UNIVERSITY OF SOUTHAMPTON

Automated Segmentation of Lumbar Vertebrae for the Measurement of Spine Kinematics

by

Yalin Zheng

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Engineering and Applied Science
Department of Electronics and Computer Science

June 2003

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE
DEPARTMENT OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

Automated Segmentation of Lumbar Vertebrae for the Measurement of Spine Kinematics

by Yalin Zheng

Low back pain is a significant problem in the western world and its associated cost is enormous. It is the second most common reason for a visit to medical practitioners. Diagnosis of the underlying causes can be extremely difficult. Since mechanical factors often play an important role, it can be helpful to study the motion of the spine. Traditional radiographic methods of obtaining spine images suffer from high radiation dosage and can yield only a limited number of images. In order to overcome this problem, digital videofluoroscopy was introduced. This can provide motion sequences of the lumbar spine with many frames, but the images obtained often suffer due to noise, exacerbated by the very low radiation dosage. Thus determining vertebra position within the image sequence presents a considerable challenge.

There have been a number of studies on vertebral extraction from fluoroscopic images, varying from manually locating the vertebral landmarks to template matching. However, the former method poses problems of repeatability which can lead to errors in kinematic analysis, while the latter cannot cope with the large changes in the illumination and contrast amongst the frames. In this thesis, extended forms of the Hough transform, together with a recent method of low level feature extraction (phase congruency), have been used to solve this problem. Phase congruency allows for better low level feature extraction, but even this is sufficient to reconcile errors which needs a higher level interpretation. The generalised Hough transform can be used to extract arbitrary shapes, but can suffer from discretisation errors. In the new approaches to the Hough transform, Fourier descriptors are used to describe the vertebral body shape. This description was incorporated within the Hough Transform algorithm from which the affine transform parameters, i.e. scale, rotation and center position can be obtained. The method has been applied to images of a calibration model and to images from a sequence of a moving human lumbar spine. The results are encouraging but sometimes difficulties can be experienced in the extraction because of the extremely poor image quality.

A new spatio-temporal Hough transform has been developed which can improve performance by incorporating the contextual information within the image sequence. An energy function that represents the compromise between the contextual information and the Hough space has been constructed. Rather than finding the maxima in the Hough space like traditional Hough transform, here the Genetic Algorithm is employed to search for maxima of the energy function. The application results on nine subjects show better performance than the traditional Hough transform. The results are promising and show potential for object extraction from poor quality images, and that models of spine movement can indeed be derived for clinical application. As a generic approach, the spatio-temporal Hough transform has potential application in other fields.

Some kinematic parameters have been determined from the data extracted by using the spatio-temporal Hough transform. In order to investigate the validity of the new method, a comparison between results from the automated segmentation and those from a manual labelling was conducted and the statistical analysis shows a good match. As such, new approaches have been developed for vertebrae extraction in spine motion analysis leading to automated extraction of parameters of clinical interest.

Contents

List of Figures

List of Tables

Symbols and Abbreviations

ω	the basic frequency
ρ	rotation
$a_{xk}, b_{xk}, a_{yk}, b_{yk}$	terms of Fourier coefficients
$\omega(t, s, \rho)$	transform kernel
$S_{DH}(\mathbf{b}, s, \rho)$	accumulator array
$E_{spatial}$	energy of spatial information
$E_{temporal}$	energy of temporal information
E_{Hough}	energy of Hough space
$w_1, w_2, w_3, \alpha, \beta, \gamma, \zeta_x, \zeta_y, \zeta_\rho$	weight factors
κ	curvature

LBP	Low back pain
BP	Back pain
SLBP	Specific low back pain
NSLBP	Non-specific low back pain
DVF	Digital videofluoroscopy
MRI	Magnetic resonance imaging
CT	Computerised tomography
SPECT	Single-photon emission computed tomography
PET	Positron emission tomography
IVA	Intervertebral angle
NZ	Neutral zone
ROM	Range of motion
HT	Hough transform
STHT	Spatio-temporal Hough transform
GHT	Generalised Hough transform
PHT	Parameterised Hough transform
RHT	Randomised Hough transform

VHT	Velocity Hough transform
FDs	Fourier descriptors
GA	Genetic algorithm

Glossary of Terms

Arthritis: an inflammatory condition that affects joints. It can be infective, autoimmune, traumatic in origin.

Coronal: having the direction of the coronal suture.

Degenerative: characterised by deterioration or change from a normally active state to a lower or less active form, especially of body tissue, as in a disease process.

Digitized videofluoroscopy (DVF): a system for acquiring low-dose motion radiographs. It consists of an X-ray source, an image intensifier and a video camera, which allows capture of dynamic images of the structure of interest (in particular the lumbar spine) at low radiation exposure. The captured images can be stored on videotape or digitised and stored on a computer.

Goniometer: an instrument for measuring angles.

Hernia: protrusion, as of an intestine or other organ from its normal position.

Image intensifier: a device for amplifying the intensity of a X-ray image for recording.

Inclinometer: an instrument used to measure the angle of thoracic prominence, referred to as angle of trunk rotation).

Invasive: of or relating to a medical procedure in which a part of the body is entered, as by puncture or incision.

***In vitro*:** in a test tube, culture dish, etc.; hence, outside a living body, under artificial conditions.

***In vivo*:** within the living organism.

Noninvasive: of a diagnostic or therapeutic procedure: not involving the disruption of body tissues.

Sagittal: of or pertaining to the longitudinal plane dividing an animal into right and left halves.

Scoliosis: lateral curvature of the spine.

Spondylolisthesis: forward displacement of one of the lower vertebrae over the vertebrae below it or over the sacrum, causing severe pain in the lower back.

Stenosis: narrowing. In spinal stenosis, the spinal canal, which contains and protects the spinal cord and nerve roots, narrows and pinches the spinal cord and nerves. The result is low back pain as well as pain in the legs.

Acknowledgements

I would like to express my profound gratitude to my supervisors Prof. Mark S. Nixon and Prof. Robert Allen for their guidance and encouragement throughout my studies. Thanks to them for the Monday coffee to inspire and stimulate me in the right direction.

Special thanks also goes to Dr. M. Kondracki for providing me with the fluoroscopic image sequence data and for his manual labelling of the data. I would also like to acknowledge my colleagues for their invaluable criticisms and suggestions. I thank Dr. Michael Grant and my research mates Mr. James B. Hayfron-Acquah and Ms. Vijay Laxmi for helping with the proof reading of this thesis.

I also gratefully acknowledge the University of Southampton for offering me full scholarship.

I am very grateful to all members of the Image, Speech and Intelligent System (ISIS) Group in the Department of Electronics and Computer Science at the University of Southampton for providing enjoyable working environment.

Finally, I wish to express my profound gratitude to my dear parents for their prayers and support.

To my wife
Tianqi

Chapter 1

Introduction

1.1 Background

Low back pain is a very common problem in the world and humans have struggled against backache for many years (Allan and Waddell 1989). Compared with other diseases, such as cancer and infection, low back pain has been (and continues to be) one of the incompletely understood, and as yet unsolved, problems in modern medicine even though considerable technical advances have been made in diagnosis, treatment and rehabilitation. Both the problem and its associated disability have appeared to escalate with time. Statistically back pain is the second most common reason for a clinical visit (North American Spine Society 2001). Nearly 80% of all people over the age of 30 will experience back problems during some period in their life. The cost of low back pain is now enormous. In the United States an estimated direct and indirect cost is \$75 to \$100 billion (Porterfield and Derosa 1998). In the United Kingdom, the situation is similar: the cost has been quoted as £6 billion annually and hence more and more attention has been paid to low back pain in the last 25 years (Department of Health 1999). In 1994 the Clinical Standards Advisory Group suggested that top priority should be given to low back pain (Clinical Standards Advisory Group 1994).

Despite the high occurrence of back pain, diagnosis of the underlying problem remains a major problem. One possible hindrance is that low back pain might be a syndrome of many different diseases and disorders, and the other is the structural complexity of the spine and the difficulty of undertaking *in vivo* diagnosis and testing. There are a number of possible sources of pain, e.g. injured muscles or ligaments, disc hernia, arthritis, cancer or bone infection. Most of them have a direct or indirect influence on the movement of the vertebral bones and can be described as mechanical disorders.

In order to understand the mechanical behaviour, specifically of the lumbar spine, a number of studies has been conducted in which different methods have been used to obtain movement information and again a large variety of parameters have been proposed to depict the motion pattern. Among the information acquisition techniques, radiographic imaging with X-rays is the most commonly used technology. After the images have been obtained, it is essential to locate the positions of the vertebrae for analysing the motion patterns and this is called "landmark location" in most studies. For a long time, this has been carried out manually. This manual work can be very tedious and exhausting while the reliability and accuracy will largely depend on the experience of the marker and thus is error-prone.

1.2 Objectives

This study attempts to bridge the spine kinematics and computer vision fields and it is expected that this multidisciplinary effort can benefit these two areas. Due to the importance of the accuracy and efficiency of the landmarking in kinematic studies, the main objective of this research is to explore an automated landmarking algorithm (the spatio-temporal Hough transform (STHT)). This method should provide better accuracy and has the ability to cope with large image sequences efficiently without the inter/intra observer errors of manual landmarking. We will also investigate the application to spine kinematics based upon the results obtained by using the newly developed STHT.

1.3 Contributions

The main contributions of this work are considered to be in the following aspects:

- The most important contribution is the ability to locate spinal objects (lumbar vertebrae in particular) automatically. The Hough transform has been introduced for spinal motion studies owing to its good performance in handling noise and partial occlusion. Results of this application have been reported in our earlier papers (Zheng et al. (2000a, 2000b, 2001a, 2003, 2002a); Allen et al. (2001)). The extraction results have also been used in a 3-D, real-time spine visualisation system (Zheng et al. (2002b)).

- Another significance of this work is the ability to study the continuous motion of the spine. First, digital videofluoroscopic (DVF) imaging has enabled the possibility of obtaining the continuous motion information of the lumbar spine without excessive radiation risk. Second, the proposed STHT is designed to be suitable for extracting the vertebrae in a motion sequence and, in fact, good performance of the STHT is achieved by incorporating the spatio-temporal information within the motion sequence. Some extraction results by using the STHT have been introduced in our recent papers (Zheng et al. (2001b, 2002a)). As a generic approach, the STHT can be applicable to other motion analyses as well.

- In this study the kinematics of the spine have been quantified, not only by widely-used indices, but also by some new concepts that might provide clinicians with a more intuitive sense of the lumbar spine motion and hence be helpful in diagnosis and treatment of the spinal disorders.

1.4 Thesis Overview

In chapter 2, the anatomy of the human spine is briefly introduced and then the problem of low back pain is discussed in detail where the possible causes of low back pain are presented. So far, spine kinematics has been widely studied for better understanding of the relationship between low back pain and the motion pattern of the spine. The last section of the chapter focuses on issues arising from spine kinematics. The emphasis is on the landmark location problem within radiographic images. This chapter will be helpful for understanding the motive and the clinical background of the research.

Although a number of techniques has been used to capture the motion of the spine, so far radiographic imaging is still the most frequently used technology. Chapter 3 introduces some imaging techniques and gives an analysis of their suitability for spine motion study. To our knowledge, to date DVF imaging is the only practical tool for acquiring continuous spine motion and therefore the focus of this chapter is on DVF. The DVF method, the radiation dosage, the performance and the data acquisition procedures are discussed in detail.

Chapter 4 deals with the edge detection problem. In many applications, edge information is of great interest and how to obtain edge information efficiently and with high quality has been a main concern in image processing. In this study, this becomes an essential part of the processing because the image quality of the DVF is limited by the low radiation dosage. This chapter begins with an overview of edge detection approaches. In many applications, phase congruency can yield good edge information and has many advantages. A comparison of the results obtained from both Canny and phase congruency on DVF images is given and the results reveal the good attributes of phase congruency.

In chapter 5, a short overview of image segmentation techniques is given and then the

Hough transform is introduced in detail. The basic idea of the Hough transform is presented by simple applications to straight lines and arbitrary shapes. How to describe the shape is very important in arbitrary shape extraction. Instead of a discrete description of target shape, here Fourier descriptors (FDs) are introduced to represent the shape in a continuous pattern. The theory of the parameterised Hough transform is given based upon the derivation of FDs from the coordinates of the target contour. To evaluate its performance, some experiments have been conducted. These experiments include applications to a synthetic image and images of a calibration model. For the synthetic image, performance with different orders of FDs, different noise levels and occlusions have been investigated. Finally the approach is applied to DVF sequences and some results are presented.

As stated earlier, the poor image quality of DVF limits the performance of edge detection that in turn affects the evidence gathering of the Hough transform. In the worst case, the Hough transform is unable to locate the vertebrae correctly. These problems have provided the inspiration to develop a new version of the Hough transform. Chapter 6 presents this new method called the "spatio-temporal Hough transform" as the spatio-temporal information are incorporated in this new approach. The mechanism of the new method is very simple, instead of locating the maxima in a single Hough space, the STHT method locates the maxima from an newly defined energy function which contains two parts: one is the Hough space and the other is the information from the contextual knowledge within the motion. A spline is used to describe the smooth trajectories of the lumbar vertebrae and spatial constraints between vertebrae are expressed by a penalty function. By this adaption, the new approach can significantly improve the performance of the basic Hough transform. Genetic Algorithms (GAs) are used to tackle the search problem because of their good performance in optimisation. The basic idea of the GA is introduced in detail. Some extraction results are reported and some issues about the STHT are discussed before the short conclusion.

Chapter 7 mainly focuses on spine kinematics. First, some kinematic parameters used to quantify the motion are discussed. Based on the results from the STHT, some parameters such as rotations and intervertebral angles in the sagittal plane view are obtained. An average motion pattern is derived from the data of nine normal subjects. Finally, a comparison between these results and those from manual landmark location is conducted. Statistically they are very close but the automated method avoids many problems inherent with the manual labelling.

Chapter 8 summarises the research and offers new directions for the future. Some potential developments for the STHT are suggested.

1.5 List of Publications

1. Yalin Zheng, Mark S. Nixon and Robert Allen, Lumbar Spine Location in Fluoroscopic Images by Evidence Gathering, Proceedings of Medical Image Understanding and Analysis (MIUA2000), pp. 45-48, 2000.

2. Yalin Zheng, Mark S. Nixon and Robert Allen, Finding Lumbar Vertebrae by Evidence Gathering, 6th Annual Scientific Conference of the Institute of Physics and Engineering in Medicine (IPEM2000), pp. 95-96, 2000.

3. Yalin Zheng, Mark S. Nixon and Robert Allen, Automatic Lumbar Vertebrae Registration Using the Hough Transform in Digital Videofluoroscopy, 7th Annual Scientific Conference of the Institute of Physics and Engineering in Medicine, pp. 172-173, Physica Medica-European Journal of Medical Physics, vol. 17(3), 2001.

4. Yalin Zheng, Mark S. Nixon and Robert Allen, Automatic Lumbar Vertebrae Segmentation in Fluoroscopic Images via Optimised Concurrent Hough Transform, 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 4 pp., 2001. [Electronic version]

5. Robert Allen, Yalin Zheng and Mark S. Nixon, Measurement of the Kinematics of the Lumbar Spine *in vivo*, International Conference on Biomechanics Combined with Annual Scientific Meeting of Taiwanese Society of Biomechanics, pp. 25, 2001

6. Yalin Zheng, Robert Allen and Mark S. Nixon, Lumbar Spine Motion Analysis via Automatic Segmentation, Proceedings of the Fourth World Congress of Biomechanics, pp. 1, 2002, Calgary, Canada. [Electronic version]

7. Yalin Zheng, Mark S. Nixon and Robert Allen, Lumbar Spine Visualisation Based on Kinematic Analysis from Videofluoroscopic Imaging, Medical Engineering & Physics, vol, 25(3), pp. 171-179, 2003.

8. Yalin Zheng, Mark S. Nixon and Robert Allen, Automatic Segmentation of Lumbar Vertebrae in Digital Videofluoroscopic Images, IEEE Transactions on Medical Imaging, (in press), 2003.

Chapter 2

Human Spine and Low Back Pain

2.1 Introduction

The main aim of this study is to develop a new method for improving the landmarking problem in spine kinematic studies. This chapter introduces the anatomy of the human spine and issues of low back pain. It provides a foundation for later chapters.

This chapter is organised into four sections. The first section provides an overview of the structure and functions of the spine for better understanding of back pain. A typical lumbar vertebra is introduced in order to understand the structure of the lumbar spine. Discussions mainly focus on the low back pain problem. A definition of low back pain is given and then the possible causes are presented. Spinal instability is discussed in detail as it is regarded as being related to abnormal motion. Methods that are used for acquiring motion information for spine kinematic study are discussed and the emphasis is on radiographic techniques as they are most commonly used. Finally the anatomical landmarking problem is introduced and an overview is given of the previous methods that aim to automate the marking procedure.

2.2 The Anatomy and Functions of the Spine

2.2.1 Anatomy of the Human Spine

The spine is a very important 3-D mechanical structure. Together with the ribs and the sternum, it makes up the axial skeleton of the human body. The spine is composed of a series of bones called vertebrae. From the top and downwards, these bones can be described in the cervical, thoracic, lumbar, sacral, and coccygeal regions. The cervical

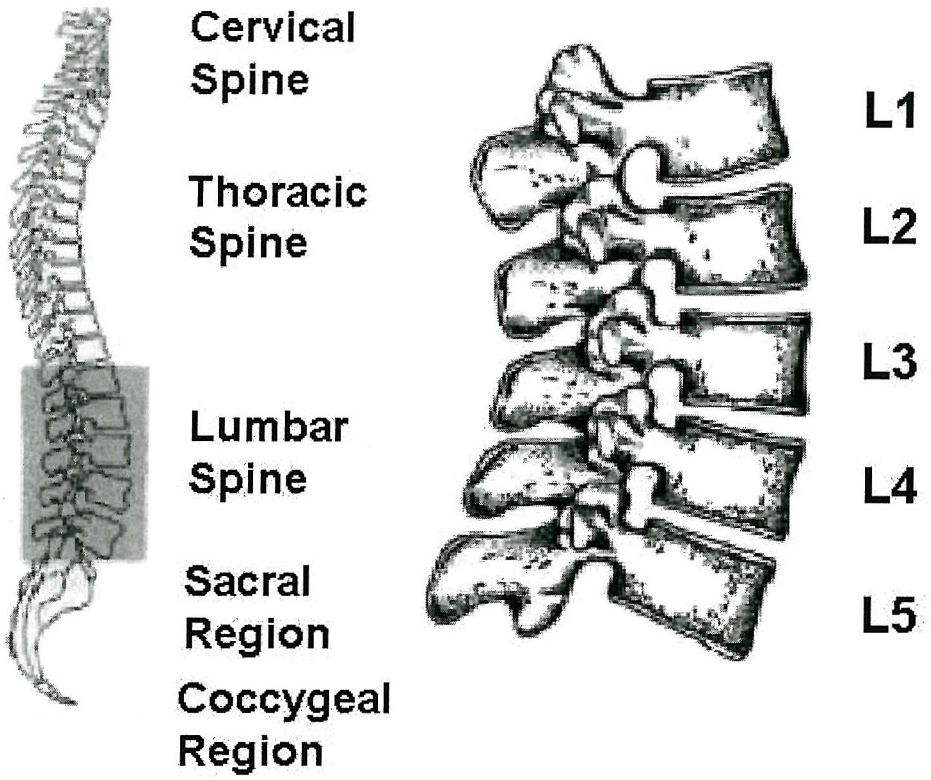


FIGURE 2.1: Illustration of the lumbar spine (adapted from Bogduk (1997) with permission of Elsevier).

spine normally comprises 7 bones, the thoracic spine has 12 bones, and the lumbar or low back region is composed of 5 bones. The sacral region consists of 5 fused bones, and the coccygeal region has 3 to 5 tiny bones, as shown in Figure 2.1. However variations do occur. These vertebrae are connected by a sequence of intervertebral disks to form a strong but flexible support for the neck and trunk. The spine is also stabilised actively by muscles and passively by ligaments that permit twisting and bending movements and also limit other possible harmful movements that might damage the spine itself or the spinal cord (Wynsberghe et al. 1995; Case et al. 1999). Without these constraints the spine would be unstable.

The spine is curved and in profile looks more like an elongated letter "S" in shape rather than a column. The cervical and lumbar regions have a forward curve whilst the thoracic region and sacral region have a backward curve when viewed from the side. The curve of the spine facilitates energy absorption and protects the spinal structures against impact by increasing its strength, helping to maintain a balanced center of gravity in the upright position and absorb shocks from actions such as walking (Wynsberghe et al. 1995).

2.2.2 Functions of the Spine

The spine is a very important, if not the most important structure, in the human body. It maintains three vital functions: support, mobility and protection of the spinal cord.

- **Support** The spinal column is the principal supporting structure. The spine bears heavy loads, in particular the cervical and lumbar regions, including the weight of the upper body and any loads being lifted, lowered, carried or held. These loads can be transferred to the lower limbs through the hips.
- **Mobility** The spine allows a wide range of movements to occur. The cervical and lumbar spine areas are particularly flexible. Forward bending (flexion) and backward bending (extension) of the lumbar spine produce the largest range of movement in the spine. Twisting (axial rotation) and side bending (lateral flexion) are also common spinal movements. In anatomical terms, spine movements can be described in three fundamental planes of motion: the sagittal, coronal and horizontal planes, shown in Figure 2.2. In biomechanics, these movements are commonly defined in relation to three imaginary axes drawn through the body, which are labelled X, Y and Z (Bogduk 1997), shown in Figure 2.3. Also, the motions of the spine are known to be "coupled", or linked each other. Lateral bending (Z rotation), for example, is coupled with axial rotation (Y rotation) and vice versa. A number of studies has been conducted in order to determine the range of these movements. Coupled movements are small, especially during the flexion/extension in the sagittal plane and thus in practice they have often been neglected (Adams 1999).
- **Protection** The spinal column protects the spinal cord and nerves from damage as they pass from the brain to the upper and lower limbs.

2.2.3 A Typical Lumbar Vertebra

The sizes of lumbar vertebrae appear to be the largest in the vertebral column because of their role as weight bearing structures. A typical lumbar vertebra is presented in this section for better understanding of the structure of the lumbar spine and is shown in Figure 2.4. There are variations in vertebral size, shape and detail; in general, however, all vertebrae have the same basic structure. Therefore, it will also be helpful in understanding other regions.

The lumbar vertebrae are irregular-shaped bones consisting of various named parts. There are three main parts to a vertebra: vertebral body (weight bearing structure);

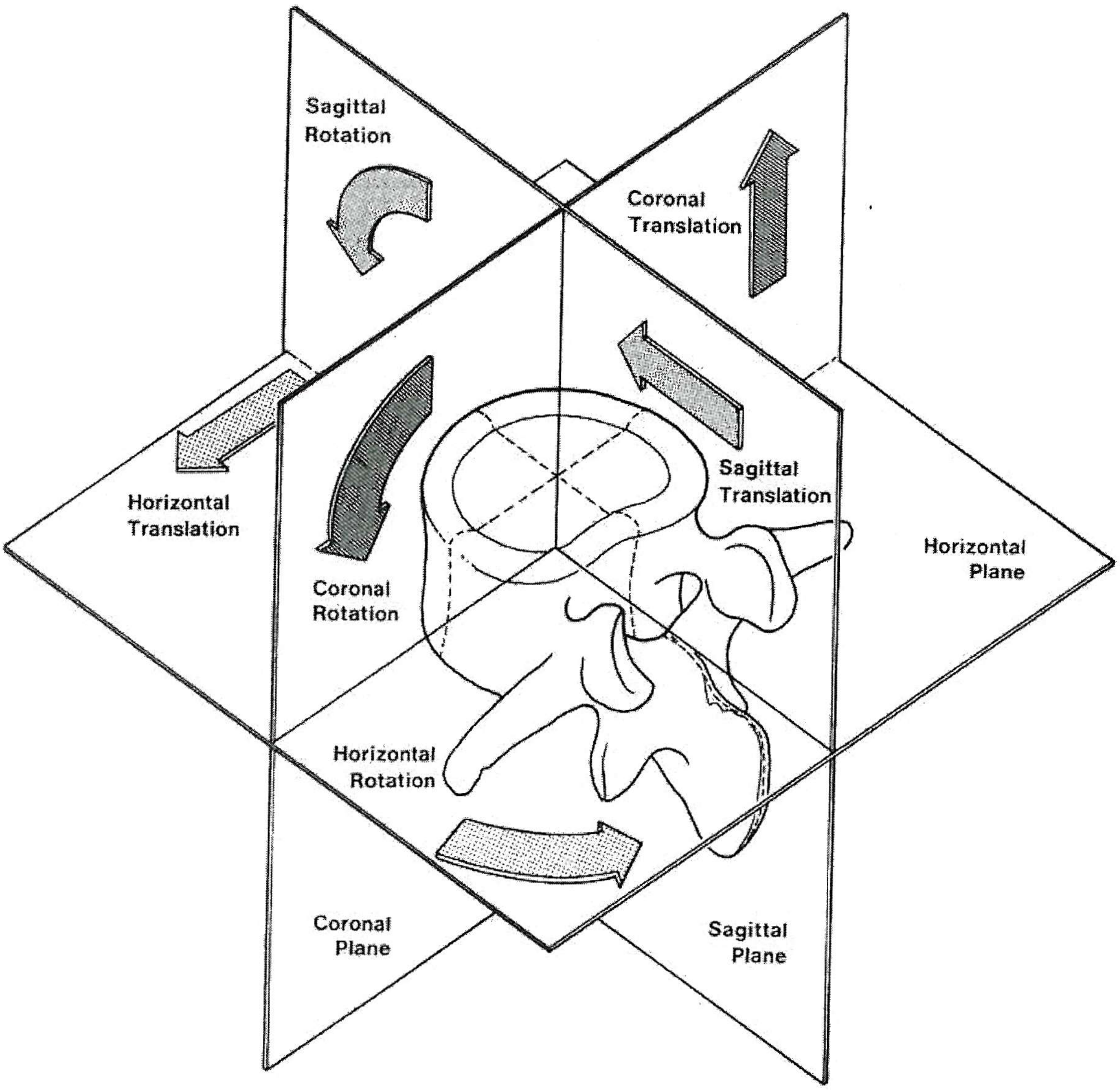


FIGURE 2.2: Planes and directions in anatomical system (adapted from Bogduk (1997) with permission of Elsevier).

vertebral arch (protective in function) and vertebral processes and these are shown in Figure 2.4.

The vertebral body is a thick, disc shaped cylindrical block of bone flattened at the back and with roughened top and bottom surfaces. The inside is made up of spongy bone called a cancellous core that enables it to resist compression, and a thin outer covering of compact bone called the cortical shell that is the weight bearing part of a vertebrae.

The vertebral arch, also known as neural arch, extends backwards from the vertebral body. It consists of two short, thick processes (called pedicles) that stick out backwards

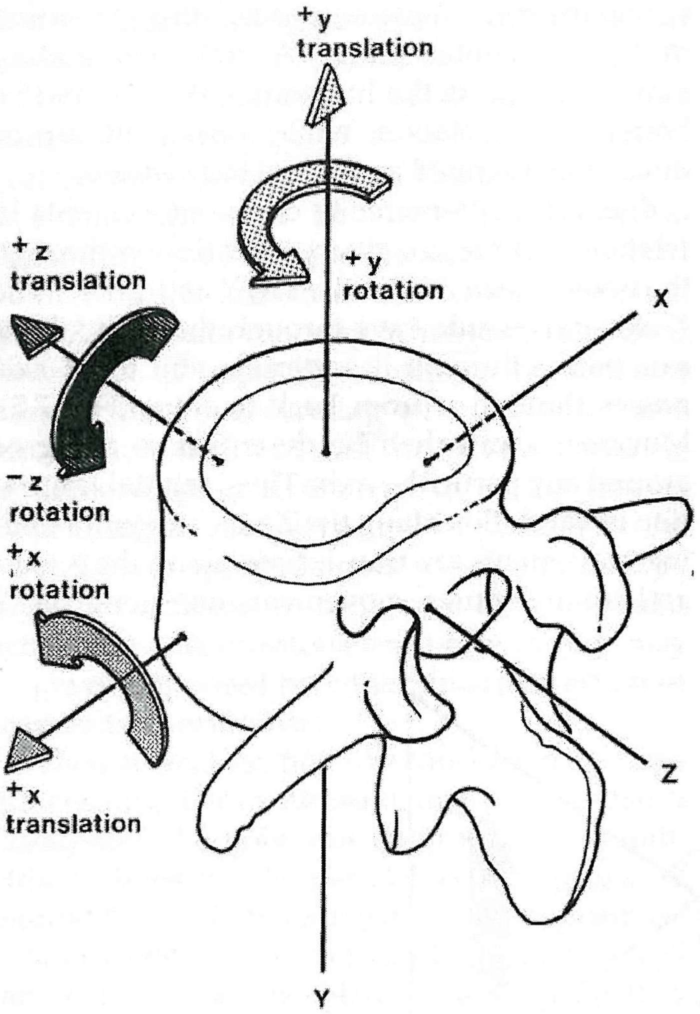
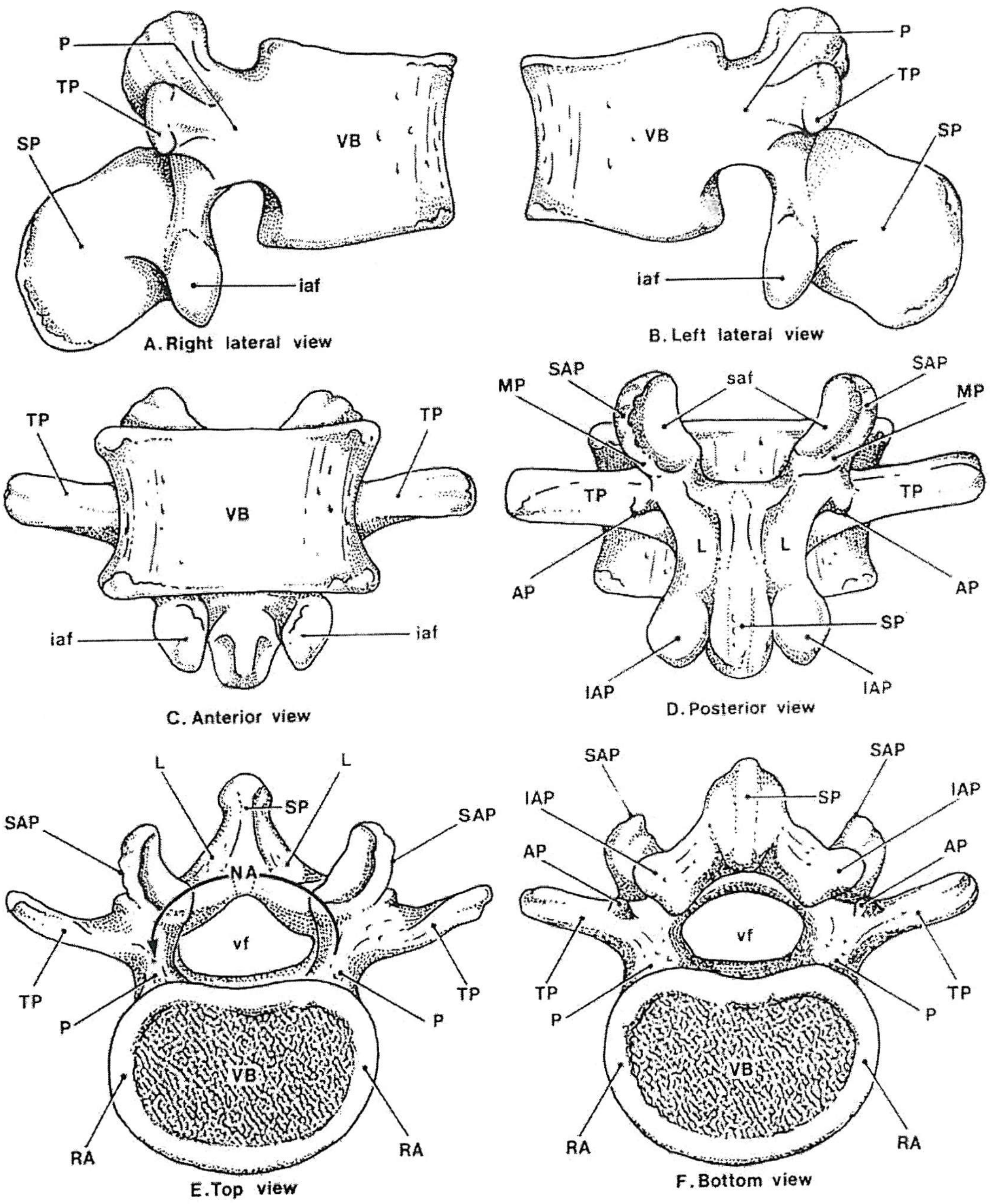


FIGURE 2.3: Axes and directions in biomechanical system (adapted from Bogduk (1997) with permission of Elsevier).

from the vertebral body and join the laminae. The laminae are flat and join together to form the back part of the vertebral arch. Together the vertebral arch and vertebral body surround the spinal column. The space occupied by the spinal column is called the vertebral foramen. The vertebral foramens, when stacked on top of each other, form the vertebral or spinal canal. On each side of the vertebral column there is an opening between each vertebra called the intervertebral foramen which enables the spinal nerves to pass through.

There are seven different processes that come from the vertebral arch. A transverse process extends sideways on each side from the junction of a lamina and pedicle. A single spinous process extends back and downwards from the junction of the laminae. These three processes have spinal muscles attached to them. The other four articular processes form joints with other vertebrae. The two articular processes on top form a



VB-vertebral body; P-pedicle; TP-transverse process; SP-spinous process; L-lamina; SAP-superior articular process; IAP-inferior articular process; saf-superior articular facet; iaf-inferior articular facet; MP-mamillary process; Ap-accessory process; vf-vertebral foramen; RA-ring apophysis; NA-neural arch.

FIGURE 2.4: Illustration of a typical lumbar vertebra. (adapted from Bogduk (1997) with permission of Elsevier.)

joint with the two articular processes on the bottom of the vertebrae above, and the two bottom processes form a joint with the two processes on top of the vertebrae below.

Adjacent vertebral bodies are attached to each other by an intervertebral disc. The disc is composed of a central mass of "pulpy" tissue called the nucleus pulposus and a tough outer covering of fibro-cartilage called the annulus fibrosis. The thinner layers of the annulus at the back make this area prone to damage. The discs are "shock absorbers" by providing resilience to the spinal column as well as flexibility. When a body is erect, the various parts of the disc are under uniform pressure; but when the spine is flexed, extended or bent to the side, one part of the disc is under increased compression whereas another part is under tension.

2.3 Low Back Pain

Low back pain (LBP) is defined as pain perceived as arising from either the lumbar spine or the sacroiliac region or from a combination of both (Bogduk 1997). LBP can be caused by many diseases and disorders and the origin of LBP from multiple organ systems complicates its differential diagnosis (D'Orazio 1999). Furthermore, understanding is also limited by the structural complexity of the spine and the difficulty of *in vivo* experiments: the lumbar spine is especially difficult to access. Up to 85% of back pain patients cannot be given accurate diagnoses (Moffett and Richardson 1995).

Compared with other diseases, such as cancer and infection, LBP has been, and continues to be one of the unsolved problems in modern medicine despite the considerable technical advances in diagnosis, treatment and rehabilitation. Both the problem and its associated disability have appeared to escalate with time. The cost of low back pain is enormous. For example, it has been estimated that chronic low back pain annually results in 225,000 to 300,000 lumbar surgeries and an estimated direct and indirect medical cost of $75 to $100 billion in the United States (Porterfield and Derosa 1998). Statistically, back pain is the second most common reason for a clinical visit in the United States (North American Spine Society 2001). Nearly 80% of all people over the age of 30 will experience back problems during some periods of their life. In the United Kingdom, the situation is similar: it costs £6 billion annually and more and more attention has been paid to LBP in the last 25 years (Department of Health 1999).

Back pain can be classified as acute or chronic. Acute back pain typically gets better within 3 months' duration whilst chronic back pain persists longer. The key distinction is not the duration of the pain, but the persistence of chronic pain longer than the expected healing time and the intractable nature of chronic pain (Waddell 1999).

Low back pain can be further divided into two main categories as specific low back pain (SLBP) and non-specific low back pain (NSLBP) (Tulder et al. 1997). SLBP means that there is a defined cause of the pain such as disc herniations, spondylolisthesis, spinal stenosis, vertebral fractures, tumors, infections and inflammatory diseases like arthritis. NSLBP means that there is no specific cause of the pain but currently there may be one or more undetermined causes. In practice fewer than 15% of people with back pain are diagnosed with a specific cause of pain. This suggests that the majority of people suffer from NSLBP where a definite cause cannot be determined. This is further confirmed by the study of White et al. (1999). Many patients with NSLBP suffer from either muscle pain, or facet joint disease or degenerative intervertebral disc. NSLBP is also known as "mechanical low back problem" (Waddell 1996; Bratton 1999; Eisenstein 1999). As the cause of pain is biomechanical in nature whilst the pain may also cause abnormal motion, many attempts have been made to establish the relationships between motion pattern and low back pain. In one study, Marras et al. (1999) used classification techniques to assess the trunk angular motion as subjects flexed and extended their trunks in five different planes of motion. The results demonstrated that the quantitative measurements derived were able to distinguish between asymptomatic subjects and those with low back disorders.

In spine motion studies, a number of concepts has been proposed for describing the relationships between motion and low back pain amongst which spinal instability is the most popular, but also the most controversial.

Instability has been regarded as one of the most common causes of LBP and indeed numerous studies have focused on it over the last fifty years (Knutson 1944; Morgan and King 1957; Dupuis et al. 1985; Gertzbein et al. 1985; Paris 1985; Pope et al. 1986; Stokes and Frymoyer 1987; Gertzbein et al. 1988; Mimura 1990; Weiler et al. 1990; Panjabi et al. 1994; Szpalski et al. 1999). Morgan and King concluded that 25% of all back pain was the result of segmental instability (Morgan and King 1957) and this has been supported by a later study (Weiler et al. 1990). Controversy surrounds the word "instability" when describing potential causes of low back pain (Szpalski et al. 1999), particularly since hypermobility is seen in many patients who have little or no pain. Similarly, many patients appearing with LBP have no demonstrable instability (Eisenstein 1999). Most likely, it seems that instability does not cause pain, at least not directly. It appears more likely that instability is only an indicator of underlying degenerative changes or preceding injuries and it is the latter that causes the pain (Okawa et al. 1998). In the case of instability defined in terms of an excessive range of movements, pain may arise from the spinal structures that normally limit that movement and whose damage leads to the increased range of movement. Links between potentially painful spinal pathology and abnormal spinal movements may then help clinicians to formulate and evaluate their own definition of spinal instability (Adams 1999).

How to quantify instability is another concern. So far, various indices have been proposed to determine the instability such as the range of movement, translations and rotations in different planes. It is unclear which of them is suitable for representation of the instability as well as the standard to divide normal motion and abnormal motion (Gertzbein et al. 1984; Ogston et al. 1986; Pearcy and Bogduk 1988; Panjabi et al. 1994; Okawa et al. 1998). These parameters involved will be discussed in detail later.

In summary better diagnosis probably relies, to some extent, on improving our understanding of the mechanics of the spine and how disorders might reveal themselves during spinal motion. This is also why so many studies have focused on spinal motion, and now it appears that studies have been extended from the motion of the bones alone to the effects of the muscles and ligaments associated with the spine, e.g. Brolin (2002) used a Finite Element (FE) model to evaluate the effect of ligament injuries on the motion of the cervical spine.

2.4 Spine Kinematics

Spine kinematics investigates the spine movement without considering the forces acting on the structures involved, as the forces are difficult to measure *in vivo*. Spine kinematics could be very useful for better understanding LBP and helping its diagnosis, treatment and rehabilitation. The motion of the human spine can be obtained from studies of living subjects or isolated cadaveric vertebral specimens. There is a compromise between them, the former can provide "realistic" but sometimes less accurate measurements because of some uncertainties caused by muscles, ligaments and other parts of the human body, whereas the latter can provide highly accurate but less "realistic" results since the conditions in the experiments can be easily controlled whilst there are differences between the living human spine and cadaveric specimens.

In spine motion studies, different technologies have been explored to capture the motion and they can also be classified as invasive and noninvasive, based on whether or not the experimental instruments entered the human body. Imaging techniques (including X-rays, magnetic resonance imaging (MRI), DVF imaging, etc.) are commonly used to obtain the motion information of the lumbar spine because they can provide intuitive and accurate information of the lumbar spine. These technologies will be discussed in the next chapter.

Apart from the radiographic imaging technique, some direct clinical measurement methods are also used, such as direct palpation, inclinometers and other sensors (position, accelerometer and electro-magnetic sensors). These approaches will be introduced next.

Normally, the spine can maintain its shape with its different curvatures. Back surface curvature obtained by palpation was used to estimate the range of the motion. Paris suggested that manual palpation was sufficient to find instability (Paris 1985). Evidence to validate these claims appears to be lacking, and it seems unlikely that accurate measurement of curvature can be achieved without the aid of other technologies such as X-rays (Stokes et al. 1987).

Goniometers and inclinometers can be used to measure the range of motion and are recommended by the American Medical Association Guides (1987) for assessing the percentage impairment in chronic low back pain, however, their validity was doubted (Nattress et al. 1999). Like palpation, it cannot measure the motion at the segmental level.

Several studies developed methods using pins placed directly in the spinous process. In one study, Gregerson and Lucas (1976) placed pins rigidly in the spinous process of volunteers and measured axial rotation in a variety of activities. Panjabi et al. (1986) and Pope et al. (1987) placed pins in the spinous processes with accelerometers mounted to them, which enables direct measurement of the lumbar spine response to vibration and impact. Steffen et al. (1997) developed a technique to measure the L3-L4 segmental motion pattern in healthy subjects by inserting pins percutaneously into the spinous processes. Electromagnetic sensors are attached to these pins. Thus 3-D real time motion can be obtained and analysed. The main limitations of the aforementioned methods are the risks of infection and pain involved in the experiments.

Mounting of transducers, markers or position sensor on the skin has also been used for the motion measurement (Stokes and Frymoyer 1977; Pope et al. 1986; Gatton and Pearcy 1999). The movement of the skin with respect to the vertebra is difficult to determine and thus the accuracy is problematic. It has not yet been proved whether or not diagnostic equipment relying on the detection of surface movements have adequate accuracy for determination of motion at the segmental level.

2.5 Landmarking Problem

2.5.1 Data Acquisition

In spine kinematics studies, once radiographic images have been obtained, reference points (also referred as landmarks) are used to describe vertebral positions in the images and have to be identified. In order to determine the positions of a 2-D rigid object in images, the positions of at least two fixed points on the object have to be located.

Following the definition of kinematics, kinematic parameters can be easily derived if landmarks have been located in each frame. For example, the intervertebral angle can be obtained by simply calculating the relative position of the landmarks on two neighboring vertebrae. These parameters will be discussed in detail in chapter 7 and their calculations can be found in the literature (White and Panjabi 1990; Frobin et al. 1996; Frobin et al. 1997; Bogduk 1997; Muggleton and Allen 1998).

From these discussions it is not difficult to find that, to date, landmarking is still the essential step in spine kinematics. Any errors caused by this procedure will propagate during the following steps, which can in turn cause large errors in kinematics. There are some basic requirements for selecting landmarks. First, they should be easily recognised and fixed through the motion sequence of interest. In the case of the lumbar vertebrae, four corners of vertebral body have been believed to be the most prominent and visible parts of the bones as they can be seen throughout the series, and hence they have been used in many studies. Second, the methods to locate them have to be easy and quick for use with good repeatability, reliability and robustness, even though the image quality is often poor.

Thus locating landmarks with accuracy and reliability has been an important aim and indeed many methods have been proposed.

2.5.2 Previous Landmarking Methods

Landmarking was originally carried out entirely manually. Due to poor image quality, it is difficult to place markers exactly on the vertebral corners and furthermore, repeatability cannot be assured. After a thorough error study, Panjabi et al. (1992) suggested that marking radiographs manually results in significantly large errors and the manual superposition and marking of radiographic films are major components of errors in kinematic parameters. For example, errors caused by the person who marks the radiographs can be very large. This error can be further classified as intraobserver and interobserver error. The former is caused by the same marker in repeated measurements and the latter is caused in marking the same radiograph by different markers. The reported errors related to them are $1.6°$ and $1.25°$ for flexion-extension, respectively (Dvořák et al. 1991). Despite these disadvantages, most studies have still been using this method since better approaches to cope with the landmark problem efficiently are yet to be introduced (Farrokhi et al. 2002; Kulig et al. 2002).

Some so-called automatic algorithms have been developed. Simonis and Allen (1993) used a template matching method wherein a template comprising the whole vertebral body was defined. It can locate the vertebrae by finding the best match between the template and the whole image. Muggleton and Allen (1997) obtained some improvements by using an annular template, which was defined by outer and inner templates. In this method, only the part between these two templates was correlated with the whole image to identify the vertebrae. This can save some computation time, but it was still based on template matching. Both of them have limitations. First, the whole image has to be searched with a large template, leading to a long computation time. Second, a major assumption is that the vertebral body will have the same dimension during the motion. However, out-of-plane motion normally happens due to the "coupled" nature in lumbar spine movement and consequently this method is problematic. Finally, the method will suffer when the image illumination changes unpredictably and considerably in a sequence. In a recent study, Cardan and Allen (2000) proposed a first pixel algorithm. In this method, four small corner areas of each vertebra are defined in the first frame, and then they are used as templates in the latter frames. Once they are located, corner points can be found by searching the first point on the edge, when a diagonal line moving towards the edge intersects the vertebral image. The authors suggested that it has good repeatability and reliability. However, it still depends on template matching and thus the inherent problems cannot be solved. Furthermore, the corner found might be incorrect since the first points found in different frames may not be the same especially in the case of poor image quality.

2.6 Conclusions

In this chapter the anatomy of the human spine was introduced and in particular a typical lumbar vertebra was presented. Then low back pain and spine kinematics were discussed in detail. Finally, an insight into the landmarking problem associated with radiographic images was given. From this analysis it appears that landmark location is the crucial step in the spine kinematics study and so far there have not yet been satisfactory methods for general application, especially automated ones. These facts encouraged us to develop an automatic method that can handle this problem effectively.

Chapter 3

Digital Videofluoroscopy

3.1 Introduction

Medical imaging techniques began with the discovery of X-rays in 1895. Since then, medical imaging has proved to be valuable in diagnosis medicine. Various imaging techniques, such as plain X-rays, Ultrasound Scan, Magnetic Resonance Imaging (MRI), computerised tomography (CT), single-photon emission computed tomography (SPECT) and positron emission tomography (PET), have been developed for different diagnostic applications and reviews on them are elsewhere (Cho et al. 1993). These imaging techniques have enabled clinicians to have more and more freedom to select a suitable approach for diagnostic tasks. In this chapter, the reason why currently only DVF is suitable for spine kinematic study is discussed by reviewing other possible imaging methodologies. Then the DVF technique is introduced in detail. Some issues within the data acquisition, radiation dosage and subject selection are presented. Possible future developments in imaging technology appropriate for spine motion analysis are then discussed.

3.2 Why DVF

Plain X-ray radiographs are in widespread use and the acquired images are usually of good quality. One lateral view of a human lumbar spine is shown in Figure 3.1(a) and the boundaries detected by Canny (an edge detection method, to be discussed later) are very clear, as shown in 3.1(b). In spine kinematics plain X-rays have been used in a number of studies (Pearcy and Bogduk 1988; Dvořák et al. 1991; Panjabi et al. 1992; Miyasaka et al. 2000). Due to the high radiation dosage, only a limited number of static images can be obtained, usually in the neutral position and at the extreme positions of mobility. Consequently, it is impossible to determine the intermediate states or to describe motion as the spine moves, say, from flexion to extension or in a lateral bending. Okawa et al. (1998) pointed out that instability cannot be accurately identified with only these terminal radiographs, because an abnormal motion may occur in the middle range of spinal motion. Bi-planar X-rays have been developed for measurement of the 3-D spine motion and regarded as the most accurate way to obtain information on spinal kinematics (Suh 1974; Brown et al. 1976; Wilder et al. 1980; Stokes et al. 1981; Pearcy et al. 1984; Mimura 1990). However, they are also limited by the radiation dosage in the same way as plain X-rays and are unlikely to be able to record the whole motion pattern.

(a) X-ray image (b) Edge map

FIGURE 3.1: X-ray image of the lumbar spine (from http://www.scar.rad.washington.edu/RadAnat/LSpineLat.html) and its edge map.

CT is a very important modality for diagnosing spinal pathological conditions (Beutel et al. 2002). Apart from its relatively lower radiation exposure compared to X-rays, due to its versatility and image quality CT provides us with more diagnostic information. It can also demonstrate more pathological changes than conventional invasive and noninvasive radiographic studies of the spine. However, the prerequisite to obtain an optimal spine image for CT is that patients should keep as stationary as possible and thus motion of the patients is strictly limited. On the other hand, CT is usually used to study the transaxial nature of the spine and thus shows "sliced spine". It is unlikely to put patients into the CT scanner sideways, sagittal or coronal plane view of the lumbar spine has to be reconstructed from multiple slices which again will significantly raise radiation risk to the patients as CT also uses ionising radiation. Therefore, CT currently cannot yield movement information. Recently a prototype 4D-CT scanner has been developed and in the future it might have potential for real-time motion study (Saito et al. 2002).

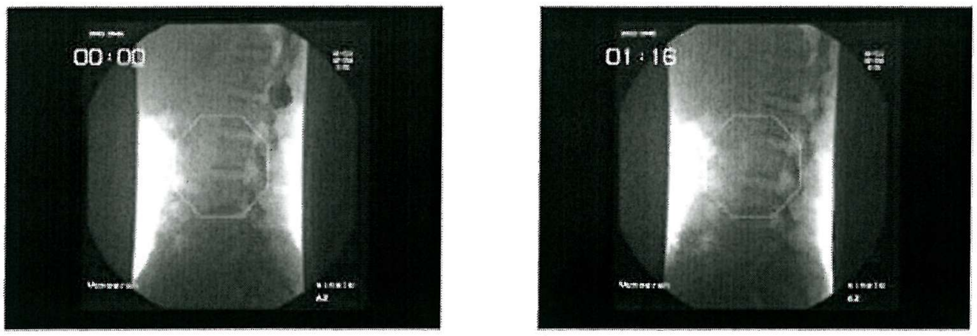
Unlike X-rays and CT, MRI does not use ionising radiation to create images. It can obtain images by measuring the response of hydrogen atoms in the human body under a powerful magnetic field together with a radio frequency pulse. MRI can acquire very clear and accurate images, particularly of soft tissues, but for a long time it has been regarded as not fast enough for motion analysis (Cho et al. 1993). Moreover, it increases the cost of diagnostic imaging for persistent low back pain and has even been considered to be an add-on rather than a substitute for other imaging modalities in the evaluation of persistent LBP (Ackerman et al. 1997). However, MRI has recently been used in some spine motion studies and it appears that it is developing rapidly (Farrokhi et al. 2002; Kulig et al. 2002).

Medical diagnostic ultrasound systems can produce images of soft tissues and internal body organs by capturing scattered echo strength of high-frequency sound waves. It is a safe and non-invasive imaging technique, but unfortunately the images are not yet of sufficient quality to allow detailed investigation of the spinal column in motion.

In short, the aforementioned techniques, so far, are not suitable for a dynamic spine motion study, as they are limited by either radiation safety or by providing only static images. DVF appears to be the only practical imaging technique for documenting the continuous motion of the spine.

3.3 Digital Videofluoroscopy

Breen et al. (1987) first introduced the DVF technique in 1987 for investigation of spinal motion. Since then the DVF has been developed and widely used (Breen et al. 1988; Breen et al. 1989; Cholewicki and Mcgill 1991; Allen et al. 1992; Breen and Allen 1993; Breen et al. 1993; Simonis et al. 1993; Simonis and Allen 1993; Muggleton and Allen 1997; Okawa et al. 1998; Muggleton and Allen 1998; Cardan and Allen 2000). DVF works according to the same principle as plain X-rays but the radiation dosage to which the patient is exposed is much lower due to the introduction of an image intensifier which can amplify the intensity of an X-ray image. A very low dose X-ray machine is used to generate the radiation and the subject is placed between the X-ray source and the image intensifier. The distance between the subject and the image intensifier is kept constant during the image capture. The image obtained by an image intensifier can be stored on videotape for future processing, or digitised and processed directly by computer. There has been considerable progress in improvement and refinement over the prototype system (Kondracki 1998). Two DVF images used in the current study are shown in Figure 3.2.



(a) Frame 1 (b) Frame 2

FIGURE 3.2: Two DVF frames from a lumbar spine sequence.

As with other techniques, there are advantages and disadvantages with DVF.

Advantages

The most important characteristic is that it provides us with the possibility of obtaining real-time, continuous motion sequences of the lumbar spine. This allows us to investigate the whole motion pattern of the lumbar spine rather than that only at the extreme positions as before. The continuous motions at segmental levels can be recorded and enable us to investigate the motion at the level of the intervertebral joint. In addition, and perhaps more important, DVF is much safer for patients than traditional X-rays owing to the great reduction in radiation dosage.

Disadvantages

Limitations of DVF are also apparent. First, the image quality is relatively poor, due to the low radiation dosage. Second, like plain X-rays the quality varies across the image. In the middle, the illumination is brighter while the intensity and contrast are very poor in other areas, which can be seen in Figure 3.2. The main reason for this is that the X-ray source is point-like and creates a cone of radiation that affects, proportionally, all the vertebrae from the middle of the image to the edges. In obtaining DVF images, the X-rays are focused on the regions of L3 in order to keep the whole lumbar spine within view. Another factor is the effect of surrounding soft tissues and bones. Generally, there are more soft tissues shown in the lateral view than in the anterior-posterior view. The composition of the soft tissue itself cannot only affect the image intensity, but also produces scatter which degrades the image further. For example, the L1 vertebral body has very low contrast against its neighboring area due to the effect of soft tissues. In the lateral view, L5 is often obscured by the pelvis and is difficult to be detected visually.

3.4 Data Acquisition

DVF image sequence data used in this work were provided by Dr. M. Kondracki [1]. For details on the data acquisition please refer to his thesis (Kondracki 2001). However, we now discuss some of the main issues relating to the acquisition procedure for better understanding of the processes.

Population Examined. Data from thirty apparently healthy male volunteers (subjects) was going to be collected in his experiment [2]. For the six months leading to the data capture, the subjects should have no history of low back pain. Subjects found to have suffered any significant back pain longer than six weeks or had received any major back surgeries such as spinal fusion, laminectomy or discectomy before the examination were excluded. In this way, and with exclusion due to acquisition errors, 20 subjects were excluded (note that acquisition was centred in a clinical referral unit) leaving the DVF image sequences of ten subjects aged between 19 and 40 years with an average age of 28 years. One subject's data has been excluded because the image quality is very poor. Therefore, results from nine subjects are evaluated in the current study.

Passive Examination. Generally, in kinematic studies of the lumbar spine, both involuntary motion (in which the subject is moved in a controlled manner) and voluntary motion (in which the subject moves himself spontaneously) are used, as shown in Table 3.1. The main objective of kinematics study is to assist in diagnosing mechanical disorders of the lumbar spine. Any technique should therefore be applicable to both patients and normal subjects. Since low back pain tends to restrain the patient from bending as far as his/her spine would allow, the examination of voluntary motion will not uncover hypermobile motion thus making it less useful as a diagnostic tool. We therefore believe that involuntary motion should be performed if a patient is to be examined by means of functional radiographic studies as has already been suggested by Dvořák et al. (1991).

TABLE 3.1: Involuntary and voluntary motion measurements.

Involuntary motion	Dvořák et al. 1991; Panjabi et al. 1992 Lee and Chen 2000; Kondracki 2001.
Voluntary motion	Dimnet et al. 1982; Pearcy et al. 1984; Pearcy and Bogduk 1988 Okawa et al. 1998; Takayanagi et al. 2001.

Passive Motion Table. During DVF acquisition Dr. M. Kondracki (2001) designed

[1]M. Kondracki now is a clinician at the Anglo-European College of Chiropractic (AECC) after he completed his PhD at the University of Southampton in 2001.

[2]Ethical approval has been given for the data collection and study.

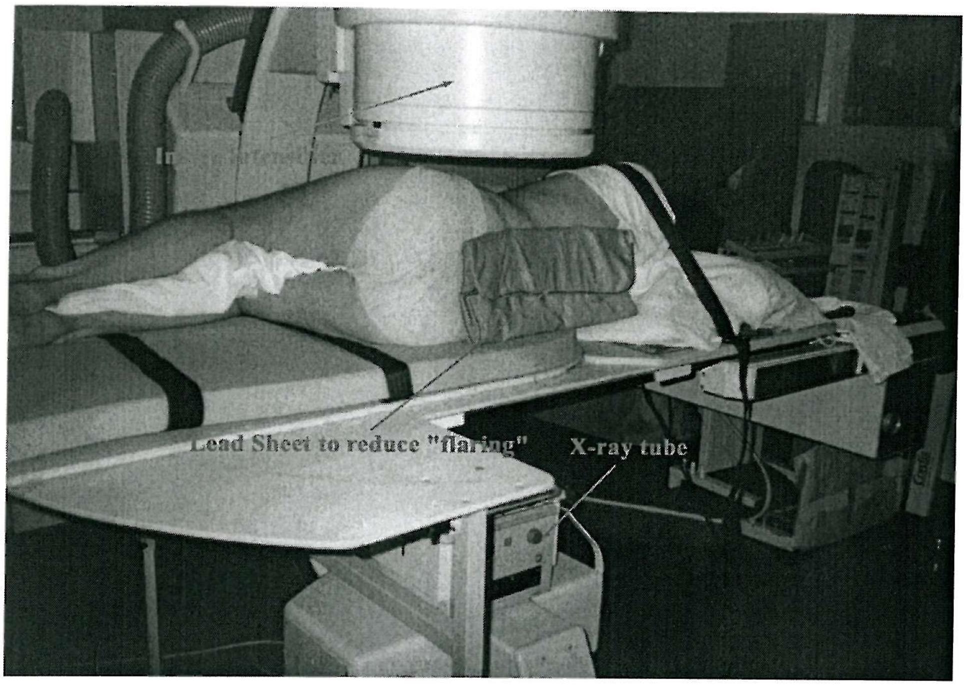
a passive motion table consisting of two main sections namely a two-piece articulating section and a base section. The former, constituting the moving and the static sections, is made of materials capable of allowing X-rays to pass through freely. The latter on the other hand is used mainly to provide support to the table and a surface on which the table can rotate. During data acquisition, subjects were placed on the two-piece articulated section. A motor is used to take the lower part of the articulating section moving through a circular arc of around 90° in approximately 12 seconds. For convenience, the arc is marked in 5° increments.

Technique of DVF Examination. The subjects were made to undergo passive flexion-extensions and lateral bending motions twice with about twenty minutes interval between the screenings. The DVF image sequences were taken from a lateral view in the former and from the anteroposterior view in the latter. The frames were captured at a rate of about 5 frames per second. In each sequence 150 frames are obtained which means that the total exposure time is about 2 minutes.

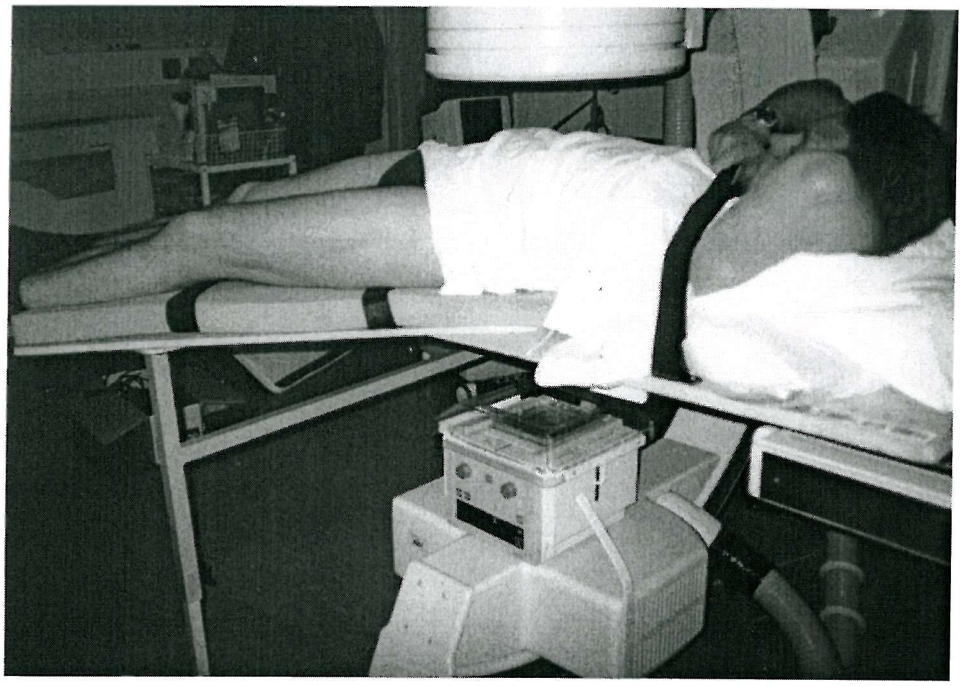
In the current study, the analysis is performed on every fourth frame and this meant 38 frames for each sequence are used, thus reducing the encompassing time span to approximately 0.8 seconds for each increment. In his experiment, the images obtained were directly digitised and saved to a hard disk. This avoided the need for an intermediate video tape as used in the early DVF to store the images. This in turn avoided the image degradation usually caused by video tape. It will therefore be proper to refer to the DVF as digitised fluoroscopy (DF).

An aluminium grid was used to calibrate the distortion in the image. This helped to calculate the resolution of DVF images. The grid was formed by precision-drilled holes of 1 mm in diameter at 1 cm spacing over an area of 14 cm by 14 cm (Kondracki 2001). The image resolution of the pixels in the horizontal and vertical directions was about 0.5 $mm/pixel$ after averaging 5 repeat measurements of pixels between two end points of each row and column.

Flexion-Extension. The subject was placed on the passive motion table on their right hand side, as shown in Figure 3.3(a). During the experiment, an attempt was made to position the L3 of the subject at the interface of two articulated parts by moving him up or down in order to maximally keep the whole lumbar spine within view during the movement (there are still some cases where lumbar spine is out of view and these have posed problem in the segmentation which will be discussed later). Further adjustments were made to make the subject feel comfortable, stable and in an attitude as close to the neutral spinal position as possible. The upper part of the body was secured by running a strap underneath the table and over the subject as shown in Fig 3.3(a). The lower body was left unsecured to alleviate excessive discomfort. Additionally, a pillow was placed



(a) Sagittal view



(b) Anteroposterior view

FIGURE 3.3: Passive motion table, with kind permission of M. Kondracki (2001).

between the knees to help keep the pelvis aligned. In order to avoid over extension of the lumbar spine the knees and hips slightly flexed.

The subject was taken through the movement a number of times before the fluoroscopic imaging to ensure they could cope with the full range of 40° for both flexion and extension. In some cases, the subjects felt quite uncomfortable in full flexion. They were then taken through the sequence under the fluoroscopic imaging. The sequence starts with an initial neutral position and then to full flexion. The subject was then taken to full extension by reversing the direction of movement of the table before the the subject was taken back to the initial neutral position. In most cases, the lumbar spine of the subjects cannot be made to return to the same "initial neutral position" and this was addressed in Kondracki's study (Kondracki 2001).

The X-ray dosage, controlled by the image intensifier, often results in image "flaring" when the subject's trunk fails to cover the extent of the X-ray beam (this often occurs in the extension phase) and the image intensifier received unattenuated rays. In order to eliminate or reduce this problem capable of causing large changes in contrast and blurring of the lumber spine, a flexible lead sheet was placed on the subject's back. This is shown in Figure 3.3(a).

Lateral Bending. A similar procedure to that discussed above was then used to adjust the position of the subject. The subject was placed in the supine position with a small pillow under the knees for comfort and to help reduce excessive extension of the lumbar spine. This helped to prevent any possible coupled motion. The subject was taken into the right lateral bending and then back into left bending before being returned to the "neutral position".

3.4.1 Dosage Analysis

The use of ionising radiation raises important issues in relation to a patient's safety. In an attempt to overcome this, Breen (1991) undertook a pioneer dosage study on the prototype system to determine the absorbed radiation dosage values for a typical patient screening sequence.

These values and, by way of comparison, the dosage associated with plain-film X-rays are shown in Table 3.2 [3]. It is clear that there is a great reduction in radiation associated with DVF when compared with a standard plain X-rays of the same region.

In DVF acquisition the total radiation dosages on different subjects have been analysed

[3] Gy is an international system unit (SI unit) of absorbed dose of radiation, 1 Gy is 1 $Joule(J)$ absorbed per kilogram of irradiated material. That is, 1 Gy = 1 J/kg.

TABLE 3.2: Absorbed radiation dosage (Breen (1991)).

Typical DVF screening (approx. 10 seconds per view)		Typical plain-film	
View	Absorbed dose (mGy)	View	Absorbed dose (mGy)
Lumbar A/P	2.9	Lumbar A/P	20
Lumbar lateral	12.6	Lumbar lateral	50

* A/P means anteroposterior

and are shown in Table 3.3. The average dose-area product across all ten subjects for one screening (anterior/posterior and sagittal) is about 2.55 $Gy\,cm^2$. There is a large reduction over the prototype and is also significantly lower than 15 $Gy\,cm^2$ as recommended by the National Radiation Protection Board (NRPB) of the United Kingdom. Nevertheless, by considering the reduction in dosage levels for common medical procedures, this value is still much lower than the median dose of 7.65 $Gy\,cm^2$ for the lumbar spine plain film (Warren-Forward et al. 1998).

TABLE 3.3: Radiation dosage and screening times for all subjects, adapted with kind permission of M. Kondracki (2001).

Subject index	Total time (min)	Total dose-area product ($Gy\,cm^2$)	Effective dose (mSV)
BM	1.7	6.82	0.99
CR	1.9	3.43	0.46
DE	1.6	5.87	0.68
DO	1.7	4.88	0.61
GD	1.7	3.85	0.44
GP	1.8	5.34	0.65
JM	1.8	4.05	0.50
JW	1.9	5.98	0.70
NW	1.9	4.77	0.57
RM	2.3	6.01	0.62
Mean values for all ten subjects	1.83	5.1	0.62
Average values for screening once	0.915	2.55	0.31

3.5 Conclusions

At present, DVF is the only practical imaging technique to obtain continuous motion images of the lumbar spine at the segmental level with a very low radiation dosage. However, the quality of DVF images is still relatively poor and requires great care in locating landmarks, which is much more difficult than in a plain radiograph. Without exception, this led to problems in the current study and this will be discussed in later chapters.

Indeed, the limitations of DVF encourage new development in radiographic science that might obtain motion sequences with improved image quality but even less radiation risk. This can be achieved in many possible areas. Currently, the technique of cineradiography is also used in spine kinematics (Kanayama et al. 1996; Takayanagi et al. 2001). Pulse X-rays in synchrony with image acquisition should ensure a progressive reduction in absorbed radiation dosage. Improvement in the sensitivity of image intensifiers and shortening of their persistence time will enhance image quality and also safeguard patients from X-ray hazards. Bi-planar fluoroscopy will provide a valuable tool for investigating 3-D motion. These changes, together with 4D-CT and high speed MRI, although by no means a reality as yet, should come into use in the future.

It is worth mentioning that the automated segmentation algorithm proposed in this study is a generic approach and should be directly applicable to the new imaging techniques which are likely to offer improvements in image quality.

Chapter 4

Edge Detection

4.1 Introduction

The boundaries of an object's surface often lead to oriented localised changes in intensity in an image, called edges. Edge detection is regarded as one of the first steps in image processing. There have long been attempts at an optimal edge detection algorithm. This effort has constituted a principal area of research in low-level vision and has led to a number of edge detection algorithms published in computer vision journals over the last 30 years. Even now, new algorithms are still being published.

The edge information is a prerequisite for the Hough transform (HT), which is foundation of this study and will be fully discussed in later chapters. To a great extent, the performance of the HT relies on the available edge information. As has been stated, the quality of DVF images is much inferior to that of X-ray images, thus extra caution is necessary to select an appropriate edge detection method for obtaining better edge information. The edge detection methods can be divided into two approaches, known as gradient-based and phase-based operators, in terms of the information used in these methods. In this chapter, gradient-based operators and phase-based operators (especially phase congruency) are introduced here. Amongst gradient-based methods, the Canny operator is widely used and is sometimes regarded as optimal. Thus, a comparison of results between Canny and phase congruency was conducted and the results showed that in most cases phase congruency can here provide superior results to those of Canny. This has encouraged us to use phase congruency for edge detection in this study.

4.2 Gradient Based Operators

Gradient-based edge operators are the most commonly used methods and they can locate the edges by detecting the changes in brightness of an image. According to Nixon and Aguado (2002) and Sonka et al. (1999), the gradient-based edge operators can be divided into three categories:

- Operators that approximate derivatives of the image function using differences such as Roberts, Sobel, Canny and Prewitt.
- Operators that use the zero-crossing of the image function second derivatives i.e. Laplace and Marr-Hildreth.
- Operators that attempt to match an image function to a parametrical model of edges.

As new algorithms have continuously been developed and these operators are often best for specific kinds of images, in practice it is very difficult to select the most appropriate edge detection method. Amongst these, the Canny operator is one of the most popular methods (Canny 1986), and is designed to provide optimal edges (Nixon and Aguado 2002). With adaptive thresholding, it can provide better results than other complex algorithms (Heath et al. 1997). However, it is virtually impossible to achieve an exact implementation of Canny given the requirement to estimate the normal direction. A common approximation usually consists of a Gaussian filter followed by the Sobel operator with non-maximum suppression and then hysteresis thresholding. A Gaussian filter is optimal for image smoothing and can significantly reduce the response to Gaussian noise. Non-maximum suppression results in thinned edges. Hysteresis thresholding connects the edge points and thus has the ability to detect major features of interest in the edge image.

In order to select a suitable edge detection method for the current study, an experiment has been conducted to compare the results of Sobel and Canny when applied to DVF images. In the experiment, a low threshold and a high threshold were used, respectively[1]. The results are shown in Figure 4.1. When the low threshold was used, too many unwanted edges remain and it is difficult to discern the desired edges (boundaries of the vertebrae) while too few useful edges were left when a high threshold was selected. In the case of high threshold, the edges of L2-4 look good but those of L1 and L5 almost disappear. The reason for this comes from the low contrast and inconsistent brightness of DVF, as discussed in the previous chapter.

[1]Hysteresis thresholding in Canny requires two thresholds, an upper one and a lower one. In this experiment both were set to the same relatively low or high values.

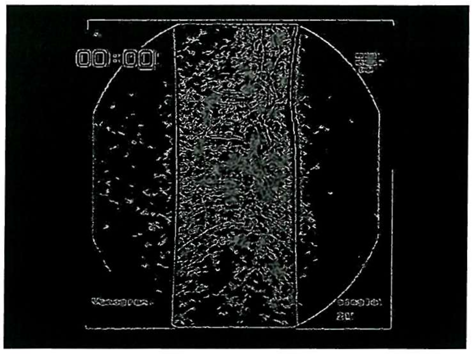
In further experiments, in order to eliminate the problem caused by large changes in brightness within the whole image, the image was cropped into small areas in which a single vertebra is contained. As a worst case, only the L1 and L5 areas were compared. This time only Canny is used as it has been justified that Canny can perform better than Sobel. Apart from low and high thresholds, a manual tuning of the thresholds was also used and is believed to provide optimal edge results. Figure 4.2(a) is the source image of the L1 and Figure 4.2(b) illustrates results of the L1 with manual optimal thresholds, Figure 4.2(c) and 4.2(d) show edges with low thresholds and high thresholds. Similarly, the L5 image and its results are shown from Figure 4.2(e) to 4.2(h). It seems that there is little improvement and problems still exist especially in the region of L5. Even though manual setting of the threshold could lead to acceptable results, it is unlikely to be useful as one cannot depend on manual effort to set appropriate thresholds for a large data set such as DVF sequences.

4.3 Phase Congruency

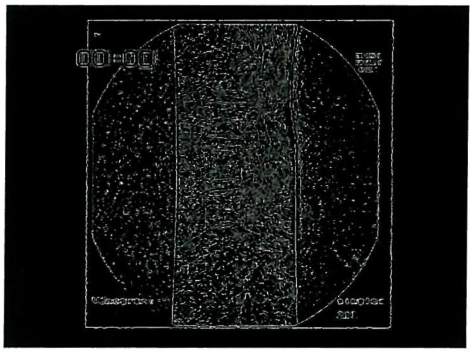
Canny, Sobel and other commonly used edge algorithms perform in the spatial domain, such as finding edges via changes in brightness. There are difficulties in selecting a threshold above which features are considered to be significant. This implies that priority may be given to edges in brighter areas. Setting too low a threshold may yield too much background information whilst setting too high a threshold may mean that features in areas of low contrast are overlooked.

Phase congruency has recently been introduced (Kovesi 1999). It is based on the fact that features are perceived at points in an image where the Fourier components are maximally in phase. A wide range of feature types can cause points of high congruency, these include step edges, line edges and roof edges. Some studies focused on the phase information before phase congruency (Morrone et al. 1986; Morrone and Owens 1987). Human vision is able to respond very successfully to large variations in image contrast and magnification. This may be due to a finely tuned adaptive threshold, but there have been suggestions that detection of phase information plays a crucial role.

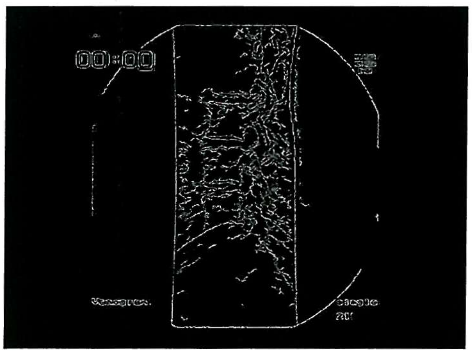
In most applications only the magnitude information of images is needed while the phase information is discarded. Despite this common practice, phase information should not be ignored but it needs more attention because phase carries considerable information of an image. The importance of phase was demonstrated by Oppenheim and Lim (1981). They constructed a synthetic image by combining the magnitude information of one image and the phase information of another. The perceived features in such an image clearly correspond to the phase data, if somewhat degraded. This is illustrated in Figure 4.3.



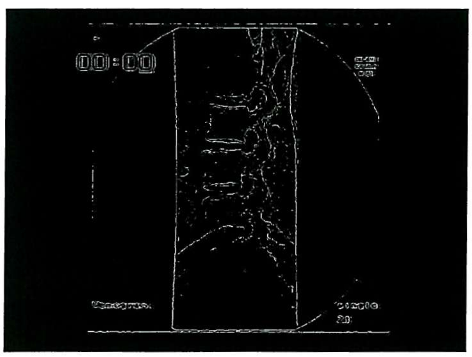
(a) Canny with low thresholds.



(b) Sobel with low threshold.

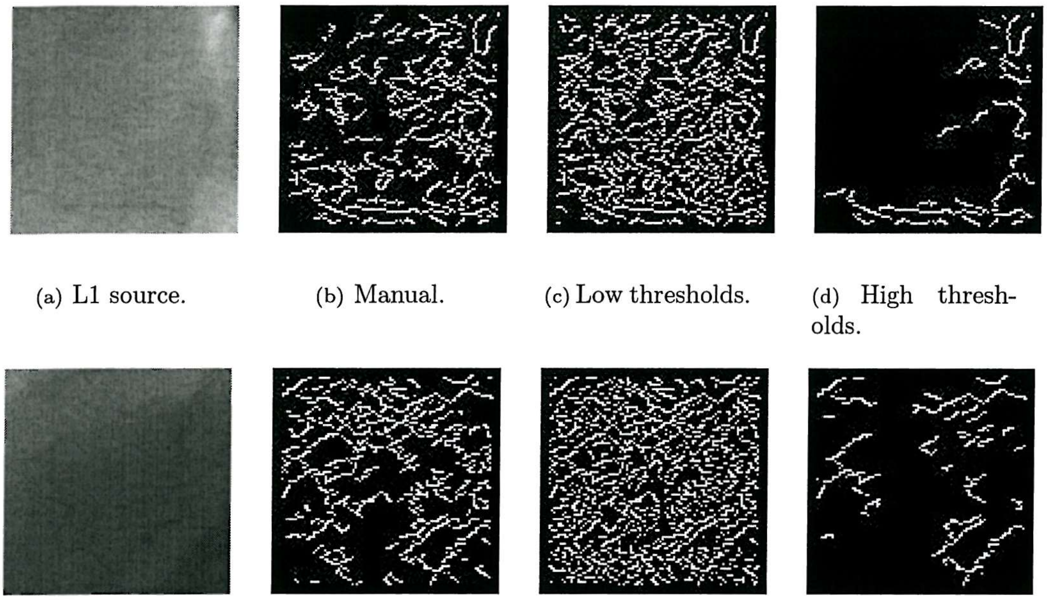


(c) Canny with high thresholds.



(d) Sobel with high threshold.

FIGURE 4.1: Edge image of a DVF image by using Canny and Sobel.



(a) L1 source. (b) Manual. (c) Low thresholds. (d) High thresholds.

(e) L5 source. (f) Manual. (g) Low thresholds. (h) High thresholds.

FIGURE 4.2: Edge detection of vertebrae L1 and L5 by using Canny.

Phase congruency has significant advantages over gradient-based methods. It is a dimensionless quantity that is invariant to changes in image brightness or contrast. Hence it provides an absolute measure of the significance of feature points, thus allows a universal threshold value that can be applied to a wide range of images. Consequently, the problem of selecting threshold for a whole image with uneven brightness, like DVF images, might be solved.

(a) Rice (b) Cameraman (c) Mixture (Magnitude (a)) plus Phase (b)

FIGURE 4.3: Importance of the phase information in image.

4.4 Comparison between Phase Congruency and the Canny Operator

Although various edge detection techniques have been proposed, because of the difficulty in obtaining ground truth for real images, the traditional technique for comparing them is to present image results, side by side, and to let readers subjectively judge the quality. This therefore is not a scientifically satisfactory strategy and makes it difficult to compare different methods. So far, there has been one study conducted by Heath et al. (1998) aiming to change this situation. In this study output edges of four well-known detectors (Canny, Nalwa-Binford, Sarkar-Boyer, and Sobel) were evaluated by a number of volunteers. During the experiment, different tuning factors (e.g. threshold values) were also considered, then the ratings were statistically evaluated so as to find any possible significance. The conclusion was a little disappointing but conformed to reality. That is, Canny usually performed better than others whether or not parameters are fixed or adapted, but in some cases this could be reversed. The important point is that the performance of the detectors largely depended on the images to which they were applied, so did the optimal parameter settings. Thus, in practice, caution is needed in selecting appropriate methods and parameters for specified images.

As many studies have suggested, in most cases, that Canny can yield more satisfactory edge information than other gradient-based methods. Here, to show the good performance of phase congruency, some comparisons have been conducted between phase congruency and Canny. The source code provided by Kovesi (1999) was used for phase congruency whilst the function provided by Matlab Image Processing Toolbox was used for Canny.

First, these two methods are applied to a cameraman picture, as shown in Figure 4.4(a). Values of two thresholds for the hysteresis thresholding used in Canny were automatically set by the function itself (where the upper threshold value is determined by assuming that only 30% of the pixels are edge points and the lower threshold value is 40% of the upper threshold value). In the current study the upper and lower threshold values for hysteresis in phase congruency were set to 0.5 and 0.3. This setting can provide more detail and probably remove more background noise. The results shown in Figure 4.4(b) and 4.4(c) are not distinct as the contrast within the source image is very sharp.

In order to show the superior ability of phase congruency against the uneven illuminations within an image, a new image was obtained from the cameraman picture by doubling the brightness in the left half whilst halving that in the right half, as shown in Figure 4.4(d). Again Canny and phase congruency are applied to it and the results are shown in Figure 4.4(e) and 4.4(f), respectively. Despite the low brightness on the right

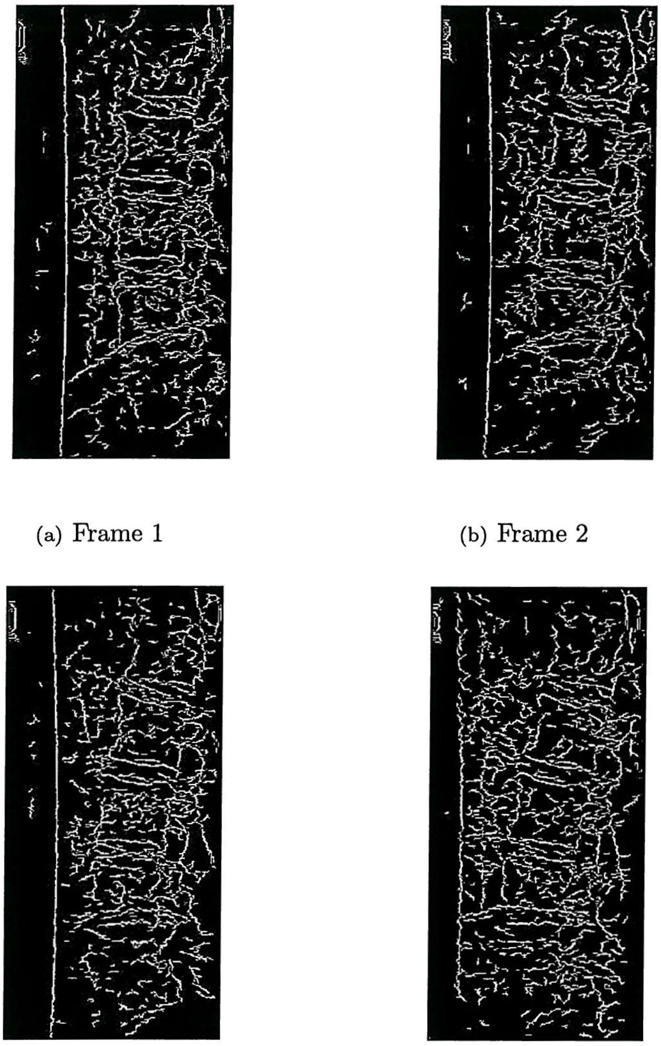
(a) Cameraman (b) Canny (c) Phase congruency

(d) Changed cameraman (e) Canny (f) Phase congruency

FIGURE 4.4: Comparisons between Canny and phase congruency on cameraman.

side, the edges of the high building were still successfully detected by phase congruency, but Canny failed. These results showed the unique feature of phase congruency. That is, there is no need to adjust thresholds for images that have wide variations in illumination level.

Phase congruency has also been applied to some DVF images and example results are illustrated in Figure 4.5. These results are encouraging compared to the results on Figure 4.1(a) and 4.1(c). They provide more edge information but with fewer unrelated noise points. Especially, there is a large improvement on the edges detected in the regions of L1 and L5. It is worth mentioning that all these results were achieved without extra effort for threshold selection. The only limitation of phase congruency might be its speed: phase congruency runs about 15 to 30 times slower than Canny. It seems unlikely to be a problem in future since computer speed continues to increase, and the gain in performance appears worthwhile.



(a) Frame 1 (b) Frame 2

(c) Frame 3 (d) Frame 4

FIGURE 4.5: Edge maps of several DVF images by phase congruency.

4.5 Conclusions

Edge detection represents an extremely important step which can facilitate subsequent high-level algorithms (e.g. segmentation and image analysis) and remains an active research area. In practice there is a compromise between operators. That is, a simpler high level algorithm can be used if good edges can be detected by the edge detection; otherwise, poor edge information will impose a more exacting requirement on the high-level techniques which should be capable of eliminating or handling the problems caused by the poor edge information.

Due to the low quality of DVF images, the most commonly used edge detection techniques are unable to yield satisfactory results even with adaptive parameters. Fortunately, phase congruency can yield better edge results. As ever, there is no panacea: better algorithms have always been desired and may be achieved in future. On the other hand, this is one of the factors that motivated development of a new spatio-temporal Hough transform.

Chapter 5

Parameterised Hough Transform

5.1 Introduction

As stated earlier, landmark location has been a crucial step in the analysis of spine kinematics and there has been no acceptable automatic approach yet. This study aims to develop an automated method by taking advantage of image segmentation which is a very active area in computer vision. Image segmentation is a very important step of image processing which ultimately leads to the image analysis. Its main goal is to find shapes of interest in an image. So far, various image segmentation methods have been developed, ranging from simple methods such as thresholding or subtraction to more sophisticated methods such as template matching, active contours (snakes) and the Hough transform. In this chapter, a brief overview of existing image segmentation methods is presented, then a review of literature on the Hough transform is conducted. Following the introduction to the Fourier descriptors, the parameterised Hough transform (PHT) is described in detail. In order to test the performance of the PHT, some experiments were conducted in which the PHT has been applied to a synthetic image with/without noise and occlusion, and to a calibration model. These tests have yielded promising results. Finally, the PHT was applied to DVF images and some extraction results are reported.

5.2 Basic Segmentation Methods

Image segmentation is one of the most important steps in image processing in which the computer attempts to separate objects from the image background and from each other. A number of algorithms has been developed to achieve the segmentation task and here only a selection is discussed.

Thresholding and subtraction are possibly the most simple segmentation methods. Thresholding assumes that the shape of interest can be identified from other objects or background by changes in brightness. Therefore, it finds the shape by locating the pixels with brightness level in a predefined range of values. Thus correct threshold selection is crucial for successful segmentation. Thresholding is sensitive to changes in the illumination. Moreover, it can only work well when the contrast between the shape and the background is large. Unfortunately, this is not the case in most real-world applications. Subtraction is also very simple to use. It works by subtracting the background from the image. In doing this the background has to be known precisely in advance. Moreover, the subtraction cannot handle noise very well as it has no noise filtering ability. Overall, thresholding and subtraction are attractive owing to their simplicity and efficiency, but they are sensitive to partial shape, noise and variation in illumination. Therefore, advanced segmentation approaches are desired to improve performance.

Template matching is another basic method of segmentation. Namely, a template is defined in which the shape of interest is contained. In implementation this template is centered on an image pixel and the degree of match between the template and the image is checked. Correlation between the template and the image is a general matching criterion. This procedure is repeated for all the pixels in the image. The maximum correlation is regarded as the best match and its position suggests where the target shape is located in the image. If the template is square and of size $M \times M$ and is matched to an image of size $N \times N$, the computational cost involved is $O(N^2M^2)$ for determining the location of the known shape. By rotating the template or by using polar coordinates, template matching can also handle the orientation change of the known shape in the image. Of course, with more parameters involved, the counting procedure becomes more and more complex and thus the computation will greatly increase. In fact, the large computation cost is one of its limitations. Furthermore, in reality the searched images are usually corrupted by noise, geometric distortion and occlusion and it appears that template matching cannot cope with these satisfactorily. These are reasons why those aforementioned automatic landmark location approaches are unattractive (Muggleton and Allen 1997; Muggleton and Allen 1998; Simonis et al. 1993; Cardan and Allen 2000) as they were based upon template matching.

An active contour (or snake) is an energy-minimising curve guided by external constraint forces and influenced by image forces that pull it toward features such as lines and edges. The weakness of a snake is that the result depends on appropriate initialisation and on parameter selection. Moreover, snakes cannot solve the correspondence problems in motion sequence analysis whereby shapes found might differ between frames because of occlusion or noise. Consequently, errors may be generated and would be propagated to the computation of spine kinematics. Furthermore, there is no evaluation of performance when applied to images of poor quality. Hence, snakes might not be an appropriate method for this type of problem.

An active shape model (ASM) is a segmentation method that combines a model of the object contour and the grey-level appearance surrounding the contours. Similar to the snake, it can handle shape variations. In the ASM, the object shape is usually derived from training data that contains the object of interest, by performing principal components analysis (PCA) which is a statistical approach widely used for image dimensionality reduction. Each training example is described by a series of landmark points. The grey-level appearance is modelled in a similar way to shape by using PCA to analyse the grey-level profile in the direction perpendicular to the edge at each landmark point. The ASM has been used to locate vertebrae (Smyth et al. 1997). However, the ASM is limited by the requirement of training data and correspondence problem between the found shapes, as for a snake.

Before moving to the discussion of the Hough transform, it appears necessary to clarify the concept of medical image registration. In medical imaging, registration is a widely used method to achieve spatial alignment between two images (normally from different modalities i.e. MRI and CT) and this is very important in the integration of useful information from these images. For example, Hamadeh et al. (1997) proposed a solution to estimate the motion based on 3-D/2-D registration of a 3-D surface model obtained from CT slices with 2-D functional radiographies. The registration was achieved by using the method of least squares to minimise the distance errors between two surfaces. The method itself might be useful but the system is very complex and requires much computation. Due to the relatively blurred nature of the images, the accuracy of image registration is limited, sometimes large displacement in the registration can be obvious to the naked eye (Maintz and Viergever 1998). Registration between image edges might be helpful to overcome this kind of problem. In fact, medical image registration describes a class of approaches of which some are based on segmentation such as registration using rigid models (point, curves, surface) and deformable models (snakes).

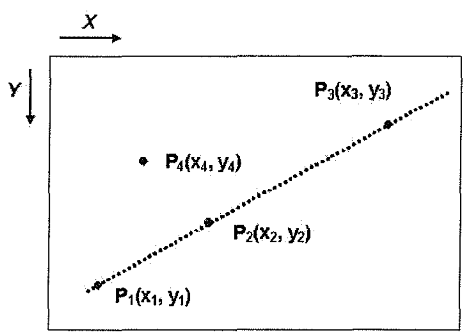
5.3 The Hough Transform

5.3.1 Straight Lines

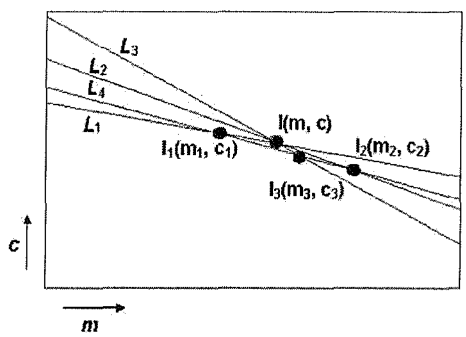
The basic principle of the Hough transform is quite simple. Take line extraction for example, the problem can be described as how to find the straight line given edge points $P_i(x_i, y_i)$ in an image as shown in Figure 5.1(a). A straight line can be expressed in the slope-intercept form: $y = mx + c$, where m is the slope of the line and c is the intercept on the y axis. Any line can be characterised by parameters m and c.

There is an infinite number of lines that could pass through a single isolated edge point $P_i(x_i, y_i)$. Each of these lines can be defined by a particular combination of (m, c). We can characterise each of the possible lines that pass through this point as having coordinates (m, c) in a slope-intercept space (also referred to as the parameter space, accumulator space or Hough space), this is shown in Figure 5.1(b). For example, the possible lines through point P_1 are defined by a line L_1 in the (m, c) space. If this procedure is repeated for other edge points, there will be more lines formed in the (m, c) space. Lines formed by any two edge points will intersect in the accumulator space, like L_1 and L_4 will go through $I_1(m_1, c_1)$. If the number of intersections (votes) at the given point (m_i, c_i) is n, this means that n feature points lie on the line defined by (m_i, c_i). In this way, the Hough transform can locate the line by finding the (m, c) pair corresponding to the maximal number of votes in the accumulator space, like $I(m, c)$ in Figure 5.1(b). In other words, the Hough transform works by letting each feature point (x, y) vote in the (m, c) space for each possible line passing through it and this is why the Hough transform is often referred to as evidence gathering. In implementation, the slope-intercept space (m, c) is converted into an accumulator array by quantising it into finite intervals or accumulator cells. As the HT proceeds, the accumulator cells that lie along the straight line (m, c) are increased. Resulting peak values in the accumulator array represent strong evidence that a corresponding straight line exists in the searched image.

Note that during these procedures, every point in image space (x, y) corresponds to a line in parameter space (m, c) and each point in (m, c) space corresponds to a line in image space (x, y), this feature is highlighted as principle of the duality in geometry (Aguado, Montiel, and Nixon 1998).



(a) Image containing a line



(b) The slope-intercept space

FIGURE 5.1: Illustration of line extraction.

An inherent weakness of the Cartesian parameterisation is that m may become infinity (i.e. a vertical straight line) and in this case no array can be set with enough elements. This problem was solved by the polar parameterisation introduced by Duda and Hart (1972), which has bounded parameter ranges.

The idea of line extraction can be easily extended to other analytical shapes such as circles and ellipses. However, the dimensionality of the accumulator space becomes larger as more parameters are required. For example, for a circle three parameters (radius and locus of the center) are needed while four parameters (locus of the center and lengths of major axis and minor axis) are needed for an ellipse (and another one is needed if there is orientation). The complexity of determining the maximal vote increases in proportion with the increase in dimensionality.

5.3.2 Arbitrary Shapes

In reality, natural shapes are not always able to be described in analytical form. Here, following the idea of Merlin and Farber (1975), an intuitive introduction will be given to show how the Hough transform can cope with the arbitrary shapes. For simplicity, only the problem where translation parameters needs to be determined was considered.

Suppose that we have a known object where its contour is used as the model (or template) in the Hough transform implementation, shown in Figure 5.2. The problem is how to determine its position within an image. As we know, for a known 2-D rigid object, if the location of a reference point to it is known and then the position of this object is uniquely determined when there is no orientation and scaling change. Such a reference point can be selected at random, but normally the center of mass is used, like O in Figure 5.2.

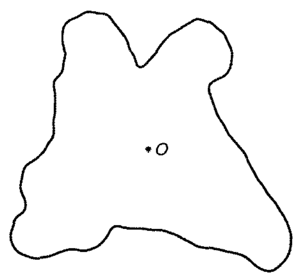


FIGURE 5.2: The curve to be found.

Given an edge point A in the image, shown in Figure 5.3(a), it can belong to an infinite number of different traces of the contour, some of them are shown in Figure 5.3(b). For each trace, there is a possible center denoted as O_i. The locus of all possible candidates O_i can be easily obtained by saving all possible positions of O_i for possible traces.

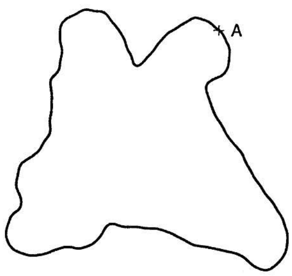
Actually this locus is just the template contour rotated by 180° and centered at A, shown in Figure 5.3(c).

By repeating this procedure for other possible feature points, more possible loci of the searched contour can be obtained, shown in Figure 5.3(d). Like straight line extraction, it can be found that there are intersections between different loci of the possible centers. The location of the center can be determined by looking for the maximal intersections and this means that the known object has been segmented from the image. Similar to straight line extraction, an accumulator array can also be introduced in order to keep the intersection numbers of the loci formed by different feature points. The translation parameters are determined by locating the maximal number of votes in the accumulator array. From this aspect, it can again be found that the Hough transform is gathering evidence from the feature points.

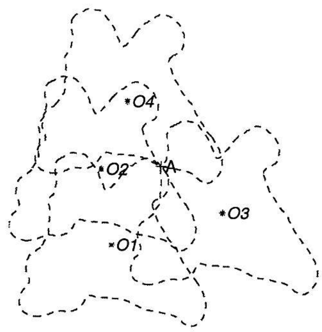
Owing to the mechanism of collecting evidence from all feature points, the Hough transform is very robust to noise. This can be intuitively explained by Figure 5.3(e). Because of the random nature of noise, noise points can only produce loci randomly, this will only increase the overall values of the accumulator array cells but is unlikely to dominate the evidence gathering procedure.

The Hough transform can be extended to more complex problems such as translation together with orientation and scaling change. This means that more parameters are needed to be determined. As stated earlier, the computation cost will increase in proportion with the increase of dimensionality.

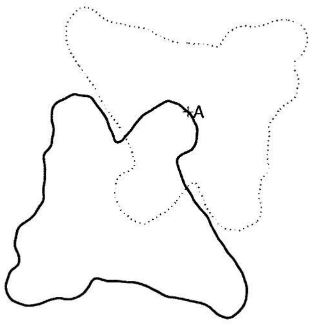
From the above discussion, it can be found that theoretically the Hough transform is capable of extracting any arbitrary shape, however, how to describe the arbitrary shape had not been solved until the generalised Hough transform (GHT) appeared (Ballard 1981). In the GHT, an R-table is introduced to describe the arbitrary shape. Given a 2-D arbitrary curve, for any feature point its edge direction ϕ and radius vector to the chosen reference point (r, α) (where r is the distance between them and α is the angle between the vector and the horizontal) can be obtained. Then the R-table is formed by arranging the points with the same edge direction values in the same row. Since the R-table is a discrete description, when the contour is scaled or rotated there can be problems with aliasing and rounding errors and thus distortions are inevitable. Recently, instead of the R-table, a continuous description has been proposed to eliminate the problems with the R-table (Aguado, Nixon, and Montiel 1998) and will be presented later.



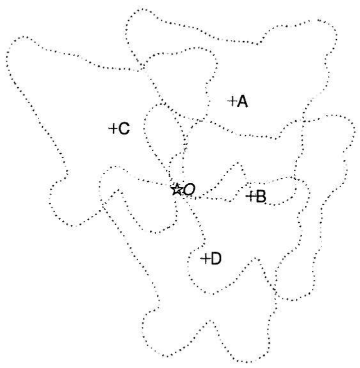
(a) The curve to be searched.



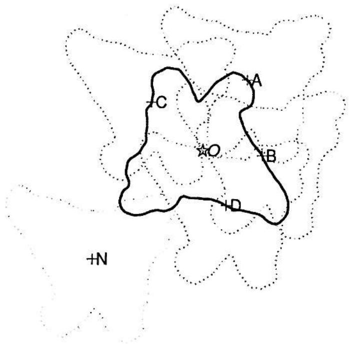
(b) Possible traces that fit the point A



(c) Locus of all possible O for the point A



(d) Accumulate the evidence



(e) Coping with an outlier

FIGURE 5.3: Illustration of arbitrary shape extraction using the Hough transform.

5.3.3 Overview

As stated earlier, the Hough transform is one of the most powerful tools in computer vision. Originally the Hough transform was designed to detect straight lines (Hough 1962). Rosenfield realised its potential as an image processing algorithm and brought it to the mainstream of image processing. Owing to its good performance, the Hough transform was first extended to detect objects that can be described analytically such as circles (Kimme et al. 1975), ellipses (Tsuji and Matsumoto 1978). To solve the problem of detecting an arbitrary shape, a generalised Hough transform has been proposed (Ballard 1981). A number of variations of the Hough transform has been proposed to handle specified problems such as 3-D shape extraction (Wang and Reeves 1990), motion extraction (Nash, Carter, and Nixon 1997) and extracting an arbitrary shape with arbitrary motion (Grant, Nixon, and Lewis 2002).

Simultaneously, a number of studies has been focused on improving the performance in different aspects. As stated earlier, the Hough transform requires much storage and extensive computation. The fast Hough transform (FHT) (Li et al. 1985) and the adaptive Hough transform (AHT) (Illingworth and Kittler 1987) have been proposed in which a coarse-to-fine algorithm is used in the accumulator space to reduce the storage and the time for locating the results. Alternatively, this objective can be achieved by parameter decomposition (or by a multi-stage strategy) so that the dimension of the Hough space can be reduced (Kimme et al. 1975; Illingworth and Kittler 1987; Muammar and Nixon 1991). For a better description of the shape to be extracted, elliptic Fourier descriptors have been introduced to describe arbitrary shapes and will be discussed later (Aguado, Nixon, and Montiel 1998). In order to improve the efficiency of peak search, Genetic Algorithms (GAs) have been used (Goulermas and Liatsis 1998; Ser et al. 1999). Another effort has been made in extracting deformable shapes and similar shapes (Samal and Edwards 1997; Brejl and Sonka 2000b).

Apart from those Hough transforms wherein all feature points are involved in the evidence gathering, there is another method called the randomised Hough transform (RHT). (Xu et al. 1990; Leavers 1992; Xu and Oja 1993; Kälviäinen et al. 1994). The RHT randomly samples the edge data and only a subset of feature points is processed. Evidence gathering stops when the peak value of the Hough space is larger than a predefined threshold. In this way the RHT can save storage and computation time. However, the results cannot be reversed to see which point has contributed to the extraction. Moreover, there is a problem in selecting an appropriate termination criterion. A performance comparison has been made on the existing varieties of the RHT (Kälviäinen et al. 1994). For more details, e.g. shape parameterisation, accumulation in parameter space, peak location, etc., we refer the reader to review two surveys of the Hough transform (Leavers

1993; Illingworth and Kitter 1988).

The unique properties of the Hough transform also provoke applications in vertebra extraction. A GHT based approach has been used to locate cervical vertebrae from X-ray images (Tezmol et al. 2002) where the criterion of reckoning votes has been modified to allow the shape information to be included in the voting process for more accurate segmentation. In this method, a mean shape averaged from 50 images has been used to represent the cervical spine. This method is proposed with the ability to identify variations of the target shape, however, the reported orientation error is 4.16° on average. It appears that this accuracy is not good enough for spine kinematics. Another method called the active Hough transform has also been applied to locate vertebrae in MRI spine images (Brejl and Sonka 2000a). This method is also based on the GHT but claims better performance by incorporating the border information and the variance of the target object contour to form the R-table. These two approaches have three limitations. First, training data is needed to form the contour model. Second, variations found in different images can result in correspondence problems. Third, as with variations of the GHT, the problem caused by discrete description of model shape yet awaits solution.

5.3.4 Further Discussion

Despite the sheer volume of studies on the HT, so far, there is no rigorous mathematical expression yet and only very few studies discussed its definition (Davies 1990; Deans 1981; Princen et al. 1992; Aguado et al. 1998; Olson 1999) amongst these the definition based on the principle of geometry duality is most elegant and attractive (Aguado, Montiel, and Nixon 1998). More surprisingly, there have been very few papers on theoretical discussions of its performance such as bias and aliasing errors (Veen and Groen 1981; Maitre 1986; Hunt et al. 1990; Grimson and Huttenlocher 1990; Kiryati and Bruckstein 1991; Lam et al. 1994; Palmer et al. 1997; Soffer and Kiryati 1998; Aguado et al. 2000).

Sklansky (1978) showed that the HT provides a result equivalent to that derived by template matching. Therefore, the HT inherits advantages such as immunity to noise and occlusion but with less computational effort. The HT, by matching only image edge points to target contour points, requires much less computation than template matching. In particular, the number of edge points increases only linearly with the image size N, not by N^2, likewise with the number of target contour points M. Thus, the complexity of the Hough transform is approximately $O(NM)$.

In summary, the Hough transform is a powerful model based approach in computer vision and its implementation relies on two conditions. First, a model of the object of interest, i.e. line, circle or any arbitrary shape, should be predefined. Second, edges of the image to be searched should be available.

5.4 The Parameterised Hough Transform (PHT)

In a revision of the GHT (Aguado, Nixon, and Montiel 1998), elliptic Fourier descriptors were used to describe the shape. This representation gives a continuous representation that can be sampled at any resolution without the aliasing problems of the R-table.

Elliptic Fourier descriptors are chosen for their completeness, simple geometric interpretation, access to frequency information and the fact that they can be easily obtained from the chain code of the model contour (Kuhl and Giardina 1982). Alternatively they can be derived by applying the Fourier transform to the coordinates of the points on the contour directly. Apart from this description, however, other analytic representations (e.g. a wavelet descriptor) might equally be used.

5.4.1 Fourier Descriptors

Usually, a 2-D curve $c(t)$ can be mathematically described by a vector function, which defines the position of the points in it by their components in two orthonormal axes. That is

$$c(t) = c_x(t)U_x + c_y(t)U_y \tag{5.1}$$

where $U_x = [1 \quad 0]^T$ and $U_y = [0 \quad 1]^T$ are two orthonormal vectors.

During the Fourier descriptor calculation, the overall length of the curve is normalised to 2π. Therefore, the position of the point on the curve t is indexed by radians in the range of $[0, 2\pi)$.

According to Fourier theory, c_x and c_y can be expressed by a Fourier expansion. That is

$$\begin{aligned} c_x(t) &= \frac{a_{x0}}{2} + \sum_{k=1}^{\infty}(a_{xk}\cos(k\omega t) + b_{xk}\sin(k\omega t)) \\ c_y(t) &= \frac{a_{y0}}{2} + \sum_{k=1}^{\infty}(a_{yk}\cos(k\omega t) + b_{yk}\sin(k\omega t)) \end{aligned} \tag{5.2}$$

In Equation (5.2), ω is the basic frequency and k is the harmonic number. The coefficients a_{xk}, b_{xk}, a_{yk} and b_{yk} which are later called Fourier descriptors (FDs), can be computed by the discrete approximation given in trigonometric form as

$$a_{xk} = \frac{2}{m}\sum_{i=1}^{m} x_i \cos(k\omega i) \quad \text{and} \quad b_{xk} = \frac{2}{m}\sum_{i=1}^{m} x_i \sin(k\omega i)$$
$$a_{yk} = \frac{2}{m}\sum_{i=1}^{m} y_i \cos(k\omega i) \quad \text{and} \quad b_{yk} = \frac{2}{m}\sum_{i=1}^{m} y_i \sin(k\omega i) \tag{5.3}$$

where m is the number of total sampling points, x_i and y_i define the values of the functions $x(t)$ and $y(t)$ at the sampling point i.

Thus Equation (5.2) can be expressed in matrix form as

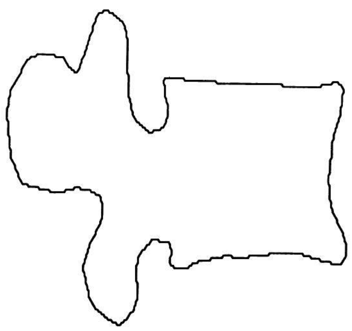
$$\begin{bmatrix} c_x(t) \\ c_y(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} a_{x0} \\ a_{y0} \end{bmatrix} + \sum_{k=1}^{m/2}\begin{bmatrix} a_{xk} & b_{xk} \\ a_{yk} & b_{yk} \end{bmatrix}\begin{bmatrix} \cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \tag{5.4}$$

Given the discrete nature, the possible number of frequencies in the expansion k should be integers between 1 and $m/2$ as suggested by sampling theory, but the determination of the maximal frequency still deserves some discussion especially when the sampling points are very few and the curve has a very sharp corner. For convenience, the DC components a_{x0} and a_{y0} in Equation (5.4) can be omitted since any curve can be defined with its center at the origin of the coordinate system. That is

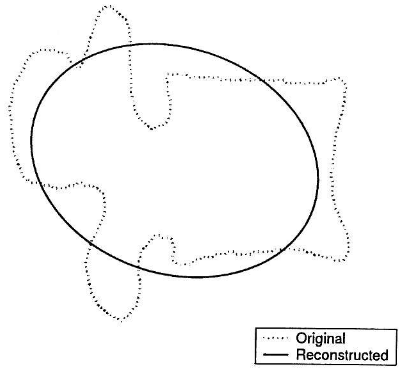
$$\begin{bmatrix} c_x(t) \\ c_y(t) \end{bmatrix} = \sum_{k=1}^{m/2}\begin{bmatrix} a_{xk} & b_{xk} \\ a_{yk} & b_{yk} \end{bmatrix}\begin{bmatrix} \cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \tag{5.5}$$

For a given curve, shown in Figure 5.4(a), its reconstructions with different orders of FDs are illustrated in Figure 5.4(b) to Figure 5.4(k). With more harmonics, the reconstruction is increasingly close to the original curve. Especially from 16 FDs, the difference between the reconstruction and the original curve becomes very small. The values of the FDs are shown in Figure 5.5 and it appears that these values decrease with increase in the FDs' order. This implies that it is possible to reconstruct the original curve by limited numbers of FDs without large errors.

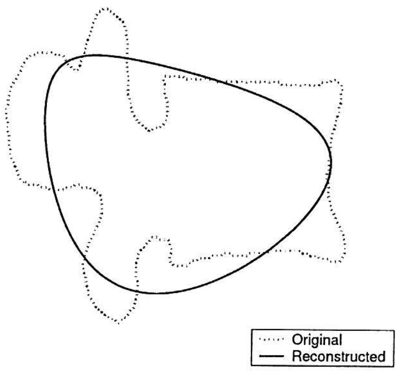
In order to illustrate the possible errors between the reconstruction and the curve itself, another experiment has been conducted. During the calculation of FDs, the indexes in radians corresponding to the sampling points along the curve can be obtained as well. Note, these indexes do not increase linearly because in generating the chain code the distance between the neighboring points is not same but with possible value of 1 or $\sqrt{2}$. After the FDs were obtained, the reconstruction is made at these index values. Therefore, differences between the curve and reconstruction can be compared and Figure 5.6 shows such comparison of the differences with 8, 24 and 48 FDs. Clearly, improvement



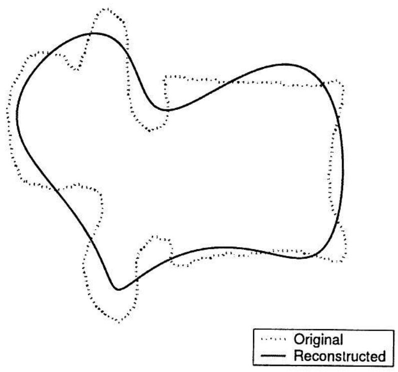
(a) Original curve



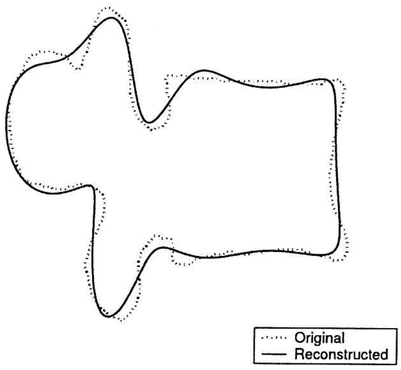
(b) 1 FDs



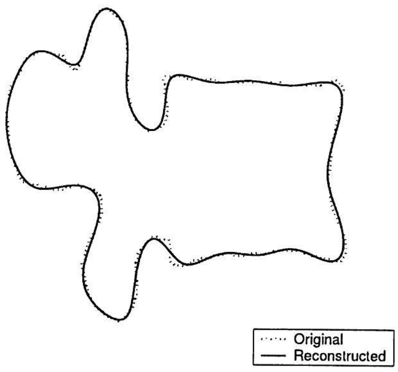
(c) 2 FDs



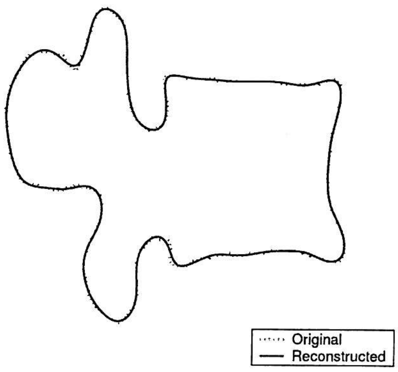
(d) 4 FDs



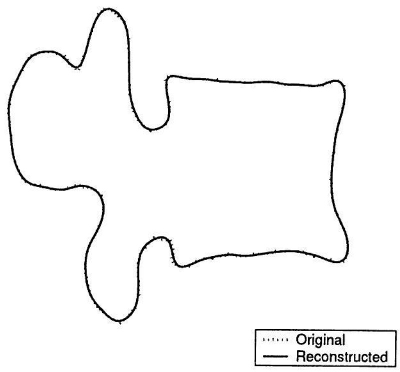
(e) 8 FDs



(f) 16 FDs

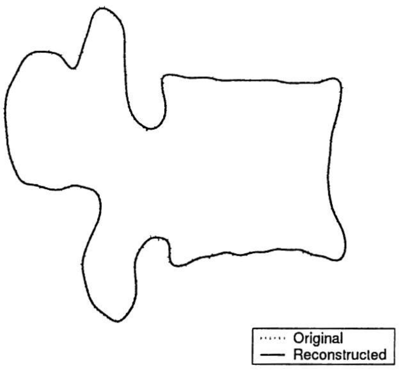


(g) 24 FDs

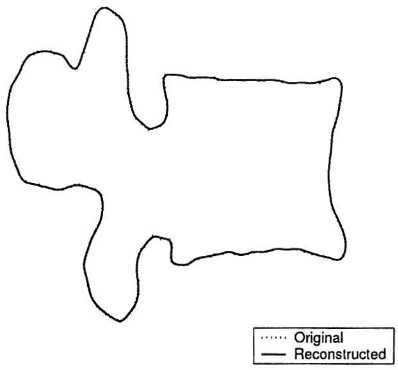


(h) 32 FDs

FIGURE 5.4: Curve reconstruction by different FDs.

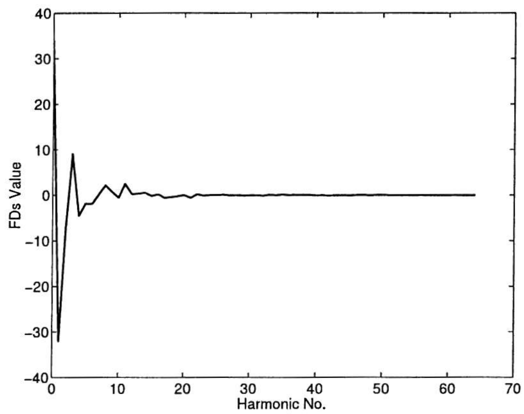


(j) 48 FDs

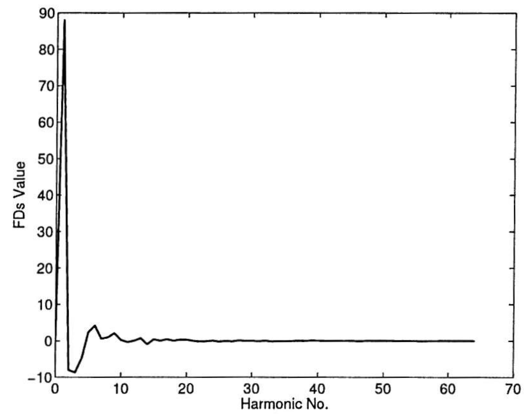


(k) 64 FDs

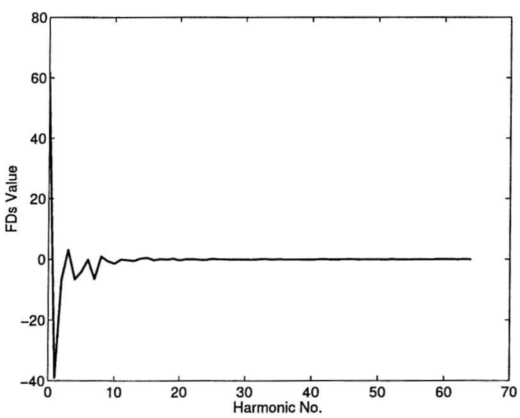
FIGURE 5.4: Curve reconstruction by different FDs (continued).



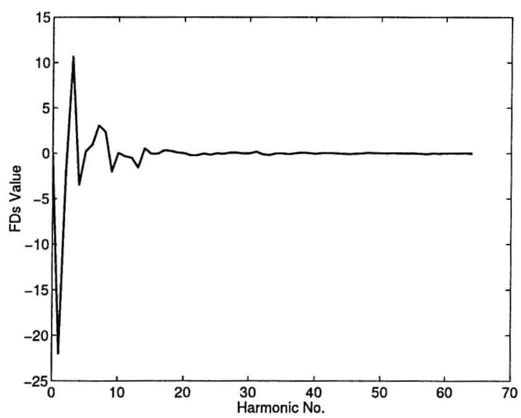
(a) Term A



(b) Term B



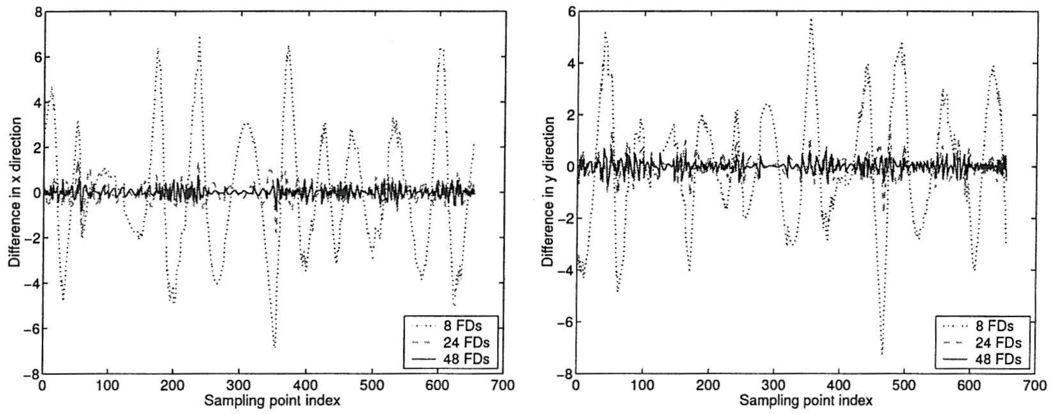
(c) Term C



(d) Term D

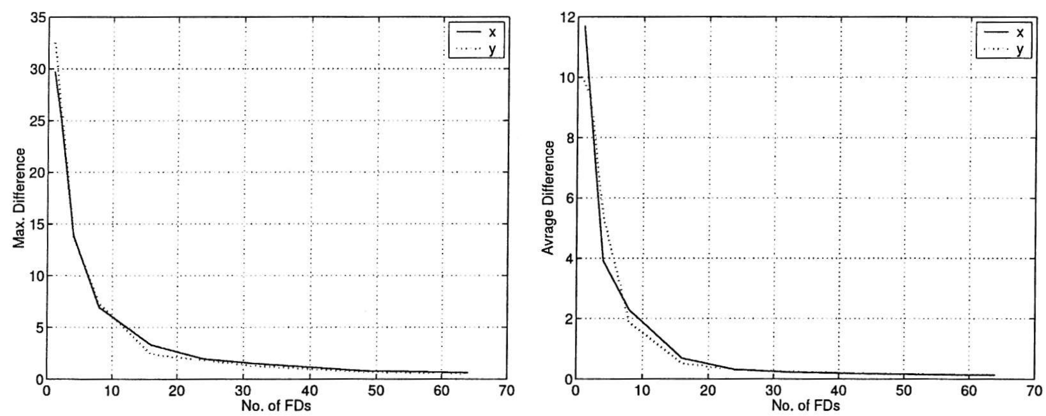
FIGURE 5.5: FDs for the curve.

in accuracy can be achieved with a larger number of descriptors. Another consideration is the possible maximal difference. For each reconstruction, the maximal absolute values of the difference at sampling points was obtained and these are shown in Figure 5.7(a). Similarly, the average differences with each reconstruction were also obtained and are shown in Figure 5.7(b). It denotes that the overall average difference has become less than unity after 16 FDs and the error continues to reduce but only slightly with continuing increase in the number of descriptors. This shows that the FDs can describe the shape well.



(a) x difference (b) y difference

FIGURE 5.6: Difference between reconstructions and the curve itself.



(a) Maximal difference (b) Average difference

FIGURE 5.7: Maximal and average differences between the reconstructions and the curve itself.

The good attributes of FDs, such as invariance to rotation, scaling and translation, and especially access to frequency components, make it a powerful tool to represent curves, whether they are closed or open. In case of an open curve like a straight line, the FDs can be derived by tracing the curve twice in opposite directions.

5.4.2 Description of the Algorithm

Any curve $A(t)$ obtained by applying the affine transformation (only translation, scaling and rotation are considered here, shears can be included as well) to a curve $c(t)$ can be expressed by its two components in the x and y directions. That is,

$$\begin{bmatrix} A_x(t) \\ A_y(t) \end{bmatrix} = \begin{bmatrix} x_t \\ y_t \end{bmatrix} + s \begin{bmatrix} \cos(\rho) & -\sin(\rho) \\ \sin(\rho) & cos(\rho) \end{bmatrix} \begin{bmatrix} c_x(t) \\ c_y(t) \end{bmatrix} \tag{5.6}$$

where s represents the scale factor, ρ is the (clockwise) rotation, and (x_t, y_t) are translations in the x and y directions.

Following the derivation of the HT (Aguado, Nixon, and Montiel 1998), the transformation kernel was defined as

$$\omega(t, s, \rho) = s \begin{bmatrix} \cos(\rho) & -\sin(\rho) \\ \sin(\rho) & cos(\rho) \end{bmatrix} \begin{bmatrix} c_x(t) \\ c_y(t) \end{bmatrix} \tag{5.7}$$

For any edge point $P(p_{x_i}, p_{y_i})$ located by edge detection methods (especially phase congruency, as used in this study), its translation vector can be expressed as

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} p_{x_i} \\ p_{y_i} \end{bmatrix} - \omega(t, s, \rho) \tag{5.8}$$

As introduced earlier, in order to form the accumulator array an evaluation criteria has to be made in order to increase the values of the array cells where there are intersections. Here, a simple matching function was introduced and it can be expressed as

$$H(a, b) = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases} \tag{5.9}$$

Here, a and b can be vectors [1]. Now the HT can be defined in discrete form as

$$S_{DH}(\mathbf{b}, s, \rho) = \sum_{i \in D_i} \sum_{t \in D_t} H(\mathbf{b}, \begin{bmatrix} p_{x_i} \\ p_{y_i} \end{bmatrix} - \omega(t, s, \rho)) \tag{5.10}$$

where $\mathbf{b}$ is the translation vector, and $S_{DH}(\mathbf{b}, s, \rho)$ is a 4-D accumulator space. D_i is the collection of the edge points located within the image and D_t is the domain of the

[1] In terms of signal processing, the procedure to form the accumulator space in line extraction can be regarded as a certain space-variant transform followed by sampling on a rectangular grid. Thus, the commonly used matching function like Equation (5.9) is a discontinuous and non-bandlimited function. Due to aliasing effects a non-bandlimited signal cannot be properly represented by a discrete set of samples and this is one weakness in the implementation of the Hough transform (Kiryati and Bruckstein 1991). In order to reduce/alleviate the aliasing effect, Kiryati et al. (1991) proposed a "bow tie" matching function. In a recent study a "top-hat" function was introduced for line extraction (Palmer et al. 1997). they reported that the new "top-hat" matching function can perform better to locate the true parameters of lines than the commonly used one even when the outliers exist. For arbitrary shape extraction, such functions can also exist but perhaps they will have more complicated forms.

points in the model shape.

In fact, Equation (5.10) is an evidence gathering procedure which was introduced earlier. Specifically, all possible instances of the model in the image, i.e. changes in size, rotation and position, are taken into account. Thus, the real translation vector $\mathbf{b}_0$, rotation ρ_0, and scale s_0 can be determined by locating the maximal value in Equation (5.10), where $\mathbf{b}_0 = (x_{t0}, y_{t0})^T$. That is,

$$(\mathbf{b}_0, s_0, \rho_0) = \arg\max(S_{DH}(\mathbf{b}, s, \rho)) \tag{5.11}$$

The expression of Equation (5.11) that defines the HT using FDs for arbitrary shape extraction was used in this study. Essentially, for a given feature point, a locus of points is plotted through the 4-D accumulator space (translation in the x and y directions, rotation ρ and scale s). The maximal value of this accumulator denotes the position of the object found[2].

5.5 Applications

5.5.1 Synthetic Image

First, the HT was applied to a simple image, shown in Figure 5.8. In this image, the center of the object is actually located at $(60.75, 55.26)$, which can be rounded to the nearest pixel position $(61, 55)$. The true results should be *scale* $= 1$, *rotation* $= 0$, $x = 61$ and $y = 55$. The total number of edge points is 377.

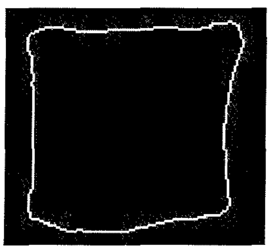


FIGURE 5.8: The synthetic image.

As discussed above, the core of the HT is forming the accumulator space. In this study, by assuming intervals for scale, rotation and translation are unity, one degree and one pixel, respectively, a 4-D array can be formed by discretising the ranges of these parameters according to the above intervals. Each edge point votes in this array. That is,

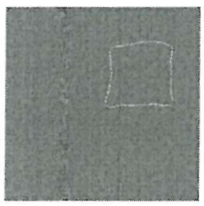
[2]Normally, there will be peak formed in the accumulator even without occurrence of the searched object. In application, a specified threshold value is often defined and only the peak with value larger than this value is regarded as the actual peak.

TABLE 5.1: The extraction results with different Fourier harmonic components.

FDs order	4	8	16	24	32	40	60
Peak	81	155	249	275	279	293	302
Scale	1	1	1	1	1	1	1
Rotation	0	0	0	0	0	0	0
X	62	61	61	61	61	61	61
Y	55	55	55	55	55	55	55

Unit for rotation is degree, unit for x and y is one pixel.

if the parameters obtained are within the range set earlier, the value of the corresponding cell will increase by one. In this way, the array cells are assigned values by all the edge points. Then the parameters can be determined by locating the maximum in the array. In order to illustrate this procedure, the array values were transformed into an image (for convenience, only the vote distribution for translation was shown with scale of unity and rotation of zero degree), in which the brightness of the image point is proportional to the number of votes that cell obtained. Here, 16 FDs were used.



(a) 1 point

(b) 50 points

(c) 100 points

(d) 377 points

FIGURE 5.9: Illustration of Hough space formation.

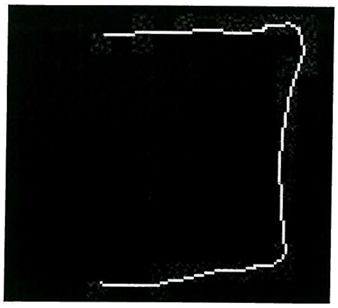
In Figure 5.9, from 5.9(a) to 5.9(d), are the Hough spaces formed by 1, 50, 100 and 377 edge points, respectively. After analysing all edge points, the position of the brightest point is the values of x and y we are looking for. In Figure 5.9(d), the coordinates (x, y) of the brightest point are (61, 55) and this means that the x and y parameter values are (61, 55) as expected. The peak value corresponding to it is 249, which is smaller than the total number of edge points. The reasons for this might be from the discretisation of the Hough space or partially from the approximation of the 16 FDs for the original model.

As we know, reconstruction with differing numbers of FDs will approximate the original model with different degrees of accuracy. An experiment was conducted to demonstrate how this number affects the extraction. Table 5.1 shows the extraction results with different numbers of FDs. Here, reconstruction by 16 FDs gives a good match to the image data. Four and eight descriptors are insufficient, as the peak values are less than half of the number of total edge points. More descriptors, such as 24 and 32 or higher, are unnecessarily complex as the results show no improvement over that by 16 FDs and the vote values increase only slightly. This implies that reducing the number of descriptors can reduce computational cost without sacrificing accuracy. Clearly, the results of 16 FDs are sufficient for a more refined analysis of the match at the vertebral corners, which is the main objective of this approach.

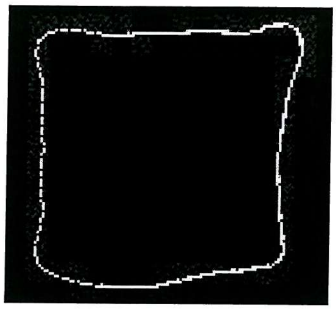
Another concern is to what degree the HT can handle occlusion and noise. Here, instances of the whole shape, 25% and 75% occlusion with different noise levels from 0 to 95% were considered. In this experiment, 16 FDs were also used.

It is difficult to add "real" noise to an image so we attempted to look at noise and occlusion in the worst case. Moreover, with powerful noise removal algorithms, noise can be reduced to minimum in most cases. Therefore carrying out the performance analysis by adding "salt and pepper" noise to the edge image appears to be reasonable. Occlusion was simulated by removing certain chosen parts from the edge image. Figure 5.10 shows 25% occlusion, its results and images with different noise levels and the corresponding Hough space. In the Hough space images, it can be seen that the relatively bright area spreads with the increase in noise level. This means that the signal/noise ratio decreases and the peak is gradually impaired and eventually inundated by the peaks generated by noise. Figure 5.11 shows 75% occlusion as well as its extraction results, images with different noise levels and the corresponding Hough space.

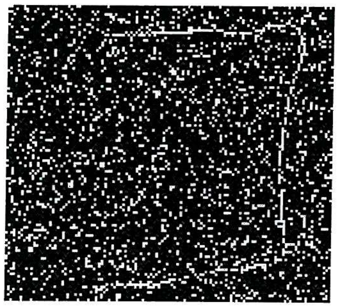
Visually when the noise level is about 60%, the shape cannot be identified by human vision. The results show that the HT can work well until the noise level is 90% for the whole shape. When there is occlusion, this ability is reduced slightly. It can succeed until the noise levels are 85% and 70% for 25% and 75% occlusion, respectively. These results confirm that the HT has a good ability to handle noise and occlusion, and can be useful for application to medical images. It is worth mentioning that the PHT should be able to handle a higher percentage occlusion. However, the percentage of occlusion may only be an effective measure for smooth shapes while it perhaps does not make sense for a large angular shape. For this kind of shape the PHT can still succeed if only one small corner area remains after occlusion. In this case, despite the high percentage of occlusion the main feature is retained and this does not result in extra difficulty to the PHT at all. Therefore, we believe that percentage of occlusion is just a measure of ability and should not be over-emphasised. Moreover, it appears more reasonable to consider occlusion together with noise.



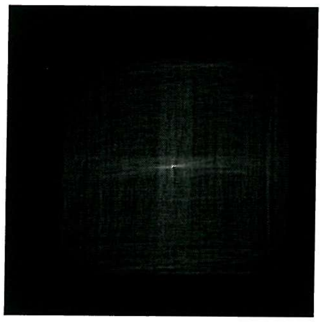
(a) 25% occlusion



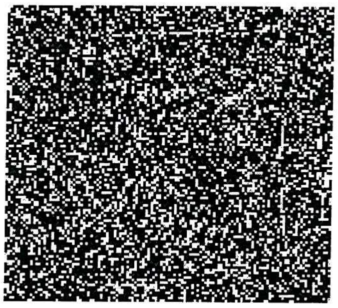
(b) Extraction result



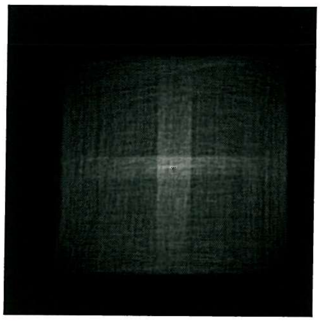
(c) 30% noise



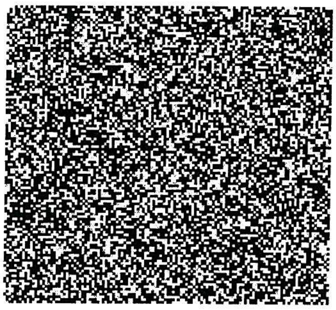
(d) Hough space (30% noise)



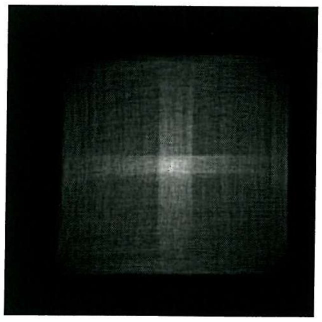
(e) 60% noise



(f) Hough space (60% noise)

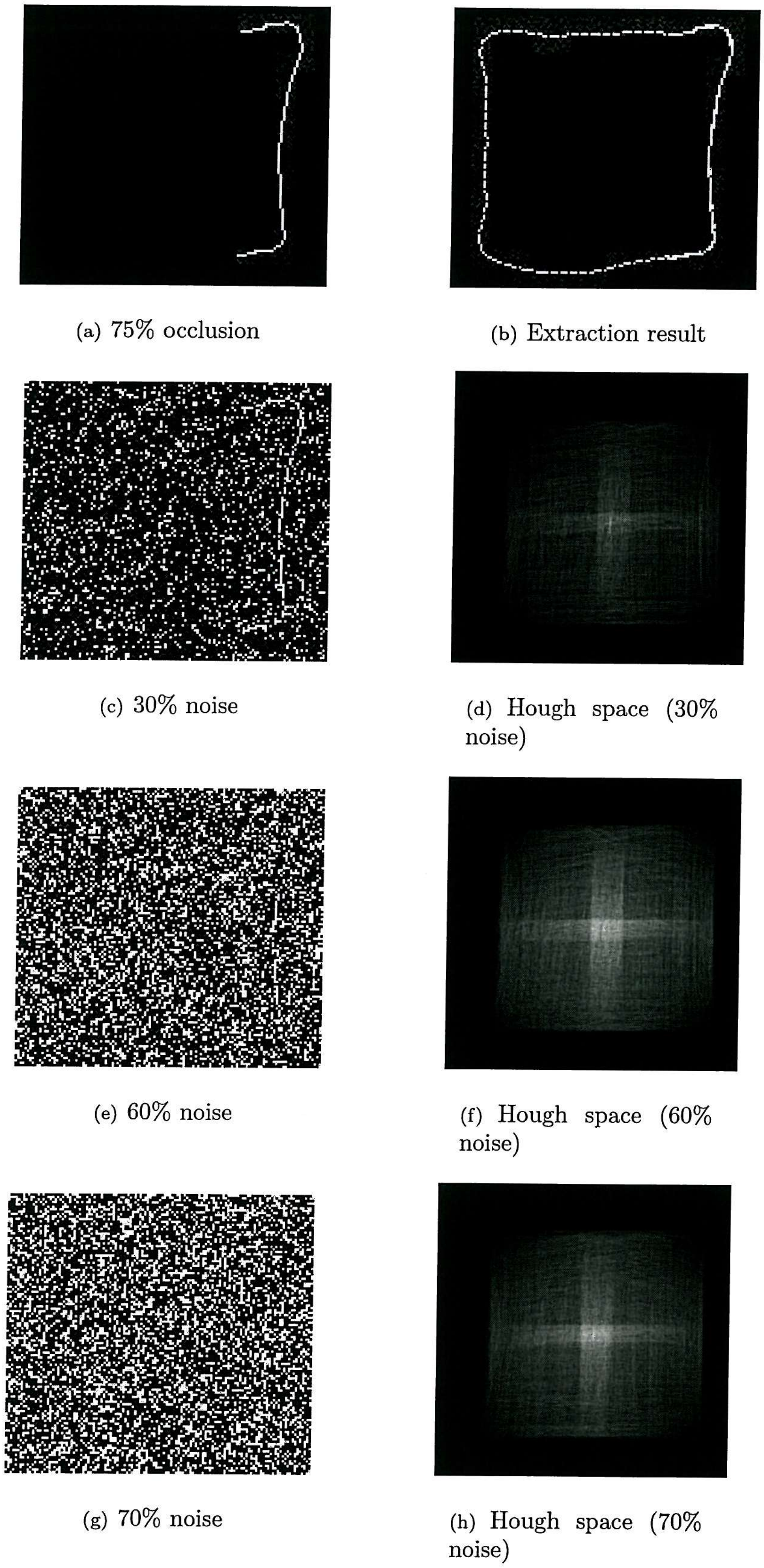


(g) 85% noise



(h) Hough space (85% noise)

FIGURE 5.10: 25% occlusion and different noise levels.



(a) 75% occlusion

(b) Extraction result

(c) 30% noise

(d) Hough space (30% noise)

(e) 60% noise

(f) Hough space (60% noise)

(g) 70% noise

(h) Hough space (70% noise)

FIGURE 5.11: 75% occlusion and different noise levels.

5.5.2 Application to the Calibration Model

When a clinical measuring technique is developed, it is essential to demonstrate its reliability before it can be used in clinical trials. In this study, a calibration model has been designed to examine the validity of the PHT.

The calibration model is comprised of two human lumbar vertebrae (L3 and L4) linked at the position of the centrum of the disc by means of a universal joint, shown in Fig 5.12. A perspex pointer was fitted to one side of the superior vertebral body in such a way as to describe arcs whose centres correspond to the universal joint. A computerised numerical control (CNC) machine tool was used to preset the angular position in 2-D and the accuracy of these preset angles are considered to be within $\pm 1^\circ$. The measurement of these arcs was obtained from a protractor attached to the base of the model. The increment of the preset angles is 5° and the useable range of the protractor was 30°.

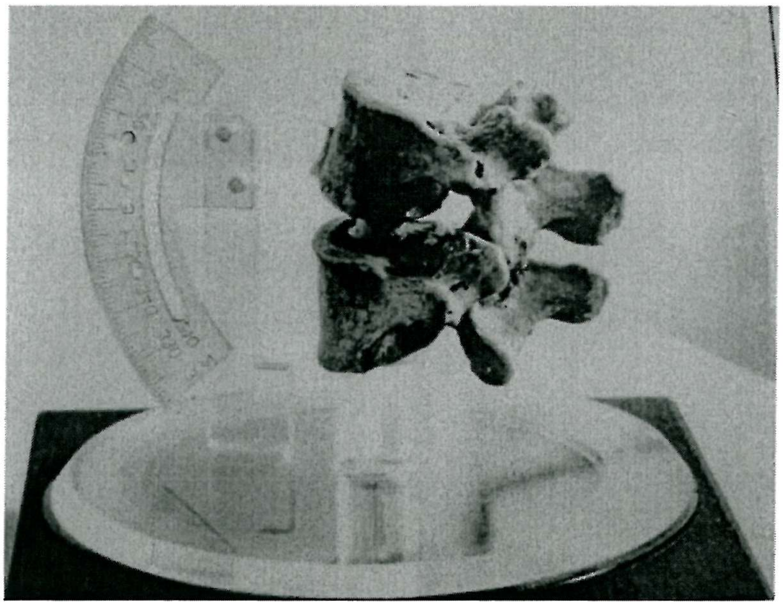
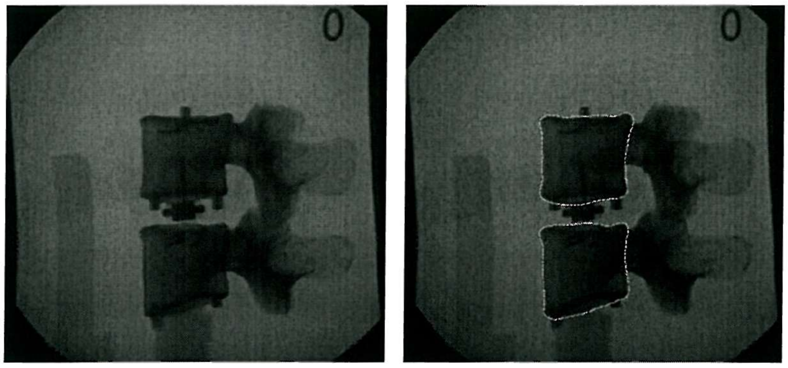
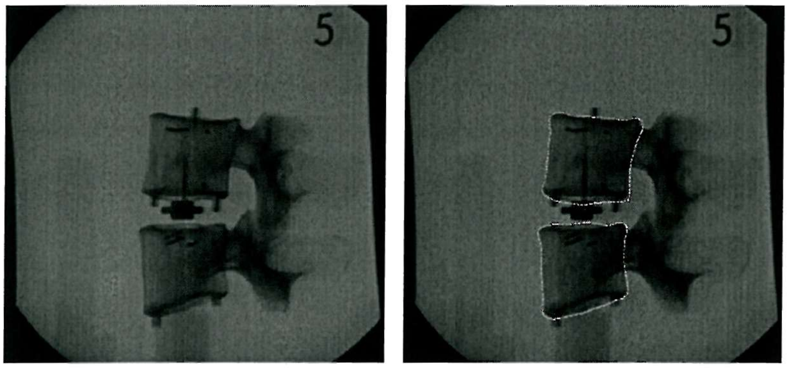


FIGURE 5.12: The calibration model.

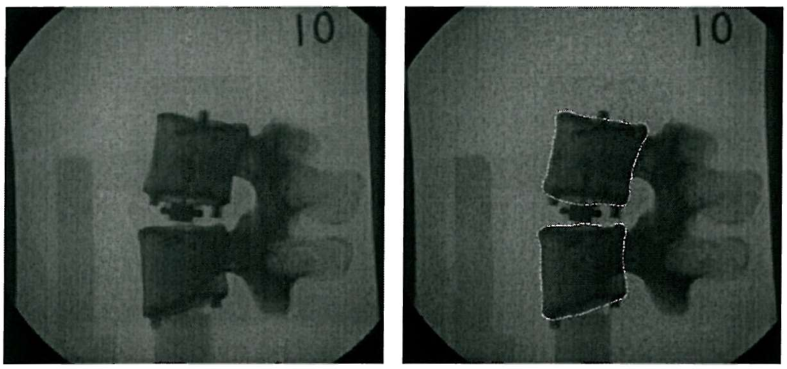
During DVF acquisition, vertebra L3 rotates with increments of 5° with reference to the universal joint whilst the vertebra L4 remains stationary. The model was X-rayed in lateral projection using a focus-to-intensifier distance of $1m$. The results of the calibration studies have been reported (Breen et al. 1993). In brief, by manual landmarking, they established the inherent ability of the technique to measure intervertebral rotations to an accuracy of $\pm 1^\circ$ over a motion range of 30° and instantaneous axis of rotation (IAR) to within $\pm 5mm$ of their true location for rotations of 7° or greater. The left column of Figure 5.13 illustrates these DVF images. The value shown on the top right of each frame is the preset rotation angle. The positive value denotes that motion is in the extension direction whilst negative means flexion.



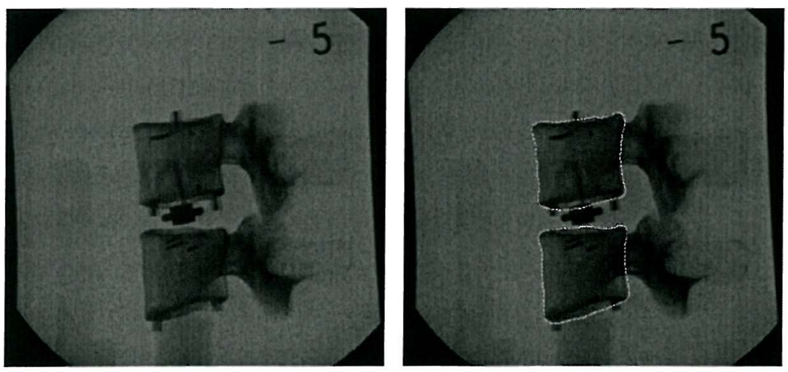
(a) $\theta = 0°$ (b) $\theta = 0°$, extraction



(c) $\theta = 5°$ (d) $\theta = 5°$, extraction

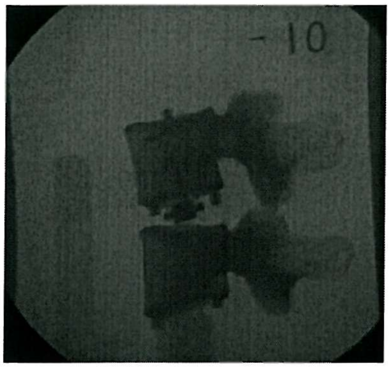


(e) $\theta = 10°$ (f) $\theta = 10°$, extraction

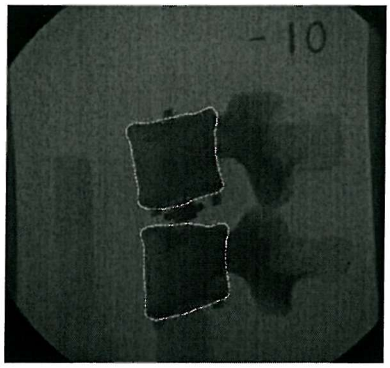


(g) $\theta = -5°$ (h) $\theta = -5°$, extraction

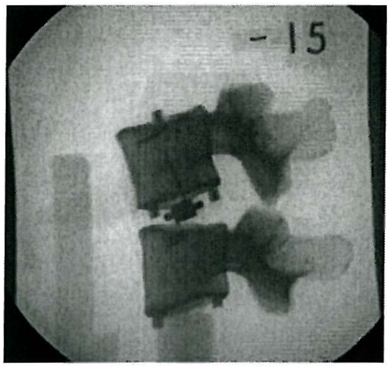
FIGURE 5.13: Extraction results of model images with 16 FDs.



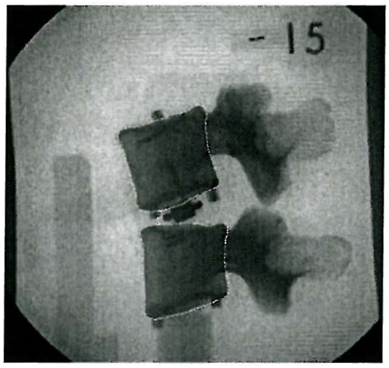
(i) $\theta = -10°$



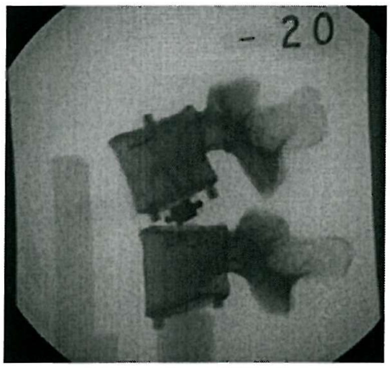
(j) $\theta = -10°$, extraction



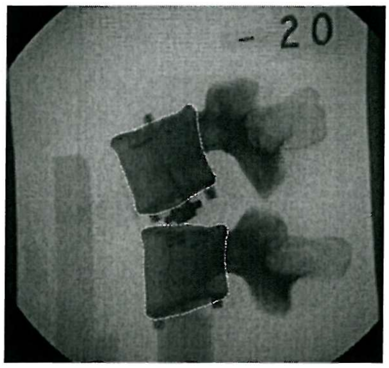
(k) $\theta = -15°$



(l) $\theta = -15°$, extraction



(m) $\theta = -20°$



(n) $\theta = -20°$, extraction

FIGURE 5.13: Extraction results of model images with 16 FDs (continued).

As has introduced earlier, the Hough transform is a model based approach which means the ability to accurately and effectively construct model for the target object is essential in applications. Since the emphasis of this study is to develop a new segmentation algorithm, the template contours are manually constructed by using the edge information within the image. For the calibration study, the template contours of L3 and L4 are obtained by manually connecting the edge points located by applying phase congruency to the frame with the preset rotation set to zero. For the case of a real DVF sequence,

TABLE 5.2: Extraction results with 16 Fourier harmonic components.

L3				L4			
Preset rot. (degree)	x (pixel)	y (pixel)	Rot. (degree)	Preset rot. (degree)	x (pixel)	y (pixel)	Rot. (degree)
10	211	168	11	0	198	289	0
5	207	169	6	0	198	290	0
0	200	168	0	0	198	290	0
-5	195	168	-5	0	199	290	0
-10	188	169	-10	0	198	290	0
-15	182	170	-16	0	198	290	0
-20	176	171	-21	0	198	290	0

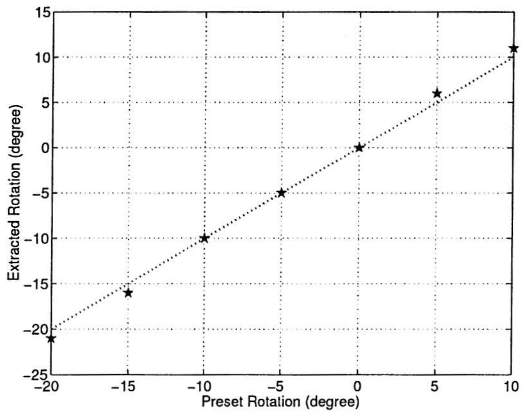


FIGURE 5.14: Comparison between preset rotation value and extracted value for L3.

since the image quality is poor, it becomes more difficult to construct them. Considering that the subjects keep stationary in the first 10 frames within a sequence, the average of these 10 frames is used in order to reduce the effect of noise. Edges are then detected by applying phase congruency to this averaged image. Finally, manual work is used to connect the edges to form the template contours. In this way we managed to minimise the possible variations and errors by using the edge information. This approach to the initialisation of template contours is somewhat time consuming, manual work and subject specific. This should be improved in the future.

Once the template models are ready, the HT was used to locate the vertebrae within images based on the edge information detected by phase congruency. Here only results with 16 FDs are presented. Note that during the extraction the changes of the unknown parameters (x, y and ρ) were limited to a certain small range in order to reduce the computational cost and requirement on storage. As such, the ranges of x and y are 64 pixels and that for ρ is 64°. This means the size of Hough space is $64 * 64 * 64$. The Hough spaces formed here will also be used to test the newly developed algorithm which will be discussed in the following chapter.

The extractions were superimposed on the source images and illustrated by highlighting, shown in the right column of Figure 5.13. Figure 5.14 shows the comparison between the extracted value and the preset rotation for L3 where the stars denote the extracted rotation values. Table 5.2 presents the extraction results for both L3 and L4. The rotation results of L3 are very close to the preset values and there is only a 1° difference in rotation. This accuracy is much better than that established in a similar study where an averaged error of 4.16° was reported (Tezmol et al. 2002). Rotated from right to left, the x coordinate of its center changes from 211 to 176 while the y coordinate slightly changes between 168 and 171 (here x increases from left to right and y increases from top to bottom). It means that the y values do not change much because of the small rotation involved. Note that in this case the interval values for the parameters are 1° for rotation and one pixel for translation. By using a finer interval, more accurate results might be obtained. Furthermore, it is worth mentioning that the PHT has an inherent advantage over a manual method.

The relative error of these extraction results, about the 30° movement range of the vertebra L3, is less than 2% according to the following Equation (5.12).

$$err(\%) = \frac{\sum_{i=1}^{N} |x_i^{'} - x_i|}{N \cdot M_r} \tag{5.12}$$

where N is the total number of experiments and M_r is the range of movement.

5.5.3 Application to DVF Images

5.5.3.1 Vertebra L3 Extraction

As discussed earlier, during acquisition of DVF images of the lumbar spine, the X-rays are usually focused on the area of the vertebra L3 and consequently this area of the image has the best quality. For this reason the PHT was applied to the vertebra L3 as the starting point.

During the extraction only areas containing vertebrae (width 100 and height 100) are analysed rather than the full size of DVF image (about 800*600) in order to reduce the computation cost. This is achieved by cropping them from the original images by the estimated positions made from several frames (i.e. neutral position and extreme flexion/extension positions). Similar to the extraction of calibration model, the ranges of the unknown parameters (x, y and ρ) were also limited. The ranges of x and y are

TABLE 5.3: L3 extraction results with different FDs.

FDs order	Scale	Rotation (degrees)	x translation (pixel)	y translation (pixel)	Peak value
4	1	2	47	40	123
8	1	1	48	39	128
16	1	1	48	40	152
24	1	1	48	40	141
32	1	1	48	40	146

32 pixels and that for ρ is 64°. This means the size of Hough space is $32 * 32 * 64$. The Hough spaces will also be used in the newly developed algorithm.

First, the vertebra L3 image was isolated from the first frame of a sequence, shown in Figure 5.15(k). The binary image of L3, shown in Figure 5.15(l), was derived after application of phase congruency. As introduced earlier, the model shape was initialised by manual work in order to form the chain code, from which the FDs can be derived, shown in Figure 5.15(m).

As there are no size changes of the vertebrae in DVF sequences, size was not considered and so only a 3-D Hough space was used in this study, unless stated. The PHT algorithm was used to determine the position parameters of the L3, that is rotation and translation in the x and y directions. Figure 5.15(n) shows the formed Hough space and Figure 5.15(o) presents the resulting image when 16 FDs were used. To compare the results of different orders of FDs, we experimented with 4, 8, 16, 24 and 32 components and the final results are shown in Table 5.3. From Table 5.3, it can be found that the peak value appears to be largest when 16 FDs are used. Afterwards it appears that there is a small decrease in the peak value. The possible reason for this is that the reconstruction model is more and more close to the chain code model with the increment of Fourier components. It will not represent the real shape if the model does not represent the objects completely.

5.5.3.2 Other Lumbar Spine Segment Extractions

The PHT was applied to other lumbar vertebrae within the same image, in which 16 FDs were used as well.

The first row of Figure 5.15 shows the original vertebra L1 image, its edge image, model template, Hough space and resulting image in turn. The second, fourth and fifth rows correspond to those of vertebrae L2, L4 and L5, respectively. Together with the FDs, the extraction has been used to form a contour which was superimposed onto the source image for accuracy evaluation. From Figure 5.15, it appears that the PHT can yield satisfactory extraction results. As it is always difficult or impossible to have the ground truth in clinical applications, this is such a case as well. Given that the ground truth of the real motion is lacking, it will be valuable to validate it if we can compare the results obtained by the HT and those from other studies on the same subjects. An evaluation of comparison between the manual labelling and the HT will be presented in Chapter 7.

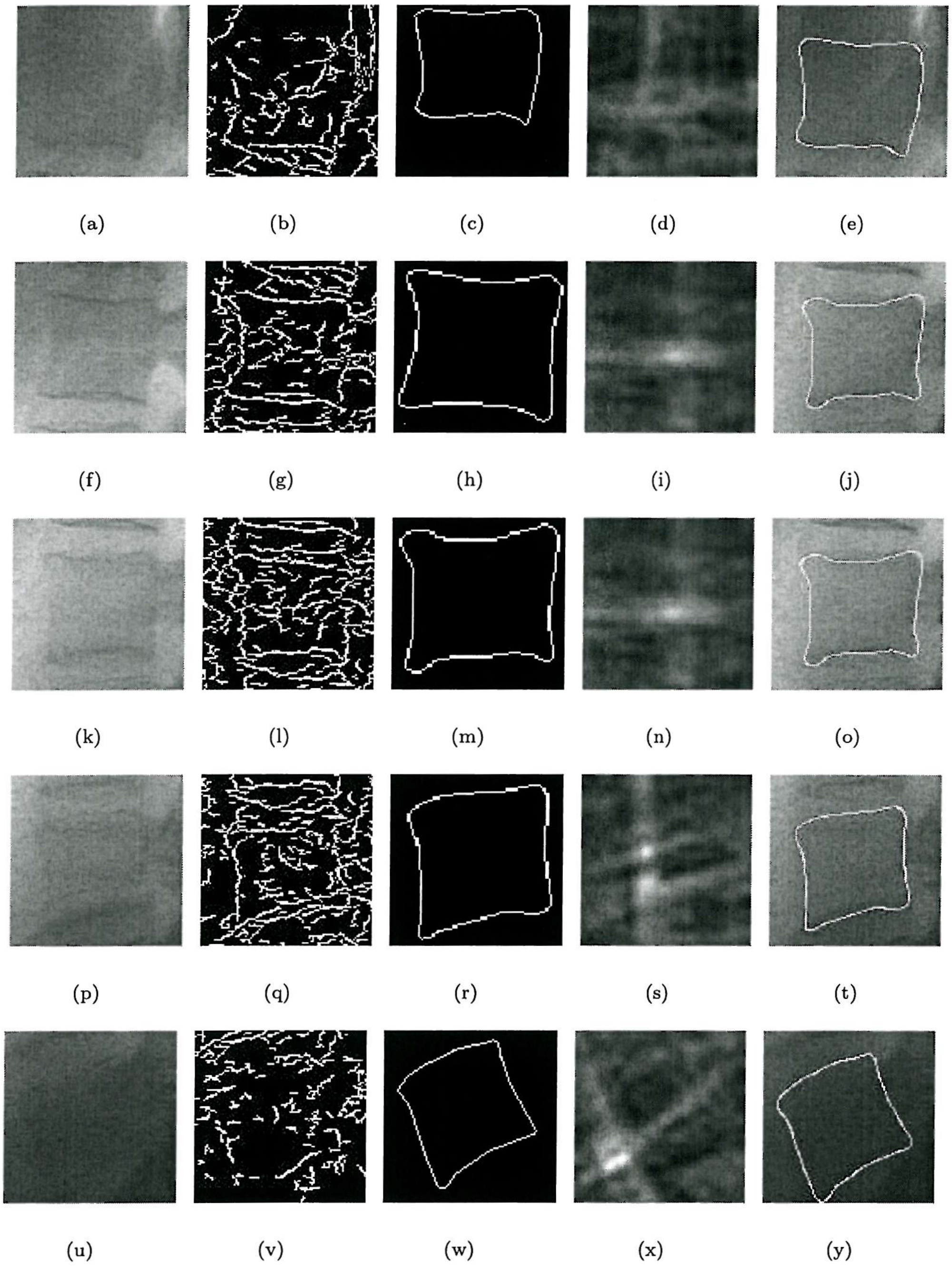
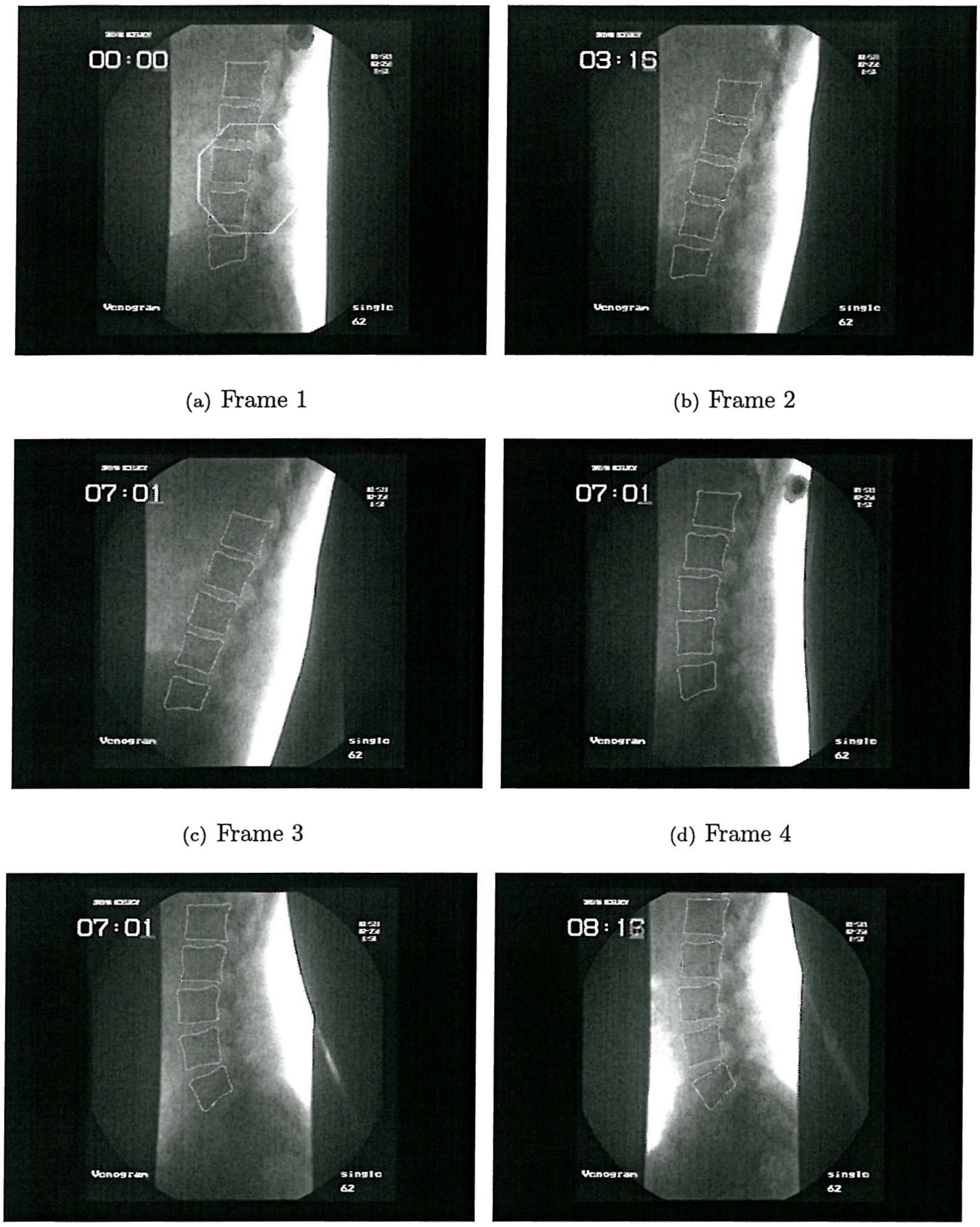


FIGURE 5.15: Extraction results on L1 - L5 of a DVF frame. Column 1: original images; column 2: edge maps; column 3: models; column 4: Hough spaces and column 5: extraction results.

Similarly, the PHT has been applied to other DVF images. The results shown in Figure 5.16 look very promising.



(a) Frame 1 (b) Frame 2

(c) Frame 3 (d) Frame 4

(e) Frame 5 (f) Frame 6

FIGURE 5.16: Extraction results of 6 frames.

5.5.4 Computational Cost

The computational time and memory required by the HT depends on the total number of edge points in the searched image, the size of the known object, the possible ranges of the wanted parameters and the intervals used for quantising these parameters. Currently, the Java program requires about five minutes to extract one DVF frame on a 450MHz PC with 256MB memory. This does not include the time consumed by phase congruency to detect the edge maps. It is necessary to note that the implementation was not optimised for speed. Some improvement would be expected if using C/C++ and optimising the source code with respect to the computational speed. However, it appears unnecessary to analyse the motion sequence in real time for clinical applications. It is worth mentioning that its inherent parallel character enables the potential for real-time implementation if high speed is required.

5.6 Conclusions

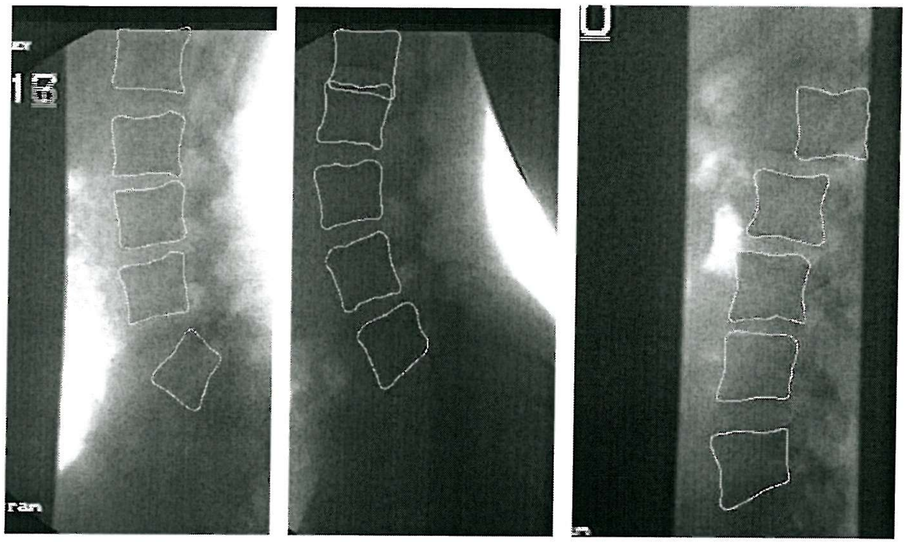
In this chapter, an overview of image segmentation was given first, then the basic ideas of the HT were introduced in detail followed by a short review on the early studies. Elliptic FDs were presented from which the PHT was derived. In order to test performance, the PHT has been applied to a synthetic image, both with and without noise and occlusion, and images of a calibration model and the results look promising. Finally, the PHT was applied to DVF images and the results are still encouraging. The calculation of the PHT reveals that this performance is not achieved without computational cost, though this cost is somewhat mitigated by continuing advance in computer technology.

Chapter 6

Spatio-Temporal Hough Transform

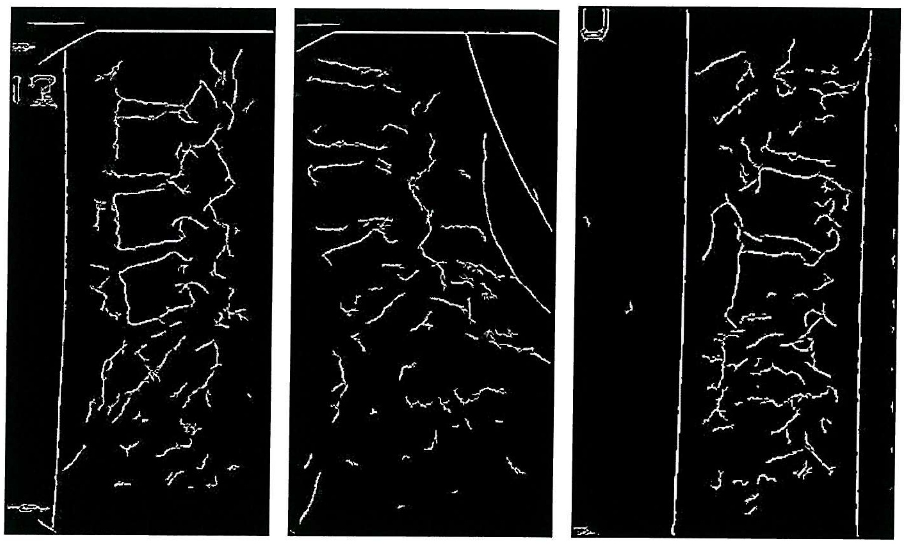
6.1 Introduction

We have shown in the previous chapter that the PHT has good performance in vertebral extraction. It sometimes fails to locate the vertebrae correctly, especially L5 and L1. Three instances are shown in Figure 6.1 where the locations of L1 and the L5 in 6.1(a), the locations of L1 in 6.1(b) and in 6.1(c) were wrongly identified: in 6.1(a) L5 is too low and also incorrectly estimated rotation; in 6.1(b) L1 appears erroneously low; and in 6.1(c) L5 is displaced. This is largely caused by poor edge information and their edge maps are shown in Figure 6.2. From Figure 6.2 it can be seen that the edges of L1 and L5 are always in discontinuous form and less clear than the neighbouring features i.e. boundaries of neighbouring vertebrae or the sharp boundaries around the image. As stated earlier, phase congruency was used to obtain edges owing to its good performance against large illumination changes and low contrast. Therefore, it appears that the problem originates from two factors during DVF image acquisition. One is the radiation dosage which has been set to a low value for the safety of subjects, whilst adversely affecting the image quality. The other is from the appearance of other structures of the human body, such as pelvis, muscles and other soft tissues, that reduced the image quality even more. These are beyond the scope of our study and the only feasible way is to investigate if there are some approaches that can help to overcome these problems. An alternative is to alleviate the problem at a higher level, that is within the Hough transform framework itself.



(a) Problem case 1 (b) Problem case 2 (c) Problem case 3

FIGURE 6.1: Illustration of some cases of incorrect extractions.



(a) Edge maps of problem case 1 (b) Edge maps of problem case 2 (c) Edge maps of problem case 3

FIGURE 6.2: Edge maps corresponding to the cases of incorrect extractions.

6.2 The PHT and Spatio-Temporal Information

Before turning to the new approach, it is worth having a close look at the conventional Hough transform. Consider the PHT as an example, from the discussion in the previous chapter there are two limitations with it in analysing motion sequences. First, for an image containing multiple objects, the PHT has to extract them separately. If these objects have the same shape, one accumulator is needed to contain the votes produced by the HT during the evidence gathering. Objects can then be located by finding maxima in the accumulator which are higher than a predefined threshold. If these objects have different shapes, then the problem becomes complex and accumulators have to be constructed for different types of shapes. In order to find multiple occurrences of the same shape one by one, a complex scheme has to be introduced to interpret the accumulator space. This implies that the HT assumes them to be isolated objects. The possible relationships (referred to as spatial information) such as the position constraints are not taken into account within the implementation of the HT. For lumbar spine extraction, there are spatial constraints on the vertebrae, i.e., these bones cannot enter each other. Second, for object extraction within a motion sequence, such as the problem to locate five lumbar vertebrae in DVF sequences, the PHT algorithm is applied frame by frame. This strategy means that the PHT does not take advantage of the possible relationships between frames (referred to as temporal information) within a motion sequence. These two aspects imply that the conventional PHT only utilises the edge information alone and does not incorporate the spatio-temporal information into itself.

However, in our view, spatio-temporal information will be valuable for improving the performance of the HT to handle motion sequence analysis. Consider multi-object extraction from a sequence as an example, there should be some relationships among these objects such as distance constraints within one image and contextual information between the frames, i.e. the trajectory of the object of interest often appears to be a smooth pattern when the sampling rate is high enough compared with the speed of the moving object. Thus if this information can be used, it will be helpful in handling noise, partial occlusion and false peaks.

Therefore, the new version of the HT should be able to incorporate this information. Hereafter, this hybrid HT will be referred to as the Spatio-Temporal Hough transform (STHT). To some degree, the STHT is designed to extend the HT's ability to handle motion analysis by combining the spatio-temporal information. Thus, it is also expected to be more robust for motion analysis with an improved ability to handle low quality image sequences.

The motion analysis has been conducted in a number of previous studies. Before introducing the STHT, a brief overview of existing methods is presented in the following section.

6.3 Existing Techniques

To locate moving objects in an image sequence, it appears that techniques that analyse the motion sequence as a whole can be more powerful than those based on analysing single static images individually. So far, there has been a substantial literature on the problem of motion analysis in a sequence.

Optical flow has been used to analyse motion in a number of studies. Optical flow is a vector field that reflects the direction and magnitude of the intensity changes from one image to the other due to motion involved. The time interval between frames must be very short in order to guarantee small inter-frame changes (Sonka et al. 1999). Optical flow is based on two basic assumptions. That is, the observed brightness of any object point is constant over time and nearby points in the image should move in a similar manner. Optical flow will be subject to error if these two assumptions are violated. Unfortunately, this is quite common in application. Error propagation is often observed in application. Furthermore, optical flow is not the exact motion field but an approximation.

Feature point correspondence is also used in motion analysis. Unlike optical flow, it can work even when the time interval between frames is not small. One application of the correspondence is to solve the problem of many objects moving simultaneously and independently (Seith and Jain 1987; Rangarajan and Shah 1991; Chetverikov and Verestóy 1999). This is based on the notion of path coherence which assumes that the motion of an object at any point in an image sequence will not change abruptly. How to design a good path coherence function is yet to be solved. Furthermore, this approach suffers from occlusion and missing data problems. To overcome these problems, additional local trajectory constraints must be considered and incomplete trajectory must be allowed.

The Kalman filter is a tracking method widely used in motion analysis. It is an efficient computational (recursive) solution of the least-squares method. It can give estimates of past, present and future states of a system even when the underlying model is imprecise or unknown. Its basic form requires that the system be linear with its observations to be a linear function of the underlying state while noise involved is assumed to be white and Gaussian. In application these assumptions often appear unrealistic. Moreover, if the parameters of the Kalman filter are not initialised properly, the filter can diverge and thus fail in tracking. This approach estimates the object's motion by using earlier motions which means errors produced in the earlier frames can be propagated during the analysis.

As mentioned earlier, the VHT is a variation of the HT that has been developed to conduct motion analysis (Nash 1999). The original form of the VHT was used to extract circular shapes with linear velocity. In implementation, apart from the shape parameters, velocities in the x and y directions are also incorporated into the accumulator space. Then the VHT collects evidence from all feature points in the whole sequence into a single accumulator. This differs to other studies where accumulators are used for each frame when the HT is applied to individual frames one after the other. This is the reason why the VHT appears to be more robust than a standard frame-by-frame tracking implementation, especially when the target is occluded or noisy as any missing or damaged structural information in one frame can be compensated for by redundant data in others. Due to the global search nature of the VHT, it is unnecessary to initialise the algorithm to search in a specific area although limiting the extent of the search can often reduce the computational cost. The VHT has also been adapted to be capable of coping with other motions such as pulsating artery. The limitation of the VHT is that it can only handle motions that can be described parametrically.

An improvement has been made in the form of the continuous VHT (CVHT) which can handle arbitrary shaped objects moving in arbitrary patterns of motion (Grant, Nixon, and Lewis 2002). In this method, a motion template was designed and also represented by FDs like the shape model. By evidence gathering, the moving object can be located by finding the possible variations of the motion as well as position parameters. The CVHT requires a motion template which means some *a priori* knowledge of the motion has to be known in advance. As discussed before, the computation involved can be prohibitive when more parameters are used.

Recently, another HT version has been proposed for motion analysis (Lappas et al. 2002). In this method, the standard HT is applied to individual frames of one sequence and accumulator spaces corresponding to these frames are obtained. Then, an energy function is designed in which the Hough space term and a smoothness term are included. The Hough space term is the sum of peak values in the accumulator space where the considered trajectory passes. The smoothness term represents the assumption that motions are smooth and is obtained by evaluating changes both in direction and in speed. By locating the maxima of the energy function, the extraction of a sequence is achieved. There are some limitations with this method, i.e. dynamic programming is used to search the optimal values but dynamic programming can converge very slowly.

From the brief review of related works in motion analysis, it is clear that these three versions of the HT only consider a single moving object and cannot be used for multiple moving objects. Feature point correspondence may be possible, but it requires that points in different frames should be matched. It seems this cannot be true in reality as occlusion and missing data often occur.

Hence, a new motion analysis technique is desirable and the STHT is one of the first approaches to fully use the spatio-temporal information within the sequence. The STHT is based on the HT and thus it inherits its good performance in noisy environments and ability to handle occlusion.

To re-iterate, the new STHT which performs a global analysis of a sequence may offer certain performance benefits, over analysis on frame by frame. When extracting objects from a very noisy image sequence, or instances where the object of interest is partially occluded by another object in the scene, a global analysis may offer improved performance by the nearby information in neighbouring frames. Temporal information may be valuable in handling noise and occlusion. For multiple objects, spatial information can also be very useful, i.e. the relationship between two objects can be utilised to eliminate the false peak location in the HT.

6.4 Description of the Algorithm

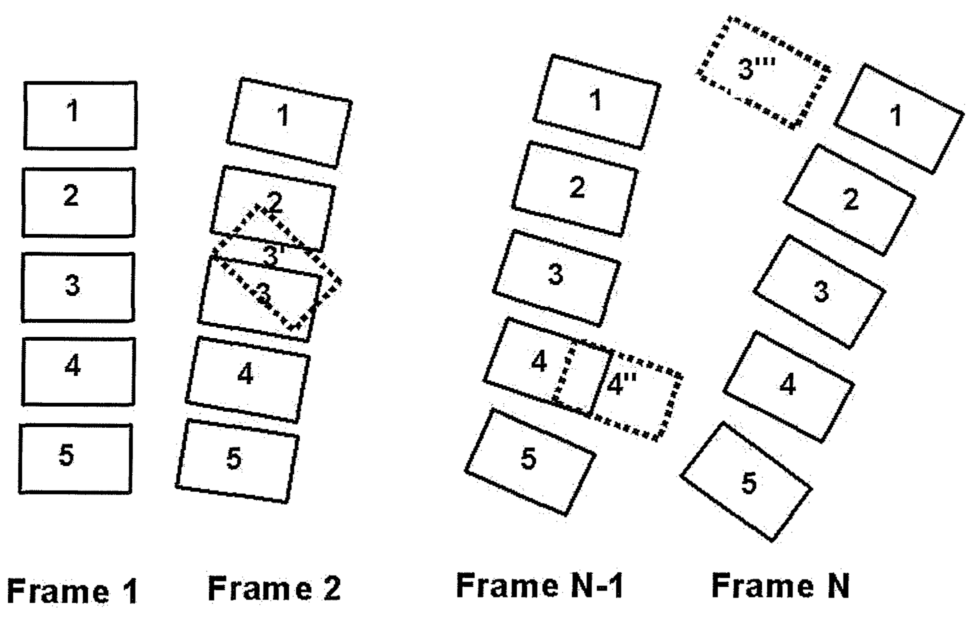


FIGURE 6.3: Illustration of the basic idea of the STHT.

6.4.1 Basic Ideas

Figure 6.3 shows the basic idea of the STHT, given a motion sequence consisting of N frames with 5 moving objects which are denoted by solid blocks. The blocks with dotted lines denote incorrect extractions with either wrong relative positions or abnormal changes in the motion. The STHT should be able to minimise these possibilities by using spatio-temporal information. Although the STHT designed here is specific for the extraction of lumbar vertebrae, it can easily be extended to other applications, i.e. kinematic study of knee or finger joints and other kinds of motion analysis. In these cases, the above conditions can be generalised or supplemented to adapt to the specified problems in order to obtain more robust results. This should be borne in mind during the development of the algorithm.

A major concern with the STHT is how to use the spatio-temporal information. This included how to translate the spatio-temporal information into parameters that can mathematically be measured, and how to combine this with the HT. In order to do this, some assumptions have to be made and these will be discussed in further detail.

6.4.2 Assumptions

In the development of the STHT, the following assumptions were made to simplify the problem.

- **Temporal Issues**
  Each object should be a rigid body and undergo an arbitrary but smooth planar motion. Assumption of rigidity can guarantee that there is no deformation in target shapes during the motion sequence. The assumption of smooth motion enables us to use the motion information in the STHT. Provided that the sampling rate is high enough while the speed of the object's motion itself is relatively slow, this assumption on motion can be reasonable. In spine motion studies these assumptions should hold. In particular, abrupt motions are unlikely to occur because in the current study only young healthy normal subjects are used. It is worth mentioning that in the current study the sampling intervals between frames are constant. As a variation of the HT, the STHT potentially has the ability to handle 3-D motion, but this is conditional upon the availability of 3-D information (i.e. multi-view imagery, but this needs synchronisation). Here only planar motion was considered. As discussed earlier, this should be reasonable because the out-of-plane motion can be negligible in the sagittal plane motions.

- **Spacial Issues**
  In each frame, there are some spatial constraints between objects which can be used. In particular for the lumbar spine, the positions of vertebrae are constrained by muscles and ligaments. The vertebral bodies should not enter each other or at least the distance should be maintain because of the existence of intervertebral discs between them. Obviously the earlier extraction of L1 in Figure 6.1(b) contradicts this and the extraction is impossible from a physiological viewpoint.

6.4.3 Energy Function

In the STHT, the spatio-temporal information and the HT are embodied in an energy function by using the following criteria: the motion must be smooth; the trajectory must pass through the points in the accumulator space with as large peaks as possible; and the spatial relationships between objects should be known in some form. Taking all these into account, we consider three terms in the energy function. Therefore, for any searched object, the energy function can be expressed as

$$E = w_1 E_{Hough} - w_2 E_{temporal} - w_3 E_{spatial} \tag{6.1}$$

where w_1, w_2 and w_3 are weighting factors that may be adjusted to vary the relative contribution of each term.

The first term is the so-called Hough term. It contributes the sum of peak values of the points in the accumulator spaces, which pass through the trajectory. It can be expressed as,

$$E_{Hough} = \sum_{i=1}^{N} H(x_i, y_i, \rho_i) \tag{6.2}$$

where N is the total number of frames in the sequence. This term is used to force the sum of peaks of possible solution to be as large as possible. $H(x_i, y_i, \rho_i)$ is the Hough space corresponding to each frame.

The second term, $E_{temporal}$, expresses the smoothness in the trajectory over the whole sequence and it will tend to prefer solutions that have a smooth trajectory. The third term, $E_{spatial}$, represents the constraints of spatial information. It will favour the expected position changes between objects. In other words, it is used to penalise those positions that violate the spatial constraints between neighboring objects. These two terms will be discussed in further detail in the following sections.

The latter two terms typically reflect the *a priori* knowledge. Equation (6.1) is, in fact, a compromise between the PHT and the underlying spatio-temporal information in the sequence. With the aid of *a priori* knowledge, the STHT is expected to have a better performance against occlusion, noise and false peak locations in the Hough spaces. Setting $w_2 = 0$ and $w_3 = 0$ allows the energy function to search the maxima among all the Hough spaces involved and the chosen solution will be the same as that of the standard PHT.

6.4.4 Temporal Term

As discussed earlier, temporal information can be very useful in improving the performance of the HT by collecting global evidence from the sequence. In implementation, however, how to describe this information and how to incorporate it into the energy function needs careful consideration.

In Lappas et al.'s approach (2002), the smooth trajectory is represented by both direction and velocity constraints. This will favour small velocity and direction changes and can only be effective for an object that moves slowly relative to the frame rate. The term that they used is a simplification of the energy function proposed by Sethi and Jain (1987) for handling point correspondence problems.

In the current study, a spline is used to describe the smooth trajectory. A spline is a smooth piecewise polynomial function. It originates from the flexible strips used to create smooth curves in traditional drafting applications. Splines are widely used in computer graphics and image processing to obtain smooth curves (or surface) owing to their good properties. When used for interpolation, they do not have the oscillatory behavior that is a characteristic of high-degree polynomial interpolation such as Lagrange or Hermite interpolation. Also, mathematically, it is the smoothest curve that passes through a set of fixed points (de Boor 1978). It appears that a spline can be a good candidate to describe the smoothness of the trajectory.

Given observations of an independent variable, y at design points x (these points are referred to as nodes), the most common piecewise polynomial approximation using cubic polynomials between each successive pair of nodes is called cubic spline interpolation. A general cubic polynomial involves four constants so there is sufficient flexibility in the cubic spline to ensure not only that the interpolation is continuously differentiable on the interval, but also that it has a continuous second derivative on the interval.

In order to guarantee continuous first and second derivatives, conditions at these nodes are introduced. With these conditions, the coefficients that are used to define those cubic polynomials can be determined uniquely. This means that the cubic spline S that passes through all these nodes is known. Normally, the points x should not have the same values in order to avoid intervals with zero values which make it impossible to determine a spline. Details of this procedure are given in Appendix A (Burden and Faires (1993)). Provided that the spline S is known, for any given x_i within the range $[x_{min}, x_{max}]$, a unique value of y_i can be determined, so do the first and second derivatives which can be expressed as $S'(x_i)$ and $S''(x_i)$, respectively. These are used to quantify the smooth trajectory.

The second problem is how to form the smoothness function. Motivated by the active contour, in which the first and second derivatives are used to represent the continuousness and smoothness of a curve, here the smoothness of a trajectory is represented in a similar way. Each vertebra involves a translation in the x and y directions and a rotation change ρ, requiring three cubic splines, S_x, S_y and S_ρ, to describe them. Thus, the $E_{temporal}$ term in the Equation 6.1 is defined as follows:

$$\begin{aligned} E_{temporal} =& \alpha \int_{t_1}^{t_N} \left(\zeta_x (S_x')^2 + (1-\zeta_x)(S_x'')^2\right) dt \\ &+ \beta \int_{t_1}^{t_N} \left(\zeta_y (S_y')^2 + (1-\zeta_y)(S_y'')^2\right) dt + \gamma \int_{t_1}^{t_N} \left(\zeta_\rho (S_\rho')^2 + (1-\zeta_\rho)(S_\rho'')^2\right) dt. \end{aligned} \tag{6.3}$$

Here, α, β and γ are factors that may be adjusted to vary the relative contributions of these terms. Similarly, ζ_x, ζ_y and ζ_ρ are used to control the relative contributions of the first and second derivatives. In implementation the integrals are approximated by summing them up at sampling points as

$$\begin{aligned} E_{temporal} =& \alpha \sum_{t_1}^{t_N} \left(\zeta_x (S_x')^2 + (1-\zeta_x)(S_x'')^2\right) \\ &+ \beta \sum_{t_1}^{t_N} \left(\zeta_y (S_y')^2 + (1-\zeta_y)(S_y'')^2\right) + \gamma \sum_{t_1}^{t_N} \left(\zeta_\rho (S_\rho')^2 + (1-\zeta_\rho)(S_\rho'')^2\right) \end{aligned} \tag{6.4}$$

6.4.5 Spatial Term

Spatial information is another important term in forming the energy function of the STHT. For the moving lumbar spine, this kind of information is plentiful, i.e. physical constraints, the possible smooth connection curve through five vertebrae etc. However, how to describe them mathematically and incorporate into the energy function is not easy. For instance, if using the smoothness of possible connection curve, it implies that five vertebrae have to be investigated simultaneously and this can lead to a problem of dimensionality.

In simplifying the problem, only the distance between centers of neighboring vertebrae is used. Specifically, the distance d between centers of two neighboring vertebrae should be within a predefined range $[D_{min}, D_{max}]$. If d is beyond this range, a penalty with large value L will be given. That is,

$$E_{spatial} = \begin{cases} 0 & D_{min} \leq d \leq D_{max} \\ L & d < D_{min} \text{ or } D_{max} < d. \end{cases} \tag{6.5}$$

Equation (6.5) implies that the spatial information serves as a means to penalise those positions that violate the physical constraints. This is possible due to the flexibility of the Genetic algorithms (GAs) that are used to search for the maxima of the energy function.

6.4.6 Energy Maximisation

A solution to the STHT is defined to be the global maximum of Equation (6.1). That is,

$$E_{opt} = \max(E). \tag{6.6}$$

From Equation (6.1), it can be seen that maximising this energy function can lead to high dimensionality. For example, if the Hough space for each frame is 3-D (translation in x and y directions and rotation ρ) and the possible size of each dimension is 10, if N (the total number of frames contained in a sequence) is 10, the STHT has to explore a space as large as 10^{3N}. In application, this inevitably results in an increasing requirement on memory and computation. Hence, it appears that the exhaustive search is not applicable in terms of complexity. GAs are used for the optimisation due to known good ability in search of multimodal accumulator spaces.

In short, the implementation can be described as follows. First, the Hough spaces of each lumbar vertebra in all frames of a sequence are obtained. Then the STHT is applied to these vertebrae one after the other. That is, for each vertebra, the solution is obtained by using a GA to maximise the energy function defined in Equation (6.1). To locate five vertebrae is a multi-step procedure. Due to the relatively good edge information in the area of L3, L3 is extracted first by considering only the first two terms. Then the other vertebrae, L2, L4, L1 and L5, can be extracted in turn and in that order. In these extractions, the results of each prior extraction are used to form spatial constraints. For example, the results of L3 are used for L2 extraction as well as L4. Then results of L4 are used for L5 extraction and so are those of L2 for L1. This appears reasonable as the extraction progresses from the area of best quality to areas of poor quality.

From the implementation, it can be seen that the STHT extracts moving objects one by one using the global information in the sequence. On the other hand, the STHT can also be adapted to extract all objects of interest by using a single energy function that can evaluate them together, this can be illustrated by Figure 6.4. The energy function can be expressed as

$$E = \sum_{j=1}^{M} w_1^j E_{Hough}^j - w_2^j E_{temporal}^j - w_3^j E_{spatial}^j \tag{6.7}$$

where M is the number of total objects. E_{Hough}^j, $E_{temporal}$ and $E_{spatial}$ can be defined as in Equation (6.1). However, with the increase of objects, more weighting factors are required to express the relationships between these objects. How to determine appropriate values for these weighting factors can be difficult because of the possible complex relationships amongst these objects. Furthermore, the dimensionality problem becomes even more pressing. This may be investigated in future work.

6.5 Genetic Algorithms

Genetic algorithms (GAs) are known for their good performance in the optimisation of high dimensionality problems. There has not, so far, been any formal definition which is often problem-oriented. GAs are based on an analogy to natural behaviour and are adaptive methods used to solve search and optimisation problems. GAs can solve the specified problem by mimicking the principles of natural selection and "survival of the fittest" (Beasley et al. 1993a; Beasley et al. 1993b). In implementation a population of individuals is introduced. Each individual represents a possible solution to a given problem, and it is also assigned a fitness value according to the extent of how good this solution is for the problem. The fitter individuals will have a higher probability of survival and of reproducing new individuals by which the most promising areas of the search space are explored. If a GA is designed well for a specified problem, it will converge to an optimal solution for the problem. Apart from GAs, there are other algorithms

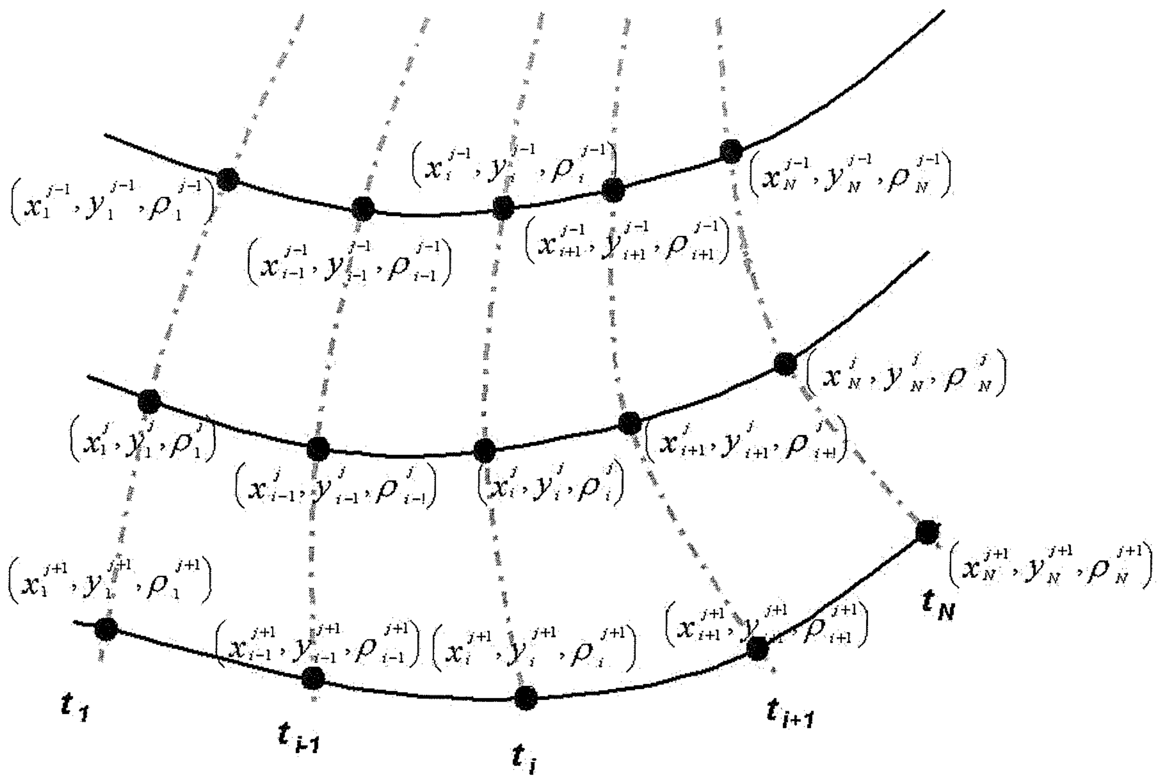


FIGURE 6.4: Illustration of the STHT.

based on analogy with nature, i.e. Neural Network and Simulated Annealing. Neural Networks are based on the behavior of neurons in the brain. They are often used for classification tasks and their areas of application partly overlap those of GAs. Simulated Annealing is another popular search technique that is based upon the analogy with the cooling of a solid.

The popularity of GAs stems from their robust and simple implementation and that they can successfully deal with a wide range of problems. As suggested, although GAs are not guaranteed to find the global optimum solution to a given problem, they are generally good at finding "acceptably good" solutions to problems "acceptably quickly" (Beasley et al. 1993a).

There are numerous books on GAs. Readers are invited to refer to the book by Goldberg (1997). Many studies have been conducted in order to improve its performance such as efficiency and convergence. For instance, cooperative coevolution was proposed to favor parallel computation and more data sets (Potter 1997).

6.5.1 Existing Applications

GAs are powerful searching tools and have various applications in pattern recognition and machine learning. In particular, GAs are used as a means of overcoming the computational cost experienced in the conventional HT.

Goulermas and Liatsis (1998) proposed a hybrid GA/HT system for fine-tuning the accumulator space of the HT for circular shape detection. In this method, the HT is embedded into the GA which is specially designed to obtain the accurate detection against the case of excessive noise.

Ser et al. (1999) also employed a GA in their implementation of the GHT. Here, a GA is used to speed up the maximum search in the parameter space. This method can avoid enormous storage for the Hough parameter space even in the case that fine resolutions are used for the parameters and thus higher accuracy can be obtained without the problem of storage limitation. This method can be applicable to the detection of both analytic and nonanalytic objects. Apart from the weakness of the GHT, its main limitation is the large time consumed to obtain accurate results.

6.5.2 Fundamental Principles

The process involves four major steps (Ser et al. 1999). These are, namely, initialisation, reproduction, crossover and mutation. During initialisation, population is initialised by values randomly selected from the space to be searched. Therefore, each individual is assigned a fitness value. After initialisation, iterative reproduction is started and continued until the required convergence which is, sometimes, predefined generation number or an error limit. After each reproduction, a new generation is produced and called a "child" while the previous generation is called a "parent". During each reproduction, the natural phenomenon of "survival of the fittest" is introduced. However, it is not the simple reproductions of the best chromosomes. To maintain the number of individuals, new individuals are also introduced by using some operators, i.e. crossover and mutation.

This section continues by discussing each of the genetic operators coding, fitness function, reproduction, initialisation and convergence in further detail.

6.5.2.1 Coding

In the implementation of a GA for a particular problem it is assumed that the potential solution to the problem can be represented as a set of parameters (known as *genes*) i.e. in our problem, the possible positions of the vertebra x_i, y_i and ρ_i for $1 \leq i \leq N$. These parameters are joined together to form a string which is often referred to as *chromosome*. In the forming of the string binary encoding is often used (other methods are also applicable.). Given the size of the Hough spaces described in section 5.5.3, our problem is to maximise Equation (6.1) with variables x_i, y_i and ρ_i, for $1 \leq i \leq N$, which are represented by 6, 6 and 7-bit binary numbers, respectively (one bit digit is used to denote that the value is positive or negative). Therefore each chromosome will have $3N$ genes arranged as $x_1 y_1 \rho_1 x_2 y_2 \rho_2 ... x_N y_N \rho_N$ and consist of $19 * N$ $(N * (6 + 6 + 7))$ binary digits.

6.5.2.2 Fitness Function

Once the coding has finished, a single function value can be derived, which depends on the possible values assigned to those variables. In GAs this is called fitness function and it is designed to represent how good a chromosome is during evolution. This fitness value is of great importance for the reproduction in the GAs. In function optimisation, the fitness function will just measure the value of the function.

6.5.2.3 Reproduction

After calculating the fitness function of all individuals in a specified generation, the reproduction phase is involved for designing a new generation according to the performance of the individuals within the current generation. As has been introduced, the selection protocol is "survival of the fittest". That is, individuals with higher fitness values will be more likely to be selected as parents to produce offspring while the least fit individuals may not be selected and will die out. Having selected two parents, offspring are generated by recombining their chromosomes in which crossover and mutation mechanisms are typically used. These operations are outlined as follows.

Selection The manner of selecting parents has important effects on producing the new generation. The behavior of the GA largely depends on how individuals are chosen. Without exception, there are a number of approaches that have been proposed, i.e. proportionate selection, fitness ranking, steady state selection and tournament selection. Consider the tournament selection as an example, pairs of individuals are randomly selected from the current generation and the individual with the higher fitness value is selected as one of parent. Another parent can be selected in such a way as well. In practice large tournaments may also be used in which the best one among n randomly selected individuals is used as one of the parent.

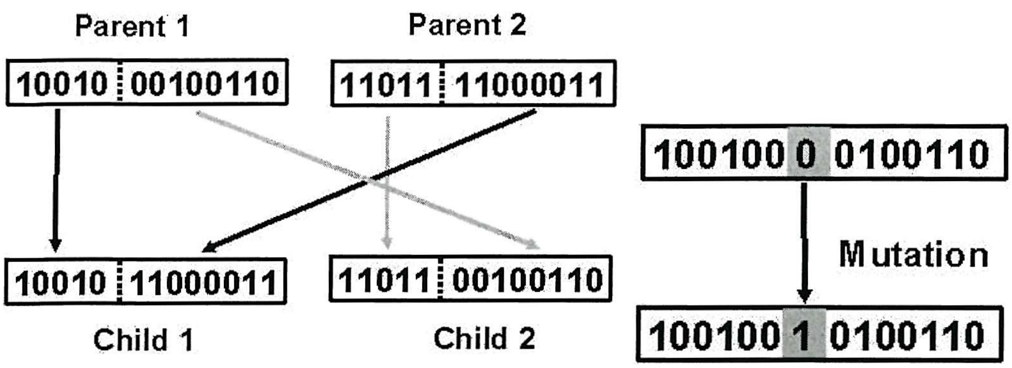
The aforementioned selection strategies (apart from proportionate selection), however, can yield similar performances if suitable adjustment of parameter is made, such that there is no absolute "best" method (Goldberg and Deb 1991). In our application, tournament selection is used.

Crossover Crossover is the most vital operator in GA as it mostly determines how the offspring are generated, given specified parents. There are three most commonly used schemes, namely, 1-point crossover, 2-point crossover and uniform crossover. Of course other multiple point crossover may also be used. The 2-point and uniform crossovers are just the special cases of multiple point crossover. The basic idea of crossover is illustrated by considering 1-point crossover. Given two parents, cut them at a certain position (often randomly selected) and form two "heads" and two "tails" segments. Heads and tails from different individuals are connected to form two new individuals, shown in Figure 6.5(a). In this way, two offspring are expected to inherit some genes from both parents and after evolution over many generations the solution can be found. 2-point and uniform crossover are different in which 2 randomly selected positions (or any binary for uniform crossover) are used to cut the chromosome into pieces for generating new offspring. Crossover is believed to maintain the goodness between generations. From this aspect, a 1-point crossover seems better as multiple crossover may cause the disruptive change on the chromosome and impair the performance of the GA. However, multiple crossover has the advantage of a more thorough search of the problem space. Again there is no conclusion on which scheme performs better, and it appears that to some degree it will depend on the specified problem.

In application, crossover is not applied to all pairs of individuals selected for mating. This means that some of them are just copied to produce the offspring rather than by crossover. This is often achieved by setting a threshold $p_{crossover}$ (typically between 0.6 and 1), and only those that have a higher probability than the threshold are selected for crossover. This is believed to give each individual a chance of passing its genes without the disruption caused by crossover.

Mutation Mutation is applied to each child individually after crossover. It can produce new individuals by reversing the binary bit selected at random in the string and each bit of a chromosome has a probability to be reversed, as shown in Figure 6.5(b).

Generally, it is believed that crossover plays a more important role than mutation for rapidly exploring a search space. However, crossover alone cannot work well because it only functions to find the better fitness values in neighboring areas of the parents and sometimes gets trapped by local maxima. In contrast, mutation provides a small amount



(a) crossover (b) mutation

FIGURE 6.5: Illustration of crossover and mutation.

of random search and can maintain the exploration capability of GAs and ensure that each point in the search space has the same probability of being examined (Beasley et al. 1993a). However, the probability of mutation $p_{mutaion}$ should be smaller than that of crossover $p_{crossover}$, otherwise, the performance of the GAs will be reduced by disruptive genes. It becomes random search when $p_{mutation} = 1$. In most applications, the value of $p_{mutation}$ is set to the reciprocal of the (individual) chromosome length.

6.5.2.4 Initialisation and Convergence

Initialisation is used to create the starting population. In application, this is often achieved by assigning random values to the starting population. Nevertheless, as always, if some *a priori* knowledge about the solution is available, it will be helpful to use this information to make the initial population as fit as possible. In some sense, the better the initialisation, the easier and quicker the search for the optima (Sonka et al. 1999).

If the GA has been correctly implemented, the population will evolve over successive generations so that the best fitness and the average fitness of the population will gradually converge to the global optimisation (Beasley et al. 1993a). In application, different criteria are used as the convergence condition to terminate the evolution. The simplest one is to terminate when an initially specified generation has been finished.

In summary, GAs can work well for most of optimisation problems, however, some special skills are needed to handle different problems in order to obtain good performance. There are many factors that can affect performance, i.e. how to select parents, how to select an appropriate crossover scheme, how to select probability values for crossover and mutation, whether or not using elitism (the fittest individual is directly inherited between generations), etc. Normally, careful tuning is needed to achieve a good performance and this often requires a considerable amount of insight into the nature of the problem to be solved.

6.5.3 Comparison with Other Approaches

There are a number of approaches in connection with search and optimisation problems.

Random (or an enumerated) search is the simplest method. It is an unintelligent strategy which looks for the optima within the whole search space. It can definitely find global optima when the search space is not extremely large. However, in most applications time is a crucial factor.

Gradient methods are generally referred to as hillclimbing and rely on gradient information of the fitness function which directs its search. If the function is discontinuous, it will fail as the gradient information is unavailable. Moreover, it only performs well on functions with a single peak (known as unimodal functions). It will be trapped by the local optimum of multimodal functions. Iterated hillclimbing is formed by combining random search and hillclimbing together in order to improve the performance. Once a peak is located, another hillclimbing search is started from a randomly selected point again. It will have better performance only if the fitness function does not have too many local maxima.

Dynamic programming is a well-known method in optimisation problem and has been widely used in image processing. For example, it was used in a recent work by Lappas et al. (2002). It is only applicable for solving multi-step optimisation problems where the overall fitness function is the sum of the fitness functions for each stage of the problem. Moreover, there should be no interactions between stages while this is often not the case in many applications.

As has been stated, simulated annealing works by mathematically simulating the physical process of annealing which is used to strengthen a material by first heating the matter and then slowly cooling it to reach a minimal energy state. It has been used successfully and continues to be an active research area. Its limitation is that it only deals with one candidate solution at a time and does not build up an overall picture of the search space. Therefore a previous step is not used to guide new moves while GAs can achieve this by inheriting the best fitness between generations.

It is unwise to assert which method is best as these methods are often problem-oriented. In many applications, however, GAs can perform very well and even better if it is combined with other approaches such as incorporating local search. Moreover, it is easy to construct and is adaptable for many different problems. All these have made GAs popular and this is also the reason why they are being used in this work.

6.6 Extraction Results

Before the STHT implementation, the performance of the developed GA is evaluated by analysing some well known optimisation functions. These functions are typically used as benchmark functions in evaluating optimisation algorithms and can be found in other studies (i.e. De Jong (1975)). During the experiments, unless otherwise stated, fitness values are obtained by averaging the results of 10 runs with different initialisations. These experiments have shown that the designed GA performs well in terms of efficiency and convergence. Both uniform and 2-point crossover were tested. It appears that the final results are very close but the GA using 2-point crossover converges slightly faster than using uniform crossover, especially at the beginning of the evolution. In the following STHT performance experiments, tournament selection is used to select parents and two-point crossover operator is used. p_c is set to 0.95 and p_m is set to the reciprocal of the length of an individual which depends on the length of each parameter, the total number of frames provided that the number of parameters corresponding to each frame is three.

6.6.1 Parameter Settings

As has been discussed, the core of the STHT is based on the compromise between the Hough space and the spatio-temporal information. In the current study we have not attempted to find the optimal weighting factors, i.e. weighting factors w_1 and w_2 are set to 0.8 and 0.2, respectively. Currently, the spatial term is used as a penalty and w_3 is set to 1 here. The α, β and γ are set equally to 1/3 while the ζ_x, ζ_y and ζ_ρ are set to 0.4. In practice, these weights can be adjusted according to the specified problems. For example, if the images are of good quality, then the weighting factor w_2 can be set to a relatively small value which means that the Hough spaces dominate the energy function. However, tuning these weights will possibly be part of future work rather than the framework proposed here. The effects of spatial information and temporal information were investigated in the following experiments.

6.6.2 Test on Calibration Model

The DVF images of the calibration model used in chapter 5 were used to provide a measure of validation of the STHT performance. The STHT is applied to extract L3 in the DVF images of this model. A synthetic dynamic motion sequence was formed by arranging seven DVF images in the order that the values of the preset rotation change from $-20°$ to $10°$. The estimated positions of L4 by the PHT were used to describe the spatial term. That is, a penalty will be introduced if the possible center position of L3 is too close or too far from that of L4.

During the evolution 100 individuals were used and each of them contains $3 * N$ parameters where N is the total number of frames in the sequence and here is equal to 7. As discussed in previous chapter, during the PHT implementation the possible range of x, y and ρ were set to small values and the size of the Hough space is 64*64*64. Thus 7 binary digits were used to represent each parameter in the GA and the total length of each individual is $(7 + 7 + 7) * N$. The results obtained from the PHT were used to initialise the GA in order to locate the optima more quickly. The GA was designed to terminate after 2000 generations.

The initialisation with the results of the PHT means that the Hough term is already global optima at the beginning of optimisation. During the optimisation, the value of the temporal term $E_{temporal}$ will gradually decrease. Meanwhile the value of Hough term E_{Hough} could become smaller. The combination of this compromise leads to a maximisation of the total energy function E that corresponds to the target results.

During the experiments on the calibration model, two different cases are considered to test the performance of the STHT with special consideration of accuracy of motion estimates. In the first case w_3 is set to 0 while w_2 is 0.2. This means the spatial term was excluded from the energy function. In the second case the full energy function of the STHT was considered by setting w_2 to 0.2 and w_3 to 1. Note that here the effect of excluding temporal information was not investigated. When temporal information is excluded, the STHT will have no effect on the results from the PHT because of the manner of initialisation: these results have already met the requirement of spatial term $E_{spatial}$ and there is no further optimisation that can be achieved by the energy function.

With these two values for w_3, the results from 10 runs are averaged and calculated to the nearest 0.1°. Together with the results from an early study using template matching (Simonis 1994) and those from the PHT, these results are presented in Table 6.1. It can be seen that the PHT and two cases of the STHT can obtain better results than the template matching method. According to Equation (5.12), their errors are within 2% while the error of Simonis' approach (1994) is about 3.7%. The results are also shown in Figure 6.7. From Figure 6.7 it can be observed that deviations of the results by using template matching sometimes appear large while the rotation angles computed by the PHT and the STHT are much close to the preset values. It can also be seen that two cases of the STHT have a slight improvement over the PHT. The reason for this may come from the fact that the PHT results are already very good and only a fine tuning can be obtained by the STHT. To some extent, this can be illustrated by the Hough space as shown in Figure 6.6. A significant peak is often observed as the quality of the

DVF images of the model is somewhat better than that of the images obtained *in vivo* from the lumbar spine. Although the differences among the results sometimes are larger than 1°, the resulting images of four methods are very similar and it seems that the naked eye is unlikely to distinguish such small differences among them, the resulting images with the preset value $-20°$ are presented in Figure 6.8.

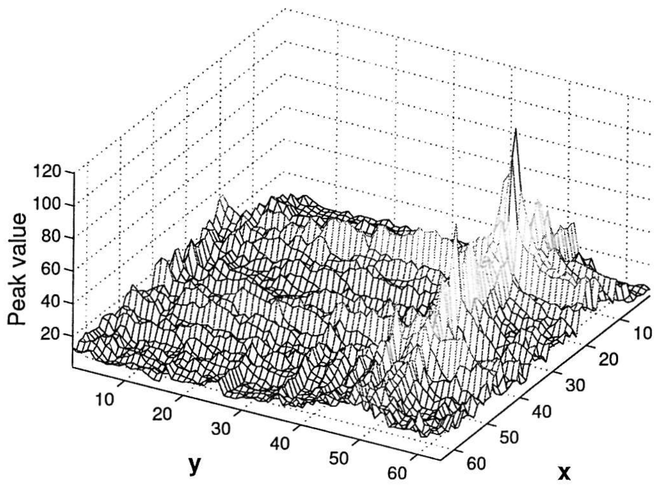


FIGURE 6.6: Hough space of the image ($\rho = 0$).

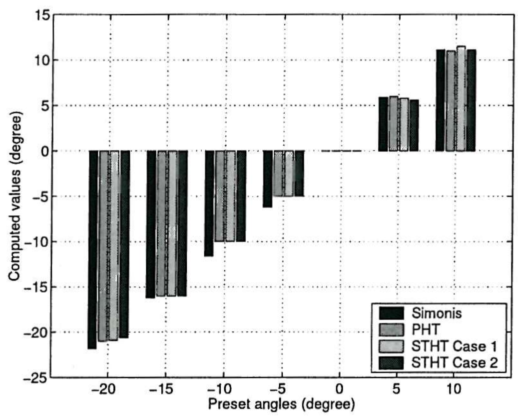
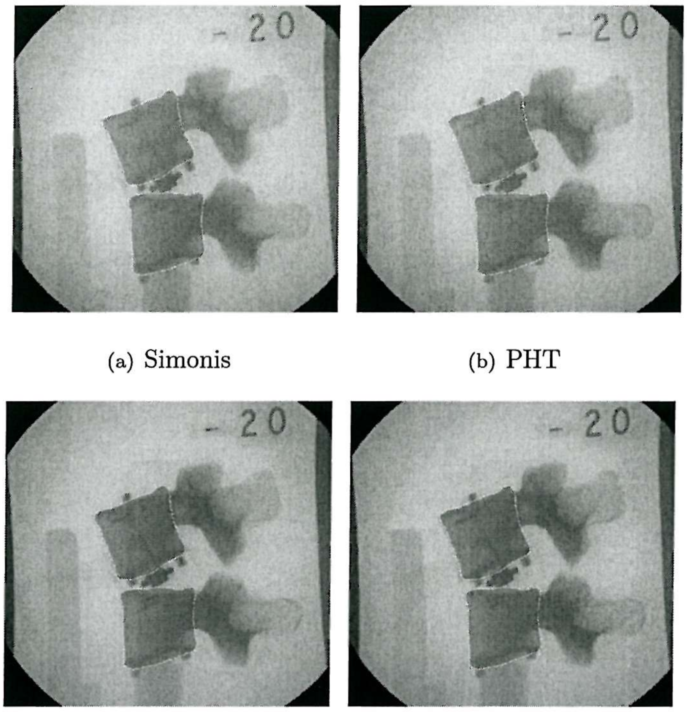


FIGURE 6.7: Result comparison of four studies on the calibration model.

6.6.3 Extractions of DVF Sequences

In the extraction of DVF sequences of nine subjects, 200 individuals are used in the GA. Each individual contains $3 * N$ parameters where the total number of frames N is 38. As described in the early part of this chapter, the length of each individual is 722 bits ($19 * N$). Similar to the studies on the calibration model, here the results from the PHT are also used to initialise the GA. The STHT algorithm has been applied to nine subjects in the order of L3, L2, L4 , L1 and L5. The overall results are promising and



(a) Simonis (b) PHT

(c) STHT case 1 (d) STHT case 2

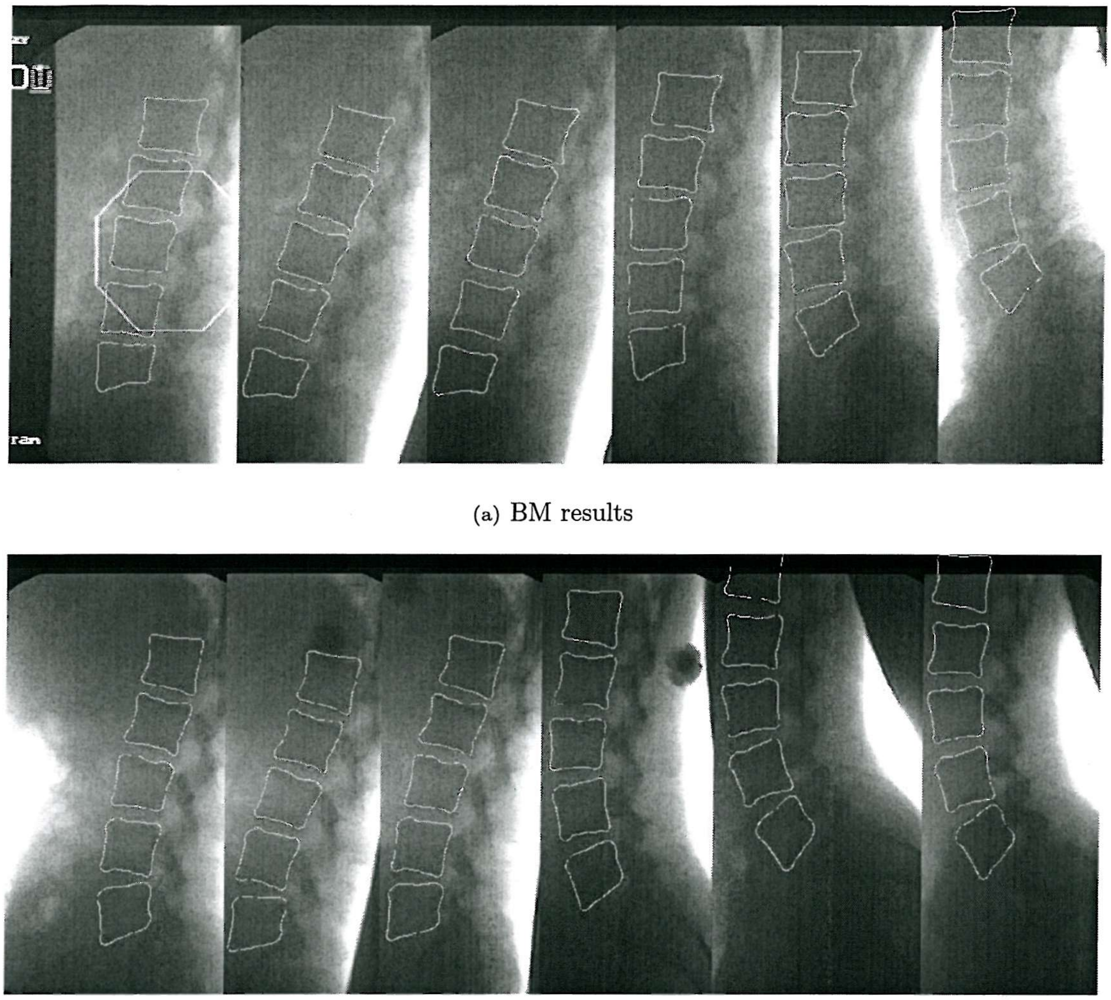
FIGURE 6.8: Illustration of the resulting images by four studies.

TABLE 6.1: The comparison of rotations amongst four studies.

Preset angle (degree)	Simonis' results (degree)	PHT (degree)	STHT case 1 (degree)	STHT case 2 (degree)
-20	-21.8	-21	-20.9	-20.6
-15	-16.2	-16	-16	-16
-10	-11.6	-10	-10	-10
-5	-6.2	-5	-5	-5
0	0	0	0	0
5	5.9	6	5.8	5.6
10	11.1	11	11.5	11.1

there is a large improvement over the PHT. Some results from two subjects are shown in Figure 6.9. More results will be provided in the following chapter where comparisons with the manual labelling are also given.

In the current study, the effect of excluding temporal and spatial information were considered by setting w_2 or w_3 to 0, respectively. Table 6.2 shows comparison of correct extraction rate of these studies for a total of 342 DVF images (38 frames for each of nine subjects). During the experiment, visual evaluation is obtained by observing the resulting images which are obtained by superimposing the reconstructed contours onto the original images. In the evaluation those resulting contours that are obviously away



(a) BM results

(b) DE results

FIGURE 6.9: Extraction of the lumbar spine by the STHT.

from the real boundaries of the vertebrae are labelled as incorrect extraction. This evaluation could be improved in future by quantitative comparison of the overlaying points if radiologists are available to help us to mark the boundaries of the vertebrae. It appears that an overall improvement has been achieved and especially there is significant improvement for the L1 and the L5. Compared to the improvement of L5, improvement of L1 is a little lower. The main reason is that most parts of the L1 are out of the scope in many DVF frames. In these cases the useful edges become too few and the STHT cannot detect it correctly. Therefore, in future DVF acquisition, extreme caution is needed to ensure that all the lumbar spine is contained within images.

Also, it seems that the introduction of the spatial information or temporal information alone with the Hough term can improve the correction rate, and the effect of the latter looks much more significant than that of the former. The reason for this may come from the way that they are used in the STHT, the latter is quantified by smoothness function while the former is just represented by a fuzzy penalty term. The combination of these two terms with the Hough space term can lead to a better result than any of them only.

TABLE 6.2: The comparison of correct extraction rate between the STHT and the PHT.

	L1	L2	L3	L4	L5
PHT	76%	93%	98%	94%	61%
STHT ($w_2 = 0, w_3 = 1$)	80%	93%	99%	95%	79%
STHT ($w_2 = 0.2, w_3 = 0$)	82%	95%	100%	98%	85%
STHT ($w_2 = 0.2, w_3 = 1$)	87%	97%	100%	100%	92%

6.7 Discussions

6.7.1 Computational Cost

Apart from the computation required in forming the Hough space, the computation is mainly consumed by optimisation using GAs. In practice there are different criteria of convergence of GAs, i.e. specified generation number, the best fitness value/the average fitness value etc, here a specified generation number 2000 is used as criteria to terminate the optimisation process. The time consumption depends on many factors such as chromosome length and fitness calculation. For our problem, it appears that the time taken by GAs is about 7 minutes for each vertebrae.

6.7.2 Accuracy

After an extraction (including landmark location) has been obtained, two questions appear paramount: How accurate is the extraction? and how can it be used? For spine kinematics, the answer to the latter question appears simple as these will be used to quantify the spine motion. Specifically, how to use the extraction results from the STHT will be discussed in detail in the next chapter. The other question concerning the segmentation concerns is the accuracy. There is a widespread quest for measure that somehow quantify the extraction accuracy, unfortunately there is no such measure in medical imaging.

In practice, accuracy can be divided into quantitative and qualitative accuracy. Quantitative accuracy needs a ground truth that is usually unavailable in clinical practice. The latter one is usually done by using simple visualisation. Compared to the former one, it is easy with a reasonable idea of accuracy. The results of the PHT and STHT are both evaluated in this way. It is worth mentioning that a further accuracy investigation is conducted in the next chapter by comparing the results of the STHT and the manual landmark location.

6.7.3 Endplate vs. Model

One concern is the model of the vertebral contour. The vertebral endplates have non-sharp edges and the projection of them on the sagittal plane may become fuzzy and may not be constant throughout the motion sequence. However, this suggests why the Hough transform has an advantage over the earlier manual methods from another aspect. As discussed earlier, the Hough transform can locate the vertebrae by collecting evidence over the whole contour rather than at several corner points and thus it will not be seriously affected by partial deformation, missing data, or inconsistent changes of the image illumination. In contrast, manual marking is difficult or impossible to cope with such cases. Although the endplate shape sometimes can be very helpful in diagnosing disease, as our main focus is the motion rather than exact shape of the endplate, we do not think it is vital in this study.

6.8 Conclusions

In this chapter the Spatio-Temporal Hough transform algorithm has been introduced. The STHT can achieve better performance than the traditional HT by the inclusion of the spatio-temporal information within the image sequence. In the STHT, an energy function is designed to combine the Hough spaces and the spatio-temporal information. A GA is used to locate the maxima of the energy function. The STHT performance was investigated by applying it to the DVF images of the calibration model and has shown acceptable results. The extraction results on nine normal subjects have shown that the STHT is a promising approach in motion analysis. Clearly, the inclusion of temporal constraints can improve the accuracy of the extracted results. Further, it has been shown how including spatio-temporal information improves the results, with the greatest increase due to the temporal constraints, as opposed to the enforcing of the spatial information. With the extraction results obtained, the spinal kinematic parameters can be determined and these will be discussed in the following chapter.

Chapter 7

Preliminary Study of Spine Kinematics

Spine kinematics is of great interest for better understanding of low back pain. A number of parameters has been used to quantify the motion of the spine, typically in the cervical and lumbar spine. This chapter provides an insight into these parameters. A preliminary study is conducted of spine kinematics based upon vertebral extraction results from DVF image sequences of nine subjects by the STHT. First, a short review on the commonly used parameters in spine biomechanics is provided. Based upon the results obtained from the complete version of the STHT (where weighting factors were set as $w_1 = 0.8$, $w_2 = 0.2$ and $w_3 = 1$, respectively), some of these parameters are calculated. Finally a comparison of results obtained from both manual marking and the STHT is made and some statistical evaluation results are presented.

7.1 Kinematic Parameters

So far, a number of parameters has been proposed in order to depict the motion pattern of the lumbar spine. One reason for the apparent range of parameters is due to the arbitrary shapes of vertebrae. Manual labelling aims to simplify this problem by finding corners that can be regarded as a simple description of the shape. If the vertebrae were conic sections, as such explicitly described by parameters, it is likely that this range would be smaller. However, the new use of shape definition could be help to reduce this range as an estimation of center positions and rotations. The other is due to the complexity of spine motion, most of these parameters are only able to partially describe the motion such as in rotation or translation. As discussed in chapter 2, the spine can rotate around its axis, move forward and backward and bend in a lateral direction. There are also coupled motions, especially in lateral bending. Some studies have suggested that the coupled rotation is not very large especially in extension and flexion (Panjabi and White 1971; Pearcy 1986). Therefore, we assume that pure planar motion is involved when the subject flexes and extends during the experiments. During DVF acquisition the patient was constrained on the passive motion table, and hence this assumption appears to be reasonable. As the focus of this study is not to evaluate these parameters, only some typical parameters are reviewed.

- Instantaneous Axis of Rotation (IAR)
  In the early studies, the IAR, also known as Instantaneous Center of Rotation (ICR), was one of the most often used parameters in spine kinematics. It can reflect the combination of the sagittal translation and simultaneous rotation. During flexion-extension, each vertebra exhibits a motion with respect to the lower one. In the IAR calculation, the lower vertebra is assumed to keep the original position while only the upper one moves in relation to it, as shown in Figure 7.1. The IAR lies at the intersection of the perpendicular bisectors of the displacement vectors (AA' and BB') of two markers A and B provided that a vertebra can be assumed to be a rigid body. The IAR is error sensitive (Dimnet et al. 1976) and the error can be unacceptably large if the magnitude of rotation is less than 5° and if the landmarks A and B are located at a distance of less than 30mm from the center of rotation. There have been a number of studies on its calculation and optimisation (Panjabi 1979; Bogduk 1997; Muggleton and Allen 1998; Challis 2001). In short, it appears that the IAR can be useful only when the magnitude of movement is sufficiently large. Also, if relative movement between two vertebrae exhibits a pure translation, two bisectors will become parallel which means that the estimate of the coordinates of the IAR will become infinite. This remains as the most significant limitation of the IAR.

- Centrode
  There is an intimate relationship between the centrode and the IAR. During spine movement, if the IARs are determined for each phase of the motion and then plotted in sequence, they will depict a locus known as the centrode of motion (Bogduk 1997). The centrode is like the path during the full range of motion of the joint, as shown in Figure 7.2. One study proposed that centrode patterns can be used to determine whether there are pathological changes in the vertebrae (Gertzbein et al. 1985). For example, in normal cadaveric specimens the centrode will be short and located within a restricted area in the vicinity of the upper endplate of the next lower vertebra, while it will exhibit longer, displaced and seemingly erratic behaviour with degenerative vertebrae. However, Pope et al. (1999) noticed

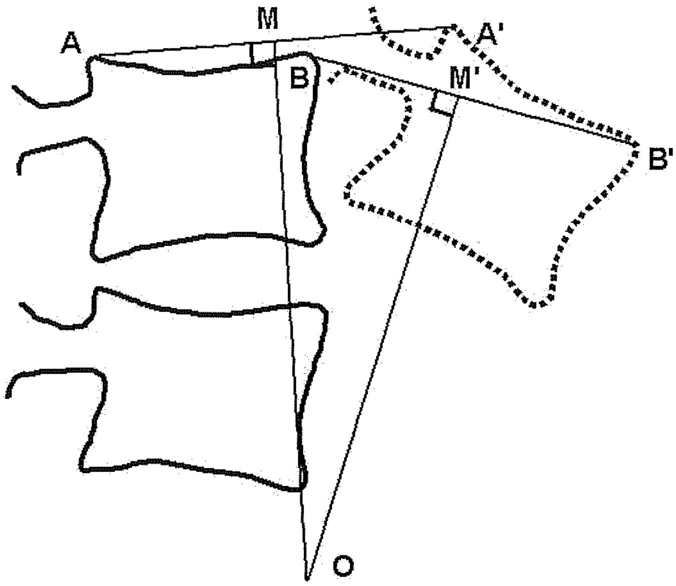


FIGURE 7.1: The location of an IAR.

Gertzbein et al. (1988) has reported that this pattern cannot be established in their later study. Unfortunately, this finding has been overlooked or neglected by most studies. A centrode is a collection of a series of IARs and as such is subject to potentially large errors for small measurements. It appears to be impractical to use the centrode to distinguish between normal and abnormal motions for the LBP patients with limited movements, since the magnitude of each movement may be very small (Pearcy and Bogduk 1988). Furthermore, the centrode is meaningless when the motion is only examined at the extreme positions.

- Intervertebral Angle (IVA)
  The intervertebral angle is a parameter to measure the rotation movement between two neighbouring vertebrae. It can be measured directly by the angle between lines across their vertebral endplates, as shown in Figure 7.3. These lines are defined by anatomical landmarks. As has been discussed in chapter 2, with manual marking it is unlikely to locate the same landmarks in different images.

- Sagittal Translation
  Translation is often used in sagittal motion analysis. There are different notations to define translation (Muggleton and Allen 1998). Figure 7.4 shows one (Frobin et al. 1996). For two neighbouring segments, four corners of each vertebral body are used to define a midplane. The sagittal translation is defined as the change along the bisectrix of two midplanes against the height of the superior vertebral body. One should notice that this definition relies on more landmarks than the IVA and is potentially subject to larger errors caused by landmarking.

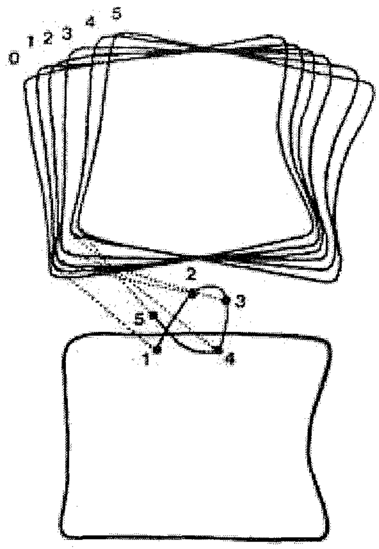


FIGURE 7.2: Illustration of the centrode (1,2,..,5 in the middle of two vertebrae denote IARs and 0,1,...,5 on the top are different positions during the motion.), adapted from Bogduk (1997) with permission of Elsevier.

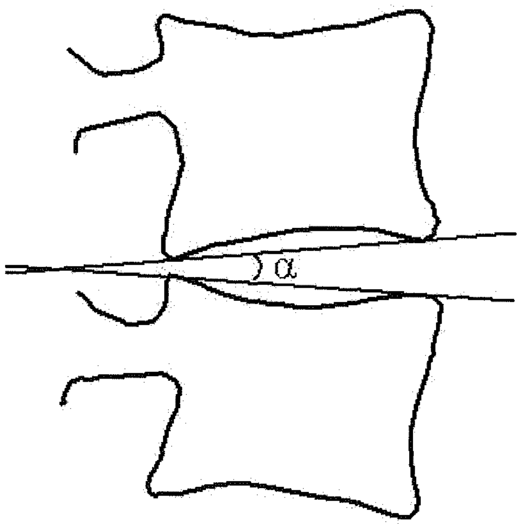


FIGURE 7.3: Illustration of the intervertebral angle.

- Axial Translation

  It is believed to be valuable in clinical application as it has potential to represent the intervertebral stiffness. It can be defined as the percentage of the average intervertebral disc depth over the average superior vertebral body depth.

- Neutral Zone (NZ)

  Neutral zone is a part of the range of motion of a vertebral body, starting from the neutral position up to the beginning of some resistance offered by the joint (White and Panjabi 1990). This is illustrated in Figure 7.5. It reflects the sense of excessive

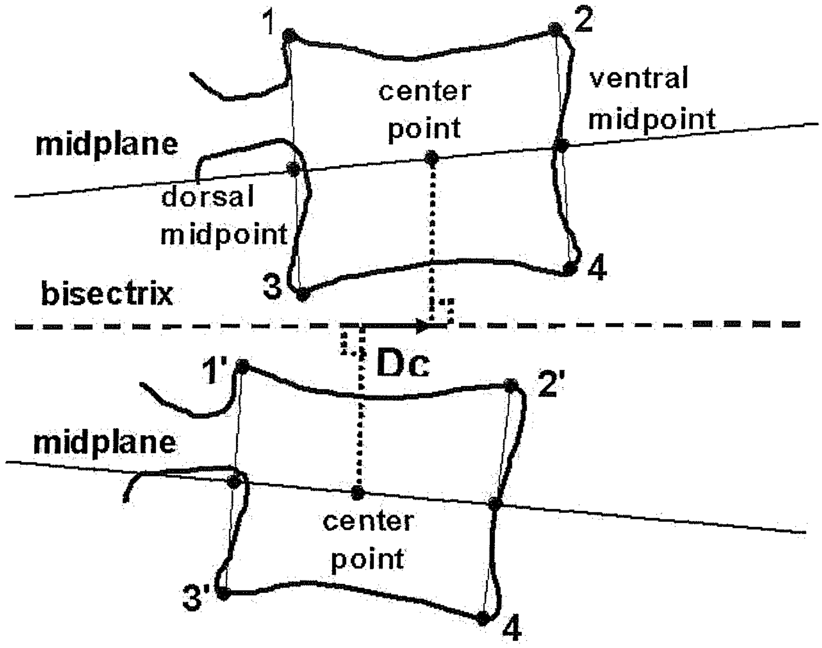


FIGURE 7.4: An illustration of the sagittal translation.

displacement under minor load. An increased value of NZ is believed to be a factor of instability.

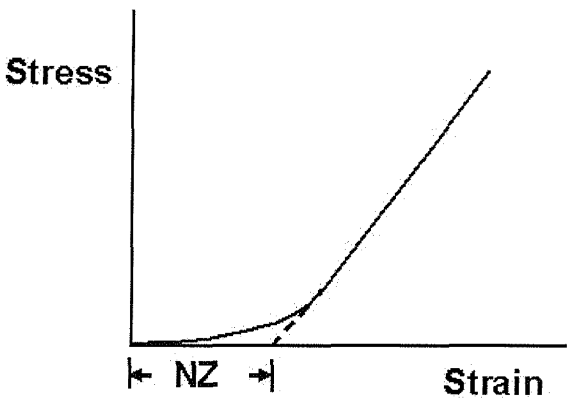


FIGURE 7.5: Illustration of the neutral zone, adapted from Bogduk (1997) with permission of Elsevier.

- Range of Motion (ROM)

  Range of motion is defined as the displacement between the extremes of the physiological range of motion or rotation of a joint. It might be useful to represent the whole range of the motion, but it is limited by many factors. For example, movements of LBP patients may be affected by pain and consequently ROM cannot represent the degree of the LBP impairment. One study suggests that neutral zone is more appropriate than the ROM in characterising instability (Panjabi et al.

1994).

There are other parameters derived by measuring position changes of vertebrae in the images. For example, the curve along the posterior longitudinal ligament between the inferior T12 (the 12*th* vertebra of the thoracic spine) to the superior S1 is assumed to be portion of an ellipse. Thus an elliptical model is used to discriminate between normal and low back patients (Harrison et al. 1998). Kanayama et al. (1996) proposed phase lag to evaluate the sagittal motion. The relative position of the pedicles (or the spinous processes) against the vertebra is also used in quantifying the vertebral axial rotation, on which there is a good comparison conducted by Russell et al. (1990).

Some parameters based on direct measurement have also been used. Fingertip-to-floor distance measures the distance from the tip of the middle finger to the floor when the patient maintains full flexion posture (Gill et al. 1988). However, it cannot reveal motions in the segmental level and measurement repeatability is problematic due to the effects of many vertebrae, shoulder, elbow, wrist etc. involved. Waddell (1999) even suggested that it is inappropriate to use it as a measure of lumbar flexion. The Schober technique is also used to measure the flexion: with the subject standing erect, make a mark at the approximate position of the lumbosacral junction, make a second one 10*cm* higher and a third mark 5*cm* lower. The distance between the latter two marks are measured when the subject flexes and extends (Gill et al. 1988). This is still unable to measure the motion in segmental level. Together with parameters directly measured by inclinometers, these are adequately reviewed by Gill et al. (1988).

From the above discussion, it can be seen that accuracy of most kinematic parameters, e.g. IAR, centrode, IVA and sagittal translation, largely depends on the accuracy with which the landmark points have been located. For example, the errors caused by manual landmarking has been regarded as the major source in the errors of the IAR (Panjabi et al. 1992). It is desirable to develop a landmark location approach that can identify landmarks with high accuracy, efficiency and reliability, as has been the focus of earlier studies (Simonis et al. 1993; Muggleton and Allen 1997; Cardan and Allen 2000). This highlights the great value of the newly developed STHT.

7.2 Calculations of Kinematics

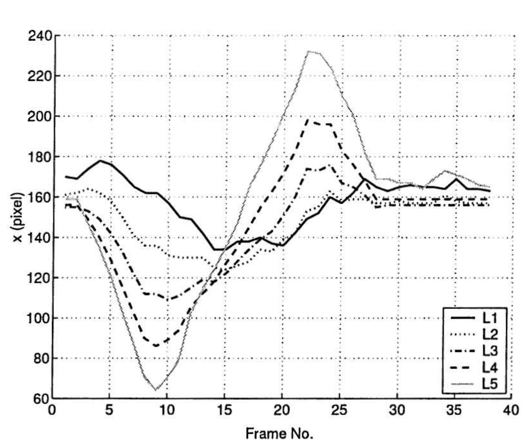
As discussed in the previous chapter, the STHT has been developed to analyse the dynamic motion of the spine in an image sequence. Its performance has been reinforced by incorporating the spatio-temporal information within the analysis. The STHT can extract vertebrae within a DVF sequence by detecting contours that are predefined as target shapes. After extraction, these contours, the positions of their centers, together with their orientations, are known. Figure 7.6 shows the extraction results of one subject. When the subject was taken in a passive flexion/extension motion as described in chapter 3, the vertebral centers depicted a sinusoidal pattern in the forward and backward directions Figure 7.6(a), a sinusoidal pattern was also observed in rotation with respect to the vertebral centers 7.6(c). It appears that the centers of the vertebrae were slightly stretched away from the head at the extreme flexion while they were clearly pushed towards to the head at the extreme extension 7.6(b).

From Figure 7.6, it seems that the lumbar spine cannot return to the initial position. This has been observed with some subjects. There are several possible causes, e.g. the subject shifting during the experiment, measurement errors, the variability of the neutral position or perhaps the nature of passive motion protocol in which the muscles and ligaments of the subject involved are still in a relaxed condition.

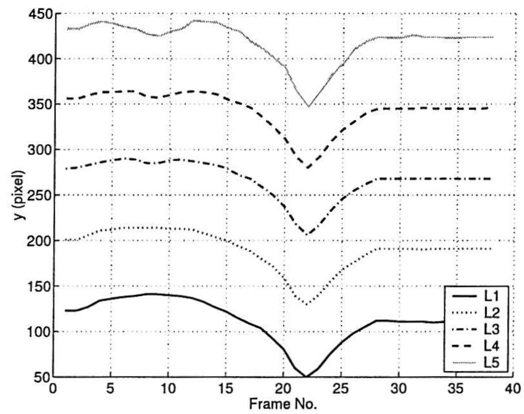
In contrast to the manual identification, where only a few anatomical landmarks (usually four corners) are located, the STHT can present much more useful information. For example, positions of any points on the contour can be determined mathematically and uniquely from the extraction. This will be discussed next. This guarantees that the same points are used in the parameter calculation and can avoid the location errors caused by inter/intra observers in the manual work. In addition, the STHT aims to locate the whole contour rather than several discontinuous points, thus it will not be greatly affected by missing data, partial occlusion and noisy frames, while manual marking appears problematic in these cases. Therefore, the STHT has a significant advantage over manual location and can be a powerful tool in spine kinematic measurement.

With the power provided by Fourier descriptors to effectively describe the arbitrary shape, a valuable point to be addressed here is that the parameters (positions of the centers and rotations) determined by the STHT may serve as measures to quantify the motion of the lumbar spine. These parameters appear more intuitive than those used before and are sufficient to describe the positions of a moving rigid arbitrary-shape object such as vertebra. This can be seen from Figure 7.7 where the squares denote the center positions of the vertebrae and the dotted lines show the virtual connections between these centers. This is further shown in Figure 7.8.

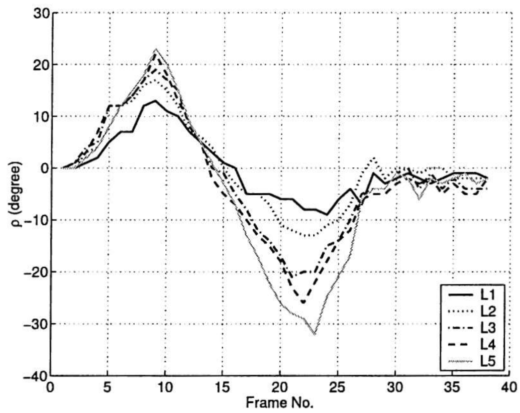
However, questions may arise, i.e. whether parameters obtained will still be meaningful when the contour constructed is not as same as the true contour of the vertebral body (Figure 7.9(b)). Two cases are shown in Figure 7.9(a) and 7.9(c). From a mathematical view, the rotation values obtained by using either Figure 7.9(a) or 7.9(c) as model will remain the same as that obtained by using the model as shown in Figure 7.9(b). As for



(a) x translation



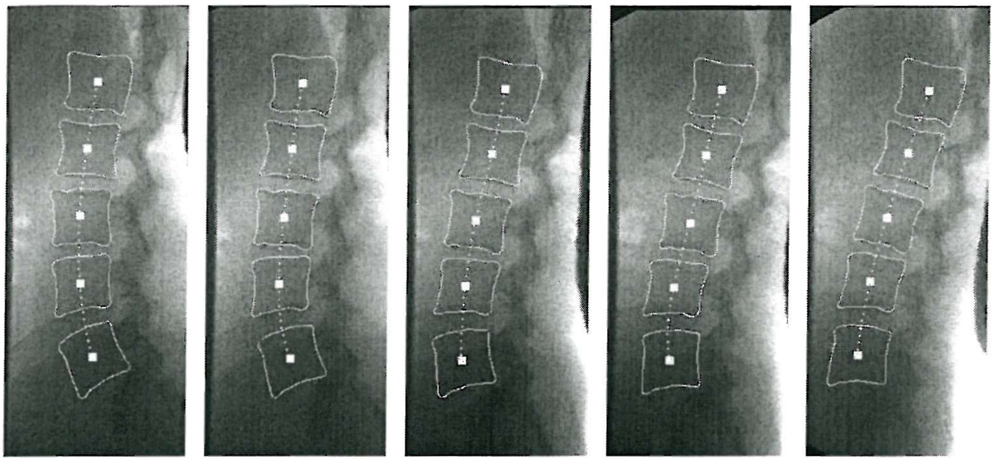
(b) y translation



(c) Rotation

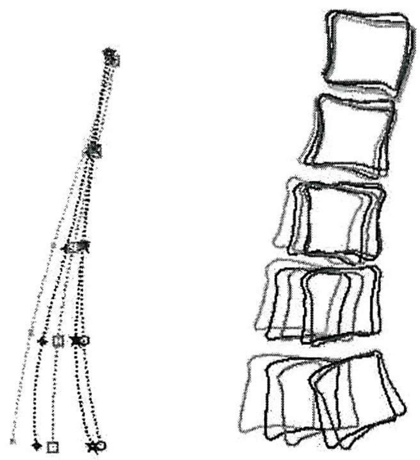
FIGURE 7.6: Results on one subject.

center positions, there will be a constant difference when using Figure 7.9(a) and 7.9(b), so do when using Figure 7.9(b) and 7.9(c) as models. However, the relative motion patterns of centers in a motion sequence with respect to the first frame will be the same. Therefore, these parameters can still be useful.



(a) Frame 1 (b) Frame 2 (c) Frame 3 (d) Frame 4 (e) Frame 5

FIGURE 7.7: Illustration of center connections.

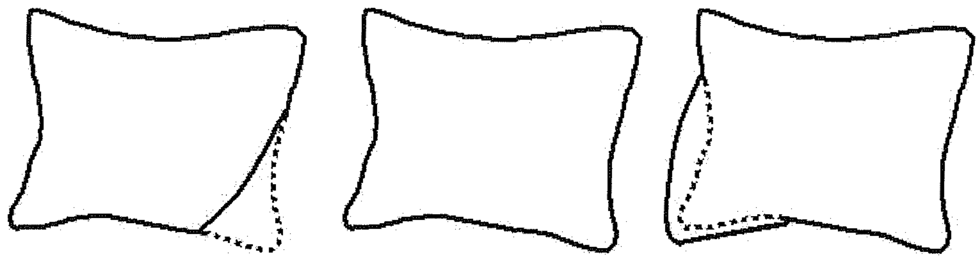


(a) (b)

FIGURE 7.8: Illustration of motion. (a) Center connection using smooth curves. (b) Motion shown by the contours.

7.2.1 Corner Derivations

In spine kinematics, most conventional parameters are derived by positions of two or more landmarks. Although the STHT can provide more information on the positions of the lumbar spine, this cannot directly be used to calculate those parameters. Here we



(a) Incomplete shape (b) Complete shape (c) Partial deformed shape

FIGURE 7.9: Illustration of variations of model shapes.

illustrate how landmarks, four corners in particular, can be obtained from the STHT extraction results.

Corner detection is achieved by examining the values of curvature κ along the contour. Given a 2-D curve, the curvature κ at any point on the curve can be expressed as

$$\kappa = \frac{\frac{d^2y}{dx^2}}{\sqrt[3/2]{1 + (\frac{dy}{dx})^2}} \tag{7.1}$$

where dy/dx and d^2y/dx^2 are the first and second derivatives. The curvature κ can be rewritten as $\kappa(t) = \frac{\dot{x}(t)\ddot{y}(t)-\dot{y}(t)\ddot{x}(t)}{[\dot{x}^2(t)+\dot{y}^2(t)]^{3/2}}$ for a curve in parametric form $y = y(t)$ and $x = x(t)$, as used in Equation (5.5) (details about curvature can be found in James (1992).). In Equation (5.5), t is used to parameterise the position along the curve and its range is between 0 and 360°. The t is the only unknown parameter, provided that the Fourier descriptors have been determined.

The curvature can represent the sharpness of the local area. That is, a point with a larger value of κ means the radius is smaller at that point. Therefore, a corner can be located by detecting a relatively large value of κ.

As each vertebral body is approximated by a closed curve on which there are roughly four corners, the curvature at the corner should be larger, at least in the local area. A simple method is used to define four corners. That is, the first corner is located by finding the point corresponding to the global maximum of the curvature, the other three corners can be detected by looking at approximately 90° intervals.

This approach can succeed in locating the corners. Figure 7.10 shows this procedure on 5 vertebral models of one subject where the stars denote the position of the corners and the circles indicate positions of peaks. In all but one case, the corners appear to form sharp peaks of the curvature function. In one case, as shown in Figure 7.10(a), the value of κ at the lower right corner appears to be close to that of the neighboring areas because of local smoothness. A more rigorous approach may be needed and some methods can be found in the book by Nixon and Aguado (2002).

After four corners of the target contour are determined, their positions in the DVF sequence can be derived from the extracted rotations and translations by the STHT. This enables us to calculate conventional kinematic parameters with potentially increased repeatability and accuracy.

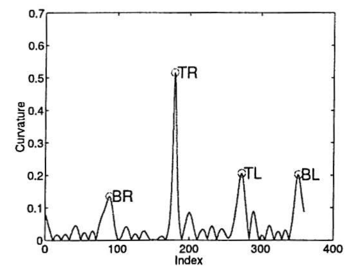
7.2.2 Results of Rotations and Intervertebral Angles (IVAs)

If landmark data are available, most conventional kinematical parameters can be easily derived by applying elementary geometric techniques. Moreover, the main aim of this study is to solve landmarking problems. Only the results of rotations and intervertebral angles are presented. The former describes the motion of single vertebra while the latter is effective for measuring the relative rotation between neighbouring vertebrae. Since all lumbar vertebrae in the image are moving, it may be more appropriate to study the motion between two neighbouring segments.

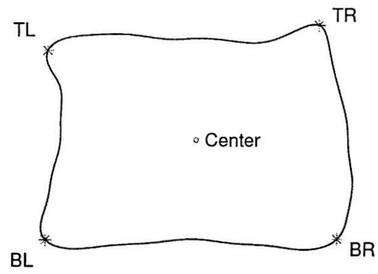
During calculation, the initial orientations of the vertebrae are normalised to zero so that all angular measurements are defined relative to the starting position, in this case, the neutral position. It should be emphasised that the results of L1 lack significance as L1 is out of view in many DVF images.

7.2.2.1 Rotations of Vertebrae

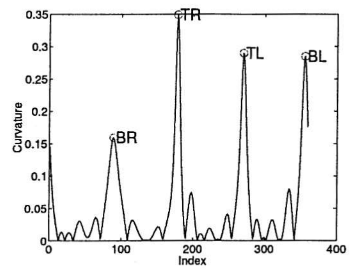
The absolute rotations of five vertebrae of nine subjects are shown in Figure 7.11(a) to 7.11(e). Average rotations across the 9 subjects were also calculated for a better understanding of the motion pattern and these are shown in Figure 7.11(f). Due to the passive motion protocol used in the experiment, L5 vertebra has the largest rotation while the rotation of L1 is relatively small. It seems that the maximal rotations occur at the extremes of flexion and extension. The average curve of the L1 rotation around the vicinity of the extreme extension appears less smooth than those of other vertebrae. This is caused by the out of view problem which has been discussed earlier. The problem of how to constrain the subject in DVF acquisition needs consideration in future work.



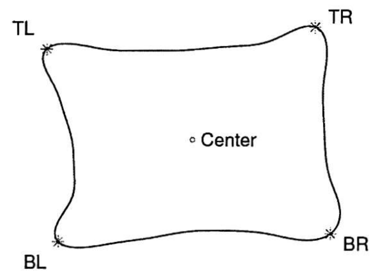
(a) κ values



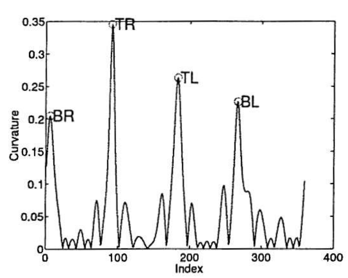
(b) Corners



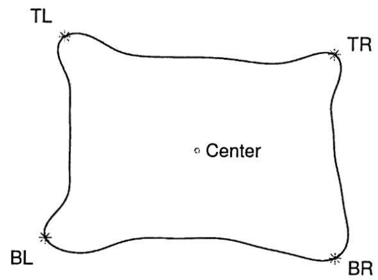
(c) κ values



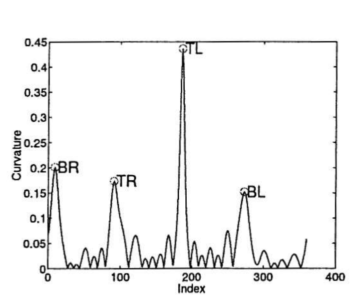
(d) Corners



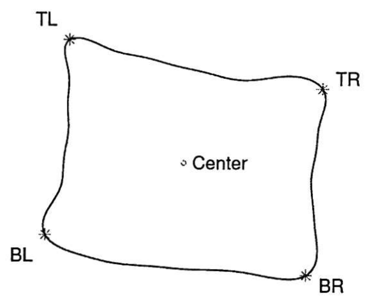
(e) κ values



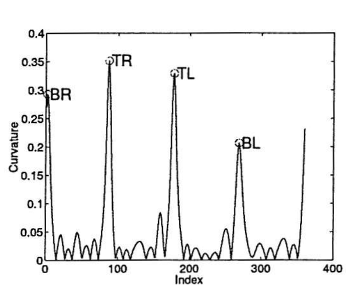
(f) Corners



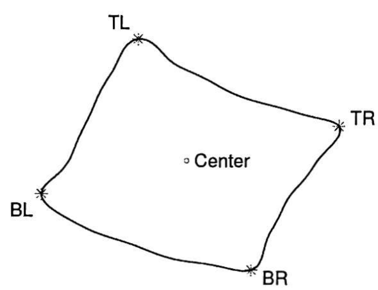
(g) κ values



(h) Corners

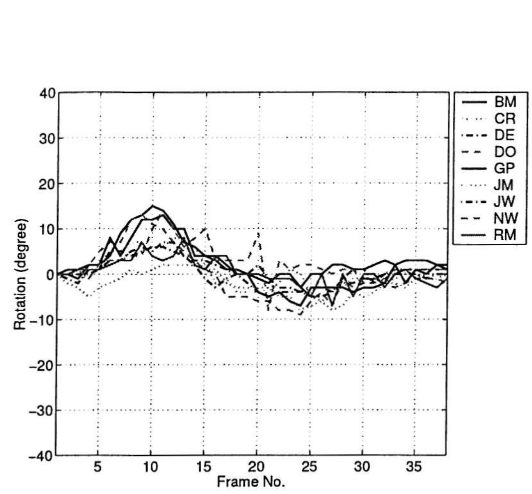


(i) κ values

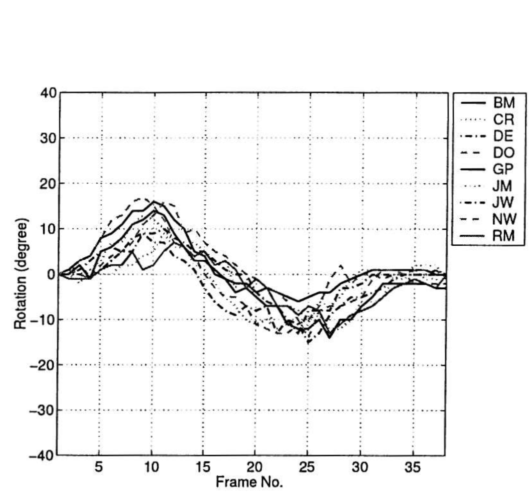


(j) Corners

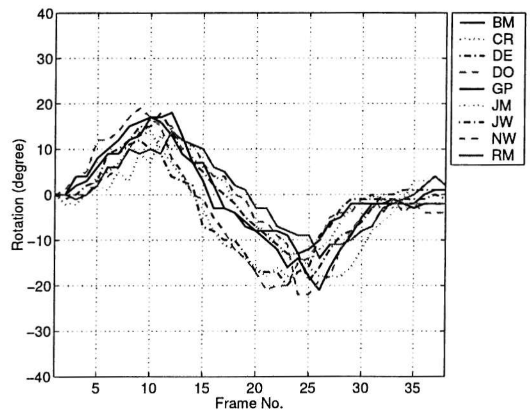
FIGURE 7.10: Illustration of how to locate corners (TR: top right corner; TL: top left corner; BL: bottom left corner; BR: bottom right corner).



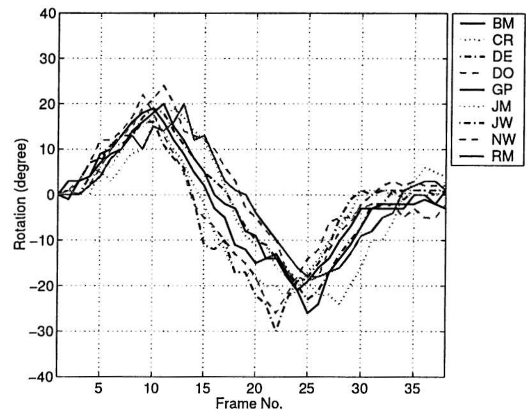
(a) L1 rotations



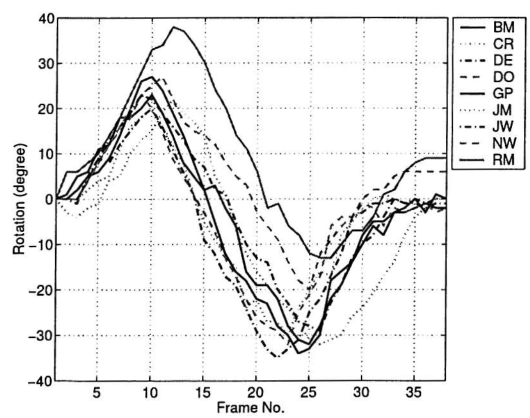
(b) L2 rotations



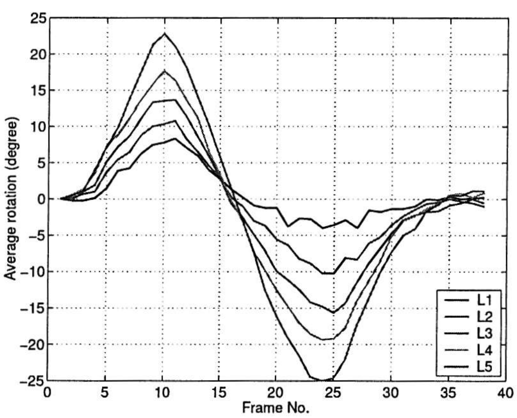
(c) L3 rotations



(d) L4 rotations



(e) L5 rotations

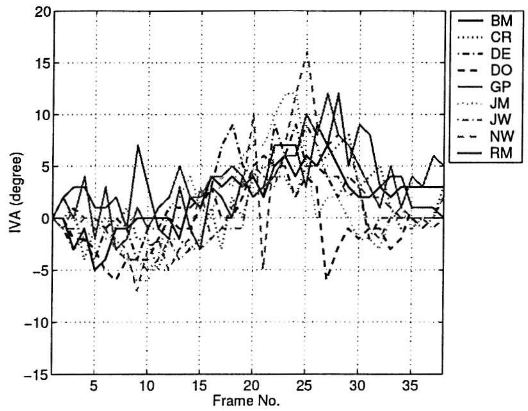


(f) Average rotations

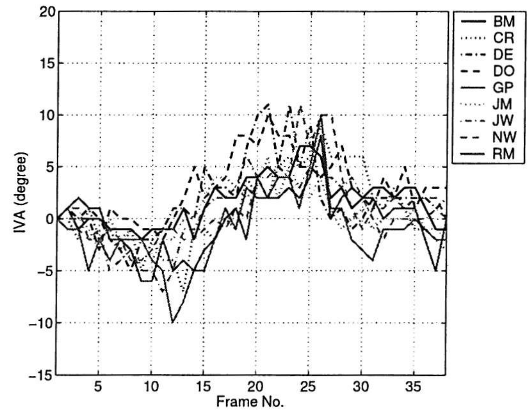
FIGURE 7.11: Rotation of the lumbar vertebrae in 9 subjects.

7.2.2.2 Intervertebral Angle (IVA)

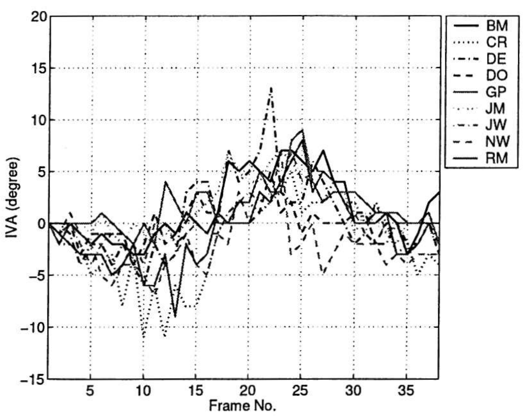
From the vertebral extraction data, we can obtain the rotations of each vertebra ρ_i^j, $j = 1, ..., 5$ and $i = 1, ..., N$, where N is the total number of frames used in a sequence. Thus the intervertebral angles can be obtained by subtracting the rotations of neighbouring vertebrae. That is, $IVA = \rho_i^k - \rho_i^{k+1}$, where $k = 1, 2, 3, 4$. The results are shown in Figure 7.12. Note that the initial IVAs are zero because of the normalisation. Overall, it appears that the maximum IVA value is less than 15° and IVAs between L4 and L5 are larger than others. It can be seen from Figure 7.12 that the IVA is not a smooth function during the motion and this implies that the motion of neighbouring vertebrae does not change in the same pattern during flexion and extension.



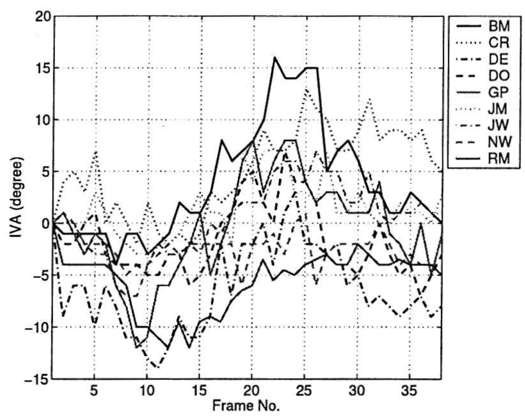
(a) L12 IVAs



(b) L23 IVAs



(c) L34 IVAs



(d) L45 IVAs

FIGURE 7.12: The IVAs of the lumbar vertebrae in 9 subjects.

7.3 Comparison with Manual Labelling

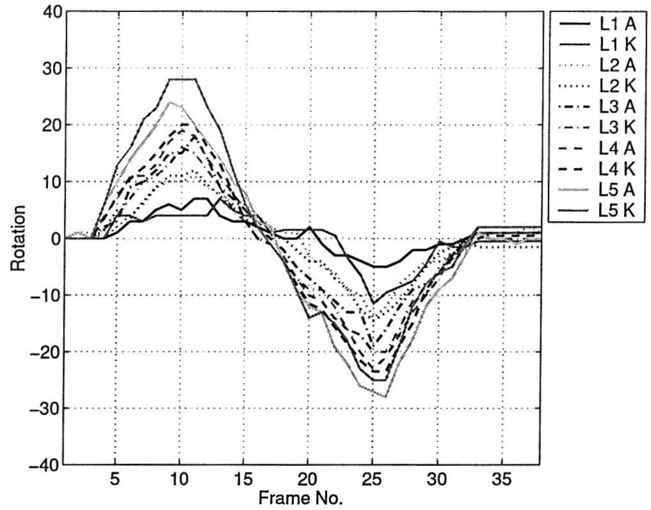
It is often difficult, or even impossible, to obtain "ground truth" in medical imaging analysis. Here is such a case. The data of M. Kondracki (2001) from manual landmark identification was used as a measure of validation.

The comparison of rotations from the STHT and from manual marking is presented in Figure 7.13(a), where 'A' means results of the STHT and 'K' means that of manual marking. The motion pattern looks very similar for most of the vertebrae of the subject. L1, however, shows large variations which appears to be due to the problem in data acquisition, as discussed earlier. Figure 7.13(b) shows the difference between two studies with respect to the maximum of the manual labelling. To further evaluate the relationships between the two methods, a statistical analysis was conducted. The value of R^2 returns the square of the Pearson product moment correlation coefficient R through data points in two variables x and y. R is defined in Equation (7.2). R^2 ranges from 0 to 1 and can reveal how closely y corresponds to x. If it is near unity, the match between two variables is deemed to be close. Here x and y denote the results obtained by manual labelling and the STHT, respectively.

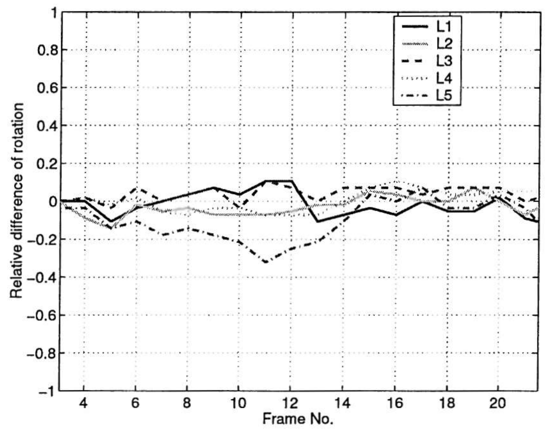
$$R = \frac{n(\sum(xy)) - (\sum(x))(\sum(y))}{\sqrt{[n\sum(x^2) - (\sum(x))^2][n\sum(y^2) - (\sum(y))^2]}} \tag{7.2}$$

For each lumbar segment across all subjects, the value of R^2 is calculated and is shown in Table 7.1. From this table, it can be seen that the results of the two methods statistically match very well except for L1. Surprisingly, despite the poor image quality in the L5 area, the value of R^2 for L5 is very good. As discussed in the previous chapter, the correct extraction rate of L5 is not very good, this might imply that the manual labelling may also have a large uncertainty for L5. In another test, the rotations for each segment across all subjects were grouped and the R^2 values are 0.465, 0.776, 0.829, 0.855 and 0.784 for L1 to L5, respectively. Again, the R^2 value for L1 is much lower than those of other segments, the reason is mainly due to the problem of L1 occasionally being out of the field of view of the fluoroscope.

The use of correlation coefficients sometimes could be inappropriate and misleading, and for this reason we looked at linear regression. Given two observations X and Y, linear regression attempts to explain this relationship with a straight line fit to them. The linear regression model postulates that $Y = aX + b + \epsilon$, where the "residual" ϵ is a random variable with mean zero. The coefficients a and b are determined by the condition that the sum of the square residuals is as small as possible. Here linear regression analysis was also used to determine the relationship between the manual labelling rotation values



(a) Comparison



(b) Relative comparison

FIGURE 7.13: Rotation comparison between automated extraction and manual marking for a subject.

TABLE 7.1: The R^2 values between rotations determined by the STHT and by manual labelling on 9 subjects.

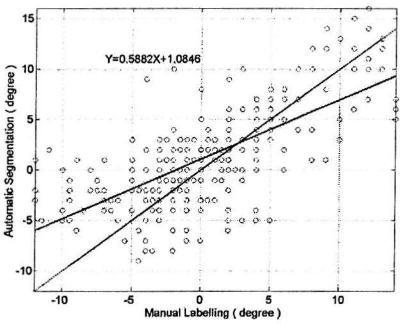
Subject	L1	L2	L3	L4	L5
1	0.784	0.954	0.968	0.973	0.943
2	0.157	0.927	0.971	0.969	0.976
3	0.750	0.916	0.979	0.978	0.984
4	0.490	0.884	0.954	0.936	0.929
5	0.767	0.923	0.965	0.983	0.937
6	0.700	0.960	0.973	0.970	0.958
7	0.720	0.947	0.982	0.991	0.975
8	0.334	0.922	0.959	0.933	0.967
9	0.219	0.897	0.938	0.952	0.929

and those obtained by the STHT on 9 subjects. That is, the 342 values obtained by the STHT and 342 values marked manually were evaluated for each vertebrae. Ideally a should be one and b will be zero if measurements obtained from these two approaches match perfectly.

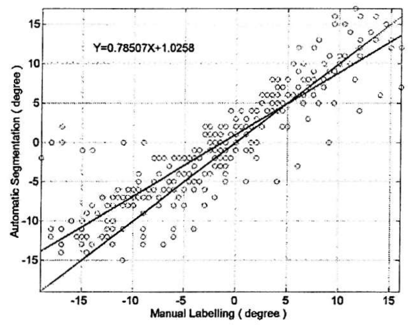
The relationships corresponding to the five vertebrae appear good, especially for L2, L3 and L4, as shown in Figure 7.14 where the thick solid lines denote the regressed lines, the dotted lines denote $Y = X$ while the "o" denotes the data points. The regression equations are also shown in Figure 7.14 with 95% confidence intervals, respectively. Some disparity was observed for L1 and the underlying reason might arise from the out-of-view problem. The regression of L5 appears different from that for L1. The slope of the regressed line is close to one while the intercept has a value over 2.5 degrees which is much higher than for the others. This could be explained by the difficulty in identifying L5 which is corroborated by visual analysis. Overall, this regression analysis confirms that the accurate extraction appears to have been achieved and we await confirmation of the potential of this new technique on a much larger dataset.

7.4 Conclusions

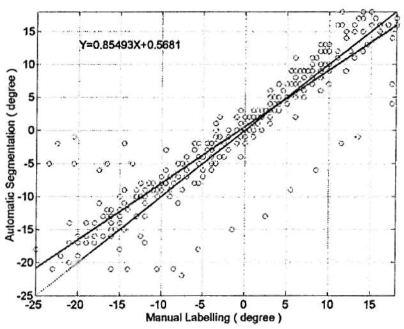
This chapter has discussed the parameters that are commonly used in spine kinematics. Most of them can be derived from landmark data and their accuracy largely depends on that of the landmarks. After the contours of the vertebral bodies are identified using the STHT, it is necessary to determine some points as landmarks to study the kinematics. Here, a simple method that can detect four corners by curvature was proposed. In this way, problems in the conventional landmarking approaches can be overcome. Finally, some results of the estimation of vertebral rotations and intervertebral angles were presented. As a validation measure, the rotations obtained from the STHT and from manual landmarking correlate quite closely and the statistical results are encouraging.



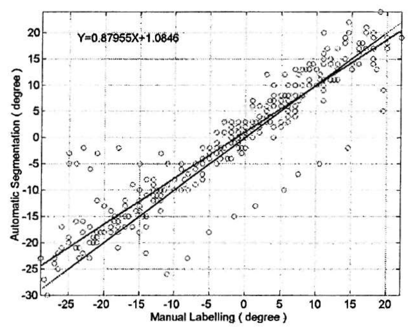
(a) Linear regression on L1



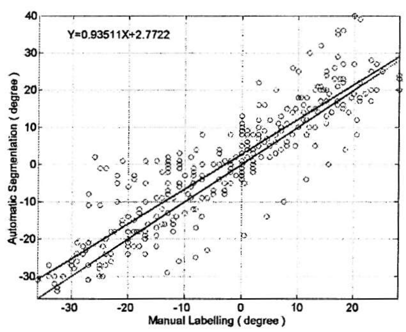
(b) Linear regression on L2



(c) Linear regression on L3



(d) Linear regression on L4



(e) Linear regression on L5

FIGURE 7.14: Linear Regression analysis on the manual labelling rotation values and those obtained by STHT on 9 subjects.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

Low back pain is a very common problem. Both the problem and its associated disability have appeared to escalate with time even though there is continuous and considerable progress being made by means of diagnosis, in appearance of new medicines and increasing knowledge of the human body. The direct and indirect cost of back pain have become a significant problem.

Despite the high occurrence of back pain, uncertainty still exists about its exact causes. The main problem is that it might be caused through multiple sources, that is, not only from the spine itself, such as through injury or pathological problems, but also through other structures such as nerves. Even psychological factors can result in low back pain. On the other hand, the complexity of the structure and the difficulty in accessing the spine also limit an exact diagnosis.

One of the main functions of the spine is to provide mobility of the human body, therefore there must be some relationships between the motion and low back pain. It is difficult to determine the relationships between them, i.e., does abnormal motion cause low back pain, or does low back pain cause abnormal motion or is there an interaction between them? Most problems, however, have a direct or indirect influence on the movement of the vertebral bones and can be catalogued as mechanical disorders. In order to understand the mechanical behaviour, specifically of the lumbar spine, a number of studies has been conducted, and again various parameters have been proposed to depict the patterns of the spinal kinematics.

In spine kinematics, radiographic techniques are the most commonly used methods to capture the motion of the spine. There are two significant issues with radiographic imaging. First, is it possible to obtain real time motion of the lumbar spine with as low risk as possible to the subject? In attempting to answer this question, DVF has been developed and indeed it can provide images of continuous motion of the lumbar spine at the segmental level, while the radiation involved is much lower than that of plain X-rays. Unfortunately, the quality of DVF is compromised by the reduction in the radiation dosage. The second question is whether or not it is possible to identify the vertebrae from the images obtained, accurately and efficiently. After a motion sequence (in particular DVF here) is obtained, it is crucial to determine the positions of the lumbar spine in all frames. In order to determine the position of an object in 2-D space uniquely, at least 2 points of the object have to be identified from the image and this procedure is referred to as "landmark location" in spine kinematics studies. Previous methods including manual work and some so-called automatic methods suffer many problems. Thus, the objective of this study was to develop a powerful tool for eliminating the problems associated with those previous methods. This has been achieved by solutions based on the Hough transform.

The Hough transform is a very powerful tool in computer vision and it can locate a target object in an image by gathering evidence from the edge information. Edge detection is also a very active area in image processing. In this work, phase congruency was employed to detect edges from DVF images, owing to its good performance in coping with images of low quality and variable illumination levels together with no need to tune thresholds.

The Hough transform has many variations and here a continuous version has been used in which Fourier coefficients are used to represent the model shape continuously in order to avoid discretisation and rounding errors. This method has been applied to synthetic images both with and without noise and occlusion, and also to a calibration model. These experiments showed good results. The extractions from the DVF sequence were also very encouraging.

In some cases, however, the PHT cannot yield correct results especially for extrema vertebrae such as L1 and L5 because of the inferior image quality. To improve the performance of the basic Hough transform, a Spatio-temporal Hough transform (STHT) has been developed. In the new method the contextual information within the motion sequence has been exploited to augment the ability in coping with noise, occlusion and partial missing data. Together with the Hough space this information is used to form an energy function. Compared with the traditional Hough transform, it is the maxima of the energy function in the STHT that correspond to the solution we are looking for rather than the maximum of the Hough space in the traditional HT. Due to the high dimensionality involved, a genetic algorithm (GA) is used to handle the search problem. Results from the DVF sequences of nine subjects are very promising.

Finally, some kinematic parameters have been calculated based on the results obtained by using the new STHT. In order to validate the new method, a comparison has been conducted in which the results derived from expert manual marking were compared with results from the STHT. Statistical comparison of these results showed considerable potential for the new approach.

In short, the landmarking problem in spine kinematics actually is motion sequence analysis. The Parameterised Hough transform (PHT) is a powerful approach that can handle the segmentation problem in a single image. When it is applied to a motion sequence, it works on the frame by frame basis and does not consider the potentially valuable information within the sequence. Thus its performance is limited in handling motion sequences. The new STHT, however, can take advantage of this useful information and thus is powerful for motion sequence analysis. That is, the STHT has extended the ability of traditional HT to handle motion sequences by collecting spatio-temporal information within the sequences.

8.2 Future Directions

Future work will lie in the following areas:

The STHT has proved to be a powerful tool in motion analysis, however, currently it is only a framework and more theoretical study is needed.

- The first question is on the spatio-temporal information. Currently, only the constraints on the trajectory and the spatial positions have been combined with the Hough transform. Additional *a priori* knowledge, e.g. gradient information, could also be included in the STHT. In order to achieve this, however, we must represent them and incorporate them in a mathematical form. These could be extremely valuable in further improvement of the existing STHT.
- The Hough transform is a model-based method and usually a unique model has to be built for each target object. Accordingly, the second question is whether or not it is possible to extend the STHT to cope with shape similarity? If so, this could be very useful in extraction of similar shapes such as the lumbar vertebrae.
- Thirdly, a thorough performance analysis of the Hough transform would be desirable. As has been stated, although a large variety of Hough transforms has been proposed, very few studies have focused on the property of the Hough transform itself. Analysis in this area might be helpful in answering some fundamental questions, e.g. how to select intervals in quantising the parameters, how to form

the accumulator space etc. This may give us new guidance for improving the performance of the HT.

- A further area is the construction of the model shape. So far, this was achieved by manual initialisation. This is the main reason why the STHT is called an "automated" method. The significance of automatic initialisation is that it can avoid possible errors in manual model construction (in particular, different initialisations may occur when different persons are involved.). In turn, this may affect the Hough transform. However, this automatic initialisation appears to be impractical for arbitrary shapes as it can be directly used for segmentation of the motion sequence effectively if such an approach exists.

- The Hough transform has the ability to handle 3-D object extraction. Thus, it has the potential to handle the out-of-plane motion problem if such information can be recorded by certain forms of images. This encourages us to develop a 3-D version of the STHT.

Another area would be to find more areas of applications. First, more analysis of the spine kinematics would be of great interest. At present only data from nine subjects were analysed. If possible, we would like to apply the STHT to a larger database. This will not only further validate the STHT, but also enable us to have an insight into the motion patterns, i.e., what is normal and what is abnormal motion? What is the relationship between the motions and low back pain? These could be very useful for clinicians as a powerful tool for better understanding of low back pain and improving its diagnosis, treatment and rehabilitation. Second, the STHT can be applied to other medical kinematics studies, e.g. pelvis, knees and finger joints etc. This will be very useful in decision making and also for evaluating the outcome of a joint surgery. As a generic method, it could be applied to motion analysis in broader areas than just spine kinematics.

New imaging techniques are also desirable. They should not only have the ability to record real time motion, but also be expected to yield images with better quality than DVF whilst minimising/eliminating the potential radiation risks to the subject.

Finally, but non-trivially, some studies might improve the edge detection, especially for moving objects.

References

Ackerman, S. J., E. P. Steinberg, R. N. Bryan, M. BenDebba, and D. M. Long (1997). Trends in diagnostic imaging for low back pain: has MR imaging been a substitute or add-on? *Radiation 203*(2), 533–538.

Adams, M. A. (1999). Biomechanics of the intervertebral disc, vertebra, and ligaments. *Lumbar Segmental Instability*, 1–13.

Aguado, A. S., E. Montiel, and M. S. Nixon (1998). On the intimate relationship between the principle of duality and the Hough transform. *Proceedings of the Royal Society-A 456*, 503–526.

Aguado, A. S., E. Montiel, and M. S. Nixon (2000). Bias error analysis of the generalised Hough transform. *Journal of Mathematical Imaging and Vision 12*(1), 25–42.

Aguado, A. S., M. S. Nixon, and M. E. Montiel (1998). Parameterizing arbitrary shapes via Fourier descriptors for evidence-gathering extraction. *Computer Vision and Image Understanding 69*(2), 202–221.

Allan, D. B. and G. Waddell (1989). An historical perspective on low back pain and disability. *Acta Orthop Scand 60*, 1–23.

Allen, R., C. Simonis, and A. C. Breen (1992). Spine kinematic analysis using digital videofluoroscopy and image processing. In *IEE Colloquium on Image Processing for the Disabled*, Edinburgh.

Allen, R., Y. Zheng, and M. S. Nixon (2001). Measurement of the kinematics of the lumbar spine *in vivo*. In *International Conference on Biomechanics combined with Annual Scientific Meeting of Taiwanese Society of Biomechanics*, Taiwan, pp. 25.

American Medical Association (1987). Guides to the evaluation of permanent impairment. Technical report, American Medical Association, Chicago. 52-59.

Ballard, D. H. (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition 13*(2), 111–122.

Beasley, D., D. R. Bull, and R. R. Martin (1993a). An overview of genetic algorithms: Part 1, fundamentals. *University Computing 15*(2), 58–69.

Beasley, D., D. R. Bull, and R. R. Martin (1993b). An overview of genetic algorithms: Part 2, research topics. *University Computing 15*(4), 170–181.

Beutel, J., H. L. Kundel, and R. L. Van Metter (2002). *Handbook of Medical Imaging: Physics and Psychophysics*, Volume 1. SPIE.

Bogduk, N. (1997). *Clinical Anatomy of the Lumbar Spine and Sacrum* (3rd ed.). London: Churchill Livingstone.

Bratton, R. L. (1999). Assessment and management of acute low back pain. *American Family Physician 60*, 2299–2308.

Breen, A. C., R. Allen, and A. Morris (1987). A computer/X-ray method for measuring spinal segmental movement: a feasibility study. *Transactions of the Pacific Consortium for Chiropractic Research E3.1-3.7.*

Breen, A. C., R. Allen, and A. Morris (1988). An image processing method for spine kinematics-preliminary studies. *Clinical Biomechanics 3*, 5–10.

Breen, A. C., R. Allen, and A. Morris (1989). A digital videofluoroscopic technique for spine kinematics. *Journal of Medical Engineering & Technology 13*(1/2), 109–113.

Breen, A. C. and R. Allen (1993). Image presentation for spinal kinematic analysis using digital videofluoroscopy. *Third International Conference on Image Processing and Its Applications.*

Breen, A. C., R. Brydges, H. Nunn, J. Kause, and R. Allen (1993). Quantitative analysis of lumbar spine intersegmental motion. *European Journal of Physical Medicine & Rehabilitation 3*(5), 182–190.

Breen, A. C. (1991). *The measurement of the kinematics of the human spine using videofluoroscopy and image processing.* Ph. D. thesis, University of Southampton, Southampton, U. K.

Brejl, M. and M. Sonka (2000a). Automated initialization automated design of border detection criteria in edge-based segmentation. In *The 4th IEEE Southwest Symposium on Image Analysis and Interpretation*, Austin, USA, pp. 26–30.

Brejl, M. and M. Sonka (2000b). Object localization and border detection criteria design in edge-based image segmentation: Automated learning from examples. *IEEE Transactions on Medical Imaging 19*(10), 973–985.

Brolin, K. (2002). Influence of ligamentous injuries on the upper cervical spinal kinematics. In *Proceedings of the 4th World Congress of Biomechanics*, pp. 1. [Electronic version].

Brown, R. H., A. H. Burstein, C. L. Nash, and C. C. Schock (1976). Spinal analysis using a three-dimensional radiographic technique. *Journal of Biomechanics 9*, 355–365.

Burden, R. L. and J. D. Faires (1993). *Numerical Analysis* (5th ed.). Boston: PWS Publishing Company.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence PAMI-8*(6), 679–698.

Cardan, C. and R. Allen (2000). Measurement of spine motion for diagnosis of mechanical problems. *Computer Simulation & Modeling in Medicine 1*(1), 15–19.

Case, K., D. C. Xiao, B. S. Acar, and J. M. Porter (1999). Computer aided modeling of the human spine. *Proc. Inst. Mech. Engrs. 213*(part B), 83–86.

Challis, J. H. (2001). Estimation of the finite center of rotation in planar movements. *Medical Engineering & Physics 23*, 227–233.

Chetverikov, D. and J. Verestóy (1999). Feature point tracking for incomplete trajectories. *Computing 62*(4), 321–338.

Cholewicki, J. and S. M. Mcgill (1991). Method for measuring vertebral kinematics from videofluoroscopy. *Clinical Biomechanics 6*, 73–78.

Cho, Z. H., J. P. Jones, and M. Singh (1993). *Foundations of Medical Imaging.* New York: John Wiley & Sons, Inc.

Clinical Standards Advisory Group (1994). Epidemiology review: The epidemiology and cost of back pain. Technical report, London: HMSO.

Davies, C. R. (1990). *Machine Vision: Theory, Algorithms, Practicalities.* London: Academic Press.

Deans, S. R. (1981). Hough transform from the Radon transform. *IEEE Transactions on Pattern Recognition & Machine Intelligence PAMI-3*(2), 185–188.

Department of Health (1999). The prevalence of back pain in Great Britain in 1998. Bulletin 1999/18, Department of Health, U. K.

de Boor, C. (1978). *A Practical Guide to Splines.* New York, Heidelberg, Berlin: Springer-Verlag.

De Jong, K. (1975). *An analysis of the behaviour of a class of genetic adaptive systems.* Ph. D. thesis, University of Michigan.

Dimnet, J., J. P. Carret, G. Gonon, and L. P. Fischer (1976). A technique for joint center analysis using a stored program calculator. *Journal of Biomechanics 9*, 771–778.

Dimnet, J., A. Pasquet, M. H. Krag, and M. M. Panjabi (1982). Cervical spine motion in the sagittal plane: Kinematic and geometric parameters. *Journal of Biomechanics 15*(12), 959–969.

D'Orazio, B. P. (1999). *Low Back Pain Handbook.* Boston: Butterworth and Heinemann.

Duda, R. O. and E. Hart (1972). Use of the Hough transform to detect lines and curves in pictures. *Communications of the ACM 15*(1), 11–15.

Dupuis, P. R., K. Yong-Hing, J. D. Cassidy, and W. H. Kirkaldy-Willis (1985). Radiologic diagnosis of degenerative lumbar spinal instability. *Spine 10*(3), 262–276.

Dvořák, J., M. M. Panjabi, D. G. Chang, R. Thiler, and D. Grob (1991). Functional radiographic diagnosis of the lumbar spine. *Spine 16*(5), 562–571.

Eisenstein, S. M. (1999). "Instability" and low back pain: A way out of the semantic maze. *Lumbar Segmental Instability*, 39–44.

Farrokhi, S., K. Kulig, J. Burnfield, and C. Powers (2002). *In vivo* assessment of the instantaneous axes of rotation of the lumbar spine in symptomatic and asymptomatic subjects. In *Proceedings of the 4th World Congress of Biomechanics*, pp. 1. [Electronic version].

Frobin, W., P. Brinckmann, M. Biggemann, M. Tillotson, and K. Burton (1997). Precision measurement of disc height, vertebral height and sagittal plane displacement from lateral radiographic views of the lumbar spine. *Clinical Biomechanics*, 1–64. Supplement.

Frobin, W., P. Brinckmann, G. Leivseth, M. Biggermann, and O. Reikeras (1996). Precision measurement of segmental motion from flexion-extension radiographs of the lumbar spine. *Clinical Biomechanics 11*, 457–465.

Gatton, M. L. and M. J. Pearcy (1999). Kinematics and movement sequencing during flexion of the lumbar spine. *Clinical Biomechanics 14*, 376–383.

Gertzbein, S. D., R. Holtby, A. Kapasouri, and B. Cruickshank (1984). Determination of a locus of instantaneous centers of rotation of the lumbar disc by Moiré fringes: A new technique. *Spine 9*(4), 409–413.

Gertzbein, S. D., J. Seligman, R. Holtby, K. H. Chan, A. Kapasouri, M. Tile, and B. Cruickshank (1985). Centrode patterns on segmental instability in degenerative disc disease. *Spine 10*(3), 257–261.

Gertzbein, S. D., N. Wolfson, and G. King (1988). The diagnosis of segmental instability *in vivo* by centrode length. In *The Annual Meeting of the International Society for the Study of the Lumbar Spine*, Miami, Florida. [Presentation].

Gill, K., M. Krag, G. Johnson, L. Haugh, and M. Pope (1988). Repeatability of four clinical methods for assessment of lumbar spine motion. *Spine 13*, 50–53.

Goldberg, D. E. and K. Deb (1991). A comparative analysis of selection schemes used in genetic algorithms. In G. J. E. Rawlins (Ed.), *Foundations of Genetic Algorithms*, Morgan Kaufmann, pp. 69–93.

Goldberg, D. E. (1997). *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley.

Goulermas, J. Y. and P. Liatsis (1998). Genetically fine-tuning the Hough transform feature space, for the detection of circular objects. *Image and Vision Computing 16*, 615–625.

Grant, M. G., M. S. Nixon, and P. H. Lewis (2002). Extracting moving shapes by evidence gathering. *Pattern Recognition 35*, 1099–1114.

Gregersen, G. G. and D. B. Lucas (1976). An *in vivo* study of the axial rotation of the human thoracolumbar spine. *Journal of Bone Joint Surgery (American) 49*, 247–262.

Grimson, W. E. L. and D. P. Huttenlocher (1990). On the sensitivity of the Hough transform for object recognition. *IEEE Transactions on Pattern Recognition & Machine Intelligence 12*(3), 255–274.

Hamadeh, A and P. Cinquin (1997). Kinematics study of lumbar spine using functional radiographies and 3D/2D registration. In J. Troccaz, E. Grimson, and R. Mösges (Eds.), *Lecture Notes in Computer Science, CVRMed-MRCAS'97*, Volume 1205, pp. 109–118. Springer.

Harrison, D. D., R. Cailliet, T. J. Janik, S. J. Troyanovich, D. E. Harrison, and B. Holland (1998). Elliptical modeling of the sagittal lumbar lordosis and segmental rotation angles as a method to discriminate between normal and low back pain subjects. *Journal of Spinal Disorder 11*(5), 430–439.

Heath, M. D., S. Sarkar, T. Sanocki, and W. Bowyer (1997). A robust visual method for assessing the relative performance of edge-detection algorithms. *IEEE Transactions on Pattern Analysis & Machine Intelligence 19*(12), 1338–1359.

Heath, M., S. Sarkar, T. Sanocki, and K. Bowyer (1998). Comparison of edge detectors: A methodology and initial study. *Computer Vision and Image Understanding 69*(1), 38–54.

Hough, P. V. C. (1962). A method and means for recognizing complex patterns. U.S. Patent 3,069,654.

Hunt, D. J., L. W. Nolte, and A. R. Reibman (1990). Hough transform and signal detection theory performance for images with additive noise. *Computer Vision, Graphics, and Image Processing 52*, 386–401.

Illingworth, J. and J. Kitter (1988). A survey of the Hough transform. *Computer Vision, Graphics, and Image Understanding 44*, 87–116.

Illingworth, J. and J. Kittler (1987). The adaptive Hough transform. *IEEE Transactions on Pattern Analysis & Machine Intelligence PAMI-9*(5), 690–698.

James, G. (1992). *Modern Engineering Mathematics* (1st ed.). Addison-Wesley.

Kälviäinen, H., P. Hirvonen, L. Xu, and E. Oja (1994). Comparisons of probabilistic and non-probabilistic Hough transforms. In *Computer Vision* (Jan-Olof Eklundh ed.), Volume 2, Stockholm, Sweden.

Kanayama, M., K. Abumi, K. Kaneda, S. Tadano, and T. Ukai (1996). Phase lag of the intersegmental motion in flexion-extension of the lumbar and lumbarsacral spine. *Spine 21*, 1416–1422.

Kimme, C., D. Ballard, and J. Sklansky (1975). Finding circles by an array of accumulators. *Communications of ACM 18*, 120–122.

Kiryati, N. and A. M. Bruckstein (1991). Antialiasing the Hough transform. *CVGIP: Graphical Models & Image Processing 53*(3), 213–222.

Knutson, F. (1944). The instability associated with disc degeneration in the lumbar spine. *Acta Radiology 25*, 593–609.

Kondracki, M. (1998). Kinematic determination of the functional integrity of the lumbar spine *in vivo*. In *3rd Interdisciplinary World Congress of Low Back and Pelvic Pain*, Vienna, Australia, pp. 32–37.

Kondracki, M. (2001). *Clinical application of digitised videofluoroscopy in the lumbar spine.* Ph. D. thesis, University of Southampton, Southampton, U. K.

Kovesi, P. (1999). Image features from phase congruency. *Videre: Journal of Computer Vision Research 1*(3), 2–26.

Kreyszig, E. (1993). *Advanced Engineering Mathematics* (7th ed.). Wayne Anderson.

Kuhl, F. P. and C. R. Giardina (1982). Elliptic Fourier features of a closed contour. *Computer Graphics & Image Processing 18*, 236–258.

Kulig, K., C. Powers, R. Landel, H. Chen, M. Fredricson, M. Guillet, and K. Butts (2002). Segmental mobility of lumbar spine during a prone press-up maneuver in symptomatic and asymptomatic subjects: assessment using dynamic MRI. In *Proceedings of the 4th World Congress of Biomechanics*, pp. 1. [Electronic version].

Lam, W. C. Y., L. T. S. Lam, K. S. Y. Yuen, and D. N. K. Leung (1994). An analysis on quantizing the Hough space. *Pattern Recognition Letters 15*, 1127–1135.

Lappas, P., J. N. Carter, and R. I. Damper (2002). Robust evidence-based object tracking. *Pattern Recognition Letters 23*, 253–260.

Leavers, V. F. (1992). The dynamic generalized Hough transform: its relationship to the probabilistic Hough transform and application to the concurrent detection of circles and ellipses. *CVGIP: Image Understanding 56*(3), 381–398.

Leavers, V. F. (1993). Survey: Which Hough transform? *CVGIP: Image Understanding 58*(2), 250–264.

Lee, Y. and Y. Chen (2000). Regressionally determined vertebral inclination angles of the lumbar spine in static lifts. *Clinical Biomechanics 15*, 672–677.

Li, H., M. A. Lavin, and R. J. Lemaster (1985). Fast Hough transform. In *Proc. 3rd Workshop Computer Vision: Representation and Control*, Bellair,MI, pp. 75–83.

Maintz, J. B. and M. A. Viergever (1998). A survey of medical image registration. *Medical Image Analysis 2*(1), 1–37.

Maitre, H. (1986). Contribution to the prediction of performances of the Hough transform. *IEEE Transactions on Pattern Analysis & Machine Intelligence PAMI-8*(5), 669–674.

Marras, W. S., S. A. Ferguson, P. Gupta, S. Bose, J. Y. Kimand M Parnianpour, and R. R. Crowell (1999). The quantification of low back disorder using motion measures, methodology and validation. *Spine 24*(20), 2091–2100.

Merlin, P. M. and D. J. Farber (1975). A parallel mechanism for detecting curves in pictures. *IEEE Transactions on Computers C-24*, 96–98.

Mimura, M. (1990). Rotational instability of the lumbar spine - a three-dimensional motion study using bi-plane X-ray analysis system. *Japanese Orthopaedics Association 64*, 546–559.

Miyasaka, K., K. Ohmori, K. Suzuki, and H. Inoue (2000). Radiographic analysis of lumbar motion in relation to lumbosacral stability. *Spine 25*(6), 732–737.

Moffett, J. A. K. and G. Richardson (1995). *Lumbar spine disorders: Current concepts*, Volume 1, pp. 225–235. World Scientific: Singapore.

Morgan, F. P. and T. King (1957). Primary instability of lumbar vertebrae as a common cause of low-back pain. *Journal of Bone Joint Surgery 39*, 6–22.

Morrone, M. C. and R. A. Owens (1987). Feature detection from local energy. *Pattern Recognition Letters 6*, 303–313.

Morrone, M. C., J. R. Ross, D. C. Burr, and R. A. Owens (1986). Mach bands are phase dependent. *Nature 324*(6094), 250–253.

Muammar, H. K. and M. Nixon (1991). Tristage Hough transform for multiple ellipse extraction. *IEE Proceeding E 138*(1), 27–34.

Muggleton, J. M. and R. Allen (1997). Automatic location of vertebrae in digitized videofluoroscopic images of the lumbar spine. *Medical Engineering & Physics 19*(1), 77–89.

Muggleton, J. M. and R. Allen (1998). Insight into the measurement of vertebral translation in the sagittal plane. *Medical Engineering & Physics 20*, 21–32.

Nash, J. M., J. N. Carter, and M. S. Nixon (1997). Dynamic feature extraction via the Velocity Hough transform. *Pattern Recognition Letters 18*, 1035–1047.

Nash, J. M. (1999). *Evidence Gathering for Dynamic Feature Extraction.* Ph. D. thesis, University of Southampton, Southampton, U. K.

Nattress, C. L., J. E. Nitschke, P. B. Disler, M. J. Chou, and K. T. Ooi (1999). Lumbar spine range of motion as a measure of physical and functional impairment: an investigation of validity. *Clinical Rehabilitation 13*, 211–218.

Nixon, M. S. and A. S. Aguado (2002). *Feature Extraction & Image Processing* (1st ed.). UK: Butterworth-Heinemann.

North American Spine Society (latest found on March 6, 2001). Facts about back pain. http://www.spine.org/bthw/FactsAboutBackPain.htm.

Ogston, N. G., G. J. King, S. D. Gertzbein, M. Tile, A. Kapasouri, and J. D. Rubenstein (1986). Centrode patterns in the lumbar spine - baseline studies in normal subjects. *Spine 11*(6), 591–595.

Okawa, A., K. Shinomiya, H. Komori, T. Muneta, Y. Arai, and O. Nakai (1998). Dynamic motion study of the whole lumbar spine by videofluoroscopy. *Spine 23*(16), 1743–1749.

Olson, C. F. (1999). Constrained Hough transforms for curve detection. *Computer Vision and Image Understanding 73*(3), 329–345.

Oppenheim, A. V. and J. S. Lim (1981). The importance of phase in signals. In *The Proceedings of The IEEE*, Volume 69, pp. 529–541.

Palmer, P. L., J. Kittler, and M. Petrou (1997). An optimizing line finder using a Hough transform algorithm. *Computer Vision & Image Understanding 67*(1), 1–23.

Panjabi, M. M., G. B. J. Anderssen, and L. Jorneus (1986). *In vivo* measurements of spinal column vibrations. *Journal of Bone Joint Surgery (American) 68*, 695–702.

Panjabi, M. M., C. Lydon, A. Vasavada, D. Greter, J. J. Crisco, and J. Dvořák (1994). On the understanding of clinical instability. *Spine 19*(23), 2642–2650.

Panjabi, M. M. (1979). Centers and angles of rotation of body joints: A study of errors and optimisation. *Journal of Biomechanics 12*, 911–920.

Panjabi, M., D. Chang, and J. Dvořák (1992). An analysis of errors in kinematic parameters associated with *in vivo* functional radiographs. *Spine 17*(2), 200–205.

Panjabi, M. and A. A. White (1971). A mathematical approach for three-dimensional analysis of the mechanics of the spine. *Journal of Biomechanics 4*, 203–211.

Paris, S. V. (1985). Physical signs of instability. *Spine 10*, 277–279.

Pearcy, M. J. and N. Bogduk (1988). Instantaneous axes of rotation of the lumbar intervertebral joints. *Spine 13*(9), 1033–1041.

Pearcy, M., I. Portek, and J Shepherd (1984). Three-dimensional X-ray analysis of normal movement in the lumbar spine. *Spine 9*, 294–297.

Pearcy, M. (1986). Measurement of back and spinal mobility. *Clinical Biomechanics 1*, 44–51.

Pope, M. H., M. Svensson, H. Broman, and G. B. J. Andersson (1986). Mounting of transducers in measurement of segmental instability motion of the spine. *Journal of Biomechanics 19*, 675–677.

Pope, M. H., D. G. Wilder, and L. Jorneus (1987). The response of the seated human to sinusoidal vibration and impact. *Journal of Biomechanical Engineering 109*, 279–284.

Pope, M., M. Ogon, and A. Okawa (1999). Biomechanical measurements. *Lumbar Segmental Instability*, 27–37.

Porterfield, J. A. and C. Derosa (1998). *Mechanical Low Back Pain: Perspective in Functional Anatomy* (2nd ed.). W. B. Saunders Company.

Potter, M. (1997). *The Design and Analysis of a Computational Model of Cooperative Coevolution.* Ph. D. thesis, George Mason University, Fairfax, Virginia, USA.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in C* (2nd ed.). Press Syndicate of the University of Cambridge.

Princen, J., J. Illingworth, and J. Kittler (1992). A formal definition of the Hough transform: Properties and relationship. *Journal of Mathematics & Imaging Vision 1*, 153–168.

Rangarajan, K. and M. Shah (1991). Establishing motion correspondence. *Computer Vision Graphics Image Processing 54*(1), 56–73.

Russell, G. G., V. J. Raso, and J. Mcivor (1990). A comparison of four computerized methods for measuring vertebral rotation. *Spine 15*(1), 24–27.

Saito, Y., H. Aradate, H. Miyazaki, Y. Kudo, K. Tsujita, N. Shimadu, and Y. Sawanaga (2002). Development and evaluation of a real-time three dimensional CT (4D-CT) scanner. In *Proceedings of SPIE*, Volume 4982, San Diego, pp. 801–808.

Samal, A. and J. Edwards (1997). Generalized Hough transform for natural shapes. *Pattern Recognition Letters 18*, 473–480.

Seith, I. K. and R. C. Jain (1987). Finding trajectories of feature points in a monocular image sequence. *IEEE Transactions on Pattern Analysis & Machine Intelligence 9*(1), 56–73.

Ser, P. K., S. T. C. Clifford, and W. C. Siu (1999). Genetic algorithm for the extraction of nonanalytic objects from multiple dimensional parameter space. *Computer Vision & Image Understanding 73*(1), 1–13.

Simonis, C., R. Allen, and R. Cloke (1993). Spatial analysis of the movement of the spine: Application of parallel computing in the field of spine biomechanics. *Transputer Applications 1*, 35–38.

Simonis, C. and R. Allen (1993). Calculation of planar spine kinematic parameters using videofluoroscopic images and parallel processing. In *Proceedings of 15th IEEE Engineering in Medicine & Biology Society*, Part 3, San Diego, California, USA, pp. 1087–1088.

Simonis, C. (1994). *Parallel calculation and analysis of spine kinematics parameters using videofluoroscopy and image processing.* Ph. D. thesis, University of Southampton, Southampton.

Sklansky, J. (1978). On the Hough technique for curve detection. *IEEE Transactions on Computers C-27*(10), 923–926.

Smyth, P. P., C. J. Taylor, and J. E. Adams (1997). Automatic measurement of vertebral shape using active shape models. *Image & Vision Computing 15*, 578–581.

Soffer, M. and N. Kiryati (1998). Guaranteed convergence of the Hough transform. *Computer Vision & Image Understanding 69*(2), 119–134.

Sonka, M., V. Hlavac, and R. Boyle (1999). *Image Processing, Analysis and Machine Vision* (2nd ed.). U. S.: PWS Publishing.

Steffen, T., R. K. Rubin, H. G. Baramki, J. Antoniou, D. Marchesi, and M. Aebi (1997). A new technique for measuring lumbar segmental motion *in vivo* method, accuracy and preliminary results. *Spine 22*(2), 156–166.

Stokes, I. A. F., T. M. Bevins, and R. A. Lunn (1987). Back surface curvature and measurement of lumbar spine motion. *Spine 12*(4), 355–361.

Stokes, I. A. F. and J. W. Frymoyer (1977). Relationship between movements of vertebra and adjacent skin markers and sections of the back. Annual report, The Oxford Orthopaedics Engineering Center.

Stokes, I. A. F. and J. W. Frymoyer (1987). Segmental motion and instability. *Spine 12*(7), 688–691.

Stokes, I. A. F., D. G. Wilder, J. W. Frymoyer, and M. H. Pope (1981). Assessment of patients with low back pain by bi-planar radiographic measurements of intervertebral motion. *Spine 6*, 233–240.

Suh, C. H. (1974). The fundamentals of computer aided X-ray analysis of the spine. *Journal of Biomechanics 7*, 161–169.

Szpalski, M., R. Gunzburg, and M. H. Pope (1999). *Lumbar Segmental Instability* (1st ed.). Philadelphia: Lippincott Williams & Wilkins.

Takayanagi, K., K. Takahashi, M. Yamagata, H. Moriya, H. Kitahara, and T. Tamaki (2001). Using cineradiography for continuous dynamic motion analysis of the lumbar spine. *Spine 26*(17), 1858–1865.

Tezmol, A., H. Sari-Sarraf, S. Mitra, R. Long, and A. Gururajan (2002). Customized Hough transform for robust segmentation of cervical vertebrae from X-ray images. In *Fifth IEEE Southwest Symposium on Image Analysis and Interpretation*, Santa Fe, New Mexico, U. S., pp. 224–228.

Tsuji, S. and F. Matsumoto (1978). Detecting of ellipses by a modified Hough transform. *IEEE Transactions on Computers C-27*(8), 777–781.

Tulder, M. W., W. J. J. Assendelft, B. W. Koes, and L. M. Bouter (1997). Spinal radiographic findings and non-specific low back pain. *Spine 22*(4), 427–434.

Veen, T. M. V. and F. C. A. Groen (1981). Discretization errors in the Hough transform. *Pattern Recognition 14*(1-6), 137–145.

Waddell, G. (1996). Keynote address for primary care forum low back pain: A twenty century health care enigma. *Spine 21*(24), 2820–2825.

Waddell, G. (1999). *The Back Pain Revolution.* United Kingdom: Churchill Livingstone.

Wang, H. L. and A. P. Reeves (1990). Three dimensional generalized Hough transform for object identification. *J. Soc. Photo-Opt. Instrum. Eng. 1192*(1), 363–374.

Warren-Forward, H. M., M. J. Haddaway, D. H. Temperton, and I. W. McCall (1998). Dose-area product readings for fluoroscopic and plain film examinations, including an analysis of the source of variation for barium enema examinations. *British Journal of Radiology 71*, 961–967.

Weiler, P. J., G. J. King, and S. D. Gertzbein (1990). Analysis of sagittal plane instability of the lumbar spine *in vivo*. *Spine 15*(12), 1300–1305.

White, A. A., M. Bernhardt, and M. M. Panjabi (1999). Clinical biomechanics and lumbar spinal instability. *Lumbar Segmental Instability*, 15–25.

White, A. A. and M. M. Panjabi (1990). *Clinical Biomechanics of the Spine* (2nd ed.), Volume 1. Philadelphia: Lippincott-Raven.

Wilder, D. G., D. Selgison, J. W. Frymoyer, and M. H. Pope (1980). Objective measurement of L4-5 instability: A case report. *Spine 5*, 56–58.

Wynsberghe, D. V., C. R. Noback, and R. Carola (1995). *Human Anatomy and Physiology* (3rd ed.). U.S.A.: McGraw-Hill, Inc.

Xu, L., E. Oja, and P. Kultanen (1990). A new curve detection method: Randomized Hough transform (RHT). *Pattern Recognition Letters 11*(5), 331–338.

Xu, L. and E. Oja (1993). Randomized Hough transform (RHT): Basic mechanisms, algorithms, and computational complexities. *CVGIP: Image Understanding 57*(2), 131–154.

Zheng, Y., M. S. Nixon, and R. Allen (2000a). Finding lumbar vertebrae by evidence gathering. In *The 6th Annual Scientific Conference of the Institute of Physics and Engineering in Medicine (IPEM2000)*, Southampton, pp. 95–96.

Zheng, Y., M. S. Nixon, and R. Allen (2000b). Lumbar spine location in fluoroscopic images by evidence gathering. In *Proceedings of Medical Image Understanding and Analysis (MIUA2000)*, London, pp. 45–48.

Zheng, Y., M. S. Nixon, and R. Allen (2001a). Automatic lumbar vertebrae registration using the hough transform in digital videofluoroscopy. In *Physica Medica-European journal of medical physics*, Volume 17, Belfast, pp. 172–173. The 7th Annual Scientific Conference of the Institute of Physics and Engineering in Medicine (IPEM2001).

Zheng, Y., M. S. Nixon, and R. Allen (2001b). Automatic lumbar vertebrae segmentation in fluoroscopic images via optimised concurrent hough transform. In *23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Istanbul, Turkey. IEEE.

Zheng, Y., M. S. Nixon, and R. Allen (2002a). Lumbar spine motion analysis via automatic segmentation. In *The Fourth World Congress of Biomechanics*, Calgary, Canada.

Zheng, Y., M. S. Nixon, and R. Allen (2002b). Lumbar spine visualisation based on kinematic analysis from videofluoroscopic imaging. *Medical Engineering & Physics 25*(3), 171–179.

Zheng, Y., M. S. Nixon, and R. Allen (2003). Automatic segmentation of lumbar vertebrae in digital videofluoroscopic images. *IEEE Transactions of Medical Imaging*. in press.

Spline Theory

Cubic splines are piecewise cubic polynomials which can be fitted to a series of (x_i, y_i) data points. A full cubic spline with *knots* at x_i (where the piecewise portions join) exactly interpolates the data points. The piecewise portions are defined so that at the knots the function and its first two derivatives are continuous.

Given a function f defined on $[a, b]$ and a set of numbers, called nodes, $a = x_0 < x_1 < \ldots < x_n = b$, according to Burden and Faires (1993), a cubic spline interpolant S for f is a function that satisfies the following conditions:

1. S is a cubic polynomial, denoted as S_j, on the subinterval $[x_j, x_{j+1}]$ for each $j = 0, 1, ..., n-1$;
2. $S(x_j) = f(x_j)$ for each $j = 0, 1, ..., n$;
3. $S_{j+1}(x_{j+1}) = S_j(x_{j+1})$ for each $j = 0, 1, ..., n-2$;
4. $S'_{j+1}(x_{j+1}) = S'_j(x_{j+1})$ for each $j = 0, 1, ..., n-2$;
5. $S''_{j+1}(x_{j+1}) = S''_j(x_{j+1})$ for each $j = 0, 1, ..., n-2$;
6. one of the following set of boundary conditions should be satisfied:
   (a) $S''(x_0) = S''(x_n) = 0$ (**free or natural boundary**)
   (b) $S'(x_0) = f'(x_0)$ and $S'(x_n) = f'(x_n)$ (**clamped boundary**)

The above conditions (2-5) are to ensure that the function and its first and second derivatives respectively are continuous at the knots while the final restriction above means that cubic spline will be a unique solution for the given data sets (proof of uniqueness and existence can be found in Kreyszig (1993)).

When the free boundary condition occurs, the spline is called a natural spline, and its graph approximates to the shape that a long flexible rod is forced to pass through each of the data points. In general, clamped boundary conditions lead to more accurate approximations since they have more information about the function. However, it is slightly difficult to use as it is necessary to have either the values of the derivative at the endpoints or an accurate approximation of those values.

Given a cubic spline interpolant S, the conditions in the definition are applied to the cubic polynomials:

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3; \tag{1}$$

for each $j = 0, 1, ..., n - 1$.

From condition (2), we can have $S_j(x_j) = a_j = f(x_j)$.

If condition (3) is applied, we have:

$a_{j+1} = S_{j+1}(x_{j+1}) = S_j(x_{j+1}) = a_j + b_j(x_{j+1} - x_j) + c_j(x_{j+1} - x_j)^2 + d_j(x_{j+1} - x_j)^3$.

If we define $a_n = f(x_n)$, equally we have

$$a_{j+1} = a_j + b_j h_j + c_j h_j^2 + d_j h_j^3 \tag{2}$$

for each $j = 0, 1, ..., n-1$. Here a simpler notation $h_j = x_{j+1} - x_j$ for each $j = 0, 1, ..., n-1$ is introduced for convenience.

In a similar manner, define $b_n = S_j'(x_n)$ and from condition (4) we will have

$$b_{j+1} = b_j + 2c_j h_j + 3d_j h_j^2 \tag{3}$$

for each $j = 0, 1, ..., n - 1$.

We can also obtain coefficients c_j by defining $c_n = S''(x_n)/2$ and applying condition (5). In this case,

$$c_{j+1} = c_j + 3d_j h_j \tag{4}$$

for each $j = 0, 1, ..., n - 1$.

Solving for d_j in Equation (4) and substituting this value into Equation (2) and (3) gives two new equations

$$a_{j+1} = a_j + b_j h_j + \frac{h_j^2}{3}(2c_j + c_{j+1}) \tag{5}$$

$$b_{j+1} = b_j + h_j(c_j + c_{j+1}) \tag{6}$$

for each $j = 0, 1, ..., n-1$.

The final relationship involving the coefficients is derived from Equation (6) by substituting the values b_j and b_{j-1} obtained by solving Equation (5) and can result in the form of linear system of equations

$$h_{j-1}c_{j-1} + 2(h_{j-1} + h_j)c_j + h_j c_{j+1} = \frac{3}{h_j}(a_{j+1} - a_j) - \frac{3}{h_{j-1}}(a_j - a_{j-1}) \tag{7}$$

for each $j = 1, 2, ..., n-1$.

Note that there are $n-1$ equations involved in the linear system described by Equation (7). Although a_j for each $j = 0, 1, ..., n$ and h_j for each $j = 0, 1, ..., n-1$ are given by the spacing of the nodes x_j for each $j = 0, 1, ..., n$ and the values of f at the nodes, it is still not enough to solve c_j, $j = 0, 1, ..., n$ where are $n+1$ unknown parameters. This is why the boundary conditions are essential to define a cubic spline. Given either of these two conditions, a unique cubic spline interpolant will be defined as the matrix is strictly diagonally dominant.

In application, there are standard programs elsewhere to solve this problem, i.e. a C program is available in *Numerical Recipes in C* (Press et al. 1992).

The cubic spline has the property that it is the interpolating function which minimises the integrated squared second derivative $\int (f''(x))^2 dx$ (for proof, see Kreyszig (1993)). The integrated squared second derivative acts as a roughness penalty, with smooth or slowly varying curves giving small values (for example a straight line gives a value of 0). However the cubic spline has a discontinuous third derivative:

$$f'''(x) = d_j; \text{ if } x_j < x < x_{j+1} \tag{8}$$

for each $j = 0, 1, ..., n-1$.