COMPUTER-AIDED CEPHALOMETRIC LANDMARK IDENTIFICATION

YANG MEI

NATIONAL UNIVERSITY OF SINGAPORE
2004

COMPUTER-AIDED CEPHALOMETRIC LANDMARK IDENTIFICATION

YANG MEI
(B.Eng.(Hons.), NTU)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2004

ACKNOWLEDGEMENT

The author would like to express her deepest gratitude to her supervisors, A/P Ong Sim Heng and A/P Kelvin W. C. Foong, for their professional guidance, precious advises and patient explanation during the process of the project. Their invaluable technical knowledge shared with me is greatly appreciated.

The author would also like to thank Doctor. Zulfikri Salikin from Faculty of Dentistry, NUS for his great advice and assistance on the project.

Special thanks to Faculty of Dentistry, NUS, without the medical images they provided, the project cannot be completed.

Last but not least, the author would like to express gratitude to the officer in the Vision & Image Processing laboratory, Mr. Francis Hoon for being so friendly, supportive and helpful during the progress of this project.

CONTENTS

ACKNOWLEDGEMENT i

CONTENTS ii

List of Figures iv

List of Tables vii

SUMMARY viii

Chapter 1 INTRODUCTION 1

1.1 Introduction to Cephalometrics 1

1.2 Motivation 8

1.3 Objective 9

1.4 Thesis Organization 10

Chapter 2 LITERATURE REVIEW 12

2.1 Overview of Automatic Cephalometric Landmarking 12

2.2 Current Techniques 13

2.2.1 Knowledged-based approach 13

2.2.2 Target recognition approach 15

2.2.3 Spatial spectroscopy approach 16

2.2.4 Fuzzy detection approach 16

2.2.5 Feature subimage extraction approach 17

2.2.6 Active shape model approach 18

2.2.7 3-D cephalometric landmarking 18

2.3 Proposed Approach 20

Chapter 3 2-D CEPHALOMETRIC LANDMARK DETECTION 23

3.1 Overview of 2-D Landmark Detection Procedure 23

3.2 Wavelet Feature Extraction 25

3.2.1 Wavelet transform 25

3.2.2 Wavelet decomposition 26

3.2.3 Feature extraction for landmark detection 29

3.3 Neural Network 32

3.3.1 Neural network fundamentals 32

3.3.2 Multilayer perceptron 35

3.3.3 MLP structure 37

3.4 Genetic Algorithm 39

3.4.1 Genetic algorithm fundamentals 39

3.4.2 Parameter setting 42

Chapter 4 DETERMINATION OF 3-D LANDMARK POSITIONS 44

4.1 3-D Coordinate System 44

4.2 Sprays of Lines from X-Ray Source to the Film 45

4.3 Correction of Magnification 47

4.4 Correction of Rotational Distortion 51

4.5 Construction Procedure 54

Chapter 5 PROTOTYPE CEPHALOMETRIC SYSTEM 57

5.1 System Overview 57

5.2 System Functions 59

5.2.1 Neural network training function 59

5.2.2 Parameter setting 60

5.2.3 3-D landmark position determination 62

5.2.3.1 Landmark line generation 63

5.2.3.2 Wire frame generation 66

Chapter 6 RESULTS AND DISCUSSION 68

6.1 Results of 2-D Landmark Detection 68

6.1.1 Selected landmarks 68

6.1.2 Testing of landmark detection on geometrical shapes 70

6.1.3 Results of 2-D automatic landmark detection 75

6.1.4 Texture of landmark pattern from wavelet decomposition 78

6.1.5 Texture of landmark patterns and automatic landmarking performances 81

6.1.6 Neural network training speed and texture of landmark patterns 83

6.1.7 Comparison with other automatic landmarking methods 86

6.2 Performance of 3-D Cephalometric Landmark Position Computation 88

6.2.1 Result of 3-D landmark position computation 88

6.2.2 Evaluation on dry skulls 90

6.2.3 Evaluation on test block 91

Chapter 7 CONCLUSION AND RECOMMENDATION 94

7.1 Conclusion 94

7.2 Recommendations 95

Reference 97

Appendix A CEPHALOMETRIC LANDMARKS

Appendix B TEST IMAGES AFTER LANDMARK DETECTION

List of Figures

FIGURE 1-1. A LATERAL VIEW CEPHALOGRAM. ---------- 2
FIGURE 1-2. PATIENT IN A CEPHALOSTAT. ---------- 3
FIGURE 1-3. CONVENTIONS IN TAKING CEPHALOGRAMS. ---------- 4
FIGURE 1-4. POSTEROANTERIOR (LEFT) AND BASAL (RIGHT) PROJECTION CEPHALOGRAMS. ---------- 5
FIGURE 1-5. BIPLANAR AND COPLANAR 3-D CEPHALOMETRY. ---------- 6
FIGURE 1-6. LANDMARKS ON LATERAL AND POSTEROANTERIOR CEPHALOMETRIC RADIOGRAPH. ---------- 8
FIGURE 2-1. GEOMETRY OF AN IDEALIZED 3-D LANDMARK RECONSTRUCTION: INTERSECTION OF TWO APPROXIMATELY PERPENDICULAR LINES IN SPACE. --------- 19
FIGURE 3-1. OVERVIEW OF 2-D LANDMARK DETECTION PROCEDURE. ---------- 24
FIGURE 3-2. THE DWT IMAGE DECOMPOSITION STEP. ---------- 27
FIGURE 3-3. (A) OUTPUT SUBIMAGES OF ONE-LEVEL WAVELET DECOMPOSITION. (B) A ONE-LEVEL WAVELET DECOMPOSITION OF A REAL IMAGE. ---------- 28
FIGURE 3-4. (A) OUTPUT SUBIMAGES OF THREE-LEVEL WAVELET DECOMPOSITION. (B) A THREE-LEVEL WAVELET DECOMPOSITION OF A REAL IMAGE. ---------- 29
FIGURE 3-5. HAAR WAVELET DECOMPOSITION OF X-RAY IMAGE. ---------- 30
FIGURE 3-6. (A) LOW FREQUENCY SUBIMAGE I_{Low}. (B) RECOMBINED HIGH FREQUENCY SUBIMAGE I_{High}. ---------- 31
FIGURE 3-7. STRUCTURE OF A NEURAL CELL IN THE HUMAN BRAIN. ---------- 32
FIGURE 3-8. STRUCTURE OF A NEURON IN A NEURAL NETWORK. ---------- 34
FIGURE 3-9. STRUCTURE OF A COMMONEST TYPE OF NEURAL NETWORK. ---------- 35
FIGURE 3-10. NN INPUT PATTERN DEFINITION. ---------- 37
FIGURE 3-11. A NEARLY IDEAL NN RESPONSE. ---------- 38
FIGURE 3-12. FLOW CHART OF THE GA. ---------- 40
FIGURE 3-13. CLASSIC CROSSOVER. ---------- 41
FIGURE 3-14. CLASSIC MUTATION. ---------- 42
FIGURE 3-15. MODIFICATION OF CROSSOVER FUNCTION. ---------- 43
FIGURE 4-1. 3-D COORDINATE SYSTEM. ---------- 45

FIGURE 4-2. SPRAYS OF LINES REPRESENTING THE X-RAY BEAMS FROM THE TWO X-RAY SOURCES TO THE LATERAL (LAT) AND POSTEROANTERIOR (PA) VIEW FILMS IN A BIPLANAR X-RAY SYSTEM. ---- 46
FIGURE 4-3. CORRECTION OF MAGNIFICATION. ---- 47
FIGURE 4-4. TRIANGULATION FOR CORRECTION OF MAGNIFICATION. ---- 48
FIGURE 4-5. ERROR CORRECTION FOR CORRECTING MAGNIFICATION DISTORTION. ---- 51
FIGURE 4-6. 2-D COORDINATE SYSTEMS. ---- 52
FIGURE 4-7. CORRECTION OF ROTATIONAL DISTORTION. ---- 53
FIGURE 4-8. OVERVIEW OF 3-D LANDMARK POSITION CONSTRUCTION PROCEDURE. ---- 55
FIGURE 5-1. OVERVIEW OF THE COMPUTER-AIDED CEPHALOMETRIC PROTOTYPE SYSTEM. ---- 58
FIGURE 5-2. NEURAL NETWORK TRAINING FUNCTION. ---- 60
FIGURE 5-3. PARAMETER-SETTING INTERFACE. ---- 61
FIGURE 5-4. COMPUTATION OF 3-D LANDMARK POSITIONS. ---- 63
FIGURE 5-5. LANDMARK LINES OF CEPHALOGRAMS. ---- 64
FIGURE 5-6. GENERATION OF LANDMARK LINES. ---- 64
FIGURE 5-7. WIRE FRAME. ---- 66
FIGURE 5-8. WIRE FRAME GENERATION RULE. ---- 67
FIGURE 6-1. LANDMARKS DETECTED BY 2-D AUTOMATIC LANDMARKING. ---- 69
FIGURE 6-2. TEST IMAGES: SET A, WITH REGULAR GEOMETRICAL SHAPES. ---- 71
FIGURE 6-3. TEST IMAGES: SET B, WITH REGULAR GEOMETRICAL SHAPES IN LOWER RESOLUTION. ---- 72
FIGURE 6-4. TEST IMAGES: SET C, WITH IRREGULAR GEOMETRICAL SHAPES. ---- 72
FIGURE 6-5. NEURAL NETWORK RESPONSES OF TESTING GEOMETRICAL SHAPES SET A. 74
FIGURE 6-6. VARIATION OF PERFORMANCE BETWEEN LANDMARKS. ---- 77
FIGURE 6-7. VARIATION OF PERFORMANCE BETWEEN CEPHALOGRAMS. ---- 78
FIGURE 6-8. T_{1L} AND T_{2L} IN I_{Low}. ---- 80
FIGURE 6-9. T_{1H} AND T_{2H} IN I_{High}. ---- 81
FIGURE 6-10. I_{low} TEXTURE FEATURE AND DETECTION PERFORMANCE OF LANDMARK. 82
FIGURE 6-11. I_{high} TEXTURE FEATURE AND DETECTION PERFORMANCE OF LANDMARK. 82
FIGURE 6-12. TRAINING ERROR CHANGES FOR LANDMARK IS DURING NN TRAINING --- 84
FIGURE 6-13. NN TRAINING SPEED AND I_{low} TEXTURE FEATURE. ---- 85

FIGURE 6-14. NN TRAINING SPEED AND I_{high} TEXTURE FEATURE. ---------- 86
FIGURE 6-15. COMPARISON WITH HAN'S ALGORITHM. ---------- 88
FIGURE 6-16. ORTHOGONAL CEPHALOGRAMS FOR 3-D LANDMARK POSITION COMPUTATION. ---------- 89
FIGURE 6-17. TEST BLOCK FOR 3-D LANDMARK POSITION COMPUTATION. ---------- 92

List of Tables

TABLE 5-1. LANDMARKS INVOLVED IN 3-D LANDMARK POSITION COMPUTATION. 67
TABLE 6-1. SELECTED LANDMARKS. 69
TABLE 6-2. DETAILS OF GEOMETRICAL SHAPE VARIATION. 72
TABLE 6-3. RESULTS OF GEOMETRICAL SHAPE TESTING 73
TABLE 6-4. SYSTEM PERFORMANCE FOR EACH LANDMARK. 76
TABLE 6-5. SYSTEM PERFORMANCE FOR EACH TEST IMAGE. 77
TABLE 6-6. TEXTURE OF LANDMARK PATTERN FROM WAVELET DECOMPOSITION. 80
TABLE 6-7 NEURAL NETWORK ERRORS REDUCED (ER) IN THE FIRST 100 LITERATIONS . 85
TABLE 6-8. SUMMARY OF EXISTING AUTOMATIC LANDMARKING METHODS. 87
TABLE 6-9. 3-D LANDMARK POSITION COMPUTATION EVALUATION ON DRY SKULL. 90
TABLE 6-10. 3-D LANDMARK POSITION COMPUTATION EVALUATION ON TEST BLOCK. .. 92
TABLE A-1. CEPHALOMETRIC LANDMARKS
TABLE B-1. 2-D LANDMARK DETECTION RESULT

SUMMARY

Cephalometrics is a technique that employs oriented radiographs to make head measurements. It makes use of a number of characteristic reference points in terms of landmarks to characterize the geometry of the skull. Locating landmarks manually is tedious and repeated for each patient. Therefore, the motivation exists to automate this process.

Automatic cephalometric landmarking is a difficult task since the cephalograms of patients vary significantly from one to another. Moreover, the difficulty of controlling the process of image acquisition can result in deformation and distortion. If automatic landmarking can be realized accurately enough, orthodontists can make treatment planning more effective by omitting the troublesome manual tracing of landmarks. Computer-aided automatic landmarking can also make significant contributions to 3-D cephalometric landmarking. Patients can be exempt from the CT scanning, which involves high radiation and high cost.

We propose an automatic landmarking approach that employs a multilayer perceptron (MLP) neural network, a genetic algorithm (GA) [1,2] and also a wavelet technique to detect the 2-D landmark positions. Wavelet decomposition is used to preprocess the subimages within the area of interest in a cephalogram to separate the high pass and low pass features. This is a feature extraction procedure. It makes neural network training easier and faster. The grey level values of the preprocessed parts of subimages are the inputs to the neural network. The MLP neural network is trained to indicate whether a landmark is present in the centre of a subimage. Then it could play the role

of a fitness function for the GA. Subimages of the original radiographs are grouped as the population set for a GA search. The one with the best fitness, which has the highest possibility of containing a landmark, will survive to the last.

3-D cephalometry makes use of two projections of cephalograms to compute 3-D cephalometric landmark positions. In this study, we use the lateral and posteroanterior projection cephalograms to develop our 3-D landmarks position computation method. Our computation is defined based on the study of the cephalometric apparatus being used.

Ten lateral view cephalograms have been used to test our 2-D landmark detection method. The detection results have confirmed the effectiveness of this method. On average, 40% of the landmarks have been detected within 1mm of their accurate positions and 80% of the landmarks are detected within 2mm of their accurate positions. Our experiment on 3-D landmark position computation shows that 3-D cephalometrics should be realized on cephalograms that all acquired simultaneously in exact orthogonal positions.

Chapter 1

INTRODUCTION

This chapter gives the background knowledge of cephalometrics, states the motivation and objective of this study, and presents the organization of this thesis.

1.1 Introduction to Cephalometrics

Clinicians are faced with the problem of developing systems of knowledge that are more powerful than common sense and more practical than intuition. Because skeletal dysplasia[1] is associated with a high percentage of severe malocclusions[2], craniofacial morphology must be analyzed carefully before treatment begins. The analysis is carried out routinely in orthodontics by means of a standardized radiographic procedure called cephalometrics [3].

[1] Abnormal development or growth of tissues, organs, or cells.
[2] Faulty contact between the upper and lower teeth when the jaw is closed.

Cephalometrics is a technique that employs oriented radiographs to make head measurements. It provides a quantitative assessment or description of craniofacial morphology and abstracts the complexities of the live human head into a geometric scheme. Its purpose is a comparison with norms, ideals or one's self. These comparisons help in the diagnosis of anomalies, monitoring of treatments, and understanding and predictions of craniofacial growth.

The current practice of cephalometrics utilizes a 2-D radiographic image on film, the **cephalogram** [3]. Figure 1-1 shows a lateral view cephalogram.

Figure 1-1. A lateral view cephalogram.

From the cephalogram, a cephalometric analysis is derived in which anatomic structures are reduced to a number of characteristic reference points. These reference points or landmarks are the output of cephalometrics. They are used to indicate shapes

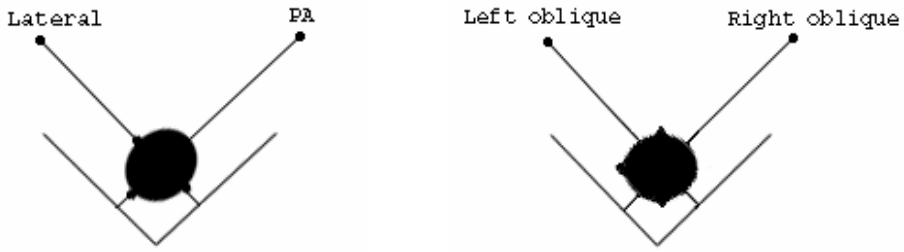
and relative locations of curves that could characterize the geometry of the skull. Cephalometric analysis could then be carried out by computing the related data (describing the geometry of the skull) based on landmark information for diagnosis or treatment.

A cephalometric apparatus consists of a cephalostat (or head holder), an x-ray source, and a cassette holder (see Figure 1-2).

Figure 1-2. Patient in a cephalostat [4].

The cephalograms can be taken in different conventions. The ***lateral projection cephalograms*** are the most commonly used in clinical orthodontic practise (see Figure 1-1). They are taken with the left side of patient's head toward the film. The central beam of the x-rays coincides with the transmeatal-axis[3]. Other conventions involve ***posteroanterior (PA or frontal) projection cephalograms***, ***basal projection cephalograms*** and ***oblique projections cephalograms***.

[3] The axis joins the two ear rods of the cephalostat.



a b

Figure 1-3. Conventions in taking cephalograms [5].

Figure 1-3 shows lateral, PA and oblique conventions in taking cephalograms. The titles of the conventions are indicated at the positions of the x-ray sources. Patients' heads are drawn with positions of their ears and noses marked. Posteroanterior projection cephalograms are taken by rotating the head 90 degrees to the lateral projection so that the central ray perpendicularly bisects the transmeatal axis. The right and left oblique projection cephalograms are taken at 45 and 135 degrees to the lateral projection, the central ray entering behind one ramus to obviate superimposition of the halves of the mandible. Basal projection cephalograms are taken by further rotating the head 90 degrees to the posteroanterior projections so that the head is facing upwards with the central ray still perpendicularly bisects the transmeatal axis [3].

Figure 1-4 shows samples of posteroanterior and basal projection cephalograms of a dry skull.

Figure 1-4. Posteroanterior (left) and basal (right) projection cephalograms.

Orthodontists paid little attention to cephalograms other than lateral projection ones shortly after the first introduction of posteroanterior cephalograms. This was because the clinical problems encountered by most orthodontists were similar on both sides of the bone structure [6] at that time. They appeared to be adequately recorded by the lateral projection alone. In recent years, as orthodontists have become “craniofacial orthopedists” treating more severe, often asymmetric craniofacial anomalies, the limitations of the lateral cephalogram have become obvious. Other conventions of cephalograms have to be taken into consideration also.

3-D cephalometry makes use of two projections of cephalograms to compute 3-D cephalometrics landmark positions. There are basically two major methods of 3-D

cephalometry, which are different in the projections of cephalograms they make use of and also the way they position the x-ray images in the system. These two methods are the biplanar and coplanar methods. Figure 1-5 shows the geometric relationship between the biplanar and coplanar x-ray system. S1 and S2 stand for the positions of x-ray sources. The patients' heads are drawn with positions of their ears and noses marked. Figure 1-5a shows the biplanar x-ray system. Figure 1-5b, c, d are coplanar x-ray systems with different x-ray tubes and skull positions [5, 7].

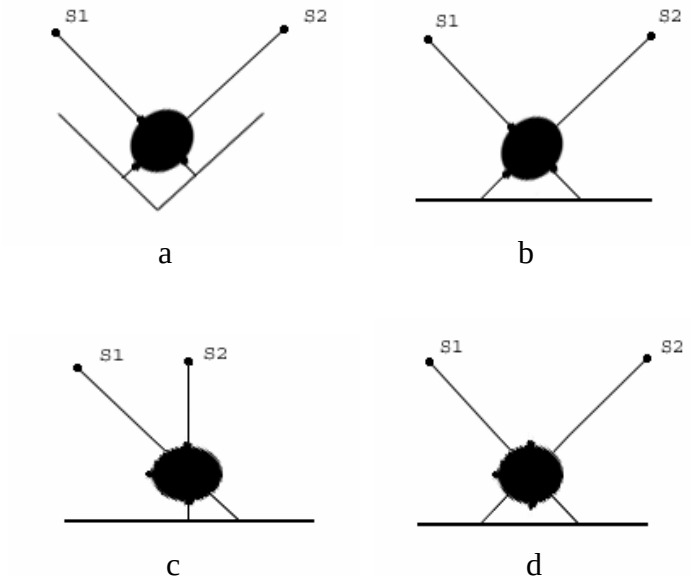


Figure 1-5. Biplanar and coplanar 3-D cephalometry [5].

In the biplanar x-ray system, lateral and posteroanterior view cephalograms are used. The paired x-ray images are oriented in two planes at right angles to each other. While for coplanar x-ray system, oblique and lateral (or posteroanterior) view cephalograms are used. Two images of each x-ray pair are positioned in the same plane rather than being at right angles to each other [5, 7].

In general, it is much easier to identify any given landmark in the two x-ray images in a coplanar system. It is true for manual identification of landmarks since the appearance on the images of any landmark differs much less in the two films of a coplanar system than it does in the two orthogonal positioned images of a biplanar pair. But in cases where the landmark has already been reliably identified on both images, the solution for 3-D landmark location is mathematically stronger for biplanar system [5].

In this study, we mainly consider the lateral and posteroanterior projection cephalograms for they are more commonly used and the craniofacial landmarks are well defined in these two projections. Our 3-D landmarks position computation method was developed based on a biplanar x-ray system using the lateral and posteroanterior view cephalograms. This is because the final target of this research is to compute 3-D landmarks positions base on automatic detected 2-D landmark positions. The advantage of the coplanar system doesn't apply since the 2-D position of landmarks on the two x-ray images are detected completely separately [5].

Figure 1-6 shows a set of standard cephalometric landmarks on lateral and posteroanterior projection cephalograms. A list of commonly used cephalometric landmarks is given in appendix A [8, 9].

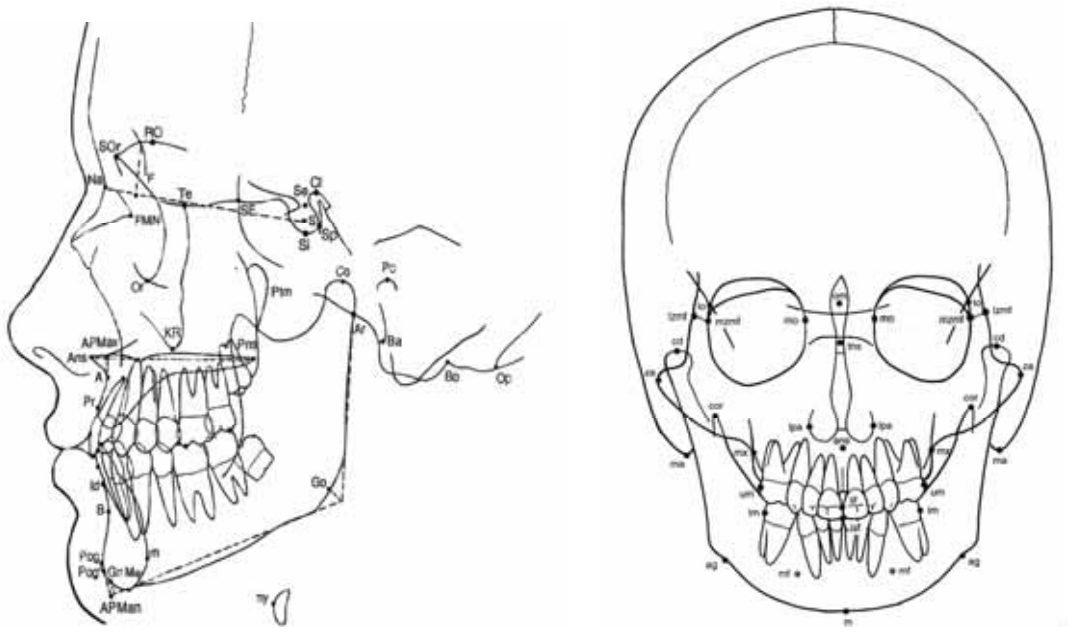


Figure 1-6. Landmarks on lateral and posteroanterior cephalometric radiograph [10].

1.2 Motivation

The identification of the craniofacial landmarks plays an essential role in treatment planning and diagnosis by orthodontists. Conventionally these landmarks (typically about 30) are located by hand or digitized using a graphics tablet by manual tracing. The process, which takes approximately 10 minutes, is tedious, operator dependent and is repeated for each patient. Therefore, the motivation exists to automate this repetitive and time-consuming process.

Image processing techniques in engineering could help to automate this process. Automatic landmark detection will free the orthodontists from the tedious manual process of landmark tracing. It is certainly more practical for quantitative or long-term serial analysis. In fact, investigations in automatic landmark recognition using different image processing techniques have been undertaken for several years. Craniofacial

landmark detection is a difficult task since no two x-ray images look alike and images are of variable quality. The first extensive use of image processing techniques to segment cephalometric lateral skull radiographs was done in 1984 [11]. A great number of other attempts have taken place after that. A review on the major approaches is given in Chapter Two.

It is important that the orthodontists make use of the best available resources to allow informed decisions on treatment to be reached. Cephalometrics is a technique of making 3-D head measurement [6]. Both orthodontic treatment and research concern the figure of the skull in all three dimensions of space. 3-D landmark information allows a more comprehensive assessment of facial changes occurring with orthodontic treatment than two-dimensional analyses. The availability of 3-D geometric information will provide a more complete picture to the clinician for research and surgery purposes. Current methods of extracting dimensional information from single planar radiographs have limited the geometric information to the clinician to 2-D. (Computed tomography (CT) scanning could possibly provide such information, but the patient would suffer not only from high radiation but also from the high cost.) Another contribution of image processing techniques in cephalometrics is the construction of the 3-D environment of cephalograms by employing more than one cephalogram convention. A review of this topic is presented in Chapter Two.

1.3 Objective

The objective of this study is to develop a computer-aided cephalometric landmark identification system for measuring skull dimensions through an automated process of

landmark detection and projection. The system should automatically detect the landmark positions on 2-D cephalograms. It should also be able to compute 3-D landmark positions using 2-D information from orthogonal radiographic projections (lateral and posteroanterior projections). This would then provide the landmark information clearly in 3-D environment for further cephalometric analysis.

1.4 Thesis Organization

The thesis is divided into 7 chapters:

Chapter 1: "Introduction": This chapter presents the introduction on cephalometrics, and describes the motivation and objective of this study. The organization of the report is also included.

Chapter 2: “Literature Review": This chapter presents an overview of automatic cephalometric landmarking, a review of the current automatic landmarking methods and an explanation of some major techniques. The approaches applied in this study are presented at the end of this chapter.

Chapter 3: "2-D Cephalometric Landmark Detection": This chapter describes our algorithm for automatic landmark detection on 2-D cephalograms. It includes an introduction to neural networks, genetic algorithms (GAs) and wavelet techniques, which are the major techniques employed in this study. It emphasizes the features of those techniques that had made them valuable in this study. It also states the principle and working procedure on how these techniques are combined to realize the detection.

Chapter 4: “Determination of 3-D Landmark Positions": This chapter explains how 3-D landmark position is computed based on the 2-D position from different cephalogram projections. The explanation is focused on the factors that could affect the computation accuracy and have to be considered in the calculations.

Chapter 5: " Prototype Cephalometric System": The major functions of the computer-aided cephalometric landmark identification prototype system are presented in this chapter.

Chapter 6: "Results and Discussion": The experimental results are presented in this chapter. The discussion is carried out by comparing our method with existing methods, both theoretically and experimentally.

Chapter 7: "Conclusion and Recommendation": This chapter gives a conclusion on the study and suggests avenues for further work.

Chapter 2

LITERATURE REVIEW

This chapter reviews different automatic cephalometric landmarking techniques and discusses the approach taken in this study.

2.1 Overview of Automatic Cephalometric Landmarking

In the sixty-year history since cephalometric radiology has been developed, many methods of analysis have been proposed, that have contributed to a better understanding of the complexity of changes associated with facial growth. The manual process of cephalometric landmarking is very tedious, time-consuming, and also heavily dependent on the orthodontist's experience. An available automated cephalometric landmarking system will offer many advantages that could benefit both the orthodontist and the patient.

Methods in computer-aided cephalometric landmark detection can be indirect (the detection of the current landmarks depends on the previous detection of other

landmarks) or direct (the detection of the current landmarks does not depend on the previous detection of other landmarks), 2-D or 3-D.

The earliest research on automatic cephalometric landmark detection works as knowledge-based automatic cephalometric landmark detection. It is considered as an indirect method since it follows a strategy similar to that used by orthodontists. The detection of the current landmark is affected by the previous detection of other landmarks, which limits detection accuracy and many even result in termination of the process by one failure. The approaches suggested later, such as the target recognition approach and feature subimage extraction approach, do not suffer from this problem. All the landmarks are located directly at the same time without relying on each other. By incorporating different projections of the radiographs, the computer-aided automatic cephalometric landmarking is now moving from 2-D to 3-D. Although research in this area is still limited, the advantages of a 3-D system have made it a promising research area. A review on the major approaches of automatic cephalometric landmarking is presented in the following section.

2.2 Current Techniques

2.2.1 Knowledged-based approach

Knowledge-based approaches of automating cephalometric analysis employ a combination of image processing techniques to extract the important edges first. Landmark locations are then found on the edges using a geometrical description of the landmark. A number of this type of systems has been developed, the most successful

one being the system implemented by Parthasarathy *et al.* [12]. It algorithmically implemented the knowledge-based algorithm proposed by Lévy-Mandel *et al.* [13]. In the system developed Lévy-Mandel *et al.* [13], edges were tracked in the image to locate landmarks on structures with well-defined outlines, such as the lower border of the mandible. *A priori* knowledge of the typical shape of the important edges was encoded in algorithms that followed the boundaries of different structures. Parthasarathy *et al.* [12] presented a similar scheme but used the pyramid method to decrease the spatial resolution of the image for improving the efficiency of their search and then enhanced the local image to extract useful edges or lines. Tong *et al.* [14] presented an extension to the work of Parthasarathy *et al.* to locate landmarks on soft tissue. Contreras-Vidal *et al.* [15], Jackson *et al.* [16] and Cohen *et al.* [17, 18] also approached the problem with some form of knowledge-based edge tracking. Davis and Taylor [19] described how knowledge-based algorithms could be integrated in the blackboard architecture, allowing backtracking in the face of contradictions.

Lévy-Mandel *et al.* tested their system on two high-quality cephalograms. Only landmarks that lay on or near to edges in the image could be located. They reported that 23 out of 36 landmarks could be determined on a good quality image [13]. Parthasarathy's system was tested on five cephalograms of varying quality. They compared the accuracy of their system to the landmarks as placed by two experts. Of ten landmarks, on average 18 percent were located to within 1 mm, 58 percent within 2 mm and 100 percent within 5 mm [12]. (It has been suggested that an error of 2 mm in landmark placement is acceptable [20] but an accuracy of 1 mm is desirable [21, 22]). Tong *et al.* indicated that their results could be combined to yield a full cephalometric analysis [14]. Five cephalograms were used for testing. On average 40 percent of 27

landmarks would be located to within 1 mm, 70 percent to within 2 mm and 95 percent to within 5 mm [14]. A formal evaluation by Forsyth and Davis [21] used 10 cephalograms to test their system. On average 63 percent of 19 landmarks were located to within 1 mm and 74 percent to within 2 mm [21].

In all these earlier works it was not stated whether the algorithms was performed with the testing images unseen. However to evaluate an image-understanding algorithm it must be shown that it will perform acceptably on new images, not just on those that have been used for designing the method. This flaw in scientific method means that direct comparison of the results of other study with those of the studies mentioned above is perhaps inadvisable.

2.2.2 Target recognition approach

An alternative scheme suggested by Cardillo and Sid-Ahmed [23, 24] attempted to locate the landmarks directly. Based on grey-scale mathematical morphology they developed a target recognition algorithm to locate landmarks. Training is required to define the structuring elements and the probability distributions of their original location. A statistical approach to training was used to overcome subtle differences in skeletal topographies, and decomposition was used to desensitise the algorithm to size differences. From a training set of 40 images, 76% of their 20 landmarks were located to within 2 mm.

2.2.3 Spatial spectroscopy approach

Rudolph *et al.* [25] described the use of spatial spectroscopy to characterize the grey-level appearance around each landmark from a training set of images. They compared their system's performance with that of an expert-using image of the same resolution. They reported that no statistical difference could be found between the manual error and the error of their automated system. This implied that 100 percent of the landmarks were located to within 4 mm. However, the images they used were just 64×64 pixels. They suggested that as the resolution increased, the landmarks would correspondingly become more accurate but this remains to be proven.

The methods mentioned in Sections 2.2.2 and 2.2.3 used essentially the same approach; i.e., generating a model of the grey-levels around each point from a training set and then matching this model to a new image to locate the points of interest. Other matching techniques, such as cross-correlation [26] may also be used for the same purpose. These methods rely on the image appearance around each landmark but do not take enough care on the nature of varying data for different skulls.

2.2.4 Fuzzy detection approach

Sanei *et al.* introduced a fuzzy detection of cephalometric landmarking in 1997 [27]. Proper matrices of suitable size were used as templates to match the selected candidates of the features and fuzzy decision-making was used to verify relative possible positions of the landmarks due to the variation of their grey-level appearance. They have announced that 30 out of 36 landmarks have been detected by their system.

They claimed an accuracy of 96% but did not report the details of how this figure was obtained.

2.2.5 Feature subimage extraction approach

Chen *et al.* used neural networks together with GAs to search for subimages that contained each of the cephalometric landmarks [1, 2]. GAs are a part of evolutionary computing, which is a rapidly growing area of artificial intelligence. The algorithm is started with a set of solutions (represented by chromosomes) called the population. Solutions from one population are taken and used to form a new population. This is motivated by a hope that the new population will be better than the old one. Solutions that are selected to form new solutions (offspring) are selected according to their fitness - the more suitable they are the more chances they have to reproduce. The GA's behaviour provides an approximation within an acceptable tolerance, instead of an exact match. A full-connected multiplayer perceptron (MLP) of three layers with error back-propagation learning is employed in their study as the fitness function. The accuracy of their system was presented and compared with the cross-correlation method [26] they used before. It was reported that the accuracy improved from over 80% to more than 90%. However, this improvement was achieved based on the detection of only 9 landmarks and again no detailed explanation on how these figures were calculated.

2.2.6 Active shape model approach

Most recently, Tim *et al.* have evaluated the accuracy of the active shape models (ASMs) approach as applied to automatic cephalometric landmarking [28]. ASMs were described by Cootes *et al.* [29, 30] and were first reported by Hill *et al.* [31]. They used a model of the spatial relationships between the important structures - a template - to help search the image for features of interest. The key innovation was that the variation in shape is modelled, enabling the synthesis of plausible new examples of the structures seen. They have concluded that this implementation did not give sufficient accuracy for completely automated landmarking but provided a framework for a range of improvements to be made.

2.2.7 3-D cephalometric landmarking

Previous research has concentrated on the detection of landmarks for the lateral cephalograms. It should be mentioned that from the very first introduction of the cephalostat, Broadbent and Bolton [32] have already stressed the importance of coordinating different conventions of cephalograms to arrive at a distortion-free definition of craniofacial form.

3-D cephalometrics is an expedient use of existing cephalostat data sets to derive certain analysis of the 3-D form. Figure 2-1 shows the geometry of an idealized 3-D landmark reconstruction: intersection of two approximately perpendicular lines (the lateral projection and posteroanterior projection) in space. The mathematics of making

3-D measurements from paired x-ray images involves correction for origin shift, correction for magnification, and correction for rotation.

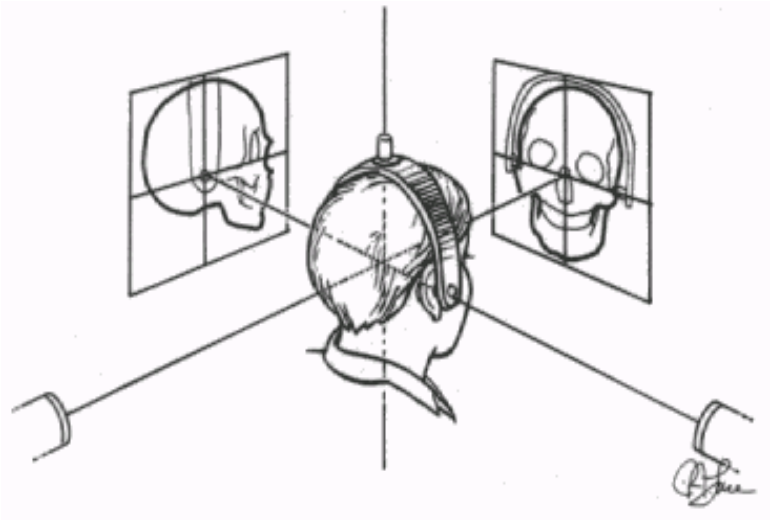


Figure 2-1. Geometry of an idealized 3-D landmark reconstruction: intersection of two approximately perpendicular lines in space [6].

Researchers working in this area [5, 6, 7, 33] have based their computations on manually traced landmark positions on the cephalograms. Some researchers from Japan [34] have developed a cephalometric system to compute the 3-D position of some landmarks, but again this is based on manually selected 2-D landmark positions.

Computer-aided method can also make significant contributions to 3-D cephalometric landmarking. If it can make accurate 3-D measurement of human head, orthodontists can make treatment planning more effective by omitting the troublesome manual tracing of landmarks; moreover patients are exempt from the CT scanning, which involves high radiation and high cost. However, research in this field has been limited.

2.3 Proposed Approach

The main problem with knowledge-based automatic landmark detection is that it requires consistent good quality cephalograms. The algorithms have difficulties in detecting the landmarks if a cephalogram contains extra lines, which is often the case created by the imperfect superimposition of the symmetric features on either side of the face. Furthermore if the cephalogram is poorly exposed, resulting in fragmented lines, the critical lines defining the landmarks may not be tracked properly.

The actual selection of landmarks used by each orthodontist varies with personal preference and experience. It is important for an automatic cephalometric landmarking system to be able to detect the significant landmarks. Typically an average of thirty-six landmarks are used roughly in cephalometrics. These previous knowledge-based methods are weak in locating landmarks which do not lie on the edges, so the number of landmarks detected is limited.

“Direct” automatic landmark detection methods, such as “target recognition approach”, “spatial spectroscopy approach”, “fuzzy detection approach”, require the definition of structuring elements or templates to represent each landmark. The problem is that the algorithms cannot be expected to recognize a wide range of possible shapes using a single model. In order words, the methods are not able to provide a good estimate with acceptable tolerance to compensate for the differences between patients’ radiographs. Another “direct” automatic landmark detection method, the “active shape models approach”, does not give acceptable accuracy.

The method of “feature subimage extraction approach” by Chen *et al.* [1,2] employs neural networks together with GAs to search for landmark positions. This method provides acceptable tolerance and accuracy but their neural network training procedure involves too much computational complexity.

Base on the above analysis of existing methods, we propose a novel approach that employs a multilayer perceptron (MLP) neural network, a GA [1,2] and also a wavelet technique to detect the 2-D landmark positions.

A wavelet-based approach is used to preprocess the subimages within the area of interest in a cephalogram. The original subimage will be decomposed and recombined into two parts that contain the high pass and low pass features separately. This is a feature extraction procedure. It makes neural network training easier and faster compared to [1,2]. The grey level values of the preprocessed parts of subimages are the inputs to the neural network. The MLP neural network is trained to indicate whether a landmark is present in the centre of a subimage. Then it could play the role of a fitness function for the GA. Subimages of the original radiographs are grouped as the population set for a GA search. The one with the best fitness, which has the highest possibility of containing a landmark, will survive to the last. One more difference between our approach and [1,2] is that the crossover function of our GA does not create “unknown” subimages. Instead the “child” is generated as an existing subimage between its parents. This is to prevent the algorithm from being confused by “unknown” noise when two far away “parents” cross over each other. It is very important especially when neural networks are not so well trained.

This method is able to give a good approximation with acceptable tolerance to compensate for the differences between patients' radiographs. Our investigations showed that the method is robust and possesses good accuracy.

The mathematics of computing 3-D landmark position involves three parts, as presented in Section 2.2.7. Our computation is also defined based on the study of the cephalometric apparatus being used. A few assumptions are made to simplify the computation procedure. The methods are detailed in Chapter 4.

Chapter 3

2-D CEPHALOMETRIC LANDMARK DETECTION

This chapter explains in detail the proposed algorithm for landmark detection in 2-D cephalograms. It presents the principle and working procedure of how various techniques have been combined together to realize the detection.

3.1 Overview of 2-D Landmark Detection Procedure

Figure 3-1 gives an overview of the proposed 2-D landmark detection procedure. There are two phases in this procedure: the training phase and the testing phase. The main techniques employed in these two phases are wavelet decomposition, neural network and genetic algorithm (GA).

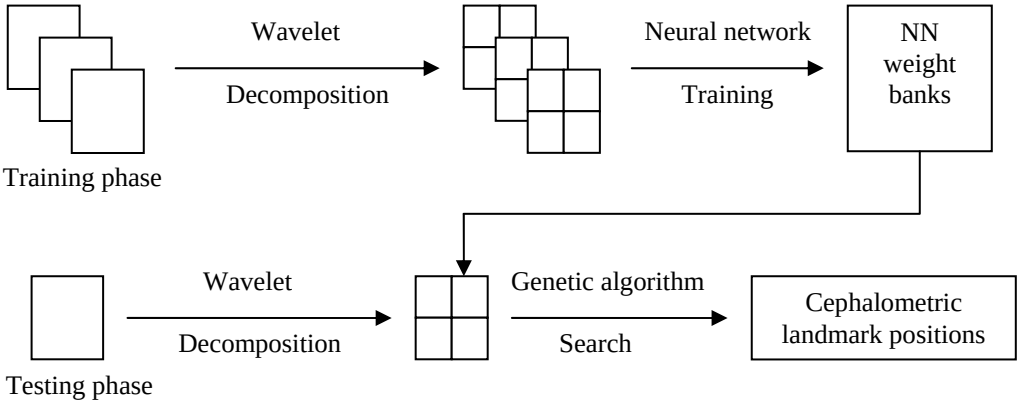


Figure 3-1. Overview of 2-D landmark detection procedure.

In the training phase, neural networks are trained to recognize the prescribed structure of each landmark pattern in order to determine the landmark position. The “correct” and “wrong” subimages are selected from the training images around each landmark. A “correct” subimage is the one that has the specified landmark in its central area, while a “wrong” subimage does not. All these subimages are then decomposed by the wavelet technique resulting in high-pass and low-pass subimages containing high frequency and low frequency information. After the recombination of the high-pass subimages, the gray-level values of the recombined high-pass subimages and low-pass subimages are used as inputs for training the neural networks.

The testing phase is the procedure for landmark detection. Here, the neural network and GA are combined and used as the search technique. In this phase, a search area, which is big enough to contain the landmark, is predefined for each landmark. The size of this search area is determined statistically based on the variation of landmark positions of the training images. The subimages of the test images within the search area are decomposed by wavelet techniques to form the first generation of population

for the GA search. The evolution of the GA aims to produce a final population with prescribed structures that characterize each landmark. The fitness function describing these prescribed structures of landmarks is provided by the neural network.

3.2 Wavelet Feature Extraction

The wavelet decomposition procedure is responsible for extracting the low frequency and high frequency information from the images and providing suitable training and testing input data for neural network. This step is able to help the neural network recognize the prescribed feature pattern of each landmark and speed up the training and testing procedure.

3.2.1 Wavelet transform

For the Fourier transform, sinusoidal waves are used as the orthonormal basis functions. These functions are nonzero over their entire domain. However, many important features in images such as edges are highly localized spatially. Such components do not resemble any of the Fourier basis functions. A new way to construct useful transforms is to use basis functions of limited duration. These basis functions vary in position as well as frequency. They are waves of limited duration and are called wavelets [35, 36].

The wavelet transform is a very useful tool in the analysis of non-stationary signals such as seismic signals. This is due to the ability of the wavelet transform to resolve

features at various scales. The continuous-time wavelet transform (CWT) of a function $x(t)$ is defined as [35, 36]

$$W_f(a,b) = \int f(x)\psi_{a,b}(x)dx \tag{3.1}$$

where $\psi(x)$ is called the basic wavelet. It is normally a oscillatory function centered on the origin and dies out rapidly as $|x| \to \infty$. Then a set of basis functions is generated by dilating (scaling) and translating $\psi(x)$. $\psi_{a,b}(x)$ is defined by [35, 36]

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \int \psi(\frac{x-b}{a}) \tag{3.2}$$

The dilation parameter a is known as the scale factor, and b is a translation factor. The wavelet is dilated or compressed by the scale factor. Thus, at low scales high frequency behaviour is localized, while at high scales (when the wavelet is stretched out) low frequency features are better resolved. This is of significant benefit when one is dealing with signals containing features with various frequency characteristics. Another advantage of the wavelet transform is that the analyzing wavelet can be chosen based on the application. By extending (3.1) to 2-D, the 2-D wavelet transform is [35, 36]

$$W_f(a,b_x,b_y) = \int_y \int_x I_c(x,y)\psi_{a,b_x,b_y}(x,y)dxdy \tag{3.3}$$

3.2.2 Wavelet decomposition

Since the data of the image is discrete-time, the DWT (discrete wavelet transform) is used instead of the continuous wavelet transform (CWT). In practice the transform is computed by applying a separable filter bank to the image. The decomposition step is shown in the following diagram [35, 36].

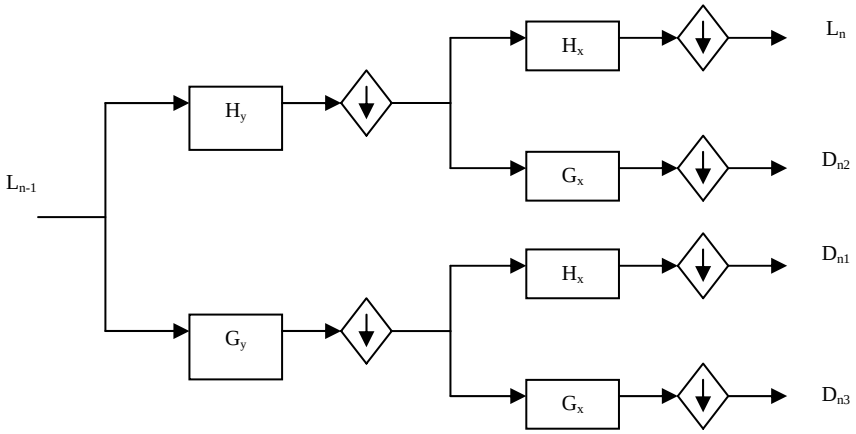


Figure 3-2. The DWT image decomposition step.

The wavelet decomposition of a 2-D image is obtained by performing the filtering consecutively along the horizontal and vertical directions. It can also be represented by the following equations:

$$L_n(\overrightarrow{x,y}) = [H_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\overrightarrow{x,y}) \tag{3.4}$$

$$D_{n1}(\overrightarrow{x,y}) = [H_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\overrightarrow{x,y}) \tag{3.5}$$

$$D_{n2}(\overrightarrow{x,y}) = [G_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\overrightarrow{x,y}) \tag{3.6}$$

$$D_{n3}(\overrightarrow{x,y}) = [G_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\overrightarrow{x,y}) \tag{3.7}$$

where $*$ denotes the convolution operator, $\downarrow_{2,1}$ ($\downarrow_{1,2}$) denotes down sampling along the rows (columns) and $L_0 = I(\overrightarrow{x,y})$ is the original image. H and G are the low and high pass filters, respectively. L_n is obtained by low pass filtering and is therefore referred to as the low resolution image at scale n. The D_{ni}'s are obtained by high pass filtering in a specific direction and this contains detailed directional information at scale n. Figure 3-3(a) shows the four output subimages of one-level wavelet decomposition and Figure 3-3(b) a one-level wavelet decomposition of a real image. Subimage I_{LL} contains the low frequency information of the original image.

I_{HL}, I_{LH} and I_{HH} contain the high frequency information in the horizontal, vertical and diagonal directions, respectively.

I_{LL}	I_{HL}
I_{LH}	I_{HH}

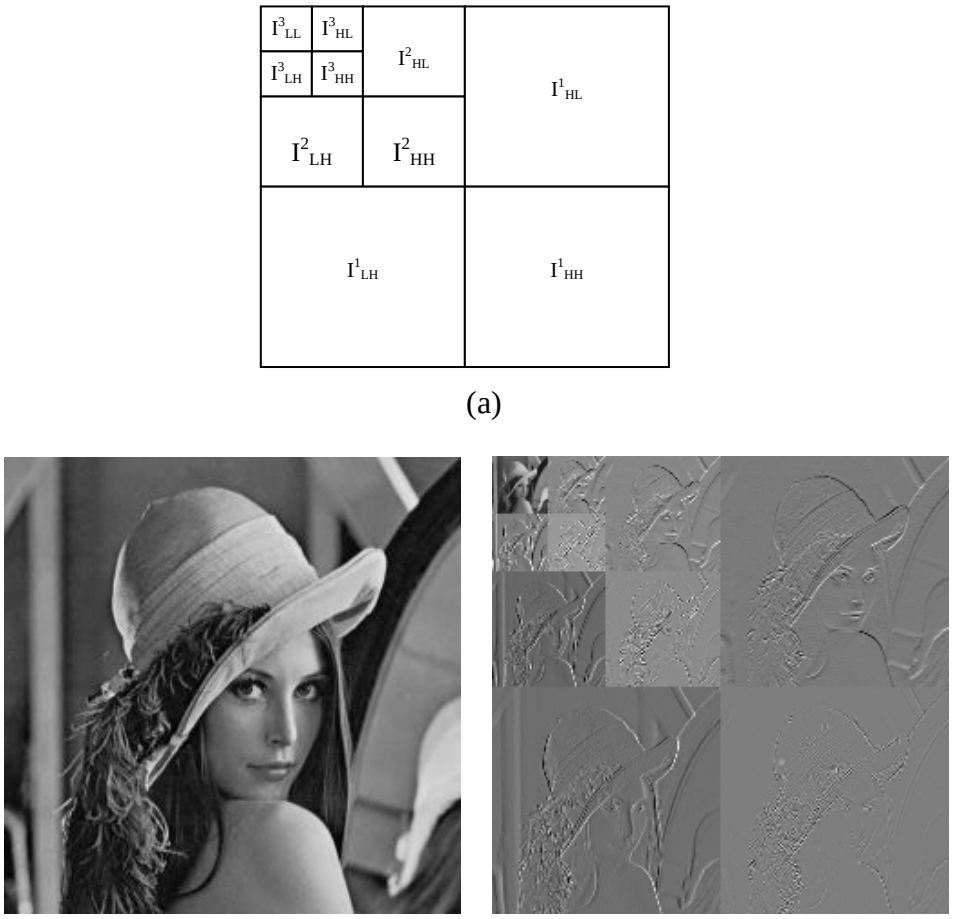
(a)

(b)

Figure 3-3. (a) Output subimages of one-level wavelet decomposition. (b) A one-level wavelet decomposition of a real image.

A one-level wavelet transform decomposes the images to a low-pass subimage and three high-pass subimages. A two-level wavelet transform will decompose an image into 8 subimages. Hence for n-level of decomposition, a total of $4n$ sub-bands is returned. The more times the discrete wavelet transform is applied to the image, the greater the amount of detail that will be separated from the image. The remaining low-pass sub-band of the image will diminish when more transformations are performed. Figure 3-4(a) illustrates the output subimages of three-level wavelet

decomposition. Figure 3-4(b) shows a three-level wavelet decomposition of a real image.



(a)

(b)

Figure 3-4. (a) Output subimages of three-level wavelet decomposition. (b) A three-level wavelet decomposition of a real image.

3.2.3 Feature extraction for landmark detection

Wavelet decomposition is able to extract low frequency and high frequency information from images. One of the most important high frequency information would be the edge information, which is also a significant clue for landmark detection since most of the important cephalometric landmarks lie on or near some edges of the

cephalogram. Figure 3-5 shows a one-level wavelet decomposition of a lateral view cephalogram. The original image has been decomposed into four subimages. The high frequency information in the horizontal (I_{HL}), vertical (I_{LH}) and diagonal (I_{HH}) directions and the low frequency (I_{LL}) information have been extracted and stored in each subimage. Edge information of the cephalogram can be clearly seen in the high-pass subimages.

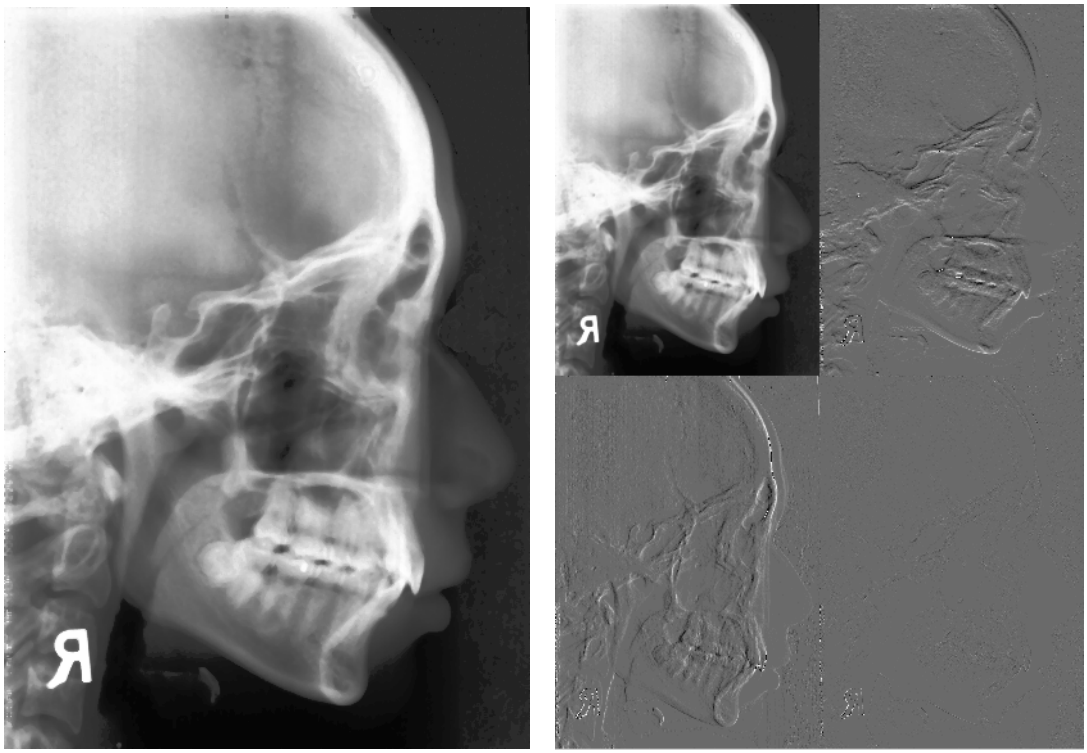
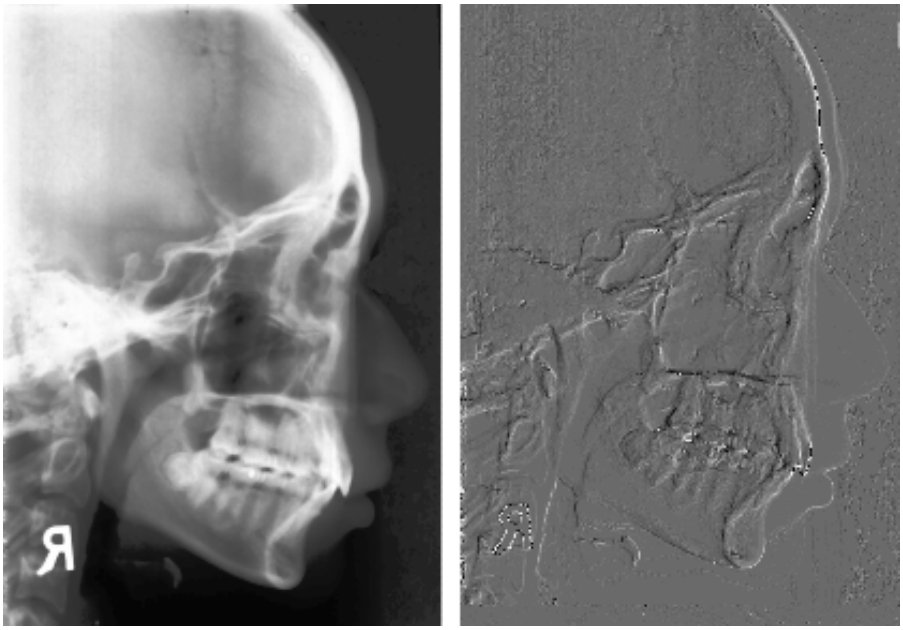


Figure 3-5. Haar wavelet decomposition of x-ray image.

During the training and testing procedures, subimages selected from the image are decomposed by the wavelet technique. One may use the decomposed subimage information directly as the input pattern to the neural network, but this would then result in an input pattern of the same size as the selected subimage. Unfortunately, the larger the input size the larger the network becomes, and the higher the computational cost. So we should reduce the input pattern size and at the same time try to retain as

much as possible the important image information. Our solution is to form the neural network input set by combining the three high-pass subimages together. So we use one decomposed low frequency subimage (I_{Low}) and one high frequency subimage (I_{High}, recombined from the three high-pass subimages) to train the neural network. Figure 3-6 shows I_{Low} and I_{High} of a cephalogram.



(a) (b)

Figure 3-6. (a) Low frequency subimage I_{Low}. (b) Recombined high frequency subimage I_{High}.

We obtain I_{High} by using the general principle that one should keep as much high frequency information as possible. Each pixel of I_{High} is set as the one that has the highest absolute value among the three corresponding pixels in the three high-pass subimages, then multipling all the pixel values by ten to compensate for the difference in values between I_{Low} and I_{High}, (As the pixel values of I_{Low} are much higher than I_{High}, the neural network may give a higher priority to I_{Low}). By doing this, we have

reduced the input data by half while still retaining the important information of the original image.

3.3 Neural Network

A neural network is utilized as the fitness function for the GA in our study. The gray-level values of the wavelet decomposed and recombined subimages (I_{Low} and I_{High}) are used as the inputs into the neural network. It is trained so that the output can indicate the possibility of the central pixel of the original subimage being a particular landmark.

3.3.1 Neural network fundamentals

The human brain consists of a large number (more than a billion) of neural cells that process information. Each cell works like a simple processor and only the massive interaction between all cells and their parallel processing makes the brain's abilities possible. Figure 3-7 shows a sketch of such a neural cell [37].

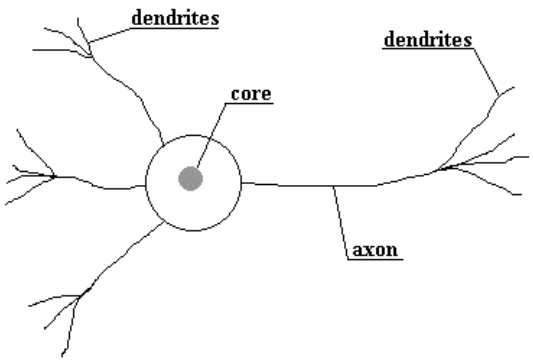


Figure 3-7. Structure of a neural cell in the human brain [37].

A neuron consists of a core, dendrites for incoming information and an axon with dendrites for outgoing information that is passed to connected neurons. Information is transported between neurons in form of electrical stimulations along the dendrites. Incoming information that reaches the neuron's dendrites is added up and then delivered along the neuron's axon to the dendrites at its end, where the information is passed to other neurons if the stimulation has exceeded a certain threshold. In this case, the neuron is said to be activated. If the incoming stimulation has been too low, the information will not be transported any further. In this case, the neuron is said to be inhibited. The connections between the neurons are adaptive, what means that the connection structure is changing dynamically. It is commonly acknowledged that the learning ability of the human brain is based on this adaptation [37].

Many tasks which seem simple for us, such as reading a handwritten note or recognizing a face, are difficult for even the most advanced computer. In an effort to increase the computer's ability to perform such tasks, programmers began designing software to act more like the human brain, with its neurons and synaptic connections. Thus the field of "artificial neural networks" was born. A neural network is an artificial representation of the human brain that tries to simulate its learning process [37, 38, 39].

A neural network is composed of interconnected "units", which serve as model neurons. The function of the synapse is modeled by a modifiable weight, which is associated with each connection. Each unit converts the pattern of incoming activities that it receives into a single outgoing activity that it broadcasts to other units. Figure 3-8 shows an idealized neuron of a neural network [37, 38, 39].

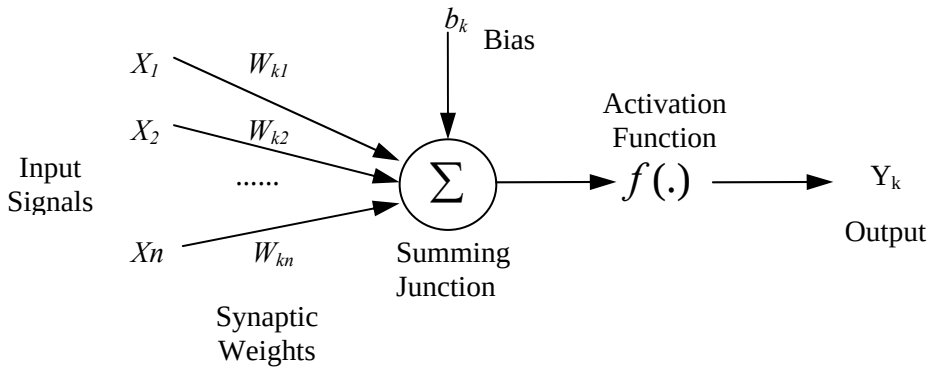


Figure 3-8. Structure of a neuron in a neural network.

A neural network performs his conversion in two steps. At first, it multiplies each incoming activity by the weight on the connection and adds together all these weighted inputs to get a quantity called the total input. Then a unit uses an input-output function that transforms the total input into the outgoing activity. The input-output function (transfer function) that is specified for the units could be linear, threshold or sigmoid. For linear units, the output activity is proportional to the total weighted input. For threshold units, the outputs are set at one of the two levels, depending on whether the total input is greater than or less than some threshold value. For sigmoid units, the output varies continuously but not linearly as the input changes. Sigmoid units bear a greater resemblance to real neurons than linear or threshold units, but all these must be considered rough approximations [37, 38, 39].

Once the data to be analyzed was collected, the network will run through them and "learn" how the inputs of each record may be related to the result. The network then continually refines itself until it can produce an accurate response when given those particular inputs. After training on a few dozen cases, the network begins to organize itself, and refines its own architecture to fit the data, much like a human brain "learns".

3.3.2 Multilayer perceptron

There are many different types of neural networks. They can be distinguished by their types (feedforward or feedback), their structures and the learning algorithms they use. The multilayer perceptron (MLP) network is a very important type of neural network. Typically it composes of a set of sensory units (source nodes) that constitute the input layer, one or more hidden layers of computation nodes, and an output layer of computation nodes. The input signal propagates through the network in a forward direction, on a layer-by-layer basis. Figure 3-9 shows a typical MLP neural network with only one hidden layer [37, 38, 39].

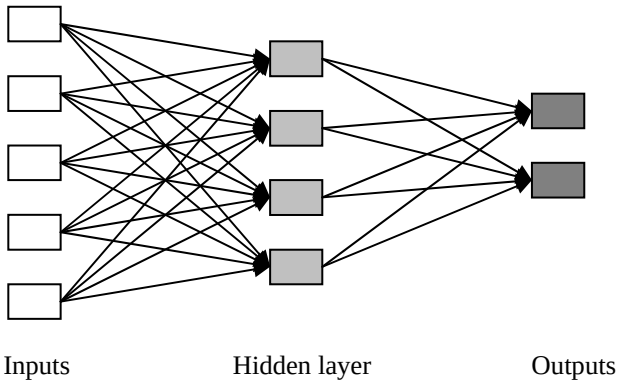


Figure 3-9. Structure of a commonest type of neural network.

The basic learning algorithm used by MLPs is called the back-propagation learning algorithm. It is one of the most powerful learning algorithms. In the application of the back-propagation algorithm, two distinct passes of computation are distinguished: the forward pass and the backward pass. In the forward pass the synaptic weights remain unaltered throughout the network, and the function signals of the network are computed on a neuron-by-neuron basis. At iteration n in layer l, the induced local

field of neuron j, $v_j^{(l)}(n)$ and the function signal appearing at the output of neuron j, $y_j^{(l)}(n)$, is computed as [1, 2]

$$v_j^{(l)}(n) = \sum_{i=0}^{m} w_{ji}^{(l)}(n) y_i^{(l-1)}(n) - \theta_j^{(l)}(n) \tag{3.8}$$

$$y_j^{(l)}(n) = \varphi_j(v_j^{(l)}(n)) \tag{3.9}$$

where $m+1$ is the total number of inputs (excluding the bias) applied to neuron j, $w_{ji}^{(l)}(n)$ is the synaptic weight connecting neuron i in layer $l-1$ to neuron j, $y_i^{(l-1)}(n)$ is the input signal of neuron j from layer $l-1$, $\theta_j^{(l)}$ is the threshold applied to node j and $\varphi(.)$ is the activation function.

The backward pass, on the other hand, starts at the output layer by passing the error signals $e_j(n) = d_j(n) - y_j(n)$ backward through the network, layer by layer, and recursively computing the local gradient δ_j for modifying the weight w and the threshold θ [1, 2]:

$$\Delta w_{ji}^{(l)} = \eta \delta_j^{(l)}(n) y_i^{(l-1)}(n) \tag{3.10}$$

$$\Delta \theta_j^{(l)} = -\eta \delta_j^{(l)}(n) \tag{3.11}$$

For node j in the output layer:

$$\delta_j^{(l)}(n) = e_j^{(l)}(n) \varphi_j'(v_j^{(l)}(n)) \tag{3.12}$$

For node j in the hidden layer:

$$\delta_j^{(l)}(n) = \varphi_j'(v_j^{(l)}(n)) \sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n) \tag{3.13}$$

where η is the learning-rate parameter of the back-propagation algorithm. $\varphi_j'(.)$ denotes differentiation with respect to the argument. The training procedure is stopped when the performance criterion (e.g., the maximum of the mean square error) is met.

3.3.3 MLP structure

A fully connected multilayer perceptron (MLP) neural network [1, 2] with one hidden layer is employed in this study as the fitness function of the GA. It has only one node in the output layer. The size of the hidden layer is 20×20. The size of the input layer is $30 \times 30 \times 2$.

During the training procedure, accurate landmark positions information is indicated by an expert orthodontist. Figure 3-10 illustrates how we determine the input pattern.

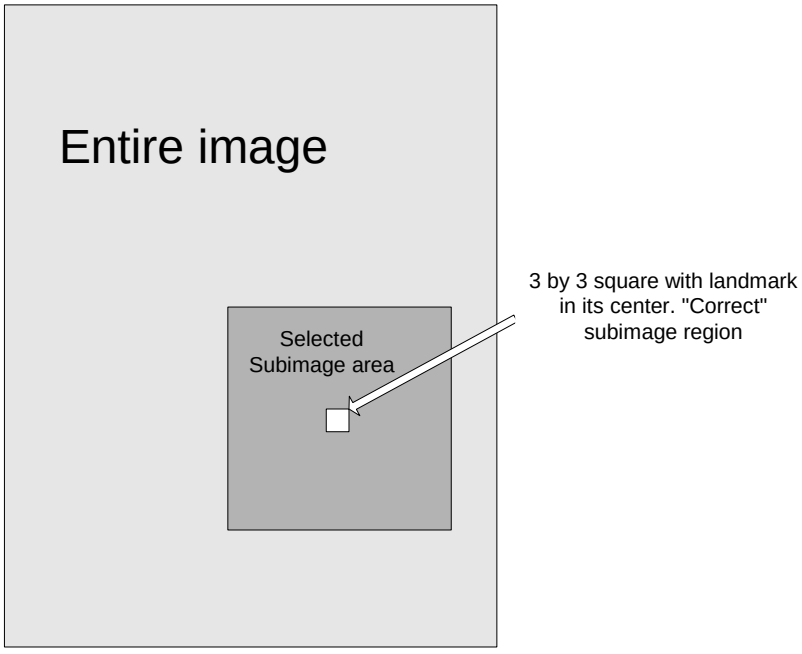


Figure 3-10. NN input pattern definition.

A small area (3×3) with the landmark in its center is considered as the "correct" subimage region. A much bigger area (300×300) around the landmark is defined as the selected subimage area. Within this area, the subimages of size 60×60 are grouped together to be decomposed by the wavelet technique and recombined to form the neural network input patterns. Any subimage with its central pixel falling into the "correct" subimage region is regarded as the "correct" input while any others are the "wrong" ones.

After training, the neural network is able to "remember" the pattern feature around the landmark and judge whether a landmark is presented in the central of a subimage so that it can act as the fitness function of the GA for landmark detection. Figure 3-11 shows a nearly ideal neural network response over a critical area (60×60) of the image. The neural network output will show a peak value at the landmark position and nearly zero elsewhere.

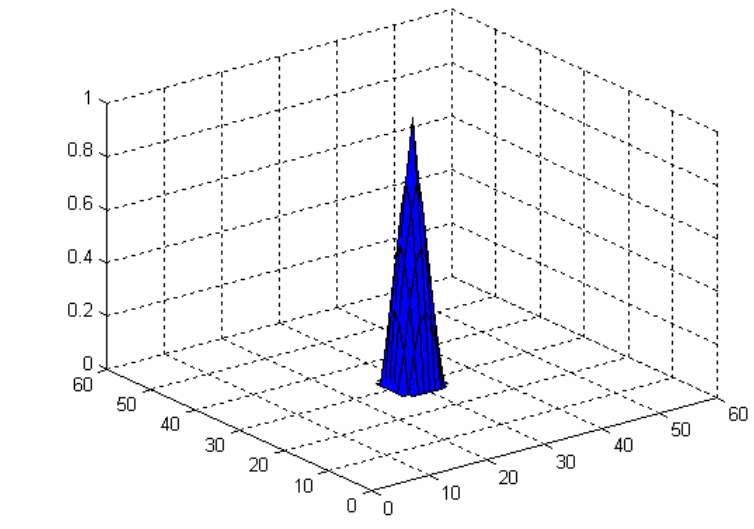


Figure 3-11. A nearly ideal NN response.

The training procedure is realized by the QuickProp algorithm of Fahlman [40] since the standard back-propagation method is very slow. The major difference between the standard back-propagation and QuickProp algorithms is that Quickprop assumes the error surface to be locally quadratic and attempts to jump in one step from the current position directly into the minimum of the parabola. It computes the derivatives in the direction of each weight. After computing the first gradient with regular back-propagation, a direct step to the error minimum is attempted by

$$\Delta w_{ji}^{(l)}(n) = \frac{S(n)}{S(n-1) - S(n)} \Delta w_{ji}^{(l)}(n-1) \qquad (3.14)$$

where $\Delta w_{ji}^{(l)}(n)$ is the actual weight change, $w_{ji}^{(l)}(n) = w_{ji}^{(l)}(n-1) + \Delta w_{ji}^{(l)}(n)$, $S(n)$ is the partial derivative of the error function by w_{ij}, $S(n) = \frac{\partial e(n)}{\partial w_{ij}}$, and $S(n-1)$ is the last partial derivative. Equation 3.14 is used instead of Equation 3.10 in our study to train the neural networks.

3.4 Genetic Algorithm

The genetic algorithm (GA) is used in our study as a search algorithm for cephalometric landmark detection.

3.4.1 Genetic algorithm fundamentals

Genetic algorithms (GAs) are adaptive heuristic search algorithms premised on the evolutionary ideas of natural selection and genetics. The basic concept of GAs is to simulate processes in natural system necessary for evolution, specifically those that

follow the principles first laid down by Charles Darwin. As such they represent an intelligent exploitation of a random search within a defined search space to solve a problem. They are modelled loosely on the principles of the evolution via natural selection, employing a population of individuals that undergo selection in the presence of variation-inducing operators such as mutation and recombination (crossover). A fitness function is used to evaluate individuals, and reproductive success varies with fitness [41, 42, 43].

To use a GA, solutions to the problem must be represented as genomes (or chromosomes). The algorithm is started with a set of solutions called the population. Current population are taken and applied genetic operators such as mutation and crossover to form a new population. This is motivated by the hope that the new population will be better than the old one. Figure 3-12 shows the basic process of the GA [41, 42, 43].

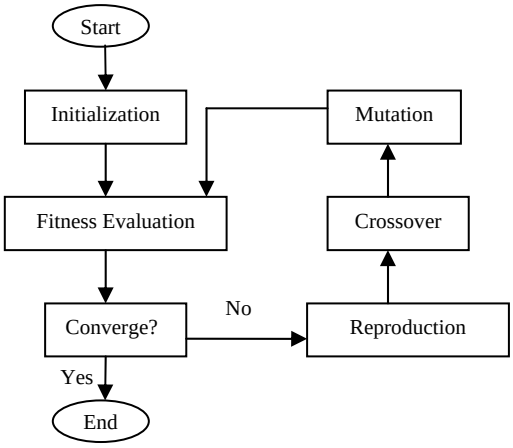


Figure 3-12. Flow chart of the GA.

Solutions that are used to form the new population (offspring) are selected according to their fitness - the more suitable they are the more chances they have to reproduce. This is the reproduction procedure, which implies the rule of natural selection, i.e. individuals with higher fitness have a greater chance to replicate themselves in the next generation, while those with lower fitness are more likely to fail.

At the crossover stage, individuals of the new population are mated, based on the probability of crossover. A crossing site was selected randomly for exchanging the partials of the mates. After reproduction and crossover, the new strings were taken from the more significant areas in the search domain. The high-performance strings will pass their characteristics to individuals in the next generation. Figure 3-13 shows the crossover procedure. Crossover and fitness-based reproduction are the processes that contribute most to evolution [43].

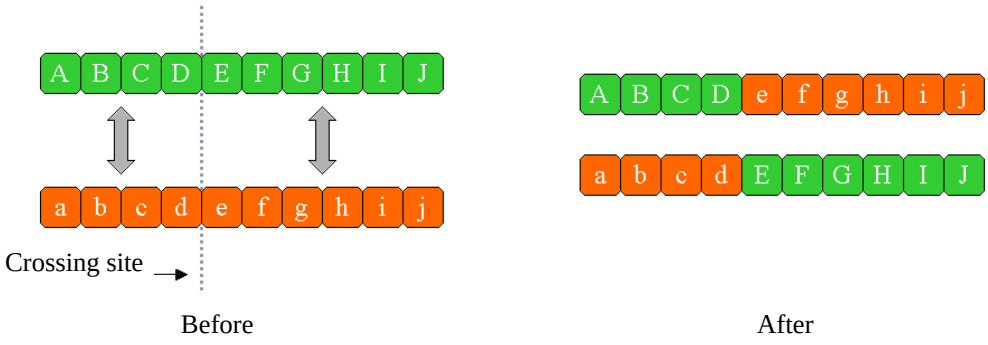


Figure 3-13. Classic crossover [43].

The mutation stage is associated with the crossover stage for alternating the value of each chromosome, based on the probability of mutation, to search for the lost potentially useful information. The importance of its role continues to be a matter of debate. Figure 3-14 shows the mutation procedure [43].

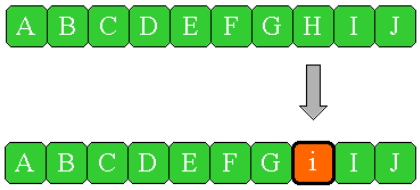


Figure 3-14. Classic mutation [43].

GAs have been widely studied, experimented and applied in many fields in the engineering worlds. Not only does the GA provide an alternative method to solving problems, it also consistently outperforms other traditional methods in many real-world problems involve finding optimal parameters, which might prove difficult for traditional methods but ideal for GAs [41].

3.4.2 Parameter setting

The evolution of the GA aims to produce a final population with prescribed structures that characterize each landmark. The fitness function describing these prescribed structures is provided by neural network training.

A GA search area from which the landmark position is to be found is predefined for each landmark before GA searching. 100 subimages inside this search area are randomly selected to form the first generation of the population. This GA search area is a critical parameter for landmark detection. It should be big enough to ensure the inclusion of each landmark of the testing images but not too big to slow down the

searching speed unnecessarily. It was evaluated statistically based on the landmark position variation between the training images.

The crossover operator of GA has been modified by us to suit our particular application. As illustrated in Figure 3-15, when subimage “a” and “c” want to crossover each other, the “child” is created as a subimage between its parents: subimage “b”. This is due to the nature of our problem in that there is not much sense to generate the new population by two “parents” far from each other; on the other hand it is quite possible to introduce noise (confusing the neural network). The probability of crossover was set to 0.6.

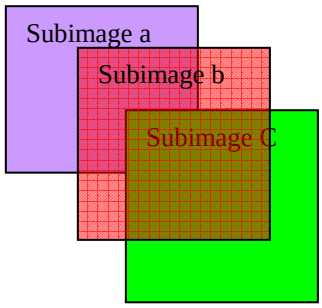


Figure 3-15. Modification of crossover function.

Based on the idea that superior chromosomes with high fitness would be more adaptable than inferior chromosomes, there is no reason for a superior chromosome to undergo a large change with a high mutation probability. So the probability of mutation is set as a variable, which is a linearly dependent on the value of fitness and is varied from 0.05 to 0.2.

Chapter 4

DETERMINATION OF 3-D LANDMARK POSITIONS

This chapter explains in detail the computation of the 3-D landmark positions based on their corresponding 2-D landmark positions in the lateral and posteroanterior cephalograms. The explanation focuses on the factors that affect computational accuracy.

4.1 3-D Coordinate System

Before the 3-D landmark position can be computed, a standard 3-D coordinate system has to be established. Figure 4-1 shows how we define it in our study.

- Porion (*Po*) is the (left or right) most superior point of the outline of the external auditory meatus (right side: *PoR*; left side: *PoL*).
- *PoC* is the midpoint of the bilateral porion.
- Orbitale (*Or*) is the (left or right) lowest point on the inferior orbital margin (right side: *OrR*; left side: *OrL*) [44].

- *OrC* is the midpoint of the bilateral orbitale.

This standard coordinate system is defined by *PoC*, *OrR* and *OrL*. The plane determined by *PoC, OrR* and *OrL* is defined as the standard plane. The coordinate system is defined as follows: *PoC* is the origin, the straight line passing through *PoC* and *OrC* is the x axis (posteroanterior direction as positive), the straight line passing through the origin and perpendicular to the standard plane is the y axis (downward as positive), and the straight line perpendicular to the x and y axes and passing through *PoR* and *PoL* is the z axis (left direction as positive). This system is defined as the standard 3-D coordinate system [34].

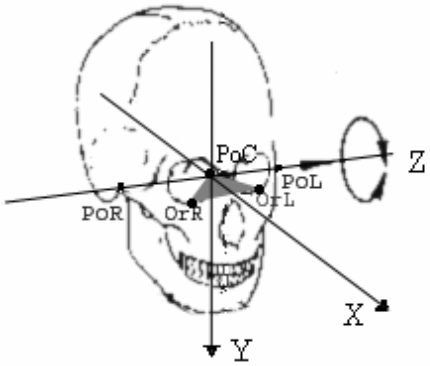


Figure 4-1. 3-D coordinate system [34].

4.2 Sprays of Lines from X-Ray Source to the Film

The cephalostatic x-ray beams connecting the x-ray sources to each landmark in the lateral and posteroanterior view films can be drawn as lines in space. Figure 4-2 shows the sprays of lines representing the x-ray beams from the two x-ray sources to the films in a biplanar x-ray system [6].

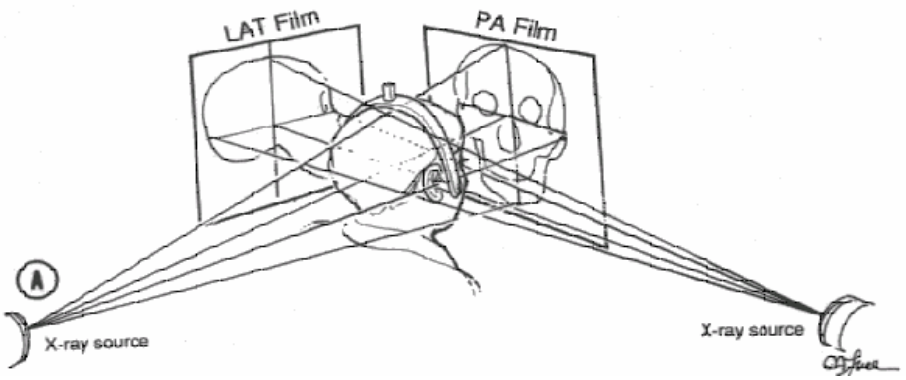


Figure 4-2. Sprays of lines representing the x-ray beams from the two x-ray sources to the lateral (LAT) and posteroanterior (PA) view films in a biplanar x-ray system [6].

The distortion inherent in the spread of the x-ray beams is known as magnification distortion. Due to magnification distortion, the 2-D landmark position projected from the patient to the cephalograms will shift from its original 2-D position. In order to build an accurate 3-D cephalometric system, this magnification distortion has to be corrected. Only the central rays of the sprays are not affected by it. The cephalometric apparatus is positioned to force the central ray of the lateral view x-ray source to pass through the two ear rods of the patient and to intersect with the central ray of the posteroanterior view x-ray source at the origin of the defined standard 3-D coordinate system. The projection of this origin on the x-ray films needs to be manually indicated on the two cephalograms before the 3-D landmark positions are computed. On the lateral view cephalogram, the 2-D projection of the origin lies at the center of the ear rod. On the posteroanterior view cephalogram, it lies on the midpoint of the line joining the two ear rods.

4.3 Correction of Magnification

Figure 4-3 shows graphically the position of the x-rays sources and the films in a biplanar x-ray system. S_1 and S_2 represent the two x-ray sources. O is the origin. $S_1^{'}$ and $S_2^{'}$ are the projections of the origin on the lateral and posteroanterior view cephalograms respectively. Then lines $S_1S_1^{'}$ and $S_2S_2^{'}$ represent the central rays from two x-ray sources. The distance from the x-ray source to the origin and the distance from the origin to the x-ray film are fixed at 60 inches (152.4 cm) and 7 inches (17.78 cm), respectively. Hence, the coordinates of S_1, $S_1^{'}$, S_2 and $S_2^{'}$ are $(0,0,-60)$, $(0,0,7)$, $(-60,0,0)$ and $(7,0,0)$ accordingly.

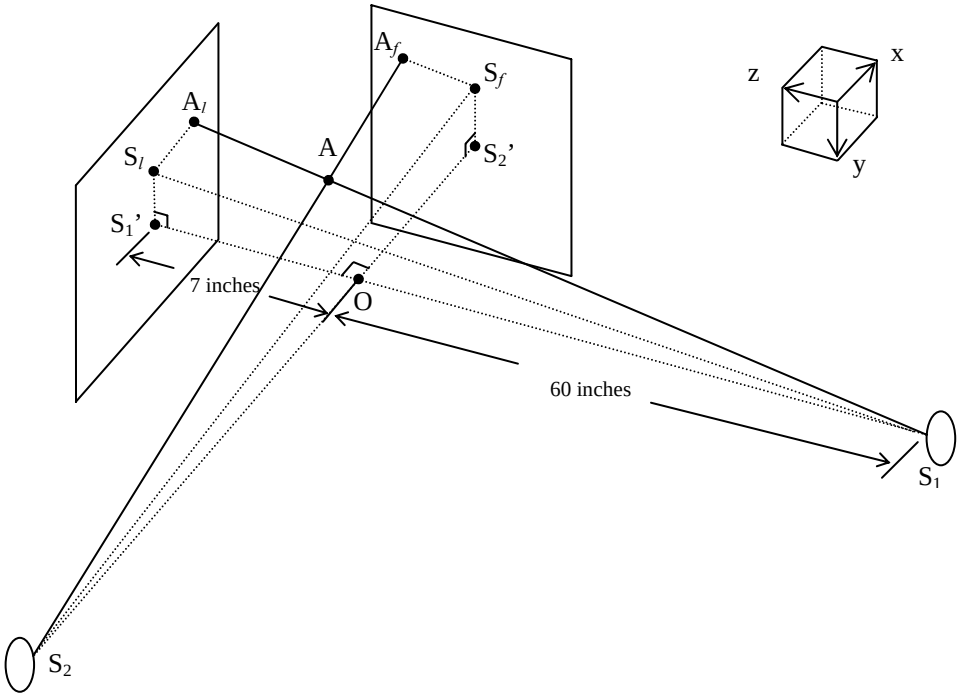


Figure 4-3. Correction of magnification.

A represents a craniofacial landmark in the patient's head. The projection of this point on the lateral and posteroanterior view cephalograms are $A_l(x_l, y_l, 7)$ and $A_f(7, y_f, z_f)$. The 3-D coordinates (x, y, z) of landmark A are the coordinates we

wish to determine without magnification distortion. $S_f A_f$ is parallel to the z axis and $S_l A_l$ to the x axis. S_f is the point that line $S_f A_f$ intersects with the mid-sagittal plane $S_2 S_2^{'} S_f$. $S_1 S_1^{'} S_l$ is the orthogonal plan of the mid-sagittal plane through the origin O. $S_l A_l$ intersects with $S_1 S_1^{'} S_l$ at point S_l.

Figure 4-4 redraws part of Figure 4-3 to emphasize the triangulation relationship in triangles $S_1 A_l S_l$ and $S_1 S_l S_1^{'}$. The coordinates of landmark A is (x, y, z). Plane $ABCD$ is parallel to the lateral view cephalogram. B is the point where plane $ABCD$ intersects line $S_l S_l$. C is the point where plane $ABCD$ intersects line $S_l S_l^{'}$.

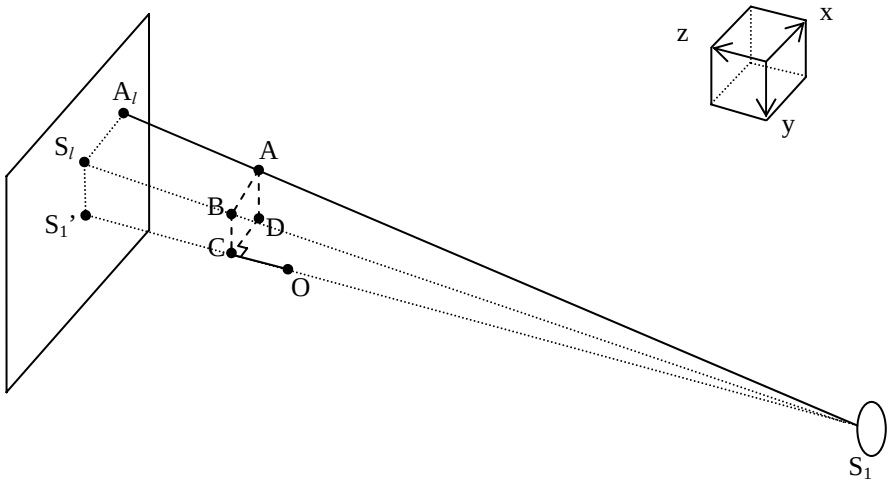


Figure 4-4. Triangulation for correction of magnification.

Based on the triangulation relationship in triangles $S_1 A_l S_l$ and $S_1 S_l S_1^{'}$, we have

$$\overline{BA} / \overline{S_l A_l} = \overline{S_l B} / \overline{S_l S_l} = \overline{BC} / \overline{S_l S_l^{'}} = \overline{S_l C} / \overline{S_l S_l^{'}} \qquad (4.1)$$

where $\overline{BA} = x$, $\overline{S_l A_l} = x_l$, $\overline{BC} = y$, $\overline{S_l S_l^{'}} = y_l$, $\overline{S_1 C} = 60 + \overline{OC}$, $\overline{S_1 S_1^{'}} = 67$ and $\overline{OC} = z$. Then we have

$$x = x_l (60 + z) / 67 \qquad (4.2)$$

$$y = y_l(60 + z)/67 \tag{4.3}$$

Similarly, based on the triangulation relations in triangles $S_2A_fS_f$ and $S_2S_fS_2^{'}$, we have

$$y = y_f(60 + x)/67 \tag{4.4}$$

$$z = z_f(60 + x)/67 \tag{4.5}$$

The coordinates (x, y, z) of A can be calculated by Equations 4.2 to 4.5. First, z is obtained with Equations 4.2, 4.3 and 4.4 [34]. Then x is calculated from Equation 4.5 and y Equation 4.3 or 4.4.

The coordinates of A can also be determined from the intersection of lines S_1A_l and S_2A_f. As lines in 3-D space, S_1A_l and S_2A_f can be represented by Equations 4.6 and 4.7

$$S_1A_l : \frac{x - x_{01}}{m_1} = \frac{y - y_{01}}{n_1} = \frac{z - z_{01}}{p_1} = t_1 \tag{4.6}$$

$$S_2A_f : \frac{x - x_{02}}{m_2} = \frac{y - y_{02}}{n_2} = \frac{z - z_{02}}{p_2} = t_2 \tag{4.7}$$

where (x_{01}, y_{01}, z_{01}) are the coordinate of point A_l, which is $(x_l, y_l, 7)$. (m_1, n_1, p_1) represents vector $\overrightarrow{S_lA_l}$, which is $(x_l, y_l, 60 + 7)$. (x_{02}, y_{02}, z_{02}) are the coordinates of point A_f, which is $(7, y_f, z_f)$. (m_2, n_2, p_2) represents vector $\overrightarrow{S_2A_f}$, which is $(60 + 7, y_f, z_f)$. The intersection point A can then be calculated by solving the parametric Equations 4.6 and 4.7 with parameters t_1 and t_2. From Equations 4.6 and 4.7, we have

$$\frac{x - x_{01}}{m_1} = t_1 \tag{4.8}$$

$$\frac{x - x_{02}}{m_2} = t_2 \tag{4.9}$$

$$\frac{z - z_{01}}{p_1} = t_1 \tag{4.10}$$

$$\frac{z - z_{02}}{p_2} = t_2 \tag{4.11}$$

Since we have four equations with four unknowns, t_1, t_2, x and z can be computed. y can then be calculated by Equation 4.12 or 4.13:

$$\frac{y - y_{01}}{n_1} = t_1 \tag{4.12}$$

$$\frac{y - y_{02}}{n_2} = t_2 \tag{4.13}$$

For the first method, there are actually four equations but three unknowns. Following the procedure mentioned above, there are two equations (Equation 4.3 and 4.4) that can be used to calculate y. When the corresponding 2-D landmark positions on the lateral and posteroanterior view cephalograms are not at their exact positions, the values of y obtained from these two equations can be slightly different. The same problem occurs for the second method if S_1A_l and S_2A_f do not exactly intersect with each other. The values of y obtained from Equations 4.12 and 4.13 are not exactly the same. We handle this problem in both methods by setting y as the average of the two values to minimize the error [6]. Figure 4-5 shows this strategy graphically.

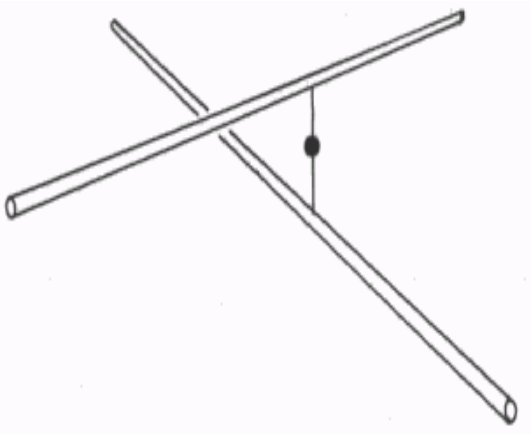


Figure 4-5. Error correction for correcting magnification distortion [6].

4.4 Correction of Rotational Distortion

The 2-D coordinate systems in the lateral and posteroanterior view cephalograms are naturally defined by setting the *x*-, *y*- and *z*-axis directions to coincide with the sides of the cephalograms. Figure 4-6 shows how they are defined. The position of the 2-D origin O, which is the projection of the 3-D origin onto the 2-D cephalograms, is manually input by the user. The 3-D landmark coordinates obtained from the magnification correction procedure is actually derived in a 3-D coordinate system, which has the same origin as the standard 3-D coordinate system and follows the *x*-, *y*- and *z*-axis directions in the 2-D cephalograms. We call this 3-D coordinate system the natural 3-D coordinate system.

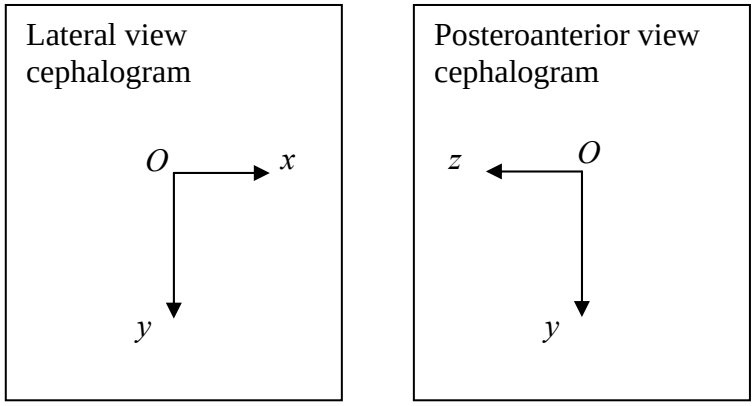


Figure 4-6. 2-D coordinate systems.

The standard 3-D coordinate system is embedded in the patient's head. The *x*-, *y*- and *z*-axis rotate together with the rotation of the patient's head. The 3-D landmarks in this system have the common basis (the standard plane defined by *PoC, OrR* and *OrL*) and landmarks of different patients (or same patient before and after treatment) can compare with each other or with the diagnostic and treatment 'norms' [45] easily. Ideally the *x*-, *y*- and *z*-axis of the standard 3-D coordinate system should coincide with the axes in the natural 3-D coordinate system. Rotational distortion occurs when they do not. Rotational distortion can exist in all the three directions (around the three axes of the coordinate system) if the patient's head rotate from its desired position during acquisition of the cephalograms. Our 3-D position-computation method is based on the assumption that the only rotational distortion we need to correct is the one around the ear rod (*z*-axis). Rotation about any other axis will cause pain in the patient's ear and is likely to be insignificant. Figure 4-7 shows how we define the rotation angle around the *z*-axis.

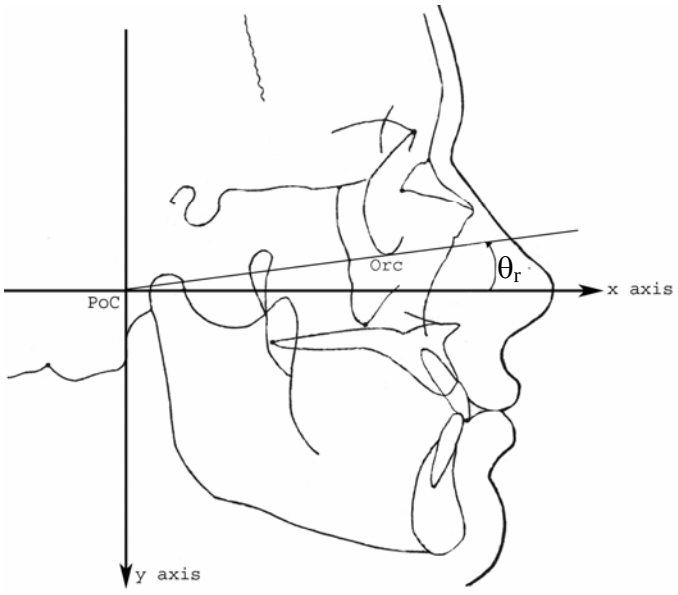


Figure 4-7. Correction of rotational distortion [10].

The rotation angle around the ear rod ($z-$axis) is defined as the angle between the $x-$axis of the cephalogram and the straight line passing through *PoC* and *OrC*, which is the $x-$ axis of the standard 3-D coordinate system. The 3-D coordinate of *OrC* under the natural 3-D coordinate system is (x_{ol}, y_{ol}, z_{ol}). The rotational angle around z-axis, θ_r (clockwise direction as positive), is

$$\theta_r = \tan^{-1}(y_{ol} / x_{ol}) \tag{4.14}$$

For a landmark A with 3-D coordinates (x_a, y_a, z_a) under the natural 3-D coordinate system, its 3-D coordinates (x, y, z) under the standard 3-D coordinate system can be calculated by:

$$x_a^2 + y_a^2 = x^2 + y^2 \tag{4.15}$$

$$\tan^{-1}(y_b / x_a) - \tan^{-1}(y / x) = \theta_r \tag{4.16}$$

$$z = z_a \tag{4.17}$$

From Equations 4.15 and 4.16, the two unknown quantities, x and y can be calculated. The value of z is z_a. With $\theta_a = \tan^{-1}(y_b / x_a)$ (from Equation 4.16) and $A \stackrel{\Delta}{=} y / x = \tan(\tan^{-1}(y_b / x_a) - \theta_r)$ (from Equation 4.16). The solution can be presented as

$$|x| = \sqrt{(x_a^2 + y_a^2)/(1 + A^2)} \tag{4.18}$$

$$y = Ax \tag{4.19}$$

$$z = z_a \tag{4.20}$$

The sign of x can be determined as

$$\text{sgn}(x) = \begin{cases} -\text{sgn}(x_a) & \pi/2 < \theta_a < \pi/2 + \theta_r \\ -\text{sgn}(x_a) & 3\pi/2 < \theta_a < 3\pi/2 + \theta_r \\ \text{sgn}(x_a) & \textit{otherwise} \end{cases} \tag{4.21}$$

4.5 Construction Procedure

There are four steps in the procedure for determining 3-D landmark positions in this study. They are “input origin”, “unit conversion”, “correction of magnification” and “correction of rotational distortion”. Most of the current methods for computing 3-D landmark position only correct for the magnification of the cephalograms but do not take into account the rotational distortion that occurs during the acquisition of x-ray images. The 3-D landmarks in the standard coordinate system (without rotational distortion) have a common basis, which is the standard plane define by *PoC, OrR* and *OrL*. Landmarks of different patients, or the same patient before and after treatment, can be compared with each other or with the diagnostic and treatment ‘norms’ [45] easily. In order to develop an accurate and meaningful 3-D cephalometric system, both

magnification and rotational distortion of the image are corrected in our system. Figure 4-8 summarizes the entire 3-D landmark position computation procedure.

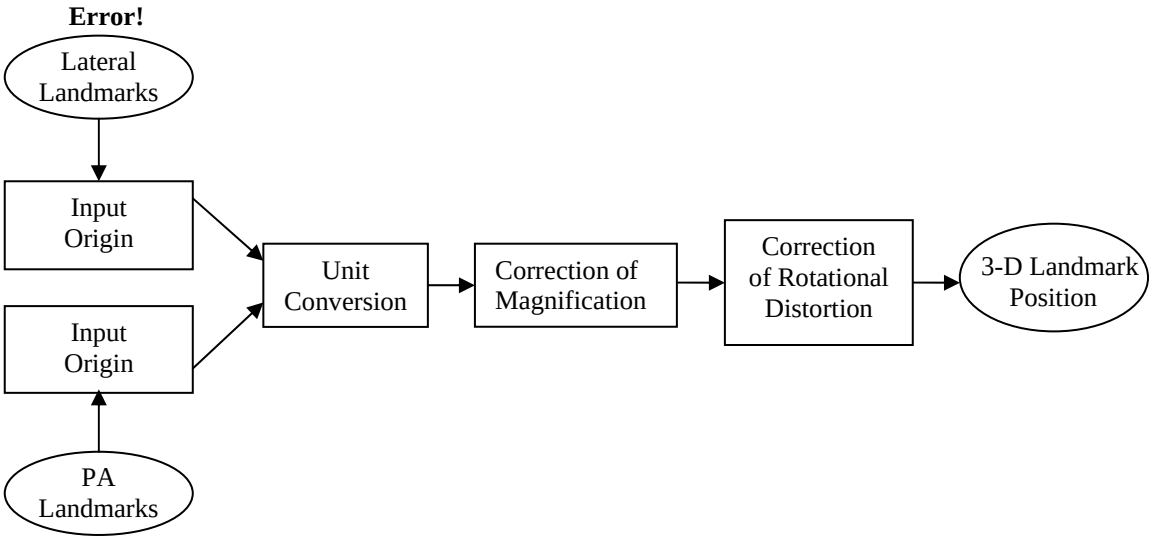


Figure 4-8. Overview of 3-D landmark position construction procedure.

“Input origin” and “unit conversion” are the preparation procedures for correction of magnification and rotational distortion. The coordinates of the 2-D landmark positions detected from the lateral and posteroanterior view cephalograms are the pixel position values offset from the upper left corner of the x-ray images. The “input origin” step corrects the origin of the images from the upper left corner to the center of the ear rod on the lateral view cephalogram and the middle point of the two ear rods on the posteroanterior view cephalogram. The origins are manually input by the user but the coordinates are still pixel position values. “Unit conversion” is the step to convert the coordinate of the landmark from the pixel position value to its corresponding real-time value. In order to do it, the resolution of the input x-ray images must be known. In this study, the resolution is 28.346 pixels/cm or 72 pixels/inch.

For a landmark A with 2-D coordinate of pixel position values (a,b) in the lateral view cephalogram and (c,d) in the posteroanterior view cephalogram, the origins input by the user are (a_0,b_0) and (c_0,d_0) respectively. The corresponding real-time position values are $((a-a_0)/72,(b-b_0)/72)$ and $((c-c_0)/72,(d-d_0)/72)$ (in inches). The projection of this landmark on the lateral and posteroanterior view cephalograms in 3-D natural coordinate system are $A_l(x_l,y_l,7)$ and $A_f(7,y_f,z_f)$ with $x_l=(a-a_0)/72$, $y_l=(b-b_0)/72$, $y_f=(d-d_0)/72$ and $z_f=(c-c_0)/72$. Then the correction of magnification and rotational distortion can proceed using the methods that have been shown in Section 4.3 and 4.4 to make the 3-D landmark position available for further analysis by the orthodontist.

Chapter 5

PROTOTYPE CEPHALOMETRIC SYSTEM

This chapter presents the main features namely, 2-D landmark detection and 3-D landmark position computation, of our computer-aided prototype cephalometric system.

5.1 System Overview

We have developed a prototype system for computer-aided cephalometric landmark identification which implements our 2-D landmark detection and 3-D landmark position computation algorithms. The computing hardware is a PC with a Pentium III 933 MHz CPU and 256 MB SDRAM. The software is developed using Microsoft VC++ 6.0 and OpenGL. The 2-D landmarks on cephalograms can be detected automatically or input manually by the user. The 2-D automatic landmark detection function is realized by training the neural networks so that it could be combined with other techniques to form the landmark-searching algorithm. The 3-D landmark position computation function computes the 3-D landmark position based on the

corresponding 2-D landmark positions (obtained from the 2-D landmark detection function of the system) on the lateral and posteroanterior view cephalograms. The technical details about these algorithms have been presented in Chapters 3 and 4. After the 3-D landmark positions have been computed, the geometry of the skull can then be presented as a wire frame model, which is generated by connecting the 3-D cephalometric landmarks following certain rules [6], for further analysis by orthodontists. The framework of this cephalometric system is given in Figure 5-1.

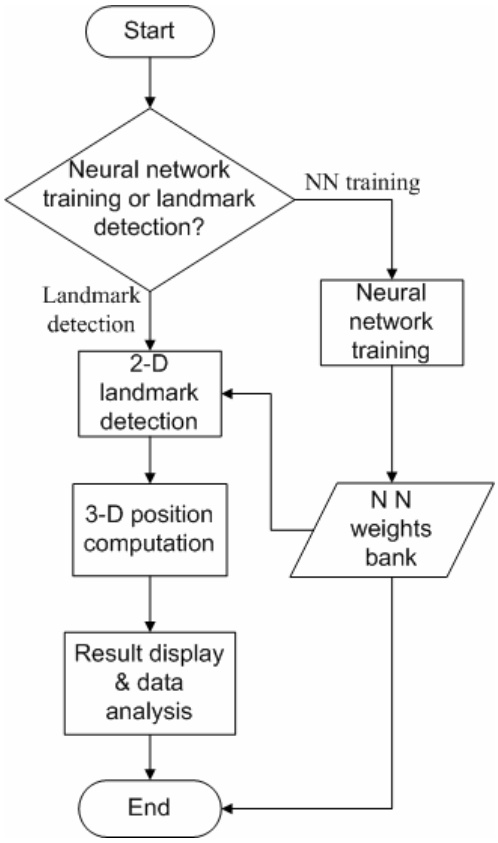


Figure 5-1. Overview of the computer-aided cephalometric prototype system.

5.2 System Functions

The computer-aided cephalometric system is actually a collection of graphical interface programs and image processing programs. The combination of these programs realizes the implementation of our cephalometric landmarking functions. As a multiple-document VC++ application, the main window of the system is created from the standard MFC class CSplitterWnd, which provides the functionality of a splitter window that contains multiple panes. Different projections of cephalograms and detected landmark position information can be viewed simultaneously to help orthodontists make the correct surgical decisions. The main functions of this system are presented in this section.

5.2.1 Neural network training function

Figure 5-2 shows the GUI (graphical user interface) of our system for neural network training. The visualization window displays the 2-D cephalogram and the interface enables users to conduct the operations related to neural network training. It allows the user to click on the cephalogram to input the accurate landmark positions and save them inside the system as a binary file. The user can view or change the 2-D landmark information stored in the system at any time. The radio buttons on the left hand side help the user specify the landmark. The neural network can be trained by the landmark information stored in the system or input instantly from the interface or the combination of both. The landmark information input here can also be used to justify the performance of the system by comparing it with the system-detected landmark information.

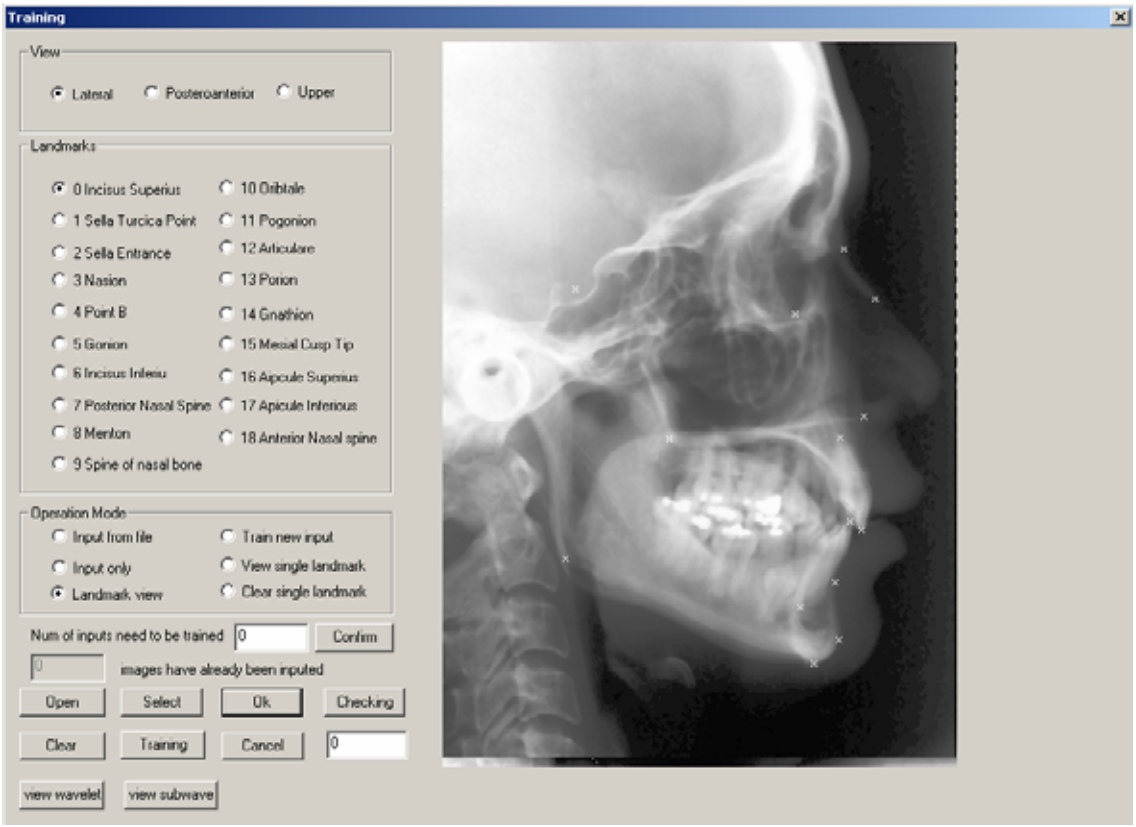


Figure 5-2. Neural network training function.

The landmarks input to or output from the system are displayed by "×" on the cephalogram for viewing. The design of this GUI makes the system more flexible for further modification. New landmark information can easily be added to the system. Existing landmark information can also be checked and modified anytime.

5.2.2 Parameter setting

Other techniques like GA and wavelet decomposition used for landmark searching do not require much user interaction. But there are some very important parameters that need to be set for these algorithms to perform correctly. Furthermore the input and output files have to be managed with care. "Parameter setting" dialog is the interface

for the user to manage these parameters. Figure 5-3 shows this parameter-setting interface.

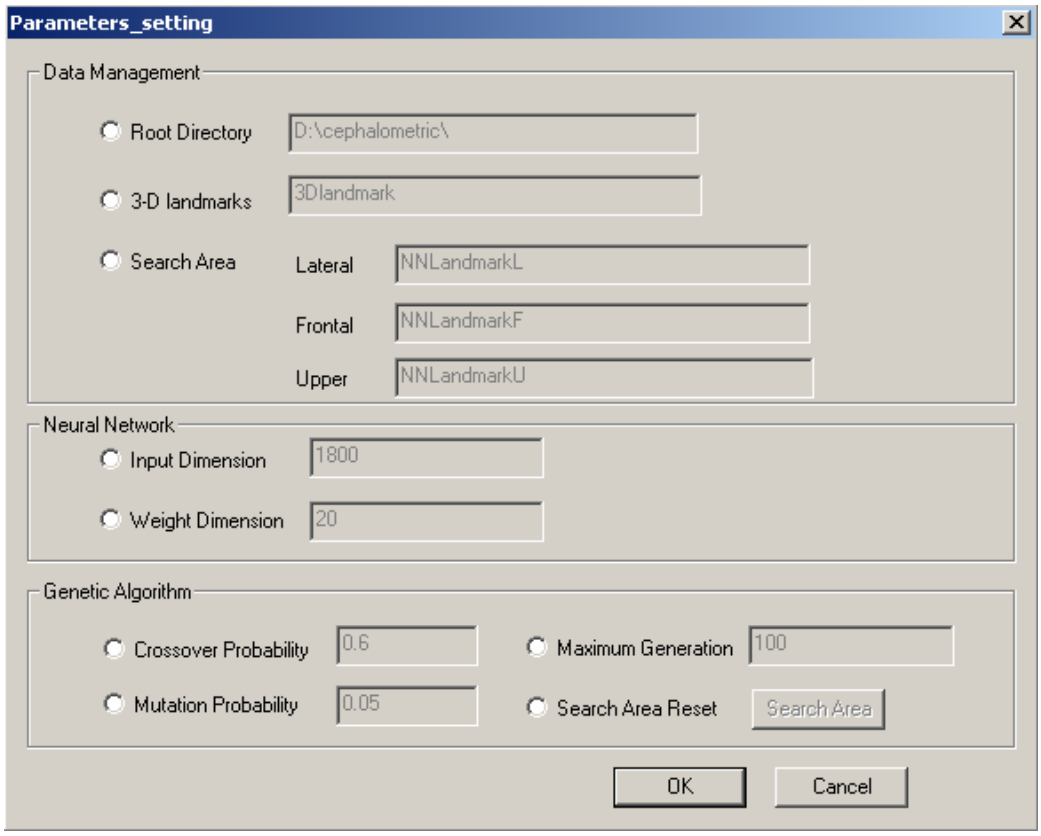


Figure 5-3. Parameter-setting interface.

Under the “Data Management” category, “Root Directory” indicates the directory where all the input and output files are located. “3-D landmarks” is the file stores 3-D landmark information. Three files under “Search area” store the search area information (location and area) for each landmark in the three views of cephalograms.

Under the “Neural Network” and “Genetic Algorithm” categories, the input and weight dimension of the neural network, crossover and mutation probability of GA and maximum generation of the genetic search are listed.

To change these parameters, the corresponding radio button on the left has to be selected. It will enable the edit box on the right for updating. The default values of these parameters are displayed in these edit boxes initially.

“Search area reset” enables the user to redefine the search area for the landmarks. The search area is evaluated statistically based on the landmark position variation between the training images. This action is normally conducted when there are new training images stored in the system.

5.2.3 3-D landmark position determination

Figure 5-4 shows the interface created from the standard MFC class CSplitterWnd, which is used for 3-D landmark position computation. Three sub-windows are engaged to display three cephalogram projections. We mainly make use of the lateral and posteroanterior views with one sub-window reserved for the basal view cephalogram. With 2-D landmark position information from a third view, the detection accuracy may be improved or more landmarks that cannot be seen clearly in posteroanterior view may be detected. Detected landmarks will be marked on them whenever the 2-D landmark detection function is completed. The fourth sub-window displays the results of 3-D landmark position computation. The results can be displayed as text information or as a wire frame (Figure 5-4).

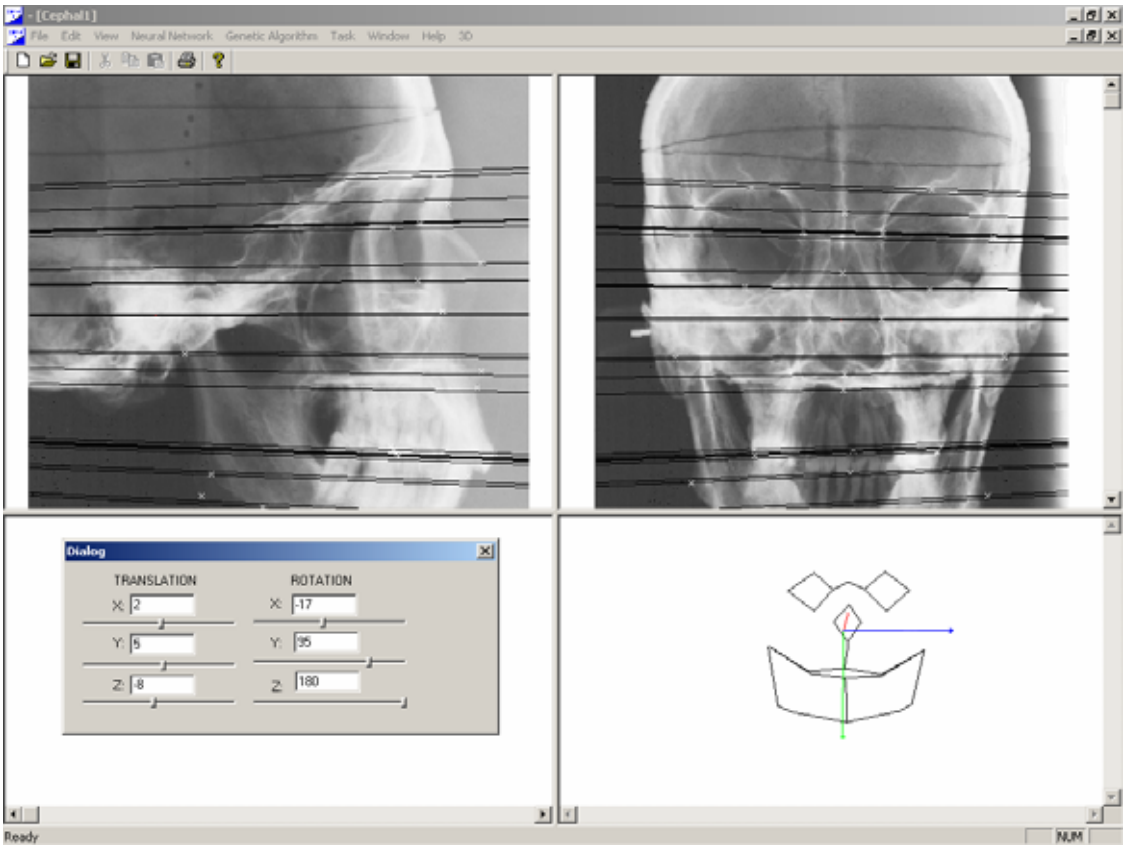


Figure 5-4. Computation of 3-D landmark positions.

5.2.3.1 Landmark line generation

The lines on the cephalograms in Figure 5-4 are landmark lines that have been drawn during the process of 3-D landmark position computation. Figure 5-5 is a clearer picture of the landmark lines with fewer landmarks on the cephalograms and Figure 5-6 explains how these landmark lines have been generated.

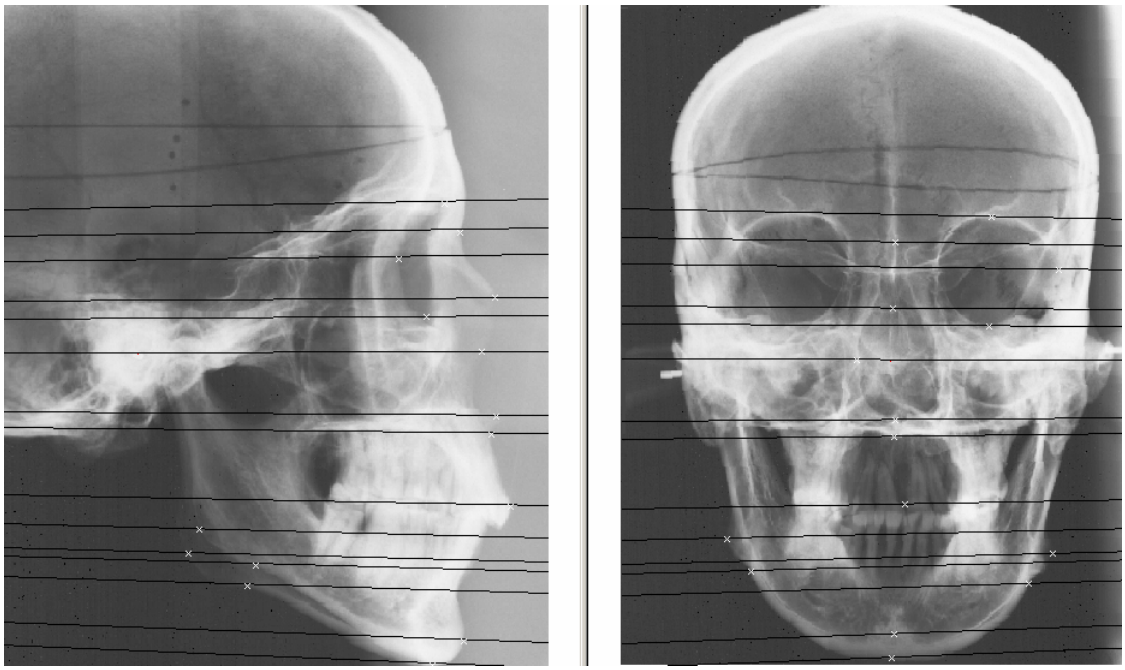


Figure 5-5. Landmark lines of cephalograms.

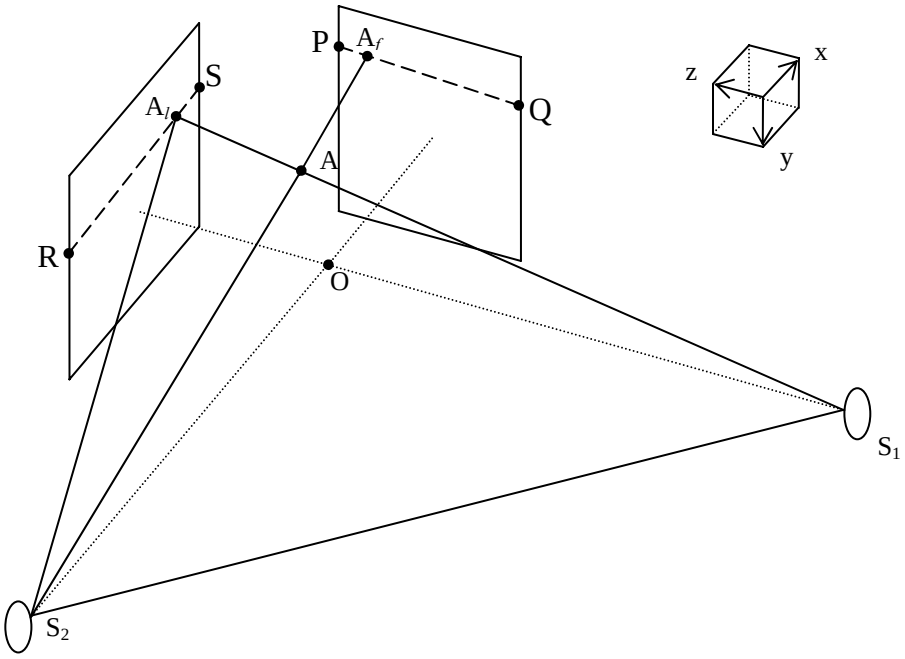


Figure 5-6. Generation of landmark lines.

As explained in Section 4.2, the rays connecting the x-ray source to each landmark of the lateral and posteroanterior view films can be drawn as lines in space. Figure 5-6

represents the 3-D geometric configuration of Figure 4-3. For a 3-D landmark A with 2-D landmarks on lateral and posteroanterior view cephalograms A_l and A_f, respectively, A_f should ideally lie on the projection of S_1A_l on the posteroanterior view cephalogram. A_l should lie on the projection of S_2A_f on the lateral view cephalogram. This indicates that each landmark in one view can be restricted to lie on one particular line in the other view. We draw these lines on the cephalograms for all the landmarks in our system as in Figure 5-4. If a landmark in one cephalogram is seen far from its corresponding line, we know that the detection is not very accurate. This can serve as an indicator of detection errors.

The projection of S_1A_l on the posteroanterior cephalogram (line PQ in the figure) can be treated as the line where plane $S_1S_2A_l$ cuts the posteroanterior view cephalogram. The equation of plane $S_1S_2A_l$ is

$$x + ay + bz + c = 0 \tag{5.1}$$

where S_1, S_2 and A_l (whose coordinates are known) are points on it. The parameters a, b and c in Equation 5.1 can be solved by substituting S_1, S_2 and A_l in it. If we set the x value in Equation 5.1 to 7, we obtain the equation of line PQ. It can then be drawn on the posteroanterior view cephalogram. The equation of the projection of S_2A_f on the lateral view cephalogram (line RS in the figure) can be obtained in the same way.

5.2.3.2 Wire frame generation

OpenGL has been used widely as a graphic interface to visualize 3-D images. The wire frame is generated by OpenGL. When the results are displayed as a wire frame, the interface for transformation (the dialog on the third sub-window of Figure 5-4) allows the user to move and rotate the wire frame. It allows users to transform the wire frame in six degrees of freedom and this includes translation in the *x*, *y*, and *z* directions and rotation about the *x*, *y* and *z* axes. This enables users to view the wire frame from any viewpoint. The standard 3-D coordinate system we presented in Chapter 4 is used here. The wire frame is generated by connecting the landmarks in 3-D that computed by our system to represent the geometry of the skull for further analysis of orthodontists. One application is to compare the wire frame of the patient with the diagnostic and treatment 'norms' such as three-dimensional Bolton standards. As on of the steps in surgery simulation, this can help the orthodontist make the correct surgical decisions [6]. Figure 5-7 shows samples of the wire frame drawn together with the x, y and z axes.

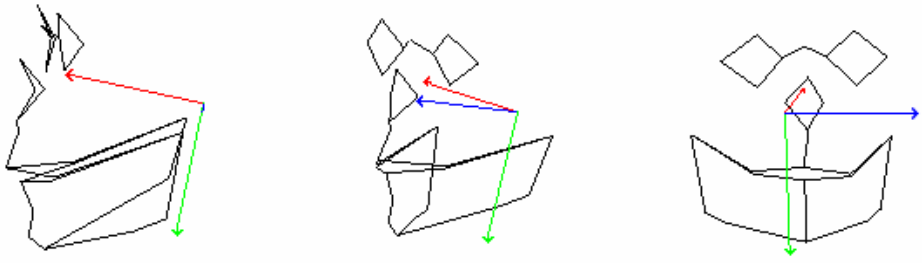


Figure 5-7. Wire frame.

Table 5-1 lists the landmarks that have been involved in 3-D landmark position computation in our study [6]. These landmarks are also the ones that used to generate the wire frame. The rule about how these landmarks are connected to generate the wire frame is shown in Figure 5-8.

Table 5-1. Landmarks involved in 3-D landmark position computation.

No	Name	Definition	No	Name	Definition
1	SoR	Supra Orital Right	16	PiR	Piriform Right
2	LoR	Lateral Orbit Right	17	PiL	Piriform Left
3	OrR	Obitale Right	18	Sdl	Supradentale
4	MoR	Medial Orbit Right	19	Idl	Infradentale
5	SoL	Supra Orbital Left	20	UmR	Upper Right 1st Molar mesiobuccal cusp
6	LoL	Lateral Orbital Left	21	UmL	Upper Left 1st Molar mesiobuccal cusp
7	OrL	Orbitale Left	22	LmR	Lower Right 1st Molar mesiobuccal cusp
8	MoL	Medial Orbit Left	23	LmL	Lower Left 1st Molar mesiobuccal cusp
9	N	Nasion	24	GoR	Gonion Right
10	A	A Point	25	GoL	Gonion Left
11	B	B Point	26	AgR	Antegonion Right
12	Pog	Pogonion	27	AgL	Antegonion Left
13	Me	Menton	28	MaR	Mastoidale Right
14	ANS	Anterior Nasal Spine	29	MaL	Mastoidale Left
15	Rhn	Rhinion			

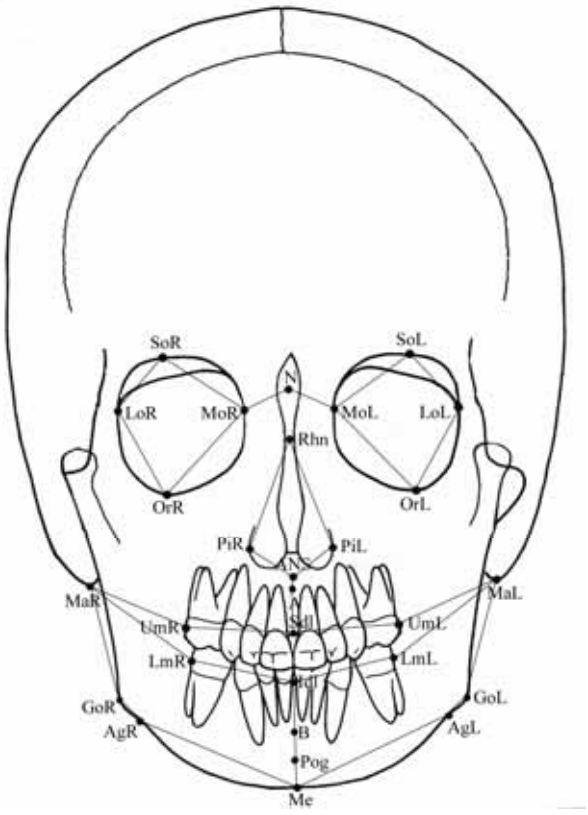


Figure 5-8. Wire frame generation rule [10].

Chapter 6

RESULTS AND DISCUSSION

Experimental results of our study are presented in this chapter. The results are analysed and compared with those obtained by other automatic landmarking methods.

6.1 Results of 2-D Landmark Detection

6.1.1 Selected landmarks

Table 6-1 lists the landmarks detected by the 2-D automatic landmarking function of our system. These are commonly used landmarks in clinical orthodontics. The actual selection of landmarks used by each orthodontist varies with personal preference and experience [23]. Our selected landmarks are of varying types and provide an adequate test base for assessing our method. Figure 6-1 shows the positions of these landmarks graphically.

Table 6-1. Selected landmarks.

No.	Name	Definition
1	Incisor superious (IS)	The tip of the crown of the most anterior maxillary central incisor.
2	Sella (S)	The midpoint of the line posterior clinoid process and the anterior opening of the turcica.
3	Nasion (N)	The most anterior point of the nasofrontal suture in the median plane. The skin nasion (N) is located at the point of maximum convexity between nose and forehead.
4	Point B (B)	The most anterior part of the mandibular base. It is the most posterior point in the outer contour of the mandibular alveolar process, in the median plane.
5	Gonion (Go)	A constructed point, the intersection of the lines tangent to the posterior margin of the ascending ramus and the mandibular base.
6	Incisor inferious (II)	The tip of the crown of the most anterior mandibular central incisor.
7	Posterior nasal spine (PNS)	A constructed radiological point marking the intersection of a continuation of the anterior wall of the pterygopalatine fossa and the floor of the nose. It marks the dorsal limit of the maxilla.
8	Menton (Me)	The lowest point of the mandible.
9	Spine of nasal bone (Snb)	The most anterior point of the nasal bone
10	Orbitale (Or)	The lowermost point of the orbit in the radiograph.

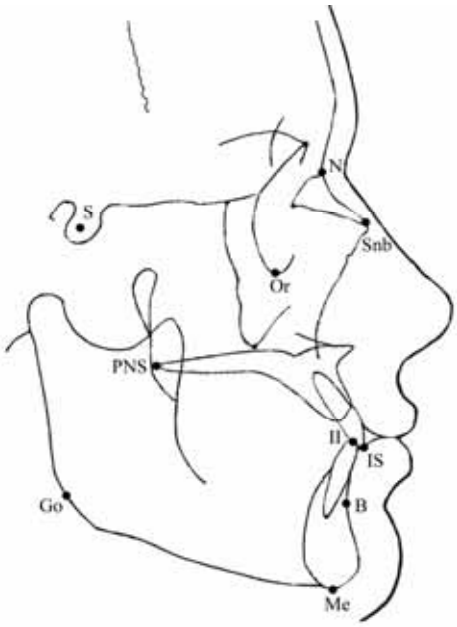


Figure 6-1. Landmarks detected by 2-D automatic landmarking [10].

6.1.2 Testing of landmark detection on geometrical shapes

Automatic cephalometric landmarking is a difficult task since the cephalograms of patients vary significantly from one to another. Moreover, the difficulty of controlling the process of image acquisition can result in deformation, rotation, scaling distortion and variation in brightness and resolution. A good automatic cephalometric landmarking method should be able to handle these distortions.

A test procedure has been developed to verify the performance of the proposed technique for application to a number of images with distorted geometrical shapes. The basic idea of this procedure is to train a neural network with an image that contains undistorted geometrical shapes only. The presumed landmark is attached to one of these shapes. The distorted images are generated from the original image by deforming or transforming the geometrical shape to see whether the system can detect the landmark on these distorted images.

Figure 6-2 shows one set of the geometrical shapes that we have used for testing. Figure 6-2a is the original image. The top vertex of the triangle is the presumed landmark for detection. The triangles in Figures 6-2b to 6-2g have been deformed or transformed in several different ways to simulate the distortion of x-ray images in the application of cephalogram analysis.

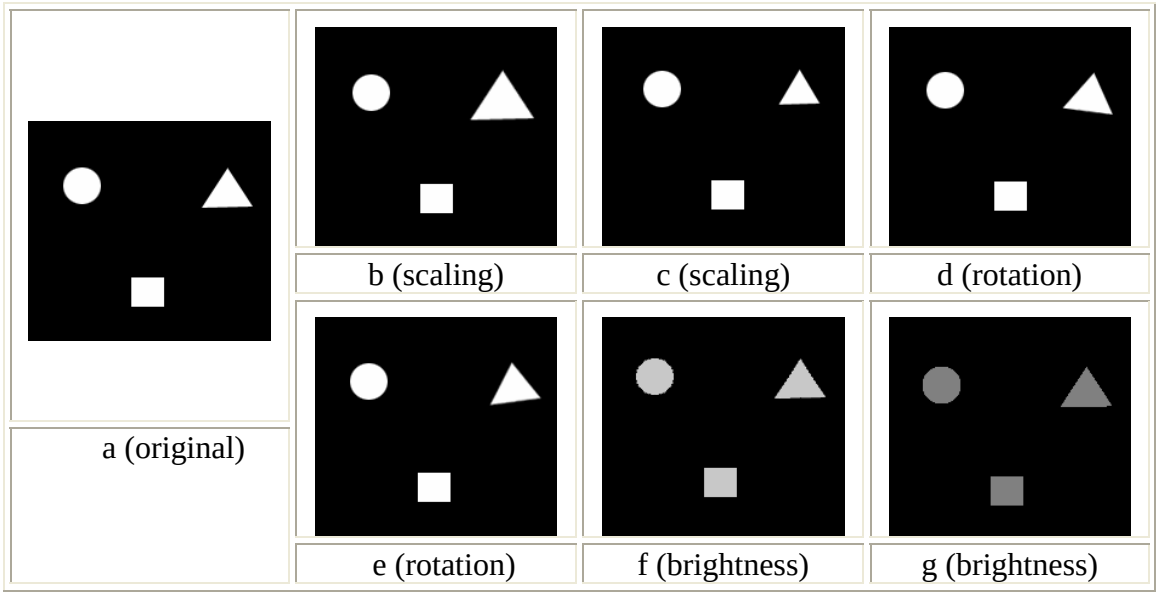


Figure 6-2. Test images: Set A, with regular geometrical shapes.

Since the geometrical shapes we used here are much simpler compared with real cephalometric landmark patterns, the hidden layer size of the MLP used is 10×10 only. During this experiment, the MLP was trained by Figure 6-2a only and was then used to detect the presumed landmark in Figures 6-2b to 6-2g. The details of the training procedure have been presented in Chapter 3.

The same experiment has been repeated for Figures 6-3 and 6-4. Figure 6-3 contains the same geometrical shapes with Figure 6-2 but in a lower resolution. We use the MLP trained for Figure 6-2 directly for Figure 6-3. Figure 6-4 contains irregular geometrical shapes. The lowest point of the "U" shape is the presumed landmark position. The details of scaling, rotation, variation in brightness and resolution in test image sets A, B and C are present in Table 6-2. The testing results are summarized in Table 6-3.

Table 6-2. Details of geometrical shape variation.

Test image set	Resolution (pixels/inch)	Scaling (%)	Rotation (degree)	Brightness (gray level)
Set A	72	±20	±6	-55/-127
Set B	36	±20	±6	-55/-127
Set C	72	±20	±6	-55/-127

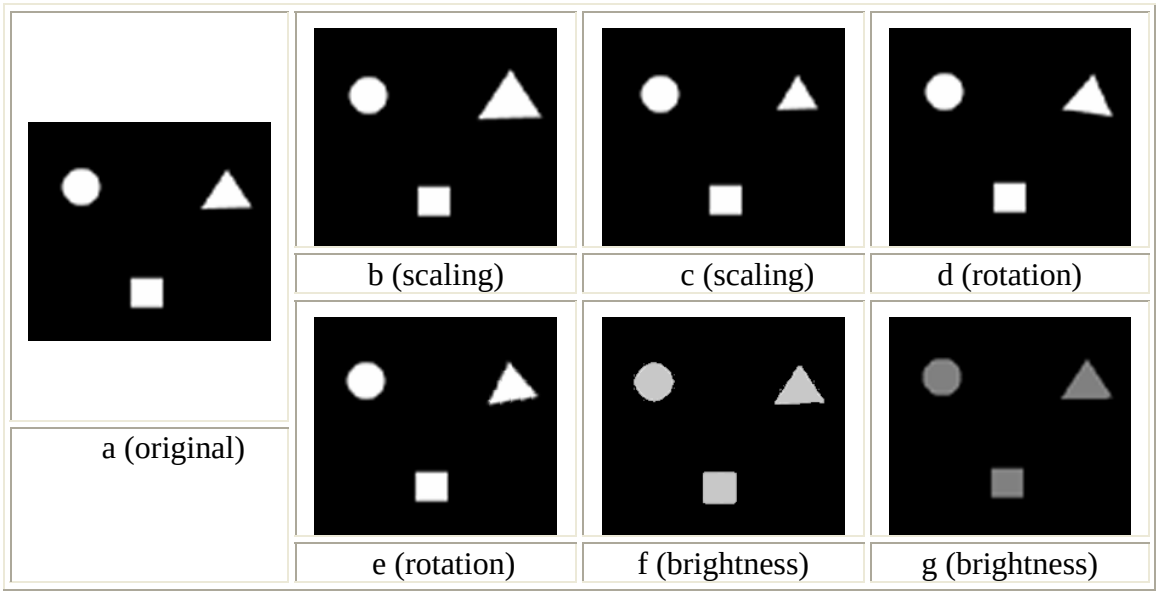


Figure 6-3. Test images: Set B, with regular geometrical shapes in lower resolution.

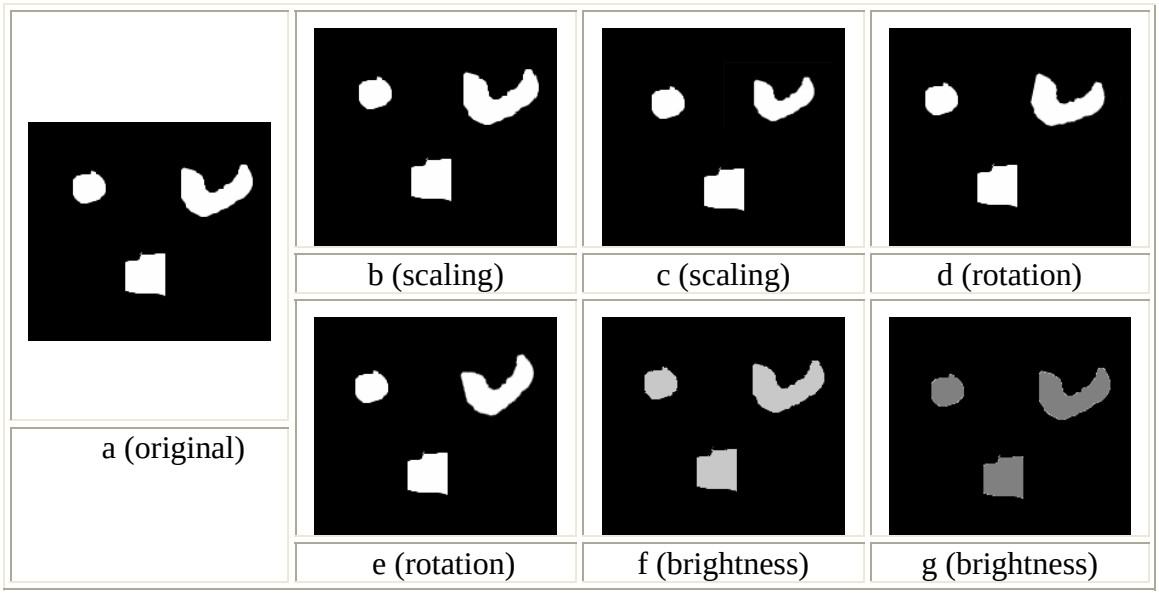


Figure 6-4. Test images: Set C, with irregular geometrical shapes.

Table 6-3. Results of geometrical shape testing.

Test image set	b	c	e	d	g	f
Set A	2	3	1	3	2	1
Set B	1	3	2	2	1	2
Set C	1	2	5	3	1	1

The errors indicated in Table 6-3 are the distances (in pixels) between the detected and corresponding accurate position of the landmarks for images b to f of the three test sets. It can be seen that most of the results are within 3 pixels of the exact landmark positions. Based on the image resolution we used in this study, which is 28.346 pixels/cm, a distance of 3 pixels is approximately 1mm. Most cephalometric measurements are acceptable with accuracies of roughly 2 mm [8]. This experiment indicates that the proposed algorithm is robust enough in dealing with these distortions.

Now we take a closer look at the results. Figure 6-5 shows the neural network response over a 60×60 area around the presumed landmark for the corresponding images in Figure 6-2. Ideally the NN output should show a peak value at the landmark position and nearly zero elsewhere. Figure 6-5a shows the output for the original image. The peak value appears at the center of the pattern since the landmark is there. In Figures 6-5b to 6-5g, the peak values lie near the presumed landmark position but the peak values in Figures 6-5f and 6-5g are rather weak. Although the detection results are good, significant noise has appeared in those two figures. Figures 6-2f and 6-2g contain geometrical shapes with brightness distortion. It implies that the detection algorithm may be sensitive to brightness distortion. Since the intensity distributions of the cephalograms acquired on different machines or at different time are likely to be

different, we add the preprocessing step of histogram equalization to all the training and test images to reduce the brightness difference between them.

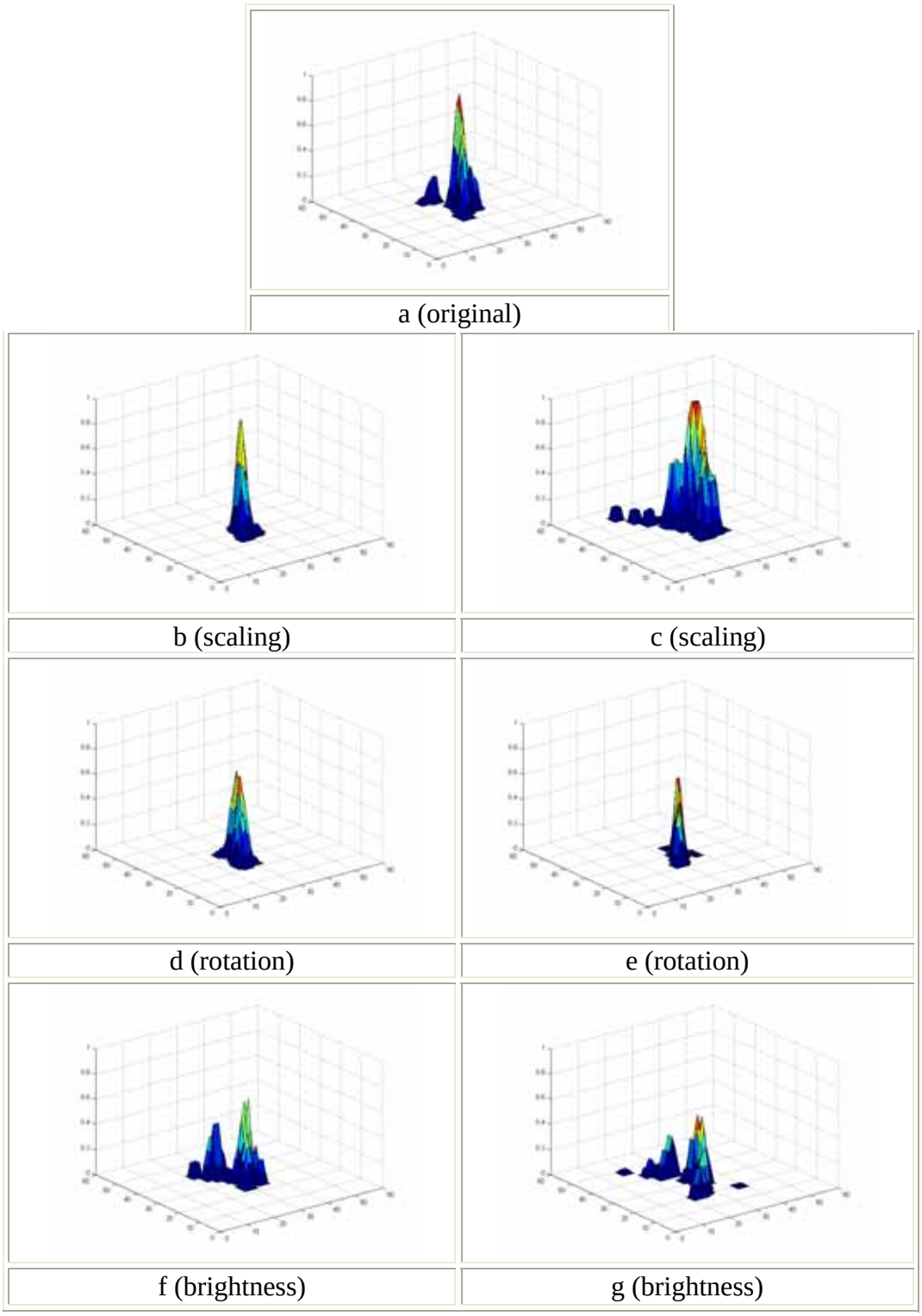


Figure 6-5. Neural network responses of testing geometrical shapes set A.

6.1.3 Results of 2-D automatic landmark detection

The experiment in Section 6.1.3 was carried out on synthetic images. In this section, we shall show the experimental results of our system when dealing with real cephalograms. Ten lateral view cephalograms were randomly selected as the test images to verify the performance of the system for 2-D landmark detection. We employed lateral view cephalograms to do this verification since we only had a limited number of posteroanterior view cephalograms taken from dry skulls and this would not be enough for neural network training. Also, with lateral view detection results we can compare the performance with other research work easily since most automatic cephalometric landmarking research deal with lateral view cephalograms only.

As stated in Section 6.1.3, it is possible for the algorithm to be sensitive to the change of brightness. To overcome this, we apply histogram equalization to all the training and test images as a preprocessing step to reduce the brightness difference between them. Median filtering is also used prior to the histogram equalization as part of the preprocessing step to reduce noise without excessive blurring.

Thirty images have been used to train the neural network for each landmark. The number of training images is relatively small and normally not sufficient for the neural network to recognize the landmark pattern. We increase the number of “wrong” inputs to help the neural network recognize the “correct” pattern and compensate for the lack of training images. 3600 subimages (9 of them are “correct” subimages) from each training image have been selected as the inputs to the neural network. The detection results have confirmed the effectiveness of this method. On average, 40% of the landmarks have been detected within 1mm of their accurate positions and 80% of the

landmarks are detected within 2mm of their accurate positions. The detection results are displayed in the images (with the detected landmarks marked by "×") and Table B-1 in Appendix B. The details of the results are provided in Tables 6-4 and 6-5 and in Figures 6-6 and 6-7.

The lateral view x-rays were digitised to 429×579 pixels with a resolution of 28.346 pixels/cm (scanned at 72 dpi). With this resolution, a position accuracy of about 0.25 mm (the pixel is in the size of 0.35mm×0.35mm with half the diagonal at about 0.25mm) is possible at best. This resolution is more than adequate considering that most cephalometric measurements are acceptable with an accuracy of roughly 2 mm [8].

We now present the detection performance of our system in the format commonly adopted by other researchers [12, 14, 19, 23, 28]. The number of detected landmarks located within 1mm, 2mm and 5mm of their actual positions are tabulated. Table 6-4 summarizes the system performance for all the test images. Table 6-5 summarizes the performance for each test image for all the detected landmarks.

Table 6-4. System performance for each landmark.

Landmarks	1mm	2mm	5mm	Performance Value
Incisor superious (IS)	60%	100%	100%	26
Sella (S)	70%	80%	80%	23
Nasion (N)	50%	90%	100%	24
Point B (B)	20%	70%	80%	17
Gonion (Go)	10%	70%	80%	16
Incisor inferious (II)	60%	100%	100%	26
Posterior nasal spine (Pns)	30%	70%	90%	19
Menton (Me)	60%	100%	100%	26
Spine of nasal bone (Snb)	20%	60%	90%	17
Orbitale (Or)	20%	60%	80%	16

Table 6-5. System performance for each test image.

Test images	1mm	2mm	5mm	Performance value
1	50%	90%	100%	24
2	50%	90%	100%	24
3	50%	100%	100%	25
4	50%	90%	90%	23
5	50%	90%	100%	24
6	30%	60%	70%	16
7	30%	70%	90%	19
8	30%	70%	90%	19
9	30%	70%	80%	18
10	30%	70%	80%	18

The performance value for the landmark and test image in the tables is obtained by the following rules (In general, a performance value higher than 18 can be considered as good performance):

- A detection of the landmark within 1mm of the accurate position scores 3 points.
- A detection of the landmark within 1mm to 2mm of the accurate position scores 2 points.
- A detection of the landmark within 2mm to 5mm of the accurate position scores 1 point.

Figures 6-6 and 6-7 present the performance values in Tables 6-4 and 6-5 graphically.

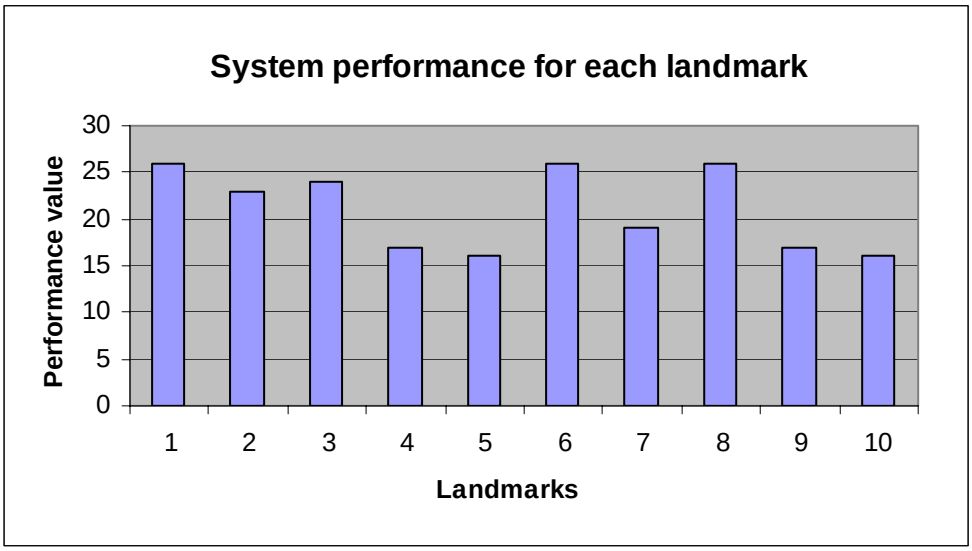


Figure 6-6. Variation of performance between landmarks.

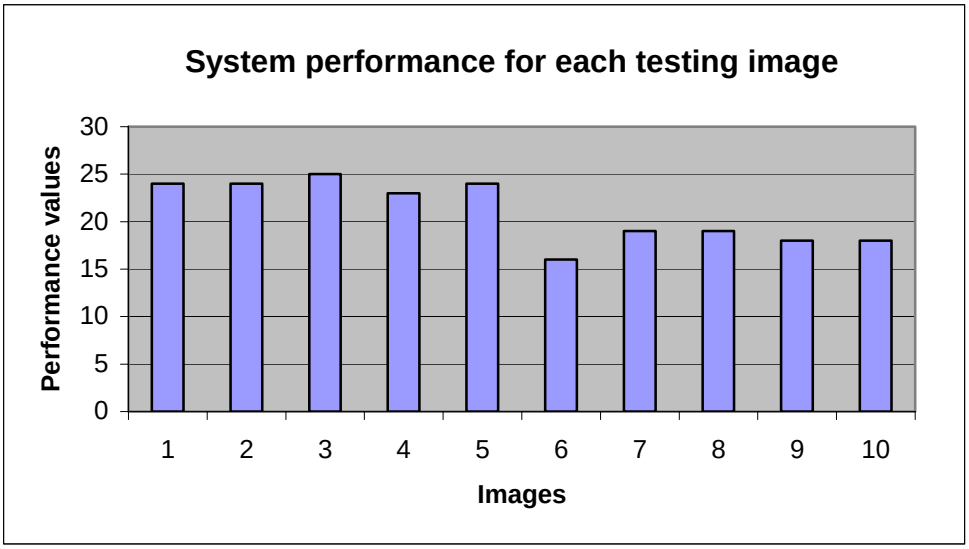


Figure 6-7. Variation of performance between cephalograms.

From Figure 6-6 one may notice that the accuracy of landmark detection varies from one landmark to another. The highest accuracy is obtained for landmarks “menton”, “incisor inferious” and “incisor superious”. The lowest accuracy is obtained for landmarks “orbitale” and “gonion”. There is relatively less performance difference between cephalograms. Further analysis of these results is presented in the following section.

6.1.4 Texture of landmark pattern from wavelet decomposition

Texture is an important element to human vision and can be a very important factor that affects the automatic landmark detection results, too. The texture feature study in this section is to find wether the texture feature of landmark patterns affects the automatic landmarking performance.

The analysis and extraction of texture is a very difficult problem and has not yet been sufficiently addressed. Wavelet subband features have been used to quantify the

texture features of an image [46]. Since we use wavelet decomposition in our study, it is also used here to extract texture features for each landmark pattern. Some simple texture features that can be extracted from the wavelet subbands are “mean”, “standard deviation” and “energy”. The first two have been used by other researchers to study wavelet texture features [47, 48] and are considered to be appropriate for our investigation.

The correct subimages (the ones with the specified landmark in the central area) of each landmark are decomposed (1-level) by the wavelet transform and recombined to give the two low-frequency and high-frequency information subimages I_{Low} and I_{High} (details in Chapter 3). The “mean” and “standard deviation” of each landmark pattern are calculated from these two subimages. The experiment here is done for all the thirty training cephalograms. The formulas for calculating the mean T_{1L}, T_{1H} and standard deviation T_{2L}, T_{2H} of an image (subscripts L for I_{Low} and H for I_{High}) are, respectively,

$$T_{1L} = \frac{1}{h_d \times w_d} \sum_x \sum_y |W_L(x,y)| \tag{6.1}$$

$$T_{1H} = \frac{1}{h_d \times w_d} \sum_x \sum_y |W_H(x,y)| \tag{6.2}$$

$$T_{2L} = \left[\frac{1}{h_d \times w_d - 1} \sum_x \sum_y (|W_L(x,y)| - T_{1L})^2 \right]^{1/2} \tag{6.3}$$

$$T_{2H} = \left[\frac{1}{h_d \times w_d - 1} \sum_x \sum_y (|W_H(x,y)| - T_{1H})^2 \right]^{1/2} \tag{6.4}$$

where $W_L(x,y)$ and $W_H(x,y)$ are the subimages of I_{Low} and I_{High}, and h_d and w_d are their height and width. The wavelet transformation with Haar basis vectors is

used here. We obtain one set of texture data for every landmark from each training image (Table 6-6).

Table 6-6. Texture of landmark pattern from wavelet decomposition.

No.	Landmark	I_{Low}		I_{High}	
		T_{1L}	T_{2L}	T_{1H}	T_{2H}
1	Incisor superious (IS)	120.45	61.57	2.33	3.67
2	Sella (S)	160.84	20.81	1.50	1.72
3	Nasion (N)	110.13	56.16	2.42	2.69
4	Point B (B)	117.67	47.04	1.49	1.75
5	Gonion (Go)	81.26	20.50	1.01	1.34
6	Incisor inferious (II)	156.13	59.35	2.64	3.64
7	Posterior nasal spine (PNS)	147.85	40.25	2.07	2.09
8	Menton (Me)	68.78	47.86	2.26	4.37
9	Spine of nasal bone (Snb)	51.35	17.63	1.13	2.29
10	Orbitale (Or)	143.25	30.82	2.14	1.97

The mean and standard deviation values in Table 6-6 are the average values that we obtained from all the training images. As we can see, T_{1L}, T_{2L} and T_{1H}, T_{2H} vary from one landmark to another. Figures 6-8 and 6-9 present the data in Table 6-6 graphically. Further analysis of these data is presented in Section 6.1.5.

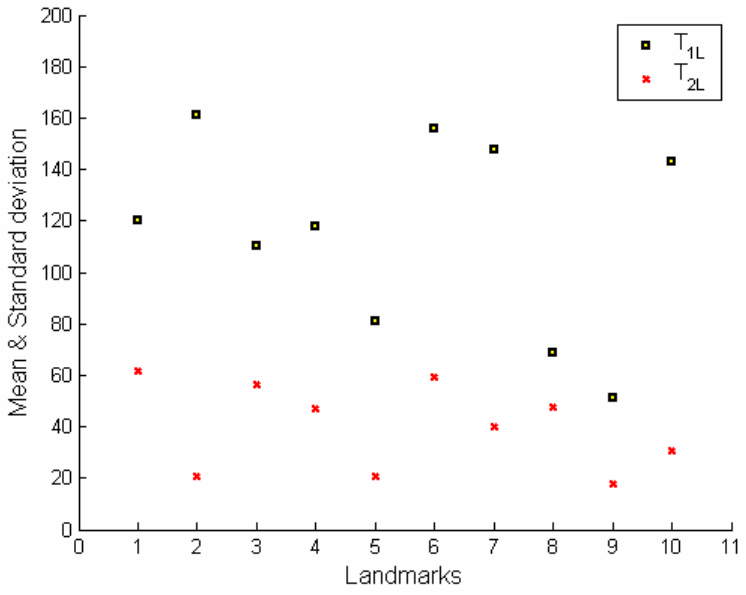


Figure 6-8. T_{1L} and T_{2L} in I_{Low}.

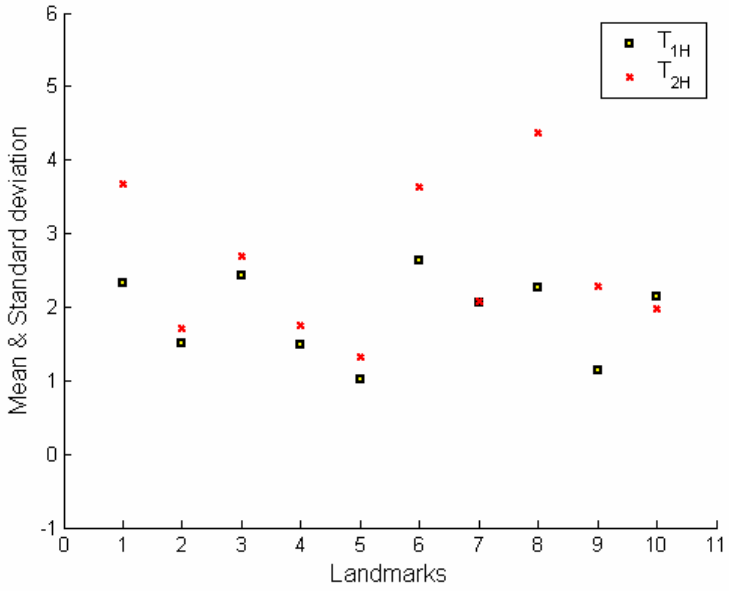


Figure 6-9. T_{1H} and T_{2H} in I_{High}.

6.1.5 Texture of landmark patterns and automatic landmarking performances

As shown in Figure 6-6, the accuracy of automatic landmarking varies from one landmark to another. One important reason for this variation can be the structural property differences between the landmark patterns. One way to study the structural property of the landmark patterns is to measure their texture features. In this section, we investigate how detection performance varies with the texture feature of the landmark pattern. Figures 6-10 and 6-11 presents the texture feature of landmark patterns (shown in Table 6-6) together with the corresponding landmark detection performances (shown in Figure 6-6) for each landmark. Since the texture feature and performances data are in different units, we compute their corresponding percentage values to the average value of each data to shown them in the same figure. The percentage values for the texture feature and performances data are computed as

$$P = 100 + 100(D - AverageD) / AverageD \qquad (6.5)$$

where P is the percentage value of D (texture feature or performances data), $AverageD$ is the average value of this data of all the landmarks. The texture features shown are I_{low} in Figure 6-10 and I_{high} in Figure 6-11.

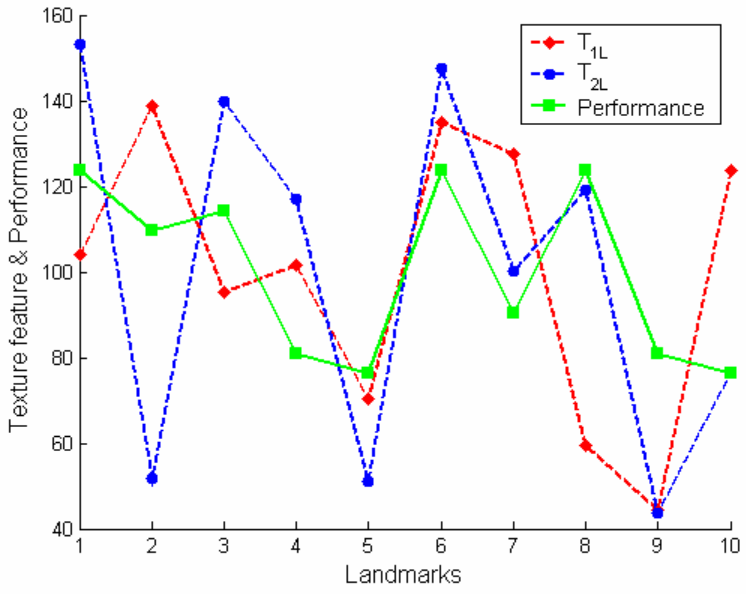


Figure 6-10. I_{low} texture feature and detection performance of landmark.

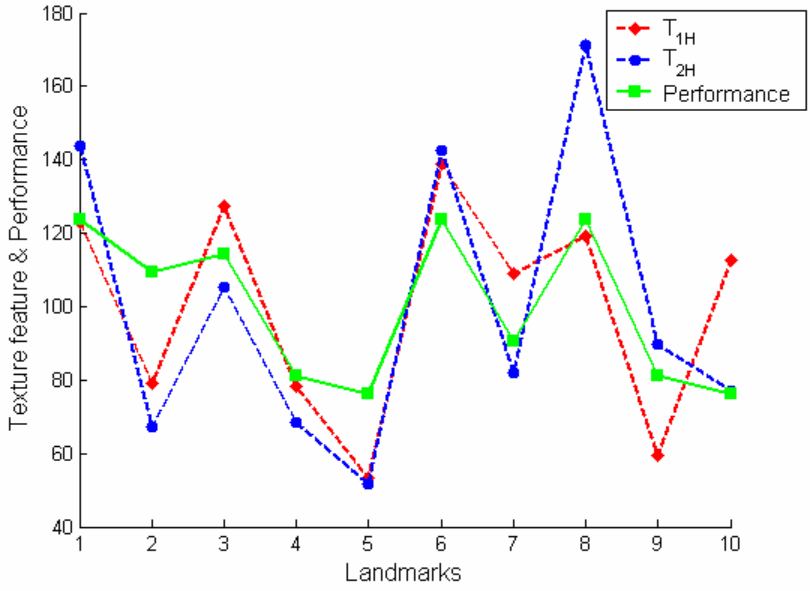


Figure 6-11. I_{high} texture feature and detection performance of landmark.

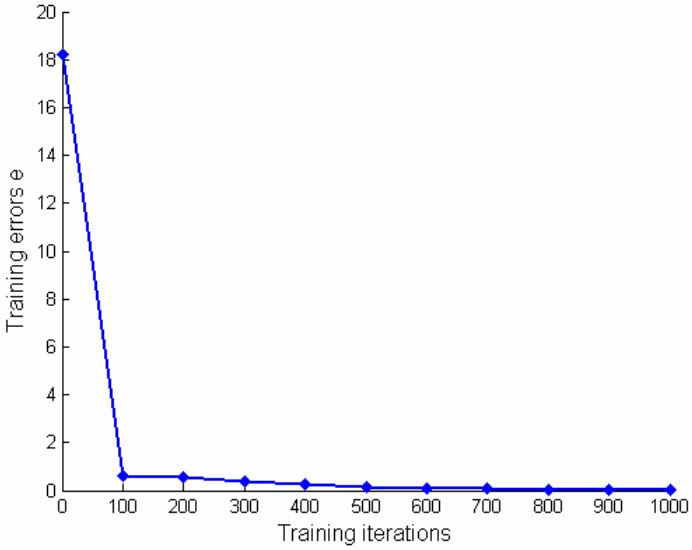
In Figure 6-10, the variation of the texture features “standard deviation” and “ mean” do not really follow the variation of the landmark detection performance very well while they correlate much better in Figure 6-11, which is for I_{high}. This means that the landmark patterns whose I_{high} are richer in texture feature tend to have a better landmark detection performance and vice versa. It also implies that I_{high} plays a more important role in landmark detection compared with I_{low}. This is in fact quite natural since most of the texture features in I_{high} comes from the edge information it contains and for human being we also try to locate the landmark with reference to some edge information. It also indicates it is a wise move to employ the wavelet technique to extract I_{high} to help the neural network learn the landmark pattern.

Our intention here is not to provide an exhaustive analysis of the correlation between landmarking performance and texture feature. The factors that affect automatic landmarking performance can be complicated. Also the texture feature we use cannot fully reflect the structural property of the landmark patterns. For example, in the case of landmark 2, Sella, the computed texture feature value is relatively low compared with its landmarking performance. This is because although the texture value is not high, the landmark itself is clear and hence easy for the neural network to demarcate.

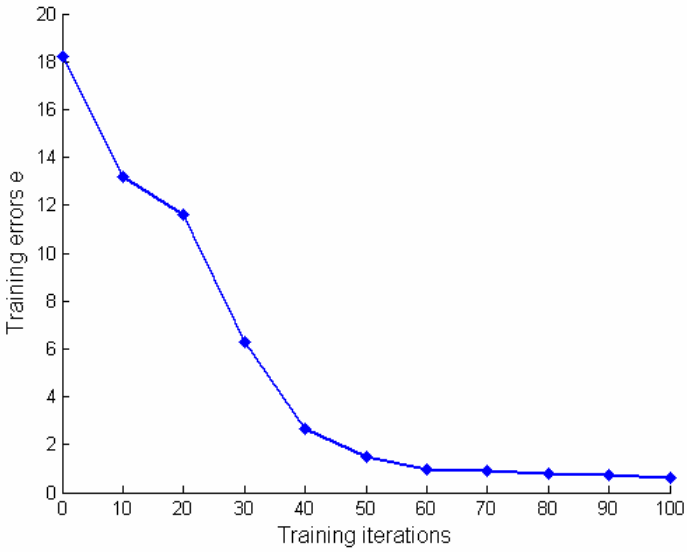
6.1.6 Neural network training speed and texture of landmark patterns

Texture feature of landmark patterns can also affect the neural network training speed. In this section, we study how neural network training speed has varied with the presence of texture feature of landmark patterns. One way to measure the training

speed is to use the rate of training errors ($e_j(n)$ in Equation 3.12), which reduces with training iterations. Figure 6-12 shows how training errors changed during the training procedure for landmark IS. Figure 6-12a shows the changing in 1000 iterations and Figure 6-12b emphasizes the first 100 iterations.



(a)



(b)

Figure 6-12. Training error changes for landmark IS during NN training

Here we measure the neural network training speed as the reduction in error over the first 100 training iterations for each landmark as shown in Table 6-7. From Figure 6-12 we can see that the correlation between the training errors and iterations is nonlinear. The error rate before and after 100 iterations is very different. So the training speed of the first 100 iterations cannot fully represent the NN training speed. Our intention here is just to have a rough analysis of the correlation between NN training speed and texture feature. Figures 6-13 and 6-14 present the training speed together with the texture features of landmark patterns (shown in Table 6-6) for every landmark. Again the data shown here are percentage values. The texture features shown are I_{low} in Figure 6-13 and I_{high} in Figure 6-14.

Table 6-7 Neural network errors reduced (ER) in the first 100 literations

Landmark	Is	S	N	B	Go	II	Pns	Me	Snb	Or
ER	17.34	11.48	13.68	5.58	10.33	13.40	14.21	13.92	7.25	11.60

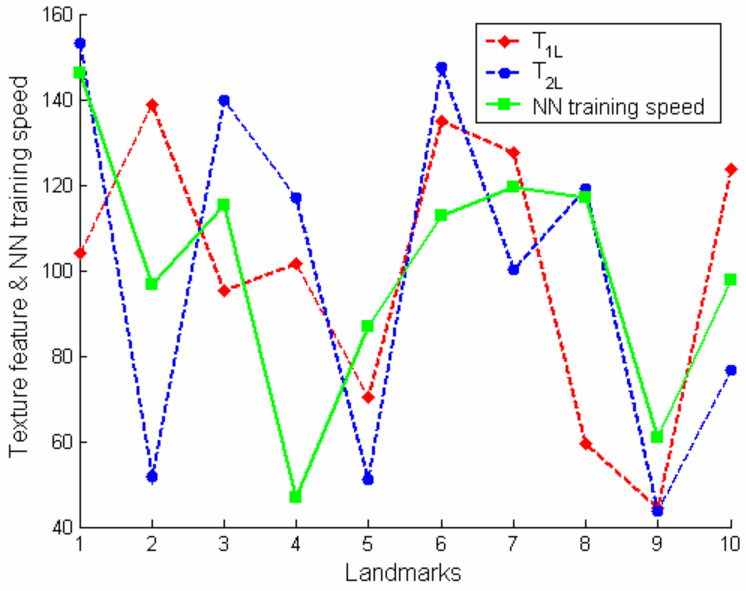


Figure 6-13. NN training speed and I_{low} texture feature.

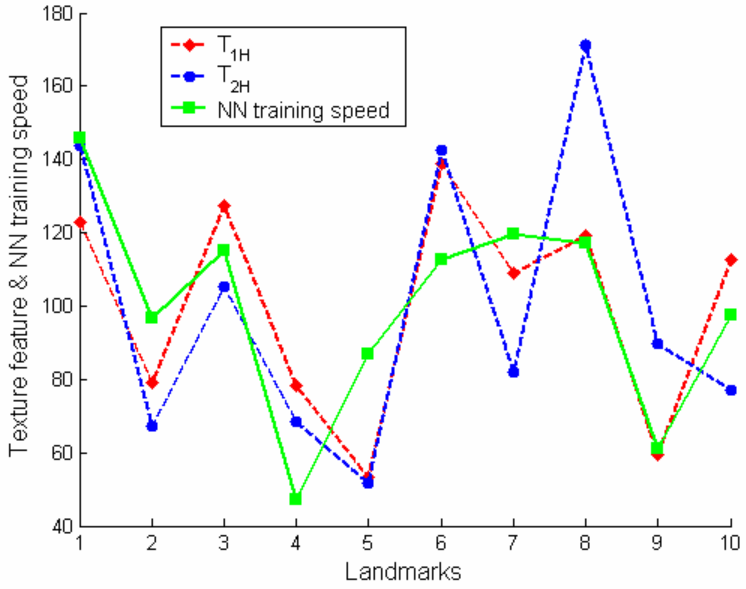


Figure 6-14. NN training speed and I_{high} texture feature.

The phenomenon we observe here is quite similar to what we got from Figures 6-11 and 6-12. The variation of neural network training speed coincides better with the variation of the texture feature in I_{high} than I_{low}. The landmark patterns whose I_{high} are richer in texture feature tend to be trained faster and vice versa. It implies that during training, neural network has given I_{high} a higher priority and also confirms the effectiveness of employing wavelet decomposition in our study.

6.1.7 Comparison with other automatic landmarking methods

Table 6-8 summaries the landmark detection performance of the algorithms listed in Chapter 2. Some authors have not described the details of how they measured their experimental accuracy. Only a rough comparison with these algorithms is possible.

Table 6-8. Summary of existing automatic landmarking methods.

Method	Author	Number of landmarks	Accuracy
Simple Knowledge-based	Lévy-Mandel A D et al. [13]	23	---
Knowledge-based pyramid method	Parthasarathy S et al. [12]	10	1mm: 18% 2mm: 58% 5mm: 100%
Knowledge-based on soft tissue	Tong. W et al. [14]	17+10	1mm: 40% 2mm: 70% 5mm: 95%
Knowledge-based blackboard architecture	Davis D N et al. [19]	19	1mm: 63% 2mm: 74%
Target Recognition	Cardillo J et al. [23]	20	2mm: 76%
Spatial spectroscopy	Rudolph et al. [25]	15	4mm: 100%
Fuzzy detection	Sanei. S et al. [27]	30	96%
Feature subimage extraction	Chen Y T et al. [1,2]	9	94%
Active shape models	TJ Hutton et al. [28]	16	1mm: 13% 2mm: 35% 5mm: 74%

As indicated in the table, some of the methods have been tested on 20 or more landmarks. We have restricted our landmarks to 10 due to time and resource limitations.

On average 80% of the detected landmarks are within 2mm of their accurate positions. Comparing with the reported accuracy of other works, the results obtained here are at least comparable to theirs and are mostly much better. A major advantage of the proposed algorithm is that there is no need to develop new algorithm for detecting new landmarks. All it takes is to train the neural network to recognize the new landmark using the standard training algorithm. Figure 6-15 compares our detection result with one of the well-known knowledge-based methods [49] by showing the percentage of the landmarks detected within 2mm of their accurate positions (This knowledge-based method does not detect the landmark “Spine of nasal bone”). We can see that the performance of our method is comparable to this knowledge-based method.

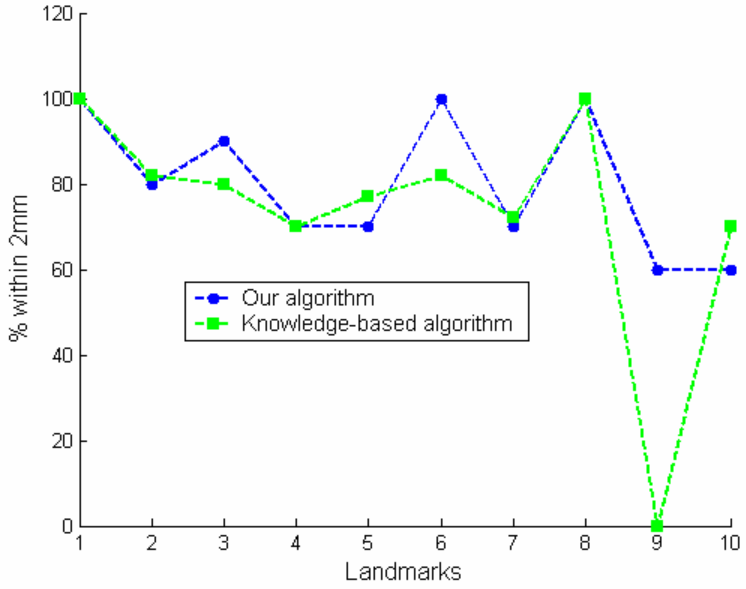


Figure 6-15. Comparison with Han's algorithm.

6.2 Performance of 3-D Cephalometric Landmark Position Computation

We wish to build a fully automatic computer-aided landmark detection system that can compute the 3-D landmark positions based on automatic detected 2-D landmark positions. Unfortunately we do not have enough posteroanterior view cephalograms for neural network training. So in this section the possibility of 3-D cephalometric landmark position computation is studied based on the manually input 2-D landmark positions.

6.2.1 Result of 3-D landmark position computation

The lateral and posteroanterior view cephalograms (provided by University of Copenhagen, Denmark) shown in Figure 6-16 were taken simultaneously in exact

orthogonal positions. Markers have been put on this patient at some landmark positions, which can be seen on both cephalograms. (The landmarks here are not standard landmarks since the locations where the markers could be put on the patients are very limited and not possible to be at the exact standard landmark positions.) We select the landmarks with markers on the posteroanterior view cephalogram. Lines passing through these landmarks representing the rays from x-ray source to the films (based on the geometric relationship explained in Section 5.2.2) have been drawn on these two cehalograms. We can see that the lines drawn for each landmark pass through the corresponding landmark markers on lateral view cephalogram. This has confirmed the possibility of 3-D landmark position computation by 2-D landmark positions. Since we do not have the physical measured values of this patient to verify our system, the experiments in Sections 6.2.2 and 6.2.3 have been done to further evaluate our algorithm.

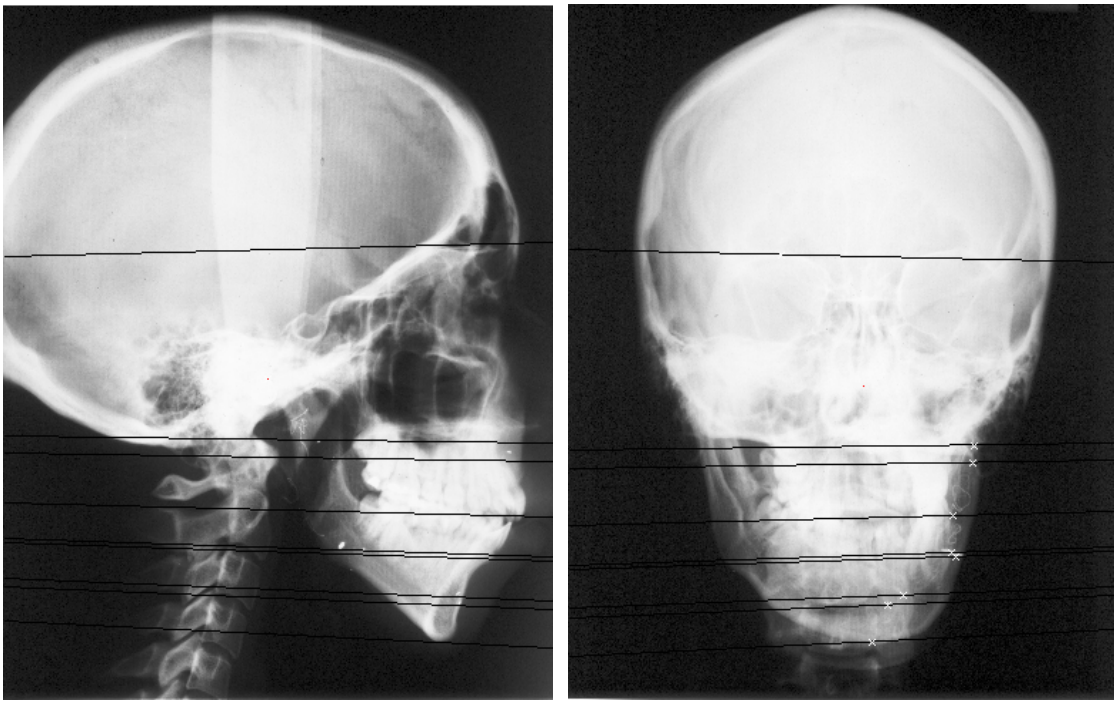


Figure 6-16. Orthogonal cephalograms for 3-D landmark position computation.

6.2.2 Evaluation on dry skulls

For better evaluation of the 3-D landmark position computation method, we should compare our results with physical measured values. This can be done by taking the cephalograms of dry skulls and computing the 3-D landmarks by our system. To verify the algorithm, we measured the distances between the landmarks by a vernier calliper from the dry skull and compared them with the values computed by our algorithm. Table 6-9 shows the results for one of the skulls.

Table 6-9. 3-D landmark position computation evaluation on dry skull.

Distance of landmarks	Computed distance (mm)	Accurate distance (mm)	Error (mm)
LoR-SoR	21.63	31.22	9.59
LoR-OrR	19.80	32.22	12.42
OrR-MoR	20.40	32.00	11.60
SoR-MoR	18.07	23.74	5.67
MoR-N	13.42	13.22	0.20
N-MoL	13.33	14.84	1.51
MoL-SoL	17.06	25.32	8.26
SoL-LoL	22.18	20.42	1.76
LoL-OrL	21.59	29.78	8.19
OrL-MoL	18.57	29.22	0.65
Rhn-PiR	17.30	25.58	8.28
Rhn-RiL	18.30	20.00	1.70
PiR-ANS	19.28	19.22	0.06
PiL-ANS	20.02	19.44	0.58
ANS-A	4.14	7.10	2.96
A-Sdl	16.71	10.70	6.01
UmR-Sdl	35.33	35.46	0.13
Sdl-UmL	33.89	34.80	0.91
Idl-B	14.11	13.52	0.59
B-Pog	11.82	12.34	0.52
Pog-Men	8.80	8.32	0.48
GoR-AgR	15.94	18.00	2.06
AgR-Men	54.73	66.28	11.55
Men-AgL	55.42	68.52	13.10
AgL-GoL	15.68	20.58	4.90

Unfortunately, due to resource limitation (i.e., we do not have an x-ray machine that can take the lateral and posteroanterior view cephalograms simultaneously at exact orthogonal position) the cephalograms we used here were not taken simultaneously and were only approximately orthogonal. It is clear to see that the errors are too big to be acceptable. The 3-D landmark position computation should perform on cephalograms that taken simultaneously in exact orthogonal positions. The experiment in Section 6.2.3 shows that the two cephalograms are not in exact orthogonal positions is the major source of errors of this experiment.

6.2.3 Evaluation on test block

The lateral and posteroanterior view cephalograms used in Section 6.2.2 cannot be confirmed in exact orthogonal positions. The headset holding the dry skull rotates 90° from its original position (lateral view position) to take the posteroanterior view cephalogram. This imperfect procedure of rotation can be a very important source of error. In order to further evaluate the 3-D landmark position computation method, a cubic wooden test block was built. Markers $M1$ to $M6$ were put on the surface of the block as the presumed landmarks. ($M1, M2$ and $M6$ are on surface *befd*, $M3, M4$ and $M5$ on surface *abdc*). Figure 6-17 shows the geometry of the test block. The dimension of the block is $87.56\,\text{mm} \times 124.18\,\text{mm} \times 196.86\,\text{mm}$.

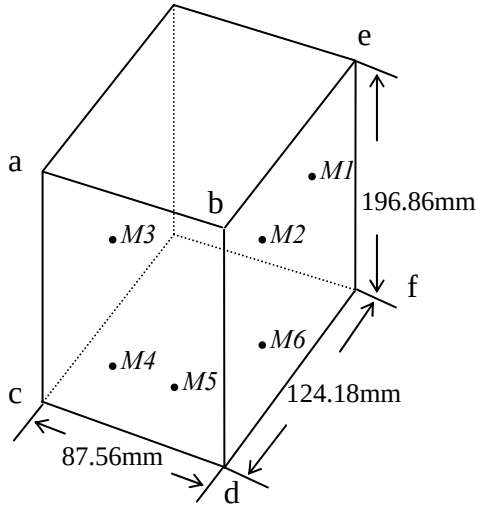


Figure 6-17. Test block for 3-D landmark position computation.

When we took its cephalograms, the positions of the test block were adjusted by a rotation table independently from the head holder of the cephalometric apparatus to ensure that they are orthogonal. To verify the algorithm, we measured distances between the presumed landmarks by a vernier calliper from the test block and compared them with the values computed by our algorithm. Table 6-10 shows the results.

Table 6-10. 3-D landmark position computation evaluation on test block.

Distance of landmarks	Computed distance (mm)	Accurate distance (mm)	Errors (mm)
1-2	29.25	29.52	0.27
1-3	82.54	83.72	1.18
1-4	126.63	127.68	1.05
1-5	123.55	124.86	1.31
1-6	113.08	114.28	1.2
2-3	71.39	70.72	0.67
2-4	120.36	118.54	1.82
2-5	116.44	115.80	0.64
2-6	111.42	110.74	0.68
3-4	75.26	73.60	1.66
3-5	89.14	87.62	1.52
3-6	110.03	108.64	1.39
4-5	30.59	28.94	1.65
4-6	68.93	68.48	0.45
5-6	42.64	43.38	0.74

We can see that the errors of the experiment have been confined to within 2 mm. The accuracy has been improved by taking the two cephalograms in the orthogonal positions in this section. This has confirmed the effectiveness of our 3-D landmark position computation method. Compared with the experiment in Section 6.2.2, it means that the major source of errors in Section 6.2.2 is the two cephalograms are not taken in exact orthogonal positions and the 3-D landmark position computation should be done on cephalograms that take simultaneously in exact orthogonal positions.

Chapter 7

CONCLUSION AND RECOMMENDATION

This chapter concludes the thesis for our study and recommends how the project can be further improved.

7.1 Conclusion

A PC-based cephalometric landmarking system, which combines wavelet technique, neural network and GA, has been implemented in this study. It realizes 2-D automatic landmark detection of cephalograms and provides the functions of computing 3-D landmark position by detecting 2-D landmarks from the lateral and posteroanterior projections of cephalograms.

The 2-D automatic landmarking function employs wavelet decomposition, neural network, GA and some other basic image processing techniques. A major advantage of the proposed algorithm is that there is no need to develop new algorithms for detecting new landmarks. All that is required is to train the neural network to recognize the new landmark. The results of 2-D automatic landmark detection confirm the effectiveness

of the system. On average 80% of the detected landmarks are within 2mm of their correct positions. The accuracy is compatible with other automatic landmarking techniques.

The 3-D landmark position computation function of this system provides the possibility of 3-D automatic computer-aided cephalometrics. Due to time limitation, there is still room for improvement for this function. The 2-D landmark information we use for 3-D computation, which get from manually detection function, could be replaced by the automatic detected landmarks in the future to build a fully automatic computer-aided cephalometric system.

7.2 Recommendations

This thesis has detailed our work made towards creating a computer-aided cephalometric system but it still got distance from the destination. The project could be further improved in the following ways.

- Due to time limitation, we did not test as many landmarks as some other research work. It is recommended that more landmarks are included to improve the system.

- Landmarks on the posteroanterior view cephalograms should be detected by the 2-D automatic landmarking function so that the 2-D landmark information we use for 3-D computation could be all from the automatic detected landmarks to make

a fully automatic computer-aided cephalometric system. We have not done so now due to time limitation.

- The possibility of using basal view cephalograms could also be studied in the future.

- The accuracy of the 3-D landmark position computation needs to be evaluated in a better way. For example, to take the cephalograms of dry skull simultaneously in exact orthogonal positions and compare the computed results with the real-time values.

Reference

[1] Y. T. Chen, K. S. Cheng, and J. K. Liu, “Improving cephalogram analysis through feature subimage extraction”, *IEEE Engineering in Medicine and Biology Magazine,* vol. 18, pp. 25-31, Jan-Feb. 1999.

[2] Y. T. Chen, K. S. Cheng, and J. K. Liu, “Feature subimage extraction for cephalogram landmarking”, *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* Hong Kong, vol. 20, no. 3, pp. 1414-1417, Oct-Nov. 1998.

[3] E. M. Robert, *Handbook of Orthodontics*, Mosby, Incorporated, 1973.

[4] J. Gordon, “Cephalometric tracing: Bonacord analysis examples”. http://www.drflash.com/ProfessionalArea/Documents/ceph_files/frame.htm

[5] B. Sheldon, H. M. Francis and C. Sean, “Three-dimensional x-ray stereometry from paired coplanar images: A progress report”, *American Journal of Orthodontics,* vol. 84, pp. 292-312, Oct. 1983.

[6] G. Barry, C. Court, L. B. Fred, K. Hiechun, G. M. Joseph, “The three-dimensional cephalogram: Theory, technique, and clinical application”, *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 94, pp. 327-337, Oct. 1988.

[7] B. Sheldon, H. M. Francis and C. Sean, “The geometry of three-dimensional measurement from paired coplanar x-ray images” *American Journal of Orthodontics,* vol. 84, pp. 312-322, Oct. 1983.

[8] T. Rakosi, *An Atlas and Manual of Cephalometric Radiography*, Wolfe Medical Publications, 1982.

[9] R. E. Moyers, *Handbook of Orthodontics*, 4th ed., Yearbook Medical Publishers, 1988.

[10] A. E. Athanasiou, *Orthodontic Cephalometry*, Mosby-Wolfe, 1995.

[11] A.M. Cohen, H. H-S. Ip, and A.D. Linney, "A preliminary study of computer recognition and identification of skeleton of skeletal landmarks as a new method of cephalometric analysis", *British Journal of Orthodontics*, vol. 11, no. 3, pp. 143-154, Jul. 1984.

[12] S. Parthasarathy, S. T. Nugent, P. G. Gregson, D. F. Fay, "Automatic landmarking of cephalograms", *Computers and Biomedical Research*, vol. 22, pp. 248-269, Jun. 1989.

[13] A. D. Lévy-Mandel, A. N. Venetsanopoulos, J. K. Tsotsos, "Knowledge-based landmarking of cephalograms", *Computers and Biomedical Research*, vol. 19, pp. 282-309, Jun. 1986.

[14] W. Tong, S. T. Nugent, G. M. Jensen, D. F. Fay, "An algorithm for locating landmarks on dental x-rays", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Seattle, WA USA, vol. 2, pp. 552-554, Nov. 1989.

[15] J. L. Contreras-Vidal and J. Garza-Garza, "A knowledge-based system for image processing and interpretation of cephalograms", *Proceedings of IEEE Canadian*

Conference on Electrical and Computer Engineering, Ottawa, Canada, pp. 75, Sep. 1990.

[16] P. H. Jackson, G. C. Dickson, and D. J. Birnie, "Digital image processing of cephalometric radiographs: A preliminary report", *British Journal of Orthodontics*, vol. 12, pp. 122-132, Jul. 1985.

[17] A. M. Cohen, F. C. Ip and A. D. Linney, "A preliminary study of computer recognition and identification of skeletal landmarks as a new method of cephalometric analysis". *British Journal of Orthodontics*, vol. 11, pp. 143-154, Jul. 1984.

[18] A. M. Cohen and A. D. Linney, "A low cost system for computer-based cephalometric analysis", *British Journal of Orthodontics*, vol. 13, pp. 105-108, Apr. 1986.

[19] D. N. Davis, C. J. Taylor, "A blackboard architecture for automating cephalometric analysis", *Journal of Medical Informatics*, vol. 16, no. 2, pp. 137-149, 1991.

[20] T. Rakosi, "An atlas of cephalometric radiography", *Wolfe Medical Publications*, London, 1982.

[21] D. B. Forsyth, D. N. Davis, "Assessment of an automated cephalometric analysis system", *European Journal of Orthodontics,* vol. 18, pp. 471-478, Oct. 1996.

[22] D. B. Forsyth, W. C. Shaw, S. Richmond, C. T. Roberts, "Digital imaging of cephalometric radiographs, part 2: image quality", *The Angle Orthodontist*, vol. 66, pp. 43-50, 1996.

[23] J. Cardillo, M. A. Sid-Ahmed, “An image processing system for locating craniofacial landmarks”, *IEEE Transactions in Medical Imaging*, vol. 13, pp 275-289, Jun. 1994.

[24] J. Cardillo, M. A. Sid-Ahmed, “Target recognition in a cluttered scene using mathematical morphology”. *Pattern Recognition*, vol. 29, pp. 27-49, Jan. 1996.

[25] D. J. Rudolph, P. M. Sinclair, J. M. Coggins, “Automatic computerized radiographic identification of cephalometric landmarks”, *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 113, pp. 173-179, Feb. 1998.

[26] Y. T. Chen, K. S. Cheng and J. K. Liu, “Automated cephalogram landmarking”, *Chinese Journal of Medicine and Biology Engineering*, vol. 16, no. 2, pp. 199-213, 1996.

[27] S. Sanei, P. Sanaei, M. Zahabsaniei, “Fuzzy detection of craniofacial landmarks”, *IEEE Sixth International Conference on Image Processing and Its Applications*, Dublin Ireland, vol. 1, pp. 376-379, July. 1997.

[28] T. J. Hutton, S. Cunningham and P. Hammond. “An evaluation of active shape models for the automatic identification of cephalometric landmarks”, *European Journal of Orthodontics*, vol. 22, pp. 499-508, Oct. 2000.

[29] T. F. Cootes, F. J. Edwards and Taylor C J, “Active appearance models”, *Proceedings of the European Conference on Computer Vision*, Freiburg, Germany, vol. 2 pp. 484-498, Jun. 1998.

[30] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham, "Active shape models-their training and application", *Computer Vision and Image Understanding*, vol. 61, no. 2, pp. 38-59, Jan. 1995.

[31] A. Hill, T. F. Cootes and C. J. Taylor, "A genetic system for image interpretation", *Proceedings of the 3rd British Machine Vision Conference*, pp. 276-285, Sept. 1992.

[32] B. H. Broadbent, "A new x-ray technique and its application to orthodontia", *The Angle Orthodontist*, vol. 1, pp. 45-66, Apr. 1931.

[33] L. Fred, G. Barry, C. K. Hie, and G. Joseph, "Landmarks in three dimensions: Reconstruction from cephalograms versus direct observation", *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 100, pp. 132-133, Aug. 1991.

[34] M. Yoshihide, M. Takahiro, M. Katsuhiro and S. Masayoshi, "An accurate three-dimensional cephalometric system: a solution for the correction of cephalic malpositioning", *Journal of Orthodontics*, vol. 28, pp. 143-149, Jun. 2001.

[35] R. Polikar, "The engineer's ultimate guide to wavelet analysis: the wavelet tutorial".
http://engineering.rowan.edu/~polikar/WAVELETS/WTtutorial.html

[36] F. Chaplais, "A wavelet tour of signal processing".
http://cas.ensmp.fr/~chaplais/Wavetour_presentation/Wavetour_presentation_US.html

[37] J. Froehlich, "Neural network with java".
http://rfhs8012.fh-regensburg.de/~saj39122/jfroehl/diplom/e-index.html

[38] H. Simon, *Neural Network, A Comprehensive Foundation*. 2nd edition, Prentice-Hall, 1999.

[39] L. Smith, “An introduction to Neural Networks”. http://www.cs.stir.ac.uk/~lss/NNIntro/InvSlides.html#why

[40] S. E. Fahlman, "Faster-learning variations on back-propagation: an empirical study", *Proceedings, Connectionist Models Summer School. Morgan-Kaufmann,* Los Altos CA, CMU-CS-88-162, Sep. 1988.

[41] T. C. Wong, “Genetic algorithm”. http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/tcw2/report.html#Content

[42] J. P. Rennard, “Genetic algorithm viewer: demonstration of a genetic algorithm”. http://www.rennard.org/alife/english/gavintrgb.html#rTit01

[43] E. Falkenauer, “Genetic algorithms”. http://www.optimaldesign.com/technology.htm

[44] Orthomail Network, “Orthodontic glossary”. http://www.orthomail.us/c.php.

[45] S. Krishna, D. David, “Scanned bi-orthogonal radiographs as a source for 3D cephalometric data”, *SPIE Proceedings*, vol. 2710, pp. 717-724, Apr. 1996.

[46] J. R. Smith, S. F. Chang, “Automated binary texture feature sets for image retrieval”. http://www.ctr.columbia.edu/~jrsmith/html/pubs/abtfsfir/node4.html

[47] N. Sebe, M. S. Lew, “Wavelet based texture classification”. http://carol.science.uva.nl/~nicu/publications/ICPR00_texture.pdf

[48] D. S. Zhang, A. Wong, M. Indrawan, G. J. Lu, “Content-based image retrieval using Gabor texture features”. http://www.gscit.monash.edu.au/~dengs/resource/papers/pcm00.pdf

[49] X. Han, “Landmark identification in later x-ray images” *Master’s Thesis*, National University of Singapore: Electrical and Computer Engineering Department, 1998.

Appendix A

Cephalometric Landmarks

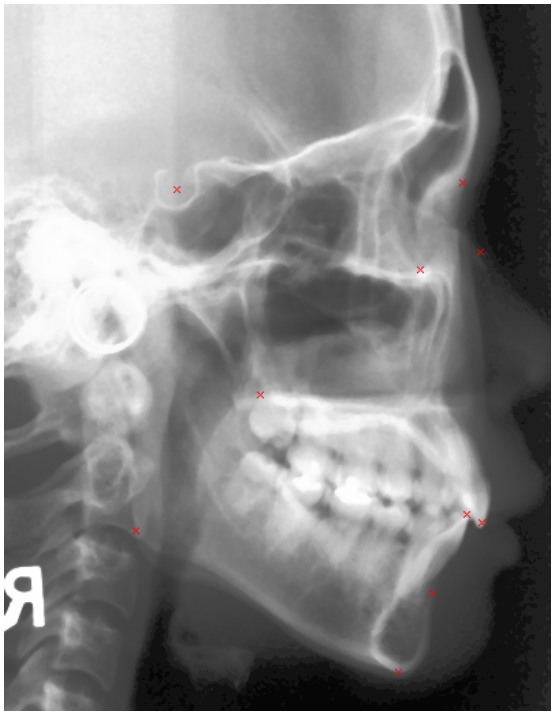
Table A-1. Cephalometric Landmarks.

No.	Name	Definition
1	Nasion	The most anterior point of the nasofrontal suture in the median plane. The skin nasion (N) is located at the point fo maximum convexity between nose and forehead.
2	Sella	The sella point is defined as the midpoint of the line posterior clinoid process and the anterior opening of the turcica.
3	Midpoint of the entrance to the sella	This point represents the midpoint of the line connecting the posterior clinoid process and the anterior opening of the turcica.
4	Subnasale	A skin point where the nasal septum merges mesially with the integument of the upper lip.
5	Point A, subspinale	The deepest midline point in the curved body outline from the base to the alveolar process of the maxilla.
6	The anterior landmark for determining the length of the maxilla	It is constructed by dropping a perpendicular from point A to the palatal plane.
7	Prosthion	Alveolar rim of the maxilla; the lowest, most anterior point on the alveolar portion of the premaxilla, between the upper central incisors.
8	Incisor superious	Tip of the crown of the most anterior maxillary central incisor.
9	Apicale	Root apex of the most anterior maxillary central incisor.
10	Incisor inferious	Tip of the crown of the most anterior mandibular central incisor.
11	Apicale	Root apex of the most anterior mandibular central incisor.
12	infradentale	Alveolar rim of the mandible; the highest, most anterior point on the alveolar process, in the median plane, between the mandibular central incisors.
13	Point B, supramentale	Most anterior part of the mandibular base. It is the most posterior point in the outer contour of the mandiblar alveolar process, in the median plane.
14	Pogonion	Most anterior point of the bony chin, in the median plane.
15	Gnathion	It is located between the most anterior and the most inferior point of the bony chin.
16	Gonion	A constructed point, the intersection of the lines tangent to the posterior margin of the ascending ramus and the mandibular base.
17	Menton	It is the lowest point fo the mandible.
18	The anterior landmark for determining the length of the mandible	It is defined as the perpendicular dropped from Pogonion to the mandibular plane.
19	Articulare	The intersection of the posterior margin of the ascending ramus and the outer margin of the cranial base.
20	Condylion	Most superior point on the head of the condyle.
21	Orbitale	Lowermost point of the orbit in the radiograph.
22	Soft tissue Pogonion	Most anterior point of the skin of the chin in the median plane.
23	Tip of the nose	The most anterior point of the skin of the nose in the median plane.
24	Anterior nasal spine	The tip of the bony anterior nasal spine in the median plane.

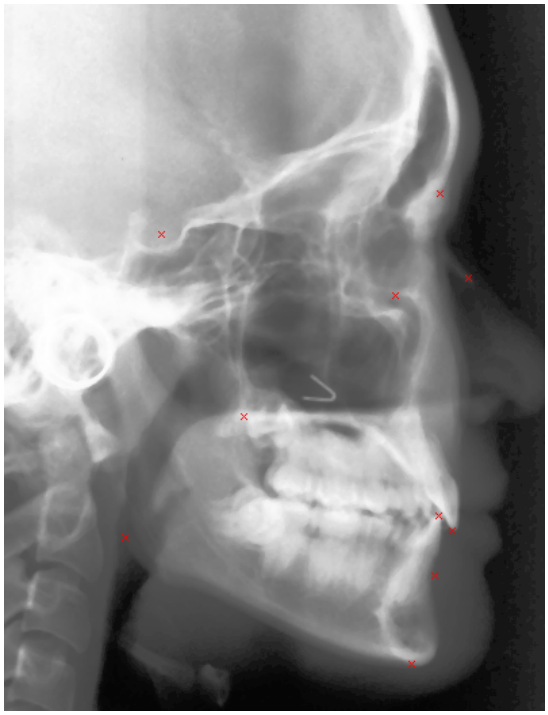
25	Posterior nasal spine	A constructed radiological point marking the intersection of a continuation of the anterior wall of the pterygopalatine fossa and the floor of the nose. It marks the dorsal limit of the maxilla.
26	Landmark for assessing the length of the maxillary base, in the posterior section	It is defined as a perpendicular dropped from point S to a line extending the palatal plane.
27	Anterior point for the occlusal plane	A constructed point, the midpoint in the incisor overbite in occlusion.
28	Posterior point for the occlusal plane	The most distal point of contact between the most posterior molars in occlusion.
29	Basion	Lowest point on the anterior margin of the foramen magnum in the median plane.
30	Pterygomaxillary fissure	The contour of the fissure projected onto the palatal plane.

Appendix B

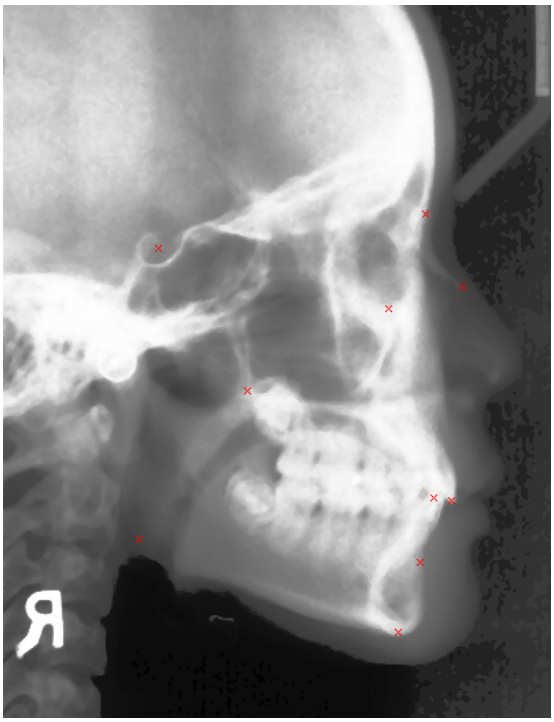
Test Images After Landmark Detection



Test image 1.



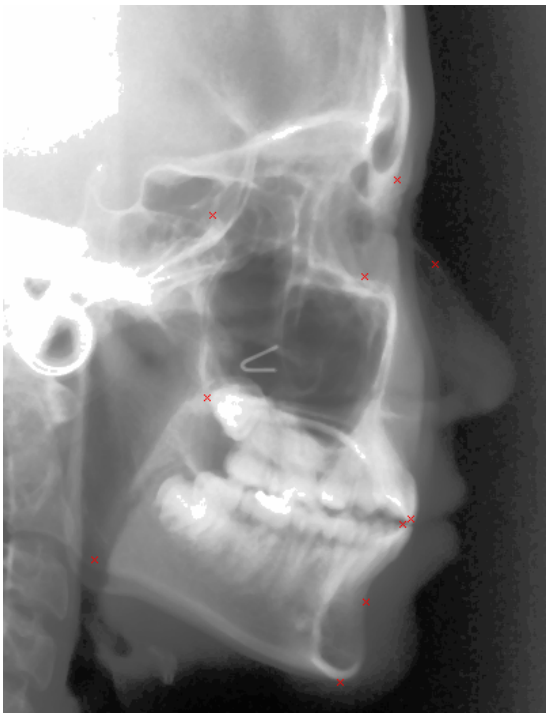
Test image 2.



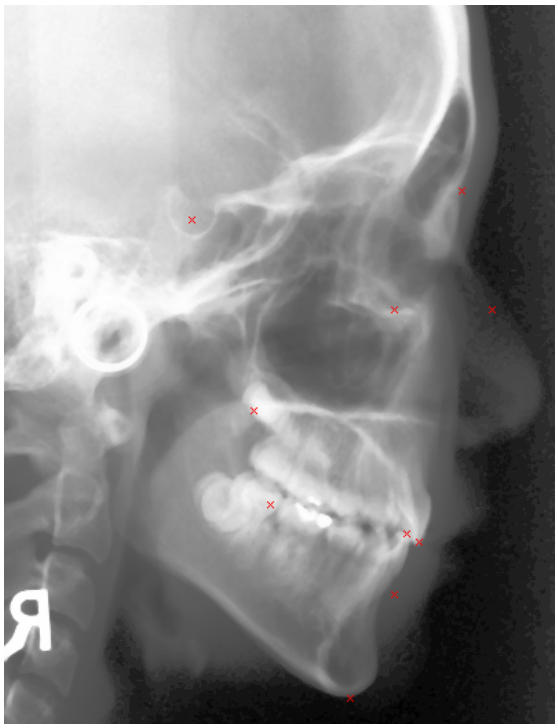
Test image 3.

Test image 4.

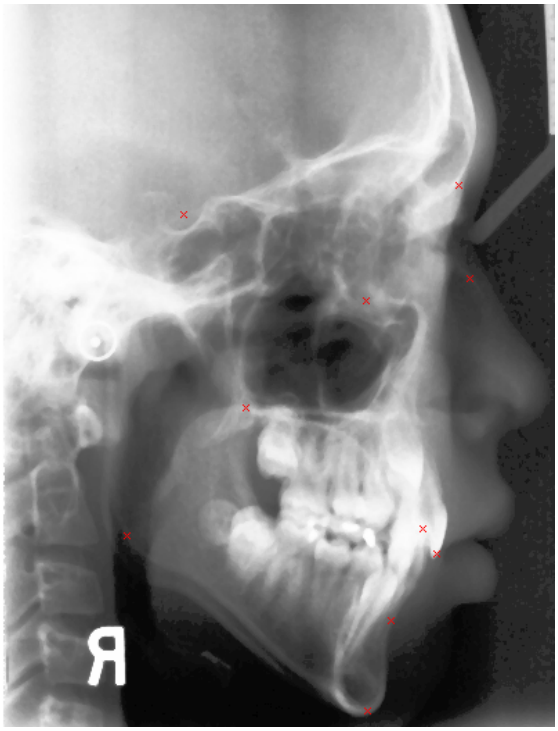
Test image 5.



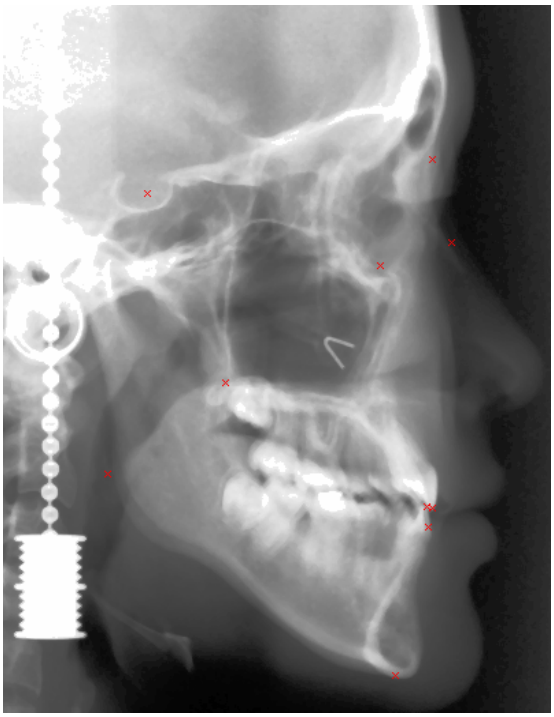
Test image 6.



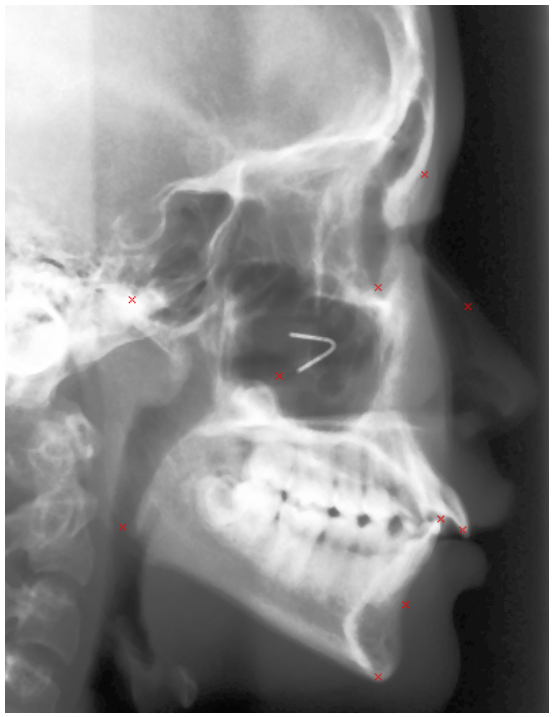
Test image 7.



Test image 8.



Test image 9.



Test image 10.

Table B-1. 2-D Landmark detection result.

Images / Landmarks	1	2	3	4	5	6	7	8	9	10
IS	5	2	3	4	2	2	4	5	3	3
S	1	2	4	2	0	71	3	3	1	54
N	1	4	3	3	3	3	5	7	6	5
B	4	5	3	25	4	3	14	4	41	4
Go	5	10	4	5	4	16	113	3	6	5
II	3	0	5	1	4	6	1	1	4	2
PNS	11	2	2	4	0	4	10	4	4	53
Me	2	5	2	3	1	4	4	5	3	3
Snb	5	3	4	4	4	17	3	7	11	9
Or	1	5	4	3	7	8	6	24	17	4

Table B-1 shows the distances between the detected landmarks and their corresponding accurate landmark positions in pixels. (3 pixels ≈ 1mm, 6 pixels ≈ 2mm, 15 pixels ≈ 5mm)