

COMPUTER ASSISTED BONE AGE ASSESSMENT

by

Mahmut Haktan

B. S. In Physics

University of Istanbul, 2000

Submitted to the Institute of Biomedical Engineering

in partial fulfillment of the requirements

for the degree of

Master of Science

in

Biomedical Engineering

Boğaziçi University

October, 2005

COMPUTER ASSISTED BONE AGE ASSESSMENT

APPROVED BY:

Assoc. Prof. Dr. Albert Güveniř
(Thesis Supervisor)

Assistant Prof. Dr. Burak Güçlü

Assoc. Prof. Dr. Halil Özcan Gülçür

DATE OF APPROVAL: 24.10.2005

ACKNOWLEDGMENTS

I would like to thank my thesis supervisor, Assoc. Prof. Dr. Albert Güveniř for his support and motivation during my thesis study. Without his valuable academic support, experience, and knowledge, this study would not be completed. I would like to thank Dr. Nihal who shared her hand radiograph archive with us, and Dr. Mehmet Ertürk for his support in řisli Etfal Hospital. I would also like to thank Prof. Dr. Tuęrul Biren M.D. who has helped me to reach the archives of the University of Marmara. I would like to thank 29 May Hospital and in particular Nihal Er who helped me to digitalize the hand radiographs. I also would like to thank Dr. Demet Bōrekçi who helped me on the every steps of this thesis.

I would like to thank Adem Altun who has helped me to develop the web-based bone age assessment software in the .Net environment. A big thank goes my friends of B.U. Biomedical Engineering Institute. Working and sharing life with you is unforgettable.

And to those who directly and indirectly help me, thank you.

ABSTRACT

COMPUTER ASSISTED BONE AGE ASSESSMENT

Bone age assessment based on the radiological examination of the left hand and wrist is a procedure frequently performed to evaluate the growth of pediatric patients. In this thesis we examined studies conducted on the skeletal age analysis and we developed a computerized system for automatic bone age analysis. Firstly we examined two different clinical methods; The Greulich and Pyle method and the Tanner and Whitehouse (TW2) method. We chose the TW2 method to develop a computerized bone age system because it is more suitable for the computer analysis than the Greulich and Pyle method. We developed the web-based TW2 skeletal age calculation software. We then introduced steps that are necessary to transform this manual clinical method into a fully automated system. We explained preprocessing methods that are needed to make a radiograph fit for analysis and steps to find regions of interest and assign them TW2 stages.

Finally we developed a computer program that assigns TW2 stages to a certain region of interest (the middle phalanx of the third finger) in the left hand. In this method we trained all TW2 stage groups for middle phalanx of the third finger and produced mean images for each TW2 stage. These mean images are used to determine the TW2 stage of queried image. We use the correlation between a mean image and a target image as an indicator of which stage should be assigned to the target.

Our results show that 70 % of all test images of the middle phalanx of the third finger have been assigned the correct TW2 stage by our method. We believe that the same method could be used to assign TW2 stages to other regions of interest in the hand.

Keywords: Computerized Bone age assessment, Web-based Tanner and Whitehouse method, Active Shape Model.

ÖZET

BİLGİSAYAR İLE OTOMATİK KEMİK YAŞI TAYİNİ

Kemik yaşı tayini sol el-bilek radiograflarının incelenmesi esasına dayanır ve pediatrik hastalıkların incelenmesinde sıklıkla kullanılır. Bu tezde kemik yaşı tayini ile alakalı yapılan çalışmalar ve bilgisayar ile otomatik kemik yaşı tayini için gerekli olan adımlar incelendi. İlk olarak Greulich&Pyle ve Tanner&Whitehouse (TW2) klinik uygulamaları incelendi ve TW2 uygulamasının bilgisayar ile analiz için daha uygun olduğu görüldü. Daha sonra web tabanlı manüel TW2 iskelet yaşı tayini yapabilen bir uygulama hazırlandı. Manüel klinik uygulamadan yola çıkarak bilgisayar ile tam olarak otomatik kemik yaşı tayini uygulaması için gerekli olan adımlar incelendi. Görüntü işleme öncesi gerekli ön analiz yöntemleri incelendi ve belirli ilgi alanlarına TW2 safhasını atamak için gerekli olan adımlar açıklandı.

Tam olarak otomatik çalışarak kemik yaşı tayini yapabilen bir yazılımın tasarımı yapıldı. Bu yazılım ile sol elin üçüncü parmağının orta boğumuna uygun düşen TW2 safhası atanabilmektedir. Bu yazılımda her bir TW2 safhası için model grafi eğitildi ve dijital atlas için ortalama görüntüler oluşturuldu. Bu görüntüler ile sorgulanan radiografin TW2 seviyesi bulundu. Sorgulanan grafi için en iyi uyuşan ortalama görüntü ile uygun TW2 safhası belirlenmiş oldu.

Deney sonuçları kullanılan yöntem ile tüm grafiğin % 70 inde, üçüncü parmağının orta boğumuna uygun düşen TW2 safhası doğru olarak hesaplanabilmiştir. Kemik yaşı tayini için gerekli olan diğer kemiklerde de bu yöntemin kullanılabileceğini düşünmekteyiz.

Anahtar Sözcükler: Bilgisayar ile Kemik Yaşı Tayini, Web-tabanlı TW2 uygulaması, Aktif Şekilleme Modeli.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
ABSTRACT.....	iv
COMPUTER ASSISTED BONE AGE ASSESSMENT	iv
ÖZET.....	v
BİLGİSAYAR İLE OTOMATİK KEMİK YAŞI TAYİNİ	v
TABLE OF CONTENTS	vi
LIST OF TABLES	viii
LIST OF FIGURES.....	ix
LIST OF ABBREVIATIONS.....	x
LIST OF SYMBOLS.....	xi
1. AUTOMATING SKELETAL AGE ASSESSMENT	1
1.1 Background and Motivation	1
1.2 Uses of Skeletal Age Assessment	1
1.3 Objectives	2
1.4 Outline of the Thesis	2
2. CLINICAL METHODS OF BONE AGE ASSESSMENT.....	3
2.1 Introduction.....	3
2.2 Skeletal Age.....	3
2.3 Bone Age Assessment Methods.....	3
2.3.1 The Greulich and Pyle Method.....	5
2.3.2 The Tanner and Whitehouse (TW2) Method:	8
3. COMPUTERIZED BONE AGE ANALYSIS	12
3.1 Introduction.....	12
3.2 Choosing the Most Suitable Clinical Method.....	12
3.3 Image Processing Steps In Detail.....	13
3.3.1 The Orientation of Hand.....	14
3.3.2 The Position of Hand and Radiograph	14
3.3.3 Finding CROI and EMROI.....	15
3.3.4 Feature Extraction from the CROI and EMROI.....	15
4. PREVIOUS WORK	18

4.1 Introduction.....	18
4.2 Segmentation of the Hand.....	18
4.3 Analysis of the EMROI	23
4.4 Discussion	26
5. WEB BASED HAND ANALYSIS SOFTWARE	28
5.1 Introduction.....	28
5.2 User Requirements	28
5.3 Design and User Interface.....	29
5.4 Performance and Experimental Results:.....	33
5.4.1 The Experiment.....	34
5.4.2 Future Developments of Application	35
6. AUTOMATIC ASSIGNMENT OF A TW2 STAGE TO THE MIDDLE PHALANX OF THE THIRD FINGER.....	36
6.1 Introduction.....	36
6.2 Method: Deformable Models	36
6.2.1 Statistical Shape Models.....	37
6.2.2 Landmarks	38
6.2.3 Modeling Shape Variation.....	39
6.3 Active Shape Models.....	42
6.4 Method: Procrustes Analysis	44
6.5 Outline of the Automatic Method	46
6.5.1 Producing Mean Images for Each Stage	46
6.5.2 Training the ASM	48
6.5.3 Assigning a TW2 stage.....	51
6.6 Experiments	53
6.7 .Discussion and Conclusions	54
6.7.1 ASM Results and Discussion.....	54
6.8 Future Work.....	55
APPENDIX A	58
REFERENCES.....	59
REFERENCES NOT CITED	62

LIST OF TABLES

Table 2.1 Skeletal Age Evaluation Sheet.....	9
Table 3.1 Some of the TW2 stage explanation examples.....	17
Table 5.1 Web-based TW2 Bone age assessment system.	30
Table 5.2 Initialization parameters to begin shape training.	31
Table 5.3 Performance tests of web based bone age assessment system	34
Table 6.1 Training of system.....	47

LIST OF FIGURES

Figure 2.1 Anatomy of left hand.	4
Figure 2.2 Radiograph of Left Hand.	5
Figure 2.3 The ossification centers on finger.....	6
Figure 2.4 Example of the descriptions of the skeletal development.....	7
Figure 2.5 Example of the images are used to find the TW2 stages.	10
Figure 3.1 Steps of fully automated bone age analysis system.....	13
Figure 3.2 All ROIs those are included in the RUS analysis.....	16
Figure 5.1 User chooses appropriate stage for each bone.....	29
Figure 5.2 Web-Based manual TW2 skeletal age calculation.	30
Figure 5.3 User selects landmarks by clicking with mouse.....	32
Figure 5.4 Initialization parameters to begin shape search in an image.....	33
Figure 5.5 Shape search process in different levels.	34
Figure 6.1 Active Contour Model-snake:	36
Figure 6.2 Landmarks choose for training sets which easily locatable in all images.	38
Figure 6.3 The first two modes of a shape model of the EMROI.....	41
Figure 6.4 A line perpendicular to the model boundary is taken through each point.	43
Figure 6.5 Procrustes analysis being applied to two triangular shapes.	45
Figure 6.6 Mean images with a mean shape projected over it.....	47
Figure 6.7 TW2 atlas reference image.....	48
Figure 6.8 A typical EMROI with the fixed model points marked on it.....	49
Figure 6.9 The segmentation.....	52
Figure 6.10 The area of the query (left) and of the mean (right) image are compared	53
Figure 6.11 Some examples of segmentations done with our ASM.	56
Figure 6.12 The marked areas exhibit the most change between stages.	56

LIST OF ABBREVIATIONS

G&P	Greulich and Pyle
TW2	Tanner and Whitehouse
RUS	Radius, Ulna and Short finger bones
ROI	Region of Interest
EMROI	Epiphyseal–Metaphyseal Region of Interest
CROI	Carpal Region of Interest
PROI	Phalangeal Region of Interest
CR	Computerized Radiology
ASM	Active Shape Model
HANDX	Semi automatic bone age assessment system was built by Michael
CASAS	Automatic system for skeletal age by Tanner
LAMF	Longitudinal axis of the middle finger

LIST OF SYMBOLS

\mathbf{x}_i	Aligned training shapes
\mathbf{b}	Vector of parameters of the model
$\bar{\mathbf{x}}$	Mean shape of the N shapes.
ΔX_i	Deviations
S_i^T	2n by 2n covariance matrix
\mathbf{p}	Unit eigen-vectors
λ	Eigenvalues of S
\mathbf{P}	Matrix of the first t eigenvectors
\mathbf{b}_i	Vector of weights (principal components)
D_m^2	Mahalanobis distance
\mathbf{g}_i	Normalized samples
L_{\max}	Number of resolutions
ρ	Correlation

1. AUTOMATING SKELETAL AGE ASSESSMENT

1.1 Background and Motivation

Developing of the computerized bone age assessment system will be useful for the pediatric radiology. It is performed about 600-1000 times per year in the Şişli Etfal Education and Research Hospital. It is not only a time consuming procedure but it is also depend on the experience of the physician. The aim of this project is to improve efficiency and objectivity of the bone age assessment by automating the process.

1.2 Uses of Skeletal Age Assessment

The bone age or skeletal age assessment is a commonly used procedure for pediatric patients to evaluate their growth disorder, examine their growth disorder, determine their growth potential, and monitor therapy effect of growth. The growth potential of an individual depends largely on the progression of ossification within the epiphysis. Bone age is a measurement of the epiphyseal center development. It is an important procedure in the diagnosis and management of endocrine disorders, diagnostic evaluation of metabolic and growth abnormalities, deceleration of maturation in a variety of syndromes, malformations, and bone dysplasias. A simple method frequently used in bone age assessment is atlas matching by a radiological examination of a left hand and wrist radiograph against a reference set of atlas patterns of normal standards. Although the hand and wrist does not contribute to the height of an individual, the radiograph of this part of the body has been proven valuable and is commonly used in assessment of bone age. [1]. After determination of skeletal age, we can compare the results with chronological age. A big difference between these two variables shows an atypical skeletal development of patient. This procedure is often used in the diagnosis and management of endocrine disorders. Generally it can indicate rate of growth of a patient. Bone age assessment is also used in forensic medicine in such conditions when there is a need to confirm chronological age of a criminal.

1.3 Objectives

There are two main objectives of this thesis. The first one is to develop a web based bone age assessment system with the Tanner and Whitehouse method. We can use this system in the clinical practice and for educational purposes. The second objective of this thesis is to automate assigning of the TW2 stages to a certain region of interest.

1.4 Outline of the Thesis

Chapter 1 introduces the subject and gives outline of this thesis work. Clinical methods of bone age assessment are discussed in Chapter 2. In Chapter 3 we will examine steps that are necessary for the fully automated bone age assessment system. In chapter four previous studies are summarized. First part of that chapter is on the segmentation of the hand and ROIs of the hand. In the second part we examine assigning the appropriate TW2 stages to a ROI. In chapter five details of the hand analysis system is examined. Finally we focused on assigning of the TW2 stage to the middle phalanx of third finger. Conclusions and future work can be found in this chapter.

2. CLINICAL METHODS OF BONE AGE ASSESSMENT

2.1 Introduction

In this chapter Clinical practice for the bone age assessment will be examined so that we can choose the most appropriate clinical method for the computerized bone age assessment system.

2.2 Skeletal Age

Skeletal age, a measure of skeletal development of a child, is determined by using standard data that includes measurement of skeletal developments of healthy children. Numerous Roentgen methods have been proposed to assess the bone age according to criteria such as the time of appearance, size and differentiation of the ossification centers. Most commonly used methods are Greulich and Pyle method (G&P), and Tanner and Whitehouse (TW2) method.

2.3 Bone Age Assessment Methods

Greulich and Pyle method (G&P) and Tanner and Whitehouse (TW2) use atlas matching methods for manual determination of skeletal age. G&P method is easier and faster to use compare to TW2 method. However TW2 method is more reliable.

Both methods rely on radiographs taken from the left hand. Prior to detailed examination of the two methods it is useful to know a little about the anatomy of the hand in order to understand some of the terms used. Figure 2.1 and 2.2 show the most important bones for skeletal age determination in the hand. In Figure 2.1 we see a schematic representation of the bones of the hand. Figure 2.2 shows an actual hand radiograph. The most important hand bones are those of the fingers, the proximal, middle and distal phalanges. The thumb has only proximal and distal phalanges. Next there are the carpal bones, Capitate, Hamate, Triquetral, Lunate, Scaphoid, Trapezium and Trapezoid. In the middle we see the five metacarpals.

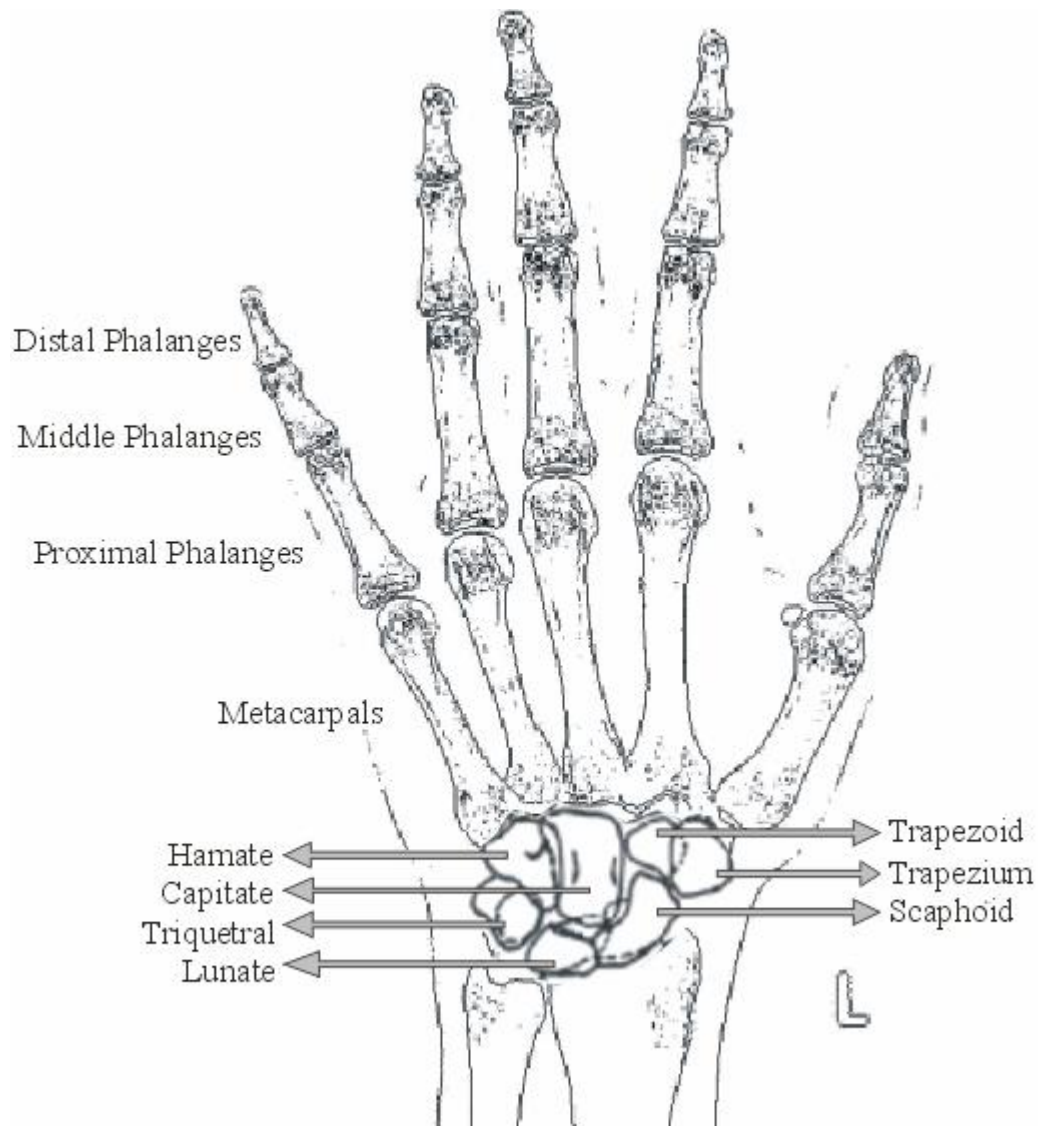


Figure 2.1 Anatomy of left hand.

Ossification centers are located between the phalanges of developing children. Because the radiograph in Figure 2.2 is of an adult subject they can not be seen. In Figure 2.3 an example of the situation in the developing children so it can be seen. We can clearly see three parts, the metaphysis, the epiphysis and the diaphysis. We call that structure as an Epiphyseal-Metaphyseal Region of Interest (EMROI). This structure changes during the development of the skeleton. The epiphysis becomes steadily wider and eventually fuses with the metaphysis.



Figure 2.2 Radiograph of Left Hand.

2.3.1 The Greulich and Pyle Method

First studies on the human growth and development started in 1929 at the Western Reserve University School of Medicine in Ohio. A Large number of children of different ages were used in these studies. Left shoulder, elbow, hand, hip and knee radiographs were taken for examination. In the first postnatal year an examination was conducted every three months, from twelve months to five years they were examined each 6 months and annually thereafter. In total the study ran from 1931 until 1942. In 1937 “Atlas of Skeletal Maturation of Hand” was published by Todd. [2]

The Greulich-Pyle has been used extensively by orthopedists and was used by Green and Anderson in compiling their data for growth remaining in children nearing skeletal maturity. Since the Moseley straight line graph was based on the Green-Anderson data, the Greulich-Pyle system is correlated with that graph also. The two methods do not give equivalent bone ages. The Tanner-Whitehouse with computer assistance may be more reproducible, but it is more cumbersome to use. For the present, the Greulich-Pyle method is still standard for the orthopedic use, even though it was derived more than a half century ago on an exclusively white upper-middle class population.

Assessment steps:

- Assess one bone at a time
- Locate the Atlas plate that most closely resembles the X-Ray bone
- Interpolate between Atlas plates.
- Assess a second time.
- Average the two twenty-eight assessments.

In the Greulich and Pyle method twenty-eight bones are compared against (normal) the reference atlas. All the bones in the hand and wrist are examined and average of the assessments gives us the bone age of patient.

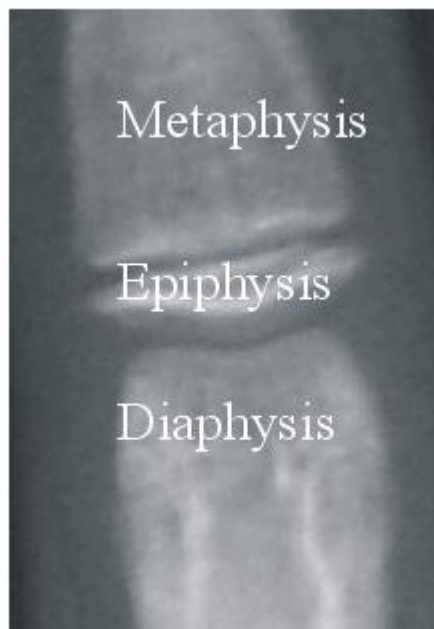


Figure 2.3 The ossification centers on finger.

Development stages of each bone are described in the atlas in Figure 2.4. The descriptions in the figure are more a general guideline to the development of each bone in the hand rather than an instruction on how to rate a bone. Most institutions are using a more rapid modified version of the original atlas, which is also potentially less accurate. This version is described below.

The atlas is divided into two parts, one for the male patients and one for the female patients because females develop quicker than males. Each part contains standard radiographic images of the left hand of children ordered by the chronological age.

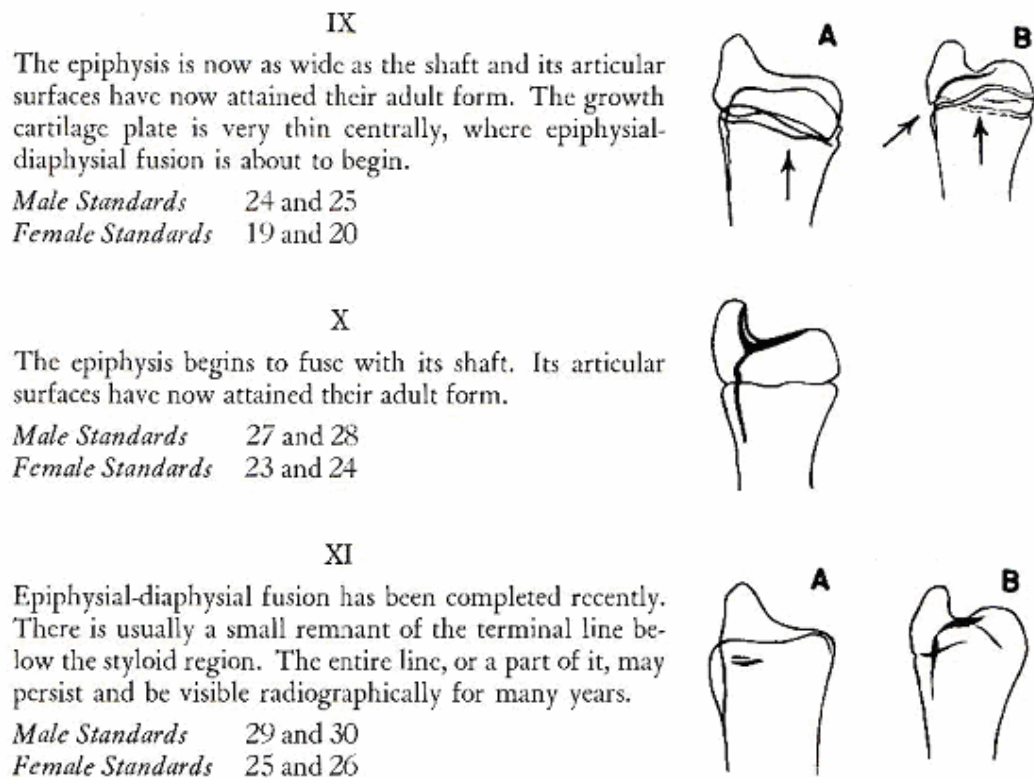


Figure 2.4 Example of the descriptions of the skeletal development found in the Greulich and Pyle atlas.

To measure skeletal age of a patient first of all, the radiograph of the patient is compared with the image in the atlas that is match with the chronological age of the patient. Next, one should compare it with the adjacent images representing both younger and older children. There are some maturity indicators when comparing the radiograph of the patient against an image in the atlas. These features can vary with the race, age and sex of the child. In the younger children the presence or absence of the certain carpal or epiphyseal ossification centers are often pointers for the physician about the skeletal age of

a child. In older children the shape of the epiphyses and the amount of fusion with the metaphysis is a good indicator for the skeletal age, carpal ossification centers did not differ at that time. Once the atlas image that most resembles the radiograph is found, the physician should conduct a more detailed examination of the individual bones and epiphyses. So we can find the skeletal age when the matching radiograph has been found.

2.3.2 The Tanner and Whitehouse (TW2) Method:

Tanner and Whitehouse knew about the Greulich and Pyle atlas and the way in which it was used. They noticed several aspects of the method that they felt needed to be improved. According to them the subjectivity of the matching process was an obvious weakness of the G&P method. Physicians generally look at the whole radiograph at once and then compare it with an image in the atlas. Often a specific radiograph does not match any of the images in the atlas exactly, and little guidance is given on how to balance out the discrepancies that arise from one bone being more or less advanced than its match in the atlas.

The scale used for expressing the maturity was another aspect of the G&P method TW2 didn't like. Each standard radiograph in the atlas has the age of the child from whom it was taken associated with it. Therefore maturity is measured on an age scale. It could even be seen as the predicted age, because the matching process gives the most likely chronological age of the child being matched, as judged from the radiograph. This most likely age is then said to be the skeletal age. Tanner and Whitehouse deemed that a new, more sophisticated system was needed which would not make use of an age scale for maturity measurements. In their view a maturity scale should be defined in a manner which does not directly relate to age. This would allow them to produce a set of "maturity standards" of any given population by studying the relationship between maturity and age.

To build standard atlas radiographs of up to 12 years old, healthy children were used in the period of six months. Changes in the shape and density markings of each bone were recorded in standard atlas. After this the different stages were identified, these stages had to be universally present in all individuals.

Table 2.1
Skeletal Age Evaluation Sheet

Skeletal Age Evaluation Sheet			
Subject: _____	Birthday: __/__/__	Date: __/__/__	
Assessment Location	1st Read	2nd Read	Average
1. Capitate	_____	_____	_____
2. Hamate	_____	_____	_____
3. Triquetral	_____	_____	_____
4. Lunate	_____	_____	_____
5. Scaphoid	_____	_____	_____
6. Trapezium	_____	_____	_____
7. Trapezoid	_____	_____	_____
8. Distal Radius	_____	_____	_____
9. Distal Ulna	_____	_____	_____
10. Metacarpal I	_____	_____	_____
11. Metacarpal II	_____	_____	_____
12. Metacarpal III	_____	_____	_____
13. Metacarpal IV	_____	_____	_____
14. Metacarpal V	_____	_____	_____
15. Proximal Phalange I	_____	_____	_____
16. Proximal Phalange II	_____	_____	_____
17. Proximal Phalange III	_____	_____	_____
18. Proximal Phalange IV	_____	_____	_____
19. Proximal Phalange V	_____	_____	_____
20. Middle Phalange I	_____	_____	_____
21. Middle Phalange II	_____	_____	_____
22. Middle Phalange IV	_____	_____	_____
23. Middle Phalange V	_____	_____	_____
24. Distal Phalange I	_____	_____	_____
25. Distal Phalange II	_____	_____	_____
26. Distal Phalange III	_____	_____	_____
27. Distal Phalange IV	_____	_____	_____
28. Distal Phalange V	_____	_____	_____
Total	_____	_____	_____
SKELETAL AGE = Total divided by 28 =			_____

Features that were only present in the bones of particular subjects were excluded, also absolute size was always ignored. The exact number of stages for each bone was chosen so that differences between consecutive stages were neither so small as to cause confusion and rating error, nor so large that significant information was lost.

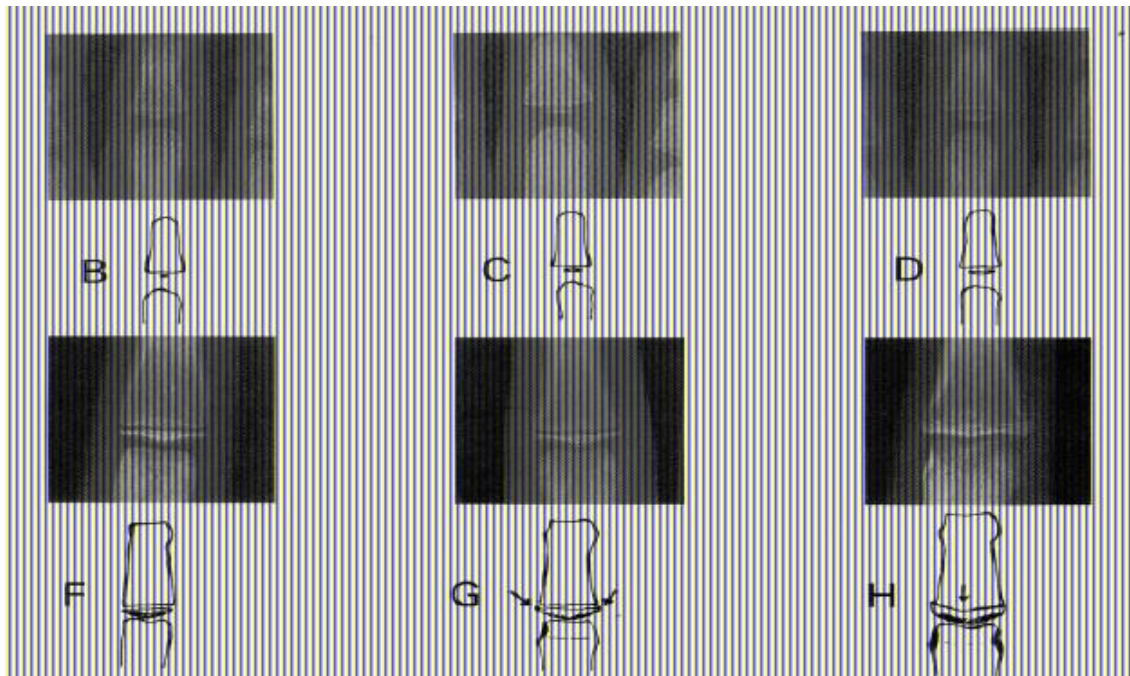


Figure 2.5 This is an example of the images are used to find the stage (B through I) of a given ROI (in this case the middle phalanx of the third finger) in the hand. When the stage is known the ROI can be scored. The development of the epiphysis is clearly shown. In our software project our first aim will be to find the stage of middle phalanx of the third finger in the hand.

A score had to be assigned to each bone in the hand for each stage. Many of the bones in the hand and wrist are giving approximately the same information about maturity. This is especially true for the bones in the fingers. Therefore it is not desirable to just take an average of all the scores in a hand since this would give the nineteen finger bones together a much greater importance than the seven carpals or the radius and the ulna. A system of weighted scores was developed in order to correct for this. Also, the scores of boys and girls are different because girls mature faster than boys. Especially in older children the carpal bones do not provide any useful information about the skeletal maturity. To allow the exclusion of the carpals or the specific examination of them, three separate scoring systems were developed:

- TW2 20-bone
- RUS
- Carpal

The first system is the most extensive one featuring 20 bones of the hand. These bones include bones of the wrist, the carpal bones, and bones of the thumb and the third and fifth fingers. The second system named RUS (Radius, Ulna and Short finger bones) contains the same bones as the first one except for the carpals. Finally the third system contains only the carpal bones. Using the TW2 method is relatively straightforward; first one chooses a suitable scoring system (TW2, RUS or carpal). A suitable scoring system is mostly determined by the calendar age of the patient.

Next one looks up the bones in the radiograph that are associated with that particular system. Those particular bones are then compared to a series (see Figure 2.5) of reference images in the TW2 book. These reference images each represent one developmental stage (B-I) and are backed up by textual descriptions. In these descriptions of a stage, important changes in the bones are described. Also, the descriptions often contain requirements that a bone must meet in order to be classified a certain stage.

The reference image that most resembles the bone from the radiograph is chosen. Of course the bone must also meet the requirements in the description. The stage corresponding to a reference image is assigned to the bone. When the stage of a certain bone is known the score associated with that particular bone and stage can be looked up. After all the required bones have been analyzed and all the needed scores have been gathered, the total score is determined by adding all the separate scores. The total score is linked to a certain skeletal age via a conversion table (see Table5.1).

3. COMPUTERIZED BONE AGE ANALYSIS

3.1 Introduction

In the previous chapter we reviewed two commonly used bone age analysis method. The most appropriate clinical method should be chosen to build an effective computerized bone age assessment system. After choosing the method, we will be able to analyze the problem further and suggest steps that will lead to a full automatic bone age assessment system.

3.2 Choosing the Most Suitable Clinical Method

The clinical method that we will choose is very important. So we have to examine the clinical methods that would be most appropriate from a computer vision point of view.

The Greulich and Pyle method depends on a global overview of the entire hand radiograph. In this method every bone in the hand needs to be compared to the bones in the atlas image. But some bones in the hand are harder to find than others so it became a disadvantage when we tried to automate because of the limited information in some of the bones in the hand. A method in which it would be possible to just examine certain well defined areas in the hand is preferred.

The Tanner and Whitehouse method has a modular structure. In this method we do not compare a general overview of the hand so it is easier to use, from a computer vision point of view. All the ROI of bones in the hand has a specific score so we can do a number of local image analyses (of the ROI) and combine their outcomes into a global skeletal age score. The Tanner and Whitehouse method has different scoring systems so the method is more flexible and adapts easily to a variety of circumstances. Negative aspects of this method are: It is slower and harder to use than other methods. This can be solved with computer analysis. To sum up the Tanner and Whitehouse method is the most suitable method for this project.

3.3 Image Processing Steps In Detail

Now that we have chosen a clinical method to underlay our approach we are able to formulate a step by step plan. Each of these steps must be solved to attain our final goal of fully automatic bone age assessment system.

First of all, we should normalize radiographs before we start any other image processing step. All images have to be in the same condition. There are several characteristics of radiographs that can be vary such as; position of hand should be aligned for each image markers and labels should be removed from image.

In pre-processing steps we should remove markers and labels and find hand position. After finding hand position we should find the ROIs in image. All ROIs in the TW2 method should be found.

After finding ROIs we can assign a TW2 stage for each one of the ROI. There are two type of region of interest. First is EMROI, epiphyseal – metaphyseal ROI, that is the ROI between phalanges of fingers, thumb and above the metacarpals. Second type is CROI, carpal ROIs, which are the ROI of wrist bones. To calculate bone age all the ROIs have to be assigned to a TW2 stage. And each of these two types of ROIs needs different approach.

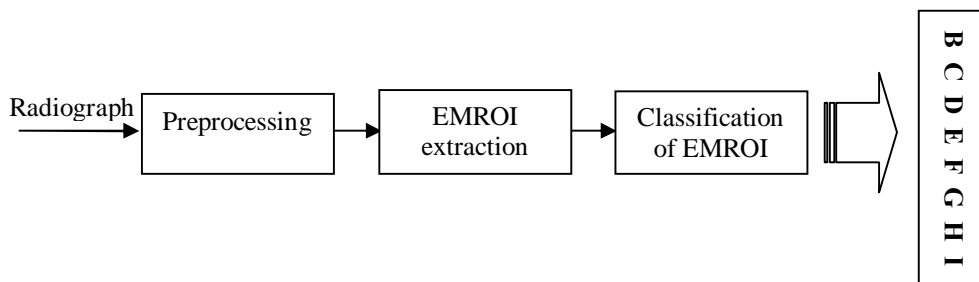


Figure 3.1 Steps of fully automated bone age analysis system.

Below is a list of steps that we are planed in order to develop a fully automated hand analyzing system.

- Pre-Processing:
- Correct orientation of hand on the radiograph
- Find (the contour) position of hand in the radiograph

ROI extraction:

- Find EMROIs and CROIs using the hand's contour
- Find features useful in assigning a TW2 stage to the EMROI and CROI

Practical Usage:

- Build a complete software framework for fully automatic hand analysis.
- Handle missing data because of incomplete analysis.

In this thesis we will focus on assigning a TW 2 stage to the extracted ROI and we will build a software framework for the hand analysis. And we have done some other researches for other steps to contribute future works.

3.3.1 The Orientation of Hand

Position of hand on the radiograph could not be on valid orientation. For initialization of image processing orientation of hand should be in standard (anterioposterior, up-right, left-hand wrist) position. When conventional X-ray units were used, some times to fit on the film hand photographed in a diagonal orientation. With the new digital X-ray machines laboratory worker usually already corrects the orientation of the hand before the radiograph is printed to film. Maas [3] developed a system, which automatically corrects the orientation of the hand and clips out any unexposed borders from the radiograph.

3.3.2 The Position of Hand and Radiograph

After correction of orientation we should locate the hand in a given radiograph. Simple threshold will not work in most cases because the background of a radiograph is never a uniform gray value but there is always a gradient. This is caused by the heel effect (see Behiels [4]). Moreover sometimes there are various markers and labels that are on the radiographs, these markers show patient data and are of no use for the automatic analysis of the radiograph. Another feature commonly found in radiographs is white borders caused by blocking of the radiation by the collimator. It is important to remove the aforementioned artifacts because they can interfere with the automatic and even the manual hand analysis. Maas [3] is researching whether Active Shape Models can be used to

segment the hand from the background. He uses the following technique: a training set of already segmented hand radiographs (trained by segmenting with hand) is used to teach the ASM algorithm to automatically segment hand contours from a radiograph. Then the area outside the hand contour is made a constant, uniform color (black) thereby removing all artifacts outside of the hand.

3.3.3 Finding CROI and EMROI

In TW2 method we have to find specific ROIs to assign a maturity stage. There are two groups of ROIs; EMROI and CROI. So we have to use two different approaches for them. CROI only provides valid information up to the age of 9 years in girls and 12 years in boys after that age some changes on the carpal bones can not be seen they overlap by the time. Therefore we decided not to find and analyze the CROI in this thesis. The TW2 method can be split into three separate types of analysis:

- Full TW2 analysis (20 ROI are examined)
- RUS analysis (13 ROI are examined, no carpal bones are involved)
- Carpal analysis (only the carpal bones are examined)

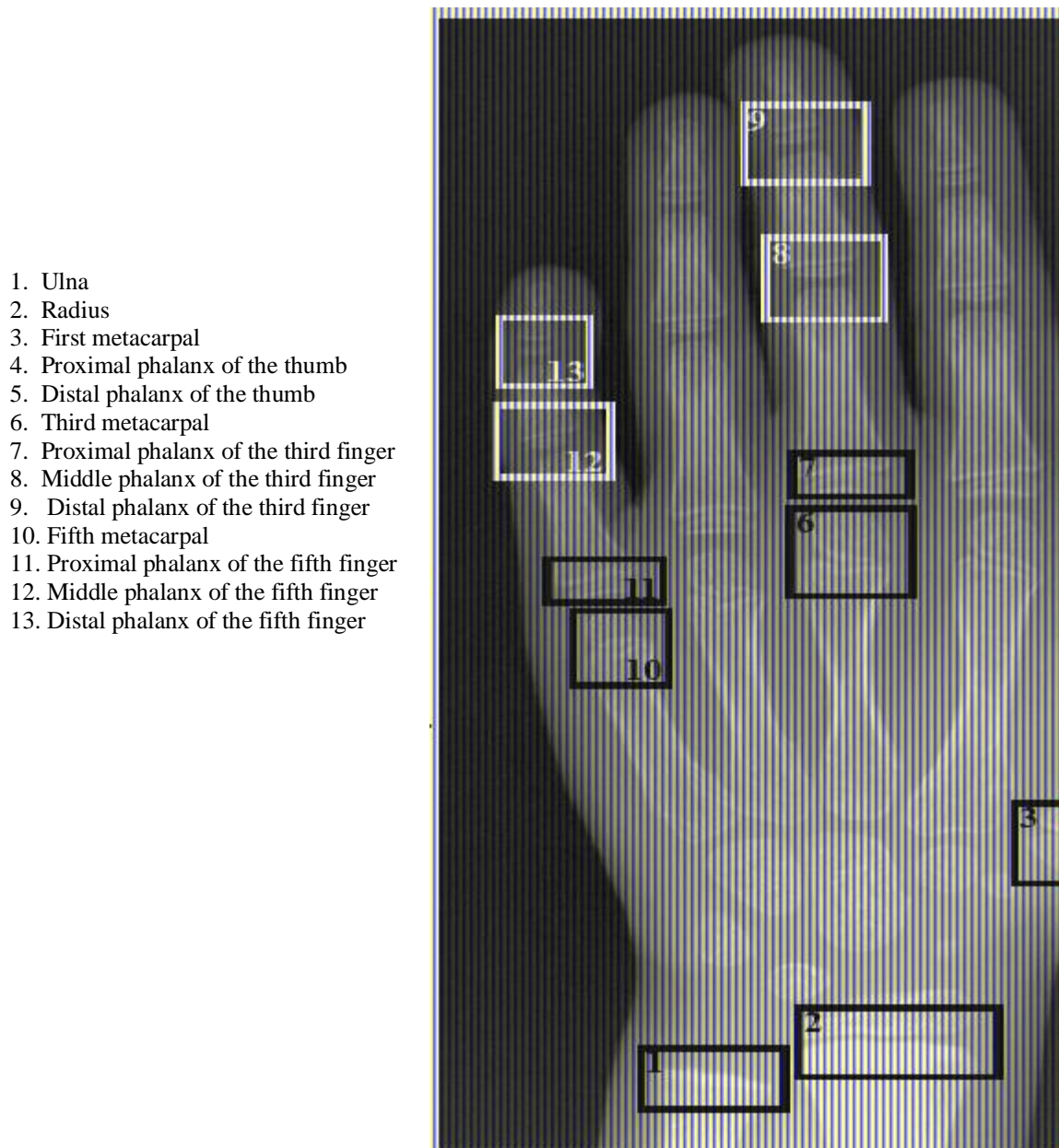
We use the second type of analysis, which avoids having to analyze the carpal bones. What we are left with is the analysis of the bones shown in Figure 3.2. These bones can be grouped into three classes, one containing the radius and the ulna (1 and 2), one containing the EMROI of the thumb (3, 6 and 11), and finally the rest of the EMROI in the fingers themselves (4, 5, 7, 8, 9, 10, 12 and 13).

Each of these groups will require a different approach in finding and analyzing them. Most of the studies that have been done by other groups focus on the bones in the fingers but not in the thumb. Because the thumb does not usually have the same orientation as the rest of the fingers and because the thumb is lying on its side its position and rotation are not as well defined as those of the rest of the fingers.

3.3.4 Feature Extraction from the CROI and EMROI

The anatomy of a typical ROI can be seen in Figure 2.3. We will focus on epiphysis and the metaphysis that are the most important indicators of the stage (and skeletal maturity) of a ROI. In the early stages the size and shape of the epiphysis is the most

important indicator. In later stages the amount of fusion between the metaphysis and epiphysis is more important.



1. Ulna
2. Radius
3. First metacarpal
4. Proximal phalanx of the thumb
5. Distal phalanx of the thumb
6. Third metacarpal
7. Proximal phalanx of the third finger
8. Middle phalanx of the third finger
9. Distal phalanx of the third finger
10. Fifth metacarpal
11. Proximal phalanx of the fifth finger
12. Middle phalanx of the fifth finger
13. Distal phalanx of the fifth finger

Figure 3.2 All ROIs those are included in the RUS analysis. On the left the names of each ROI in the image are shown.

Since these two situations are quite different from each other it might be useful to know in advance what stage we can expect to find in the radiograph. In our case we can use the calendar age that is recorded when the radiographs are taken.

TW2 atlas has a textual description of each stage which indicates what the most important measures are for a certain stage. For example:

Table 3.1 Some of the TW2 stage explanation examples.

<p>Stage D: The maximum diameter of the epiphysis is half or more the width of the metaphysis. In this stage the width of the epiphysis in relation to the width of the metaphysis is the most important measure.</p>	<p>Stage E: The central portion of the proximal border has thickened and grown towards the end of the adjacent phalanx, shaping to its trochlear surface.</p>
---	---

Stage E follows stage D, this means that everything that indicates stage D (the width of the epiphysis is half or more the width of the metaphysis) must also hold for stage E. The striking thing here is that to determine whether an EMROI 18 is stage D one has to measure the width of the epiphysis and metaphysis while to find a stage E EMROI one has to look at the shape of the epiphysis. This alternation between size and shape measures is typical in the TW2 method. This means that we cannot just measure the sizes of some of the objects in the EMROI and determine a stage from there. However, we could use these measures as an indicator of the stage and use some other method to check the shape of the epiphysis. Our approach to this problem is detailed in Chapter 6.

4. PREVIOUS WORK

4.1 Introduction

Previous studies will be discussed in this chapter. There are two parts in this chapter; first part on segmentation. The second is on assigning TW2 stages to ROIs.

4.2 Segmentation of the Hand

Most of the studies are about locating EMROI in the fingers. There are not many studies on locating thumb bones. Because finding and analyzing EMROI in finger is the easiest part of bone age assessment.

A semi automatic bone age assessment system was developed by Micheal [5] around 1989, this system was named HANDX. The author claims that the system was able to automatically segment the bones in a hand radiograph but large scale tests were not done. Firstly preprocessing steps applied to images to normalize the image gray scale so that the later segmentation step will be more robust. The program first segments the entire hand from the background using a threshold operation. After this a model-based method is used to find the bones in the hand. This method uses knowledge of the relative positions of the bones in the hand with respect to each other and to the contour of the hand to find the approximate position of a bone. Then its contour is given by an adaptive contour.

Another method developed by Pietka [6] is about to find the third finger and measure the lengths of the distal, middle and proximal phalanx. To find PROI, the ROI which contains all the phalanges and epiphyses, a number of steps are needed: First the lower boundary of the PROI is detected by scanning a horizontal line over the hand image to search for the soft tissue between the first index finger and the thumb. The upper boundary is a horizontal line at the tip of the third finger. After this two vertical lines are scanned from the middle of the hand to both the left and the right boundary of the hand. Each of the lines stop on the last pixel belonging to the hand, now the upper, lower, left and right boundary of the PROI have been defined. To segment the bones and epiphyses from the PROI it is first turned into a gradient image using two Sobel kernels. The result is thresholded using an empirically determined value to find the edges of the bones and

epiphyses. The concentration of pixel values at the end of a phalanx is up to 50% higher than in the central part. Once a window has been determined that contains the epiphysis, a horizontal line is scanned over the window. The location of the smallest intersection with the segmented finger is marked as the line that separates the phalanges. In this fashion the borders between the third distal, middle and proximal phalanges are marked. These can now be measured. This method was tested on 50 pediatric hand CR (computed radiography) images and the results were compared to the measurements of an independent radiologist. The mean difference between these two measurements was 0.02mm with a mean standard error of 0.08mm. One year later this updated method was published by Pietka [7]. In this method the author attempts to determine the ratio between the width of the metaphysis and the width of the epiphysis. These measurements play a role in determining the stage of an EMROI. Finding the epiphysis and metaphysis happens in the same way as in Pietka [6]. After these two objects are found, their width is measured. The width is detected by calculating the derivative along a perpendicular line to the phalangeal axis. Maxima in this function indicate the borders of the epi- and metaphysis. The author reports that in a test with 80 hand CR images, accurate measurements of the width of the metaphysis and epiphysis were found in 83% of the cases.

Efford [8] presents an analysis method that is driven, where possible, by knowledge of the hand/wrist anatomy. First the contour of the hand is segmented from the background manually. Regions of interest like the fingers are detected using curvature analysis of the contour of the segmented hand. Places in which the curvature is high mark the top of the fingers and the location of the inter-digital web. By fitting a polynomial through the inter-digital web locations (the inter-digital web is the piece of skin connecting two fingers at their base) each of the fingers can be isolated. When the fingers have been isolated the phalangeal axis of the finger is determined which can, when lengthened, give an indication of the location of the metacarpal centers. Using these metacarpal centers, the author test a number of deformable models to edge-detect the metacarpals. These include snakes, parametric Fourier models and active shape models. The author suggests using an expert system to automate skeletal age assessment, the written criteria found in the TW2 method should be translated into rules for the expert system. Measurements needed by the expert system could then be obtained directly from the parameters of the deformable models used in the segmentation. This system is not worked out in detail.

Manos [9] developed a segmentation method for the bones in the wrist using region growing and merging. The technique that is used is basically a bottom up approach. This means that the results of each stage of processing yield the data for the next. The method can be subdivided into several stages:

- Preprocessing stage
- Region growing stage
- Region merging stage
- Region merging by fusion of edge and region information
- Region labeling

In the preprocessing stage a Canny edge detector is used to find edges in the image. After these had been found the original image was smoothed with an edge-preserving smoothing filter. In the first region growing stage simple regions are formed by an algorithm that includes neighboring pixels in a region if their grey level difference is less than an empirically determined value. The resulting regions are then merged using a technique which is based on combining three merging scores representing region similarity, size and connectivity into a global merge score. The technique basically encourages regions with the same grey level characteristics to merge, larger regions are preferred over small ones while regions with little common boundary are discouraged to merge. The next stage involves comparing the edges found by the previous steps to region information from a separately edge-detected (from the preprocessing stage) image. Only the edges that are approximately the same in both images are kept. Only 15-20% of the edges that the first stage yielded are now left. The final stage involves the labeling of the remaining stages. Each region is labeled either they were 'bone' or 'background'. Labeling is performed through the application of a set of heuristic rules which are mainly based on the grey level characteristics of radiographs. Limited success in a test of 20 radiographs is reported by the author but no numerical data is presented.

Morris [10] presents an algorithm which uses domain knowledge in combination with a local statistical enhancement process. Because the difference in grey levels between bone and non-bone pixels is not consistent throughout the image simple threshold on the original image will not work. To alleviate this problem a statistical enhancement process

is deployed. The most important result of this process is that the grey level of bone regions throughout the image becomes approximately the same. In the statistically enhanced image the five fingers are found by examining the peaks of a histogram taken over a horizontal line through the center of gravity of the hand. The five highest peaks represent the location of the four fingers and the thumb. Next domain knowledge of the anatomy of the hand is used to find the regions where the phalanges should approximately be. Once these regions have been found the edges of the phalanges in them are enhanced using a Sobel operator. To ensure a closed contour a line following algorithm is used on the edge enhanced phalanges. The author tested the algorithm on six x-rays of hands of children in different age groups and shows a test on the radiograph of a young child.

Mahmoodi [11] describes a method to find the edges and location of the phalanges using Active Shape Models. To determine appropriate starting points for the ASM algorithm first the entire hand is segmented from the background. This is done using a valley seeking algorithm to determine an appropriate intensity threshold from the image histogram. Next the fingers are found by calculating the Euclidian distance between a point in the wrist and every point on the hand contour. This produces a hand contour signature in which the maxima represent the finger tips and the minima the finger bases. Using knowledge of the hand anatomy, seed points are determined for the phalanges in the fingers. The seed points are used to position ASM in the image and these find the contours of the phalanges. An algorithm for the identification of joint space and phalanx margin locations on hand radiographs is presented by Duryea [12]. This algorithm was not specifically designed for use in determining the skeletal age; instead it is focused on measuring the distances between the phalanges. Measuring this distance can help in the diagnosis of Rheumatoid arthritis. However, the ROI in the fingers are the same for both applications. The first step is the manual creation of a mask image marking the location of non anatomical structures like labels. All radiographs were manually turned so that the fingers were pointing to the right and the wrist was on the left side of the image. The algorithm employs a multiple resolution approach so for each radiograph a set of 4 lower resolution images were produced. First a rough outline of the hand is determined by an edge detection routine. Proceeding along the outline of the hand groups of pixels forming vertical lines are tested as potential fingertips. From these the five best candidates are selected. The finger locations are then found by moving in the negative x direction

following the path of the largest signal starting at each fingertip. This produces an array of points along the finger; these are then fitted to a straight line using the least squares method. The gray value profile along this line shows three peaks which mark the space between the phalanges. The author claims a 100% success rate in identifying the digits (54 out of 54 images) and a 99% success rate in determining the joint space locations. However, a number of images were excluded from the test. These included images of hands wearing rings or digits being touched by other objects in the image (e.g.: Label or collimator edge).

Vogelsang [13] presents an algorithm that can automatically find the ROI in a hand radiograph using Active Shape Models (see section 6.2.4). First a Canny edge detector is used to detect the edges of the bones and of the hand. Using the detected edges of the bones starting positions for the ASM are determined. Each bone in the hand has a specific ASM specially trained to detect its shape. Next, these ASM determine the exact positions and shapes of the bones. Results are not presented in the article.

A method to extract the EMROI from a hand radiograph is presented by Pietka [14]. Before the actual extraction of the EMROI takes place, the hand radiograph is first preprocessed. The orientation of the image is corrected so that the hand is always in a standard position. Then the background is removed from the image so only the hand remains. The analysis starts by extraction of the phalangeal tip. Once the tips are found the phalangeal axis is determined by moving downwards through the middle of the phalanges (middle determined by distance to edges of phalanges to left and right). To smooth the resulting line a median filter is used and then a third-order polynomial is fitted to it. The derivative of the pixel values along the central axis can be analyzed and extreme values of the derivative indicate the locations of the EMROI. The author used 200 clinical hand radiographs in a test; in 4% of radiographs the phalangeal tip was not detected. The gap between the epiphysis and metaphysis was not detected in 4% of all cases (24 out of 588).

4.3 Analysis of the EMROI

There has been less research on the determination of the skeletal age once the EMROI have been found. Again all the presented research is ordered by date of publication.

In 1991 Pietka [6] presents a simple method to determine the skeletal age using the length of several phalanges. First the phalanges are segmented using an edge detection algorithm. After this, the lengths of the distal, middle and proximal phalanges are measured. These lengths are looked up in the phalangeal bone length table by Garn [15]. The approximate skeletal age can be read from this table. The conclusion of the author is that the standard phalangeal length table can only be used to give a rough estimation of the skeletal age. The mean difference with the actual skeletal age is 1.5 years. One of the original authors of the TW2 method has also been working on automatically assigning TW2 scores to bones in the hand.

Tanner [16] presents an automatic system for skeletal age assessment called CASAS. The system uses a high-resolution monochrome video camera which captures the radiographic hand images that are put on a specially adapted light box. The user sees a stage analysis template on the screen and has to try to fit a bone to the template using the zoom feature of the camera and by moving the radiograph. Each stage of each RUS-TW2 ROI (see section 2.3.2) has its own template. These templates don't influence the computer rating but are used for placing the ROI in a consistent orientation and position. The digital analysis is done by representing the changes in density from point to point in the part of the radiograph that is displayed by a Fourier transformation. A Fourier series of 64 degrees is fitted to each horizontal line of the shown radiograph, and there is 64 such lines one above the other vertically. The shown radiograph is represented by $64 \times 64 / 4 = 1024$ image coefficients, though most of the information is contained in a smaller (manageable) subset. The average of several reference images of a certain stage can be used to make an average stage image. Using the Fourier coefficients, a new image of a stage can be classified. This is done by comparing the coefficients of the image with the average coefficients characterizing each successive stage and choosing the stage with the minimal average deviation of coefficient values. The results presented by the author show that the CASAS system rates radiographs with a high repeatability and less than one

percent reversal (reversal occurs when the analysis of the same patient later produces a lower skeletal age). However, several drawbacks are also noted by the author. The system is sensitive to ill-positioned hands and it prefers X-rays with superior definition. The program rejects between 5% and 10% of all ratings due to poor matching of the coefficients. A number of later studies further evaluated the program.

A study by Albanese [17] indicates that assessment of a radiograph is still operator dependent because the operator has to decide to keep or reject a result supplied by the system. Misclassification was also at an unacceptable level for some bones like the radius and the ulna according to the author.

Tanner [18] also noted the high misclassification of the radius and ulna (15% were rejected). An evaluation study by Teunenbroek [19] concludes that although the variability of the TW2 ratings produced by CASAS is low, the variability of the results when expressed in skeletal age seem considerable. They therefore recommend skipping the skeletal age conversion step and just use the TW2 maturity score.

Pietka [20] presents a classification algorithm that uses a fuzzy classifier to assign a bone age to the PROI (Phalangeal ROI) in the second, third and fourth finger. The features used by the classifier are automatically extracted from a hand radiograph. There are three types of features that are used in the analysis. The first feature is the length of the distal, middle and proximal phalanges. The second feature is the ratio between the diameters of the epi and metaphysis. The third feature is the amount of fusion between the epi and metaphysis. A fuzzy classifier is then used. After defining a membership function, features are processed yielding a matrix which maps the features to a year of age within a predefined range. The grades of membership are defined as membership function values in the interval [0, 1]. A classification rule based on a max-sum operator is used to process the matrix assessing the skeletal age. The results presented by the author show that a correct extraction of epiphyseal diameter / metaphyseal diameter ratios and phalangeal length has been obtained in 90% percent of all 1080 test cases.

In the later stages of skeletal development the size and shape of the epiphysis becomes less important. The amount of fusion between the epiphysis and metaphysis becomes the more important measure of skeletal age (see also section 3.4.4). In order to evaluate the amount of fusion Pietka [21] introduces the idea of using a wavelet

decomposition analysis as a quantitative measure to estimate the amount of fusion. Four levels of fusion are distinguished: no fusion, early fusion, advanced fusion and fusion completed. The author claims a successful rating of the fusion in 83% of all 90 test cases. Evaluation of the actual amount of fusion has been done by visual inspection. The method is sensitive to under and overexposure of radiographs, if these are thrown out in a preprocessing step the accuracy of the analysis increases to 90%.

In [12] Gross presents the results of a study that indicate simple neural networks could assist radiologists in the assessment of skeletal age. A database of 521 hand radiographs and a total of four features that were calculated from seven measurements were used to train the network. The measurements included the widths of several epiphyses of the proximal phalanx and the lengths of those phalanxes. Also the length of the second metacarpal was included. The parameters were then derived from 4 ratios between several of the linear measurements. All measurements were obtained by hand. The parameters were used to train the neural network. The author claims that the results produced by the neural network can be compared with those of an experienced radiologist who uses a standard pediatric skeletal atlas (Greulich and Pyle [2]).

Marques da Silva [28] presents a method to determine a digital signature for skeletal maturity. The end goal of the presented research is to determine a signature of the physiological process of skeletal maturity in terms of events that can be used to objectively assess bone age. The first step in the method determines the longitudinal axis of the middle finger (LAMF). A scale space technique is used to find the LAMF. When the image is viewed at an appropriate scale the LAMF is visible as a ridge in the image intensity landscape. In the next step the intensity transitions along the LAMF are examined using a gradient operator at a certain scale. The author claims that this can show the presence of the epiphysis and in the later TW2 stages the amount of fusion between epiphysis and metaphysis. By using the distance between the peaks and valleys in the intensity profile the gaps between the phalanges and the lengths of the phalanges can be measured. The author uses the ratio between the length of the proximal phalanx and the length of the gap between this phalanx and the metacarpal as an indicator of skeletal maturity. Some preliminary results are presented.

4.4 Discussion

Much of the studies that have been done into finding the EMROI in the hand has focused on smart thresholding techniques (e.g.: Micheals [5], Mahmoudi [11], Pietka [6] and Morris [10]) to find the outline of the hand. Once the outline has been found a variety of techniques can be used to find the EMROI. However, this thresholding step is not trivial and some researchers skip it altogether and revert to a manual threshold operation such as Efford [8]. The problem with using a thresholding operating on a radiograph is that it is almost never successful for the entire radiograph. Heel effect causes a gradient over the entire radiograph. The effect itself is caused by the fact that not all X-rays emitted from the X-ray tube are perpendicular to the target. The non perpendicular X-rays must travel a larger distance through the target (see Farr [23] and Behiels [24]). Because of this the grey value for a certain substance like bone or flesh is not constant over the entire radiograph. This makes segmentation using thresholding very difficult and less robust.

Another approach is to detect the fingertips of the hand and use these as starting points for the segmentation of the fingers and the phalanges (Duryea [12] and Pietka [14]). Although both authors report excellent results one still is more or less dependent on a well defined border both left and right of the phalanges or finger in order to be able to segment it. Also these methods can be sensitive to disturbances in the background of the radiograph like labels or markers.

The advantage of using some type of deformable model and ASM in particular is that it is less sensitive to local disturbances in the contour of the object. It manages to segment an object even when part of the contour is not well defined because the model "knows" the shape of an object from a training set. Some research has been done using these types of models (e.g.: Mahmoudi [11], Vogels [13] and Efford [8]). Extensive evaluations have not been performed but indications are positive. The research into actually assigning a certain skeletal development stage or score to an EMROI can be subdivided into two approaches. One approach is to find measurements, for example the width or length of certain bones, and relate these measurements to a certain skeletal age. Then one can train a classifier and use it to assign the measurements of new cases to a certain skeletal age or TW2 score (e.g.: Pietka [6] [20] and Gross [22]). A disadvantage of this method is that at a certain point in the skeletal development, structures like epiphyses tend to gradually

disappear because they fuse with the metaphysis. Epiphyses are often used in these measurement based techniques because their width supplies a lot of information about the skeletal development (until fusion starts). So most of these methods only yield accurate results until fusion starts.

Another approach is to use some kind of similarity measure for the general appearance of an EMROI. And use this measure to compare a new instance of an EMROI with a set of averages of each TW2 or skeletal development stage. Tanner [16] demonstrated this with their CASAS system which represents the changes in density from point to point of the part of the radiograph containing the EMROI by a Fourier transformation in two dimensions. Of each TW2 stage the average representation is known. Tanner uses these average representations to compare with new images of EMROI. The results of this system presented by the author are good, but evaluation studies showed some problems. The largest problem with the CASAS system was that a human operator was still involved in the process. This operator had to align each EMROI with a template and if this was not done correctly it had a large impact on the quality of the classification. Because of the precise alignment the analyses still took a lot of time and the operator still needed to know the TW2 system well enough to recognize wrong classifications by the system.

5. WEB BASED HAND ANALYSIS SOFTWARE

5.1 Introduction

Our first aim is to develop a manual bone age assessment system to overcome complicated calculations because TW2 method has so many complex score calculations. We built a web based manual skeletal maturity test web site working with the TW2 clinical method. The second part of our project was to develop the core of the automatic bone age assessment system by using active shape model. So with this software we could compare different tw2 stages by using interfaces of the Mathlab 7.0. In this project we focus on assigning TW2 score to the middle phalanx of third finger. Because of time limitation we exclude the entire hand bones that are necessary for calculation of the bone age. In this chapter we will discuss about important features of fully automatic bone age assessment software.

5.2 User Requirements

When writing software we have to consider end-user's [physicians, (medical) students and researchers] usage. Important software requirements are listed below:

- Software should not consume time of the user and should work with little human interaction.
- The skeletal age software should contain the complete TW2 method.
- Analysis results should be recorded for future reviews.
- Fully automatic bone age assessment software should contain analysis of the all the bones in TW2 method.
- Results should be in a standard format and show all steps of the analysis.
- The interface should be user friendly.

Reducing computation time is very important for practical usage. The final goal would be eliminate all human input to improve speed and accuracy over the manual methods. Software can be used for education, research and for calculation of bone age of the pediatric patients. Results should be in the standard format so it could be available in these environments.

Moreover software should analyze all bones in the hand and should include the carpal bone analysis. Result of analysis should include details of the image processing steps. The physician then can verify each one of the step if needed.

5.3 Design and User Interface

In this project we first designed a web site for manual determination of bone age with the TW2 method. In this web site we showed 20 bones of the hand at 7 or 8 stages. Explanations of stages can be read when the mouse is scrolled over the images. (Figure 5.1) Every one of the bone for both sex has different individual score and sum of them give us bone maturity score. Boys and Girls have different maturity scale. Maturity score can be found from Table 5.1. Bone age of the patient can be calculated with the maturity score. Bone age= $1 + \text{Maturity score} / 10$

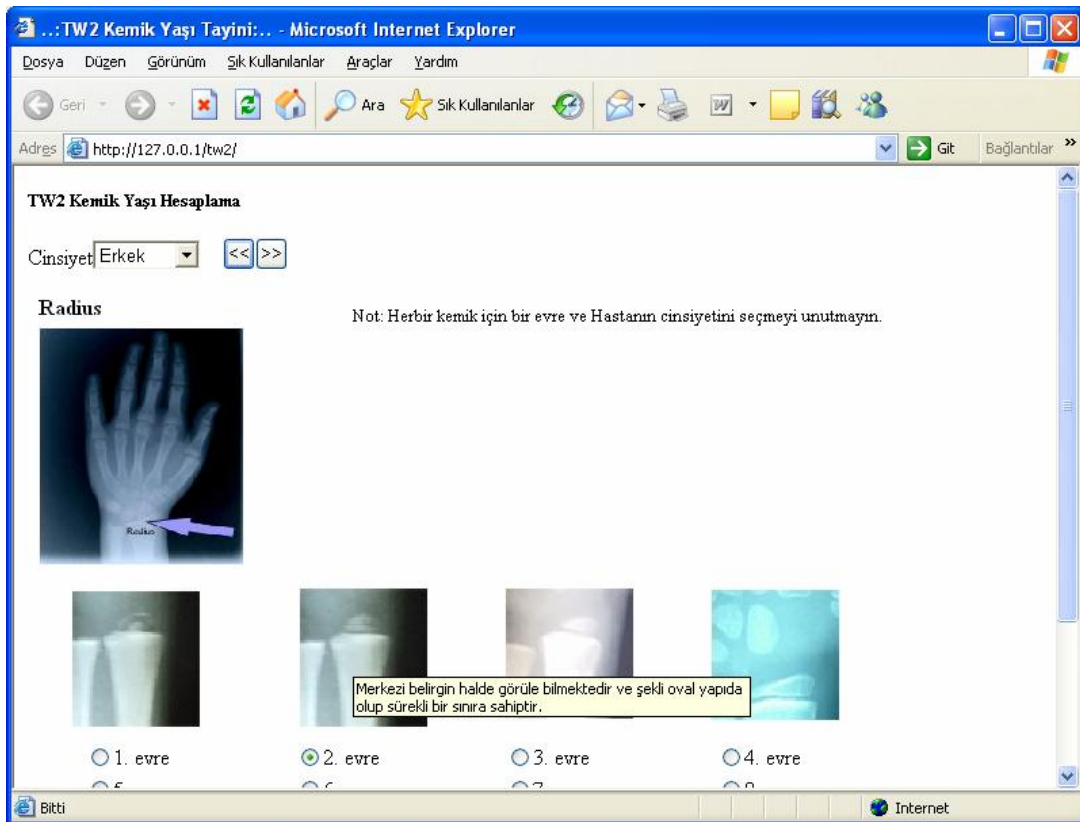


Figure 5.1 User chooses appropriate stage for each bone.

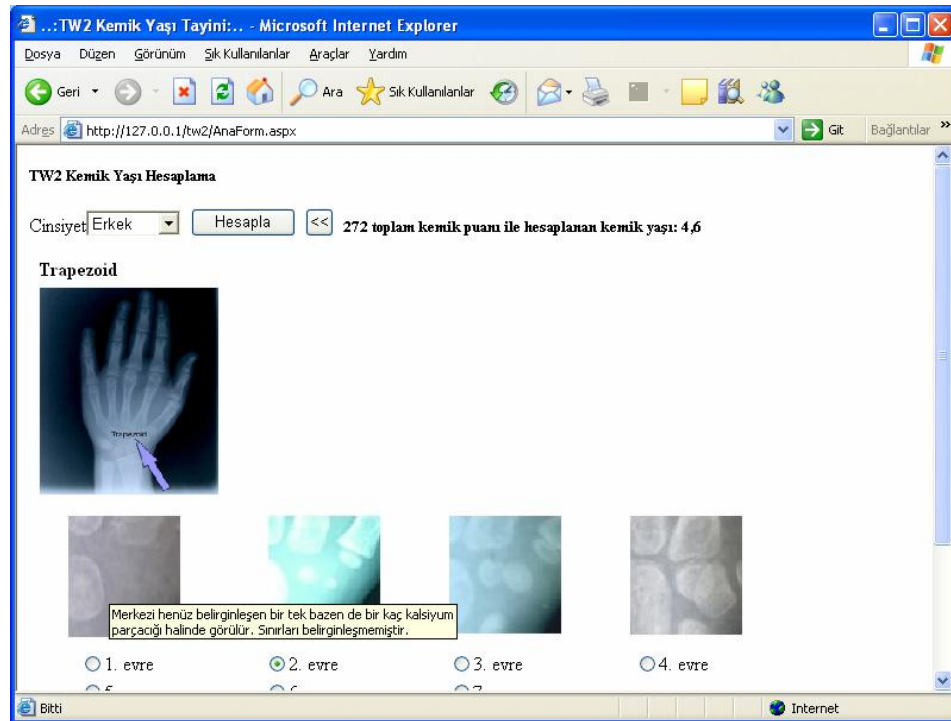


Figure 5.2 Web-Based manual TW2 skeletal age calculation. After choosing stages for each bone tw2 stage calculated.

Table 5.1 User chooses one stage for each one of the bones that correspond one individual maturity score for each one of them. With the sum of the individual maturity scores we can define the maturity score from two different scales one for male and one for female.

Individual bone maturity score			
bone	stage	sex	score
radius	B	M	15
radius	B	F	17
radius	C	M	17
radius	C	F	19
radius	D	M	21
radius	D	F	25
radius	E	M	27
radius	E	F	33
radius	F	M	48
radius	F	F	54
radius	G	M	77
radius	G	F	85
radius	H	M	96
radius	H	F	99
radius	I	M	106
radius	I	F	106
ulna	B	M	22
ulna	B	F	22
...

	Maturity score		
	girls	boys	
0	131	0	114
1	136	1	116
2	140	2	119
3	146	3	123
4	152	4	126
5	159	5	129
6	165	6	133
7	172	7	136
8	179	8	139
9	186	9	142
...
150	1000	170	1000

Second part of our project was developing the automatic bone age analysis software. For this purpose we use ASM framework and we add some modules to assign TW2 stage to the certain region of interest. Programming of this framework was done by the Ghassan Hamarneh from Chalmers University of Technology.

Software has two important stages. First we train images that are in the same TW2 stage for each bone by using landmarks which mark certain physical properties of the bone. Then training sets are saved for future comparison with query image. The second part is the analysis engine. In this part trained images are compared with the query image. Details will be introduced in the chapter 6.

To begin analysis first some of the image sets for which we know the TW2 stage should be trained. During the training stage user enter some parameters:

Table 5.2 Initialization parameters to begin shape training.

- Enter Contours Ending Points. (Last entry is no of landmarks, ex. [8 16 27])
- Enter The number of training set images (2,3,4,..)
- Enter Percentage of explained variance (0-1):
- Enter Number of points along training profile above landmark (1,2,3,...)
- Enter Number of points along training profile below landmark (1,2,3,...)

Then the user selects training landmarks. (Figure 5.3) Same physical locations should be selected in every training image. Then software aligns training sets. And we record training results for future analysis. Training procedure should be done for every one of the bone and every one of the TW2 stage.

After preparing that training sets we can search the image. To begin we need initial parameters (see Figure 5.4). In future initialization would be done automatically. With initial parameters the trained image should be in the same location with the query image. We can scale, rotate and translate trained image to see whether it is fit to the queried image or not.

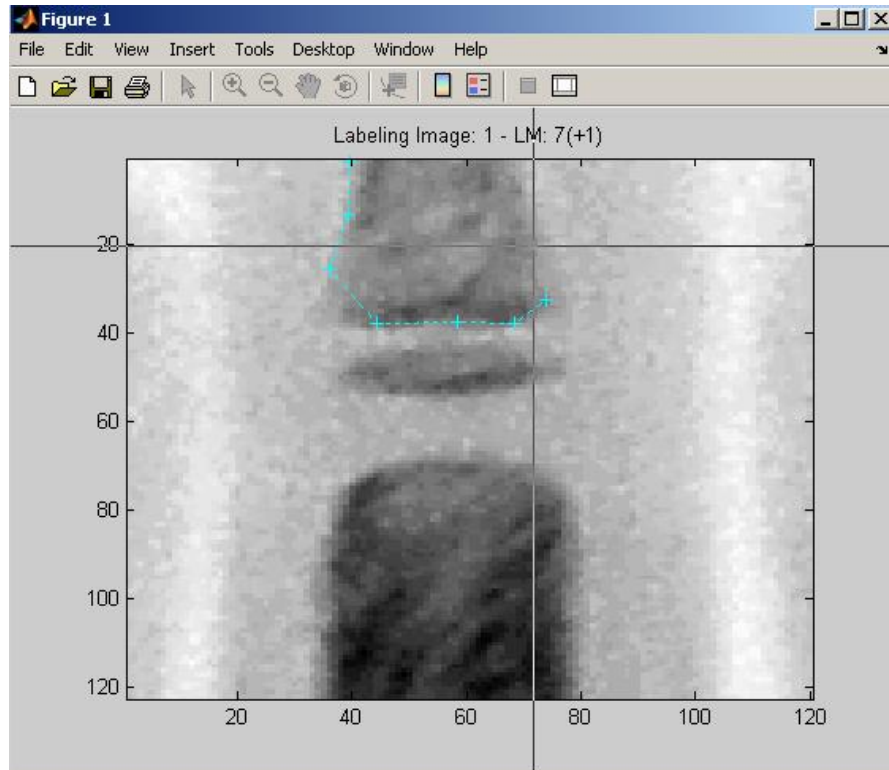


Figure 5.3 User selects landmarks by clicking with mouse and clear by pressing space key

Number of above and below points for search can be defined. And we can choose maximum loop number and multi resolution number (max 2) for different images. After initial parameters are satisfactory for the user, search process can be started. (Figure 5.4)

In the search process software try all trained TW2 stages until queried image is converged with the most closed stage. This procedure should be repeated for the every region of interest. Each bone of the hand will have one individual maturity score and summation of that give us maturity score. So we can calculate the bone age from maturity score.

5.4 Performance and Experimental Results:

Two criteria should be tested in our software, the accuracy in comparison to the manual TW2 method and difference in speed when using the software. In current software we do not examine speed tests because steps such as automatic stage determination and initialization of training sets have not been implemented yet.

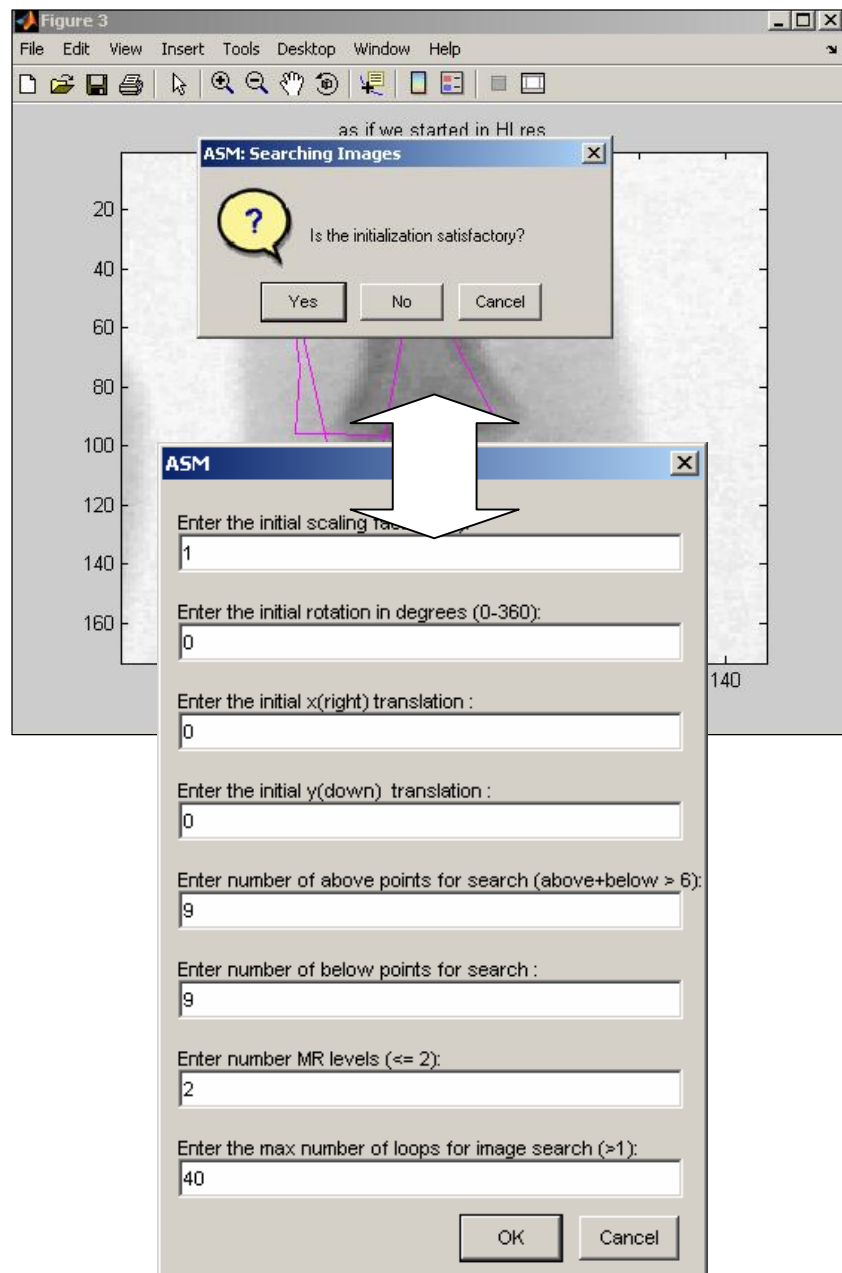


Figure 5.4 Initialization parameters to begin shape search in an image.

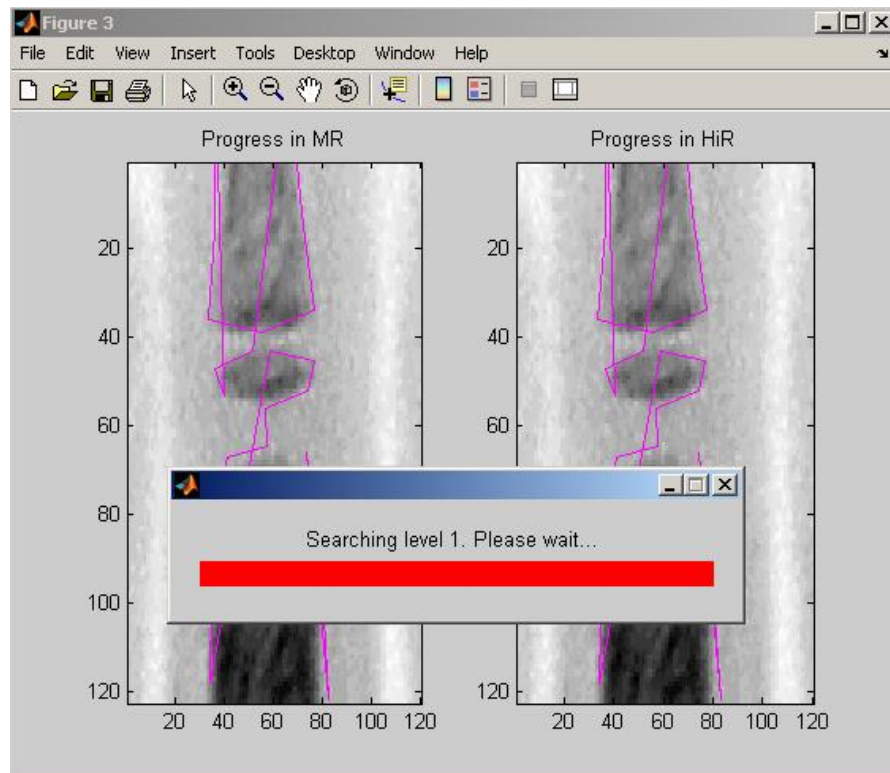


Figure 5.5 Shape search process in different levels.

5.4.1 The Experiment

We tested the performance of web-based TW2 scoring system. Five radiographs have been chosen so that they represent different stages in the skeletal development. Experiments were done by rating each radiograph according to the complete TW2 method. In this way for each radiograph twenty bones were examined.

Table 5.3 performance tests of web based bone age assessment system

Patient No	Bone age with Manuel method	Web-based TW2 Bone age calculation
11	11,6	11,1
13	8,5	8,0
15	9,5	9,9
4	7,8	7,2
6	4,8	4,4

Tanner [16] states that the manual method shows a 30% or more between-observer difference rate. On average 33.8% of the bones rated by an observer were rated differently

by another observer. This number is comparable to the 30% when using the manual method and therefore we can conclude that at least in this aspect our program performs approximately the same as the manual method.

5.4.2 Future Developments of Application

Second objective of this project is to develop software for fully automatic bone age assessment. In this chapter we will discuss new beneficial functions for the future development.

There are several projects where the software can be deployed. In hospitals results of the examinations can be stored in a database so it will be easy to analyze and access. Another possibility would be educational use. The program provides a lot of information for each stage of the TW2 method.

One of the most important features which still need to be implemented is the ability to automatically assign a stage to an ROI. While the program is still semi automatic it would be beneficial if the system could locate the ROI in the hand automatically. Manual selection of the ROI takes some time and effort for the user. In the current situation the user needs to locate each ROI in the hand.

At the moment, it is only possible to rate an ROI as one of the eight TW2 stages. This is quite limiting and it is basically a remnant from the manual TW2 method. It should become possible to rate the ROI on a continuous scale instead of the current discrete one.

As a result of our work we developed software frame work for the analysis of hand radiographs. Software can successfully calculate bone age but some of the time saving features for the software need to be implemented:

Preprocessing

- Locate the ROI in the hand automatically
- Locate the initialization of the search automatically
- Access results from the database

6. AUTOMATIC ASSIGNMENT OF A TW2 STAGE TO THE MIDDLE PHALANX OF THE THIRD FINGER

6.1 Introduction

To build the kernel of the fully automatic bone age assessment it is very important to develop software which automatically assigns stages to EMROI. In this chapter we will analyze how to assign TW2 stages to the middle phalanx of the third finger by using the Active Shape model. I will restrict myself to just the middle phalange of the third finger because of time restrictions and its clear characterization in almost all of our test radiographs. The method should be applicable to other EMROI.

The first section of the chapter is about deformable models; Active Shape Models (see section 6.2.1) in particular. The second part of the chapter contains the outline of our automatic method.

6.2 Method: Deformable Models

In recent years the model-based approach towards image interpretation named deformable models has proven very successful. This is especially true for images which contain objects with large variability like the EMROI we are studying. Among the earliest and most well known deformable models is the Active Contour Model, also known as Snakes. Snakes are model-based methods for localization and tracking of the image features.

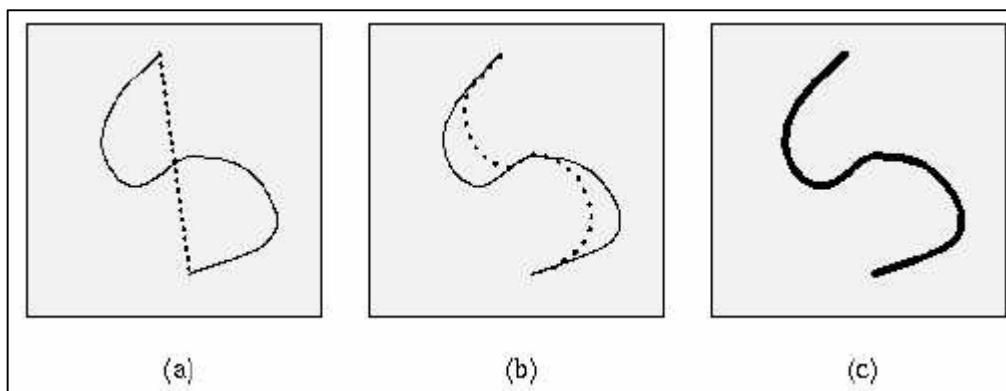


Figure 6.1 Active Contour Model-snake: a) Initial snake position (dotted) interactively defined near the true contour. b), c): Iteration steps of the snake energy minimization; the snake is pulled toward the true contour.

Snakes can be matched to an image contour by energy minimization. The energy of the snake depends on its shape and location within the image. From a given starting point it deforms itself to conform to the nearest salient contour [25]. In order to improve the specificity of deformable models many attempts at adding a priori knowledge into them have been carried out. A more general approach called Active Shape Models (ASM) that still preserved specificity was proposed by Cootes [26]. ASM are statistical models of the shapes of objects which iteratively deform to fit to an example of the object in a new image. The shapes are constrained by a statistical shape model (see section 6.2.1) to vary only in ways seen in a training set of labeled examples.

In addition to the shape model, we require models of the image appearance around each point in the model. The simplest is to assume the points lie on edges in the image. More complex models can be built to represent the statistical variation of the gradient along profiles through the points, normal to the boundary curve at that point. In this work we used active shape model for assigning TW2 stages to EMROI.

6.2.1 Statistical Shape Models

One important aspect of ASM is the statistical model of shape which is used to represent objects in images. The shape of an object is represented by a set of n points, which may be in any dimension but we will focus on the two dimensional case (points in a two dimensional space). Shape is usually defined as that quality of a configuration of points that is invariant under some transformation. In two and three dimensions usually the Similarity transformation is considered. Translation, rotation and scaling don't change the shape of an object. Statistical shape models allow both analysis of new shapes and synthesis of shapes that are similar to those in a training set. Manual segmentation yields the training set. Next, the variations in shape of the object over the training set are analyzed, and then a model can be built which can mimic these variations.

Statistical shape models are a very broad subject. Here we will briefly review only those parts of the statistical shape model theory that relate directly to the project. These parts are the choice of landmarks for training the shape model, Procrustes analysis and some of the theory behind ASMs. For further information I refer to Cootes [26].

6.2.2 Landmarks

When we want to define the shape of an object in an image we use landmarks to delineate its contour. The best positions for these points are locations which can consistently be found from one image to the next. As an example consider, Figure 6.1 in which you can see a choice of landmarks for the segmentation of the hand from a radiograph. Each landmark has been placed so as to be easily locatable in all images. We call these landmarks fixed points. In two dimensions, fixed points should be placed at clear corners of object boundaries or easily located biological landmarks.

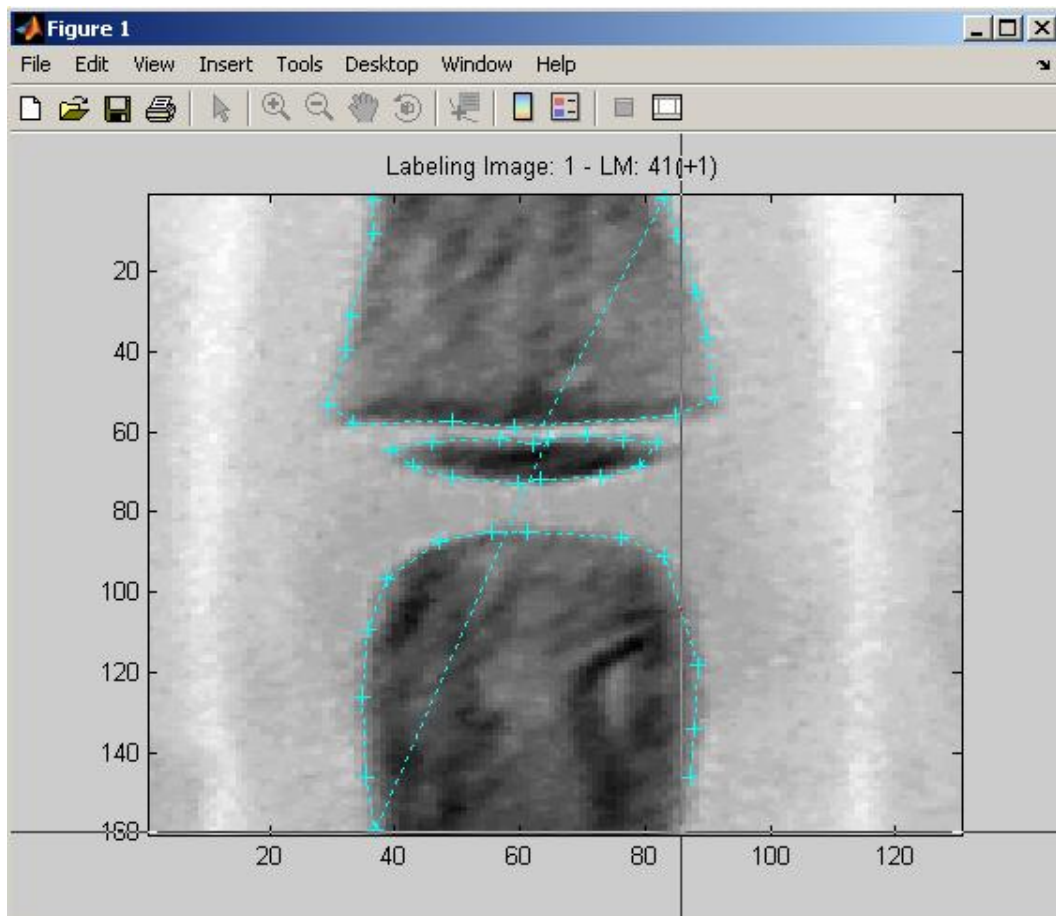


Figure 6.2 Landmarks choose for training sets which easily locatable in all images. Landmarks have been placed in a position of high curvature in the shape which is easily found across the training images. The dots are fixed points; the points between the dots are points on the contour of the hand indicated during manual segmentation.

In a two dimensional image we can represent the n points, $\{(x_i, y_i)\}$, that make up a shape as a $2n$ element vector, \mathbf{x} , where each of the aligned training shapes gave rise to a vector describing the n boundary points:

$$\mathbf{x}_i = (x_{i,0} y_{i,0}, x_{i,1} y_{i,1}, \dots, x_{i,n-1} y_{i,n-1})^T \quad i=1, \dots, N \quad (6.1)$$

Given s training examples, we generate s of these vectors \mathbf{x}_j .

6.2.3 Modeling Shape Variation

We now have a set of shapes, \mathbf{x} these forms a distribution of vectors in nd dimensional space (n =number of points in shape and d =number of dimensions).

When we model this distribution we can generate new examples that are similar to those in the original training set. We seek a model of the form $\mathbf{x} = \mathbf{M}(\mathbf{b})$. New vectors, \mathbf{x} can be generated using this model. In order to simplify the problem we lower the dimensionality using Principal Component Analysis. This allows approximation of any of the original points using a model with fewer than nd parameters. We do this by first calculating the mean shape of the N shapes.

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad (6.2)$$

The modes of variations, i.e.: the ways in which the points of the shape tend to move together, can be found by applying a principal component analysis to the deviations, $\Delta \mathbf{x}_i = \mathbf{x}_i - \bar{\mathbf{x}}$ ($i=1, \dots, N$) from the mean. From these deviations, the $2n$ by $2n$ covariance matrix can be calculated:

$$\mathbf{S} = \frac{1}{N} \sum_{i=1}^N \Delta \mathbf{x}_i \Delta \mathbf{x}_i^T \quad (6.3)$$

$$\mathbf{S} = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})^T \quad (6.4)$$

The modes of variation of the point of the shape can be described by the $2n$ unit eigen-vectors, $\mathbf{p}_1, \dots, \mathbf{p}_{2n}$, of \mathbf{S} . The eigenvectors are defined by

$$\mathbf{S}\mathbf{p}_k = \lambda_k \mathbf{p}_k \quad \text{and} \quad \mathbf{p}_k^T \mathbf{p}_k = 1 \quad (k = 1, \dots, 2n) \quad (6.5)$$

Where $\lambda_1, \dots, \lambda_{2n}$ are the corresponding $2n$ eigenvalues of \mathbf{S} ($\lambda_1 \geq \lambda_2 \dots \geq \lambda_{2n}$).

The k th principal component corresponding to the vector \mathbf{x}_i is defined as a weighted sum of the elements of this vector:

$$b_{i,k} = \mathbf{p}_k^T (\mathbf{x}_i - \bar{\mathbf{x}}), \quad i = 1, \dots, N, \quad k = 1, \dots, 2n \quad (6.6)$$

The principal components represent a linear independent decomposition of the variation of the training shapes. The first principal component, which is associated with the largest eigenvalues, λ_1 , describes the largest part of the shape variation (the first mode of variation). In fact, the proportion of the total shape variance described by the k th principal component is equal to the λ_k . Most of the variation can usually be represented by a small number of principal components, say t ($t < 2n$). The value of t can, for instance, be chosen in such a way that the first t principal components explain a sufficiently large proportion (e.g. 99%) of the total variance, $\lambda_T = \sum_k \lambda_k$, of the shapes. Any shape in the aligned training set can be approximated as a sum of the mean shape and a weighted sum of the first t eigenvectors:

$$\mathbf{x}_i \cong \bar{\mathbf{x}} + \mathbf{P}\mathbf{b}_i, \quad i = 1, \dots, N \quad (6.7)$$

Where $\mathbf{P}=(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_t)$ is a matrix of the first t eigenvectors, and $\mathbf{b}_i = (b_{i,1} b_{i,2} \dots b_{i,t})^T$ is a vector of weights (principal components), which is calculated as

$$\mathbf{b}_i = \mathbf{P}^T (\mathbf{x}_i - \bar{\mathbf{x}}), \quad i = 1, \dots, N \quad (6.8)$$

Eqn. (1) permits generation of a new shape example by replacing \mathbf{b}_i by a new vector of weight values, $\mathbf{b} = (b_1 b_2 \dots b_t)^T$. The vector \mathbf{b} defines a set of parameters of a deformable model. Shape \mathbf{x} can be varied by changing the elements of \mathbf{b} . The elements of \mathbf{b} are constrained to lie in the range $\pm m \sqrt{\lambda_i}$ where m usually has a value between 2 and 3. Provided that the weights are not too far from zero, the new synthetic example will be similar to those in the training set, as the change in shape will be determined by the modes of variation represented by the training shapes. Suitable limits for the weight vector, \mathbf{b} , are derived by examining the distributions of the weight values to generate the training set. If

Gaussian distributions are assumed, the set of weights can be chosen, so that the Mahalanobis distance, D_m^2 from the mean shape is less than a suitable value, D_{\max}^2 :

$$D_m^2 = \sum_{h=1}^t \left(\frac{b_h^2}{\lambda_h} \right) \leq D_{\max}^2 \quad (6.9)$$

The number t of eigenvectors to retain can be chosen so that the model explains some proportion f_v (e.g.: 98%) of the variance of the training set. λ_i are the eigenvalues of the covariance matrix of the training data. The total variance in the training data is given by the sum of all the eigenvalues. $V_T = \sum \lambda_i$. The desired number of modes is given by the smallest t for which the following holds:

$$\sum_{i=1}^t \lambda_i \leq f_v V_T \quad (6.10)$$

An example showing the first two modes of a shape model is given in Figure 6.2.

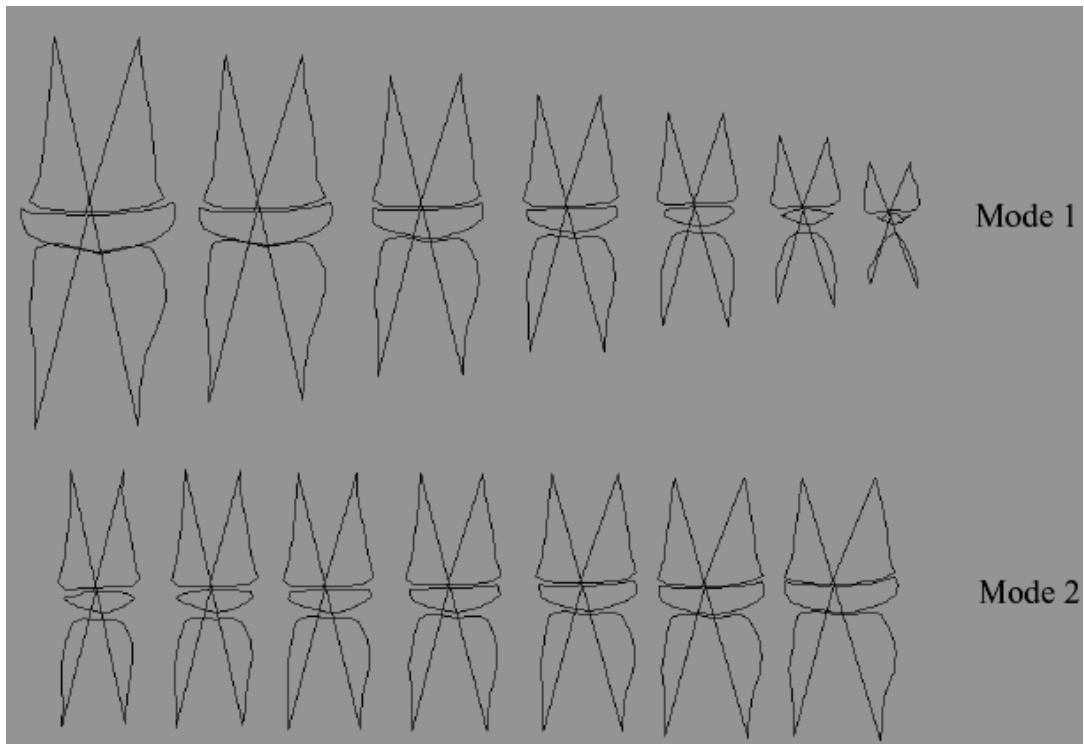


Figure 6.3 The first two modes of a shape model of the EMROI in the hand are shown here. The shape parameters are varied in turn from left to right from -3 till 3 standard deviations. Since this shape model was built from images that were not scaled the first mode represents the size differences occurring in the training set. The development of the epiphysis is clearly visible, this is logical because the size of an EMROI is usually directly related to its developmental stage.

6.3 Active Shape Models

When fitting a model to an image an iterative approach proposed by Cootes [5] is used to improve the fit of an instance X of the model to the image.

1. A region around each point X_i is examined to find the best nearby match for the point X'_i
2. Update the segmentation to best fit the new found points X (see formula 6.5)
3. Repeat until convergence (The software we use uses a fixed number of iterations)

The pixels that are examined around the points are on a line perpendicular to the contour (see Figure 6.4). This does require a notion of connectivity between the points from which the perpendicular direction can be computed. When we want to calculate the direction perpendicular to a point (x_n, y_n) we rotate the vector that runs from (x_{n-1}, y_{n-1}) to (x_{n+1}, y_{n+1}) over 90° . All contours are supposed to be closed contours in the implementation we use, so for the first point on a shape, the last point and the second point are used; for the last point, the one but last landmark and the first landmark are used.

Now the pixel profile along the perpendicular line is used as an appearance model that describes the typical image structure around the points of the shape. The pixel profiles which are made up of k pixels on either side are sampled using a fixed step size.

The profiles thus have a length of $2k + 1$. Cootes proposes using the normalized first derivatives of the profiles to build the appearance model. The profiles are normalized such that the sum of the absolute values of the elements in the derivative profile is 1. For each image in the training set the procedure is repeated to get a set of normalized samples $\{g_i\}$ for each model point.

Now the mean \bar{g} and covariance S_g are calculated for each set of samples. We end up with statistical models of the gray-level profile about each point. The Mahalanobis distance between a new profile g_i and the profile model can now be calculated using:

$$f(g_i) = (g_i - \bar{g})S_g^{-1}(g_i - \bar{g}) \quad (6.7)$$

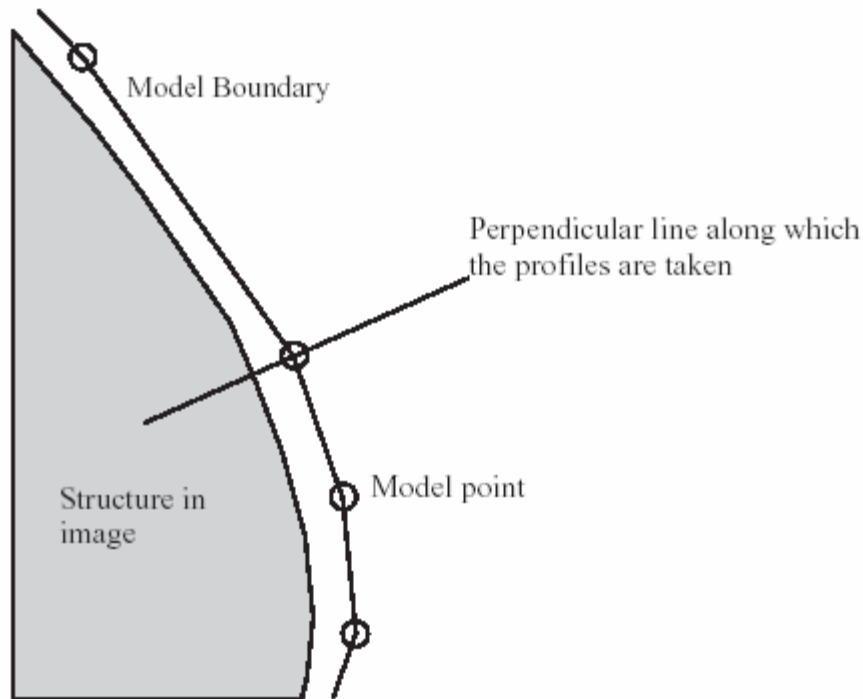


Figure 6.4 A line perpendicular to the model boundary is taken through each point.

When the Mahalanobis distance $f(g_i)$ is minimized the probability that originates g_i from the original profile is maximized. The profile models given by \bar{g} and S_g are calculated for multiple resolutions. L_{\max} denotes the number of resolutions. The original image is the finest resolution and a step size of 1 pixel when sampling the profiles. The next resolution is the image observers at scale $\sigma = 1$ and a step size of 2 pixels. Lower resolutions are constructed by doubling the image scale and the step size. The software we used does not subsample the images as proposed by Cootes. When we look again at the iterative approach detailed at the beginning of this section we can now see how we can determine a new best match for a point.

Each model point is moved along the direction perpendicular to the contour to ns positions on either side, so total of $2ns+1$ positions are evaluated. The model point is put at the position with the lowest Mahalanobis distance. When all the model points have been moved, the shape model is fitted to the new set of points and this yields an updated segmentation. This is repeated a fixed number of times at each resolution, from coarse to

fine. Cootes et al use a convergence criterion to move to a finer resolution if convergence is reached before the fixed number of iterations has been made.

6.4 Method: Procrustes Analysis

Using Procrustes analysis it is possible to align a set of n shapes so that the sum of distances of each shape to the mean shape ($D = \sum |x_i - \bar{x}|$) is minimized. Constraints must be placed on the alignment of the mean (e.g. centered on the origin, has unit scale and a fixed, arbitrary rotation). There are analytic solutions for this problem but we employ the iterative approach given by Cootes in [5].

1. Translate each shape so that its center of gravity is at the origin.
2. Choose one example as an initial estimate of the mean shape and scale so that $|\bar{x}| = 1$
3. Record the first estimate as \bar{x}_0 to define the default reference frame.
4. Align all the shapes with the current estimate of the mean shape.
5. Re-estimate mean from aligned shapes.
6. Apply constraints on the current estimate of the mean by aligning it with \bar{x}_0 and scaling so that $|\bar{x}| = 1$.
7. If not converged, return to 4.

(Convergence is declared if the estimate of the mean does not change significantly after iteration)

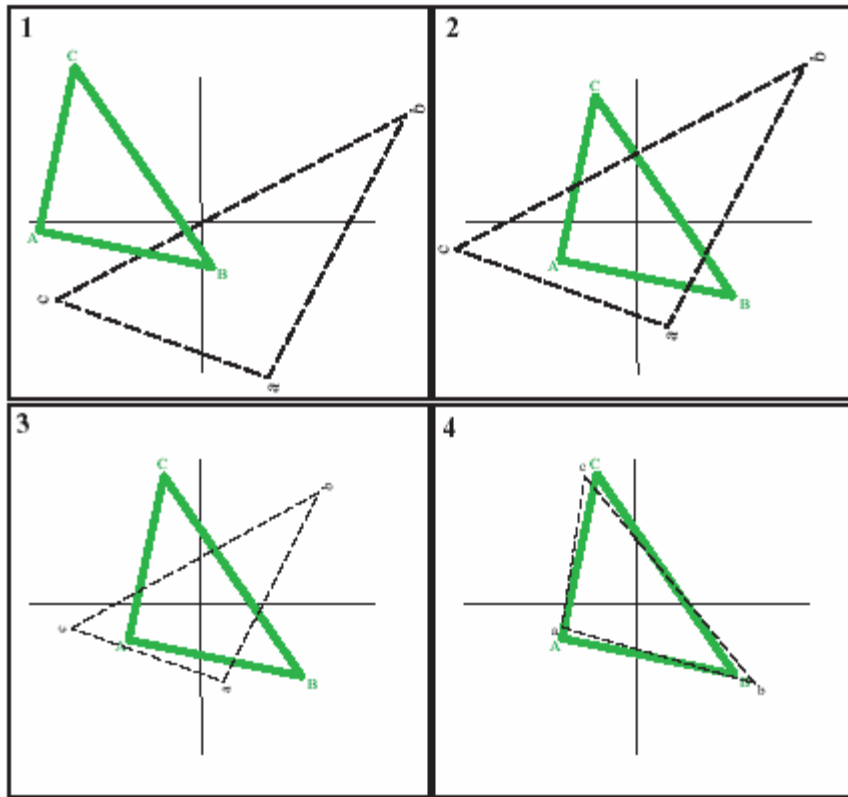


Figure 6.5 In this image the Procrustes analysis being applied to two triangular shapes is shown. The thick shape (ABC) is the shape we want the dashed shape (abc) aligned to. In 1 we see their respective starting positions. First we translate them so that they each have the same centroid (2). After this they are both scaled so that $|\bar{x}| = 1$ (3) and finally rotated so as to minimize D (4).

During the alignment we can allow certain operations, this will affect the shape of the final distribution. An approach that is commonly used is the one shown in Figure 6.5. The scaling constraint in that method means that the aligned shapes x lie on a hypersphere, which can introduce non-linearities if large shape changes occur. Therefore Cootes et al add another operation, a projection into tangent space. This operation preserves the linear nature of the shape variation.

The simplest way to achieve this is to apply the method from Figure 6.4 and add the projection into tangent space as a final step. The actual projection can be done by scaling x by $\frac{1}{|x - \bar{x}|}$.

6.5 Outline of the Automatic Method

Most of the research done on the subject of automatically assigning TW2 stages to EMROI depended mainly on measurements of the bones in the hand (see section 4). We wanted to see how a method based on the appearance of the EMROI would perform. The method we devised is to use mean images of each TW2 stage to determine the stage of a query EMROI. We use an ASM model to segment the EMROI from query images and once the segmentation has been completed we use the shape information it provides to align the query image with the mean image of each stage. Then the correlation is measured between a part of the mean and a part of the query image. The results of these measurements are then used to indicate a stage of the query EMROI. Below we will discuss each part of our method in more detail.

6.5.1 Producing Mean Images for Each Stage

The mean images for each stage are produced from a series of clipped out EMROI that have been rated as the same stage. The cutting of the EMROI is done manually for now but in the future these should be clipped out automatically. Each of these EMROI is segmented by hand so as to get a "perfect" segmentation. Once we have a certain number of EMROI of a particular stage segmented we can use Procrustes analysis to align all the segmented shapes in the images to the mean shape. Then we use an orthogonal projection, to project the mean image onto a plane. What this basically does is it adds up all the intensity values of a certain pixel in all the images and the result value is the pixel value in the mean image (see Figure 6.6).

The resulting mean images can be seen in Figure 6.6. I've placed the reference images from the TW2 atlas next to them and the match is good. Do keep in mind that the reference images from the TW2 method are single radiographs that have been chosen by the authors of the TW2 method as the "perfect" example of a stage. Because the mean images are made up of about 15 images each they are more blurred as a result of the variation within the radiographs from which they were composed.

As we had only a limited amount of images we had to make some choices regarding the amount of images we could use in each average image. We decided to use an amount of images that seemed to yield decent results. Of course some stages had more images

than others and a number of images was still needed as test images therefore we were sometimes forced to use a smaller amount of images.

The exact amount of images we had for each stage and the amount we used to make the mean images is shown in Table 6.6.1. It is obvious from the table why we did not make mean images for the first two stages B and C. There were not enough test images to make an ASM.

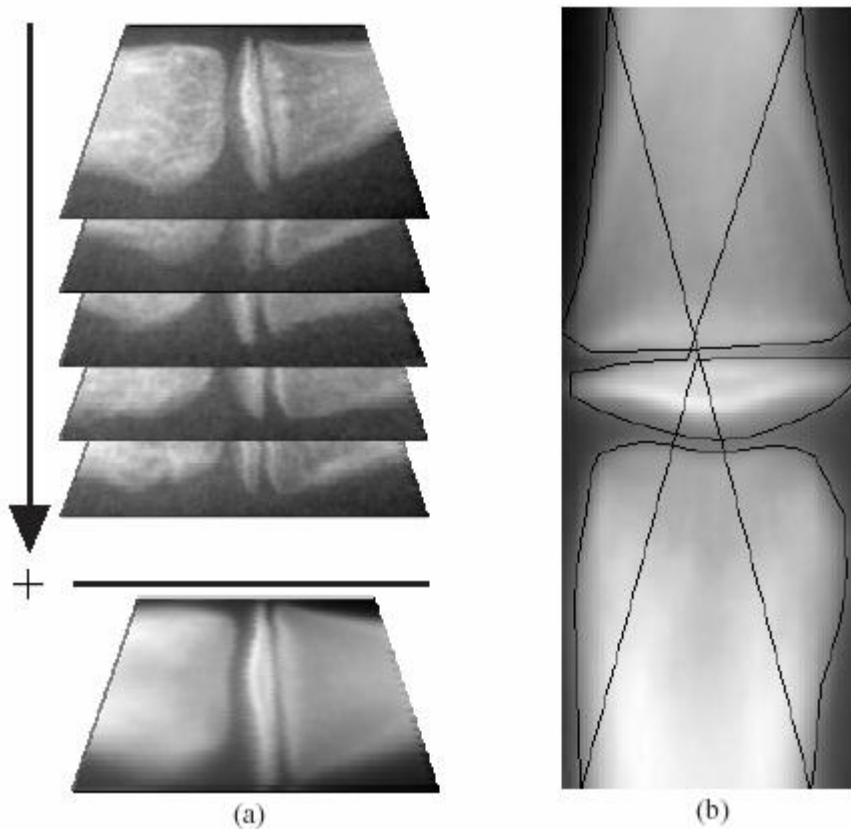


Figure 6.6 (a) Here we see the orthogonal projection of a number of different images of stage F resulting in a mean image for that same stage. All pixel intensities are added to one another so the mean image consists of the total intensity. (b) A mean image with a mean shape projected over it. The mean shape consists of points; the points have been connected in this image.

Table 6.1 The total number of images and the number used for each mean image case of stage H we have a problem. Only five images are available, this meant that we had to use all available images in our mean image.

Stage	B	C	D	E	F	G	H	I	All
Total Number of images	3	4	4	29	11	14	5	1	84
Number of Images Used in Mean	0	0	4	15	8	8	5	0	40

When calculating the mean images a mean shape is also calculated. This means shape consists of a collection of points that delineate the EMROI in the mean image (see Figure 6.6b)

6.5.2 Training the ASM

The method to automatically locate and clip out the EMROI is still being developed and tested by Maas [3]. Therefore we assumed that the images of the EMROI were given and used a program to manually clip them out (see Figure 6.7).

In order to be able to bring a random EMROI image and a mean image of a stage in registration we need to find the shape of the EMROI. We can find the shape of an EMROI by segmenting it from the background of its clipped out image.

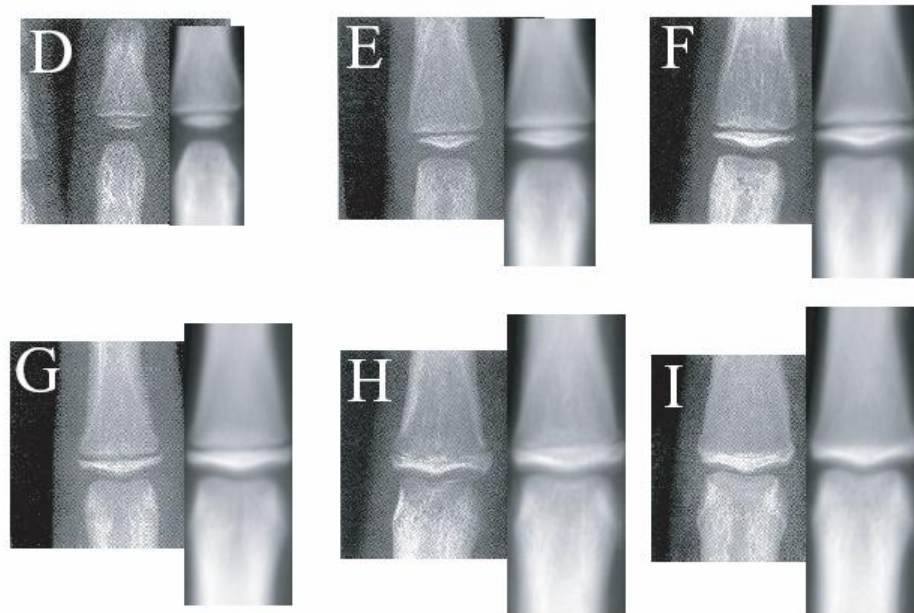


Figure 6.7 The left image of each pair is the TW2 atlas reference image while the right image is the mean image used in our experiments. They are constructed from several EMROI. Note that the mean images have been scaled to approximately the same scale as the reference images.

In section 6.2.4 the concept of ASMs has been introduced. These self adapting statistical shape models can find objects in images. In our case these objects are the EMROI. When using ASM the choice of the position of the points used to delineate the objects in the training set is very important. One can use the fixed points to accentuate certain aspects of a shape and when a lot of intermediary points are used in a certain area this heightens the chance that it is able to find a certain edge.

Selecting the positions of the fixed points that will be used to delineate the shape of the EMROI in the training set is the first step in building the ASM. According to Cootes [5] these should be placed at points on the edges of the object that are easily located throughout the training set. We have selected the points in Figure 6.8. Some points like 4, 11 and 18 are located where the longitudinal axis of the finger crosses the object border. Others like 1, 7, 15 and 21 are located at the border of the image. Points like 3, 5, 9, 10, 12 and 13 are set at the outer most points of the meta and epiphysis. These points could be used in determining the ratio of the width of the Meta and epiphysis which is an important indicator of skeletal age in the earlier stages. All the points are easily located in each image. Point 14 connects the contour of the epiphysis with that of the diaphysis. This is needed because of a limitation in the software that we used all points in the shape need to be part of a single object.

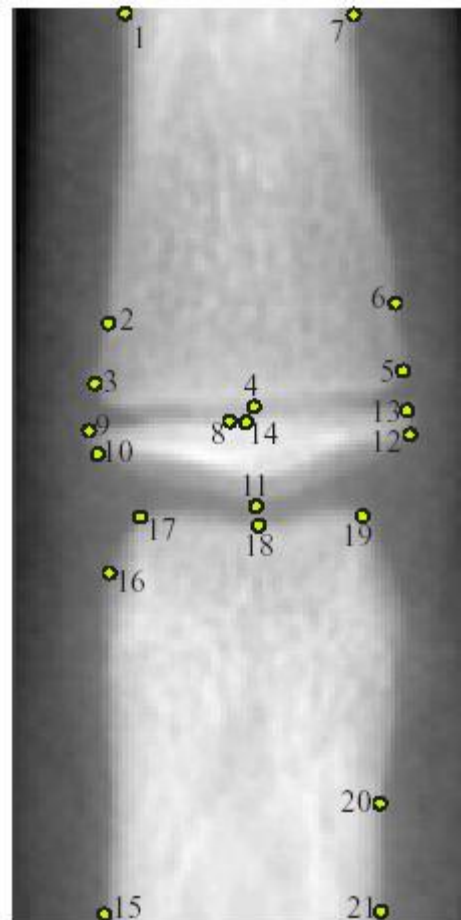


Figure 6.8 A typical EMROI with the fixed model points marked on it.

When we have determined all the fixed points we add the equidistant intermediary points. Varying the amount of intermediary points on an edge allows us to heighten or lower the "weight" of that edge. In our case we add a total of 70 intermediary points to the shape, added to the 21 fixed points this makes a total of 81 points. We add the relatively largest number of intermediary points to the epiphysis because that is a very important structure. Also at the bottom of the metaphysis and at the top of the diaphysis extra points are added.

There are some parameters that can be varied while building an ASM in order to change its behavior and the quality:

- Amount of variance to explain, f_v
- Aligned or unaligned training set
- k the number of points in the profile on either side of the landmark point.

Parameter f_v controls the number of eigenvectors to retain so that the model explains some proportion of the variance of the training set. When less variance of the training set is explained the ASM will generally not adapt well to subtle shape changes that may exist in the training set. Since the images we want to segment contain some small detailed differences we used a value of 0.995 for this parameter. One example of a subtle difference is the difference between stage F and stage G where only the outward edges of the epiphysis curl up (see Figure 6.6). A value of 0.995 means that 99.5% of the variation in the training set is explained by the model.

An aligned training set can be very useful when there are not too many training images available. By aligning the training images before building the ASM the shape model can only generate shapes with a size, position and orientation that is consistent with the supplied examples. When there are enough training examples available however it can be beneficial to use an unaligned model. Parameter k controls the length of the profiles that are taken perpendicular to the points of the model. A standard length is 4 pixels on each side thus the profile is a total of 9 pixels long ($2 \cdot 4 + 1$). After experimenting with this parameter, visual inspection of the segmentation results showed that this value also worked best for our application.

When segmenting an image with an ASM there are several parameters that can be varied influencing its segmentation:

- Number of new positions to evaluate during each iteration on either side of the current landmark position, n_s
- The number of iterations to do per scale level, N_{max}
- The bounds on the eigenvalues of λ_i , m

During each one of iteration of the algorithm a number of positions on either side of the current landmark are evaluated for all points in the shape. A standard value is 4 points, but tests showed that a value of two worked better. Because of the linearity of the ASM and the normalization of the grey value profile the ASM has difficulty distinguishing between certain edge structures. If we examine Figure 6.9 the first edges that would be detected would be the ones in the left image. Using a value of $n_s = 4$ the lower contour would jump to the stronger edge, a position shown in the right image of Figure 6.9. When using a value of 2 for n_s the ASM would stick to the correct edge because it does not detect the stronger edge above it. The number of iterations per scale level N_{max} is set at a standard amount of 10. Changing this value has little impact on the end result.

A high bound signifies that the eigenvalues can vary more and the shape model thus tolerates more extreme variations. The variation is set to 2.0 normally and that value seemed to produce the best results in our test set also. However, no extensive tests were conducted.

6.5.3 Assigning a TW2 Stage

Given a set of mean images of several stages and a segmented EMROI (query image) we propose to use the correlation between parts of these images as an indicator of similarity. The mean image that the query image has the highest correlation with indicates the stage of the EMROI in the query image. Correlation ρ of two jointly distributed variables X and Y is defined as

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} \quad (6.11)$$

The jointly distributed variables in this case are the grey values of the pixels in both the mean and the query image. Before we can use correlation we need to bring both the

mean and the query image into registration. This can be accomplished using the shape of the mean image and the segmented shape in the query image. By minimizing the distance between these two shapes we can register both images. Our first plan was to use a warping algorithm described by Ruprecht and Mueller [27] to accomplish this. However this transformation completely warps the shape in the mean image to the shape in the query image. This meant that the results were not usable because the correlation was always very high.

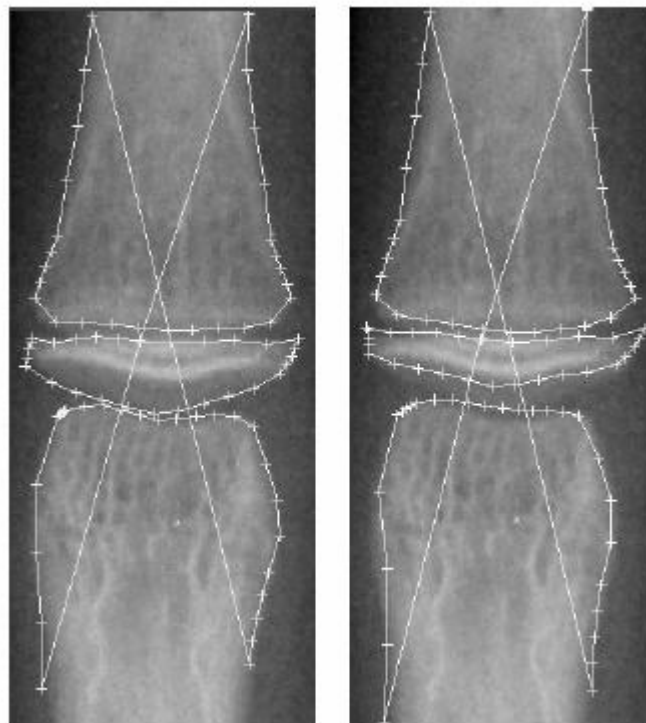


Figure 6.9 Varying ns between 2 and 4 had the effect shown here. On the left we see the segmentation result with $ns = 2$, on the right with $ns = 4$. The segmentation on the right might appear better but the dim shadow under the epiphysis must also be segmented.

Both the shapes as well as the appearance play an important role in determining the TW2 stage. Thus we wanted to use only transformations that would leave the appearance and shape of the original EMROI mostly intact (rotation, translation and scaling).

Procrustes analysis (see section 6.3) only uses rigid transformations plus scaling so it fulfills the criteria we set for a shape aligning method. Therefore we have chosen to use that analysis method to align the mean images with the query image. This means we leave the query image as it is and change the mean image so that the shape of its EMROI matches the shape in the query image.

Once a mean image and a query image have been aligned we use the correlation between parts of both images to indicate the similarity between both images. This part is specified as follows, the Y-coordinate of the 14th and 70th point in the query image shape is used as an indicator of the location of the EMROI. Then we take an area that is about one third of the total image surface and center this on the EMROI. Figure 6.10 illustrates the resulting areas.

6.6 Experiments

Using the ASM and the mean images discussed in the previous section a number of experiments were conducted. These experiments were designed to show that appearance based measurements could be used to assign a TW2 stage to an EMROI.

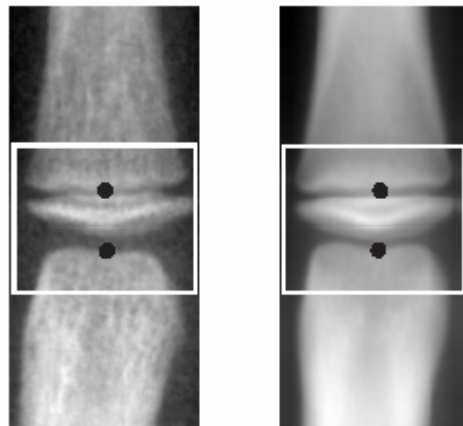


Figure 6.10 The area of the query (left) and of the mean (right) image that are compared using correlation. The two black dots represent the 14th (upper) and 70th (lower) point on the shape.

Because of practical and time limitations we have selected one EMROI for the experiments: the middle phalanx of the third finger. That EMROI is the subject of the experiments in this section. Below, when an EMROI is discussed, the middle phalanx of the third finger is meant.

A total of 40 query images were used in the test. Each of these images had a known TW2 stage. Classes E-I were represented in the query images. The distribution over the different stages was as in Table 6.1. We determined the correlation of each of these query images with each of the mean images of the stages E through I. We have not included stage B through D because we did not have enough images of the earlier stages (B and C) to train an ASM or make mean images. Since we previously determined that we needed

two ASMs one for the earlier and one for the later stages we decided to do the test with just the ASM for the later stages (E through I).

6.7 Discussion and Conclusions

There is one important factor which makes evaluating our results difficult: There is currently no method for skeletal age assessment which is uniformly accepted in clinical practice. This means that it is impossible to find a "golden standard" to which to compare our results. We have used the ratings from an experienced physician as a standard for this research. This doesn't mean that those ratings necessarily are the correct ratings. Research has shown that a difference of one stage between observers can occur in 30% or more of all analyzed bones when using the manual system (see Tanner[18]). Our own tests (see 5.5.3) yielded approximately the same number (on average 33.8%) of differently rated EMROI. For the middle phalanx of the third finger the average between observers variability was even 48%. This shows that assigning a TW2 stage to an EMROI is a process which is subject to considerable inter-observer variability. Our results should be interpreted with this in mind.

6.7.1 ASM Results and Discussion

The segmentation results we obtained with the ASM we developed seem good. Still it is important to know how often the ASM fails to find and correctly segment the epi, meta and diaphysis. We would also like to know how much this influences the assignment of TW2 stages later on in the method. Some qualitative tests have been done and in Figure 6.14 some results are shown rated from very good to average to bad to complete failure. The quality of the segmentation is determined by the fit of the end shape. Bad quality segmentation like the one in Figure 6.11 (3) exhibits the problem that the top of the epiphysis is not correctly segmented. In this case it is probably due to the curvature of the epiphysis which is quite extreme. A total failure like in Figure 6.14 is usually a problem with the initial position of the ASM. When the initial position is too far from the middle of the EMROI the ASM cannot find the edges of the epi- and metaphysis. This problem could be solved by trying multiple starting positions in the image. The starting position which results in the lowest Mahalanobis distance is the best starting position.

Another problem we encountered was that we wanted to make one ASM that could segment the entire range of TW2 stages. Because we did not have enough training images for the first stages (B and C) we could not construct an ASM that could segment those stages. The problem with this single model approach turned out to be that the difference between the later and earlier stages was so big that the ASM produced bad results for stage D and lower.

To solve this we propose to use two ASMs which overlap on one stage (stage E) to improve the results. The calendar age could then be used to give an indication of what range of stages is to be expected. Then the suitable ASM for the segmentation can be selected. We have not done any extensive testing of this method however.

6.8 Future Work

In the future further development of the method presented in this chapter can be expected. This future development could start with further testing of the proposed method on other EMROI in the fingers. Since EMROI in the fingers and hand follow the same skeletal development (epiphysis grows wider and then fuses with metaphysis) it should be possible to apply the method on those bones as well. There are some bones in the hand which would probably be less suited for our method. Some of the carpal bones would probably not work that well, also the metacarpals might present a problem but more testing is needed to confirm this.

The robustness of the method could be further improved by extending the use of the information provided by the ASM. Now we only use the segmentation to bring the target EMROI into registration with a mean image. Other information that could be derived from the ASM includes ratios between the width of the epi- and metaphysis and the location of certain sub ROI within the EMROI.

The ratio of the width of the epiphysis and the width of the metaphysis could provide a lot of information in the earlier stages. There, a better separation between stage D, E and F could be achieved because the assigning these stages in the original TW2 method depends heavily on these ratio. Another application for the extra ASM information is locating the borders of the epiphysis where the subtle change takes place when transforming from stage F to stage G and from stage H to stage I.

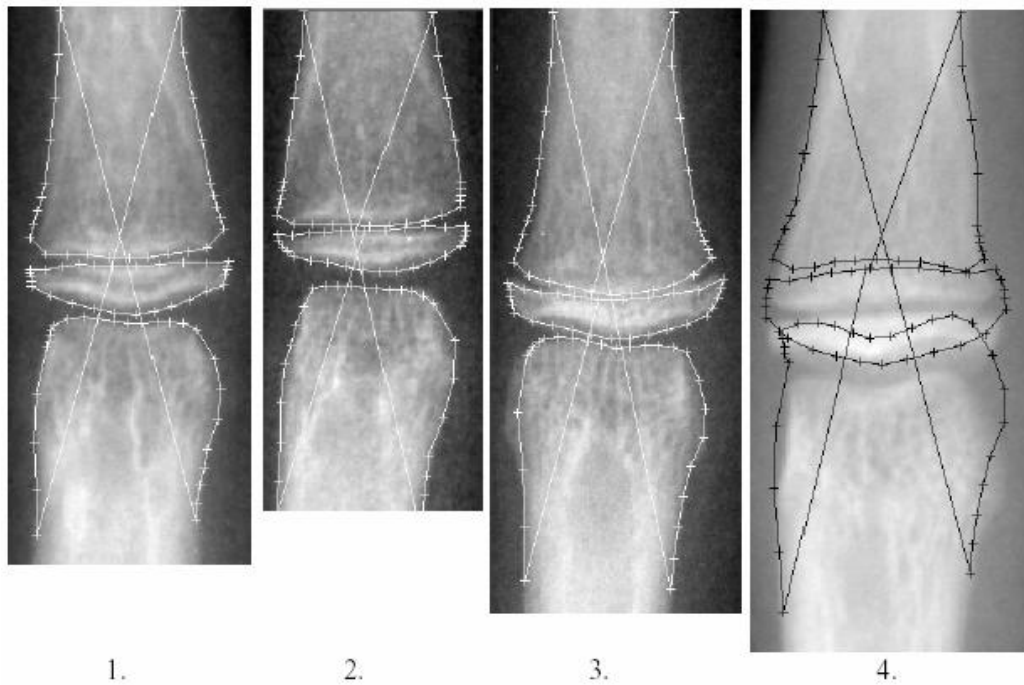


Figure 6.11 Some examples of segmentations done with our ASM. The results are ranked by quality of segmentation.

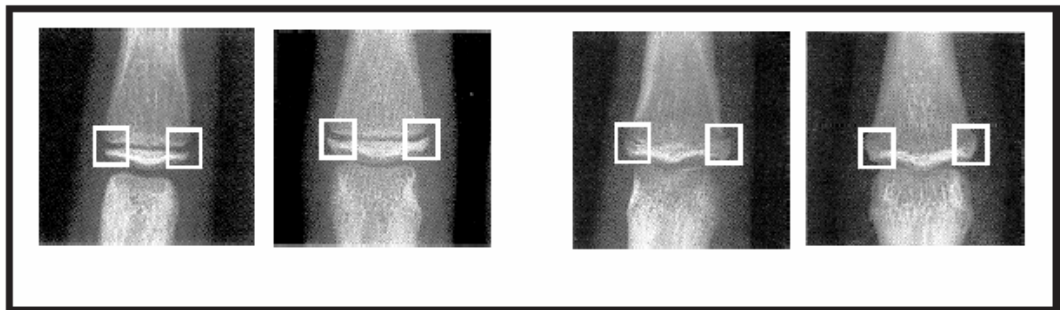


Figure 6.12 The marked areas exhibit the most change between stages F and G and stages H and I. Using the correlation between these smaller areas and the same part of the mean image could lead to better results or at least provide extra robustness in separating these highly similar stages.

Once these specific areas within the EMROI have been found, the correlation of these areas could be used to strengthen the choice for a certain stage. In Figure 6.12 these sub-ROI are marked. Using the correlation between these more specific areas could improve the performance of the method but also places higher requirements on the segmentation accuracy of the EMROI by the ASM.

Related to this is another interesting topic where further research could be expected, the segmentation quality of the ASM. It would be interesting to know how much the quality of the segmentation influences our method. I have done some qualitative testing

which looked good. Also the results of the alignment procedure using Procrustes seem to be good. Nevertheless, with the approach we are using now, good alignment seems crucial and a test could be devised which compared the results using a "perfect" manual segmentation and a segmentation done by the ASM.

Another possibility is to use a principal component analysis on the grey values of each EMROI image to build an appearance model of an EMROI. This appearance model can represent both shape and texture variability seen in a training set. The training set could consist of the images of the EMROI that have been aligned using Procrustes analysis. It could be possible to link the parameters of the appearance model to a certain TW2 stage and use this to assign stages to EMROI.

APPENDIX A

Explanation of the Matlab files of the Bone age assessment software.

Baa_asm.m

Main file used to assign TW2 stage to ROI with active shape model.

AlignShapeToShape.m

Used to align shape 2 to shape 1; by changing x and y coordinate.

AlignTrnSetCoor.m

Aligning training shapes to first shape.

FindShapeInImage.m

Search image until converged with training set

GetMatchingPosition.m

Search through the search vector and find a portion of it which is the most similar to the target vector.

GetMaxNumPyramidLevels.m

Get max needed levels for multi resolution search.

landmark.m

User chooses training landmarks then this function get training coordinates.

GetWeights.m

Use 3 dimensional matrix to contain the distances from landmark1(x) to landmark2(y) in a shape (z)

ScaleRotateTranslate.m

This function first scale and rotate then translate coordinates.

REFERENCES

1. F. Cao, H.K. Huang, E. Pietka, V. Gilsanz. "Digital hand atlas and web-based bone age assessment: system design and implementation," *Computerized medical imaging and graphics*, 2000.
2. W.W. Greulich and S.I. Pyle. "Radiographic atlas of skeletal development of hand and wrist." Stanford University Press, 2nd edition, 1971.
3. C. Maas. *Automatic detection of skeletal age based on hand x-rays of children*. Master's thesis, Image Sciences Institute, 2002.
4. G. Behiels, F. Maes, D. Vandermeulen, and P. Sueten. "Retrospective correction of the heel effect in hand radiographs. In W. Niessen and M. Viergever, editors," *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2001*, pages 301–308. Springer-Verlag, 2001.
5. D.J. Michael and A.C. Nelson. "HANDX: a model-based system for automatic segmentation of bones from digital hand radiographs," *IEEE Transactions on Medical Imaging*, 8(1):64–69, 1989.
6. E. Pietka, M.F. McNitt-Gray, M.L. Kuo, and H.K. Huang. "Computer assisted phalangeal analysis in skeletal age assessment," *IEEE Transactions on Medical Imaging*, 10(4):616–620, 1991.
7. E. Pietka, M. F. McNitt-Gray, T. Hall, and H. K. Huang. "Computerized bone analysis of hand radiographs," *SPIE Medical Imaging VI: Image Processing*, 1652:522–528, 1992.
8. N.D. Efford. "Knowledge-based segmentation and feature analysis of handwrist radiographs," Technical report, University of Leeds, School of Computer Studies, report 94.31. 1994.
9. G. K. Manos, A. Y. Cairns, I. W. Ricketts, and D. Sinclair. "Segmenting radiographs of the hand and wrist," *Computer Methods and Programs in Biomedicine*, 43:227–237, 1994.
10. D.T. Morris and C.F. Walshaw. "Segmentation of the finger bones as a prerequisite for the determination of bone age," *Image and Vision Computing*, 12(4):239–246, 1994.

11. S. Mahmoodi, B.S. Sharif, E.G. Chester, J.P. Owen, and R.E.J. Lee. "Automated vision system for skeletal age assessment using knowledge based techniques," *IEEE conference publication, ISSN 0537-9989*, 443:809– 813, 1997.
12. J. Duryea, Y. Jiang, P. Countryman, and H. K. Genant. "Automated algorithm for the identification of joint space and phalanx margin locations on digitized hand radiographs," *Medical Physics*, 26(3):453–461, 1998.
13. F. Vogelsang, M. Kohnen, H. Schneider, F. Weiler, M.W. Kilbinger, B.B. Wein, and R.W. Günter. "Skeletal maturity determination from hand radiograph by model based analysis," *Proceedings SPIE*, 3979:294–305, 2000.
14. E. Pietka, A. Gertych, S. Pospiech, F. Cao, H. K. Huang, and V. Gilsanz. "Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction," *IEEE Transaction on Medical Imaging*, 20(8):715–729, 2001.
15. S.M. Garn, K.P. Herzog, A.K. Poznanski, and J.M. Nagy. "Metacarpophalangeal length in the evaluation of skeletal malformation," *Radiology*, 105:375–381, 1972.
16. J.M. Tanner and R.D. Gibbons. "A computerized image analysis system for estimating Tanner- Whitehouse 2 bone age," *Hormone Research*, 42(6):282– 287, 1994.
17. A. Albanese, C. Hall, and R. Stanhope. "The use of a computerized method of bone age assessment in clinical practice," *Hormone Research*, 44(3):2–7, 1995.
18. J.M. Tanner, D. Oshman, G. Lindgren, J.A. Grunbaum, R. Elsouki, and D. Labarthe. "Reliability and validity of computer-assisted estimates of Tanner- Whitehouse skeletal maturity (casas): comparison with the manual method," *Hormone Research*, 42:288–294, 1994.
19. A. van Teunenbroek, W. de Waal, A. Roks, P. Chinafo, M. Fokker, P. Mulder, S. de Muinck Keizer-Schrama, and S. Drop. "Computer-aided skeletal age scores in healthy children, girls with turner syndrome, and in children with constitutionally tall stature," *Pediatric research*, 39(2):360–369, 1996.
20. E. Pietka. "Computer-assisted bone age assessment based on features automatically extracted from a hand radiograph," *Computerized Medical Imaging and Graphics*, 19(3):251–259, 1995.

21. E. Pietka and H. K. Huang. "Epiphyseal fusion assessment based on wavelets decomposition analysis," *Computerized Medical Imaging and Graphics*, 19(6):465–472, 1995.
22. G.W. Gross, J.M. Boone, and D.M. Bishop. "Pediatric skeletal age: Determination with neural networks," *Radiology*, 195(3):689–696, 1995.
23. R.F. Farr and P.J. Allisy-Roberts. *Physics for Medical Imaging*. W.B. Saunders Company Ltd, London, 1st edition, 1997.
24. G. Behiels, F. Maes, D. Vandermeulen, and P. Sueten. "Retrospective correction of the heel effect in hand radiographs. In W. Niessen and M. Viergever, editors," *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2001*, pages 301–308. Springer-Verlag, 2001.
25. M.B. Stegmann, R. Fisker, B.K. Ersbell, H.H. Thodberg, and L. Hyldstrup. "Active appearance models: Theory and cases," In *Proc. 9th Danish Conference on Pattern Recognition and Image Analysis*, volume 1, pages 49–57, Aalborg, Denmark, 2000.
26. T.F. Cootes and C.J. Taylor. "Statistical models of appearance for computer vision," Technical report, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, 2001.
27. D. Ruprecht and H. Muller. "Image warping with scattered data interpolation," *IEEE Computergraphics and Applications*, 15(2):37–43, 1995.
28. A.M. Marques da Silva, S.D. Olabbarriaga, C.A. Dietrich, and A.A. Schmitz. *Initial steps towards a digital signature for skeletal maturity*. 2001.

REFERENCES NOT CITED

1. R.K. Bull, P.D. Edwards, and P.M. Kemp. “Original articles – bone age assessment: A large scale comparison of the Greulich and Pyle, and Tanner and Whitehouse (TW2) methods,” *Archives of disease in childhood*, 81(2):172–173, 1999.
2. R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification*. Academic Press, New York, 2nd edition, 2001.
3. A.M. Marques da Silva, G.L. De Oliveira, G.P. Noal, C.A.A. Schmitz, L.S.B. Haeffner, P.S.P. Antunes, and S.D. Olabbarriaga. “EasyTW: computer-aided system for bone age assessment,” *Proceedings of CARS*, 2002.
4. Pietka and H. K. Huang. “Image processing techniques in bone age assessment,” *In Medical Imaging Systems Techniques and Applications*, chapter 5. Gordon and Breach Science Publishers, Amsterdam, 1st edition, 1997.
5. J.M. Tanner, R.H. Whitehouse, N. Cameron, W.A. Marshall, M.J.R. Healy, and H. Goldstein. *Assessment of skeletal maturity and prediction of adult height (TW2 method)*. Academic Press, London, 2nd edition, 1975.
6. van Ginneken. *Computer-Aided Diagnosis in Chest Radiography*. PhD thesis, Image Sciences Institute, 2001.
7. L Wicke. *Roentgen - Anatomie Normalbefunde*. Urban und Swartzenberg, Munchen, 5th edition, 1995.