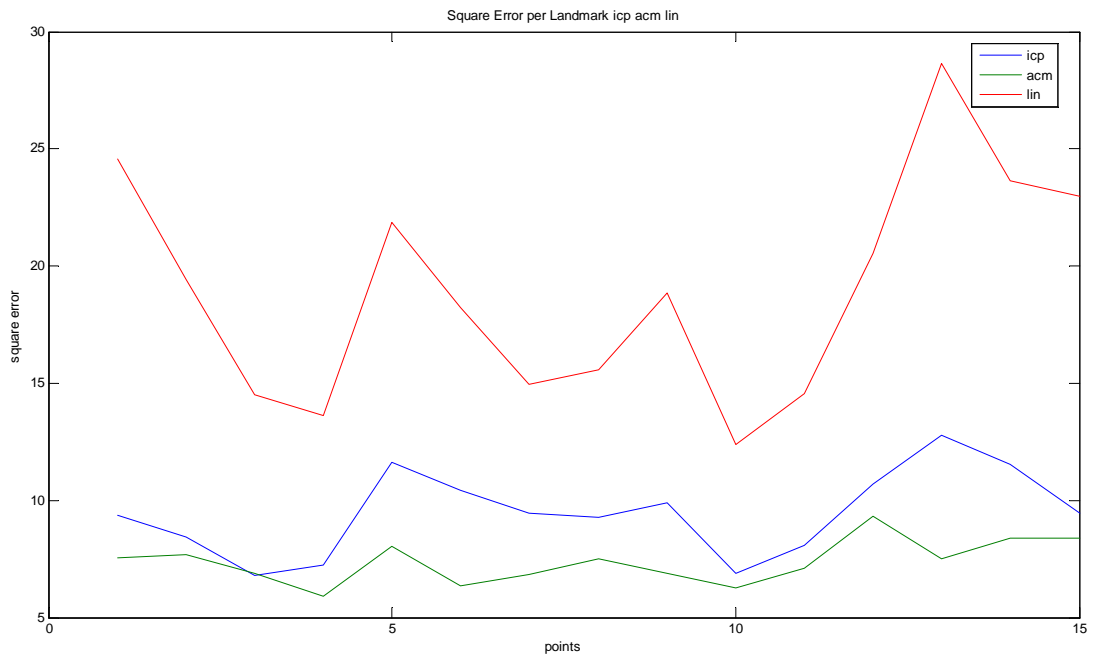
## 8.2 Error Curves

The second criterion follows. In this section is calculated the mean square error given from the following equation:

$$1/2N \sum_{i=1}^N \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \quad (1)$$

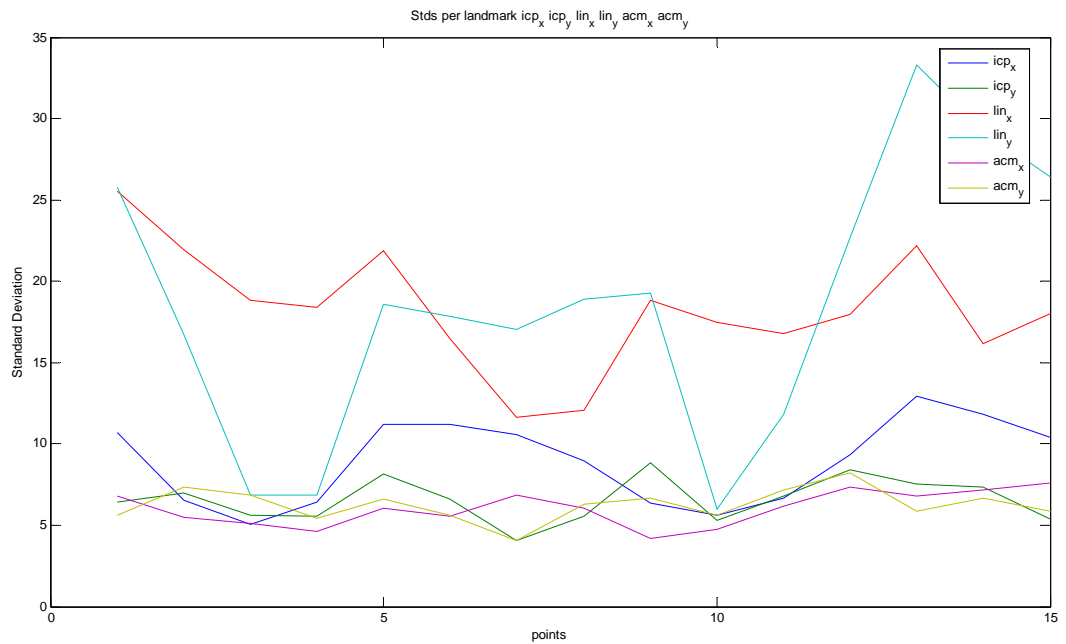
Where  $x_i$  and  $y_i$  are the coordinates of all the shapes (here 20), and  $x_m$  and  $y_m$  are the coordinates of the model shape.

- The investigated case is the ‘normal case’. In figure 8.8 is shown the mean square error for ASM, ICP and Nelder-Mead algorithms



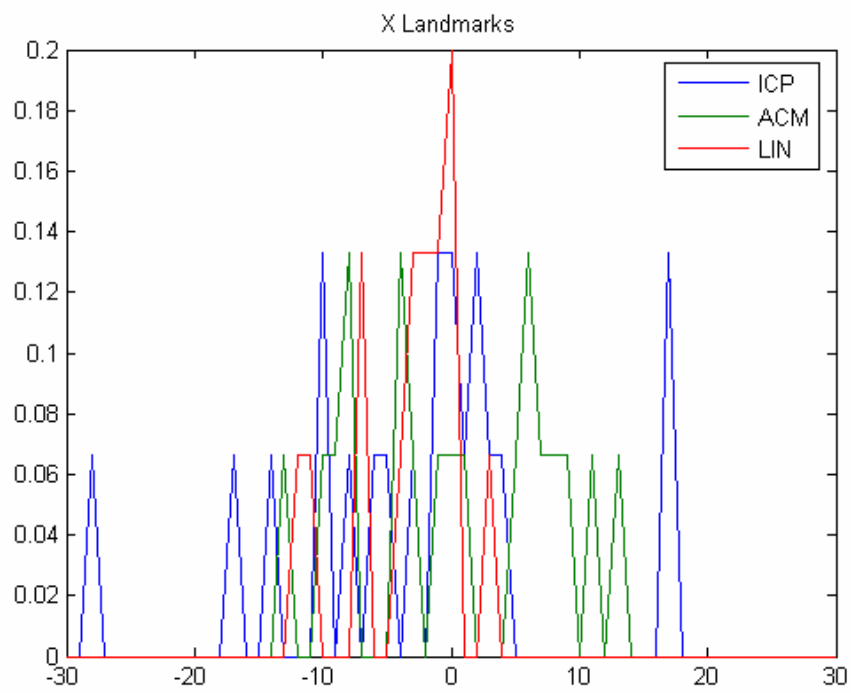
**Figure 8.8** Square Error

In Figures 8.9 can be seen the standard deviation of each landmark.

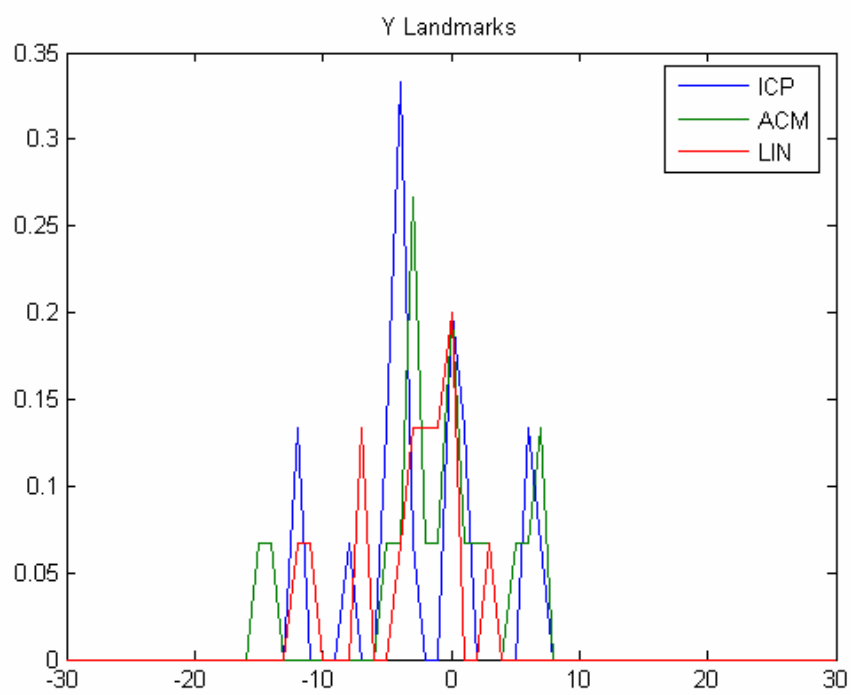


**Figure 8.9** Standard Deviation per landmark

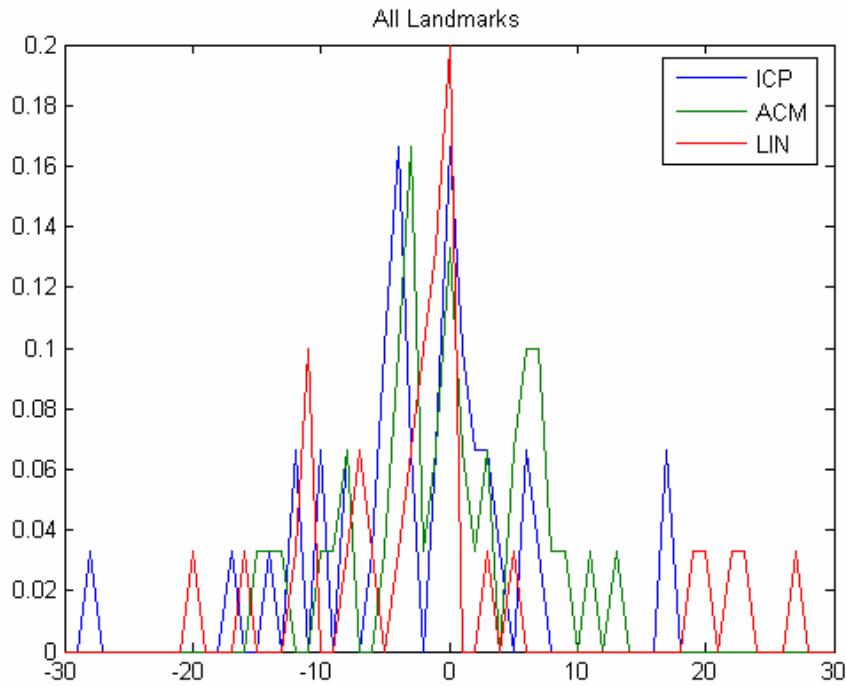
- When all the algorithms align with the mean shape as the 'model shape', ASM and ICP algorithms seem to have the minimum square error and standard deviation.
- Nelder-Mead algorithm seems to have worst square error than the other algorithms. The third criterion used for the comparison of the algorithms elaborates on how Nelder – Mead treats the problem. Let's see the distribution of error for each algorithm, for the X points (8.10), for the Y (8.11) points and for both (8.12) when the square error falls from -30 to 30.



**Figure 8.10 X-Landmarks**



**Figure 8.11 Y-landmarks**



**Figure 8.12** All-landmarks

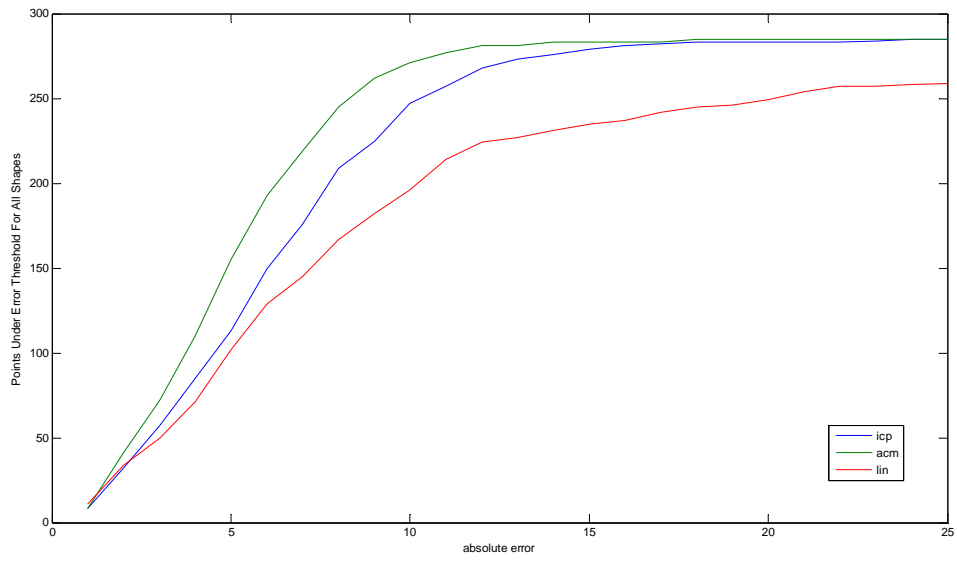
- The Nelder – Mead works in another way than the least square approach algorithms, trying to keep as many as possible points in ‘small error’. This is manifested via the concentration of more error values close to zero for the Nelder – mead algorithm than the other two algorithms.

A more thorough investigation of the distribution of errors is attempted through a study of the cumulative distribution of the absolute error. In Figure 8.13 we can see again the absolute error, but this time focusing in how many points have error < T (where T is the absolute error), and in Figures 8.14 and 8.15 are the mean absolute error and the variance of error, respectively.

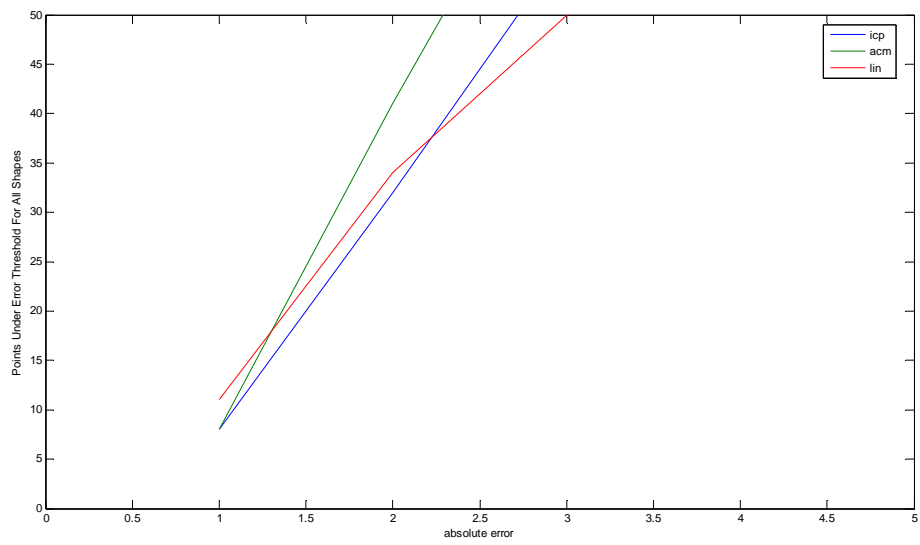
The equation that was used for the designing of these cumulative distributions diagrams was the one that measures the absolute error and is the following:

$$1/2N \sum_{i=1}^N |(x_i - x_m)| + |(y_i - y_m)|$$

We can see that the Nelder – Mead reacts with more accuracy when the error is smaller than 2, at the first 10 points. Also, ACM and ICP algorithms have same behavior; they work well after the first 10 points.

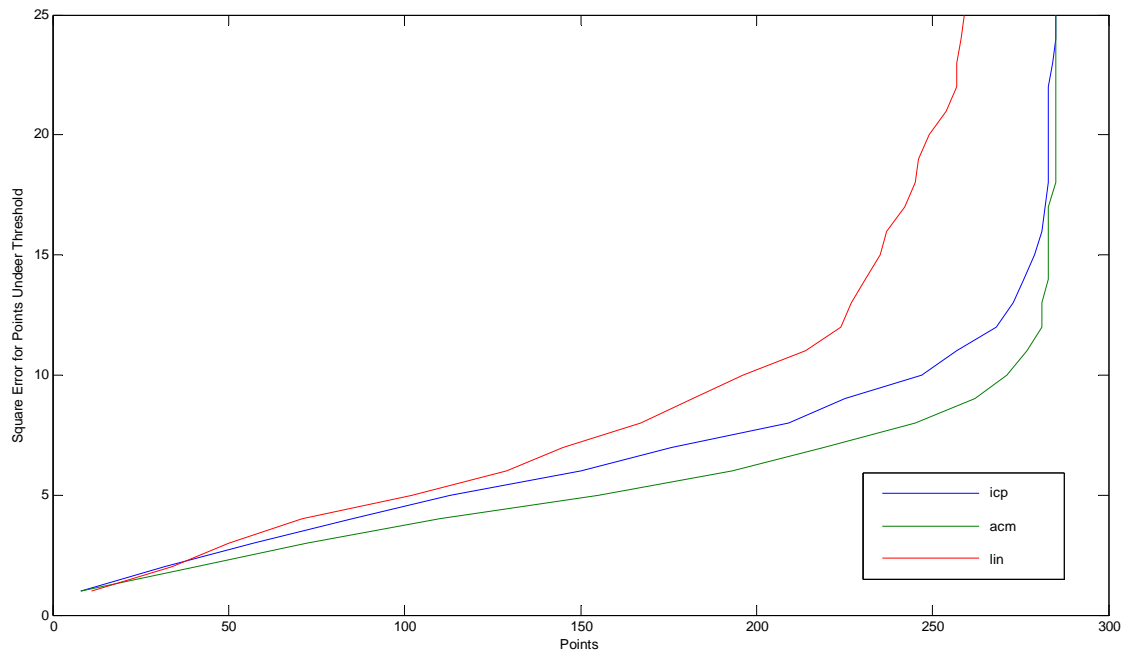


**Figure 8.13** Points under the Threshold

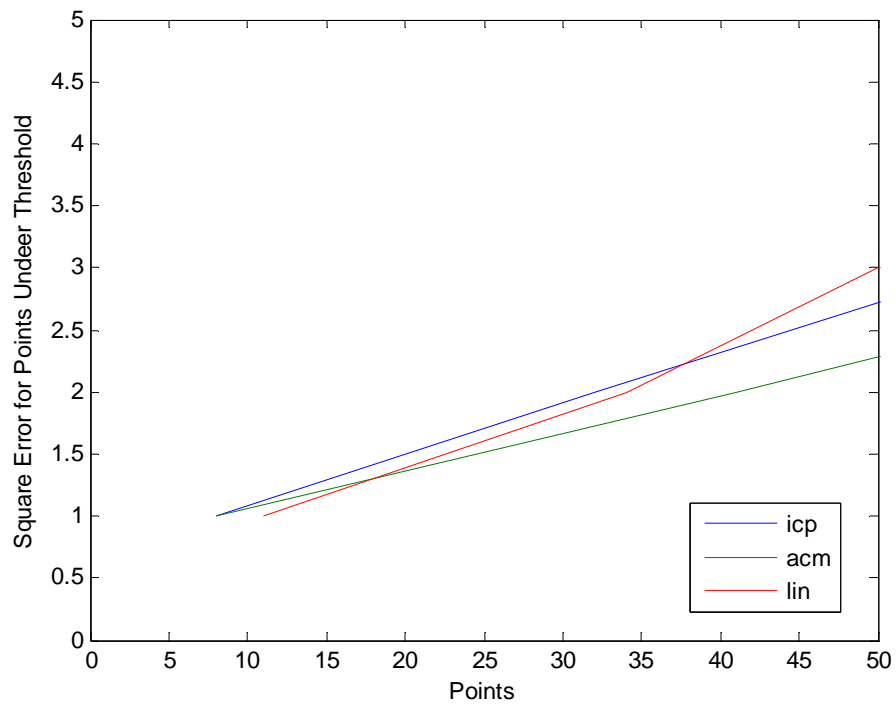


**Figure 8.14** Focusing in a part of Fig. 8.13

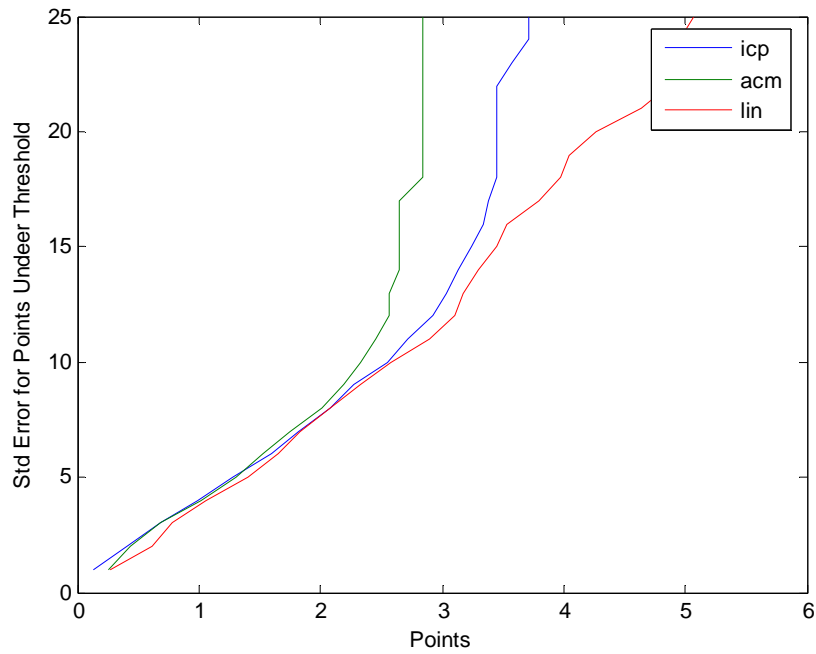




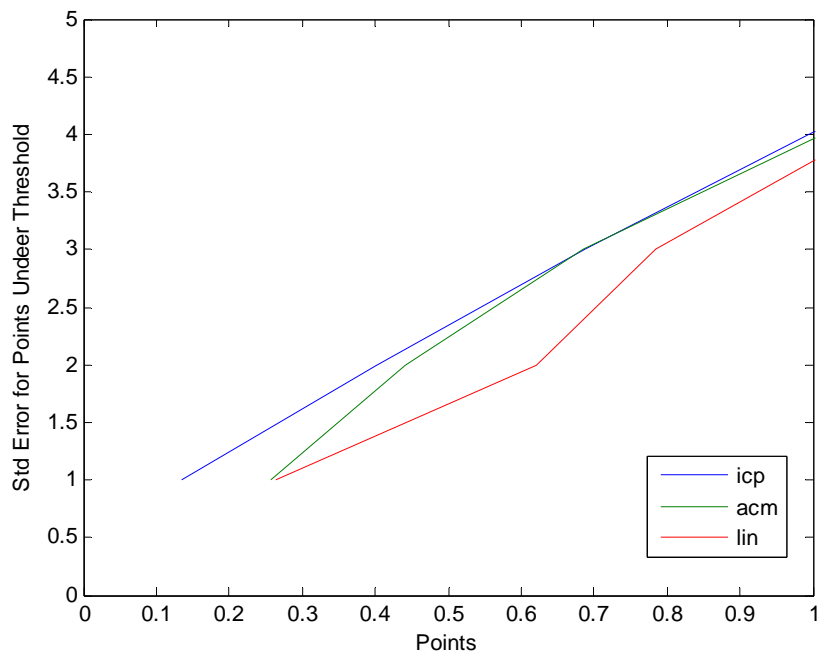
**Figure 8.15** Mean Error for point Under Threshold



**Figure 8.16** Focusing in a part of Fig.8.15



**Figure 8.17** Std Error for points under Threshold



**Figure 8.18** Focusing in a part of Fig. 8.17

- As we can see Nelder – Mead has the least mean error if the error is under the threshold of 5. This means that it forces a few points to very small errors, without being affected from larger errors at others points. This can be seen more clearly if we focus in these particular parts of the diagrams. Figures 8.13, 8.15 and 8.17 which are parts of figures 8.14, 8.16 and 8.18 respectively.

Three criteria were used in terms of the comparison of the approaches. Firstly, from figures 8.3, 8.5 and 8.7 algorithms look to work well, but with a more detailed investigation, with the use of the square error and the absolute error, ASM and ICP work with more accuracy than the Nelder – Mead overall (Fig. 8.8 and 8.9). Square error reaches almost the value of 30 in some cases for the linear algorithm, which means that it doesn't has precision at these points. The comparison of the standard deviation of the three algorithms for each individual landmark in Fig. 8.10, which shows the deviation of each point from the mean one, shows the same results as previous; the least squares approaches work more efficiently. The expected result from the beginning of this thesis was that the absolute error algorithms work in the following way: 'leave the remote points (which have the most absolute errors) and align the others with more accuracy'. In those terms a more detailed investigation followed and the distribution histograms were derived (Fig. 8.11, 8.12 and 8.13). In these diagrams we can see how the algorithms react in a specific of error (from -30 to 30). We can see that the Linear algorithm has the larger distribution in 0, which agrees with the assumption discussed before. In Figures 8.14, 8.15 and 8.16 we present the absolute error of the points that belong under a threshold of error, the mean error and the standard deviation of these points. Figures 8.17, 8.18 and 8.19 focusing in the same specific part of the three diagrams where the Nelder – Mead reacts better than the others.

As conclusion we can say that the least square approach trying to minimize the error for all the points involved and in the other hand the absolute error approach minimize the error of the points with small error without being matched affected by others. The least squares solution is optimal in case that the residuals obey a Gaussian distribution. If this is not the case, as it is in this problem, the assumptions of the least squares formulation are not valid, leading to suboptimal formulations of the problem. In such cases other norm formulations could lead to more robust results. In our problem we investigated the distribution of residuals from the least square formulation and found them to deviate from the Gaussian distribution. For this reason we also tested the least absolute error

## 9. Conclusion and further work

The purpose of this thesis was to examine registration based on landmarks. In particular we were interested in medical X-ray images. A general presentation of existing CBIR systems was done focusing in IRMA and NLM-CEB projects, also an overview about the two main parts of the CBIR systems, was presented, which are the indexing and the retrieval parts. A part of the indexing process is the registration problem. This thesis is focused on the landmark-based image registration problem.

A brief overview of the existing registration methods is presented. Extrinsic, intrinsic methods and non-image based registration techniques is discussed. Also norm-based approaches in orientation problems are presented, existing solutions in least squares and in absolute error approaches.

Generally the registration problem is an over determined problem, so least squares approach is a very common way of treating these kinds of problems. In this term two algorithms were selected: Active Shape Models (ASM) and the Iterative Closest Algorithm (ICP). These two algorithms were implemented and compared. It was found that they aligned N shapes to a 'model shape' with a lot of accuracy in three investigated cases.

Least squares is optimal in case that the residuals obey a Gaussian distribution. If this is not the case, the assumptions of the least squares formulation are not valid, leading to suboptimal formulations of the problem. In such cases other norm formulations could lead to more robust results. In our problem we investigated the distribution of residuals from the least square formulation and found them to deviate from the Gaussian distribution. For this reason we also tested the least absolute error formulation. Nelder – Mead algorithm was selected and implemented and compared with the other two.

The comparison of least squares approach and least absolute error approach was made using three criteria. The first was the implementation of these three algorithms with 19 shapes aligned to a 'model shape', represented with 15 landmarks each. Each of the algorithms worked very well. The second one was the measure of the mean square error, the absolute error and the standard deviation per landmark. Also histograms distribution diagrams presented as third criterion, and showed that the 'least squares algorithms' worked in more precision general but 'the absolute error algorithm' in few points is less than in the other approach.

Overall, least squares is the optimal solution for these kind of problems in case that the residuals obey a Gaussian distribution, if this not valid the absolute error solution is more efficient.

## 10. References

- [1]. Niblack W, Barber R, Equitz W, Flickner M, Yanker P, Ashley J: 'The QBIC-project-Querying images by content using color, texture and shape'. Proceedings SPIE 1993;1908:173-187
- [2]. Wu KJ, Narasimhalu AD, Mehtre BM, Lam CP, Gao YJ: CORE- 'A content-based retrieval engine for multimedia information systems'. Multimedia Systems 1995;3:25-41
- [3]. Carson c, Belongie S, Green span H, Malik J: 'Blobworld-Image segmentation using expectation-maximization and its application to image quering'. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002;24(8):1026-1038
- [4]. Korn P, Sidiropoulos N, Faloutsos C, Siegel E, Protopapas Z: 'Fast and effective retrieval of medical tumor shapes'. IEEE Transactions on Knowledge and Data Engineering 1998; 10(6):889-904
- [5]. Shyu CR, Brodley CE, Kak AC, Kosaka A, Aisen AM, Broderick LS: 'ASSERT-A physician in the loop content-based retrieval system for HRCT image databases'. Computer Vision and Image Understanding 1999;75(1/2)"111-132
- [6]. Long ER, Antania S, Leeb DJ, Krainakc DM, Thomaa GR: 'Biomedical information from a national collection of spine x-rays – Film to content-based retrieval'. Proceedings SPIE 2003;5033:in press
- [7]. Chu WW, Hsu CC, Cardenas AF, Tiara PK: 'Knowledge-based image retrieval with spatial and temporal constructs'. IEEE Transactions on Knowledge and data Engineering 1998;10(6):872-888
- [8]. Lehmann TM, Schubert H, Keysers D, Kohnen M, Wein BB: 'The IRMA code for unique classification of medical images'. Proceedings SPIE 2003;5033:in press
- [9]. Keysers D, Dahmen J, Ney H, Wein BB, Lehmann TM: 'A statistical framework for model-based image retrieval in medical applications'. Journal of Electronic Imaging, 2003;12(1):59-68
- [10]. M.O. Guild, M. Kohnen, D. Keysers, H. Schbert, B.B. Wein, J. Brendo, T.M. Lehman, 'Quality of DICOM header information for image categorization, in: Proceedings of the International Symposium on Medical Imaging', vol. 4685, San Diego, CA, USA, 2002, pp.280-287
- [11] A report to the Board of Scientific Counselors, September 26-27, 2002. Communications Engineering Branch, Lister Hill National Center for Biomedical Communications, National Library of Medicine
- [12] Tull L. Cooperative digital imaging prolects: the Greater memory projects. The Electronic Library. 2002;20(1):43-48
- [13] Goodrum AA, Rorving ME, Jeong KT, Suresh C. An open source agenda for research linking text and image content features. Journal of the American Society for information Science and Technology. 2001;52(11):948-953

- [14] Stan D, Sethi IK. Mapping low-level image features to semantic concepts. Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Media Databases. 2001;4315:172-179
- [15] Cootes TF, Taylor CJ. Statistical models of appearance for computer vision. Tech. rep., University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, Manchester, M12 9P, U.K. 2001 February
- [16] Hoffman ME, Wong EK. Content-based image retrieval by scale-space object boundary shape representation. IS&T/SPIE Conference on Storage and Retrieval for Media Databases. 2000;3972:86-97
- [17] Meilhac C, Nastar C. Relevance feedback and category search in image databases. IEEE International Conference on Multimedia Computing Systems, vol. 1 1999;512-517.
- [18] Zaniolo C, Ceri S, Faloutsos C, Snodgrass RT, Subrahmanian VS, Zicari R. Advanced Database Systems. Morgan Kaufmann Publishers, Inc. 1997
- [19] Guttman A. R-trees: A dynamic index structure for spatial searching. ACM SIGMOD Record- Proceedings of Annual Meeting. 1984;14(2):45-57
- [20] Qian X, Tagare HD. Optimally adapted indexing trees for medical image databases. IEEE International Symposium on Biomedical Imaging. 2002
- [21] Duda RO, Hart PE, Stork DG. Pattern Classification. John Wiley & Sons, Inc., New York, NY, second ed 2000.
- [22] Kurniawati R, Jin JS, Shaperd JA. The SS-tree: An Improved Index Structure for similarity Searches in a High Dimensional Feature Space. Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases V. 1997;3022:110-120
- [23] Pun T, Squire D. 'Statistical structuring of pictorial databases for content-based image retrieval systems'. Pattern Recognition Letters. 1996;17(12):1299-1310
- [24] White D, Jain R. Similarity indexing: 'Algorithms and performance'. Proceedings of IS&T/SPIE Conference on Storage nad Retrieval for Image and Video Databases IV Vol. SPIE 2670. 1996; 62-73
- [25] L. D. Lunsford. 'Modern stereotactic neurosurgery'. Martinus Nijhoff, Boston, MA, 1988.
- [26] D. Vandermeulen. 'Methods for registration, interpolation and interpretation of three-dimensional medical image data for use in 3-D display', 3-D modelling and therapy planning. PhD thesis, University of Leuven, Belgium, 1991.
- [27] L. Lemieux, N. D. Kitchen, S. W. Hughes, and D. G. T. Thomas. 'Voxel-based localization in frame-based and frameless stereotaxy and its accuracy'. Medical physics, 21(8):1301–1310, 1994.
- [28] L. Lemieux and R. Jagoe. 'Effect of fiducial marker localization on stereotactic target coordinate calculation in CT slices and radiographs'. Physics in medicine and biology, 39:1915–1928, 1994.
- [29] S. C. Strother, J. R. Anderson, X. Xu, J. Liow, D. C. Bonar, and D. A. Rottenberg. 'Quantitative comparisons of image registration techniques based on highresolution MRI of the brain'. Journal of computer assisted tomography, 18(6):954–962, 1994.
- [30] P. F. Hemler, P. A. van den Elsen, T. S. Sumanaweera, S. Napel, J. Drace, and J. R. Adler. 'A quantitative comparison of residual error for three different

- multimodality registration techniques'. In Y. Bizais, C. Barillot, and R. di Paola, editors, *Information processing in medical imaging*, pages 389–390. Kluwer, 1995.
- [31] D. Vandermeulen, A. Collignon, J. Michiels, H. Bosmans, P. Suetens, G. Marchal, G. Timmens, P. van den Elsen, M. Viergever, H. Ehrlicke, D. Hentschel, and R. Graumann. 'Multi-modality image registration within COVIRA'. In L. Beolchi and M. H. Kuhn, editors, *Medical imaging: analysis of multimodality 2D/3D images*, volume 19 of *Studies in health, technology and informatics*, pages 29–42. IOS Press, Amsterdam, 1995.
- [32] T. Peters, B. Davey, P. Munger, R. Comeau, A. Evans, and A. Olivier. 'Three-dimensional multimodal image-guidance for neurosurgery'. *IEEE Transactions on medical imaging*, 15(2):121–128, 1996.
- [33] K. P. Gall and L. J. Verhey. 'Computer-assisted positioning of radiotherapy patients using implanted radioopaque fiducials'. *Medical physics*, 20(4):1152–1159, 1993.
- [34] K. Leung Lam, R. K. ten Haken, D. L. McShan, and A. F. Thornton. 'Automated determination of patient setup errors in radiation therapy using spherical radio-opaque markers'. *Medical physics*, 20(4):1145–1152, 1993.
- [35] C. R. Maurer, J. J. McCrory, and J. M. Fitzpatrick. 'Estimation of accuracy in localizing externally attached markers in multimodal volume head images'. In M. H. Loew, editor, *Medical imaging: image processing*, volume 1898, pages 43–54, Bellingham, WA, 1993. SPIE Press.
- [36] S. Li, C. A. Pelizzari, and G. T. Y. Chen. 'Unfolding patient motion with biplane radiographs'. *Medical physics*, 21(9):1427–1433, 1994.
- [37] C. R. Maurer, G. B. Aboutanos, B. M. Dawant, S. Gadamsetty, R. A. Margolin, R. J. Maciunas, and J. M. Fitzpatrick. 'Effect of geometrical distortion correction in MR on image registration accuracy'. In M. H. Loew, editor, *Medical imaging: image processing*, volume 2167, pages 200–213, Bellingham, WA, 1994. SPIE Press.
- [38] C. R. Maurer, J. M. Fitzpatrick, R. L. Galloway, M. Y. Wang, R. J. Maciunas, and G. S. Allen. 'The accuracy of image-guided neurosurgery using implantable fiducial markers'. In H. U. Lemke, K. Inamura, C. C. Jaffe, and M. W. Vannier, editors, *Computer assisted radiology*, pages 1197–1202, Berlin, 1995. Springer-Verlag.
- [39] C. R. Maurer, G. B. Aboutanos, B. M. Dawant, R. A. Margolin, R. J. Maciunas, and J. M. Fitzpatrick. 'Registration of CT and MR brain images using a combination of points and surfaces'. In M. H. Loew, editor, *Medical imaging: image processing*, volume 2434, pages 109–123, Bellingham, WA, 1995. SPIE Press.
- [40] D. A. Simon, R. V. O'Toole, M. Blackwell, F. Morgan, A. M. DiGioia, and T. Kanade. 'Accuracy validation in image-guided orthopaedic surgery'. In *Medical robotics and computer assisted surgery*, pages 185–192. Wiley, 1995.
- [41] R. E. Ellis, S. Toksvig-Larsen, M. Marcacci, D. Caramella, and M. Fadda. 'A biocompatible fiducial marker for evaluating the accuracy of CT image registration'. In H. U. Lemke, M. W. Vannier, K. Inamura, and A. G. Farman, editors, *Computer assisted radiology*, volume 1124 of *Excerpta medica - international congress*

series, pages 693–698, Amsterdam, 1996. Elsevier.

[42] A. C. Evans, S. Marrett, J. Torrescorzo, S. Ku, and L. Collins. ‘MRI-PET correlation in three dimensions using a volume of interest (VOI) atlas’. *Journal of cerebral blood flow and metabolism*, 11:A69–A78, 1991.

[43] G. Q. Maguire, M. Noz, H. Rusinek, J. Jaeger, E. L. Kramer, J.J. Sanger, and G. Smith. ‘Graphics applied to medical image registration’. *IEEE Computer graphics and applications*, 11(2):20–28, 1991.

[44] R. T. Malison, E. G. Miller, R. Greene, G. McCarthy, D. S. Charney, and R. B. Innis. ‘Computer-assisted coregistration of multislice SPECT and MR brain images by fixed external fiducials’. *Journal of computer assisted tomography*, 17(6):952–960, 1993.

[45] M. Y. Wang, J. M. Fitzpatrick, C. R. Maurer, and R. J. Maciunas. ‘An automatic technique for localizing externally attached markers in MR and CT volume images of the head’. In M. H. Loew, editor, *Medical imaging: image processing*, volume 2167, pages 214–224, Bellingham, WA, 1994. SPIE Press.

[46] R. L. Wahl, L. E. Quint, R. D. Cieslak, A. M. Aisen, R. A. Koeppe, and C. R. Meyer. ‘“anametabolic” tumor imaging: fusion of FDG PET with CT or MRI to localize foci of increased activity’. *Journal of nuclear medicine*, 34:1190–1197, 1993.

[47] R. D. Bucholz, K. R. Smith, J. Henderson, and L. McDurmont. ‘Intraoperative localization using a three dimensional optical digitizer’. In *Medical robotics and computer assisted surgery*, pages 283–290, 1994.

[48] P. J. Edwards, D. J. Hawkes, D. L. G. Hill, D. Jewell, R. Spink, A. Strong, and M. Gleeson. ‘Augmented reality in the stereo operating microscope for otolaryngology and neurosurgical guidance’. In *Medical robotics and computer assisted surgery*, pages 8–15. Wiley, 1995.

[49] P. J. Edwards, D. L. G. Hill, D. J. Hawkes, R. Spink, A. C. F. Colchester, A. Strong, and M. Gleeson. ‘Neurosurgical guidance using the stereo microscope’. In N. Ayache, editor, *Computer vision, virtual reality, and robotics in medicine*, volume 905 of *Lecture notes in computer science*, pages 555–564, Berlin, 1995. Springer-Verlag.

[50] W. D. Leslie, A. Borys, D. McDonald, J. O. Dupont, and A. E. Peterdy. ‘External reference markers for the correction of head rotation in brain single-photon emission tomography’. *European journal of nuclear medicine*, 22(4):351–355, 1995.

[51] S. J. Stapleton, C. B. Caldwell, L. E. Ehrlich, C. L. Leonhardt, S. E. Black, and M. J. Yaffe. ‘Effects on non-linear flow and spatial orientation on technetium-99m hexamethylpropylene amin oxime single-photon emission tomography’. *European journal of nuclear medicine*, 22(9):1009–1016, 1995.

[52] M. Y. Wang, J. M. Fitzpatrick, and C. R. Maurer. ‘Design of fiducials for accurate registration of CT and MR volume images’. In M. Loew, editor, *Medical imaging*, volume 2434, pages 96–108. SPIE, 1995.

[53] M. Fuchs, H. Wischmann, A. Neumann, J. Weese, W. Zylka, J. Sabczynski, M. H. Kuhn, T. M. Buzug, G. Schmitz, and P. M. C. Gieles. ‘Accuracy analysis for image-guided neurosurgery using fiducial skin markers, 3D CT imaging, and an optical localizer system’. In H. U. Lemke, M.W. Vannier, K. Inamura, and A. G. Farman, editors, *Computer assisted radiology*, volume 1124 of *Excerpta medica - international congress series*, pages 770–775, Amsterdam, 1996. Elsevier.



- [54] T. Greitz, M. Bergström, J. Boëthius, D. Kingsley, and T. Ribbe. 'Head fixation system for integration of radiodiagnostic and therapeutic procedures'. *Neuroradiology*, 19:1–6, 1980.
- [55] L. V. Laitinen, B. Liliequist, M. Fagerlund, and A. T. Eriksson. 'An adapter for computer tomography guided stereotaxis'. *Surgical neurology*, 23:559–566, 1985.
- [56] L. R. Schad, R. Boesecke, W. Schlegel, G. H. Hartmann, G. H. Sturm, L. G. Strauss, and W. J. Lorenz. 'Three dimensional image correlation of CT, MR, and PET studies in radiotherapy treatment of brain tumors'. *Computer assisted tomography*, 11:948–954, 1987.
- [57] D. J. Hawkes, D. L. G. Hill, and E. C. M. L. Bracey. 'Multi-modal data fusion to combine anatomical and physiological information in the head and heart'. In J. H. C. Reiber and E. E. van der Wall, editors, 'Cardiovascular nuclear medicine and MRI', pages 113–130, Dordrecht, the Netherlands, 1992. Kluwer.
- [58] A. C. Evans, S. Marrett, L. Collins, and T. M. Peters. 'Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems'. In R. H. Schneider, S. J. Dwyer III, and R. G. Jost, editors, *Medical imaging: image processing*, volume 1092, pages 264–274, Bellingham, WA, 1989. SPIE press.
- [59] D. L. G. Hill, D. J. Hawkes, J. E. Crossman, M. J. Gleeson, T. C. S. Cox, E. C. M. L. Bracey, A. J. Strong, and P. Graves. 'Registration of MR and CT images for skull base surgery using pointlike anatomical features'. *British journal of radiology*, 64(767):1030–1035, 1991.
- [60] D. L. G. Hill, D. J. Hawkes, and C. R. Hardingham. 'The use of anatomical knowledge to register 3D blood vessel data derived from DSA with MR images'. In *Medical imaging: image processing*, volume 1445, pages 348–357, Bellingham, WA, 1991. SPIE press.
- [61] H. Zubal, G. Tagare, L. Zhang, and J. Duncan. '3-D registration of intermodality medical images'. In *proceedings of the annual international conference of the IEEE engineering in medicine and biology society*, volume 13, pages 293–294, 1991.
- [62] C. J. Henri, A. Cukiert, D. L. Collins, A. Olivier, and T. M. Peters. 'Towards frameless stereotaxy: anatomical-vascular correlation and registration'. In *Visualization in biomedical computing*, volume 1808, pages 214–224, Bellingham, WA, 1992. SPIE press.
- [63] J. Bijhold. 'Three-dimensional verification of patient placement during radiotherapy using portal images'. *Medical physics*, 20(2):347–356, 1993.
- [64] G. X. Ding, S. Shalev, and G. Gluchev. 'A \_\_\_ technique for treatment verification in radiotherapy and its clinical applications'. *Medical physics*, 20(4):1135–1143, 1993.
- [65] R. W. Fright and A. D. Linney. 'Registration of 3-D head surfaces using multiple landmarks. *IEEE Transactions on medical imaging*', 12(3):515–520, 1993.
- [66] G. Gluchev and S. Shalev. 'Robust registration in case of different scaling'. In M. H. Loew, editor, *Medical imaging: image processing*, volume 1898, pages 126–133, Bellingham, WA, 1993. SPIE Press.
- [67] D. L. G. Hill, D. J. Hawkes, Z. Hussain, S. E. M. Green, C. F. Ruff, and G. P. Robinson. 'Accurate combination of CT and MR data of the head: validation and applications in surgical and therapy planning'. *Computerized medical imaging and graphics*, 17(4/5):357–363, 1993.

- [68] E. D. Morris, G. J. Muswick, E. S. Ellert, R. N. Steagall, P. F. Goyer, and W. E. Semple. 'Computer-aided techniques for aligning interleaved sets of non-identical medical images'. In M. H. Loew, editor, *Medical imaging: image processing*, volume 1898, pages 146–157, Bellingham, WA, 1993. SPIE Press.
- [69] P. Neelin, J. Crossman, D. J. Hawkes, Y. Ma, and A. C. Evans. 'Validation of an MRI/PET landmark registration method using 3D simulated PET images and point simulations'. *Computerized medical imaging and graphics*, 17(4/5):351–356, 1993.
- [70] Y. Ge, J. M. Fitzpatrick, J. R. Votaw, S. Gadamsetty, R. J. Maciunas, R. M. Kessler, and R. A. Margolin. 'Retrospective registration of PET and MR brain images: an algorithm and its stereotactic validation'. *Journal of computer assisted tomography*, 18(5):800–810, 1994.
- [71] L. Harmon, A. Vayda, E. Erlandson, J. Taren, and D. Ross. '3D laser scanning for image-guided neurosurgery'. In *Applications of computer vision in medical image processing*, AAAI spring symposium series, pages 106–109, 1994.
- [72] J. Moseley and P. Munro. 'A semiautomatic method for registration of portal images'. *Medical physics*, 21(4):551–558, 1994.
- [73] U. Pietrzyk, K. Herholz, G. Fink, A. Jacobs, R. Mielke, I. Slansky, M. Wörker, and W. Heis. 'An interactive technique for three-dimensional image registration: validation for PET, SPECT, MRI and CT brain studies'. *Journal of nuclear medicine*, 35:2011–2018, 1994.
- [74] Y. Ge, J.M. Fitzpatrick, R. M. Kessler, M. Jeske-Janicka, and R. A. Margolin. 'Intersubject brain image registration using both cortical and subcortical landmarks'. In M. Loew, editor, *Medical imaging*, volume 2434, pages 81–95, Bellingham, WA, 1995. SPIE.
- [75] A. Hamadeh, P. Sautot, and P. Cinquin. 'A unified approach to 3D-2D registration and 2D images segmentation'. In H. U. Lemke, K. Inamura, C. C. Jaffe, and M. W. Vannier, editors, *Computer assisted radiology*, pages 1191–1196, Berlin, 1995. Springer-Verlag.
- [76] A. Hamadeh, P. Sautot, S. Lavallée, and P. Cinquin. 'Towards automatic registration between CT and X-ray images: cooperation between 3D/2D registration and 2D edge detection'. In *Medical robotics and computer assisted surgery*, pages 39–46. Wiley, 1995.
- [77] C. R. Meyer, G. S. Leichtman, J. A. Brunberg, R. L. Wahl, and L. E. Quint. 'Simultaneous usage of homologous points, lines, and planes for optimal, 3-D, linear registration of multimodality imaging data'. *IEEE Transactions on medical imaging*, 14(1):1–11, 1995.
- [78] B. J. McParland and J. C. Kumaradas. 'Digital portal image registration by sequential anatomical matchpoint and image correlations for real-time continuous field alignment verification'. *Medical physics*, 22(7):1063–1075, 1995.
- [79] M. Soltys, D. V. Beard, V. Carrasco, S. Mukherji, and J. Rosenman. 'FUSION: a tool for registration and visualization of multiple modality 3D medical data'. In M. H. Loew, editor, *Medical imaging: image processing*, volume 2434, pages 74–80, Bellingham, WA, 1995. SPIE Press.
- [80] A. Savi, M. C. Gilardi, G. Rizzo, M. Pepi, C. Landoni, C. Rossetti, G. Lucignani, A. Bartorelli, and F. Fazio. 'Spatial registration of echocardiographic and

- positron emission tomographic heart studies'. *European journal of nuclear medicine*, 22(3):243–247, 1995.
- [81] I. G. Zubal, S. S. Spencer, J. S. Khurseed Imam, E. O. Smith, G. Wisniewski, and P. B. Hoffer. 'Difference images calculated from ictal and interictal technetium-99m-HMPAO SPECT scans of epilepsy'. *Journal of nuclear medicine*, 36(4):684–689, 1995.
- [82] G. E. Christensen, A. A. Kane, J. L. Marsh, and M. W. Vannier. 'Synthesis of an individual cranial atlas with dysmorphic shape'. In *Mathematical methods in biomedical image analysis*, pages 309–318, Los Alamitos, CA, 1996. IEEE computer society press.
- [83] R. J. Evans, A. J. Strong, A. C. F. Colchester, J. Zhao, and K. Holton-Tainter. 'Clinical accuracy of the VISLAN surgical navigation and guidance system'. In H. U. Lemke, M. W. Vannier, K. Inamura, and A. G. Farman, editors, *Computer assisted radiology*, volume 1124 of *Excerpta medica - international congress series*, pages 799–804, Amsterdam, 1996. Elsevier.
- [84] A. C. Evans, D. L. Collins, P. Neelin, and T. S. Marrett. 'Correlative analysis of three-dimensional brain images'. In R. H. Taylor, S. Lavall'ee, G. C. Burdea, and R. M'osges, editors, *Computer-integrated surgery, Technology and clinical applications*, chapter 6, pages 99–114. MIT Press, Cambridge, MA, 1996.
- [85] H. Erbe, A. Kriete, A. J'odicke, W. Deinsberger, and D. B'oker. '3D-ultrasonography and image matching for detection of brain shift during intracranial surgery'. In H. U. Lemke, M. W. Vannier, K. Inamura, and A. G. Farman, editors, *Computer assisted radiology*, volume 1124 of *Excerpta medica - international congress series*, pages 225–230, Amsterdam, 1996. Elsevier.
- [86] S. Fang, R. Raghavan, and J. T. Richtsmeier. 'Volume morphing methods for landmark based 3D image deformation'. In M. H. Loew and K. M. Hanson, editors, *Medical Imaging: Image processing*, volume 2710, pages 404–415, Bellingham, WA, 1996. SPIE.
- [87] R. Rubinstein, H. Karger, U. Pietrzyk, T. Siegal, J.M. Gomori, and R. Chisin. 'Use of 201thallium brain SPECT, image registration, and semi-quantitative analysis in the follow-up of brain tumors'. *European journal of radiology*, 21:188–195, 1996.
- [88] Z. He, J. C. Maublant, J. C. Cauvin, and A. Veyre. 'Reorientation of the left ventricular long-axis on myocardial transaxial tomograms by a linear fitting method'. *Journal of nuclear medicine*, 32:1794–1800, 1991.
- [89] F. Fontana, A. Crovetto, M. Bergognoni, and A. M. Casali. 'Multiresolution registration for volume reconstruction in microscopical applications'. In M. H. Loew, editor, *Medical imaging: image processing*, volume 1898, pages 55–60, Bellingham, WA, 1993. SPIE Press.
- [90] T. Ault and M. W. Siegel. 'Frameless patient registration using ultrasonic imaging'. In *Medical robotics and computer assisted surgery*, pages 74–81, 1994.
- [91] K. Eilertsen, A. Skretting, and Tennvassas. 'Methods for fully automated verification of patient set-up in external beam radiotherapy with polygon shaped fields'. *Physics in medicine and biology*, 39:993–1012, 1994.
- [92] J. Thirion. 'Extremal points: definition and application to 3D image registration'. In *Computer vision and pattern recognition*, pages 587–592, Los Alamitos, CA,

1994. IEEE Computer Society press.

[93] T. Ault and M. W. Siegel. 'Frameless patient registration using ultrasonic imaging: a preliminary study'. *Journal of image guided surgery*, 1:94–102, 1995.

[94] M. Uenohara and T. Kanade. 'Vision-based object registration for real-time image overlay'. In N. Ayache, editor, *Computer vision, virtual reality, and robotics in medicine*, volume 905 of *Lecture notes in computer science*, pages 13–22, Berlin, 1995. Springer-Verlag.

[95] Y. Amit and A. Kong. 'Graphical templates for model registration'. *IEEE Transactions on pattern analysis and machine intelligence*, 18(3):225–236, 1996.

[96] C. S. Chua and R. Jarvis. '3D free-form surface registration and object recognition'. *International journal of computer vision*, 17:77–99, 1996.

[97] J. Thirion. 'New feature points based on geometric invariants for 3D image registration'. *International journal of computer vision*, 18(2):121–137, 1996.

[98] P. J. Besl and N. D. McKay. 'A method for registration of 3-D shapes'. *IEEE Transactions on pattern analysis and machine intelligence*, 14(2):239–256, 1992.

[99] K. S. Arun, T. S. Huang, and S. D. Blostein. 'Least-squares fitting of two 3-D point sets'. *IEEE Transactions on pattern analysis and machine intelligence*, 5:698–700, 1987.

[100] C. Chen, C. A. Pelizzari, G. T. Y. Chen, M. D. Cooper, and D. N. Levin. 'Image analysis of PET data with the aid of CT and MR images'. In *Information processing in medical imaging*, pages 601–611, 1987.

[101] D. N. Levin, C. A. Pelizzari, G. T. Y. Chen, C. T. Chen, and M. D. Cooper. 'Retrospective geometric correlation of MR, CT, and PET images'. *Radiology*, 169(3):817–823, 1988.

[102] A. Gu'eziec and N. Ayache. 'Smoothing and matching of 3-D space curves'. In R. A. Robb, editor, *Visualization in biomedical computing*, volume 1808 of *Proc. SPIE*, pages 259–273. SPIE Press, Bellingham, WA, 1992.

[103] H. Jiang, R. A. Robb, and K. S. Holton. 'A new approach to 3-D registration of multimodality medical images by surface matching'. In *Visualization in biomedical computing*, volume 1808, pages 196–213, Bellingham, WA, 1992. SPIE press.

[104] N. Ayache, A. Gu'eziec, J. Thirion, A. Gourdon, and J. Knoploch. 'Evaluating 3D registration of CT-scan images using crest lines'. In D. C. Wilson and J. N. Wilson, editors, *Mathematical methods in medical imaging*, volume 2035, pages 60–71, Bellingham, WA, 1993. SPIE Press.

[105] A. Collignon, T. G'eraud, D. Vandermeulen, P. Suetens, and G. Marchal. 'New high-performance 3D registration algorithms for 3D medical images'. In M. H. Loew, editor, *Medical imaging: image processing*, volume 1898, pages 779–788, Bellingham, WA, 1993. SPIE Press.

[106] D. F. Fritsch. 'Registration of radiotherapy images using multiscale medial descriptions of image structure'. PhD thesis, University of North Carolina, Chapel Hill, NC, 1993.

[107] J. C. Gee, M. Reivicj, and R. Bajcsy. 'Elastically deforming 3D atlas to match anatomical brain images. *Journal of computer assisted tomography*', 17(2):225–236, 1993.

[108] J. C. Gee, C. Barillot, L. le Bricquer, D. R. Haynor, and R. Bajcsy. 'Matching structural images of the human brain using statistical and geometrical image features'.

- In Visualization in biomedical imaging, volume 2359, pages 191–204, Bellingham, WA, 1994. SPIE press.
- [109] J. C. Gee, L. le Bricquer, and C. Barillot. ‘Probabilistic matching of brain images’. In Y. Bizais, C. Barillot, and R. di Paola, editors, Information processing in medical imaging, pages 113–125. Kluwer, 1995.
- [110] J. C. Gee, L. le Bricquer, C. Barillot, D. R. Haynor, and R. Bajcsy. ‘Bayesian approach to the brain image matching problem’. In M. H. Loew, editor, Medical imaging: image processing, volume 2434, pages 145–156, Bellingham, WA, 1995. SPIE Press.
- [111] J. C. Gee and D. R. Haynor. ‘Rapid coarse-to-fine matching using scale-specific priors’. In M. H. Loew and K. M. Hanson, editors, Medical Imaging: Image processing, volume 2710, pages 416–427, Bellingham, WA, 1996. SPIE.
- [112] K. G. A. Gilhuijs and M. van Herk. ‘Automatic on-line inspection of patient setup in radiation therapy using digital portal images’. Medical physics, 20(3):667–677, 1993.
- [113] D. L. G. Hill, D. J. Hawkes, N. A. Harrison, and C. F. Ruff. ‘A strategy for automated multimodality image registration incorporating anatomical knowledge and imager characteristics’. In H. H. Barrett and A. F. Gmitro, editors, Information processing in medical imaging, volume 687 of Lecture notes in computer science, pages 182–196, Berlin, 1993. Springer-Verlag.
- [114] J. Kittler, W. J. Christmas, and M. Petrou. ‘Probabilistic relaxation for matching problems in computer vision’. In International conference on computer vision, pages 666–673, Los Alamitos, CA, 1993. IEEE computer society press.
- [115] M. I. Miller, G. E. Christensen, Y. A. Amit, and U. Grenander. ‘Mathematical textbook of deformable neuroanatomies’. In National academy of sciences, volume 90 of Medical sciences, pages 11944–11948, 1993.
- [116] H. Rusinek, W. Tsui, A. V. Levy, M. E. Noz, and M. J. de Leon. ‘Principal axes and surface fitting methods for three-dimensional image registration’. Journal of nuclear medicine, 34:2019–2024, 1993.
- [117] W. Tsui, H. Rusinek, P. van Gelder, and S. Lebedev. ‘Fast surface-fitting algorithm for 3-D image registration’. In M. H. Loew, editor, Medical imaging: image processing, volume 1898, pages 14–23, Bellingham, WA, 1993. SPIE Press.
- [118] T. G. Turkington, R. J. Jaszczyk, C. A. Pelizzari, C. Harris, J. R. MacFall, Hoffman J. M., and R. E. Coleman. ‘Accuracy of registration of PET, SPECT and MR images of a brain phantom’. Journal of nuclear medicine, 34:1587–1594, 1993.
- [119] W. Zhao, T. Y. Young, and M. D. Ginsberg. ‘Registration and three-dimensional reconstruction of autoradiographic images by the disparity analysis method’. IEEE Transactions on medical imaging, 12(4):782–791, 1993.
- [120] R. Bajcsy, R. Lieberman, and M. Reivich. ‘A computerized system for the elastic matching of deformed radiographic images to idealized atlas images’. Journal of computer assisted tomography, 7(4):618–625, 1983.
- [121] A. Guéziec. ‘Large deformable splines, crest lines and matching’. In International conference on computer vision, pages 650–657, Los Alamitos, CA, 1993. IEEE computer society press.
- [122] G. Taubin. ‘An improved algorithm for algebraic curve and surface fitting’. In International conference on computer vision, pages 658–665, Los Alamitos, CA,

1993. IEEE computer society press.

[123] C. Davatzikos and J. L. Prince. 'Brain image registration based on curve mapping'. In IEEE workshop on biomedical image analysis, pages 245–254, Los Alamitos, CA, 1994. IEEE computer society press.

[124] D. MacDonald, D. Avis, and A. C. Evans. 'Multiple surface identification and matching in magnetic resonance images'. In Visualization in biomedical computing, volume 2359, pages 160–169, Bellingham, WA, 1994. SPIE press.

[125] S. Sandor and R. Leahy. 'Matching deformable atlas models to preprocessed magnetic resonance brain images'. In International conference on image processing, pages 686–690, Los Alamitos, CA, 1994. IEEE computer society press.

[126] B. C. S. Tom, S. N. Efstratiadis, and A. K. Katsaggelos. 'Motion estimation of skeletonized angiographic images using elastic registration'. IEEE Transactions on medical imaging, 13(3):450–460, 1994.

[127] M. Bro-nielsen. 'Modelling elasticity in solids using active cubes – application to simulated operations'. In N. Ayache, editor, Computer vision, virtual reality, and robotics in medicine, volume 905 of Lecture notes in computer science, pages 535–541, Berlin, 1995. Springer-Verlag.

[128] E. Bainville, G. Champlédox, P. Cinquin, V. Dessenne, A. Hamadeh, J. Troccaz, S. Lavallée, O. Péroia, P. Sautot, and R. Szeliski. 'Anatomical surfaces based 3D/3D and 3D/2D registration for computer assisted medical interventions'. In L. Beolchi and M. H. Kuhn, editors, Medical imaging: analysis of multimodality 2D/3D images, volume 19 of Studies in health, technology and informatics, pages 53–69. IOS Press, Amsterdam, 1995.

[129] J. Mangin, V. F. Tupin, I. Bloch, R. Rougetet, J. Régis, and J. L'opez-krahe. 'Deformable topological models for segmentation of 3D medical images'. In Y. Bizais, C. Barillot, and R. di Paola, editors, Information processing in medical imaging, pages 153–164. Kluwer, 1995.

[130] S. R. Sandor and R. M. Leahy. 'Towards automated labelling of the cerebral cortex using a deformable atlas'. In Y. Bizais, C. Barillot, and R. di Paola, editors, Information processing in medical imaging, pages 127–138. Kluwer, 1995.

[131] G. Borgefors. 'Hierarchical chamfer matching: a parametric edge matching algorithm'. IEEE Transactions on pattern analysis and machine intelligence, 10:849–865, 1988.

[132] P. A. van den Elsen, E. J. D. Pol, and M. A. Viergever. 'Medical image matching—a review with classification'. IEEE Engineering in medicine and biology, 12(1):26–39, 1993.

[133] M. Kass, A. Witkin and D. Terzopoulos. Snakes: 'Active contour models' Int. Journal of Computer Vision, 1(4):312–331, 1988

[134] G. Behiels, D. Vandermeulen, P. Seutens. 'Statistical shape model-based segmentation of digital X-ray Images' IEEE, 2000

[135] A. L. Yuille, D. S. Cohen and P. Hallinan 'Feature extraction from images using deformable templates' Proc. Comp. Vision Patt. Rec. 1989 ; pp.104–109

[136] L. H. Staib and J. S. Duncan, 'Parametrically Deformable Contour Models' IEEE Computer Society conference on Computer Vision and Pattern Recognition, San Diego, 1989

- [137] D. G. Lowe, 'Fitting Parameterized Three Dimensional Models to images' IEEE PAMI 1991;5, pp.441-450
- [138] T.F. Cootes and C. J. Taylor 'Active Shape Models – Smart Snakes' Department of Medical Biophysics, University of Manchester
- [139] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham ' Training Models of shape from sets of examples', Department of Medical Biophysics University of Manchester
- [140] J. C. Gower, 'Generalized Procrustes Analysis' Psychometrika, 40, 1975, 33-51
- [141] R. Bajcsy and F. Solina, 'Three-dimensional object representation revisited' Proc. 1<sup>st</sup> Int. Conf. Comput. Vision (London), June 8-11, 1989, pp. 231-240
- [142] E. Hall, J. Tio, C. McPherson, and F. Sadjadi 'Measuring curved surfaces for robot vision' Comput. Vol. 15, no.12, pp. 42-54, Dec. 1982
- [143] O. D. Faugeras and M. Hebert, 'The representation, recognition and locating of 3-D shapes' Int. J. Robotic Res. Vol. 5, no. 3, pp. 27-52, Fall 1986
- [144] G. H. Golub and C. F. Van Loan, Matrix Computations. Baltimore, MD: Johns Hopkins Univ. Press
- [145] K. S. Arun, T.S. Huang and S.D. Blostein 'Least square fitting of two 3-D point sets' IEEE Trans. Patt. Anal. Machine Intell. Vol. PAMI-9, no. 5, pp. 698-700
- [146] P.H. Schonemann ;A generalized solution to the orthogonal procrustes problem' Psychometrika, vol. 31, no. 1,1966
- [147] J. T. Schwartz and M. Sharif ,Identification of partially obscured objects in two and three dimensions by matching noisy characteristic curves' Int. J. Robotics Res., vol. 6, no. 2, pp. 29-44, 1987
- [148] R. M. Haralick et al. 'Pose estimation from corresponding point data' in Machine Vision for Inspection and Measurement(H. Freeman, Ed.) New York: Academic, 1989
- [149] 'Closed-form solution of absolute orientation using unit quaternions' J.Opt. Soc. Amer. A vol. 4, no.4, pp. 629-642, April. 1987
- [150] R. Szeliski 'Estimating motion from sparse range data without correspondence' 2<sup>nd</sup> Int. Conf. Comput. Vision(Tarpon Springs, FL,) Dec. 5-8, 1988, pp. 207-216
- [151] B. K. P. Horn and J. G. Harris 'Rigid body motion from range image sequences; Comput. Vision Graphics Image Processing, 1989
- [152] B. Kamgar-Parsi, J. L. Jones and A. Rosenfeld, 'Registration of multiple overlapping range images: Scenes without distinctive features' Proc. IEEE Comput. Vision Pattern Recognition Conference (San Diego,CA), June 1989
- [153] S. Z. Li. 'Inexact matching of 3-D surfaces 'VSSP-TR-3-90, Univ. of Surrey, England, 1990
- [154] E. G. Gilbert and C. P. Foo 'Computing the distance between smooth objects in 3-D space' RSD-TR-13-88, Univ. of Michigan, Ann Arbor, 1988
- [155] E. G. Gilbert, D. W. Johnson, S.S.Keerthi, 'A fast procedure for computing the distance between complex objects in 3-space' IEEE J. Robotics Automat., vol. 4, pp. 193-203, 1988
- [156] D. N. Levin, C. A. Pelizzari, G. T. Y. Chen, C.T. Chen, and M. D. Cooper. 'Retrospective geometric correlation of MR, CT, and PET images'. Radiology, 169(3):817–823, 1988.
- [157] D. L. G. Hill, D. J. Hawkes, and C. R. Hardingham. 'The use of anatomical knowledge to register 3D blood vessel data derived from DSA with MR images'. In

Medical imaging: image processing, volume 1445, pages 348–357, Bellingham, WA, 1991. SPIE press.