

Chapter 2

Review on Segmentation of Computer-Aided Skeletal Maturity Assessment

**Yan Chai Hum, Khin Wee Lai, Nugraha Priya Utama,
Maheza Irna Mohamad Salim and Yin Mon Myint**

Abstract Bone age assessment (BAA) is an examination of ossification development with the purpose of deducing the skeletal age of children to monitor their skeletal development and predict their future adult height. Conventionally, it is performed by comparing left-hand radiographs to standard atlas by visual inspection; this process is subjective and time-consuming; therefore, the automated inspection system to overcome the drawbacks is established. However, the automated BAA system invariably confronts with problem in segmentation, which is the most crucial procedure in the computer-aided BAA. Inappropriate segmentation methods will produce unwanted noises that will affect the subsequent processes of the system. The current manual or semi-automated segmentation frameworks have impeded the system from becoming truly automated, objective, and efficient. The objective of this thesis is to provide a solution to the mentioned unsolved technical problem in segmentation for automated BAA system. The task is accomplished by first applying the modified histogram equalized module, then undergoing the proposed automated anisotropic diffusion, following by a novel

Y. C. Hum (✉)

MIMOS Berhad, Technology Park Malaysia, 57000 Kuala Lumpur, Malaysia

e-mail: yc.hum@mimos.my

K. W. Lai

Biomedical Engineering Department, University Malaya, Kuala Lumpur, Malaysia

e-mail: lai.khinwee@um.edu.my

N. P. Utama · Y. M. Myint

Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, Level 3, V01, Block A, Satellite Building, Skudai-Johor, Malaysia

e-mail: utama@biomedical.utm.my

Y. M. Myint

e-mail: yinmontt@biomedical.utm.my

M. I. Mohamad Salim

Universiti Teknologi Malaysia, Level 3, V01, Block A, Satellite Building, Skudai-Johor, Malaysia

e-mail: maheza@biomedical.utm.my

fuzzy quadruple division scheme to optimize the central segmentation algorithm, and finally, the process ends with an additional quality assurance scheme. The designed segmentation framework works without the need of resources such as training sets and skillful operator. The quantitative and qualitative analysis of the resultant images have both shown that the designed framework is capable of separating the soft tissue and background from the hand bone with relatively high accuracy despite omitting the above-mentioned resources.

2.1 Introduction

Having a distinct definition to physical maturity is itself a problem, what an to accurately measure it with precise quantitative value; traditional stature measurement does not assure common end points and this complicates the measurement of maturity. Therefore, we are not certain about a child maturity growth by his or her chronological age; therefore, stature measurement is unsuitable for maturity measurement. However, there are some defined events that are certain to normal individuals. Those events are considered to be suitable for maturity measurement. Events during puberty throughout adolescence such as eruption of a certain tooth, occurrence of first menstrual period, degree of testicular, and appearance of pubic hair can be used as indicator for the maturity. The maturity is deducible by the events occurrence. For example, a child that has had one particular event occurred is matured than the other has not had the particular event occurred. Moreover, the extent to which the event has developed can be treated as an indication for the maturity development. These events are regarded as developmental ‘milestones’ to represent the degree along which an individual has travelled in the maturity growing pathway.

The maturity measurement based on the events sequence does suffer from certain pitfalls. For example, the events are loosely spaced (Tanner 2001) such that it fails to coverage evenly and completely along the developmental age span (Tanner 2001). Therefore, hand and wrist bones started to be utilized for maturity measurement since hand and wrist bones contain enough sequences to cover the development age span. The appearance of certain bones of hand and wrist occur only during fetal stage and some during puberty stage, and these are all the mentioned sequences that can indicate maturity growth (Aicardi et al. 2000). The measurement of maturity by hand bones sequences involve total of twenty bones: radius, ulna, metacarpals, phalanges, and seven carpals bones develop in all stages along the pathway to full maturity. Therefore, analysis of the hand skeletal development enables the maturity to be measured. The mentioned analysis from hand skeletal bones development in deducing the bone maturity age is named ‘bone age assessment (BAA)’.

BAA is used clinically to assess the skeletal development especially in children and adolescents (Cao et al. 2000). As mentioned, estimation of maturation age by chronological age is inefficient. Skeletal maturity or bone age is regarded as a

diagnosis indication of growth disorders and also it can be used to predict future adult height of the child (Martin et al. 2011). Conventionally, the left-hand radiograph is utilized to measure skeletal maturation because it is proven that the skeletal growth of hand represents the biological maturity. Such analysis of maturation is predicated on important growth features such as ossification area and calcium position in the ossification area (Roche and French 1970). Diseases such as endocrine disorders (Chemaitilly and Sklar 2010), chromosomal disorders, and early sexual maturation can be indicated through the calculated difference between the skeletal age and biological age (Heinrich 1986).

Currently, there are two types of bone age evaluation systems (Peloschek et al. 2009). First one is regarded as the Greulich-Pyle atlas (Tristán-Vega and Arribas 2008) method (Greulich and Pyle 1959) and the second one is regarded as the Tanner-Whitehouse (TW2) methods (Tanner and Whitehouse 1975). For the first method, patient's hand bone radiograph will be compared with the atlas and the conclusion is drawn; the second method is method where all the clues of skeletal growth are gathered as evidence and then converted into point which then will be used to draw conclusion. The second method TW2 system has been improved into TW3 system (Tanner 2001) in which TW2 gathers the maturity evidences from twenty bones score consisting of the combination of radius, ulna, and short bones (RUS) and carpals whereas TW3 gathers maturity evidences separately from RUS or Carpal scores because research shows that conclusion drawn from the combination of RUS and carpals is not performing as good as analyzing them separately in BAA (Aja-Fernández et al. 2004).

Both methods' reliability and accuracy have been controversial since both of them are inspected visually by physicians or radiologists. This visual inspection by human is considered to be overly dependent on the operator knowledge background and perspective; moreover, it is too time-consuming (Acheson et al. 1963; Tanner and Gibbons 1994; Ontell et al. 1996; Martin et al. 2009). Therefore, in recent years, a lot of computerized system of BAA has been developed especially for TW2 method due to its nature that is more suitable for computational execution (Pietka et al. 2001; Hsieh et al. 2007; Thodberg et al. 2009a, b). However, this kind of computerized system is still developing and far from being perfect due to the system instability and the need for manual operator to execute the system (Jonsson 2002).

In this chapter, the aim is to discuss the very first step of this computerized system: the segmentation of hand bone. This segmentation has to be subject to several constraints in order to achieve stability and autonomous property. The purpose of this segmentation is to capture the bones of the hand skeleton and delineate the hand anatomical outline in the context where the system must be autonomous and the computational efficiency must be high enough to execute instantly. In next part, let us first discuss how other previous researcher in this areas has performed in addressing the segmentation problem and to what extent the current conventional approaches can be used to solve the problem.

2.2 The Attempts for the Hand Skeletal Bone Segmentation

The automated BAA system always go through a preprocessing step, namely the segmentation to eliminate the background, noise, soft tissue region since this information contains no pertinent clues in assessing the bone age in the computerized system (Pietka et al. 1993, 2001; Zhang and Huang 1997; Niemeijer et al. 2003; Aja-Fernández et al. 2004; Somkantha et al. 2011a, b). Not only this information provides no clues, it deteriorates the subsequent stages of computerized BAA system that will affect the accuracy. Nonetheless, most of the conventional techniques adopted in this preprocessing stage are not effective in using the computational resources. Most critically, the segmentation step always involves human operation to perform. Moreover, many techniques execute the segmentation after getting the region of interest (ROI) to ease the process of segmentation (Han et al. 2007; Hsieh et al. 2008). However, this accuracy and performance of ROI search is improvable by executing the algorithm after hand bone segmentation from the soft tissue region. As the significant initial step of the system, the output accuracy and practicality of segmentation are critical since the quality of the computerized BAA system output depends heavily on it (Sotoca et al. 2003).

Substantial efforts have been devoted to research on the preprocessing of hand skeletal bone from background and soft tissue region. Most of the works involve the application of threshold setting which is considered ineffective in the hand bone segmentation due to the fact that the soft tissue region contains pixel intensity that similar to spongy bone of the hand skeletal bone (Smyth et al. 1997; Pietka et al. 2001; Aja-Fernández et al. 2004; Buie et al. 2007). Besides, most of the work, after obtaining the region of interest (ROI), implements the active contour model (Kass et al. 1988) that has inherent weaknesses such as high sensitivity toward intensity gradient, high dependency on initiation location and low ability in growing into concavity (Mahmoodi et al. 1997, 2000; Sebastian et al. 2003). Some works implement the statistical analysis to determine the membership of each pixel, whether it is belong to the bone or the soft tissue region (Tristán-Vega and Arribas 2008; Giordano et al. 2010). Some works combine various techniques segmentation in other field into the hand skeletal bone segmentation (Jong-Min and Whoi-Yul 2008; Somkantha et al. 2011a, b). The development of the study has been summarized the following paragraphs:

As early attempt, Michael and Nelson (1989) propose a CAD system for BAA consists of preprocessing, segmentation, and measurement. They preprocess the image using the histogram equalization, and then it is followed by converting the image into binary image and implementing the threshold method of pixel's intensity to remove the background using the model parameters. By using the model parameter, the main drawback is that the problem of overlapping of pixel intensity in bone and background could not be resolved. Moreover, high sensitivity to illumination change and the soft tissue region surrounding the hand bone further

deteriorates the result. Manos et al. (1993, 1994) proposed a framework for the automatic hand-wrist segmentation; they have implemented a technique of region growing and region merging after performing the edge detection during the pre-processing. Along this technique, thresholds are used to determine the edge and growing and merging algorithms. Besides, region growing result relies heavily on the performance of edge detection. Lastly, the region merging depends on gray-level similarity size and connectivity that bear a risk of combining the epiphysis sites that are situated around the metaphysis.

A group of well-known researchers for computerized BAA system Pietka et al. (1993) has conducted a number of studies on computer-aided system of hand bone analysis. Segmentation process in their early studies, thresholding, and dilation technique are used for the bones extraction. The algorithm discussed involves dilation that might ruin the result when bones are approaching each other. In the following attempt by Pietka (1994), she has started to extensively focus on the preprocessing procedure on the bone segmentation from the background using windowing technique to compute the local statistical properties followed by finding the centroid from each peak of the histogram of local window. As further attempt, Pietka et al. (2001) conduct a study on image preprocessing and Epiphyseal/Metaphyseal ROI Extraction in BAA automated system. The proposed method is about employing the method of adaptive thresholding. The statistical value of mean and variance of each window is then computed to determine the ROI utilizing the technique of star-shaped median filter and Lee filtering to segment the bone from soft tissue region after obtaining the ROI. However, the method does not address the problem of segmentation with high reliability. The number of peaks found in each local window is uncertain. Errors of computing would occur in some part of the image.

Sharif et al. (1994) have published a paper on bone edge detection: Segmentation of bone employing edge detection base on the intensity by the derivative of Gaussian (Drog) followed by the employment of thresholding technique. The preprocessing technique implemented by Mahmoodi et al. (1997) involves changing the image into binary, and performing the thresholding method using image histogram to obtain the ROI, the subsequent segmentation of epiphysis within the ROI is implemented through the technique of active shape model (ASM). Similarly, the drawbacks of the method are the sensitivity in illumination change and the soft tissue region. The preprocessing method is used in Mahmoodi et al. (1999) for segmentation of bone using deformable models and a hierarchical bone localization scheme. The method background removing process is performed only after obtaining the ROI. Mahmoodi et al. (2000) adopt binary thresholding to acquire the delineation of the hand, followed by location searching of concave-convex; finally, the segmentation is performed by the method of ASMs.

Sebastian et al. (2003) work on segmenting the carpal bones from CT images using deformable models, and the preprocessing combines the strength of all popular segmentation techniques such as active contour models, region growing and the global competition in seeded region growing, and also the local competition in region competition. The result is satisfying, but it involves complicated

and heavy computing consumption while computing the partial differential equation. Active contour model (Pietka et al. 2001) has been used in segmenting the bone; the methods (Cao et al. 2000) c-means clustering algorithm, Gibbs random fields, and estimation of the intensity function have been proposed by Pietka et al. They also proposed Gao et al. (2010) segmentation of hand bone during preprocessing using the analysis on histogram. By inspecting the peak of the histogram, the authors identify the soft tissue region and the background.

Hsieh et al. (2007) incorporate adaptive segmentation method with Gibbs random field at the preprocessing stage. Zhang et al. (2007) suggest segmenting the carpal by non-linear filter as preprocessing follows by adaptive image threshold setting, binary image labeling, and small object removal. However, it involves user-specified threshold and Canny edge detection that are not robust in segmentation. Similarly, Somkantha et al. (2011a, b) segment the carpals bone using a combination of vector image model and Canny edge detector. Han et al. (2007) propose to implement watershed transform and gradient vector flow (GVF) to perform the segmentation where the performance of watershed transform and GVF depends heavily on edge gradient strength. Tran Thi My Hue et al. (2011) proposes to implement watershed transform with multistage merging for the segmentation task. Liu et al. (2008) implement only primitive image processing technique such as edge detection and template matching on the preprocessing segmentation. Giordano et al. (2010) perform the segmentation utilizing the derivative difference of Gaussian (DrDog) techniques followed by thresholding using mean and standard deviation. The impracticality of thresholding, edge detectors, and watershed transform will be discussed latter in next chapter.

The utilization of the state-of-the-art technique of deformable model such as ASM and active appearance model (AAM) of hand bone segmentation has gained considerable attentions in recent years (Thodberg 2002; Thodberg and Rosholm 2003; Thodberg et al. 2009a, b). The strength of this method is that it is well founded on statistical learning theory. However, the main drawback of this technique is that it is not yet developed into a fully automated system. The initialization of the technique is to delineate the hand bone shape, and this thus far is accomplished manually. Manual shape delineation is extremely time-consuming especially for ASM and AAM that require a substantial number of training samples before giving enough information to the algorithm about the changes of the hand bone shape. Furthermore, the training samples have to be sufficient to have a comprehensive coverage of hand shape changes and number of bones in different age groups. In other words, this type of segmentation framework is effective only when both human operators and complete training set are available. Thus, the practicality is limited in situation when those resources are limited.

In conclusion, the current existing segmentation methods and frameworks either are involve in threshold settings or are too dependent on certain resources and image features. This indicates that improvement on hand bone segmentation is necessary in order to practically realize the fully automated computerized BAA system. Thus, this research is to explore this improvement aiming to establish a

fully automated segmentation framework that is accurate yet remains less dependent on external resources.

2.3 The Problem of Research Problem

The computerized BAA relies heavily on the performance of the segmentation in order to provide the accurate ROI of the hand for further analysis to assess the bone age. Distinct boundaries of anatomical structures of the ossification sites have implications in features localization and analysis which are essential to assure high reliability of the computerized system. Human visual capability, generally, is able to accomplish this visual recognition task effortlessly, but it is not the case for computational visual system due to various obstacles such as the lacking of prior knowledge to reason about the semantics of the object. There are abundant of hand bone segmentation techniques found in the literatures, but very few of which, if ever, is functioning as an effective and yet remains fully automated. The research problem, therefore, is to explore the question: Is there any method that can realize the goal of performing the automated segmentation task that is relatively much more effective than fundamental segmentation techniques but yet is unaffected by constraints such as training sets and human intervention that are invariably pertain to sophisticated techniques?

The factors that associate with the problem are presented as following:

1. Variability in the radiograph image attributes such as deviations in terms of illumination, number of bones, size of bones, and locations of bone. This variability deviates across different input sources and different age groups of the subjects in radiographs. The problem with this variability is that it impairs the performance consistency or precision of segmentation technique or frameworks once the input radiographs are not as expected.
2. The nature of the preprocessing module or the central algorithm in the framework that demands high degree of human cognitive ability such as visual perception or observation on certain patterns of input. This problem is attributable to devoid of prior knowledge of computational algorithm in recognizing pattern that can be easily perceived by human such as the shape of hand bone and the variability of luminance and illumination. As a consequence, most segmentation framework necessitates explicit labors and hence this problem violates automaticity.
3. The inherent bone intensity property in radiograph that stems primarily from the variations in anatomical density of different parts of the hand bone. As a consequence, two adverse properties for segmentation performance take place:
 - (a) The overlapping range of pixel intensity for the cancellous bones, the soft tissue regions, and the compact bones. This overlapping intensity range is the main complication that thwarts any global image processing techniques. For instance, the low mineral density in cancellous bone and soft tissue

regions leads to similar degree of penetrations and absorptions of X-ray protons resulting in overlapping intensity effect. Most of the intensity-based image processing that postulates on distinct separation of intensity distribution fails to address this problem.

- (b) The non-uniformity within the same category of bone such as cancellous bone or the cortical bone. The radiograph intensity is not evenly distributed. This problem stems primarily from the inhomogeneity of mineral density inside the cancellous bone due to porous structure embodying a variety of spaces that are different in size and mineral density. As a consequence, the degree of absorptions is different and hence giving it an irregular spongy textured appearance which in turn results in large variations of pixel intensity within the same structure of bone. This non-uniformity problem ultimately becomes the problem of overlapping range of intensity problem. Hence, additional processing stages containing higher-dimensional input data are in demand to differentiate the pixel into correct labeling to avoid misclassification.
4. The uneven brightness intensity difference between the edge border of compact bone and soft tissue regions and also between soft tissue regions and background further complicates the problem. This is the main problem that deteriorates the performance of edge-based segmentation technique that depends solely on edge information as main input.
5. The existing segmentation methods, as illustrated in next chapter, either are too simplified to have adequate considerations encompassing various aspects of being a comprehensive technique to perform the segmentation tasks or being too dependent on limited resources such as the availability of training hand bone samples over all age ranges, computational complexity, and the knowledge background of operators.
6. Lack of quality assurance process in the hand bone segmentation frameworks to consider possible artifacts. The segmented hand bone at the first stage of computerized BAA system should not proceed to subsequent processing stage before assessing the quality of the segmented hand bone. Most of the existing works do not incorporate such functionality in the designed segmentation frameworks. As a consequence, the subsequent processing stage such as feature analysis for BAA will accept inferior-quality segmented hand bone as input and hence produces final result that is not reliable.

Partitioning hand bone from radiograph background and soft tissue region is the first stage of computerized BAA system; the performance of this stage underpins the success of subsequent procedures of BAA, which in turn affects the final result. Despite being the significant step in computerized BAA, the automated segmentation remains a challenging problem owing to above-mentioned problems. Conventional segmentation methods and the currently developed segmentation framework designed for hand bone segmentation are generally impractical to be implemented attributable to high dependency on various resources. The most

problematic of which is the number of user-specified parameters in developing a fully automated segmentation without any human intervention. In addition, to the best of our knowledge after reviewing a vast number of literatures, the comprehensive critical appraisal on the research area of hand bone segmentation can rarely be found, if ever. Therefore, in this research, an automated and practical segmentation framework is to be proposed techniques and implemented to solve the above-mentioned problem after critically evaluating various existing segmentation techniques and frameworks.

2.4 Authors' Proposed Segmentation Framework

The contribution of authors in this problem is to develop an effective segmentation frameworks consisting of several modules that are fully automated and independent from the completeness of training samples and availability of skillful operators. The technical details can be found in Hum (2013). This proposed segmentation framework produces superior segmentations, yet it remains computationally feasible. The performance edges are summarized as follows:

1. Present the critical appraisal of existing segmentation techniques and existing hand bone segmentations frameworks by first elucidating the postulation of each technique based on and the technical details of each technique, followed by illustrating the strengths and weaknesses of which, and finally reasoning their suitability to perform the segmentation task of hand bone. Despite an abundance of segmentation reviews, we could not find any relevant and comprehensive critical evaluations on the aspect of theoretical and technical in the hand bone segmentation for computerized BAA. Hence, this critical appraisal can contribute by serving as the basis for future reference on the research of this field.
2. Extend the comprehensiveness of existing histogram equalization technique by first assessing the current theoretical and technical architecture of existing histogram equalization methods, and then, we contribute our new insight into revolutionize the conventional perception toward the ultimate goal of histogram equalization by proposing the new histogram equalization framework; then, based on the revolutionized insight, we develop a holistic histogram equalization in terms of luminance preservation, contrast, and detail preservation based on the beta function to preprocess the hand bone radiograph serving the purposes of standardizing and equalizing the non-standardized illumination among radiographs that contain high variations in luminance, improving luminance difference across edge borders in radiographs, reducing variations in luminance difference across edge borders among radiographs and most importantly, enhancing the visual perceptual effect of ossification sites for improving the accuracy in ossification localization and BAA.

3. Extend the body of knowledge of anisotropic diffusion by exploiting the potential of being fully autonomous and adaptive to input radiograph instead of being subjectively tuned by operators to solve the problem of non-uniformity and mitigate the undesired effect of overlapping intensity range. Both contributions as following have profound implication for advancing the field of anisotropic diffusion and provide adequate ground for framework that requires autonomous anisotropic diffusion.
 - (a) Address the problem of manual diffusion strength by designing an automated diffusion strength scheme based on the diffusion coefficient function of speckle reducing anisotropic diffusion (SRAD) grounding on the well-founded statistical theory of the relation between sample variance and global variance. The main strength is its computationally attractiveness and practical applicability.
 - (b) Address the problem of manual scale selection by designing an automated scale selection. The main strength of which compared with limited existing automated scale selection schemes is that it requires no excessive filtered image before determining to halt the diffusion iteration.
4. Transform the manual and rigid adaptive division scheme into an automated adaptive quadruple division scheme that embodies human cognitive ability. This transformation is significant not only in a narrow sense of hand bone segmentation, but most importantly, it is of a generic breakthrough in the field of image segmentation. This implicit modeling of human intuition and prior knowledge solves the problem of high dependency on explicit human resources to operate the algorithm. Furthermore, the scheme itself is a building block or framework for other segmentation algorithm to determine the optimum region size for algorithm implementation.
5. Incorporate quality assurance module in the segmentation framework to evaluate the appropriateness to serve as input for subsequent stages of computerized BAA. The step is important to eliminate over-segmented regions of hand bone and restore under-segmented regions to further improve the quality of segmented hand bone. This concept provides insight about a feedback system of most of the imperfect segmentation framework system should contain a stage that capable of analyzing the current output and patch up the incompleteness accordingly.
6. In conclusion to contributions, this thesis provides the contrary perception to conventional concept that prone to complicating the segmentation algorithm to seek for enhancement in segmentation performance. Instead, the proposed segmentation framework pioneers the insight postulating that combinations of several customized modules are capable of achieving result that tantamount to result achieved by complicated algorithm or algorithm that demands scarce and limited resources. The conception lies in the strategy to identify, target, and analyze the principal adversities that impede the performance of existing methods; then, based on the analyzed result, we contribute by advancing the concept of adapting the input information to the central segmentation algorithm

(in this thesis, it refers to unsupervised clustering). This concept is of contrast to conventional segmentation framework that tends to complicate the fundamental segmentation algorithm in order to adapt the input information. The strength of each module and the idea of concatenating each of them by utilizing the by-product of each module are breaking new ground in the field of automated image processing and pattern recognition. These instructive insights embrace the potential to spark attentions and generate new research grounds that will in turn contribute in other fields of applications.

2.5 Conventional Segmentation Techniques Performance Discussion

This section provides an overview for traditional or conventional segmentation techniques that are usually adopted in applications related to images segmentation to give the reader an overview of the development of segmentation techniques. The fundamental concept of each technique is presented and the pros and cons of each technique are discussed. Besides, we illustrate the unsuitability of traditional segmentation techniques as hand bone segmentation technique in the context of computerized BAA by analyzing the nature of the technique and by implementing the analyzed technique in hand bone segmentation. This evaluation and implementation of previous techniques in hand bone segmentation are crucial motivate the objective and justify the contribution of this thesis. This section ends with the conclusion that a more advanced technique of hand bone segmentation should be derived instead of using the traditional segmentation techniques.

Thresholding is one of the earliest image segmentation techniques, and yet it remains to be the most widely applied segmentation technique attributable to its simplicity and intuitiveness (Sezgin and Sankur 2004). Thresholding segmentation is normally conducted in spatial domain based on the postulation that both object and background are represented by different range of pixel intensity (Gonzalez and Woods 2007). Basically, there are three categories of thresholding: global thresholding, local thresholding, and dynamic thresholding (Bernsen 1986).

Undoubtedly, the simplest method in thresholding techniques to segment an image is through single global thresholding: this technique based on the concept that if object in the image and other object or background are mutually exclusive in terms of intensity range, then it could be separated in different partition using a single or multiple values of pixels intensity (Lee et al. 1990). In the case of single threshold, it can be represented as following:

$$f(x, y) = \begin{cases} g_1 & \text{if } f(x, y) < T \\ g_2 & \text{if } f(x, y) \geq T \end{cases} \quad (2.1)$$

where g_i is group of pixel that represents an object or background; if a pixel value is less than T , which is the threshold value, then it is grouped into g_1 ; if a pixel value is more than T , then it is grouped into g_2 . The $f(x, y)$ is image pixel intensity in 2D grayscale image in coordination (x, y) . The concern of the technique is to classify an image into object and background; this type of grouping is called binarization.

The single thresholding depends on the T . This T value determines the intensity range of an object and the intensity range of the image background. For instance, (if the object is brighter than the background) if a pixel intensity value is more than the threshold value, then the pixel will be classified as object; for the pixels which possesses intensity value less than or equal the threshold value, they will be considered as background. This kind of threshold method is considered as ‘threshold above’; another type is ‘threshold inside’ where the object value is in between two threshold values; similarly, another variant is ‘threshold outside’ where the value in between the two threshold values would be classified as background (Shapiro and Stockman 2001).

The efficiency of thresholding technique in segmentation mainly depends on two factors: first factor is the property of the image intensity distribution of both object and background. Thresholding technique performs most efficiently when the intensity of input image has distinct bi-modal distribution without any overlapping range of intensity for object and background (Liyuan et al. 1997). Overlapping range of intensity occurs often due to uneven illumination. Besides, the nature of the object itself can lead to overlapping range in which some regions within the objects in input image has overlapping range of intensity to background. As mentioned in previous chapter, one of the natures of X-ray hand bone radiograph is its uneven illumination throughout the image as well as its overlapping range of intensity distribution among soft tissue region, trabecular bone, and cortical bone due to the nature of hand bone and uneven background illumination as well.

The reasons for the inferior quality of segmented hand bone by thresholding can be summarized as follows:

1. Assumption that the whole targeted object (which is the hand bone in our case without soft tissue region) contains similar intensity range. This is always not true for hand bone radiograph as within the hand bone, there are regions of trabecular bone and cortical bone which have different bone density and hence are represented by different range of pixel intensity values in digital image.
2. Assumption that the histogram of targeted object and background (black regions and soft tissue regions) is of perfectly separation into two groups of intensity distributions. This is always not true for hand bone that the histogram of hand bone radiograph is not bi-modal distributed. This can be explained from the nature of hand bones that are formed by three classes of regions: bone, soft tissue regions, and background instead of two.
3. Assumption that there is no overlapping of intensity range between background and targeted object. This is always not true for hand bone as the some of the

intensity in soft tissue regions are identical to the regions in trabecular bones. The global thresholding neglects this intensity overlapping problem.

4. Assumption that the illumination is even in input image. This is always not true in for hand bone radiograph that lower region of hand bone radiograph has more intense illumination relative to upper region of the radiograph. The global thresholding neglects this uneven illumination and this affects the segmentation result.

Another critical problem of single global thresholding is the choice of the threshold value to obtain favorable segmentation result (Baradez et al. 2004). In fact, even the 'best' threshold value is selected, the resultant segmented image in the context of hand bone radiograph and in other medical image processing remain inferior. This fact is inevitable due to the nature of global thresholding and the nature of hand bone segmentation: only one threshold. One improvement for this limitation is by adopting multiple global thresholding (Yan et al. 2005). multilevel thresholding classifies the image into multiple classes (>2) (Tsai 1995). The multiple thresholding can be represented as follows:

$$f(x, y) = \begin{cases} g_1 & \text{if } f(x, y) > T_1 \\ g_2 & \text{if } T_1 < f(x, y) \leq T_2 \\ \vdots & \vdots \\ g_{n-1} & \text{if } T_{n-3} \leq f(x, y) \leq T_{n-1} \\ g_n & \text{if } f(x, y) \geq T_{n-1} \end{cases} \quad (2.2)$$

where g_i is group of pixel that represents an object or background. T_i is the threshold values. The $f(x, y)$ is the image pixel intensity in 2D grayscale image in coordination (x, y) .

Multiple thresholding might solve the problem arises from the assumption that the input image is of bi-modal type but solve not the problem arises from assumption that the input image is of even illumination. In next subsection, we would review and examine the local/adaptive thresholding that is claimed to be more effective in tackling the problem of uneven illumination.

Adaptive thresholding is segmentation using different thresholds in different sub-images of input image (Zhao et al. 2000). The input image is firstly divided into a number of sub-images; then in each sub-image, suitable threshold is chosen to perform the segmentation, and this process repeats until all sub-images undergo the thresholding segmentation. Adopting different threshold in different region of the input image is proven to be more effective than global thresholding that it is easier to obtain well-separated bi-modal or multiple-modal distributions in the sub-images, and hence, it improves the segmentation result (Shafait et al. 2008). In addition, sub-images are more likely to have uniform illumination implying that as it could resolve the problem that arises from the non-uniform illumination (Huang et al. 2005).

Undoubtedly, it is a fact that adaptive thresholding performs better than global thresholding in tackling the problem of uneven illumination. There are some

difficulties in applying the technique effectively in hand bone segmentation due to the problems as follows:

1. The problem arises from making the assumption there is no intensity overlapping between target object and background.
2. The size of each sub-image is difficult to determine. If the size is smaller or larger than it should be, then the result might be even more inferior than using global thresholding.
3. The size of the sub-images is globally set and is fixed throughout the entire image. Some regions need smaller sub-image whereas some regions need larger sub-image in adaptive thresholding to optimize the segmentation and the computational efficiency.
4. The number of thresholds needed in each sub-image is difficult to determine.
5. The computational cost increases in comparison with global thresholding.

The threshold values are difficult to be set manually as the number of sub-images increases (Buie et al. 2007). In global thresholding as well, the threshold value need to be correctly set in order to optimize the result. We afford to set single global threshold using human inspection. However, when we are dealing with multiple thresholding or adaptive thresholding, automated thresholding is more suitable to decrease repetitive threshold setting by human which is subjective and yet time-consuming. In next subsection, we explore and study about the automated threshold value setting techniques which can be applied in both global thresholding and adaptive thresholding. The implementation of multiple thresholding and adaptive thresholding in hand bone segmentation is illustrated in next subsection using automated threshold values selection to demonstrate that the sole implementation of these technique fail to provide good segmented hand bone.

In global thresholding, each pixel is compared with the global threshold; in local thresholding, each pixel in sub-image is compared with each local threshold which is computed from each sub-image; in dynamic thresholding, each pixel is compared with each dynamic threshold which is computed from sliding a kernel over the input image (Shafait et al. 2008). One of the popular dynamic thresholding methods is Niback method (Niblack 1990).

Generally, dynamic thresholding performs better than global thresholding and local thresholding. However, it has similar drawback as local thresholding that we need to determine the kernel size; the threshold has to be selected manually depending on application. Only suitable selection of kernel size and threshold can produce optimum result of segmentation. In addition, dynamic thresholding consumes much more computational resources relative to local thresholding and global thresholding due to its pixel-wise nature. Besides, in performing the neighborhood operations for dynamic thresholding, the padding problem arises when the kernel approaches the image borders where one or more rows or columns of the kernel are placed out of the input image coordinates.

The main technical issue being frequently discussed is the threshold value selection: the decision to determine the threshold value in which the object and the background could be separated as accurate as possible or the decision to select the threshold value so that the object and the background misclassification rate are lowest. The result of thresholding segmentation process depends heavily on this value. An inaccurate or inappropriate setting of this value will produce disastrous result in thresholding segmentation.

For the choice of threshold value, basically, there are two main methods: the manual threshold selection and the automated threshold selection. Manually determined threshold value heavily relies on human visual system. Threshold value is selected using Visual perception to partition the object from the background; the main drawback of this threshold selection is that it involves human subjective perception toward image quality. Besides, the process itself is extremely time-consuming if the operation involves multiple thresholds. Therefore, it is not practical to determine the threshold value of a large number of images. In short, the manually determined value is not effective.

For automated thresholding method, various methods exist: the simplest method is to utilize the image statistics such as mean, median (second quartile), first quartile, and third quartile, to act as threshold value (De Santis and Sinisgalli 1999): this method performs only relatively well in an image free of noises; the reason is that the noise in the image has influenced the statistic of the image. Typically, if the mean of an image used as threshold value, then it can separate a typical image with object brighter than background into two components; however, while noises exist, the noises have altered the nature that the pixels with intensity more than mean are belonged to the object. Besides, this kind of thresholding method assumes that the object and the background are themselves homogenous. In other words, the object is a group of pixels containing similar pixel intensity; the background is a group of pixels with similar intensity. This assumption has serious limitation especially in medical image segmentation where the target objects like organs or bone are not inherently homogenous. Besides using simple aforementioned statistic in input image, there are other methods to choose the threshold value. In next paragraph, we explore and study different types of automated thresholding techniques that have been developed.

Attributable to the limitations of using simple statistics, various more sophisticated types of thresholding methods based on different techniques in determining the threshold value are proposed: one of the methods is the threshold value selection based on histogram: instead of choosing the mean or median of the image as the threshold value to separate the object and the background, the histogram-based thresholding method determines the threshold value based on the histogram shape assuming that there are distinct range for object and background themselves. The value of a valley point is set as threshold.

In image processing, when the histogram of an image is mentioned, typically we mean a histogram of the values of pixel intensity; the graph of the histogram represents the number of pixels in an image at each intensity value of the pixel in the image. If say in an 8-bit grayscale image, there will be 2^8 possible values and it

means that the histogram shows the occurrence frequency of each intensity in the image. In other words, it is a representation of the image statistics based on the number of the specific intensity's occurrence.

Histogram analysis is a popular method in automated thresholding (Whatmough 1991). The postulation is that the information obtained from the physical shape of the histogram of the input image signalizes the suitable threshold value in dividing the input image into meaningful regions (Luijendijk 1991). Conventionally, the intensity bin in the valley between peaks is chosen as threshold to reduce the segmentation error rate. Instead of using manual inspections, by only analyzing the shape of the histogram and compute the intensity bin that represents the valley, the relatively good threshold value can be found (Guo and Pandit 1998).

However, the main drawback of this technique is that it depends too heavily on the shape of pixel intensity distribution. Besides, it has no consideration on the pixels location and the pixel surroundings and this leads to the failure in recognizing the semantic of the input image. This method fails when the input image does not have distinctly separated intensity distribution between the foreground and background due to overlapping of intensity as mentioned in last subsection of global thresholding. This category of automatic threshold selection performs thresholding in accordance with the intensity histogram's shape properties. Utilizing basically the histogram's convex hull and curvature, the intervening valley and peaks are identified (Whatmough 1991).

This concept is based on the facts that regions with uniform intensity will produce apparent peaks in the histogram. If only the image has distinct peaks on each objects in the images, then multiple thresholding is always applicable via histogram-based thresholding. The favorable shapes of the histogram for the purpose of segmentation are tall, narrow and contain deep valleys. This method is less influenced by the noise, but it has drawbacks like assuming the pixels intensity range of the object and background has a certain degree of distinction. If the image has no distinct valley point in the histogram, this method would fail to separate the object and the background. The main disadvantage of this histogram-based thresholding method is the difficulties they meet when they have to identify the important peaks or valleys in the image used for segmentation and classification. In next paragraph, we would explore another main automated thresholding based on clustering.

The edge-based segmentations discussed in the previous subsection attempt to perform object boundaries extraction in accordance with the identified meaningful edge pixels. Region-based segmentations, on the contrary, seek to segment an image by classifying image into two sets of pixels: interior and exterior, based on the similarity of selected image features. In this subsection, we explore and study several classic methods belong to this category.

The region-based segmentation is based on the concept that the object to be segmented has common image properties and similarities such as homogenous distribution of pixel intensity, texture, and pattern of pixel intensity that is unique enough to distinguish it from other object (Gonzalez and Woods 2007). The

ultimate objective is to partition the image into several regions where each region represents a group of pixels belong to a particular object.

Another popular region method is seeded region growing; this method grows from seeds which can be regions or pixels; then, the seeds expand to accept other unallocated pixel as its region member according to some specified membership function (Kang et al. 2012).

In comparison with deformable model-based segmentation, region-based segmentation is considered relatively fast in terms of computational speed and resources. Besides, it is certain that segmentation output is a coherent region with connected edges. Simplicity in terms of concept and procedures is an advantage of region growing for immediate implementation.

Region-based segmentation is insensitive to image semantics; it does not recognize object but only predefined membership function. Besides, the design of the region membership is as difficult as setting a threshold value; region-based segmentation is unable to separate multiple disconnected objects simultaneously. The assumption that the region within a group of object is homogenous has low practical value in hand bone segmentation due to the fact that the bone is formed by cancellous bone and cortical bone that has high variations on texture and intensity range. Besides, in the presence of noise or any unexpected variations, region growing leads to holes or extra-segmented region in the resultant segmented region and thus has low accuracy in certain condition (Mehnert and Jackway 1997). The number and the location of seeds and membership function in seeded region growing, as well as the merging criteria in split-merge region growing, depend on human decisions which are subjective and laborious.

One of the famous region growing methods is the split and merge algorithm; split and merge is an algorithm splitting the image successively until a specified number of regions remain (Tremeau and Borel 1997). To perform the split and merge region growing algorithm, firstly, the entire image is considered within one region. Then, the splitting process begins in the region in accordance with the homogeneity criterion; if the criterion is met, then it splits (Gonzalez and Woods 2007). This splitting process repeats until all regions are homogenous. After the splitting process, the merging process begins. Initially, comparison among neighborhood regions is performed. Then, the region merges to each other according to some criterion such as the pixels' intensity value where regions that are less than the standard deviation are considered homogenous.

We have reviewed the essential concept of region-based segmentation. The purpose is to identify coherent regions defined by pixel similarities. The main challenge of this type of segmentation is often related to the pixel similarities: what are the features that should be adopted as similarities measurement and how are the thresholds of chosen features should be set in defining the similarity. The selection of features is difficult as they depend on application. For example, if the targeted object is not a connected object, pixel intensity is not suitable as pixel similarities measurement. The setting of threshold is another tricky challenge as it manipulates the trade-offs in terms of flexibility. For example, if the threshold is set too low, the inferior effect of over-segmentation occurs because pixels easily

surpass the threshold leading to larger coherent regions than the actual objects; if the threshold is set too high, the otherwise occurs. Region-based segmentation is unable to segment objects of multiple disconnected regions, and therefore, in the context of hand bone segmentation, applying only region-based segmentation is inappropriate as children hand bones for BAA involve different numbers of bones regions at different ages.

Deformable model refers to classes of methods that implement an estimated model of the targeted object using the model constructed by the prior information such as the texture and shape variability of the specific class of object as flexible two-dimensional curves or three-dimensional surfaces. In two-dimensional cases, these curves deform elastically to by satisfying some constraints to match the borders of the targeted object in a given image. The word ‘active’ stems primarily from the nature of the curves in adapting themselves to fit the targeted object. There are three main classes of deformable model: active contour model, active shape model, and AAM.

Deformable models assemble the mathematical knowledge from physics in limiting the shape flexibility over the space, geometry in shape representation, and optimization theory in model-object fitting. These mathematical foundations work together by playing their roles to establish the deformable model. For instance, the geometric representation with certain degree of freedoms is to cover broader shape changes; the principle in physics, in accordance with forces and constraints, controls the changes of shape to permit only meaningful geometric flexibility; optimization theory adjusts the shape to fulfill the objective function constituted by external energy and internal energy; the external energy is associated with the deformation of model to fit the targeted object due to external potential energy, whereas the internal energy constrains the smoothness of the constructed model in terms of internal elasticity forces.

Kass et al. (1988) proposed Active contour model or known as ‘snake’ as a potential solution to segmentation problem (Leymarie 1986). From the perspective of geometry, it is an embedded parametric curve represented as $v(s) = (x(s), y(s))^T$ on image plane $(x, y) \in R^2$, where $x(\cdot)$ and $y(\cdot)$ denote coordinates functions, and $s \in [0, 1]$ denotes the parametric domain. A snake in this context illustrates an elastic contour that fits to some preferred features in image.

To apply active contour model in segmentation, first, establish the initial location of point s in image planes adjacent to targeted object. These points collect ‘evidence’ locally in their territories and feedback to the contour energy. Next, search the update of each point using local information by solving the Euler–Lagrange equation when the contour is in equilibrium according to calculus of variation. Conventionally, numerical algorithm is applied to solve the equation in discrete approximation framework. Lastly, these steps repeat until stopping criteria has been achieved.

Since the active contour model is proposed, a lot of variations have been introduced by scholars. We have summarized some of them which are highly cited as following:

The advantages of active contour compared with previously discussed methods:

1. Process the image pixels in specific areas only instead of the entire image and thus enhance the computational efficiency.
2. Impose certain controllable prior information.
3. Impose desired properties, for instance, contour continuity and smoothness.
4. Can be easily governed by user by manipulating the external forces and constraints.
5. Respond to image scale accordingly with the assistance of filtering process.

Disadvantages of classical active contour model

1. Not specific enough to be implemented in specific problem as the shape of the targeted object is often not recognized by the algorithm.
2. Unable to segment multiple objects.
3. High sensitivity to environmental noises in image.
4. High dependency toward intensity gradient along the edges.
5. Do not consider the region information of the targeted objects.
6. High dependency on initial guessed point location. If the initial snake is not sufficiently close to targeted object boundaries, then points in snake can hardly attach the boundaries.
7. Difficult to grow into concavity.
8. Do not have a global shape controller that constraints the shape of contour from deviating from allowable shape of the targeted object.

2.5.1 Balloon Snake

Classical snake suffers from two drawbacks. The first drawback is that the contour model shrinks by searching a point of equilibrium based on the internal energy and boundary conditions if external energy is absent; the second drawback is that a contour model that is not close enough to the targeted object boundaries, the attraction of the model to the boundaries is very low. Balloon snake attempt to solve the problem by introducing the inflating force that makes the contour model behaves similarly as a balloon in two dimensions so that the contour model stops not at spurious edges by considering edges point extracted from Canny-Deriche edge detector (Cohen 1991; Cohen and Cohen 1993).

2.5.2 Level Set

It is first introduced by Osher and Sethian (1988) in the area of fluid dynamics, then being applied in computer vision for segmentation by Caselles et al. (1993) and Malladi et al. (1995). Later, this method has been incorporated with region

information and boundary information by Paragios and Deriche (1999). The level set methods, different from classical snake, model the contour in terms of implicit surface extracted from initial curve and then establish the connection between the curve propagation flow and implicit function deformation flow. The curvature and image gradient are then used to evolve the surface.

2.5.3 Active Contour Without Edges

Chan-Vese snake (Chan and Vese 2001) is an exceptionally popular extension of level set by combining level set method with Mumford-Shah functional segmentation technique (Mumford and Shah 1989). The prime utility of this snake is that it can detect objects even without using its gradient information. They minimize the energy in the level set formulation to evolve a curve attaching to boundary in such a way that the stopping term relies not on the boundary, the final result requires no smoothing procedure, and the initial contour need not to be around the targeted object.

2.5.4 Geodesic Snake

Classical snake often faces problems associated with incapability to detect multiple objects and incapability to detect interior and exterior boundaries simultaneously. An extension to level set snake, Geodesic snake (Caselles et al. 1997), with the implementation of geodesic computational approach, curve evolution theory, and geometric flows, improve the contour models so that they can split and merged without additional prior information or additional topology processes. This geodesic snake is then being upgraded by Leventon et al. (2002) by incorporating prior shape model. Firstly, analysis of variance of a set of shapes is performed. Then, maximum a posterior (MAP) position is estimated at every evolution step in curve. This extension has improved the robustness of boundary convergence in active contour despite noisy inputs.

2.5.5 Gradient Vector Flow Snake

The limited respond of classical snake in contour initiation and convergence associated with concavity has been addressed by GVF snake by applying an external field extracted from diffusion of gradient vectors derived from gray level of binary image (Xu and Prince 1997, 1998a, b). The main idea is to diffuse the forces to far situated contour model from the object by minimizing the an energy functional after solving two coupled partial differential equations. This is important to attract contour model that is initiated far from targeted object and can at the

same time resolve the problem of discontinuity of object boundaries. Besides, the GVF produces forces in large capture range to hence able to attract and progress contour into concavity.

To sum up, the regularizing terms adopted in active contour model is useful in stabilizing the contour, but the robustness is limited as the imposed constraints generally tend to smooth and shorten the contour unless stronger external energy is involved; this scheme is often too general and inadequate. Therefore, a more specifically designed scheme that capable of incorporating more finely tuned prior knowledge about the class of targeted object is required and is explained in next subsection.

Active shape model is a model founded on statistical theory where the variations of the shape of the objects can be captured via training procedure using labeled object's contour in the image in set points representation. Activating the trained contour will deform the contour fitting the targeted object in the image. Cootes et al. (1995) developed the model. Generally, it works by searching the best position of initial points that are surrounding the object and then updating these positions until the stopping criteria are achieved through iterations. Ever since the technique is proposed until recently, it has been extensively applied in various fields such as facial recognition (Xue et al. 2003; Zheng et al. 2008; Sukno et al. 2010), object tracking (Jang and Choi 2000; Kim and Lee 2005; Nuevo et al. 2011; Liu et al. 2012), and medical image processing (Smyth et al. 1997; Hodge et al. 2006; Aung et al. 2010; Toth et al. 2011).

The deformable models, ASM and AAM, undoubtedly are powerful segmentation methods. However, they are not without weaknesses and are not best method for automated hand segmentation. The reasons are summarized as following:

1. The landmarks placement has to be manually annotated by users. Incorrect landmarks placements lead to unreliable capture of shape variability.
2. The number of landmarks has to be specified by user manually. Insufficient landmarks lead to failure in obtaining the shape of the targeted structures; excessive landmarks lead to computational inefficiency.
3. The training phase requires a lot of training examples in database which is not necessarily available in many applications. Insufficient training examples lead to failure in generalizing the mean structure's shape.
4. The nature of hand bone development of children: different numbers and sizes of bones in different ranges of age complicated the implementation ASM and AAM especially in establishing the general form of mean shape.
5. The alignment phase is uncertain in terms of its numerical stability: the convergence of the mean model in the iterative method has not been devised mathematically and prone to errors.
6. The choice in retaining the number of eigenvectors in principal component analysis has to be determined correctly by user. Incorrect decision leads to failure in capturing the representative points of the shape; consequently, inaccurate model is constructed and leads to undesired segmentation result.

7. Variations in hand structural positions are often largely deviated and this devotes to nonlinear parameter relations that invariably impede the accurate segmentation as a whole.

ASM has been applied by Thodberg and Rosholm (2003) to address the problem of hand bone segmentation. Extensive training has to be done to complete the model in order to imitate the recognition understanding of human beings in segmenting the hand bone. Note that the initiation of set points placement to mark the spatial position of hand bone shape demands expert to be the operators. Both requirements of training set and human expert are the main weakness of this model in addressing the problem. It is tedious, subjective and time-consuming to delineate the shape from a large training set, not to mention the critical issue of the availability of these resources. Therefore, an alternative segmentation framework has to be established when the resources are limited and this motivates the research of this thesis.

AAM is a statistical model of shape and gray-level appearance of the targeted object proposed by Edwards et al. (1998). The final aim is to generalize the model to all valid example (Cootes et al. 1996). The relationship between the model parameter displacements and the errors between training example and a model instance is learned during the training phase. By computing the errors of fitting and using the previously obtained parameters, the current parameters with the intent of improving the current fitting can be updated.

AAMs and the closely related concepts are found in the methods of active shape model. The AAMs are most frequently being adopted in the application related to face modeling. Besides face modeling (Butakoff and Frangi 2010), it has been implemented in other applications as well such as in medical image processing (Roberts et al. 2007; Patenaude et al. 2011). The typical first step of AAM is to fit the AAM to an input image using model parameters that maximize the matching criteria between the model instance and the input image. The model parameters are then passed to a classifier to yield classification tasks.

To fit the AAM to an input image involves solving an nonlinear optimization problem. The conventional method of solving the problem is by updating the parameters iteratively. This update has to be incremental additive and the parameters refer to shape and appearance coefficients. The input image can be warped onto the model coordinate frame by using the current shape parameters estimations. The error between the model instance and the fitting of AAM onto the image can be computed. This error is then acted as feedback in next iteration that would affect the updates of the parameters. The constant coefficients in this linear relationship between the updates and errors can then be found either by linear regression or by other numerical methods.

Although the AAM appears to be the useful model-based system in medical image segmentation, it has constraints that impede its performance in practical application (Gao et al. 2010).

1. Low efficiency in real-time systems: current algorithm of AAM consumes a lot of time and space computational costs. Thus, it is of prime importance to minimize the complexities in time and space needed to perform the algorithm in order to realize it in real-time system. The efficiency is mainly affected by the following factors: manual landmarks placement, complex texture representation in high-resolution medical image, iterative procedure in solving the optimization problem.
2. Low discriminative ability for recognition and segmentation systems: only a group of object is being modeled, and thus it is considered as a generative model which possesses no ability to classify different objects. This ability depends on the accuracy of model fitting which are affected by how the prior shape is chosen; how the texture is represented; how the texture is modeled. It is crucial to improve this discriminative ability to perform segmentation tasks effectively.
3. Inconsistent robustness under different circumstances: the performance of the system is influenced by different conditions such as the existence of pose variations, uneven illumination, the absence of features, low resolution, and the presence of noises.

AAM is a very useful model as it can capture the mode of variations of deformable objects given a set of training examples. The mode of variations includes shape and texture as a whole. Besides, it can perform the projection of object onto low-dimensional subspace to reduce redundancy and capture main component of variations. Thus, it has been implemented in a lot of applications especially medical image segmentation. Nonetheless, it has limitation in efficiency, discriminative ability, and robustness in different condition. In the problem of hand bone segmentation, the same group of researchers that adopted ASM has extended their works by applying AAM (Thodberg 2002; Thodberg et al. 2010). The weaknesses discussed in applying ASM remain because the technical differences between AAM and ASM enhance only the robustness in terms of prior knowledge and the information around the object that have been incorporated into the model, not the practicality in terms of availability of training set and expert operators.

2.6 Desired Properties of Segmentation

Top-down strategy is adopted in designing the proposed segmentation framework. Firstly, the overview of the desired system is obtained through literature reviews by reviewing the existing techniques and analyzing the factors leading them to failure as effective hand bone segmentation technique. After gaining some insights into constituting a desired hand bone segmentation framework, we then identify the desired characteristics, only then we propose each sub-framework to satisfy each requirement.

The desired characteristics of hand bone segmentation framework adopted in computerized BAA should comprise the followings:

1. Contrast, illumination, orientation invariance: to ensure consistent segmentation robustness under different conditions of X-ray settings and devices.
2. Relatively low computational complexity: to ensure practical execution time for automated BAA system. Ideally, it is comprehensive enough to tackle with image complexities and uncertainties, yet it is simple enough to be executed in a reasonable time frame.
3. No complicated ‘training’ procedures: to ensure no dependency on availability of training samples of hand bone radiographs. However, simple parameter tuning procedures without depending on availability of training hand bone radiographs have to be established to capture the variations of uncertainties in image nature.
4. Utilization of prior knowledge: to ensure the usage of available information to optimize the result on hand bone segmentation. Besides, making use of ‘by-products’ of image preprocessing is preferable.
5. Relatively high resistance to noises: to ensure good performance of segmentation despite the inevitable random signals in the hand bone radiographs.
6. Automated or minimum dependency on human interventions: to ensure objectivity, to enable reproducibility, and to avoid laboriousness.
7. Consistent accuracy: to ensure relatively high precision in segmentation on resultant hand bone for subsequent processing in automated BAA system.
8. Resemblance to manual segmentation: to ensure a certain level of artificial intelligence in the designed algorithm to emulate human visual perception.
9. No overdependence on certain image feature: to ensure segmentation robustness under the absence of any certain property such as intensity discontinuity or edges.
10. Adaptability: to ensure robustness under the presence of variability in different regions of hand radiographs.
11. Optimality: all parameters are chosen based on the direction of finding the optimum solution and not arbitrarily preset. However, this criterion should not violate the second criterion.

To facilitate the subsequent explanations on our propose framework, henceforth, aforementioned desired property is referred as P1, P2, P3 ... and so forth. For example, the first property of contrast, illumination, orientation invariance is referred as P1 and the tenth criterion of adaptability is referred as P10.

References

- Acheson RM, Fowler G, Fry EI, Janes M, Koski K, Urbano P, Werfftenboschjj VA (1963) Studies in the reliability of assessing skeletal maturity from x-rays: part III. Greulich-Pyle Atlas and Tanner-Whitehouse method contrasted. *Hum Biol Int Rec Res* 35:317–349
- Aicardi G, Vignolo M, Milani S, Naselli A, Magliano P, Garzia P (2000) Assessment of skeletal maturity of the hand-wrist and knee: a comparison among methods. *Am J Hum Biol* 12(5):610–615

- Aja-Fernández S, De Luis-García R, Martín-Fernández MÁ, Alberola-López C (2004) A computational TW3 classifier for skeletal maturity assessment. *A Computing with Words approach*. *J Biomed Inform* 37(2):99–107
- Aung MSH, Goulermas JY, Stanschus S, Hamdy S, Power M (2010) Automated anatomical demarcation using an active shape model for videofluoroscopic analysis in swallowing. *Med Eng Phys* 32(10):1170–1179
- Baradez MO, McGuckin CP, Forraz N, Pettengell R, Hoppe A (2004) Robust and automated unimodal histogram thresholding and potential applications. *Pattern Recogn* 37(6):1131–1148
- Bernsen J (1986) Dynamic thresholding of grey-level images. *IEEE*, pp 1251–1255
- Buie HR, Campbell GM, Klinck RJ, MacNeil JA, Boyd SK (2007) Automatic segmentation of cortical and trabecular compartments based on a dual threshold technique for in vivo micro-CT bone analysis. *Bone* 41(4):505–515
- Butakoff C, Frangi AF (2010) Multi-view face segmentation using fusion of statistical shape and appearance models. *Comput Vis Image Underst* 114(3):311–321
- Cao F, Huang HK, Pietka E, Gilsanz V (2000) Digital hand atlas and web-based bone age assessment: system design and implementation. *Comput Med Imaging Graph* 24(5):297–307
- Caselles V, Catté F, Coll T, Dibos F (1993) A geometric model for active contours in image processing. *Numer Math* 66(1):1–31
- Caselles V, Kimmel R, Sapiro G (1997) Geodesic active contours. *Int J Comput Vision* 22(1):61–79
- Chan TF, Vese LA (2001) Active contours without edges. *IEEE Trans Image Process* 10(2):266–277
- Chemaitilly W, Sklar CA (2010) Endocrine complications in long-term survivors of childhood cancers. *Endocr Relat Cancer* 17(3):R141–R159
- Cohen LD (1991) On active contour models and balloons. *CVGIP: Image Underst* 53(2):211–218
- Cohen LD, Cohen I (1993) Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Trans Pattern Anal Mach Intell* 15(11):1131–1147
- Cootes TF, Page GJ, Jackson CB, Taylor CJ (1996) Statistical grey-level models for object location and identification. *Image Vis Comput* 14(8):533–540
- Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models-their training and application. *Comput Vis Image Underst* 61(1):38–59
- De Santis A, Sinisgalli C (1999) A Bayesian approach to edge detection in noisy images. *Circ Syst I: Fundamental Theory Appl*, *IEEE Trans* 46(6):686–699
- Edwards GJ, Taylor CJ, Cootes TF (1998) Interpreting face images using active appearance models. In: FG'98. Proceedings of the 3rd international conference on face and gesture recognition, IEEE computer society
- Gao X, Su Y, Li X, Tao D (2010) A review of active appearance models. *Syst, Man, Cybern, Part C: Appl Rev*, *IEEE Trans* 40(2):145–158
- Giordano D, Spampinato C, Scarciofalo G, Leonardi R (2010) An automatic system for skeletal bone age measurement by robust processing of carpal and epiphysal/metaphysal bones. *IEEE Trans Instrum Meas* 59(10):2539–2553
- Gonzalez R, Woods R (2007) *Digital Image Processing*, 3rd edn. Prentice Hall, Upper Saddle River
- Greulich W, Pyle S (1959) Radiographic atlas of skeletal development of hand wrist. *Am J Med Sci* 238(3):393
- Guo R, Pandit SM (1998) Automatic threshold selection based on histogram modes and a discriminant criterion. *Mach Vision Appl* 10(5–6):331–338
- Han C-C, Lee C-H, Peng W-L (2007) Hand radiograph image segmentation using a coarse-to-fine strategy. *Pattern Recogn* 40(11):2994–3004
- Heinrich UE (1986) Significance of radiologic skeletal age determination in clinical practice. *Die Bedeutung der radiologischen Skelettalterbestimmung für die Klinik* 26(5):212–215

- Hodge AC, Fenster A, Downey DB, Ladak HM (2006) Prostate boundary segmentation from ultrasound images using 2D active shape models: optimisation and extension to 3D. *Comput Methods Progr Biomed* 84(2–3):99–113
- Hsieh CW, Jong TL, Chou YH, Tiu CM (2007) Computerized geometric features of carpal bone for bone age estimation. *Chin Med J* 120(9):767–770
- Hsieh CW, Jong TL, Tiu CM (2008) Carpal growth assessment based on fuzzy description. In: *Soft computing in industrial applications, 2008. SMCia'08. IEEE conference on*. pp 355–358
- Huang Q, Gao W, Cai W (2005) Thresholding technique with adaptive window selection for uneven lighting image. *Pattern Recogn Lett* 26(6):801–808
- Hum YC (2013) Segmentation of hand bone for bone age assessment. Springer, London. Limited, 2013. ISBN: 9814451657, 9789814451659. 125 pp (SpringerBriefs in Applied Sciences and Technology Series)
- Jang D-S, Choi H-I (2000) Active models for tracking moving objects. *Pattern Recogn* 33(7):1135–1146
- Jong-Min L, Whoi-Yul K (2008) Epiphyses extraction method using shape information for left hand radiography. In: *Convergence and hybrid information technology, 2008. ICHIT '08. International conference on*, 28–30 Aug 2008, pp 319–326
- Jonsson K (2002) Fundamentals of hand and wrist imaging. *Acta Radiol* 43(2):236
- Kang C-C, Wang W-J, Kang C-H (2012) Image segmentation with complicated background by using seeded region growing. *AEU Int J Electron Commun* 66(9):767–771
- Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. *Int J Comput Vision* 1(4):321–331
- Kim W, Lee J-J (2005) Object tracking based on the modular active shape model. *Mechatronics* 15(3):371–402
- Lee SU, Yoon Chung S, Park RH (1990) A comparative performance study of several global thresholding techniques for segmentation. *Comput Vision, Graph, Image Process* 52(2):171–190
- Leventon ME, Grimson WEL, Faugeras O (2002) Statistical shape influence in geodesic active contours. *Computer vision and pattern recognition, 2000. In: Proceedings. IEEE conference on*. pp 316–323
- Leymarie FF (1986) Tracking and describing deformable objects using active contour models. Master thesis, McGill University
- Liu J, Qi J, Liu Z, Ning Q, Luo X (2008) Automatic bone age assessment based on intelligent algorithms and comparison with TW3 method. *Comput Med Imaging Graph* 32(8):678–684
- Liu Z, Shen H, Feng G, Hu D (2012) Tracking objects using shape context matching. *Neurocomputing* 83:47–55
- Liyuan L, Ran G, Weinan C (1997) Gray level image thresholding based on fisher linear projection of two-dimensional histogram. *Pattern Recogn* 30(5):743–749
- Luijendijk H (1991) Automatic threshold selection using histograms based on the count of 4-connected regions. *Pattern Recogn Lett* 12(4):219–228
- Mahmoodi S, Sharif BS, Chester EG, Owen JP, Lee REJ (1997) Automated vision system for skeletal age assessment using knowledge based techniques. *IEE*, pp 809–813
- Mahmoodi S, Sharif BS, Chester EG, Owen JP, Lee REJ (1999) Bayesian estimation of growth age using shape and texture descriptors. *IEE*, pp 489–493
- Mahmoodi S, Sharif BS, Graeme Chester E, Owen JP, Lee R (2000) Skeletal growth estimation using radiographic image processing and analysis. *IEEE Trans Inf Technol Biomed* 4(4):292–297
- Malladi R, Sethian JA, Vemuri BC (1995) Shape modeling with front propagation: a level set approach. *Pattern Anal Mach Intell, IEEE Trans* 17(2):158–175
- Manos G, Cairns AY, Ricketts IW, Sinclair D (1993) Automatic segmentation of hand-wrist radiographs. *Image Vis Comput* 11(2):100–111

- Manos GK, Cairns AY, Rickets IW, Sinclair D (1994) Segmenting radiographs of the hand and wrist. *Comput Methods Progr Biomed* 43(3–4):227–237
- Martin DD, Deusch D, Schweizer R, Binder G, Thodberg HH, Ranke MB (2009) Clinical application of automated Greulich-Pyle bone age determination in children with short stature. *Pediatr Radiol* 39(6):598–607
- Martin DD, Heckmann C, Jenni OG, Ranke MB, Binder G, Thodberg HH (2011) Metacarpal thickness, width, length and medullary diameter in children-reference curves from the First Zürich Longitudinal Study. *Osteoporos Int* 22(5):1525–1536
- Mehner A, Jackway P (1997) An improved seeded region growing algorithm. *Pattern Recogn Lett* 18(10):1065–1071
- Michael DJ, Nelson AC (1989) HANDX: a model-based system for automatic segmentation of bones from digital hand radiographs. *IEEE Trans Med Imaging* 8(1):64–69
- Mumford D, Shah J (1989) Optimal approximations by piecewise smooth functions and associated variational problems. *Commun Pure Appl Math* 42(5):577–685
- Niblack W (1990) An introduction to digital image processing. Prentice Hall, Upper Saddle River
- Niemeijer M, Van Ginneken B, Maas CA, Beek FJA, Viergever MA (2003) Assessing the skeletal age from a hand radiograph: automating the tanner-whitehouse method. In: *Proceedings of the 2003 SPIE medical imaging*. pp 1197–1205
- Nuevo J, Bergasa LM, Llorca DF, Ocaña M (2011) Face tracking with automatic model construction. *Image Vis Comput* 29(4):209–218
- Ontell FK, Ivanovic M, Ablin DS, Barlow TW (1996) Bone age in children of diverse ethnicity. *Am J Roentgenol* 167(6):1395–1398
- Osher S, Sethian JA (1988) Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J Comput Phys* 79(1):12–49
- Paragios N, Deriche R (1999) Geodesic active regions for supervised texture segmentation. In: *Proceedings of the international conference on computer vision*, vol 2, IEEE computer society. p 926
- Patenaude B, Smith SM, Kennedy DN, Jenkinson M (2011) A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage* 56(3):907–922
- Peloscsek P, Nemec S, Widhalm P, Donner R, Birngruber E, Thodberg HH, Kainberger F, Langs G (2009) Computational radiology in skeletal radiography. *Eur J Radiol* 72(2):252–257
- Pietka E (1994) Computer-assisted bone age assessment based on features automatically extracted from a hand radiograph. *Comput Med Imaging Graph* 19(3):251–259
- Pietka E, Kaabi L, Kuo ML, Huang HK (1993) Feature extraction in carpal-bone analysis. *IEEE Trans Med Imaging* 12(1):44–49
- Pietka E, Gertych A, Pospiech S, Cao F, Huang HK, Gilsanz V (2001) Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction. *IEEE Trans Med Imaging* 20(8):715–729
- Roberts M, Cootes T, Pacheco E, Adams J (2007) Quantitative vertebral fracture detection on DXA images using shape and appearance models. *Acad Radiol* 14(10):1166–1178
- Roche AF, French NY (1970) Differences in skeletal maturity levels between the knee and hand. *Am J Roentgenol* 109(2):307–312
- Sebastian TB, Tek H, Crisco JJ, Kimia BB (2003) Segmentation of carpal bones from CT images using skeletally coupled deformable models. *Med Image Anal* 7(1):21–45
- Sezgin M, Sankur B I (2004) Survey over image thresholding techniques and quantitative performance evaluation. *J Electron Imaging* 13(1):146–168
- Shafait F, Keysers D, Breuel TM (2008) Efficient implementation of local adaptive thresholding techniques using integral images. *Proceedings of SPIE* 6815, Document Recognition and Retrieval XC, 681510 (January 28, 2008)
- Shapiro L, Stockman G (2001) Computer vision. Prentice Hall, Upper Saddle River
- Sharif BS, Zaroug SA, Chester EG, Owen JP, Lee EJ (1994) Bone edge detection in hand radiographic images. In: *IEEE*. pp 514–515

- Smyth PP, Taylor CJ, Adams JE (1997) Automatic measurement of vertebral shape using active shape models. *Image Vis Comput* 15(8):575–581
- Somkantha K, Theera-Umpon N, Auephanwiriyaikul S (2011a) Bone age assessment in young children using automatic carpal bone feature extraction and support vector regression. *J Digit Imaging* 24(6):1044–1058
- Somkantha K, Theera-Umpon N, Auephanwiriyaikul S (2011b) Boundary detection in medical images using edge following algorithm based on intensity gradient and texture gradient features. *IEEE Trans Biomed Eng* 58(3 PART 1):567–573
- Sotoca JM, Iñesta JM, Belmonte MA (2003) Hand bone segmentation in radioabsorptiometry images for computerised bone mass assessment. *Comput Med Imaging Graph* 27(6):459–467
- Sukno FM, Guerrero JJ, Frangi AF (2010) Projective active shape models for pose-variant image analysis of quasi-planar objects: application to facial analysis. *Pattern Recogn* 43(3):835–849
- Tanner J, Whitehouse R (1975) Assessment of skeletal maturity and prediction of adult height (TW2 method)
- Tanner JM (2001) Assessment of skeletal maturity and prediction of adult height (TW3 method). W.B. Saunders, London
- Tanner JM, Gibbons RD (1994) Automatic bone age measurement using computerized image analysis. *J Pediatr Endocrinol* 7(2):141–145
- Thodberg HH (2002) Hands-on experience with active appearance models. pp 495–506
- Thodberg HH, Jenni OG, Caflisch J, Ranke MB, Martin DD (2009a) Prediction of adult height based on automated determination of bone age. *J Clin Endocrinol Metab* 94(12):4868–4874
- Thodberg HH, Kreiborg S, Juul A, Pedersen KD (2009b) The BoneXpert method for automated determination of skeletal maturity. *Med Imaging, IEEE Trans* 28(1):52–66
- Thodberg HH, Rosholm A (2003) Application of the active shape model in a commercial medical device for bone densitometry. *Image Vis Comput* 21(13–14):1155–1161
- Thodberg HH, Van Rijn RR, Tanaka T, Martin DD, Kreiborg S (2010) A paediatric bone index derived by automated radiogrammetry. *Osteoporos Int* 21(8):1391–1400
- Toth R, Tiwari P, Rosen M, Reed G, Kurhanewicz J, Kalyanpur A, Pungavkar S, Madabhushi A (2011) A magnetic resonance spectroscopy driven initialization scheme for active shape model based prostate segmentation. *Med Image Anal* 15(2):214–225
- Tran Thi My Hue MGS, Kim JY, Choi SH (2011) Hand bone image segmentation using watershed transform with multistage merging. *J Korean Instit Inf Technol* 9(5): 59–66
- Tremeau A, Borel N (1997) A region growing and merging algorithm to color segmentation. *Pattern Recogn* 30(7):1191–1203
- Tristán-Vega A, Arribas JI (2008) A radius and ulna TW3 bone age assessment system. *IEEE Trans Biomed Eng* 55(5):1463–1476
- Tsai D-M (1995) A fast thresholding selection procedure for multimodal and unimodal histograms. *Pattern Recogn Lett* 16(6):653–666
- Whatmough RJ (1991) Automatic threshold selection from a histogram using the “exponential hull”. *CVGIP: Graph Models Image Process* 53(6):592–600
- Xu C, Prince JL (1997) Gradient vector flow: a new external force for snakes. *IEEE*, pp 66–71
- Xu C, Prince JL (1998a) Generalized gradient vector flow external forces for active contours. *Sig Process* 71(2):131–139
- Xu C, Prince JL (1998b) Snakes, shapes, and gradient vector flow. *IEEE Trans Image Process* 7(3):359–369
- Xue Z, Li SZ, Teoh EK (2003) Bayesian shape model for facial feature extraction and recognition. *Pattern Recogn* 36(12):2819–2833
- Yan F, Zhang H, Kube CR (2005) A multistage adaptive thresholding method. *Pattern Recogn Lett* 26(8):1183–1191
- Zhang A, Gertych A, Liu BJ (2007) Automatic bone age assessment for young children from newborn to 7-year-old using carpal bones. *Comput Med Imaging Graph* 31(4–5):299–310

- Zhang J, Huang HK (1997) Automatic background recognition and removal (ABRR) in computed radiography images. *IEEE Trans Med Imaging* 16(6):762–771
- Zhao M, Yang Y, Yan H (2000) An adaptive thresholding method for binarization of blueprint images. *Pattern Recogn Lett* 21(10):927–943
- Zheng Z, Jiong J, Chunjiang D, Liu X, Yang J (2008) Facial feature localization based on an improved active shape model. *Inf Sci* 178(9):2215–2223

Advances in Medical Diagnostic Technology

lai, K.W.; Hum, Y.C.; Mohamad Salim, M.I.; Ong, S.-B.;
Utama, N.P.; Myint, Y.M.; Mohd Noor, N.; Supriyanto, E.
2014, VIII, 188 p. 37 illus., 20 illus. in color., Hardcover
ISBN: 978-981-4585-71-2