ABSTRACT

Title of Document:	X-RAY IMAGE SEGMENTATION AND AN INTERNET-BASED TOOL FOR MEDICAL VALIDATION
	Jing Cheng, Master of Science, 2006
Directed By:	Professor Ramalingam Chellappa Department of Electrical and Computer Engineering

Segmentation of vertebrae in X-ray images is a difficult task that requires an effective segmentation procedure. Noise, poor image contrast, occlusions and shape variability are some of the challenges in many of the spine X-ray images archived at the U.S. National Library of Medicine (NLM). In this thesis, we propose a curvature-based corner matching approach, which exploits the posterior corners of the vertebra to estimate the location and orientation of the vertebrae. The key advantage of the proposed approach is execution time, roughly about one-fifth of the previous approach that uses the generalized Hough transform when tested on a sizeable set of cervical spine images.

This thesis also presents the first ever effort to develop a prototype internet-based medical image segmentation and pathology validation tool, which enables radiologists to validate computer generated image segmentations, modify existing or create new segmentation in addition to identifying pertinent pathology data.

X-RAY IMAGE SEGMENTATION AND AN INTERNET-BASED TOOL FOR
MEDICAL VALIDATION

By

Jing Cheng

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2006

Advisory Committee:
Professor Ramalingam Chellappa, Chair
Professor K.J.Ray Liu
Professor Adrianos Papamarcou

Acknowledgements

Table of Contents

List of Tables

List of Figures

v

Chapter 1: Introduction

1.1 Motivation

The Lister Hill National Center for Biomedical Communications (LHNCBC), an intramural research and development division of the U.S. National Library of Medicine (NLM) maintains a collection of 17,000 digitized spine X-ray images of the cervical and lumbar spine in sagittal view from the second National Health and Nutrition Examination Survey (NHANES II). Figure 1.1 (a) and (b) show radiographic images of a cervical and a lumbar spine. The images come with over 2000 text fields of survey data that include demographic information, health questionnaire, and medical reports.

(a)	(b)

Figure 1.1 Illustration of (a) Cervical spine and (b) Lumbar spine in radiographic images

A key feature of spine disease manifested in these images is the presence of osteophytes which are bony processes that alter the shape of vertebrae [2]. The segmentation of vertebral bodies in such spine X-ray images is of great interest to bone morphometrists and radiologists since they exhibit pathologies such as anterior osteophytes (AO), disc space narrowing (DSN) [3], subluxation [4] and spondylolisthesis that are detectable consistently and reliably by vertebra boundary shape. Other pathologies such as vertebral fractures, ossification of posterior longitudinal ligament (OPLL), spinal stenosis caused by posterior osteophytes, tumors, and osteoporosis may also be detected from this dataset.

Advances in medical imaging have led to growth in large image collections. Hence, computer assisted techniques is becoming a more promising approach in medical image analysis. The feasibility of computer assisted techniques for the segmentation of vertebral bodies in spine X-ray images has been of great interest [1, 2] to biomedical researchers, in particular the osteoarthritic research community. Reliable extraction of vertebrae boundaries is a prerequisite for subsequent pathology validation and Content-Based Image Retrieval (CBIR) research. However, fully automated segmentation of spine X-ray images is a very challenging problem. Many published segmentation algorithms with results using our collection of spine X-ray images [5, 6, 7, 8, 9] depend on user intervention or some initial guess to achieve satisfactory performance. The techniques published in the research literature may be roughly categorized into methods based on Active Contour Segmentation [5], Active Shape Models [6, 7], and Generalized Hough Transform [8, 9]. In general, the quality of computer-assisted segmentation algorithms is affected by three important factors,

i.e. region of interest, image quality, and image size/resolution. Particular challenges regarding vertebra boundary segmentation in spine X-ray images may be outlined as follows:

- The region of interest, i.e., the vertebra has a wide range of shape, size and orientation.

- As X-ray images are usually of poor image quality, segmentation methods often confuse tissue and vertebra boundaries. In some cases, vertebra shows double edges as illustrated in Figure 1.2. The double edge is the distant (or nearer) edge of a 3D object seen in a projective image, such as an x-ray. This feature is critical for medical reviewers but very challenging for segmentation algorithm.

- Large image size: original size for cervical spine is 1462 by 1755; for lumbar spine, it is 2048 by 2488.

Thus developing a robust and efficient approach capable of robustly determining the location of the spine and the contour of the vertebrae is essential.

Besides assessing vertebral abnormality, the segmentation of vertebral bodies helps to find the contour of vertebrae to categorize vertebrae shapes, facilitate shape indexing and shape retrieval in the CBIR System. Hence image segmentation plays an important role in the CBIR System. The ultimate goal of the CBIR system is to develop algorithms for segmenting the vertebral bodies, to accurately measure geometric features of the individual vertebrae and inter-vertebral areas; and, by using training data provided by medical experts, to classify the spine anatomy into normal

and abnormal classes for conditions of interest, including anterior osteophytes and disc-space narrowing.
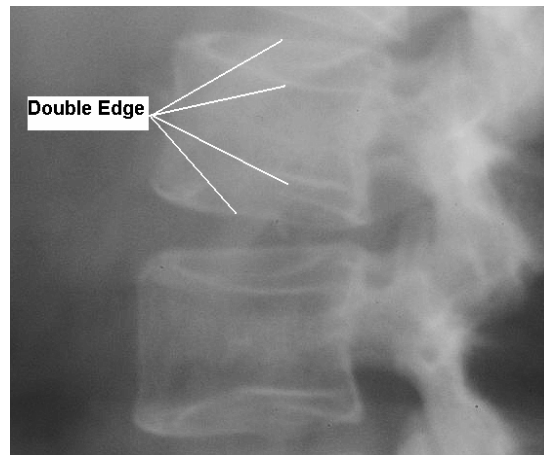
Figure 1.2. An example of cropped section from a lumbar X-ray showing two vertebrae exhibiting double edges.

The digitized X-ray images from NHANES II represent a research resource for bone morphometry and osteoarthritis. The goal is to make both the NHANES II image and the text data available to researchers and educators through a variety of visualization and retrieval tools accessible via the internet. An additional goal is to provide a comprehensive CBIR tool for such large biomedical image collections that permits indexing of these images by vertebral shape which is segmented through computer assisted segmentation technique, and their retrieval by specifying image features defining the anatomy, and in particular, the pathology of interest in them.

To provide ground truth for evaluating segmentation performance, a part of the image collection was delineated manually by medical experts. The shapes of the vertebrae are also validated to indicate type and severity of anterior osteophytes. In order to develop robust algorithms for segmentation and pathology detection, image indexing, and content-based image retrieval, it is necessary to obtain a large

4

collection of medically validated images from multiple experts. Our research goal [1] is to enable a comprehensive medical validation of the entire image set by cross validation from multiple geographically separated board-certified medical experts to include type and severity for all three pathologies mentioned above, and multi-resolution vertebral boundary segmentations.

1.2 Background

NHANES II was conducted by the National Center for Health Statistics during 1976-1980 and included participants aged 6 months to 74 years. For the NHANES II survey, the records contain information for approximately 20,000 participants. Each record contains about two thousand data points, including demographic information, answers to health questionnaires, anthropometric information, and the results of a physician's examination. In addition, approximately 10,000 cervical spine and 7,000 lumbar spine X-ray films were collected on survey participants aged 25-74. No X-rays were taken on pregnant women, and no lumbar X-rays were taken on women under 50. The pathologies of interest on these X-rays were osteoarthritis and degenerative disc disease.

Figure 1.3 (a) and (b) illustrates an example of applying segmentation technique to detect osteophyte, a bony protuberance on normal bone surface, which is a characteristic feature of degenerative joint disease of the spine. The boundary points of vertebra shape helps to determine the existence of superior and inferior anterior osteophytes (AO), which are pathologies along the anterior superior or inferior edges of the vertebral bodies. Points 2-3-8-7 describe a superior anterior osteophyte on a sagittal spinal X-ray and points 7-9-6-5 describe an inferior anterior osteophyte. In

such cases, only a small interval along the vertebral boundary is pertinent to the pathology.

<div align="center">(a)   (b)</div>
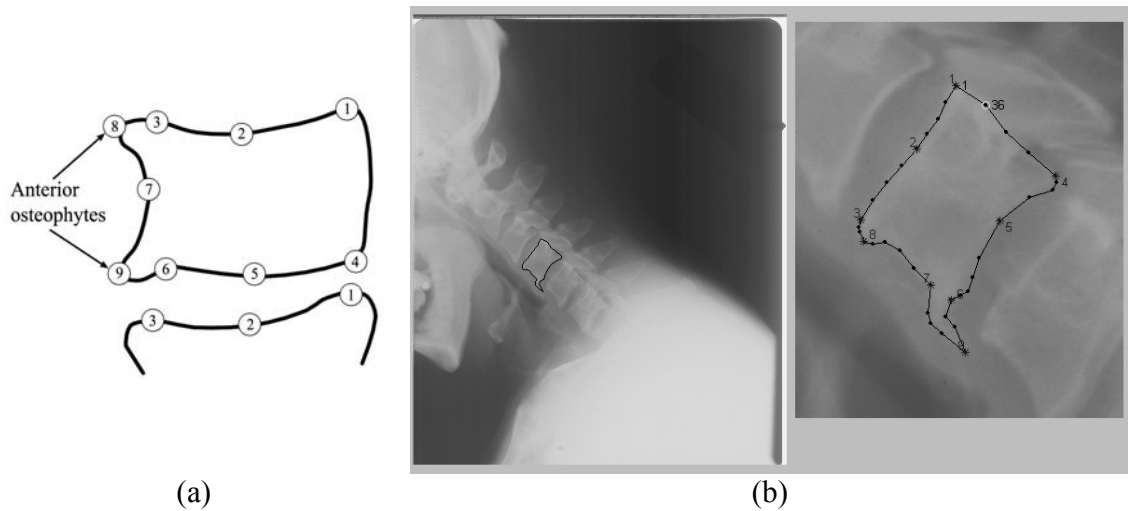
Figure 1.3. (a) Anterior Osteophytes illustrated on vertebral body outline. Points marked indicate regions of interest. (b) Spinal X-ray image showing segmented vertebra and a localized view showing inferior AO on vertebra with 36 boundary points superimposed.

Several methods have been developed in the past to segment the cervical and lumbar Vertebrae in radiographic images. Tagare [5] proposed an Active Contour Segmentation (ACS), which is a classical snake model combined with an initial contour created from a priori information and a search constraint on the contour. The algorithm minimizes an objective function by seeking a contour that maximized gradients along the normals to the contour, and minimized the contour length. The ACS handles one vertebra at a time and is semi-automatic. The accuracy of ACS depends closely on the initial shape contour, a template created by averaging manually-segmented vertebral shapes. As a result, the ACS needs user-intervention to achieve fine segmentation.

Zamora [6] implemented a customized version of Active Shape Models (ASM) for fine segmentation of the cervical and lumbar vertebrae. ASM, which was originally formulated by Cootes [7], is basically an optimization algorithm that requires an initial guess of the position, orientation and the scale of the vertebrae and this initialization is termed as the pose of the vertebra. Tezmol [8] proposed pose estimation or coarse segmentation of the cervical vertebrae using a customized version of the Generalized Hough Transform (GHT). The GHT is an efficient template matching technique, which is invariant to variations in scale and rotation and is also robust to occlusions. The GHT also requires prior knowledge of the shape to be segmented. This method gave acceptable segmentation results for the cervical vertebrae in terms of initialization for the customized ASM [6]. Gururajan [9] extended the customized version of the Generalized Hough Transform to segment lumbar vertebrae. The ASM and GHT methods automatically process multiple vertebrae at a time.

It is important to obtain medical validation data on the segmented boundaries of vertebral bodies and pathologies indicated by them. Collection of this data enables validation of retrieval performance and assists in the development of robust algorithms that reflect the typical queries posed by the targeted user community. Till recently the anterior osteophytes were our primary focus for shape-based retrieval in this project. Our interactions with a group of board certified radiologists have revealed that with use of image enhancement techniques, it may be possible to discover the existence of other pathologies such as spinal stenosis caused by posterior osteophytes. Correlating these pathology labels with segmented image features and

available participant health data can be used to build knowledge models and further enhance features of this multimedia database. These steps can also help lead the traditional CBIR research to a tool that can help further research and education in a particular field of medicine.

1.3 Thesis Organization

This thesis focuses on image segmentation and the development of a medical validation tool. The chapters are organized as follows. In Chapter 2, additional background on the characteristics of the spine X-ray image, as well as the pathology of interest is presented. Then, an investigation of various segmentation approaches for spine X-ray image segmentation including a shape operator, the active contour model, and steerable filters are presented. One major challenge is that the segmentation methods often confuse tissue and vertebra boundaries due to poor image quality. It has been shown that edge detection methods which have yielded very little promise of success in segmenting non-rigid, irregular, noisy vertebral shapes. In contrast, the model-based approach, i.e., snake, yields better results but depends more on user intervention. For each segmentation method investigated, we summarize the concept, main features, level of user interventions, and results obtained. Then, the proposed curvature-based corner matching approach for coarse segmentation of cervical vertebrae images is discussed. Chapter 3 discusses the performance of the proposed approach, and compares it with two previous methods in terms of execution time. Chapter 4 discusses the development of a medical validation tool that operates over the internet. It discusses the design considerations, features, capabilities, and technical

challenges faced when developing the tool. Finally, conclusions and future perspectives are drawn in Chapter 5.

Chapter 2: Segmentation of Spine X-ray Image

This chapter presents algorithms for segmentation of cervical and lumbar vertebrae images. In this research, segmentation is the process of locating and extracting the shape of the vertebra in an image. First, the spine vertebrae and pathology interests are introduced, and the objectives of segmentation algorithm are discussed. This is followed by a brief review of the previous methods for the segmentation of cervical vertebrae. Then various general segmentation approaches including those based on the shape operator, active contour model, and steerable filters are discussed and applied to our X-ray image collection. Pertinent issues are brought up, and discussed as well. Finally the current approach is introduced with a flow diagram and then the individual blocks of the segmentation process are taken up and discussed in detail.

2.1 Spine Vertebra and Pathology Interests

Before going into the details of the image segmentation process, it would also be of interest to examine the nomenclature for the naming of the vertebra, and the region of interest for segmentation.

The spine is made up of vertebrae. Figure 2.1 shows the human spine, and the nomenclature for naming the different parts of the spine. At the top are the cervical vertebrae. There are 7 of them and they are referred to as C1-C7, starting at the top. The seventh one joins to the first of the thoracic vertebrae. These are the 12 that run down the back and provide a place for our ribs to attach. They are referred to as T1-T12, again from the top down. The lower inward curve of our back is made up of the

five lumbar vertebrae. They are called L1-L5. Below that comes a set of 5 fused vertebrae called the sacrum that lies between the hip bones. Lastly comes the coccyx or the tailbone, another set of fused vertebrae. The vertebrae join together at what are called the facet joints. Between each of the vertebrae are the disks, which provide cushioning and act as shock absorbers. Traveling down the center of the spine is the spinal cord and at each of the vertebrae, nerves branch out through what are called the foramen to the rest of the body.

The regions of interest for segmentation are the cervical vertebrae (C3 to C6) and the lumbar vertebrae (L1 to L5). The reason for leaving out C1, C2 and C7 is because these structures are often not visible on the radiograph and hence it becomes a tough task to characterize them. The ground truth for the vertebra shape is a set of coordinate points present on and around the vertebra for a subset of 1800 images. This data was generated manually by radiologists, and both 9 and 36 points were collected per vertebra. Typically, for the cervical spine images, these included the vertebrae boundaries from C3 to C6 and for the lumbar images they spanned the boundaries from L1 to L5. For a subset of the cervical spine images, a special point marking the center of gravity of the C1 vertebra was collected additionally. These points are termed as morphometric points.

The problem with the 9-point shape contour descriptor is its inability to fully represent the shape variation in the corners of the vertebrae. Therefore, 36-point or variable-point shape contour descriptors are more widely used in image segmentation research. With the morphometric points as the ground truth, the performance of computer assisted segmentation techniques can be evaluated. The segmentation

techniques generally place significantly more points per vertebra than the number of morphometric points, and provide a better representation of the shape variations. These points are termed as landmark points (LPs).
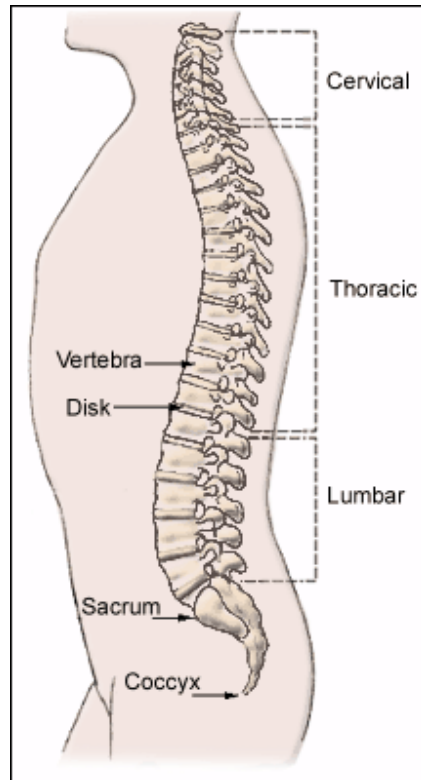
Figure 2.1 Human spine nomenclature

2.2 Investigation of Contemporary Segmentation Approaches

2.2.1 Shape Operator

Moon and Chellappa [10] proposed an approach to accurately detect 2-D shapes. In their work, the cross section of the shape boundary is modeled as a step function. A shape operator is obtained by extending the double exponential (DODE) filter [11] along the shape's boundary contour; the responses are accumulated at the centroid of the operator to estimate the likelihood of the presence of the given shape.

In fact, using a shape operator to detect the shape is a natural extension of the task of edge detection at the pixel level to the problem of global contour detection. This simple filtering scheme has been applied to the problems of vehicle detection in aerial images [10, 29], and human facial feature detection [10]. We applied this algorithm to detect vertebrae. It has been shown that the shape operator can successfully detect rigid shapes, such as vehicle (rectangular shape) and human eyes (eclipse shape). However, it failed to detect the non-rigid, irregular, noisy vertebral shapes. Noisy X-ray images produced many weak, small and discontinuous edges inside, outside, and along the actual boundary. Therefore false shape detection occurs.

2.2.2 Active Contour Segmentation (ACS)

Active contours, or snakes [12, 13, 14], are computer-generated curves that move within images to find object boundaries. The 3D version is often known as 3D deformable models or active surfaces [30, 31] in the literature. Active contours are often used in computer vision and image analysis to detect and locate objects, and to describe their shape. Tagare [5] proposed an Active Contour Segmentation (ACS), which is a classical snake model combined with an initial contour created from a priori information and a search constraint on the contour. The algorithm minimizes an objective function by seeking a contour with maximized gradients along orthogonal curves to the contour, and minimized the contour length.
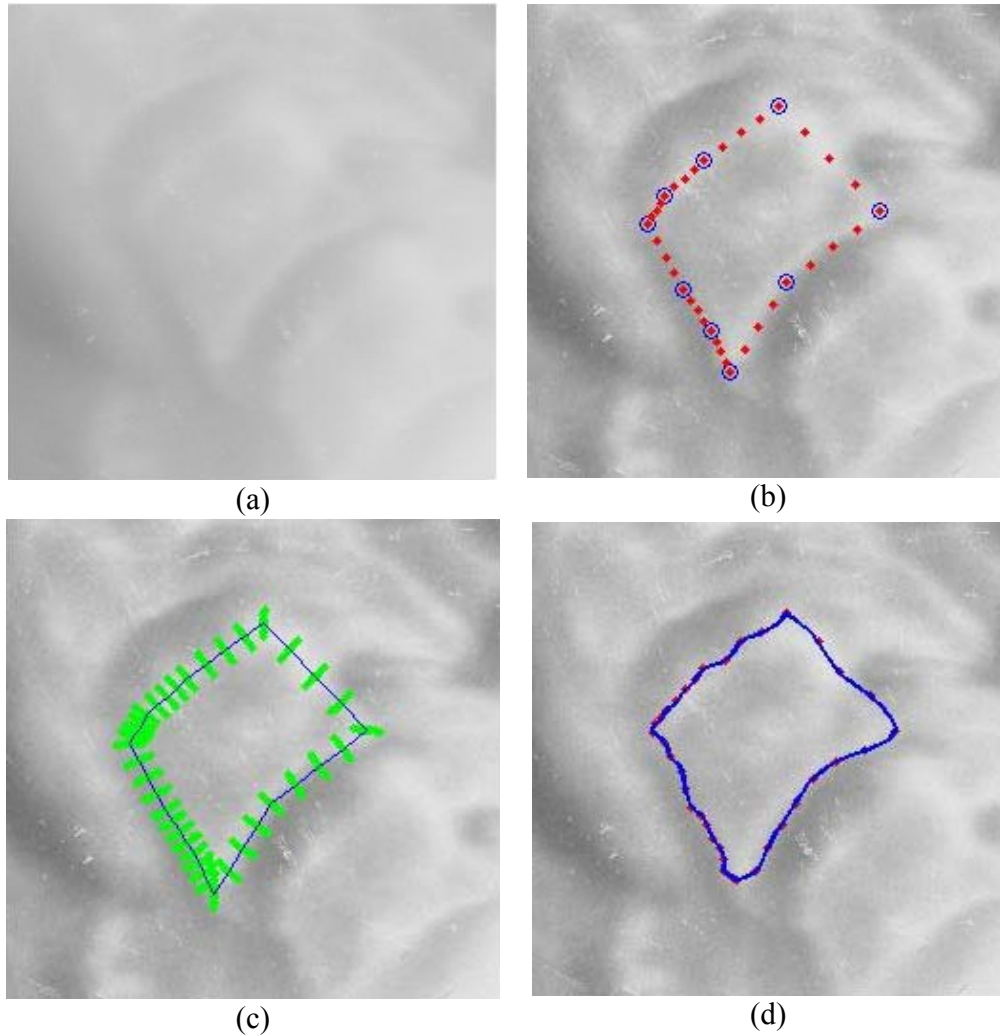
Figure 2.2. (a) A section of the original X-ray image of cervical vertebra. (b) Initial contour created by inputting 9 points (in blue circle) along the edges of the vertebral bodies. (c) Orthogonal curves in defining the search grid used in Active Contour Segmentation. (d) Final contour after deformation. The image section has been enhanced to show the details in (b), (c) and (d).

Tagare's ACS [5] is implemented as a semi-automatic algorithm, and processes one vertebra at a time. The algorithm constrains the solution to lie on a grid between an 'inner contour', inside the template, and an 'outer contour', outside the template. The 'inner contour' and 'outer contour' are user-defined. A novel feature of this algorithm is that the search space is the orthogonal curves to the contour,

14

calculated by numerically solving a boundary-value partial differential equation. The reason for implementing a search grid of orthogonal curves is to avoid having self-intersecting normal line segments to the contour where the vertebra has a narrow protrusion.

<center>(a)   (b)</center>

<center>(c)</center>
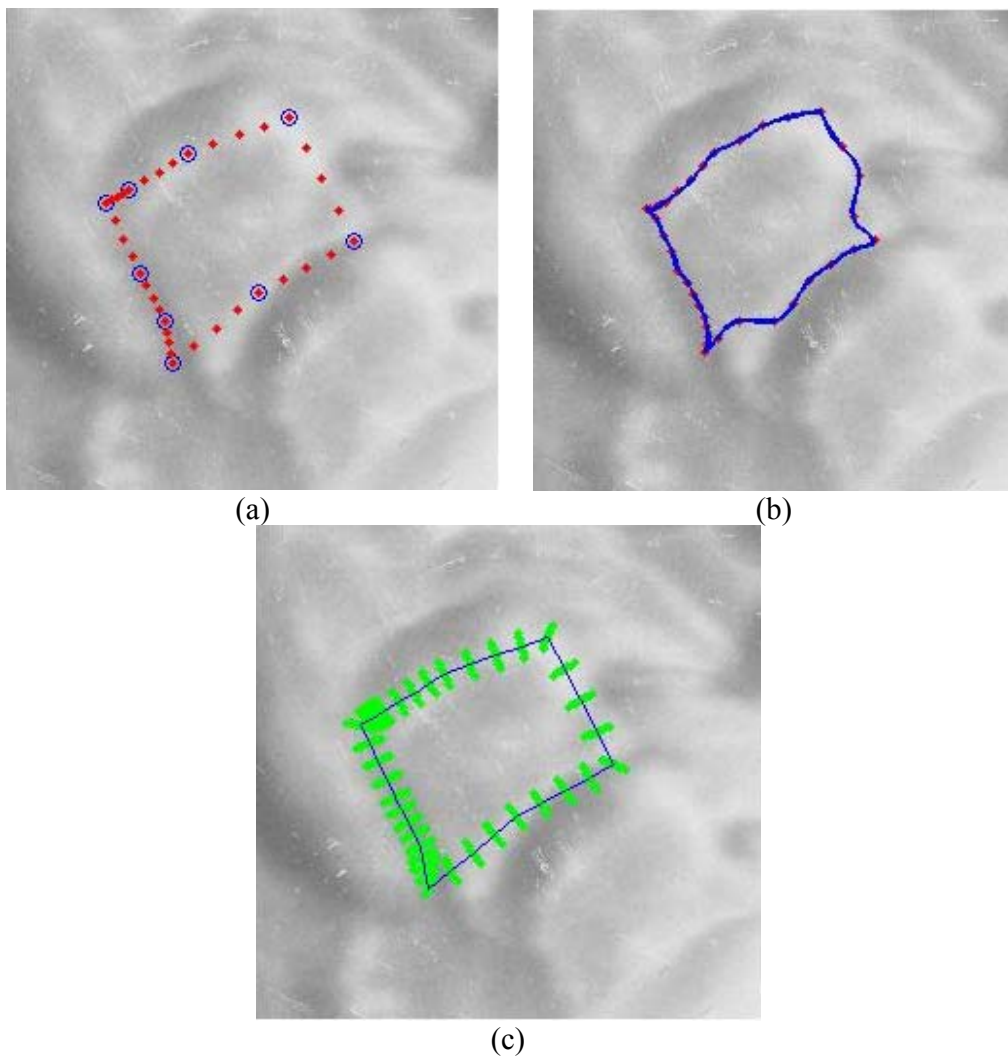
Figure 2.3. (a) Initial contour created by inputting 9 points (in blue circle) along the edges of the vertebral bodies. (b) Final contour after deformation. (c) Orthogonal curves in defining the search grid used in Active Contour Segmentation. The image section has been enhanced to show the details.

It has been shown that the accuracy of ACS depends closely on the initial shape contour. Figure 2.2 shows the results when using the 9-point shape template as the initial guess. The 9-point shape template is based on Figure 1.3(a), and created manually by inputting 9 points along the edges of the vertebral bodies. Figure 2.2(b) illustrates the initial contour where the user manually inputs 9 points depicted in blue circles, and then the ACS generates more points uniformly in between the 9 landmark points depicted in red dots. Points on the solution contour are constrained to lie on the non-intersecting curves shown in Figure 2.2(c). And Figure 2.2 (d) shows the final contour after deformation. It can be seen that the final contour convincingly captures small variations in the shape of vertebra. However, an inaccurate initial contour provided by the user, can result in bad segmentation in many cases. An example of incorrect initial contour, which is rotated by an angle from the actual vertebra shape, is shown in Figure 2.3(a). The final contour is unable to capture the shape of vertebra. In particular, it misses the inferior anterior and posterior corners of the vertebra as shown in Figure 2.3(b).

Four other active contour models have been investigated, including the traditional snake [12], the snake using a local minimization algorithm [13], template deformation along the orthogonal curves of contour [5], and the snake using Gradient Vector Flow [14]. In general active contour models (also know as snakes) represents an object boundary as a parametric curve that is allowed to deform from some arbitrary initial shape towards the desired final shape. The problem of finding this final contour is cast as an energy minimization problem. Ideally, the energy function of the contour attains a minimum when the contour is spatially aligned with the

boundary of the object in the image. A summary of commonly used energy function

is presented in Table 2.1.

Table 2.1. A summary of snake implementation

	Major terms in energy function	Level of interaction
Traditional snake	Internal energy (elastic force and bending force) External energy (gradient energy)	low
Snake using local minimization algorithm	Internal energy (continuity energy and balloon force) External energy (gradient energy and intensity energy)	low
Template deformation	Internal energy (proximity energy and smoothness energy) External energy (gradient energy)	high
Gradient Vector Flow	Internal energy (elastic force and bending force) External force (gradient vector flow)	low

In this thesis, we implemented snake-based methods using a local search

algorithm [13], template deformation [5], gradient vector flow [14], and applied the

techniques to the spine X-ray images. Simulation results show that the template

deformation [5] method finds the best final contour, but need the most human

interaction. In fact, the user needs to place 4 points marking the four corners of the

vertebra, 2 points marking the median along the upper and lower vertebra edge, 1

point marking the median along the anterior vertical edge of the vertebra, and

additional 2 points marking the upper and lower anterior osteophytes if present. So

totally 9 points are needed to be placed initially if osteophytes present, and 7 points

need to be placed otherwise. Knowing the initial positions of the 9 landmark points,

an arithmetic interpolation approach is used to generate 36 control points by adding 3

intermediate points between two original landmark points. Then the deformation

algorithm is applied. Since the contour is deformed by restricting every control point

to move only along its orthogonal curves, the optimal deformation cannot be large. The other three methods need low user interactivity, (user only provides the approximate location of the vertebra center), but they cannot find the actual contour.

Low image quality is the main reason for poor behavior of snakes as noisy X-ray images produced many weak, small and discontinuous edges inside, outside, and along the shape boundary. These edges can attract the contour points to local minima locations from which the snake cannot escape. Even worse, it has been found that the magnitude of gradient at some area inside the vertebra is greater than the magnitude of gradient on the actual boundary, leading to termination of contour deformation at these areas; as a result, the snake often misses the actual boundary of vertebra. The snake using the Gradient Vector Flow [14] performs the worst. Instead of using the gradient force as the external force, this method defines the external force as the gradient vector flow which features a larger capture range at the actual boundary as well as the areas with false edge. A larger capture range does not help in this case.

Generally speaking, user intervention and some prior knowledge of the contour lead to finer segmentation. However inaccurate position, scale, and rotation of the initial template can result in bad segmentation in many cases.

2.2.3 Generalized Hough Transform (GHT)

Tezmol [8] proposed pose estimation or coarse segmentation of the cervical vertebrae using a customized version of the Generalized Hough Transform. The GHT is an efficient template matching technique, which is invariant to variations in scale and rotation and is also robust to occlusions. The GHT also requires prior knowledge of the shape to be segmented. This method gave acceptable segmentation results for

18

the cervical vertebrae in terms of initialization for the customized ASM [6], which was originally formulated by Cootes [7]. GHT is basically an optimization algorithm that requires an initial guess of the position, orientation and the scale of the vertebrae and this initialization is termed as the pose of the vertebra. There have also been other efforts to customize the GHT for this application. Gururajan [9] extended the customized version of the Generalized Hough Transform to segment lumbar vertebrae.

It can be shown that the GHT is effective for finding the approximate location of the target vertebrae or Region of Interests (ROI) at the first step [6, 8, 9]. The basic GHT formulation for detection of arbitrary shapes in images was given by Ballard [15], and the effectiveness of GHT for the NHANES II images has been extensively investigated by Tezmol and Gururajan [8, 9].

The GHT uses a lookup table termed the R-table for an arbitrary shape, so no analytical description for the shape is necessary. Figure 2.4 illustrates the geometry for building the R-table and the format of R-table. $P_r(x_r, y_r)$ is a reference point which is the origin of an axis system fixed in the template shape. An arbitrary point on the template boundary $P_i(x_i, y_i)$ is specified by equation (2.1) and (2.2).

$$r = \sqrt{(y_i - y_r)^2 + (x_i - x_r)^2} \qquad (2.1)$$

$$\alpha = \tan^{-1}\left(\frac{y_i - y_r}{x_i - x_r}\right) \qquad (2.2)$$

Where r is the Euclidean distance from the reference point to the boundary point, and α is the angle of the line connecting $P_i(x_i, y_i)$ and $P_r(x_r, y_r)$. The pairs of (r, α) are then indexed by local edge direction angle θ_i, which is determined by the

intersection of the tangent line through $P_i(x_i, y_i)$ and the horizontal axis. The format of the R-table is then defined in terms of discrete value θ_i and the corresponding (γ_i, α_i) pairs, as illustrated in Figure 2.4(b).

θ_i	(r, α)
$\Delta\theta$	(r_1, α_1), (r_2, α_2)
$2\Delta\theta$	(r_3, α_3)
$3\Delta\theta$	(r_4, α_4), (r_5, α_5), (r_6, α_6)
…	…

(a)   (b)
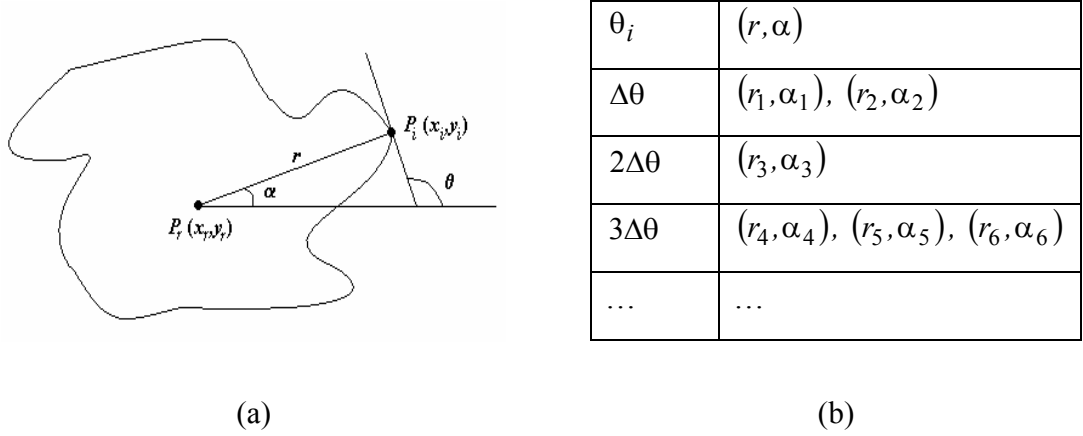
Figure 2.4 Generalized Hough Transform: (a) Illustration of the geometry for building the R-table. (b) The R-table format

The GHT methods in [8, 9] segment multiple vertebrae at a time. Figure 2.5 shows the template for the bottom part of vertebra C2 and the whole front shape of vertebrae C3-C6 with minimum bounding box in red rectangle. Although theory does not set constraints for the location of the reference point, a common practice is to choose P_r to be inside the template. In our study, P_r is positioned as the center of the minimum bounding box that encloses the template [16]. The left-most and right-most sides of the bounding box are vertical lines passing through the minimum and maximum x-coordinate of the template. While the top-most and bottom-most sides of the bounding box are horizontal lines passing through the minimum and maximum y-coordinate of the template respectively.
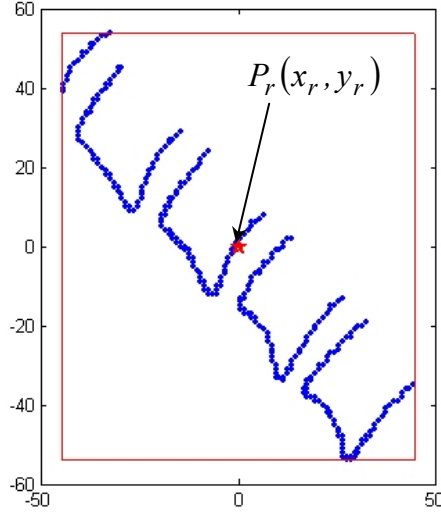
Figure 2.5. Example cervical spine shape template with location of the reference point marked in the template. The minimum vertical bounding box is depicted in red rectangle.

The GHT is a template matching method. Matching is done up to position, orientation and scale, but the shape of the template does not deform. If the template is rotated through an angle φ, relative to an image-fixed coordinate system, and scaled by a quantity s, then the relationships between a boundary point and the reference point are specified by equation (2.3) and (2.4).

$$x_i = x_r + s \cdot r_i \cdot \cos(\alpha_i + \varphi) \tag{2.3}$$

$$y_i = y_r + s \cdot r_i \cdot \sin(\alpha_i + \varphi) \tag{2.4}$$

$$\begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f(x,y)}{\partial x} \\ \dfrac{\partial f(x,y)}{\partial y} \end{bmatrix} \tag{2.5}$$

In a given image, the gradient direction is computed as $\tan^{-1}\left(\dfrac{G_y}{G_x}\right)$, where (G_x, G_y) is the gradient vector of the image $f(x,y)$ at (x_i, y_i) and defined in equation (2.5).

G_x and G_y can be computed using gradient operators, such as Canny [26], Prewitt [32] or Sobel operators [33]. And the edge direction differs from gradient direction by $\pi/2$.

Using the predefined template, the GHT algorithm tallies votes for uniquely identified bin (x_r, y_r, s, φ) in the 4-D Hough space corresponding to the candidate boundary point (x_i, y_i). At the end of all processing, the bin with the largest number of votes identifies the position, scale, and rotation of the template in the image. The candidate boundary points in the image for which we compute the Hough bins are edge points that we have detected in the image; hence the detection of edges in the images is crucial to the performance of the algorithm. More details on image enhancement and edge detection will be discussed in sections 2.3.2.

In recent research, the cervical spine vertebrae segmented are bottom of C2 through C5 or C6, while for lumbar spine the vertebrae are L1–L5. The template is usually created by finding the mean of the training set of manual segmentations. In [9], Gururajan carried out a measure of performance by creating a bounding box which closely fit the spine template, and computes the number of ground truth boundary points lying within the template's bounding box when the template had been placed on an image by the GHT algorithm. The mean percentage and standard deviation percentage of points within the bounding box were also computed. However, these numbers alone give little insight into the usefulness of the GHT even for approximate segmentation of the images. In fact, the mean percentage of points within the bounding box can be higher than 82%, but the GHT still can misplace the template or slide along the vertebral column by one vertebra in some cases. That is,

instead of being placed correctly on C2–C6, the template may be placed on C3–C7, for example. Figure 2.6 shows one of the cases of misplacement, and one of the examples where GHT outputs good segmentation. Many of these misplacements were determined to correspond to failure of the algorithm to properly distinguish the correct template bin in the Hough space from among the top three bins.

(a)   (b)
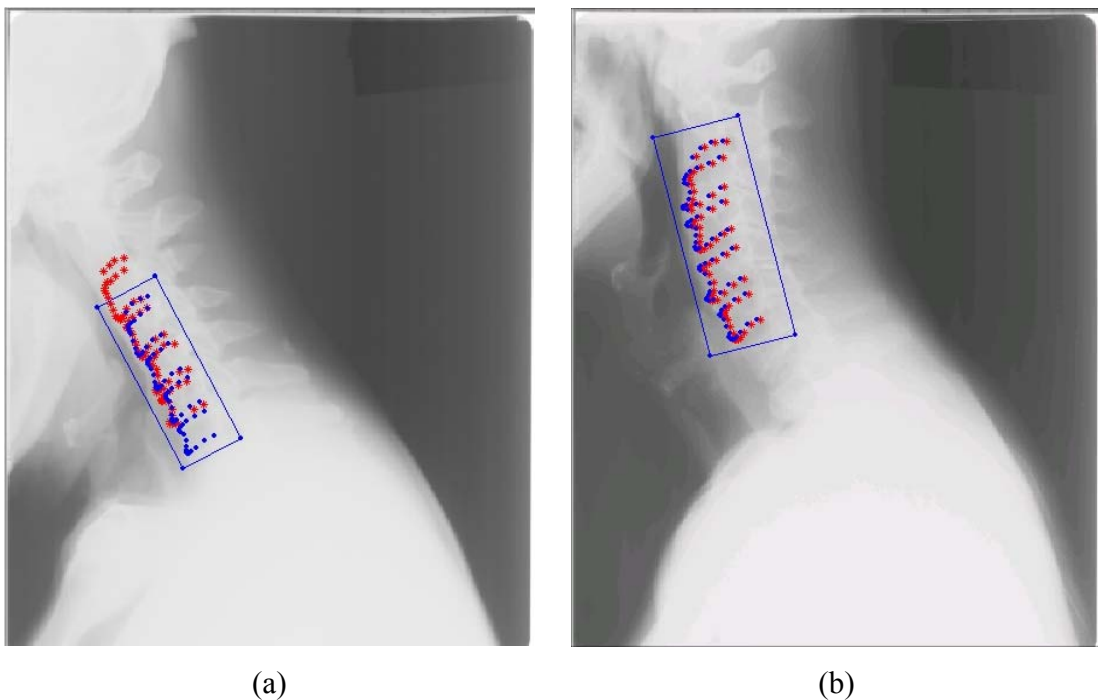
Figure 2.6. Example of GHT results: (a) misplaced the template by one vertebra down, (b) placed the template approximately close to ground truth. Output of GHT is denoted in blue marks, and the ground truth landmark points (LPs) is in red.

Like the ACS or snake, the performance of GHT can be improved by returning the template placements corresponding to the top three bins, and allowing the user to choose among these best three candidates. When the user was allowed to consider these top three candidates, more GHT results were judged acceptable [9]. However, for GHT, computation time is an issue. For example, on a 2.1 GHz AMD microprocessor using MATLAB, the processing time for cervical spine image was 55

seconds; for lumbar spine image, 110 seconds, even after the images were reduced to one quarter of their original size horizontally and vertically (original size for cervical image was 1462 by1755; for lumbar image, 2048 by 2488) shown in [9]. Moreover, the GHT allows template matching up to position, orientation and scale, but does not allow non-rigid shape deformation. This makes GHT unable to capture the prominent variation of vertebra boundary, for example when severe osteophytes are present. The GHT fails to achieve the fine segmentation for detecting and representing pathology or abnormality of vertebrae shapes. In general it can be used for coarse segmentation or finding the orientation of the spine.

2.2.4 Steerable Filter

In the literature, oriented filters are used in many vision and image processing tasks, such as image enhancement, edge detection, texture analysis, motion analysis, and image data compression. The concept of steerable filter was introduced in [23] and extended in [24]. Freeman et al. [23] first proposed to synthesize filters of arbitrary orientations from linear combinations of basis filters. It is useful to allow to "steer" a filter to arbitrary orientation under adaptive control, and to examine the filter output as a function of both orientation and phase. As a result, steerable filters provide directional edge detection as they behave as band-pass filters in a particular orientation [23]. The edge located at different orientations in an image can be detected by splitting the image into orientation sub-bands obtained by basis filters having these orientations. The steering constraint [23] is stated as

$$h^{\theta_a}(x, y) = \sum_{i=1}^{M} K_i(\theta_a) \cdot h^{\theta_i}(x, y) \qquad (2.6),$$

24

where the basis filters are represented by $h^{\theta_i}(x,y), 1 \le i \le M$, which are rotated versions of impulse response $h(x,y)$ at filter orientation θ_i, and $K_i(\theta_a), 1 \le i \le M$, are interpolation functions which control the filter orientations. In designing and implementing steerable filters, we want to know what functions $h(x,y)$ can satisfy equation (2.6), how many terms, M, are required in the sum, and what the interpolation functions, $K_i(\theta_a), 1 \le i \le M$ are.

Freeman et al. [23] point out that steerable filters work with discretely sampled functions as well as continuous ones. They also give several sample applications, such as determining dominant local orientations, filtering noise on angularly aligned features, contour detection, and shape-from-shading computations. Freeman also [23] showed an example of applying steerable filters to vessel image enhancement. The second derivative of the Gaussian (G2) along any direction is expressed as a combination of the derivatives at orientations 0, 60, and 120 degrees. The G2 filter is steered adaptively along the directions of dominant vessel orientation. The result is less noisy than the output of an isotropic filter of the same frequency passband. It is computed at a single scale.

To the best of our knowledge, no previous work on steerable filters has investigated their applications to the digitized spine X-ray image collection. Using an edge detector operator in combination with steerable filters usually achieves higher performance in boundary detection compared to using the edge detector directly on the image because steerable filter can remove noise and enhance oriented structures by angularly adaptive filtering. Thus we propose to use steerable filters to enhance oriented structures, and incorporate it into the proposed segmentation scheme to

detect the dominant orientation of the spine. In fact, the region of interests of spine for segmentation is the cervical vertebrae (C3 to C6) and the lumbar vertebrae (L1 to L5). In most cases, the anterior and/or posterior edges of the vertebrae structure approximate parallel lines except at the corner. The idea is to detect the corners of the vertebrae, match the corners to prior shape template, and then use snakes for finer segmentation. The proposed method will be discussed in the next section.

2.3 A Proposed Image Segmentation Scheme

2.3.1 Scheme Outline

In this section, we propose a curvature-based corner matching approach for coarse segmentation of cervical vertebrae and discuss some issues in this approach. The flow chart of the segmentation algorithm is illustrated in Figure 2.7, where the target image is taken to be the radiographic cervical spine or lumbar spine, and the target shape to be segmented is the cervical vertebrae C3-C6 or lumbar vertebrae L1-L5. The template of the shape is also shown in the figure. The first step determines a bounding box around the vertebrae. This is a significant step because once we have a bounding box, our domain of operation decreases in size and hence the total computational intensity involved in image enhancement, edge detection, corner detection and template matching decrease. This region of interest is then passed to the image enhancement module. The subsequent step is the detection of edges followed by corner detection. Then the detected corners are matched to a shape template so that non-vertebral corners can be identified and removed. The output image shows the corner points overlaid on the enhanced image. Using corner matching, the

26
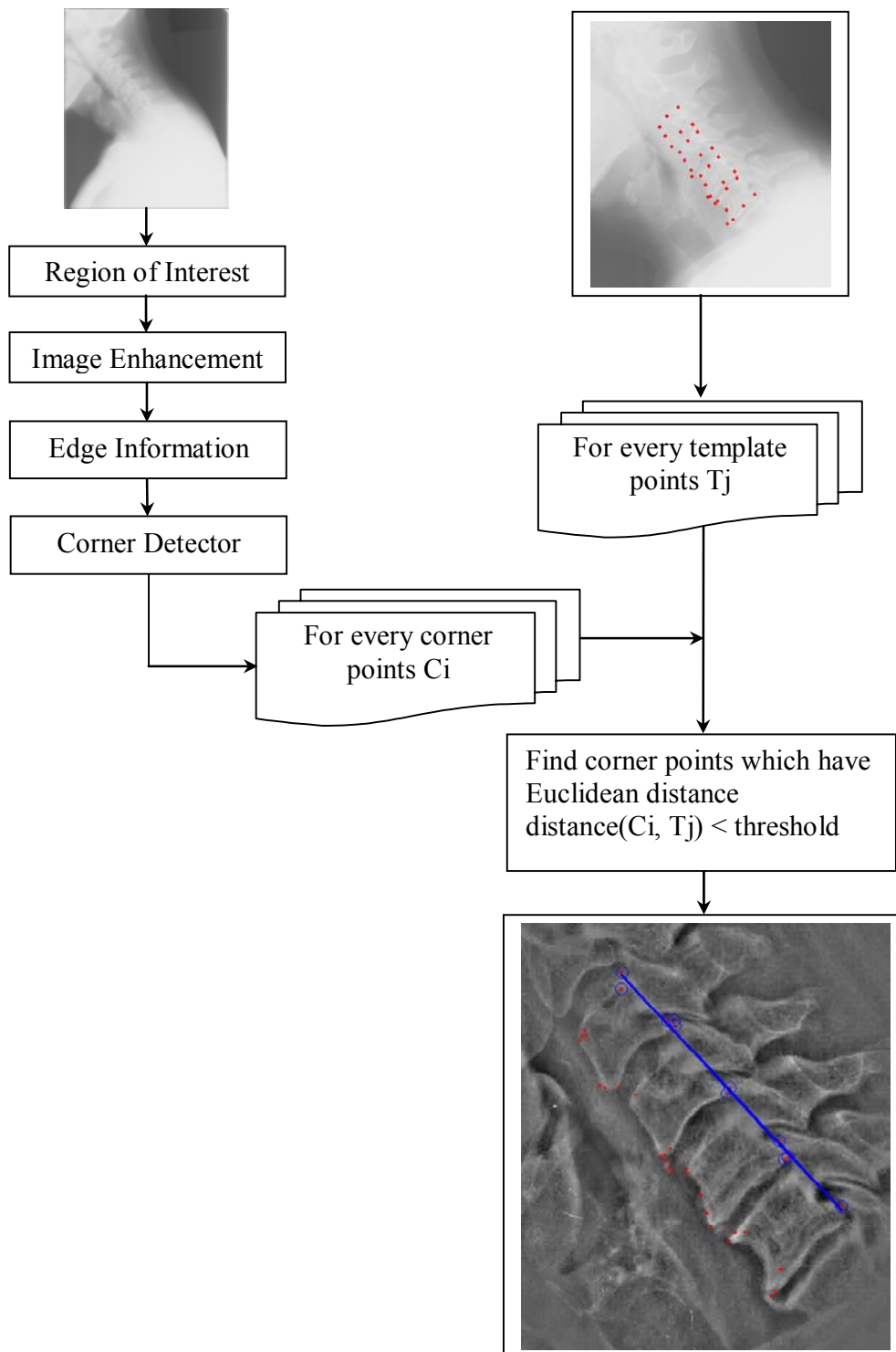
Figure 2.7. Flow chart of the proposed segmentation scheme.

approximate orientation of the anterior and inferior boundaries of the vertebrae can be detected, and then prior shape template and snakes used for fine segmentation.

2.3.2 Image Enhancement

Typically, digitized X-ray images are corrupted by additive noise. De-noising can improve the visibility of some structures in medical X-ray images, thus improving the performance of computer assisted segmentation algorithms. However, image enhancement algorithms generally amplify noise [17, 18]. Therefore, higher de-noising performance is important in obtaining images with high visual quality. The most important part of the corrupting noise is the spatially correlated Gaussian noise whose variance may vary with the signal level due to sensor non-linearity [18]. There exists different adaptive image enhancement methods, such as adaptive unsharp masking, adaptive neighborhood filtering and enhancement, adaptive contrast enhancement, and various adaptive filtering approaches using directional wavelet transform [19, 20, 21, 22]. Some methods require a priori information about the image [19].

Adaptive histogram-based equalization can be applied to aid in the viewing of key cervical and lumbar vertebrae features. However, since the method was applied to the entire image where areas including the skull and shoulders biased the histogram with a large number of high gray levels, the resulting enhancement often yielded poorly contrast enhanced vertebrae near these areas. For example, Figure 2.8(a) shows an original cervical X-ray image and Figure 2.8(b) the corresponding image after adaptive histogram equalization with C7 distorted. The vertebrae C1, C2 and C7 are basically left out because these structures are often not visible on the radiograph

28

and hence it is difficult to characterize them. Thus the regions of interest for segmentation are the cervical vertebrae (C3 to C6) and the lumbar vertebrae (L1 to L5) showed in Figures 2.8(c) and (d).

For some images, adequate contrast enhancement may not be achieved using only adaptive histogram equalization. Unsharp masking was further applied on the histogram-equalized cropped portion of the image. As shown in Figure 2.9, the main idea is to subtract the smoothed image section filtered by a Gaussian filter with a large standard deviation, σ_{large}, from the smoothed version filtered by a Gaussian filter with a small standard deviation, σ_{small}. This operation will cancel any subtle variation in the gray scale, preserving only abrupt changes (edges). The parameters of the standard deviation of the Gaussian mask σ_{small} and σ_{large} have been fixed respectively for cervical and lumbar images to achieve better enhancement performance. An example of the enhanced image using only adaptive histogram equalization is shown in Figure 2.10(b), and the enhanced image processed by both histogram equalization and unsharp masking is shown in Figure 2.10(c).
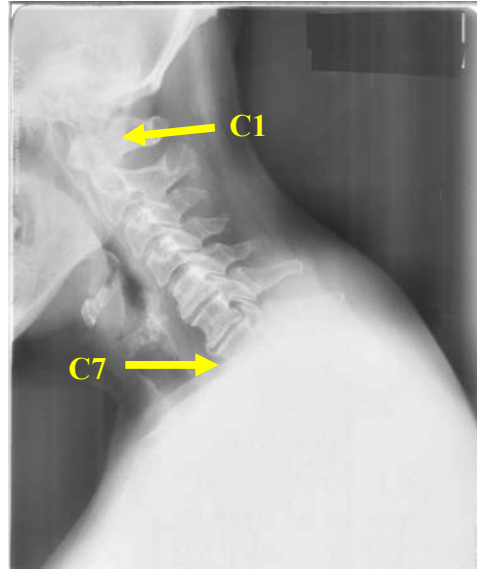
(a)   (b)
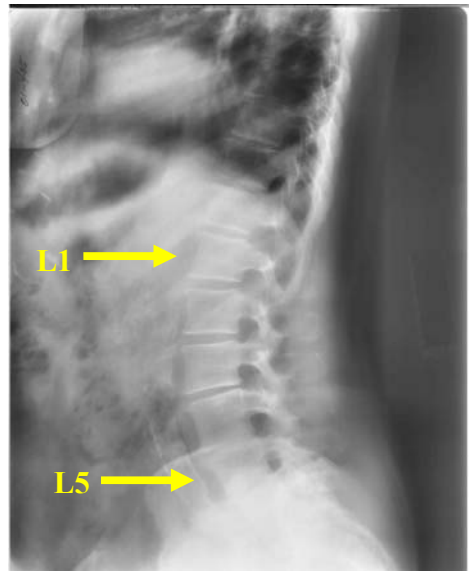
(c)   (d)

Figure 2.8. (a) Example of original cervical image. (b) Enhanced cervical image after adaptive histogram equalization. (c) Example of original lumbar image. (d) Enhanced lumbar image after adaptive histogram equalization.
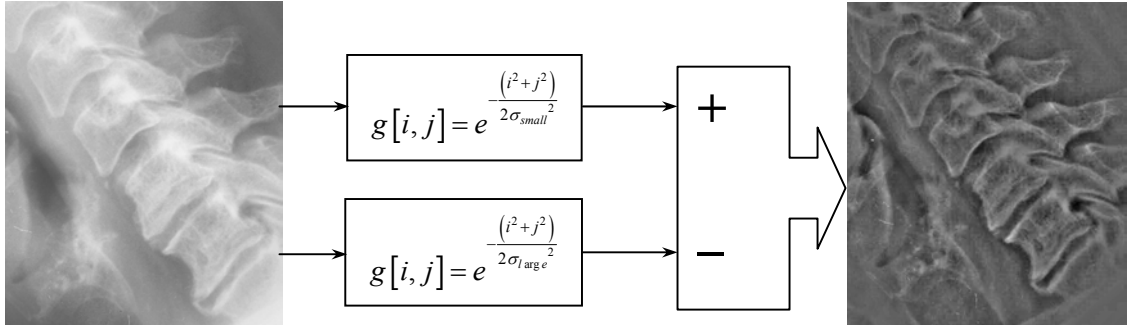
Figure 2.9. Unsharp masking applied after adaptive histogram equalization.

The two-dimensional, zero-mean discrete Gaussian function is defined by equation (2.7), where the standard deviation, σ, determines the width of the Gaussian.

$$g[x,y] = \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right) \qquad (2.7)$$

The degree of smoothing of a Gaussian filter is controlled by the parameter σ. Large σ implies a wider Gaussian filter and enhanced smoothing effect. And the amount of smoothing by Gaussian filters will be the same in all directions.

The steerable filter [23] can also be applied to assist in the viewing of key cervical and lumbar vertebrae features. It is an efficient process for removing noise and enhancing oriented structures by angular adaptive filtering. We used the steerable properties of derivatives of Gaussians for image enhancement. Figure 2.10(d) shows the result of filtering with a 3rd derivative of Gaussian separable steerable filters ($\mu = 0.09, \sigma = 1.5$).

Another important issue with image enhancement is to remove unwanted noise from the available edge image. This objective can be achieved using a great variety of algorithms and, consequently, the different algorithms that this particular approach uses can be substituted for any other that yield similar or better results. We

assume that the image is corrupted by independent additive white Gaussian noise of known variance. We then used a Bayes Least Squares - Gaussian Scale Mixture (BLS-GSM) de-noising method proposed by Portilla et al. [25]. This specific de-noising approach uses the same top-level structure as most approaches: (1) decompose the image into pyramid subbands at different scales and orientations; (2) denoise each subband, except for the lowpass residual band; (3) invert the pyramid transform to reconstruct the de-noised image from the processed subbands and the lowpass residual. The main features of the method are that the coefficients within each local neighborhood around a reference coefficient of a pyramid subband are characterized by a Gaussian Scale Mixture (GSM) model. Second, a statistical model of the coefficients of an overcomplete multi-scale oriented basis was used, and the Bayesian least squares estimate of each coefficient was reduced to a weighted average of the local linear estimate over all possible values of the hidden multiplier variable under this model. Figure 2.11 shows an example of first de-nosing the image, then applying the Canny edge detector [26] to obtain the binary edge information. The BLS-GSM de-noising method reduced the amount of unwanted high frequency components that appear in the form of small broken edges but retained most long edge information.

For various de-noising and image enhancement methods, one important issue to be considered is computational efficiency. The algorithm should be executable in a reasonable time, since the number of human experts is limited and the workloads of radiological units are increasing especially due to the screening policies. The

accuracy and resolution of X-ray images are also increasing, thus requiring more computations to be performed.
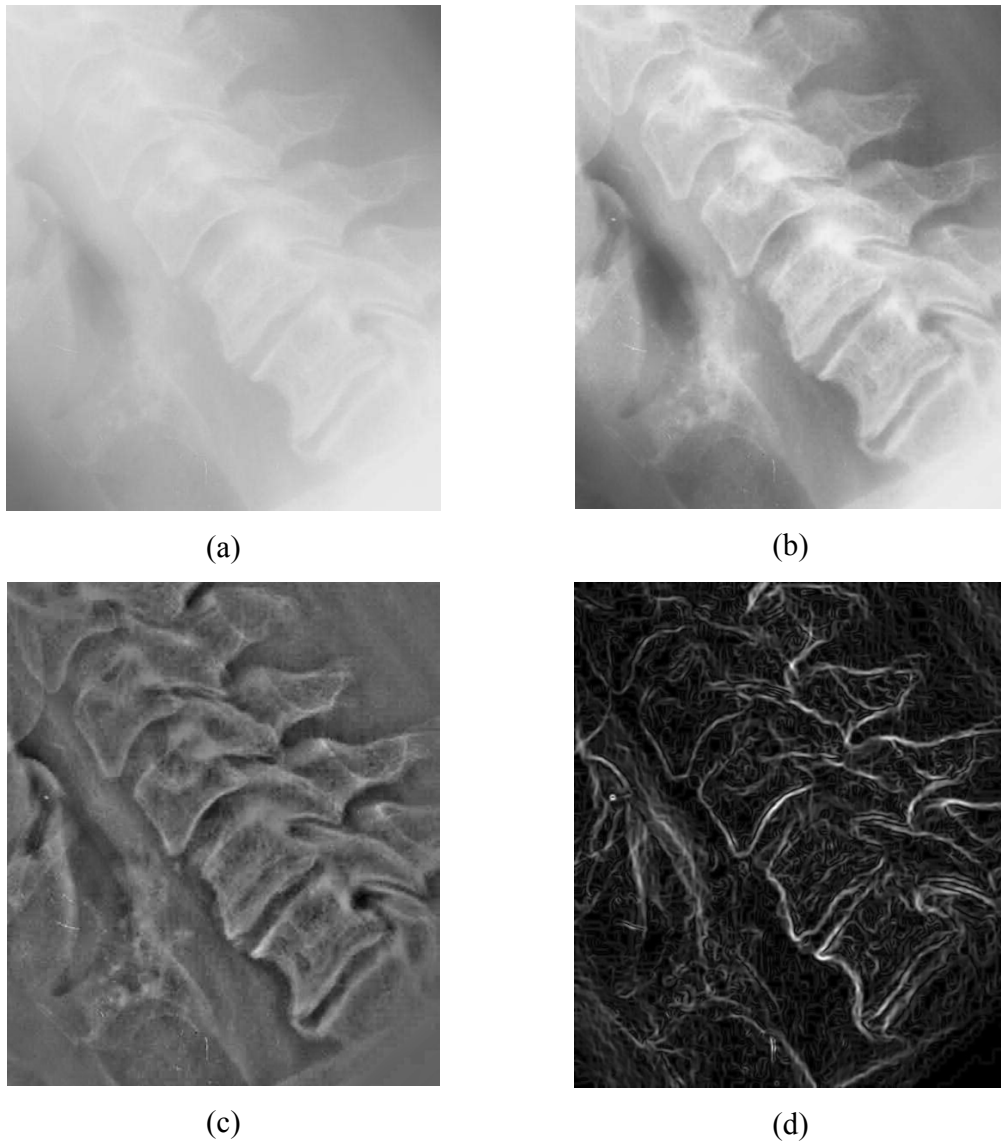
Figure 2.10. (a) A section of original X-ray image of cervical vertebra. (b) Enhanced image after adaptive histogram equalization, (c) Enhanced image after adaptive histogram equalization and unsharp masking, (d) Enhanced image after applying 3$^{\text{rd}}$ Derivative of Gaussian separable steerable filters ($\mu = 0.09, \sigma = 1.5$).

(a)   (b)

(c)   (d)
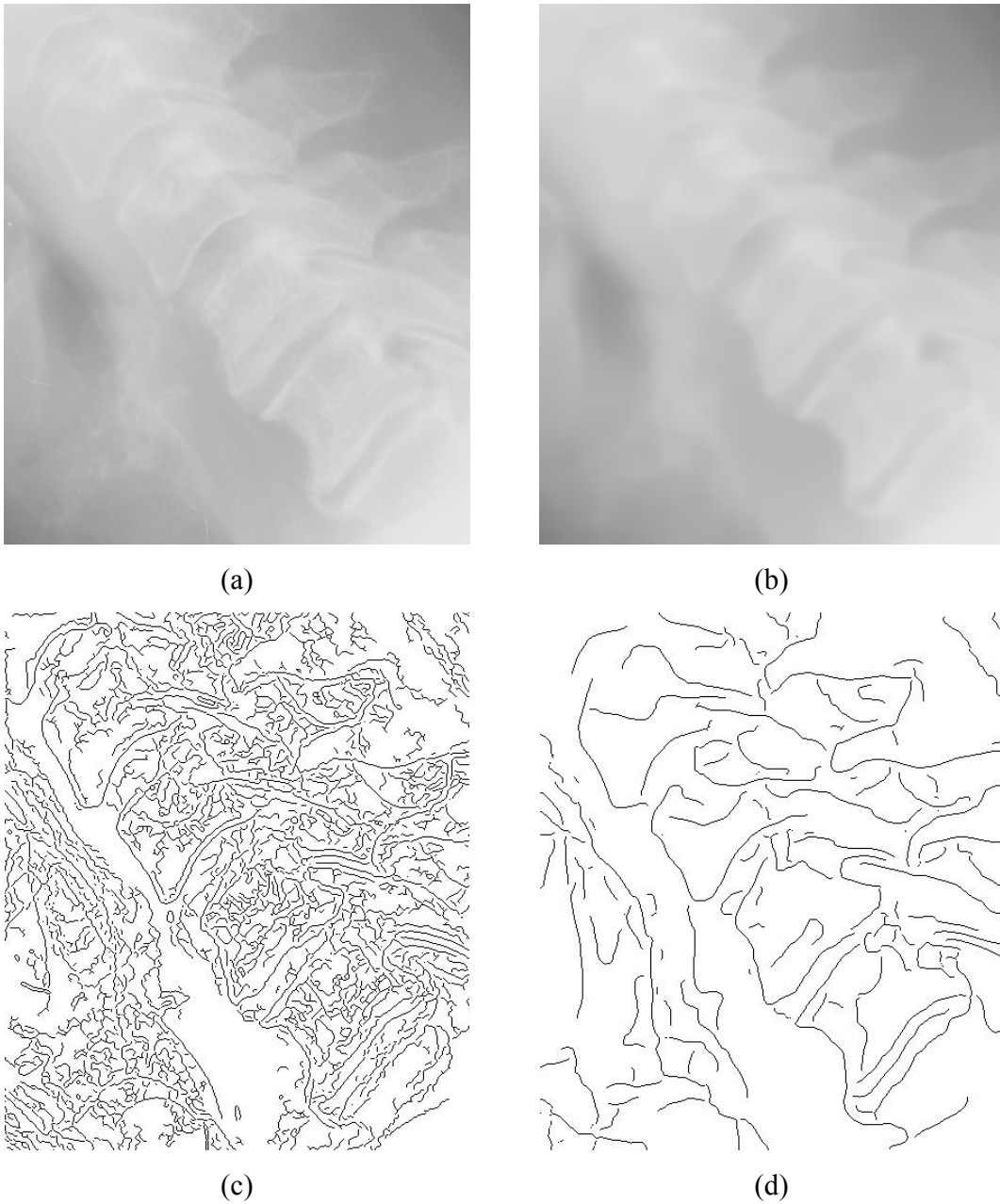
Figure 2.11. (a) Region of interest of original cervical X-ray image. (b) Resulting image after de-noising. (c) Edge image after applying Canny detector on (a). (d) Edge image after applying the Canny detector on (b).

2.3.3 Corner Detection

Corners play an important role in object identification methods used in computer vision. The vertebra generally has a rectangle shape, and the four corners of

it convey important information indicating osteophytes pathology. Since anterior and posterior osteophytes are of our primary interest in validation and diagnosis, we use corner points as the local features. In order to detect corners, we first apply the Canny edge detector [26] on the gray level image to obtain a binary edge-map. The edge contours are extracted from the edge-map, and then the gaps in the contours are filled. The curvature of each edge contour is computed based on the definition in [27], which is specified by the equation (2.8).

$$k = \frac{X_u(u,\sigma)Y_{uu}(u,\sigma) - X_{uu}(u,\sigma)Y_u(u,\sigma)}{(X_u(u,\sigma)^2 + Y_u(u,\sigma)^2)^{3/2}}$$

$$X_u(u,\sigma) = \frac{\partial}{\partial u}(x(u) \otimes g(u,\sigma)) = x(u) \otimes g_u(u,\sigma)$$

$$X_{uu}(u,\sigma) = \frac{\partial^2}{\partial u^2}(x(u) \otimes g(u,\sigma)) = x(u) \otimes g_{uu}(u,\sigma) \tag{2.8},$$

$$Y_u(u,\sigma) = y(u) \otimes g_u(u,\sigma)$$

$$Y_{uu}(u,\sigma) = y(u) \otimes g_{uu}(u,\sigma)$$

where $\otimes$ is the convolution operator, and $g(u,\sigma)$ denotes a Gaussian function with deviation σ, which is also referred to as the *scale* parameter. And $g_u(u,\sigma)$, $g_{uu}(u,\sigma)$ are the first and second derivatives of $g(u,\sigma)$ respectively.

We adopt a multi-scale corner detector based on the Curvature Scale Space (CSS) technique [27]. First the curvature of each edge contour is computed at a relatively low scale to retain all the true corners. All of the curvature local maxima are considered as corner candidates, then rounded corners and false corners due to boundary noise and trivial details are eliminated by comparing with an adaptive local curvature threshold instead of a single global threshold as in [27]. Note that the angles of corner candidates are checked in a dynamic region of support for removing

falsely detected corners. End points of line-mode curve are added as corners, if they are not close to the detected corners.

We adopt the two criteria which were proposed for eliminating falsely detected corner candidates in [28]. First, the adaptive local curvature threshold is set for a candidate according to its neighborhood region's curvature. The local maxima whose absolute curvatures are below a local threshold are eliminated, while the local maxima whose absolute curvatures are above the local threshold are declared as corner candidates. This adaptive threshold is given by:

$$T(u) = C \times \bar{k} = 1.5 \times \frac{1}{L1 + L2 + 1} \sum_{i=u-L2}^{u+L1} k(i) \qquad (2.9)$$

In equation (2.9), u is the position of corner candidate in the curve, $L1$ and $L2$ are sizes of the region of support (ROS), which is defined as from one of the neighboring local curvature minima to the next, where the curvature is strictly decreasing from the candidate point to both ends. C is a coefficient, and the mean value $\bar{k}$ is used to represent the curvature of a neighborhood region. Note that if C is set to 1, no corner is removed, and for the purpose of retaining a corner whose curvature function waveform is a standard triangle, the boundary value of C is 2. By observation, the round corner has a convex waveform in absolute curvature function, and is not sharper than a triangle. So, theoretically C should be greater than 1 and less than 2. Based on extensive experiments testing, the value of 1.5 is chosen for the coefficient C, which works well for cervical images.

The second criterion is the corner checking criterion, which states that corner C_i is a false corner if $160° \leq angle(C_i) \leq 200°$; otherwise, C_i is a true corner. The

value $160°$ is determined after further analyzing the angles of the vertebra shape corners from our cervical and lumbar images. The angles of corner candidates C_i are given by,

$$angle(C_i) = \left| arc\tan\left(\Delta Y1/\Delta X1\right) - arc\tan\left(\Delta Y2/\Delta X2\right) \right| \qquad (2.10), \text{ where}$$

$$\Delta X1 = \frac{1}{L1}\sum_{i=u+1}^{u+L1} X(i) - X(u); \Delta Y1 = \frac{1}{L1}\sum_{i=u+1}^{u+L1} Y(i) - Y(u);$$

$$\Delta X2 = \frac{1}{L2}\sum_{i=u-L2}^{u-1} X(i) - X(u); \Delta Y2 = \frac{1}{L2}\sum_{i=u-L2}^{u-1} Y(i) - Y(u).$$

As the set of corner candidates may change after applying the criterion of corner checking, further iterations are required till convergence. Using this criterion, the isolated corner candidates due to boundary noise and trivial details can be removed.

2.3.4 Definition of the Template

The vertebral shape has been described using 6, 8, 9, and 10 point models by the bone morphology community. These landmark points are placed at pathologically significant points on the vertebra boundary. We use the 9-point representation of vertebra shape since a portion of the NHANES II spine image collection was marked using the 9-point model by a board certified radiologist as ground truth. The locations of these template points, shown in Figure 1.3(a), are indicative of the pathology found to be consistently and reliably detectable in the image collection. Points 1, 3, 6, and 4 are indicative of the four "corners" of the vertebral body as seen in a projective sagittal view. Points 1 and 4 mark the upper and lower posterior corners of the vertebra, respectively; Points 3 and 6 mark the upper and lower anterior corners of the vertebra, respectively. Points 2 and 5 are the median along the upper and lower vertebra edge in the sagittal view; Point 7 is the median along the anterior vertical

edge of the vertebra in the sagittal view. Note that Points 8 and 9 mark the upper and lower anterior osteophytes, so if osteophyte(s) are not present on the vertebra, then these points coincide with points 3 and 6, respectively.

We use the mean of 50 templates randomly selected from the images which were marked using 9-point model. The target shape to be segmented is the cervical vertebrae C3-C6, thus the template is defined by a total of 36 morphometric points for the four vertebrae. As shown in equation (2.11), each template T_i contains the x and y-coordinates of a single 36-dimensional vector respectively, and $\overline{T}$ denotes the mean of the 50 templates.

$$T_i = \begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} x_{i1}, x_{i2}...x_{i36} \\ y_{i1}, y_{i2}...y_{i36} \end{bmatrix}$$

$$\overline{T} = \frac{1}{50}\sum_{i=1}^{50} T_i$$

(2.11)

2.3.5 Template Corner Matching

A matching procedure generally follows after a local feature extraction procedure to recognize and locate objects. We use corner points as the local features. These corners have sufficient information to recognize and locate objects. The shape template also has corner points, and by matching these corner points with corner points in a real image, it is possible to recognize and locate objects.

The template points 1, 3, 6, 4, 7, 8 and 9 were used as anchors in our corner matching stage. The median points 2 and 5 which reside at the middle of the upper and lower vertebra edge in the sagittal view are not used in corner matching. We overlaid the template on one image and correlate corner points to the seven points in the template. We examined every corner point C_i in the image, and compute the

Euclidean distance from every template corner point T_j. If the Euclidean distance $d(Ci, Tj)$ is below a certain threshold, the corner point C_i is extracted as a vertebra corner point, and included in the object model. The threshold value is chosen as 18 pixels from many trials to ensure that falsely detected corners due to noise and unwanted corners not belonging to vertebrae are not included in the object model. Figures 2.12 (a) and (b) illustrate the results of corner detection and template corner matching. It can be seen that corner matching picks up candidate corners which are most correlated to vertebrae corners.

(a)   (b)
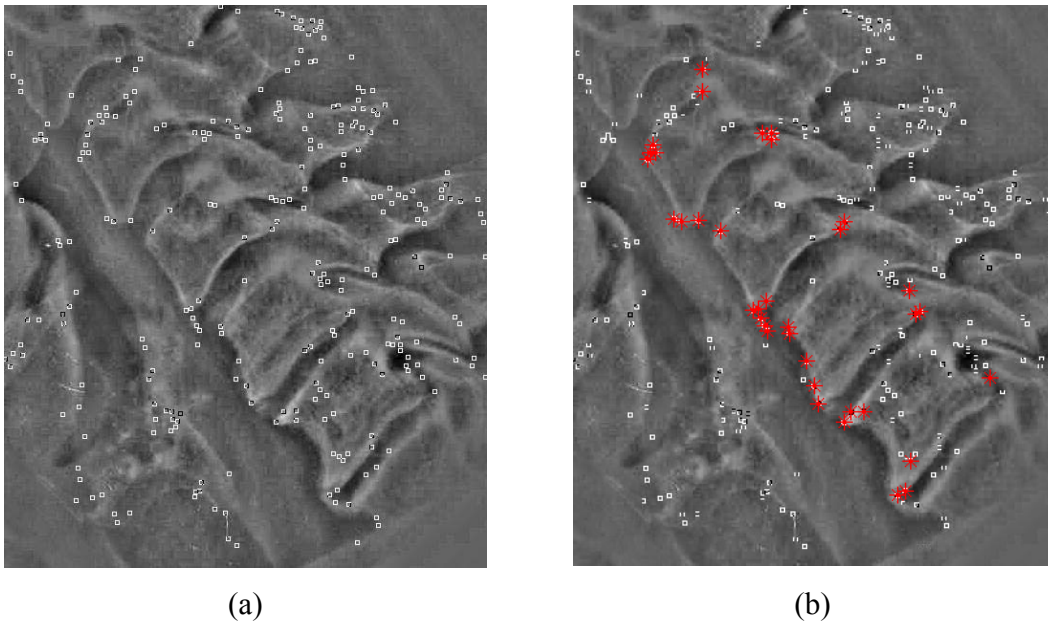
Figure 2.12. Illustration of (a) corner detection, (b) template corner matching.
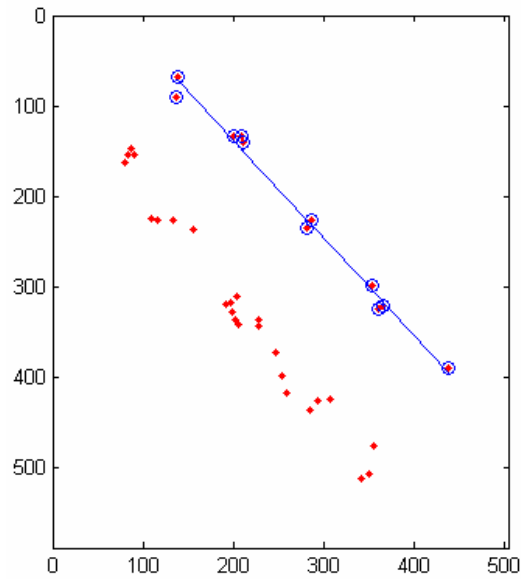
Figure 2.13. Calculating the orientation of the object model from the corner points passing through the posterior corners of the cervical vertebrae.

After obtaining the object model points by corner matching, the approximation of the orientation of the object model can be computed from the corner points passing through the upper and lower posterior corners of the cervical vertebrae. Let ϕ to be the angle that describes the orientation of the template. As shown in Figure 2.13, the posterior corner points define a curve passing through the cervical spine. The value of ϕ is defined as the inverse tangent of the slope of the polynomial of first order (straight line) that best fits that series of points. The angle ϕ and the location of the top-rightmost corner point of the object model determine the final location of the prior shape template.

2.3.6 Fine Segmentation

The 9-point model is sufficient for corner matching and finding the initial spatial location of the shape template. However it falls short of providing a

40

sufficiently rich description of the shape contour for the purposes of fine segmentation. Thus we use arithmetic interpolation to increase the 2-D points from 9 to 36 by adding 3 intermediate points between two original landmark points. And then the denser 36-point shape contour for each vertebra is allowed to deform according to the following energy function.

Let $v_k(x_x, y_k), k = 1, 2, ..., N$ be the N deformable vertices on the polygonal contour, and let $p_{k,j}, k = 1, ..., N$ and $j = -M, ..., 0, ..., M$ be the uniformly sampled points along the orthogonal curves of the contour, where $p_{k,0}, k = 1, ..., N$ denote the base point. And $f(x_k, y_k)$ is the image intensity at (x_k, y_k). The associated energy function is given by,

$$E = E_{\text{gradient}} + E_{\text{proximity}} + E_{\text{smoothness}}$$

$$E_{\text{gradient}} = -\rho_1 \sum_{k=1}^{N} |\nabla f(x_k, y_k)|$$

$$E_{\text{proximity}} = \rho_2 \sum_{k=1}^{N} D(v_k - p_{k,0}) \tag{2.11}$$

$$E_{\text{smoothness}} = \rho_3 \sum_{k=1}^{N} |D(v_k - p_{k,0}) - D(v_{k+1} - p_{k+1,0})|$$

The proximity energy $E_{\text{proximity}}$ measures the displacements of the contour vertices from the base points $p_{k,0}, k = 1, ..., N$, and the smoothness energy $E_{\text{smoothness}}$ measures the dissimilarity of distances of consecutive vertices from their base points. The resulting contour after deformation should lead to more accurate segmentation results.

Chapter 3: Experiments and Results

3.1 Data Set and Ground Truth

Based on the 9-point model, the template for cervical image is defined by a total of 36 morphometric points including the whole shape of vertebrae C3-C6, and the template for lumbar image is defined by a total of 45 morphometric points including the whole shape of vertebrae L1-L5. These morphometric points were previously marked by expert radiologists, and represented our ground truth about the correct location of points defining the shape of the cervical vertebrae. Figure 3.1 shows the location of points of the defined template for both cervical and lumbar images.

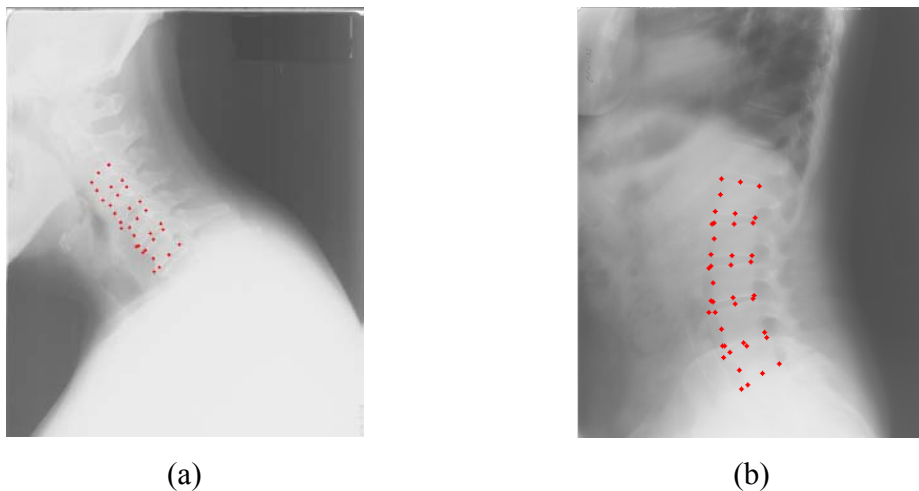(a)   (b)

Figure 3.1. Illustration of the ground truth morphometric points for (a) cervical spine C3-C6. (b) lumbar spine L1-L5.

We tested our proposed scheme on a data set of 50 randomly selected cervical images from NHANES II database used. Each cervical image is 1,755 rows by 1,462 columns with 256 levels of gray, and stored in JPG format. Although our current

efforts concentrate on cervical X-ray images, future work will also include lumbar X-ray images. The important issues in applying our scheme to lumbar images are discussed in the next section.

3.2 Experiments and Performance Measurements

Edge detection, corner detection and template corner matching are applied sequentially to every image on the data set using its own template. This experiment aims to verify the hypothesis that if the correct template is used, the proposed scheme is a suitable approach for segmenting the cervical vertebrae from radiological images. The last statement would be proven given that if the correct answer (a template extracted from every image) were used, our curvature-based corner matching approach would be able to find adequate estimates of the location, orientation and scale of the cervical spine. Results from this experiment can be summarized as follows,

- In most cases, all posterior corner points of vertebrae were detected, and further used in finding the orientation of the cervical spine. Although the corner points can be missing due to poor image contrast in some cases, the remaining corner points provide enough information to design the polynomial of first order (straight line) that best fits that series of points. An example of poor image contrast is presented in the second row of Figure 3.2. It can be seen that vertebrae C5 and C6 appear in the image of low contrast, which hinders the corner detection for posterior edges.

- An average difference in orientation is less than $1°$ and a corresponding $\sigma = 0.41°$ was measured. The displacement between the location of the

top-rightmost corner point of the object model and the template model is on average 5.4 pixels along the x-coordinate and 8.8 pixels along the y-coordinate. The location displacement does not greatly affect the result since the orientation of the template at its final location has nearly the same orientation as the template model.

These results clearly show that the curvature-based corner matching is an effective technique for coarsely segmenting the cervical vertebrae if the correct template is used. In other words, if a reasonable template is provided, our approach will yield acceptable segmentation results. Figure 3.2 shows some of the results of our proposed scheme. The first column shows the captured region of interest for C3 to C6 on cervical images. The second column shows the region of interest after applying image enhancement, which is overlaid with the detected corners and the line passing through the upper and lower posterior corners of the cervical vertebrae. The third column presents the final segmentation using a snake with initial estimates of location and orientation of the cervical vertebrae C3-C6. It can be seen that the successful pose estimation of the vertebrae leads to improved segmentation.
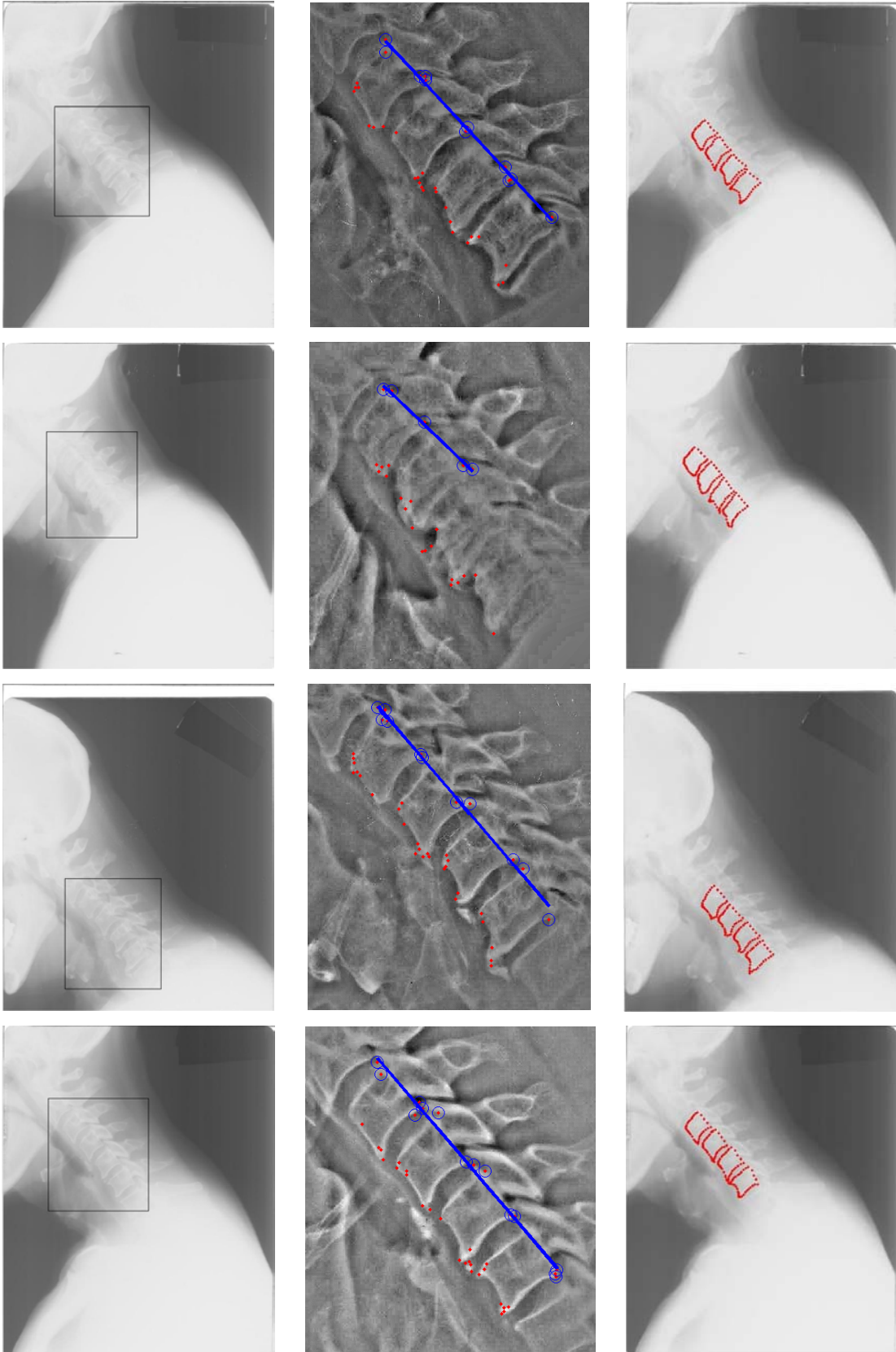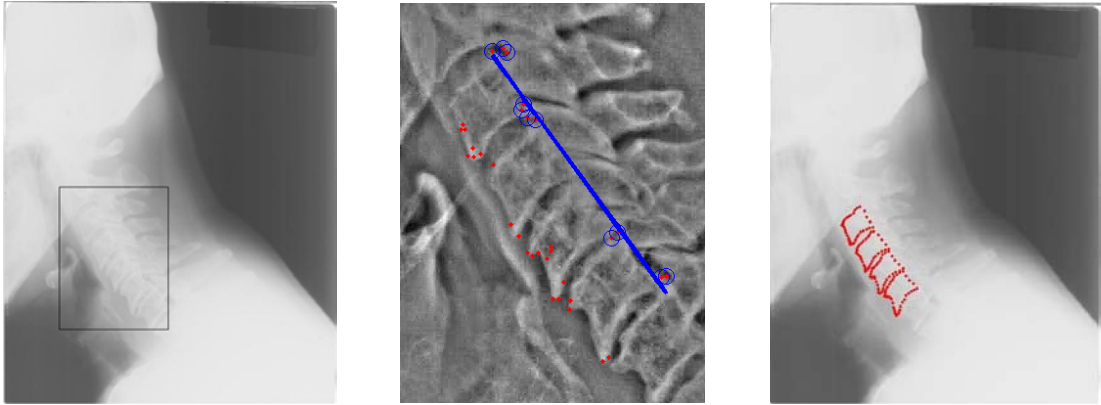
Figure 3.2. Results of applying the proposed scheme on 50 randomly selected cervical images using their corresponding templates. The region of interest is marked by the rectangular bounding box. The detected corner points are marked by red dots, the posterior corner points are additionally marked by blue circles, and the final contours after deformation are indicated by red marks in the third column.

3.3 Computational Issue

The biggest advantage of the template corner matching approach is execution time. During the development of this thesis, all algorithms were implemented using MTLAB 7.0.4 on a Pentium M 2.0 GHz computer with 512 MB of RAM. Although a number of preprocessing steps including finding region of interest, adaptive histogram equalization, and unsharp marking needs to be carried out before corner detection and matching, they are performed quickly. Moreover, these preprocessing steps are computed only once, and are saved. The total preprocessing of each image takes an average of 3.1092 seconds. During edge detection and corner matching, it takes 1.1194 seconds for edge detection, 4.5296 seconds for corner detection, and 0.0132 seconds for corner matching on average. Hence, the execution time will be the sum of the time spent in four major steps: (1) preprocessing, (2) edge detection, (3) corner detection, and (4) template corner matching. Hence, the total execution time of the proposed approach for coarse segmentation on a cervical spine image from the

data set is an average of 10.42 seconds. Table 3.1 shows the average execution time of the proposed approach compared with GHT approaches reported in [8, 9] for cervical vertebrae.

Table 3.1. Comparisons of execution time for previous and current coarse segmentation scheme

EXECUTION TIME	CERVICAL X-RAY IMAGES (seconds)
Previous approach [8]	78
Previous approach [9]	55
Current approach applied on region of interest (ROI)	10.42

The execution time of both corner matching approach and GHT depend on several factors such as the size of the target image, size of the template and some factors associated with implementation. GHT also depends on the quantization of the Hough space. The size of the target images is one of the key aspects in determining the execution time. Based on our experiments, the execution time is directly proportional to the size of target image. It should be noted that the execution time for detecting the edges, extracting the contour curve and detecting the corners accounts for more than half of the total execution time. The execution time has a greater dependence on the number of edge pixels to be processed by the algorithm than the size of the image itself, although these two quantities are related. Thus, extracting gradient information from thin edges that contain all the shape information from the cervical spine is crucial in keeping the execution time low.

Note that edge detection and corner matching are applied to the region of interest instead of on the entire image. This is one of the basic preprocessing steps but yet a very important step. The region of interest is basically a bounding box which is placed around the vertebrae and for all further processing only that particular region is processed. The algorithm should also make use of some a priori knowledge about the location of the vertebrae. To generate a proper bounding box, we examined the 939 cervical images which have been marked by radiologists using the 9-point shape model. The average location and size of the bonding box corresponding to cervical vertebrae C3-C6 are obtained statistically. And then an additional border of width 150 pixels is included to make sure that the bonding box will not be placed in a wrong position which does not correspond to the vertebrae. The first column of Figure 3.2 shows the bounding box placed around the region of interest for C3 to C6 on cervical images.

The number of critical points in the template also plays an important role. In our current approach, the number of points that the template is made of has been reduced to 36, while in previous approaches [8] and [9] each template is made out of 80 landmark points. Our approach uses 9-point for each vertebra which conveys enough information for the corner matching algorithm to obtain the posterior corners of vertebrae C-3-C6, and thus determine the orientation and location of template. However, the GHT [8, 9] needs more points in the template to represent the anatomic features of the cervical spine, since more points in the GHT accumulator will be related to the same reference point. The execution time of GHT depends on the search

procedure which is directly proportional to the number of points in the template. Therefore, a bigger template results in increased execution times.

Chapter 4: Image segmentation and Pathology Validation Tool

The medical image segmentation and pathology validation tool was developed to collect ground truth data on the digitized spine X-ray images from NHANES II. In addition, it enables verification of the segmented vertebral boundaries and collection of pertinent pathologies from the image as a whole, in addition to those local to the vertebral body. The tool, shown in Figure 4.1, is designed to allow content experts to login into our database and review image and pathology data and provide necessary updates. The software can operate in an offline or online mode and need only connect to exchange data. It has been developed using JAVA and continues to evolve based on suggestions from experts.

4.1 Design Considerations

In developing a medical validation tool, it is essential to take into account typical practices of radiologists and other medical experts providing validation data. For example, a brightness-contrast leveling tool is often assumed to exist in such software as a response to vertical drag movements of the mouse with a mouse-button press. Other image enhancement tools are unsharp masking, histogram equalization and image negation. Our interaction with experts has enabled us to develop a software that provides a traditional interface and could help in minimizing the time it takes to use the software. It also takes into account the fact that the radiologists are providing this data while caring for their patients and may be interrupted during the validation process. The system regularly saves their work and if necessary can return them to their last completed action.
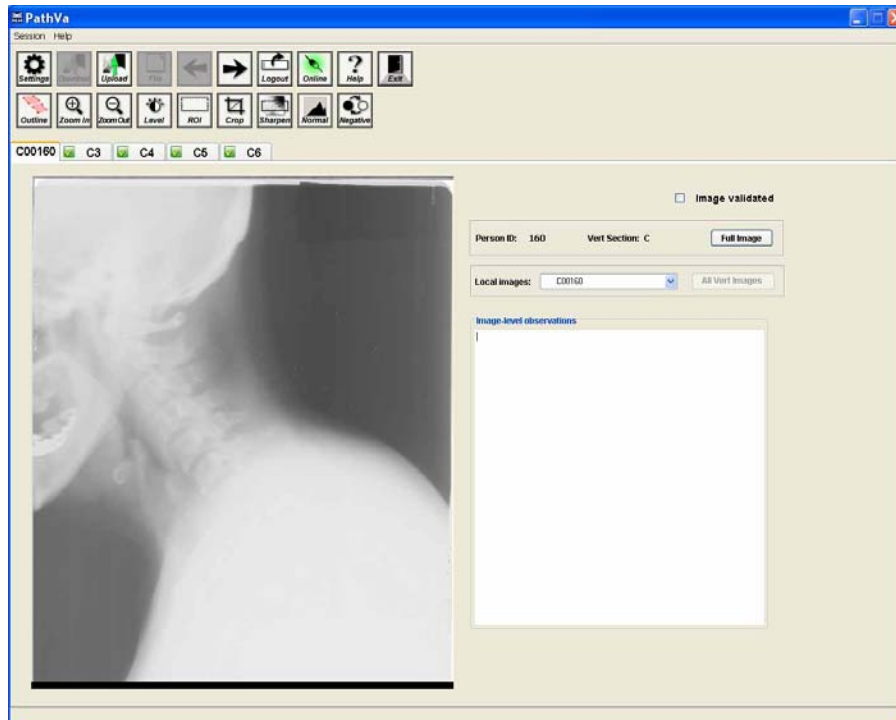
Figure 4.1: Screen shot of the main window in the medical validation tool showing a whole cervical spine X-ray and some image enhancement controls.

4.2 Features and Capabilities

The system is shown in Figure 4.2. It consists of a MySQL database server and an image server on the server side and the validation tool/applictaion on the client side. The client tool operates in both stand-alone and online modes. It starts in the stand-alone mode and the user can login to the remote server to download image and text records. These records can be obtained as individual items or can be downloaded in selectable batch sizes. If the latter is selected, then future operations of the client can be performed in the stand-alone mode until all images and records have been reviewed. After completing the batch the user uploads all validated records which are verified by the service manager. Only the validated text records and image segmentation data are uploaded. Such a batch mode operation allows the medical

experts to validate images on their own time while disconnected from the Internet. In this case, the user may need to login to the server again for performing data transactions.
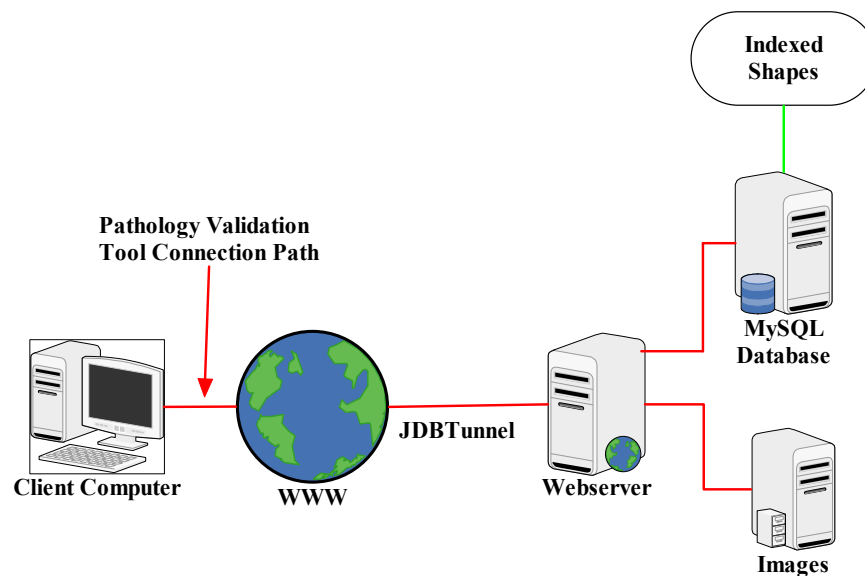
Figure 4.2: A schema depicting the various components enabling the collection of medical validation data.

For each cervical or lumbar image, the user is shown the image on a tabbed interface with the entire image on the main tab, and the corresponding vertebrae appear on the other tabs, as in Figure 4.1. Here the user can mark regions of interest (ROI) and perform image enhancements to better view the pathology of interest. Image enhancement tools include a leveling tool, un-sharp masking tool that operates in a cumulative fashion, adaptive histogram equalization tool and image negation tool. These tools are commonly used by radiologists in their typical workflow. This tab also permits image level comments to be recorded. Individual vertebral images are visible through a button on this main panel. The viewport for each vertebra is

computed from segmented vertebral boundary data coordinates that are stored in the server database. Since the vertebral segment is not cut and saved as a separate image, it is possible to allow the user to expand his/her view to include neighboring vertebrae by adjusting the coordinates of the cropping rectangle. In consultation with the medical experts, we enabled a viewing technique of multiple vertebrae while highlighting the vertebra under investigation. The user is provided with four sliders to manually adjust the dimensions of the viewport. The maximum width and height of the viewport are set to 400 and 700 pixels respectively. The top left corner coordinates, and width and height of the cropping rectangle are also displayed separately to facilitate usage, seen in the middle of the screen shot in Figure 4.3. This view would enable the medical experts to make a better judgment on the pathology.

In the vertebra pathology examination panel, shown in Figure 4.3, the medical expert can enter data about four types of pathologies and comment about others not covered by the provided drop-down boxes. The pathology categories provided are anterior and posterior osteophytes, disc space narrowing, and subluxation and spondylolisthesis. For the osteophytes, they can either be normal or have three grades of severity, viz.: slight, moderate, and severe. Each can be of traction or claw type. It is possible to mark both types of pathology for vertebrae with osteophytes of slight severity. This indicates that the expert was unable to determine the type of degenerative growth, so that he/she choose to mark the type of pathology undecided. For moderate or severe pathology, only one of the two types may be marked since it is expected that a clear judgment can be made for these severities. Overall, each osteophyte location can have only one severity grade. For the case of narrowing disc

space, the expert can indicate either mild, moderate, or severe. The tool also provides a measuring tool for measuring distances between objects. In certain cases, this tool could be used to measure the narrowing. For subluxation and spondylolisthesis, both of which indicate disk slippage, similar mild, moderate, and severe selection criteria are provided.

Often, a vertebral boundary is computed by computer assisted image processing techniques and may not be medically valid. Thus, an essential task of this validation process is verifying that these boundaries are indeed a true representation of the anatomy. Additionally, for the fixed point boundary distributions, viz., the 9-point and 36-point boundaries, it is essential that the points are located at appropriate locations. For example, in the 9-point boundary layout, shown in Figure 1.3(a), points 1, 3, 6, and 4 are indicative of the four "corners" of the vertebral body as seen in a projective sagittal view. These and other points are often used as anchors in our shape matching step and their correct location is important to proper system performance. The experts can modify boundary-point position by moving each point as necessary. In addition, the medical expert can also select one or more valid boundary representations from the multiple segmentations often available for each vertebra. A segmented boundary shape is identified by the segmenting person identifier and creation date. The software provides a way to show the experts each shape on the X-ray image and up to three shapes can be overlapped. It is also possible to turn off the lines connecting the boundary points to get a clearer view of the outline with respect to the image edge.
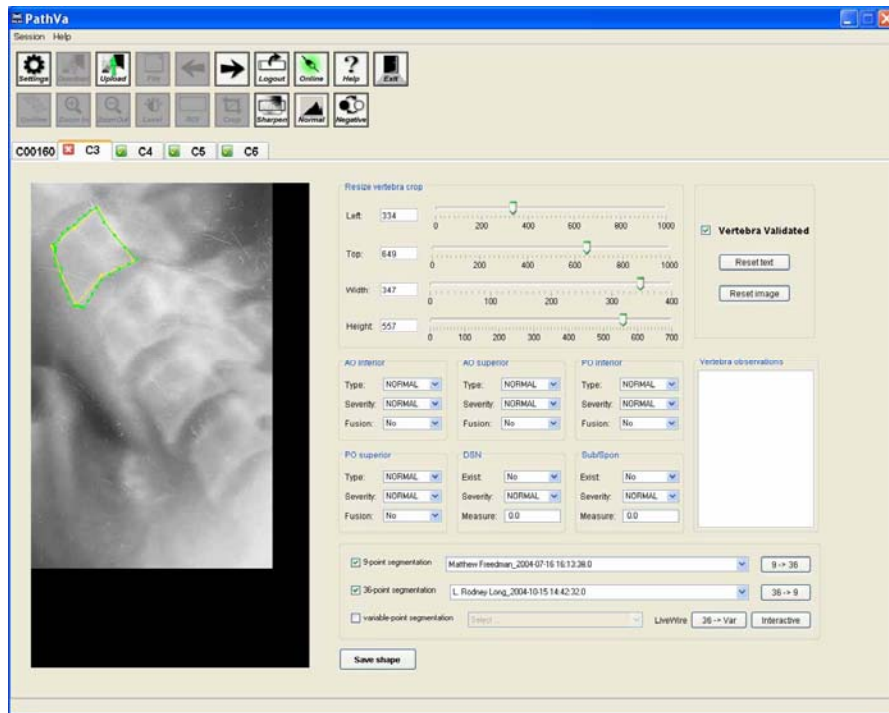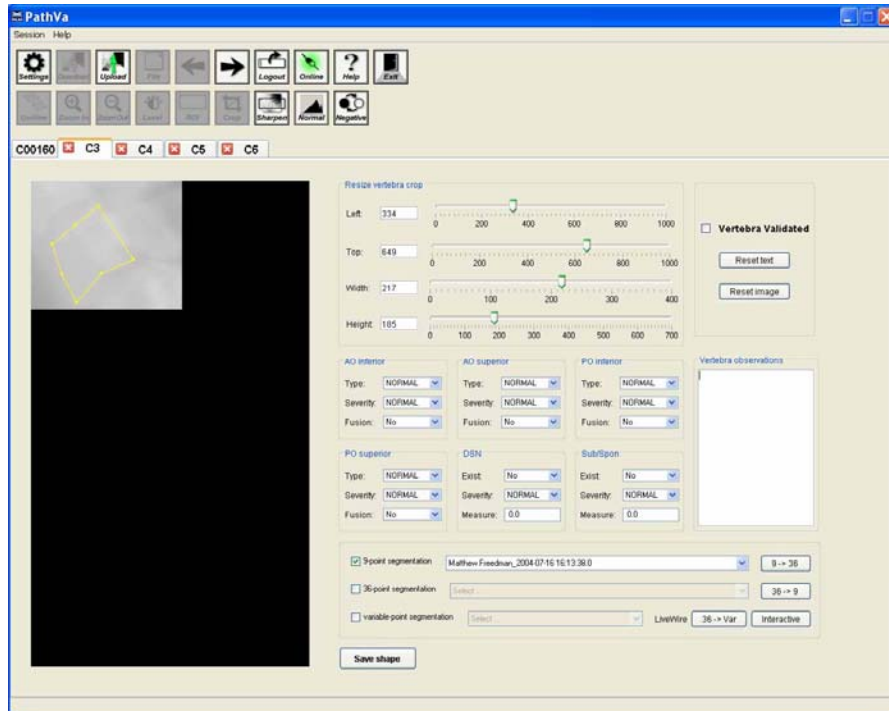
Figure 4.3: Screen shots of vertebral tab panel depicting a cropped vertebra with the segmented boundary as overlay. It also shows various pathology check boxes and notes area.

4.3 Lessons learned

In this chapter, we discussed the design and development of a Web-enabled medical validation tool. Note that the tool cannot be currently generalized for use with a variety of image types; however, it is possible to build such tools that can be easily modified for different image types by learning from the design and use of the medical validation tool. Lessons were learned in both design and developmental aspects of the tool and are important for any project of similar nature. They include concepts in project and workflow planning, software and GUI design, and software engineering. Involving primary users, who may be domain experts, early on in the development and planning of such software is very helpful. Obtaining information about their typical workflow and obtaining a list of desirable features significantly improves long-term benefits from data collection and validation efforts. Some examples of add-ons and modifications in the software resulting from such efforts include mouse-based image enhancement tools, batch-mode operation, and tools for reorienting users interrupted during the validation process.

Future enhancements are planned for the tool in order to reduce expert burden in medical validation. Several image segmentation algorithms are expected to be included after a thorough on-line testing and performance evaluation. A major hurdle in this task will be the redevelopment of specialized routines in a cross-platform language such as JAVA. In general, we expect that the contributions from this work will enable a medically validated dataset valuable for evaluating the performance of various segmentation algorithms and CBIR related research.

Chapter 5: Conclusion and Future Perspective

The success of our curvature-based corner matching scheme for coarse segmentation of cervical X-ray image mainly depends on three factors: (1) Availability of edges for corner detection, (2) Correctness of the template, and (3) Correlation between template points and detected corner points. The Canny detector is adopted to detect the edges in the cervical vertebrae, and then the multi-scale corner detector extracts vertebral contour curves using the detected edges. The corners of the contour curves are further detected based on Curvature Scale Space (CSS) technique.

Although more study is still needed, experiments showed that the proposed scheme is a suitable approach for coarse segmentation of cervical vertebra if the correct template is used. Moreover, the proposed scheme lessened segmentation's dependence on gradient information. The corner matching approach detected only part of the corners when the available gradient information is poor. However, the detected corners could be successfully related to the template, thus orientation of the vertebrae could still be estimated by finding the polynomial of first order (straight line) that best fits the series of detected corner points along the posterior edges of the vertebrae in the sagittal view. Bad segmentation occurs when gradient information is poor and the template does not represent appropriately the target object. Although current efforts concentrate only on cervical X-ray images, future work will also include lumbar X-ray images.

This thesis is also the first reported effort to develop a prototype Internet-based medical image segmentation and pathology validation tool, which aims to

enable radiologists to validate computer generated image segmentations, modify existing or create new segmentation in addition to identifying pertinent pathology data. The tool collects validation data on the digitized spine X-ray images in a uniform and consistent manner. In addition, it allows multiple expert opinions on each image, and online feedback without the physical presence of the users at NLM facility or replication of the entire collection at the remote facility.

In general, we expect that the medical image segmentation and pathology validation tool will contribute to a medically validated dataset valuable for evaluating the performance of various segmentation algorithms and CBIR related research. Moreover, our work enables the formation of general principles for design and development of similar tools that could allow validation of images of different anatomy in different modalities which may have different set of pathologies.

Bibliography

[1]   S. Antani, J. Cheng, and *et al*. Medical Validation and CBIR of Spine X-ray Images over the Internet. *Proc. IS&T/SPIE Electronic Imaging Science and Technology 2006: Internet Imaging VII*, Jan 2006, SPIE vol. 6061 pp. 60610J (1-9).

[2]   R. Long, S. Antani, G. Thoma. Image informatics at a national research center. *Computerized Medical Imaging and Graphics 29*, Feb 2005, 171-193.

[3]   P. Chamarthy, R.J. Stanley, G. Cizek, R. Long, S. Antani, G. Thoma. Image analysis techniques for characterizing disc space narrowing in cervical vertebrae interfaces, *Computerized Medical Imaging and Graphics*, 2004, 28(1-2):39-50.

[4]   R.J. Stanley, S. Seetharaman, R. Long, S. Antani, G. Thoma, E. Downey. Image analysis techniques for the automated evaluation of subaxial subluxation in cervical spine X-ray images, *Proc. 17th IEEE Symposium on Computer-Based Medical Systems*, June 2004, pp. 204-209.

[5]   Tagare H. Deformable 2-D template matching using orthogonal curves. *IEEE Transact Med Imaging*, 1997, 16(1):108–17.

[6]   G. Zamora-Camarena. Automatic segmentation of vertebrae from digitized x-ray images. *PhD thesis*, Texas Tech University, 2002.

[7]   T. F. Cootes, C. J. Taylor. Active Shape Models-smart snakes. *Proceedings of the British Machine Vision Conference*, pp. 266-275, BMVA Press, UK, 1992.

[8]   A.Tezmol. Customized Hough transform for robust segmentation of cervical vertebrae from X-ray images. *Master's thesis*, Texas Tech University, 2001.

[9]   A. Gururajan. Coarse Segmentation of Cervical and Lumbar Vertebrae Using a Customized Version of the Generalized Hough Transform. *Master's thesis*, Texas Tech University, 2003.

[10] H. Moon, R. Chellappa, and A. Rosenfeld. Optimal. Edge-based Shape Detection. *IEEE Transaction on. Image Processing*, vol. 11, pp. 1209-1227, 2002.

[11] Ben-Arie and Rao. Optimal Edge-Detection Using Expansion Matching and Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 16, pp. 1169–1182, 1994.

[12] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, pages 321-331, 1987.

[13] A. P. Paplinski and J. F. Boyce. An implementation of the active contour method for noisy images using a local minimization algorithm. *Technical Report TR95-1*, Monash University, January 1995.

[14] C. Xu and J. L. Prince. Snakes, Shapes, and Gradient Vector Flow. *IEEE Transactions on Image Processing*, 7(3), pp. 359-369, March 1998.

[15] Ballard DH. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition,* 13(2), pp. 111–22, 1981.

[16] P.-F Fung, W.-S Lee and I. King. Randomized Generalized Hough Transform for 2-D Grayscale Object Detection. *Pattern Recognition and Signal Analysis, 13th International Conference on Pattern Recognition*, vol. 2, pp. 511-515, August 1996.

[17] T. Aach, U. Schiebel, G. Spekowius. Digital image acquisition and processing in medical X-ray imaging. *Journal of Electronic Imaging*, 8(1), pp. 7-22, January 1999.

[18] K.K. Shung, M.B. Smith, B.M.W. Tsui. Principles of Medical Imaging. *Academic Press*, 1992.

[19] M.J. Carreira, D. Cabello, A. Mosquera, M.G. Penedo, I. Facio. Chest X-ray image enhancement by adaptive processing. *Annual international conference on the IEEE Engineering in medicine and Biology Society*, vol. 13, No.3, 1991.

[20] S. Guillion, P. Baylou, M. Najim, N. Keskes. Adaptive non-linear filters for 2D and 3D image enhancement. *Signal processing*, pp. 237-254, 1998.

[21] L. Li, W. Qian, L.P. Clarke. X-ray Medical Image Processing Using Directional Wavelet Transform. *Proc. of IEEE. Int. Conf. on Acoustics, Speech, and SignalProcessing*, pp. 2251-2554, 1996.

[22] Z. Hua, M.-N. Chong. A wavelet de-noising approach for removing background noise in Medical Images. *International Conference on the Information, Communication and Signal Processing*, pp. 980-983, 1997.

[23] W.T. Freeman and E.H. Adelson. The design and use of steerable filters. *IEEE Trans. on Patt. Anal. and Machine Intell.*, vol. 13, pp. 891–906, Sept. 1991.

[24] E. Simoncelli and H. Farid. Steerable wedge filters for local orientation analysis. IEEE *Transactions on Image Processing*, 5(9), pp. 1377–1383, 1996.

[25] J Portilla, V Strela, M Wainwright, E P Simoncelli. Image Denoising using Scale Mixtures of Gaussians in the Wavelet Domain. *IEEE Trans. on Image Processing*, vol. 12, no. 11, pp. 1338-1351, Nov. 2003.

[26] J.F.Canny. A Computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, No.8, pp. 679-698, 1986

[27] F. Mokhtarian, R. Suomela. Robust image corner detection through curvature scale space. *IEEE Trans. on Pattern Analysis and Machine Intelligence,* 20(12):1376-1381, 1998.

[28] X.C. He, N.H.C. Yung. Curvature scale space corner detector with adaptive threshold and dynamic region of support. *Proceedings of the International Conference on Pattern Recognition (ICPR)*, vol. 2, pp. 791-794, 2004.

[29] H. Moon, R. Chellappa, and A. Rosenfeld. Performance analysis of a simple vehicle detection algorithm. *Image and Vision Computing*, vol. 20, pp. 1–13, 2002.

[30] Demetri Terzopoulos, Andrew Witkin, and Michael Kass. Constraints on deformable models: recovering 3d shape and nonrigid motion. *Artificial Intelligence Journal*, vol. 36, pp. 91–123, 1988.

[31] L. Cohen and I. Cohen. Finite-Element Methods for Active Contour Models and Ballons for 2D and 3D Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 15, no. 11, pp.1131-1147, Nov. 1993.

[32] J. M. S. Prewitt. Object enhancement and extraction. *Picture Processing and Psychopictorics*, New York: Academic, 1970.

[33] L. S. Davis. A survey of edge detection technique. *Computer Graphics Image Processing*, vol. 4, pp. 248-270, 1975.