Bournemouth University

AUTOMATED SHAPE ANALYSIS AND VISUALIZATION OF THE HUMAN BACK

BRIGHT OSEI TWUMASI

A thesis submitted in partial fulfilment of the requirements of Bournemouth University for the degree of Doctor of Philosophy.

July 2007

Bournemouth University

PAGE

NUMBERING

AS ORIGINAL

Abstract

Spinal and back deformities can lead to pain and discomfort, disrupting productivity, and may require prolonged treatment. The conventional method of assessing and monitoring the deformity using radiographs has known radiation hazards. An alternative approach for monitoring the deformity is to base the assessment on the shape of back surface. Though three-dimensional data acquisition methods exist, techniques to extract relevant information for clinical use have not been widely developed.

This thesis presents the content and progression of research into automated analysis and visualization of three-dimensional laser scans of the human back. Using mathematical shape analysis, methods have been developed to compute stable curvature of the back surface and to detect the anatomic landmarks from the curvature maps. Compared with manual palpation, the landmarks have been detected to within accuracy of 1.15mm and precision of 0.8mm. Based on the detected spinous process landmarks, the back midline which is the closest surface approximation of the spine, has been derived using constrained polynomial fitting and statistical techniques. Three-dimensional geometric measurements based on the midline were then computed to quantify the deformity. Visualization plays a crucial role in back shape analysis since it enables the exploration of back deformities without the need for physical manipulation of the subject. In the third phase, various visualization techniques have been developed, namely, continuous and discrete colour maps, contour maps and three-dimensional views. In the last phase of the research, a software system has been developed for automating the tasks involved in analysing, visualizing and quantifying of the back shape.

The novel aspects of this research lie in the development of effective noise smoothing methods for stable curvature computation; improved shape analysis and landmark detection algorithm; effective techniques for visualizing the shape of the back; derivation of the back midline using constrained polynomials and computation of three-dimensional surface measurements.

List of Publications

Conferences

Twumasi, B.O. and Claremont, D.J. (2005). Extraction of useful clinical information from 3D laser scans of the surface of the human back. LSI Conference, University of Southampton, 2005, UK.

Twumasi, B.O. & Lefley, M. (2005), Analysis and visualization of the human back shape using surface curvature. In: Nicolaou N. et al (eds) Proceedings of the 4th IEEE EMBSS UK & IR Postgraduate Conference in Biomedical Engineering & Medical Physics, 18-20 July, University of Reading, UK.

Twumasi, B.O., Lefley, M., Crook, S. (2005), Curvature Analysis of the human Back Shape. Academic Biomedical Engineering Research Group Workshop, Bournemouth, UK.

Twumasi, B.O., Lefley, M. and Crook, S. (2004), Automated Analysis of the Human Back Shape. Academic Biomedical Engineering Research Group Workshop, Bournemouth, UK.

Twumasi, B.O. (2005), Automated Analysis of the Human Back Shape. Presented at Annual Workshop of British Computer Society Dorset Branch, 17 February, 2005, Bournemouth.

Awards and Achievements

William James Memorial Biomedical Engineering Student Award (UK&IR)—'In recognition of your outstanding paper and presentation, and for contributions to Biomedical Engineering' July, 2005.

Participation and presentation for Britain's Top Younger Engineers Award, Special Reception and Competition at the House of Commons, UK, December 14, 2004.

Acknowledgements

This research would not have reached a successful completion without the help, support and encouragement of many individuals, to whom I am forever indebted.

First of all, I thank the Almighty God, my Creator and Guide, who gave me the strength and knowledge to execute this work.

I would like to express my deep and sincere gratitude to my supervisory team, Dr Martin Lefley, Prof Denzil Clarement, Prof Jian Zhang and Dr Steve Crook for their support, suggestions, criticisms and contributions throughout the research period. From the beginning to the end of the research, they provided invaluable support and were like family to me. I could go to them any time, sometimes without prior appointment to discuss my problems. I greatly admire their patience, effort and friendliness I enjoyed during my work with them.

I thank my research colleagues at Bournemouth University and Salisbury District Hospital who volunteered for data collection. I am also indebted to Nigel Bryant and Mathias from the Media School for their support during data collection.

I am grateful to Bournemouth University who funded this research, thereby giving me the opportunity to develop my potential.

Finally, I thank all my friends, who shared the hard times with me during the research period, especially Alinco, Marryam, Thelma, Abed and Linda. God bless you all.

This work is dedicated to my Mum, Hannah Appiah, who has always been a source of inspiration to me!

Bright

July, 2007

Table of Contents

List of Figures

List of Tables

List of Acronyms

1D, 2D, 3D	One-, Two- and Three-dimensional Euclidean Space
ADCM	Absolute Deviations about Class Medians
AIS	Adolescent Idiopathic Scoliosis
ASCII	American Standard Code for Information Interchange
ATI	Angle of Trunk Inclination
BSR	Back Surface Rotation
CAD	Computer Aided Design
CCD	Charged Coupled Device
CIE	Commission Internationale de l'Eclairage (International Commission on Illumination)
CIELAB	Perceptually uniform colour model defined by CIE
CIELUV	Perceptually uniform colour model defined by CIE
COM	Component Object Model
CTAS	Crude Trunk Asymmetry Scores
DXF	Drawing Interchange Format or Drawing Exchange Format
EOP	External Occipital Protuberance
ESRI	Environmental Systems Research Institute (developer of ArcGIS Engine)
GIS	Geographic Information Systems
GOF	Goodness of Fit
GUI	Graphical User Interface
GVF	Goodness of Variance Fit
ISIS	Integrated Shape Imaging System
LLD	Leg Length Difference
LUT	Look-up Table
NIAMS	National Institute of Arthritis and Musculoskeletal and Skin Diseases (USA)
PAX	Principal Axis Orientation
PSIS	Posterior Superior Iliac Spine
RGB	Red-Green-Blue Colour Model

RMS	Root Mean Square
RMSE	Root Mean Squared Error
SQL	Structured Query Language
SRS	Scoliosis Research Society (USA)
SSE	Sum of Squared Deviations
STAS	Standardised Trunk Asymmetry Score
TAS	Trunk Asymmetry Score
TXT	Text File Extension

1. Introduction

1.1. Background

Spinal deformity or curvature can lead to pain and discomfort of the individual involved, disrupting productivity; and may require expensive and prolonged treatment. The human spine can exhibit three types of spinal curvature, namely, scoliosis, kyphosis and lordosis. Scoliosis refers to a curvature of the spine in the frontal plane and is associated with rotations of the vertebrae, alterations in the rib cage and often results in corresponding deformity of the trunk surface. Kyphosis refers to the outward curvature of the thoracic spine in the sagittal plane and lordosis refers to the inward curvature of the lumbar spine in the sagittal plane (see Section 2.4). The normal spine always exhibits a small degree of both kyphotic and lordotic curvatures; scoliosis on the other hand, is always abnormal. However, mild non-structural scoliosis may pose no real health threat to the patient and thus may require no treatment other than monitoring for progression. In over 80% of the cases, the cause is unknown, the so-called idiopathic scoliosis (SRS, 2000). The surface manifestations of scoliosis typically include shoulder height difference, rib prominens, scapula asymmetry, pelvic asymmetry, a skin fold on the side of the concavity and twisted trunk (see Section 2.4.1.3). The surface deformity is usually more important to the patient and family, and is what frequently motivates them to seek treatment (Mahood et al, 1995).

Although spinal deformities such as scoliosis are generally thought to be associated with back pain, the strength of this associated is not well understood. Many studies have reported equal incidence of back pain and mortality in the general population and in people with spinal deformities (Lawrence, 1969; Branthwaite, 1986; Ettinger et al, 1992; Virta and Ronnemaa, 1993; Cordover, 1997; Mayo, 1994). On the other hand, other studies have found higher incidence of back pain among subjects with structural and postural deformities than in the general population (Ross et al, 1987; Bowen, 1995; Cockkerill et al, 2000). In a resent study (Weinstein et al, 2003), a 50-year follow-up of 117 patients diagnosed with scoliosis in adolescence revealed that the incidence of back pain in untreated scoliotic patients was twice that of 62 normal volunteers of similar age. Similar results were reported by Mayo et al (1994) in a 10-year follow-up study involving 1,476 scoliosis subjects compared with 1,755 volunteers

without scoliosis. These studies suggest that in the long term, back pain is responsible for a considerable amount of disability in scoliosis subjects.

The costs associated with back pain in general are huge. In the UK, the NHS expenditure has been estimated to be £265—£383 million, comprising 0.7—0.9% of the total expenditure in 1995 (Klaber Moffett et al, 1995). In another study, the direct health care cost of back pain in 1998 was estimated to be £1632 million of which about 35% relates to services provided in the private sector and thus is most likely paid for directly by patients and their families (Maniadakisa and Gray, 2000). This study also estimated the costs of informal care and the production losses related to back pain, which totalled £10,668 million. Back pain is thus one of the most costly conditions and research that help identify, prevent or treat the condition is of utmost importance to society.

1.1.1. Screening for Spinal Deformities

In common with many conditions, if a spinal deformity is detected in its early stages, there is likelihood that an effective preventive treatment may be administered. The higher incidence of scoliosis in the adolescent population has prompted several screening studies to be undertaken in many countries (Viviani et al, 1984; Burwell, 1990; Bremberg, 1986; Dickson, 1980; Pin 1985, Willner, 1984a, 1984b; 1999; Morais, 1995; Grossman, 1995). The purpose of these screening programs is to detect the deformity early enough to allow effective preventive actions to be taken. While some of these studies demonstrate benefits in screening programs, others question the effectiveness of such programs. To date, there is no consensus as to the effectiveness of routine screening for spinal deformities. In 1986, the British Scoliosis Society produced the following recommendations concerning scoliosis screening programs (Burwell, 1989).

I. The Natural History Committee of the British Scoliosis Society is unanimous in the opinion that screening for scoliosis should continue as (a) an epidemiological study, also involving (b) case finding.

II. It should still not be a national policy, at present, to screen children routinely throughout the UK.

III. More data are required.

IV. National screening may be possible after further investigation

Recently, the National Screening Committee (NSC) of the UK advised that "the introduction of screening should not be recommended" (NSC, 2005). Similar recommendations have been made by the U.S. Preventive Services Task Force (USPSTF) which grades its recommendations as (A, B, C, D, or I) and the quality of the overall evidence for a service as (good, fair, poor). The recommendations of the USPSTF regarding scoliosis screening are as follows (USPSTF, 2004):

> The USPSTF did not find good evidence that screening asymptomatic adolescents detects idiopathic scoliosis at an earlier stage than detection without screening. The accuracy of the most common screening test—the forward bending test with or without a scoliometer—in identifying adolescents with idiopathic scoliosis is variable, and there is evidence of poor follow-up of adolescents with idiopathic scoliosis who are identified in community screening programs.
>
> The USPSTF found fair evidence that treatment of idiopathic scoliosis during adolescence leads to health benefits (decreased pain and disability) in only a small proportion of people. Most cases detected through screening will not progress to a clinically significant form of scoliosis. Scoliosis needing aggressive treatment, such as surgery, is likely to be detected without screening.
>
> The USPSTF found fair evidence that treatment of adolescents with idiopathic scoliosis detected through screening leads to moderate harms, including unnecessary brace wear and unnecessary referral for specialty care. As a result, the USPSTF concluded that the harms of screening adolescents for idiopathic scoliosis exceed the potential benefits.

As the above quotations demonstrate, the main arguments against routine screening for spinal deformities can be summarised as (i) lower accuracies of current measurements methods, leading to false detections, unnecessary brace wear and referrals; and (ii) discomfort to patients especially with tactile measurement devices like the Scoliometer (see Section 3.3.1). With the development of more accurate, fast, automated and non-contact measurements systems, it is likely that there will continue to be interest in routine mass screening for spinal deformities.

1.2. Rationale for the Research

The conventional method for diagnosing spinal deformities is to acquire a full length radiographic image of the spine, from which the magnitude of the curvature is measured, usually by the Cobb method (see Section 2.4.1.7). However, there are problems with the conventional approach. First, the use of radiographs carries accumulating risks of radiation exposure over the patient's lifetime, especially if there is the need to monitor the deformity for possible progression over an extended period. In particular, since scoliosis have high incidence in adolescents; continuous exposure of young subjects to radiation risks for the purpose of monitoring is undesirable. Second, scoliosis is a three-dimensional deformity and the Cobb angle alone does not provide complete description of the condition. It is an angle obtained from a two-dimensional radiographic image of the spine which does not take into account vertebral rotation. Three dimensional measurements are therefore required. Moreover, various studies have shown wide variations in measurement of the Cobb angle (up to 10^0) (Goldberg et al, 1988; Morrissy et al, 1990).

An alternative approach for monitoring the deformity is to base the assessment on the shape of the back surface. Since deformity of the spine will almost certainly result in corresponding deformity of the trunk surface, this approach should work well for a large proportion of the population. Surface measurement has been the subject of research in the past few decades. Various imaging modalities have been proposed to address the problem and various surface measurements have been derived. These surface measurements methods have been discussed in Chapter 3. All these methods have been compared to the radiographic 'gold standard'—Cobb angle—yielding moderate correlations, and with many authors concluding that, the frequency of radiographs could be reduced. From the literature, it appears that no single surface measurement has been consistently related to the Cobb angle.

The main problem with the surface measurements systems that have been developed so far is that the measurements are susceptible to posture changes during data capture, which largely account for the lower accuracies obtained (see Chapter 3). In practice, patients are often brought into standardised positioning pads or other fixing elements in an attempt to standardise posture during data capture. However, in the case of deformities, this procedure may actually result in modification of the deformity being measured (Frobin et al, 1983). A free natural position of the patient is therefore required for proper acquisition of patient data. Moreover, to improve the accuracy of the results, processing techniques should be able to effectively mini-

mise the effect of random errors arising from the mobile and flexible nature of the human body.

For a mathematical analysis of the back shape, accurate and high resolution three dimensional coordinates of the back are required. Most of the imaging methods that have been developed can only yield 3D coordinates indirectly after further processing. As a result, these systems are slow, less accurate and/or have low resolution. Recent advances in technology have now enabled simultaneous acquisition of 3D coordinates and colour image data of objects. Three-dimensional laser scanning can directly yield high density data of the back surface at high rates, with high accuracies, achieving the result in almost real time (see Section 3.3.5). This makes the technique fast, reliable and non-invasive, and thus a natural method for acquiring 3D data of the surface of the back.

Though 3D data acquisition methods exist, techniques to extract relevant information for clinical use have not been widely developed. To convert the 3D point data to useful information for assessment of back deformities requires continuous research and development of methods and algorithms for interpretation and modelling. This investigation addresses this need and focuses on the development of processing techniques to extract useful information for assessment of spinal and back deformities.

1.3. Research Objectives and Methodology

The overall objective of this investigation is to develop methods to extract useful information from three-dimensional data of the back surface to enable quantification of the deformity. This information is required for assessment of spinal and back deformities, monitoring the deformity for progression and auditing effectiveness of treatment methods. The specific objectives are:

- Detection of anatomic landmarks of the back using mathematical shape analysis.
- Derivation of the back midline and computation of three dimensional surface measurements to quantify the deformity.
- Development of methods for visualizing and exploring back deformities.
- Development of a software system to automate the tasks involved in back shape analysis.

1.3.1. Detection of Anatomic Landmarks of the Back

Certain features of the normal human spine project to form bump or dimples on the back surface, especially the spinous processes (see Section 2.3). Some of these bony landmarks can be seen or at least palpated. Spinal deformity such as scoliosis will almost certainly result in corresponding deformity of back the surface. This phase of the investigation focuses on detection of these bony landmarks of the back using shape analysis. Successful detection of these landmarks will allow the derivation of the back midline and computation of surface measurements to quantify the deformity. The objective of this phase is to develop a method to detect the landmarks using mathematical shape analysis. An important property of surface shape (or surface curvature) is that it is invariant under rigid-body transformations such as rotation and translation. This implies that the measurements will be less dependent on patient positioning during data capture, and thereby improve the accuracy of the results.

1.3.2. Midline Derivation and Surface Measurements

The midline of the back represents the closest surface approximation of the underling spine. With the successful detection of the spinous process landmarks, the midline can be derived by fitting a smooth curve. Surface measurements can then be computed based on the midline. As already pointed out, the Cobb angle alone does not account for the three-dimensional nature of the deformity. The Scoliosis Research Society (SRS) in the USA has developed a terminology for three-dimensional description of scoliosis. This phase of the investigation builds upon this work and derives surface measurements consisting of distances, angles, areas and volumes.

1.3.3. Development of Visualization Tools

Visualization constitutes an important aspect of the assessment of spinal and back deformities. Effective visualization tools will enable exploratory examination of the extent of the deformity and allow changes to be tracked over time. The effect of treatment methods such as bracing or surgery can also be explored using effective visualization tools without the need for physical manipulation of the patient with the problem The objective of this phase of the investigation is to develop effective visualization tools to aid the assessment process.

1.3.4. Development of Automated System

The last objective of the research is to automate as much as possible the tasks involved in back shape analysis. Such a system could be deployed at the centre of the efficient management of patient data, where patient images, radiographs, back shape measurements, together with other relevant clinical data could be combined and used for decision making. An automated system could be applied to workforce or population screening. At risk people can be identified and appropriate correction can be given. Patients with progressive spinal curvature can be examined at regular intervals to determine the degree and timing of intervention required.

1.4. Scope of the Investigation

This investigation was not intended to comment extensively upon the nature, cause or treatment of spinal deformities but was to focus on the measurement problem. Only relevant background information has been given in Chapter 2 to lay the foundation for the study. Neither was it the aim of this investigation to develop a new surface data acquisition system. Several systems have already been developed for acquiring back surface data (see Chapter 3). Recent advances in technology have now enabled direct acquisition of three dimensional coordinates of objects at high rates and with high accuracies. The main focus of this investigation is to develop methods for extracting useful measurements for the assessment of spinal and back deformities, as outlined in Section 1.3.

1.5. Structure of the Thesis

This thesis is organised into seven chapters. Chapter 1 introduces the background, motivations, objectives and methodology of this investigation.

Chapter 2 gives a general background to three-dimensional spinal and back deformities. Beginning with the basic anatomy of the spine and the vertebrae, the anatomic landmarks of the back are described. The detection of these landmarks is one of the objectives of this work. Background to the three types of spinal curvatures (i.e. scoliosis, kyphosis and lordosis) is given; including possible aetiology, statistics, conventional assessment and treatment options. Finally, three-dimensional description of spinal deformities based on the terminology of the Scoliosis Research Society (SRS) is described.

Chapter 3 describes the fundamental surface measurements systems that have been developed for back shape assessment. The motivations for assessing spinal deformities from the shape of the back are outlined and various surface measurement methods that have been developed in past few decades are reviewed. These include tactile devices (such as the Scoliometer), moiré topography, structured light methods, phase measuring methods and three-dimensional scanning systems. The Minolta laser scanner which has been used for data acquisition is also discussed.

Chapter 4 discusses the mathematical foundation of the shape analysis techniques used in this study. The concept of surface curvature is discussed, including the various measures of curvature and methods used to compute them. The Least Squares method has been used extensively in this study, consequently, the basic principles are discussed, including the applications to which the method has been committed in this study, namely, surface fitting, curve fitting and coordinate transformation.

Chapter 5 covers three aspects of this research, namely, error reduction, curvature visualization and implementation issues. Error reduction constitutes a crucial aspect since all measurements contain errors which are propagated to any derived quantities. The problem of errors is addressed through scanner calibration and empirical techniques to estimate the nature of the errors and to determine an effective method for reducing their effect. Since visualization plays an important role in back shape analysis, a major part of this chapter is devoted to developing methods for visualizing the shape of the back. Lastly, major implementation and design issues regarding data structures and conversion, choice of development tools and system design are addressed in this chapter.

Chapter 6 focuses on the detection of anatomic landmarks of the back, derivation of the midline and computation of surface measurements. A mathematical method is developed to detect the landmarks from curvature maps and the results are compared with palpation. Based on the detected landmarks, the back midline is derived using constrained polynomial fitting via the Least Squares method. Surface measurements are then computed and the relationships among them analysed via correlation.

Chapter 7 summarises the contributions of this research, the conclusions drawn and recommendations for further work.

2. Three-Dimensional Spinal and Back Deformities

2.1. Introduction

This chapter provides general background to spinal and back deformities. Beginning with the basic anatomy of the spine, the visual and palpable landmarks of the back are described. These landmarks are of a particular interest since they are the parts of the spine that generally protrude to the surface of the back forming bumps or dimples. The three curvatures of the spine, namely, scoliosis, kyphosis and lordosis are described, with background information on their definition, aetiology, treatment options and incidence statistics. The conventional approach for assessing spinal deformities based on physical examination and radiographs is also described. The traditional approach of describing scoliosis as a lateral spinal curvature does not provide complete description of the three-dimensional nature of the deformity. Based on the terminology developed by the Scoliosis Research Society, we discuss three-dimensional measurements that can serve as alternative or complementary quantification for spinal and back deformities.

2.2. Basic Anatomy of the Spine

The human spine is a flexible column made up of a series of bones called vertebrae that support the entire upper body. The spinal column is divided into five regions of vertebra: cervical, thoracic, lumbar, sacrum and coccyx (Figure 2.1). There are 7 vertebrae in the cervical region, 12 thoracic vertebrae, 5 lumbar vertebrae, 5 in the sacrum (fused) and 4 in the coccyx. Thus, there are a total of 33 vertebrae in the spine. Each vertebra is distinguished by using a combination of a letter that describes the spinal region and a number that describes its location within that region. For example, C7 is the seventh cervical vertebra, T4 is the fourth thoracic vertebra and L5 is the fifth lumbar vertebra. The sacrum is simply referred to as S1 and the coccyx is not abbreviated or numbered. Inter-vertebral disks separate the individual vertebrae enabling spinal articulation and also act as dynamic shock absorbers.

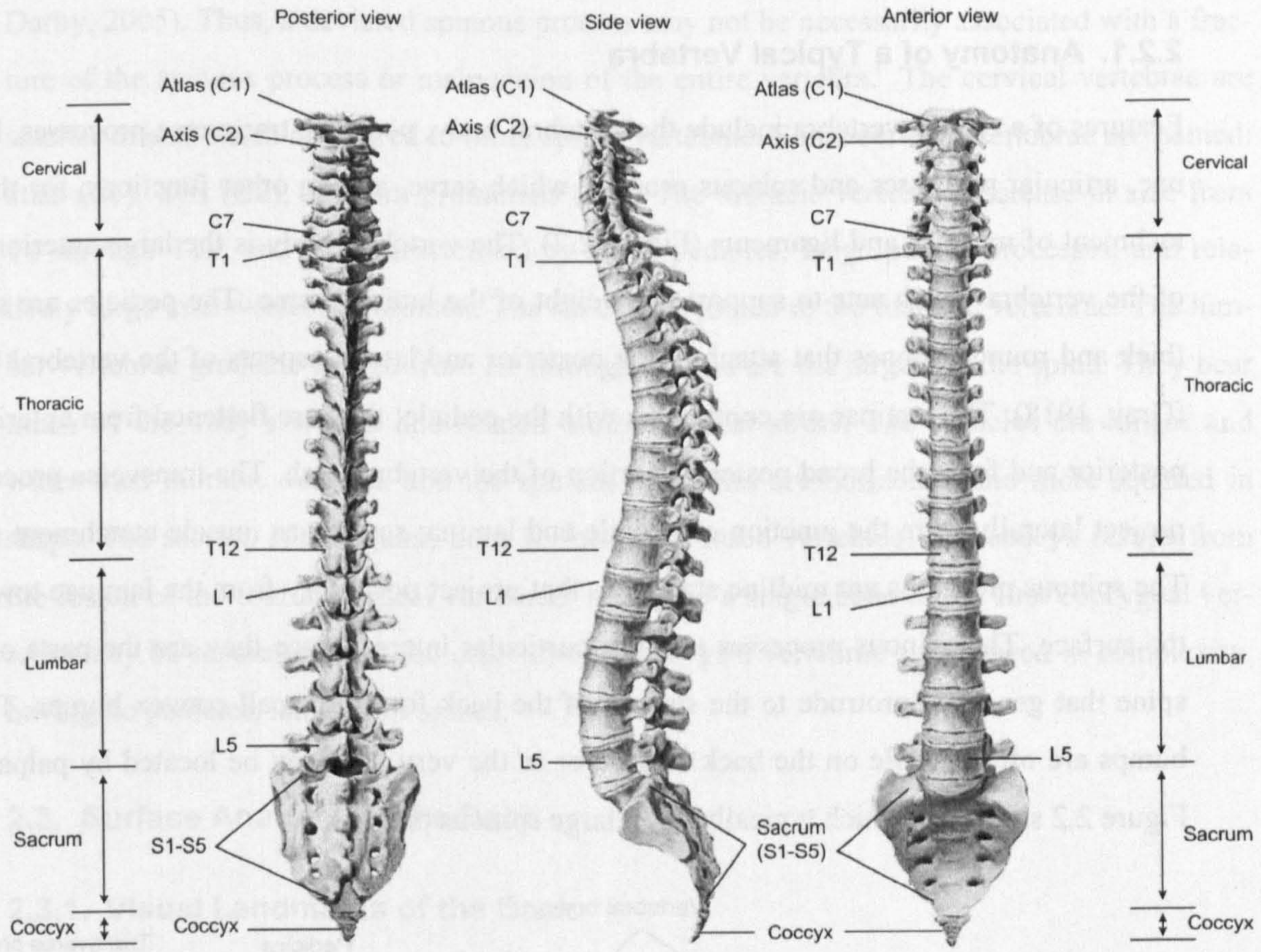


Figure 2.1: Anatomic regions of the spine. Illustration was prepared based on original images from Arai & Read (2005)

Table 2.1 summarises the divisions of the spine, body area and naming conventions.

Table 2.1: Regions of the spine

Spine Region	No. of Vertebrae	Body Area	Abbreviation
Cervical	7	Neck	C1 – C7
Thoracic	12	Upper Back (Chest)	T1 – T12
Lumbar	5	Low Back	L1 – L5
Sacrum	5 (fused)	Pelvis	S1 – S5
Coccyx	4	Tailbone	None

2.2.1. Anatomy of a Typical Vertebra

Features of a typical vertebra include the vertebral body, pedicles, transverse processes, laminae, articular processes and spinous process, which serve, among other functions, for the attachment of muscles and ligaments (Figure 2.2). The vertebral body is the large anterior part of the vertebra which acts to support the weight of the human frame. The pedicles are short, thick and rounded bones that attach to the posterior and lateral aspects of the vertebral body (Gray, 1918). The laminae are continuous with the pedicle; they are flattened from anterior to posterior and form the broad posterior portion of the vertebral arch. The transverse processes project laterally from the junction of pedicle and lamina; serving as muscle attachment sites. The spinous processes are midline structures that project posteriory from the laminae towards the surface. The spinous processes are of a particular interest since they are the parts of the spine that generally protrude to the surface of the back forming small convex bumps. These bumps are often visible on the back surface or in the very least can be located by palpation. Figure 2.2 shows C7 which typically has a large spinous process.

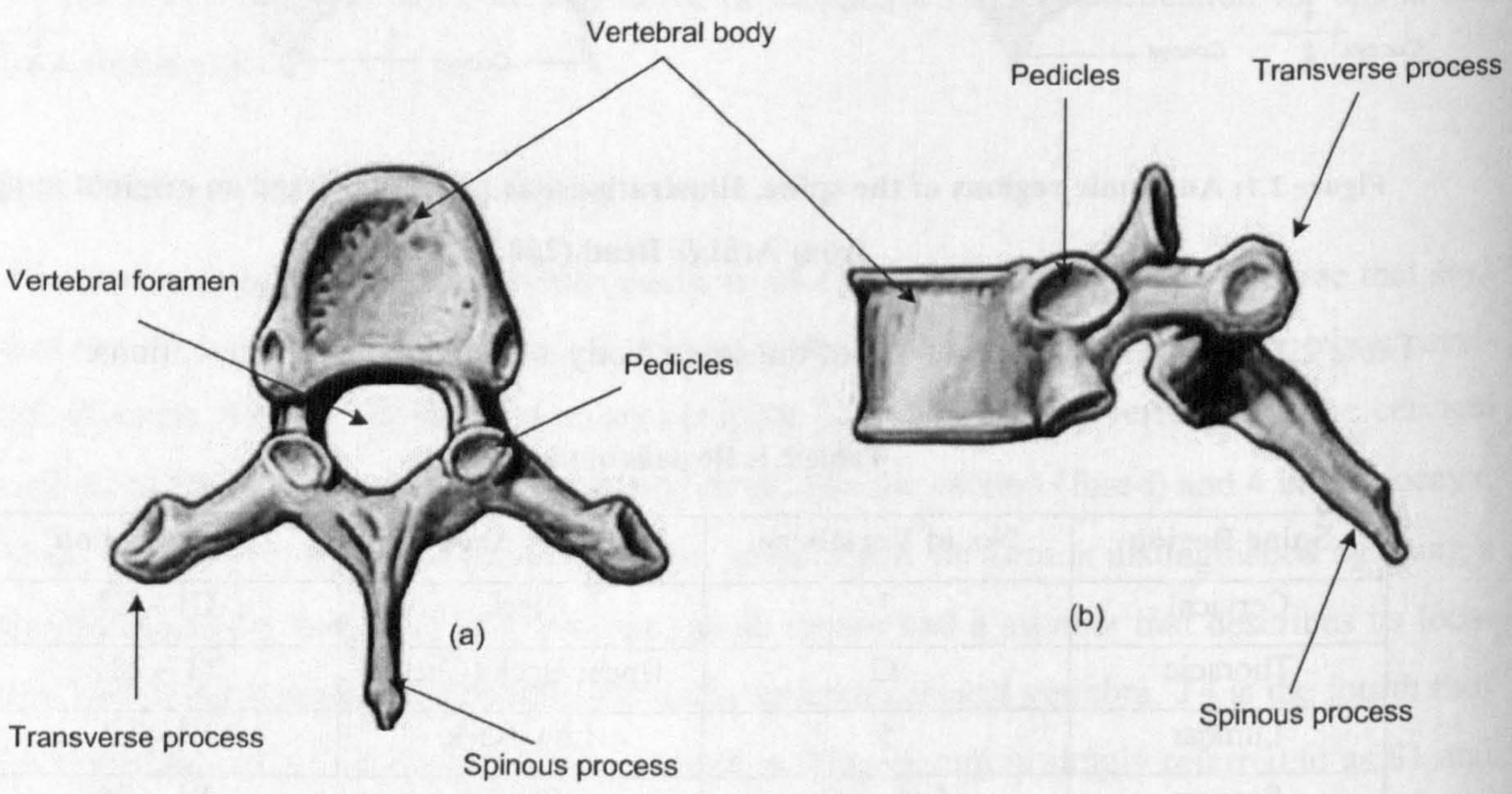


Figure 2.2: A typical thoracic vertebra (a) axial or overhead view (b) lateral or side view. Illustration prepared based on original images taken from Bridwell (2001b)

The size, shape and direction of the spinous processes vary from one region of the spinal column to the next (Gray, 1918). A spinous process may also normally deviate to the left or right of the midline, and this may be a cause of confusion in clinical examination (Cramer &

Darby, 2005). Thus, a deviated spinous process may not be necessarily associated with a fracture of the spinous process or malposition of the entire vertebra. The cervical vertebrae are smaller in size when compared to other spinal vertebrae. Some cervical vertebrae are named: atlas (C1), axis (C2), vertebra prominens (C7). The thoracic vertebrae increase in size from T1 through T12; and are characterised by small pedicles, long spinous processes, and relatively large inter-vertebral foramen. The rib cage is joined to the thoracic vertebrae. The lumbar vertebrae graduate in size from L1 through L5 and are the largest in the spine. They bear much of the body's weight and related biomechanical stress. The pedicles are longer and wider than thoracic vertebra and the spinous processes are horizontal and more squared in shape. The sacrum is triangular, been formed by 5 fused vertebrae. The coccyx results from the fusion of the four coccygeal vertebrae; it may be a single bone or the first coccygeal vertebra may be separated from the other three. Coccygeal vertebrae are reduced in complexity, having no pedicles, laminae or spines.

2.3. Surface Anatomy of the Back

2.3.1. Visual Landmarks of the Back

The subcutaneous parts of the spine are the apices of the spinous processes. These are distinguishable at the bottom of a furrow which runs down the middle line of the back from the external occipital protuberance (EOP) to the middle of the sacrum. The furrow is shallow in the lower cervical region and deepest in the lumbar region, widening inferiorly to form a triangle with a line connecting the posterior superior iliac spines (PSISs), forming the base above and the gluteal cleft forming the apex of the triangle below. The PSISs are visible as a pair of dimples (denoted in thesis as LPSIS and RPSIS respectively) located about 3cm lateral to the midline at the level of the S2 spinous tubercle (Cramer & Darby, 2005).

Several bony landmarks are usually visible (see Figure 2.3). The spinous process of C7 is usually visible in the lower cervical region. In 60-70% of the population, the spinous process of C7 is the most prominent while the spinous process of T1 is more prominent in the other 30-49% (Cramer & Darby, 2005). When the patient's head is flexed, the spinous process of C6 is usually visible as well. The other cervical spinous processes are usually sunken. In the thoracic region the furrow is shallow and during stooping disappears, and then the spinous processes appear more or less visible. In the lumbar region the furrow is deep and spinous processes are frequently indicated by little pits or depressions, especially when the muscles in

the loins are well-developed. In the sacral region the furrow is shallower, presenting a flattened area which ends below at the most prominent part of the dorsal surface of the sacrum.

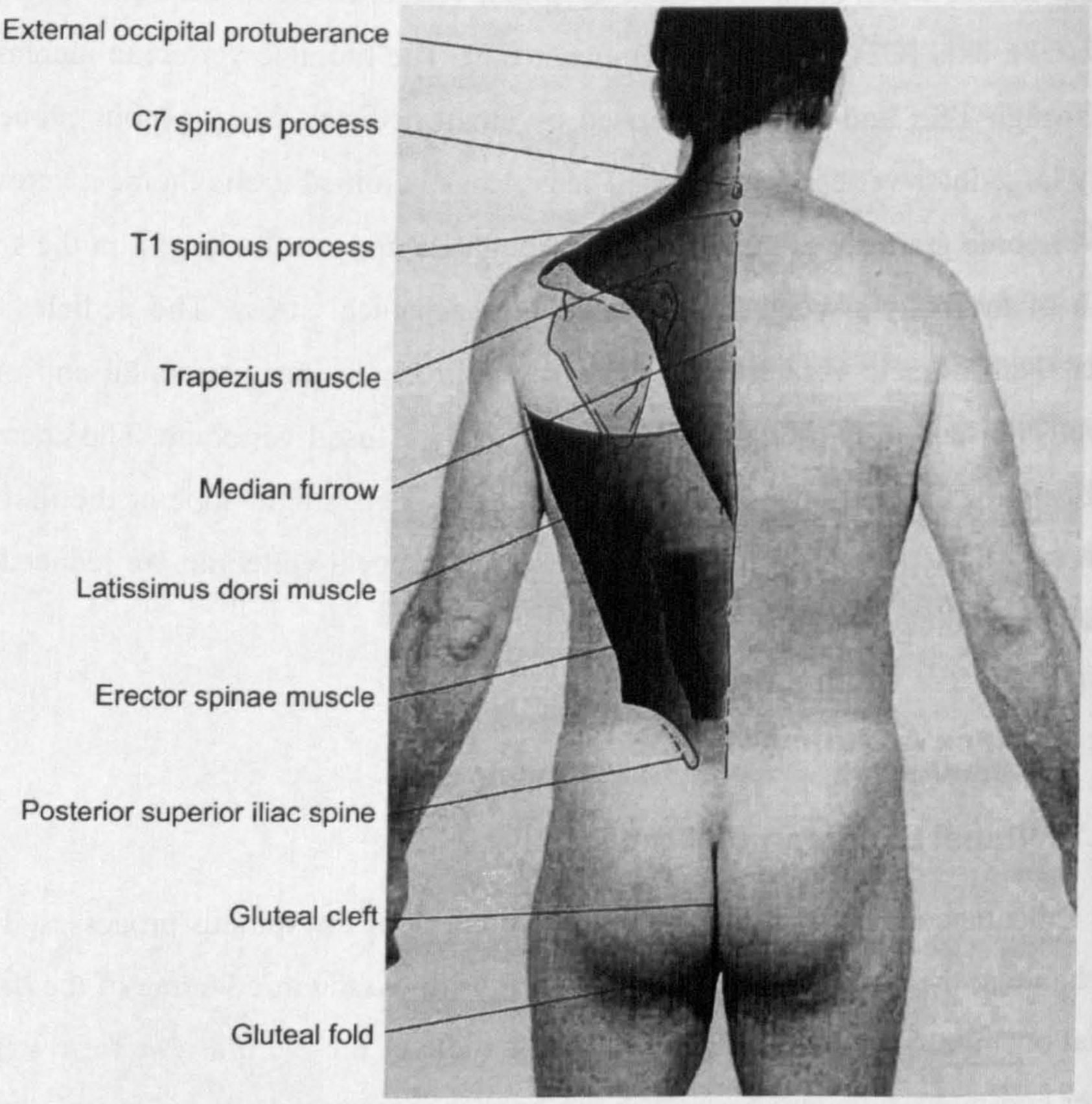


Figure 2.3: Visible landmarks of the back (Cramer & Darby, 2005)

Besides bony landmarks, several muscles are commonly visible in the back region. The *trapezius* is a large, flat, triangular muscle that originates in the midline from the EOP to the spinous process of T12 and inserts laterally onto the spine of the scapula. The *latissimus dorsi* extends from the region of the iliac crest to the posterior border of the axilla and forming the lateral border of the lower thoracic portion of the back. Between the trapezius medially and latissimus dorsi laterally, the inferior angle of the scapula may be seen at approximately the level of T7 spinous process. The *erector spinae* muscles form two large longitudinal masses in the lumber region that extends approximately 10 cm laterally from the midline.

2.3.2. Palpable Landmarks of the Back

The following structures are not usually visible, but can be located by palpation (see Figure 2.4). The spinous process of the C2 is the first readily palpable bone in the posterior midline below the EOP; the second being C7. The spinous process of C6 is usually palpable with full flexion of the neck. The other cervical spinous processes are generally more difficult to palpate. In the thoracic region, T1 is the third palpable spinous process. The spinous process of T4 is located at the extreme convexity of the thoracic kyphosis and is usually the most prominent spinous process below the root of the neck. The spinous processes of T9 and T10 are often palpably closer together; and the spinous process of T12 is roughly located between the level of the inferior angle of the scapula and the superior margin of the iliac crest. In the lumbosacral region, the spinous processes are horizontal and more squared in shape. The spinous processes of L4 and L5 are shorter than other spinous processes and difficult to palpate. The last palpable bone is the tip of the coccyx which can be found in the gluteal cleft approximately 1 cm posterior to the anus.

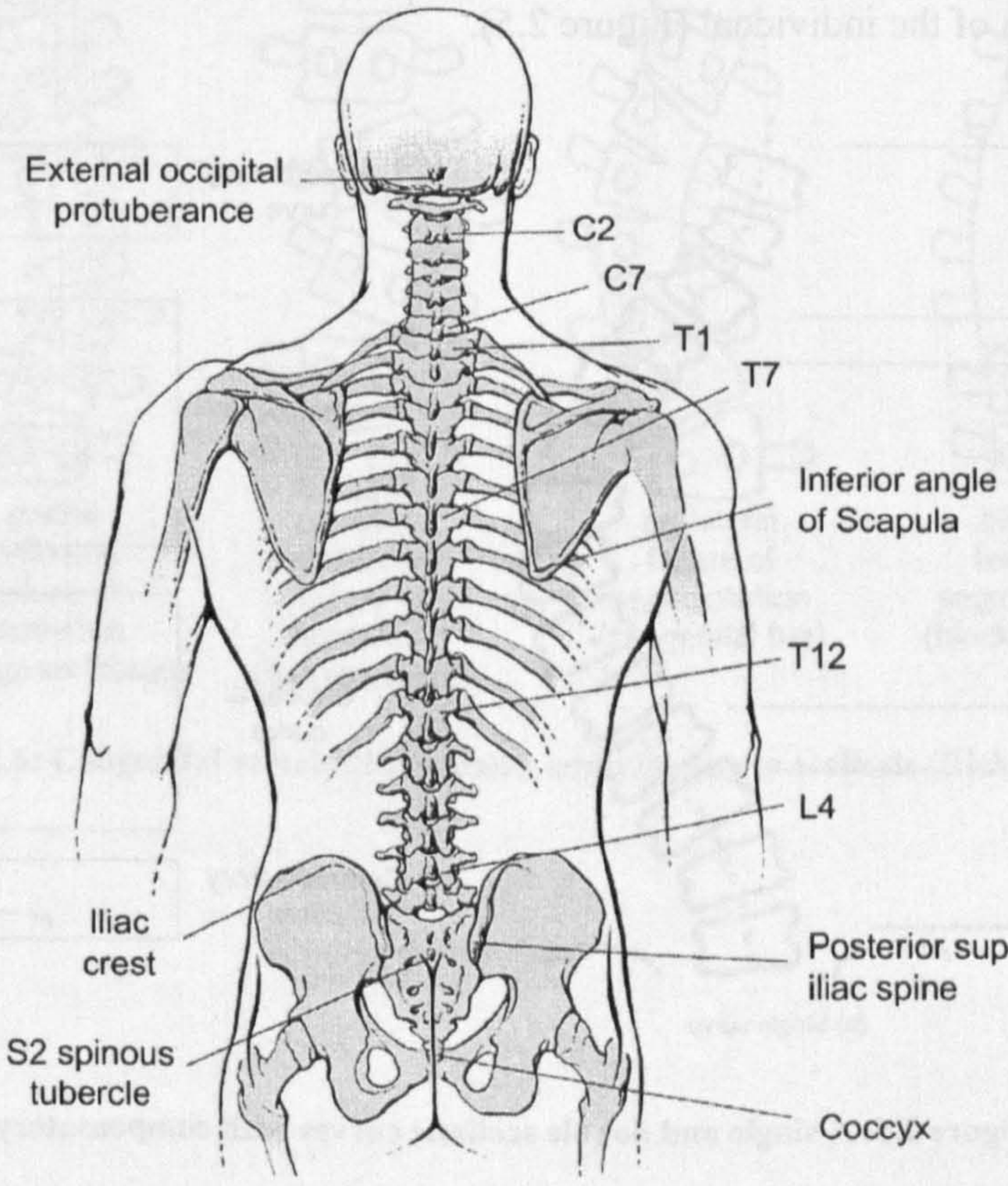


Figure 2.4: Palpable landmarks of the back (Cramer & Darby, 2005)

2.4. Curvatures of the Spine

There are three types of spinal curvatures, namely, *scoliosis*, *kyphosis* and *lordosis*. Scoliosis refers to a lateral curvature in the frontal plane; kyphosis refers to the outward curvature of the thoracic spine in the sagittal plane; and lordosis refers to the inward curvature of the lumbar spine in the sagittal plane (SRS, 2000). Scoliosis is always abnormal, although mild, non-structural scoliosis may pose no real health threat to the patient and may require no treatment other than monitoring for progression. The normal spine always exhibits a small degree of both kyphotic and lordotic curvature. Exaggeration of the kyphotic curve results in what is known as *hunchback*. Exaggeration of the lordotic curve is often called *swayback*. These three types of spinal curvature are further discussed in the following sections.

2.4.1. Scoliosis

Scoliosis, in its simplest definition, refers to a lateral curvature of the spine in the frontal plane. Typically, scoliosis will exhibit either a single or double structural curve and will be accompanied by compensatory curves above and below the structural curves to maintain the general position of the individual (Figure 2.5).

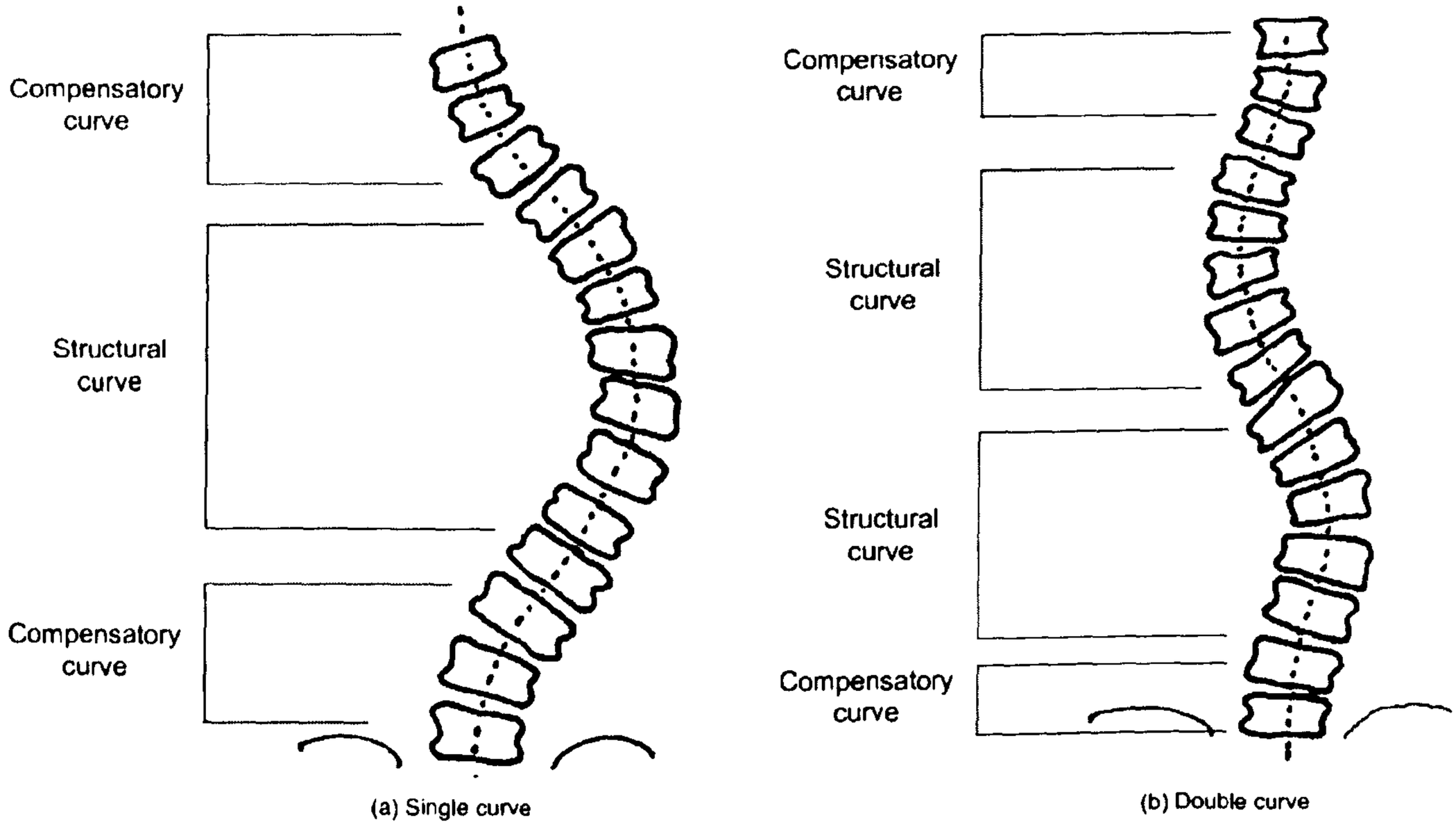


Figure 2.5: A single and double scoliotic curves with compensatory curves

This simple definition of scoliosis does not provide a complete description of the deformity. For this reason, the Scoliosis Research Society (SRS) in the USA developed a three-

dimensional terminology of scoliosis covering the geometrical properties of the vertebrae and the spine (Stokes, 1994). This is discussed in Section 2.5.

2.4.1.1. Developmental Anomalies of the Vertebrae

Due to the complicated nature of the developmental processes of the vertebrae, various anomalies can occur. Although many of these anomalies do not pose any health problems and are never detected, some of them can have a profound impact on the health of the individual (Gregory et al, 2005). Scoliosis may result from morphological deformation of the vertebrae during development. Hemivertebra or vertebral wedging is the situation where only one half of the vertebral body develops resulting in a wedge-shaped vertebra, causing the spine to curve laterally at the side of the defect. Failure of vertebral formation can be partial or complete (Figure 2.6). Block vertebra or congenital bar represents a non-segmentation of the somites and results in the fusion of two adjacent vertebrae to form a single vertebra twice as high than normal.

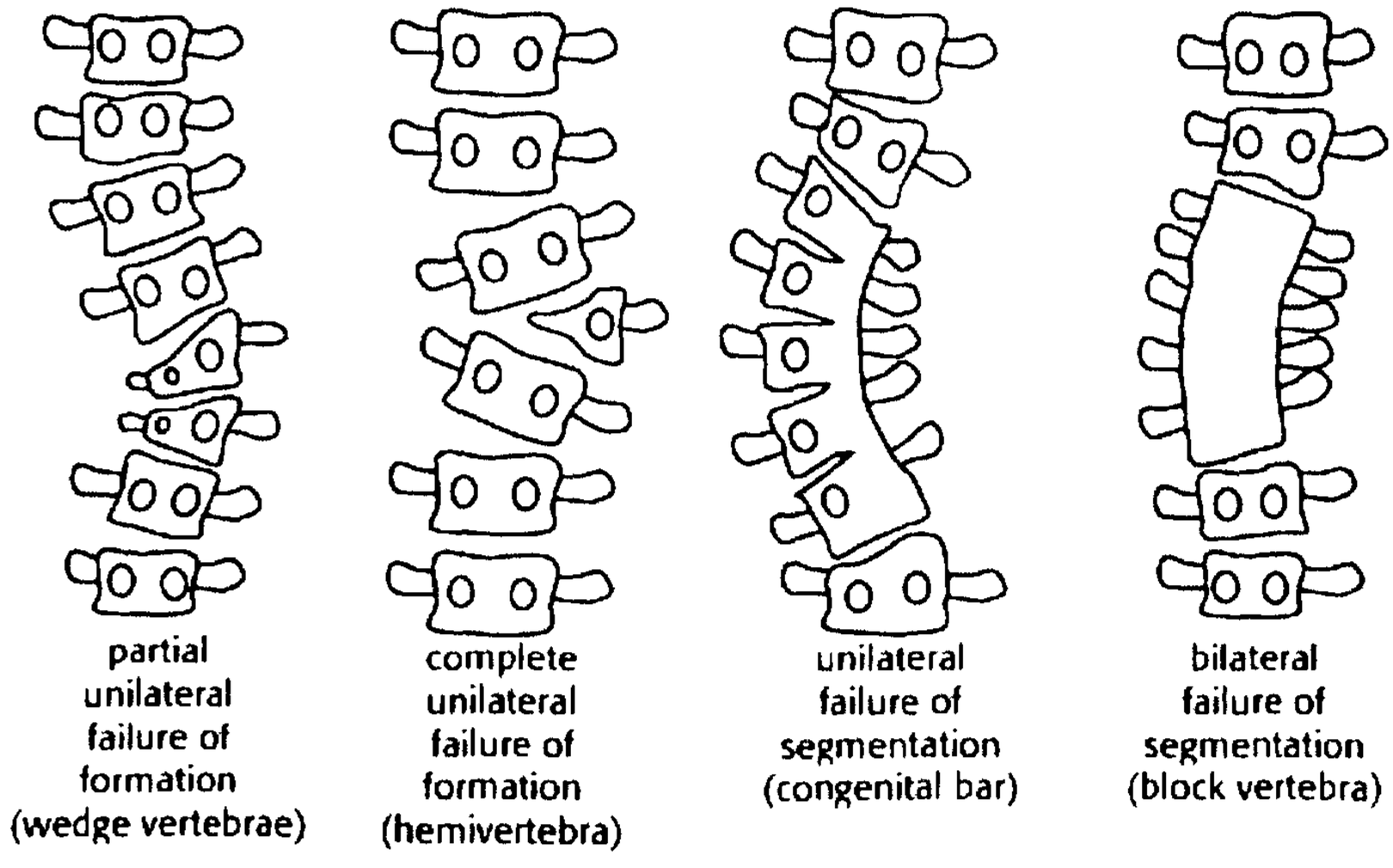


Figure 2.6: Congenital vertebral anomalies contributing to scoliosis (Richardson, 2000)

2.4.1.2. Vertebral Rotation and Rib Cage Distortion

Structural scoliosis is a deformity that is characterised by both lateral curvature and vertebral rotation (McAlister and Shackelford, 1975). As the deformity progresses, the vertebrae and spinous processes rotate toward the concavity of the curve. On the concave side of the curve, the ribs are close together. On the convex side, they are widely separated. As the vertebral bodies rotate, the spinous processes deviate more and more to the concave side and the ribs follow the rotation of the vertebrae (Figure 2.7). The posterior ribs on the convex side are pushed posteriory, causing the characteristic rib hump seen in thoracic scoliosis. The anterior ribs on the concave side are pushed anteriorly (Richardson, 2000).

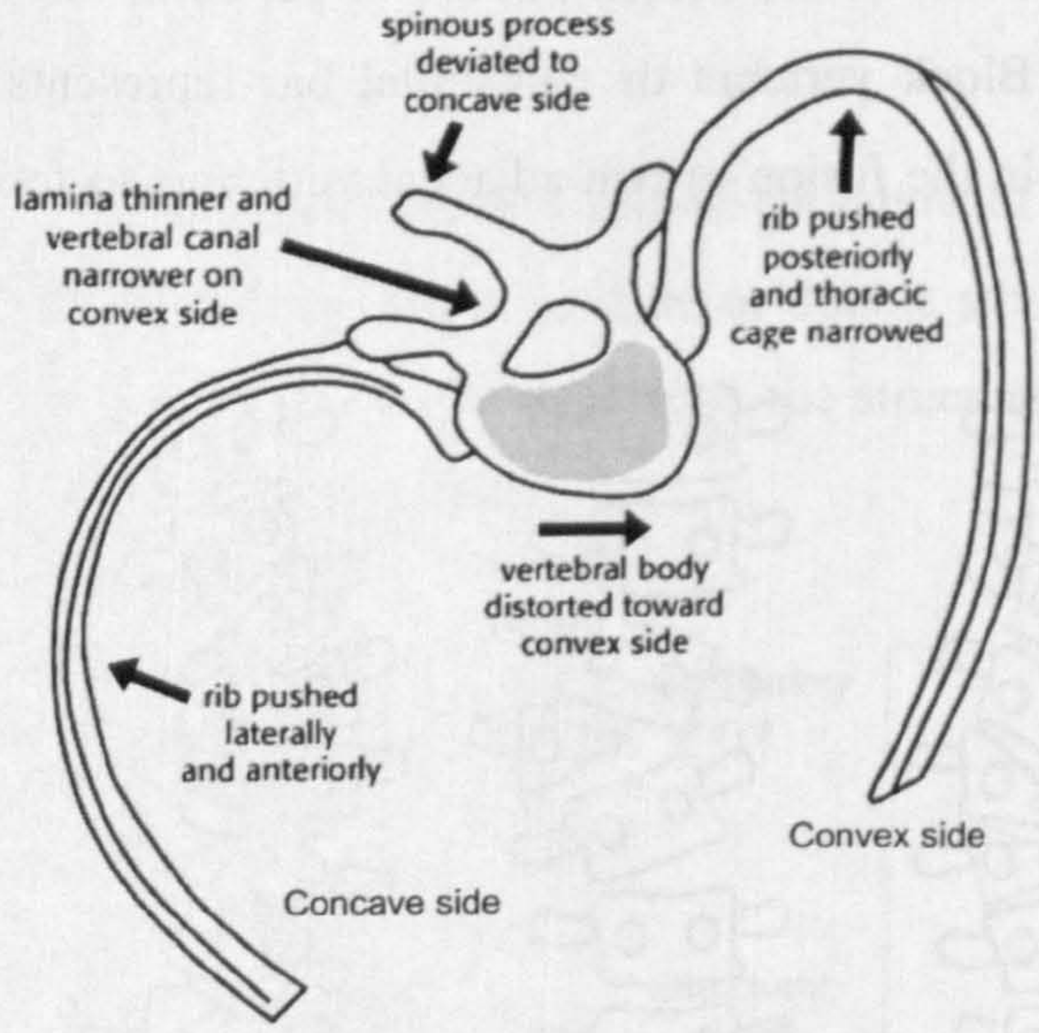
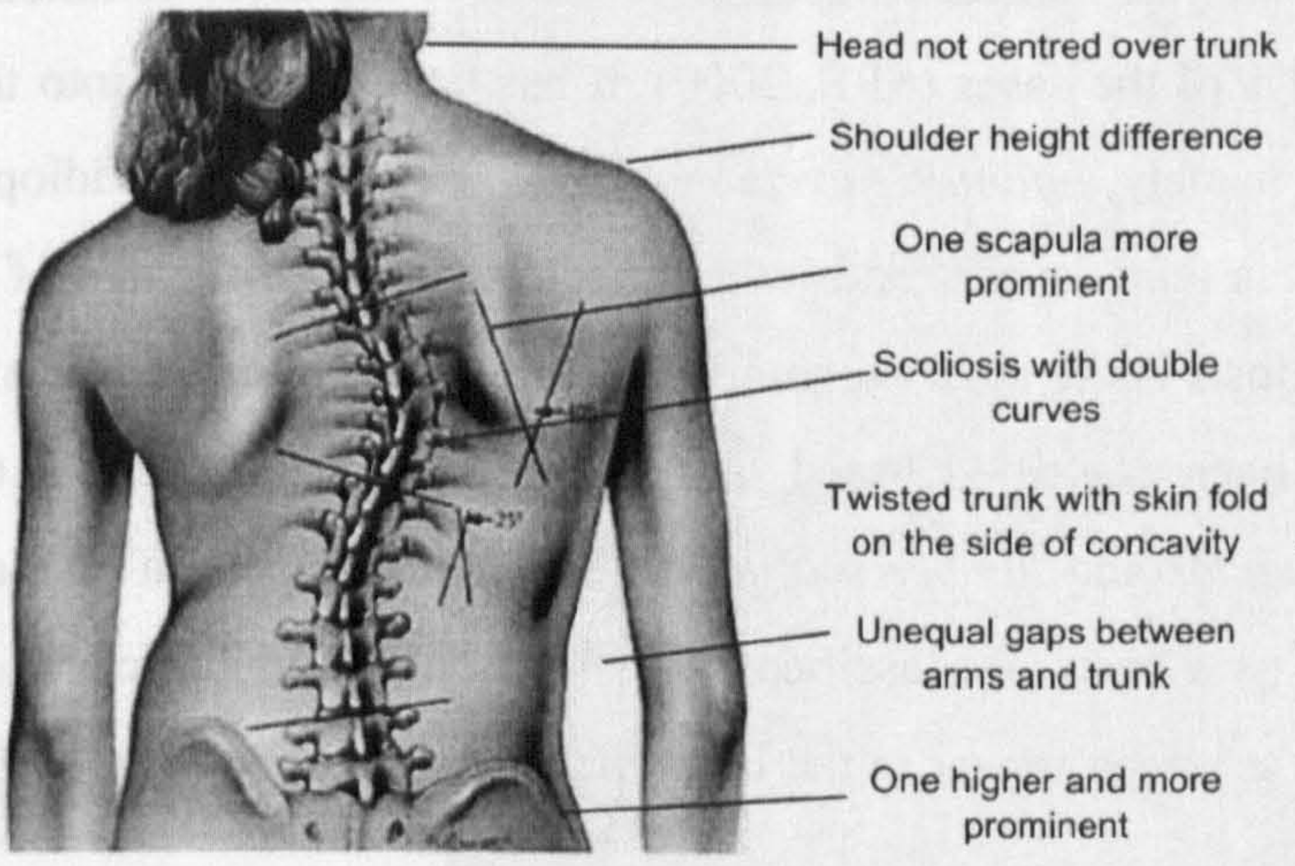


Figure 2.7: Vertebral rotation and rib distortion (Richardson, 2000)

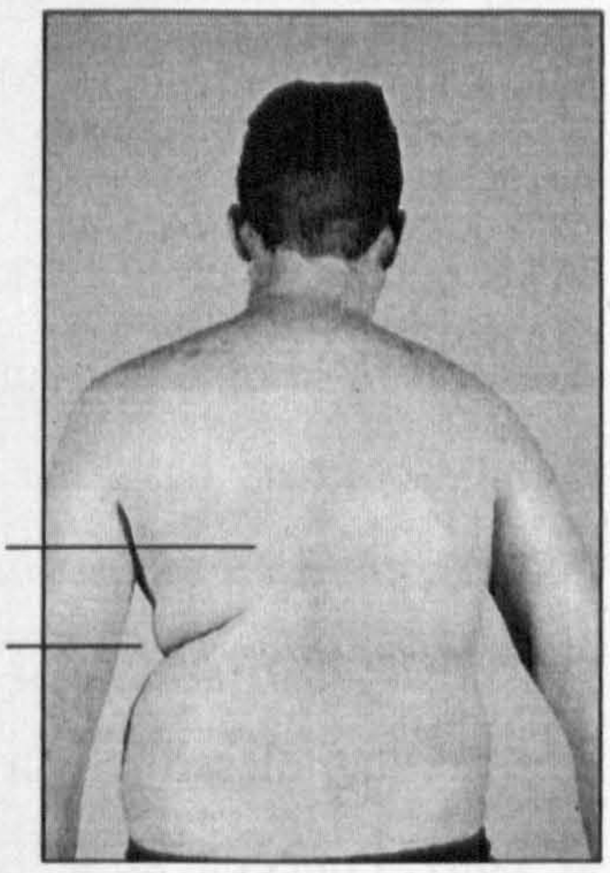
2.4.1.3. Surface Deformity Caused by Scoliosis

The skeletal distortions caused by scoliosis will almost certainly result in the deformity of the trunk surface. The surface deformity typically include shoulder asymmetry, rib hump, scapula asymmetry, pelvic asymmetry, a skin fold on the side of the concavity and a subtle twisting of the trunk (Mahood et al, 1995). In Figure 2.8, we observe various surface deformities caused by scoliosis: twisting of the whole trunk toward the side of the concavity; head not centred over trunk; right shoulder higher than the left; left pelvis higher than the right; vertebra rotation causing rib cage distortion and the rib hump. The surface deformity enables surface assessment to be carried out. The cosmetic deformity of the trunk surface is actually more im-

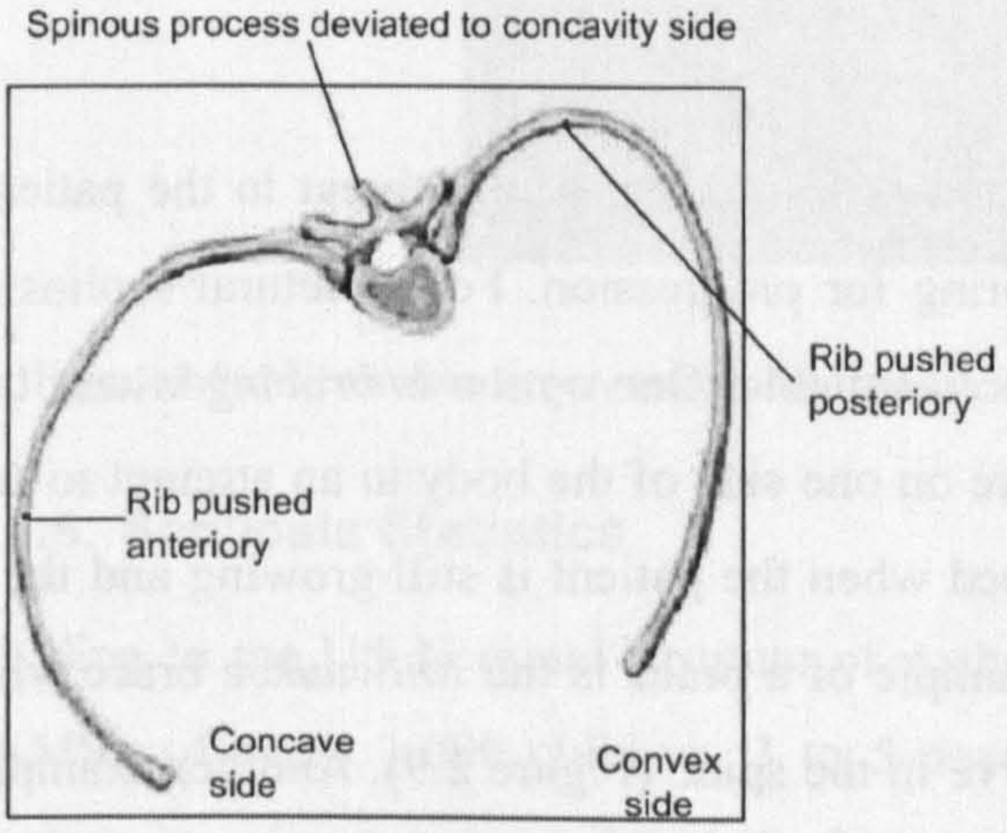
portant to the patient and family, and is what frequently motivates them to seek treatment (Mahood et al, 1995).



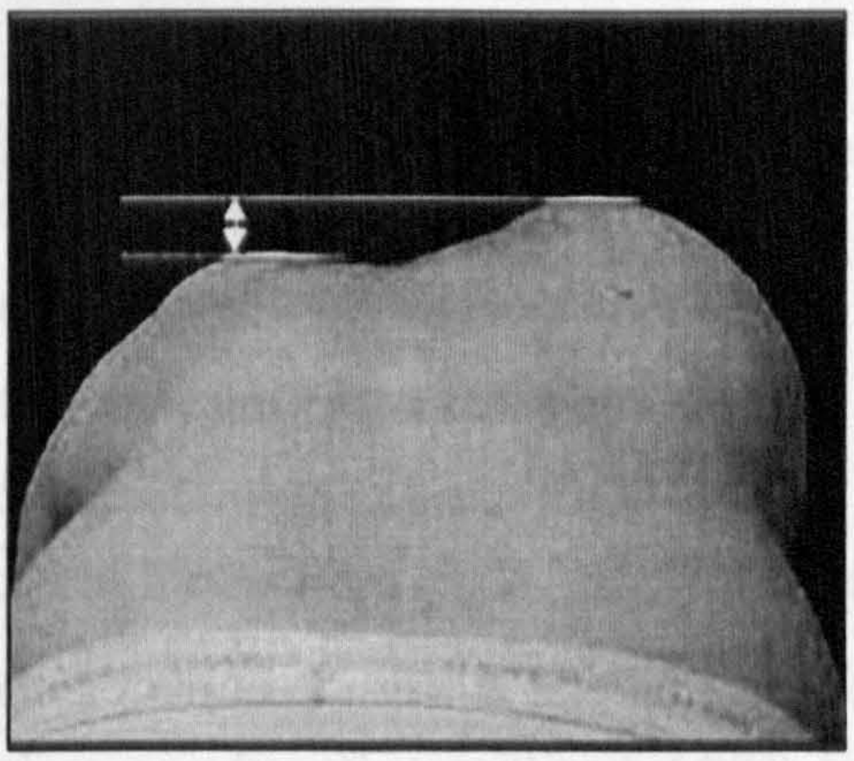
(a) Scoliosis showing surface deformities



(b) Scoliosis showing surface deformities



(c) Rib hump – skeletal deformity



(d) Rib hump – surface deformity

Figure 2.8: Skeletal and surface deformity in scoliosis. Illustration is based on original images taken from (Wasserman, 1992; Dennis, 2003)

2.4.1.4. Aetiology and Classification of Scoliosis

Clinically, there are two main types of scoliosis, namely, *structural* and *non-structural* (or functional) scoliosis (McAlister and Shackelford, 1975). Non-structural scoliosis is mild, non-progressive and correctable by ipsilateral bending. It is a temporary, changing curve caused by some underlying condition such as leg length difference (LLD), muscle spasm or inflammatory conditions such as appendicitis. In structural scoliosis, the vertebrae undergo morphological changes such as wedging and rotations. Based on its aetiology, structural scoliosis has been classified as *idiopathic*, *congenital* or associated with miscellaneous conditions includ-

ing developmental, neuromuscular and tumoural causes (McAlister and Shackelford, 1975). The cause of idiopathic scoliosis remains unknown but it is thought to be associated with many factors including those just named. Idiopathic scoliosis is the most common type and accounts for more than 80% of the cases (SRS, 2000). It has been classified into three types based on the age of onset, namely, *infantile*, *juvenile* and *adolescent*. Infantile idiopathic scoliosis occurs in children less than 3 years and juvenile occurs between the age of 3 and 10. Adolescent idiopathic scoliosis is the most common type and occurs after the age of 10. Idiopathic scoliosis has also been classified based on the location of the apex of the curve, namely, *thoracic*, *thora-lumbar* and *lumber* scoliosis. Moreover, the pattern of the scoliotic curve has also been used as a basis of classification: single curve, double curve and triple curve. Other classification schemes appear in the literature (Cruickshank et al, 1989; Lenke et al, 2001; Poncet et al, 2001).

2.4.1.5. Treatment of Scoliosis

When scoliosis is mild, non-structural and poses no real health threat to the patient, it may require no treatment other than monitoring for progression. For structural scoliosis that requires treatment, a number of options are available. One option is *bracing* where the patient wears a rigid brace which exerts pressure on one side of the body in an attempt to correct the deformity. Bracing is normally prescribed when the patient is still growing and the scoliosis has the tendency of progressing. An example of a brace is the *Milwaukee* brace which has a neck ring and is worn to correct any curve in the spine (Figure 2.9). Another example is *Thoracolumbosacral orthosis* which fits under the arm and around the rib cage, lower back, hips, and is typically worn to correct curves whose apex is at or below the eighth thoracic vertebra (Howard et al, 1998). In cases where leg length difference (LLD) is the underlying cause of functional scoliosis, it may be corrected simply by shoe raise in the shorter leg (Baylis et al, 1988; Zabjek et al, 2001). The final alternative, but more severe option is to intervene with spinal surgery to arrest progression of the deformity and where possible to straighten the curve without injury to the spinal chord. Many surgical techniques exist; the main procedure being correction, stabilization, and fusion of the curve.

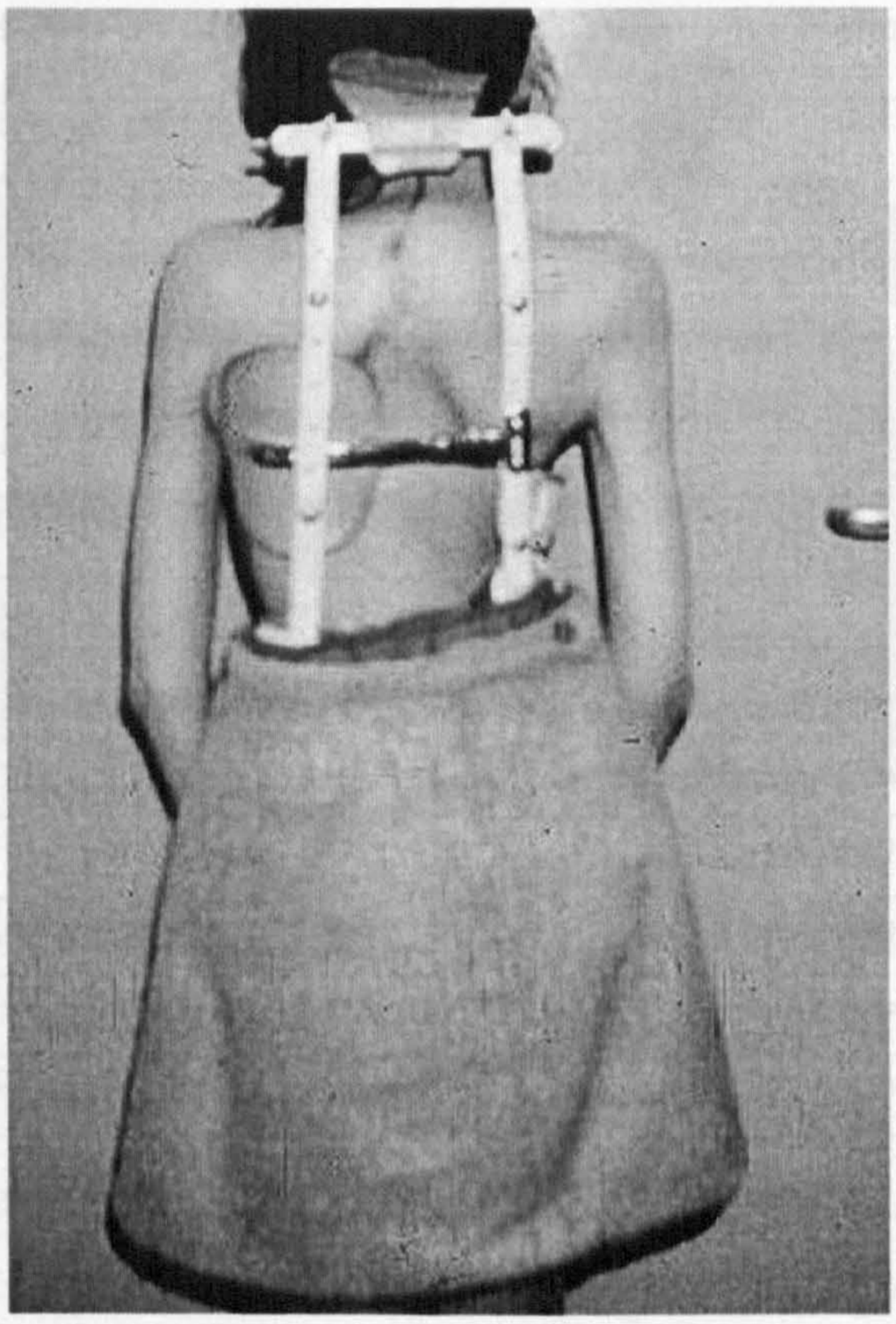


Figure 2.9: Milwaukee brace, worn after spinal fusion to hold the spine in place (SAUK, 2000)

2.4.1.6. Scoliosis Statistics

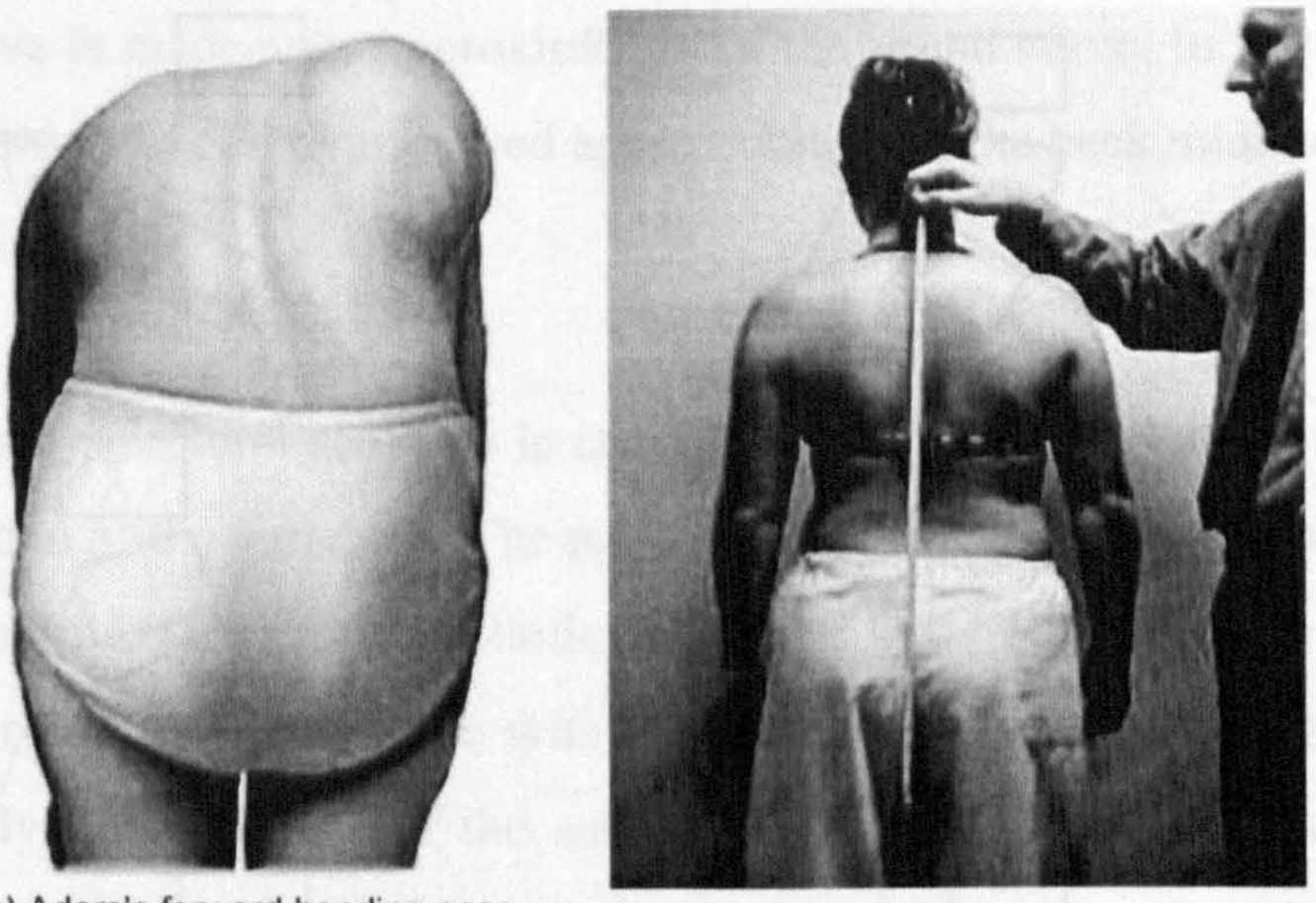
According to the US National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), of every 1,000 children, 3 to 5 develop spinal curves that are considered large enough to require treatment (NIAMS, 2001). Adolescent idiopathic scoliosis (AIS) is the more common type accounting for more than 80% of scoliosis cases. Table 2.2 shows scoliosis statistics over a period of six years taken from the UK National Episode Statistics. The mean number of scoliosis episodes from 1998-2004 was 2371. Out of this total, about 90% required hospital admission. These statistics also confirm that females are more likely to develop scoliosis than males.

Table 2.2: Scoliosis statistics (Source: Hospital Episode Statistics, 1998-2004, Department of Health, UK)

Year	98-99	99-00	00-01	01-02	02-03	03-04	Mean
Hospital consultant episode	2110	2307	2259	2129	2559	2866	2017
Percentage of hospital adm. (%)	91	92	91	91	89	88	90
Percentage of Males (%)	33	33	33	32	32	34	33
Percentage of Females (%)	67	67	67	68	68	66	67
Percentage emergency adm. (%)	7	7	8	7	11	6	8
Waiting list (%)	61	68	63	67	63	62	64
Mean waiting time (days)	163	165	173	185	186	154	171
Mean length of stay (days)	11	10	11	11	10	10	11
Percentage of age group 0-14 (%)	35	35	32	37	37	44	37
Percentage of age group 15-59 (%)	47	46	46	43	42	37	44
Percentage of age group 60-74 (%)	10	11	13	11	11	10	11
Percentage of age group 75+ (%)	8	8	9	9	8	9	9
Mean age (years)	31	32	33	31	30	29	31

2.4.1.7. Conventional Assessment of Scoliosis

The conventional approach for assessment of scoliosis usually begins with physical examination followed by radiographic assessment if required. Physician examination will typically include palpating the back surface or the trunk in attempt to assess the extent of the deformity or asymmetry of the back. The examiner will typically look for physical indications such as uneven shoulders or hips, humpback, or prominent hip. The patient is often instructed to bend forward at the waist, with arms extended forward (Adam's forward bending test) which emphasises asymmetry of thoracic prominence or rib hump (Figure 2.10a). The rib hump is typically measured using the Scoliometer (see Section 3.3.1.1) while the patient is in Adam's forward bending pose. Any discrepancy in the lengths of the two legs is also measured, usually with a tape. To obtain a general alignment of the spine, a plumb line is sometimes 'dropped' from the C7 vertebra and is allowed to hang below the buttocks (Figure 2.10b). In the case of scoliosis, the line will normally not hang between the buttocks, indicating a curved spine.



(a) Adam's forward bending pose

(b) Plumb line

Figure 2.10: (a) Patient in Adam's forward bending position (b) Plumb line used to measure spinal alignment. Illustration based on original images from (Allrefer, 2006)

The patient's ability to perform a range of motion such as flexion, extension, bending, and rotation movements is also usually assessed. In addition to testing reflexes, any symptoms experienced by the patient such as back pain, numbness, extremity weakness, muscle spasm, etc, are also recorded. These factors are taken into account in deciding whether or not the patient requires further assessment. If the patient requires further examination, radiographic assessment is carried out. Two parameters are typically measured from the radiographs, namely, the angle of curvature and vertebral rotation, which are discussed below.

Angle of Curvature

The most commonly used method for measuring the angle of curvature of the spine is that proposed by Cobb (Cobb, 1948). To measure the Cobb angle, one must first decide which vertebrae are the end-vertebrae of the curve. These end-vertebrae are the vertebrae at the upper and lower limits of the curve which tilt most severely toward the concavity of the curve. Once these vertebrae have been selected, lines are then drawn along the upper endplate of the upper body and along the lower endplate of the lower body as shown in Figure 2.11a. The Cobb angle is then the angle between these two lines. The Cobb angle thus depends on the *subjective* selection of the end vertebrae bodies, which may account for the variation in its measurement (see Section 3.2.3). Another method, proposed by Ferguson (1930), for estimating the angle of curvature measures the angle between lines drawn from the centroids of the end vertebrae to the centroid of the apical vertebra or disc (Figure 2.11b).

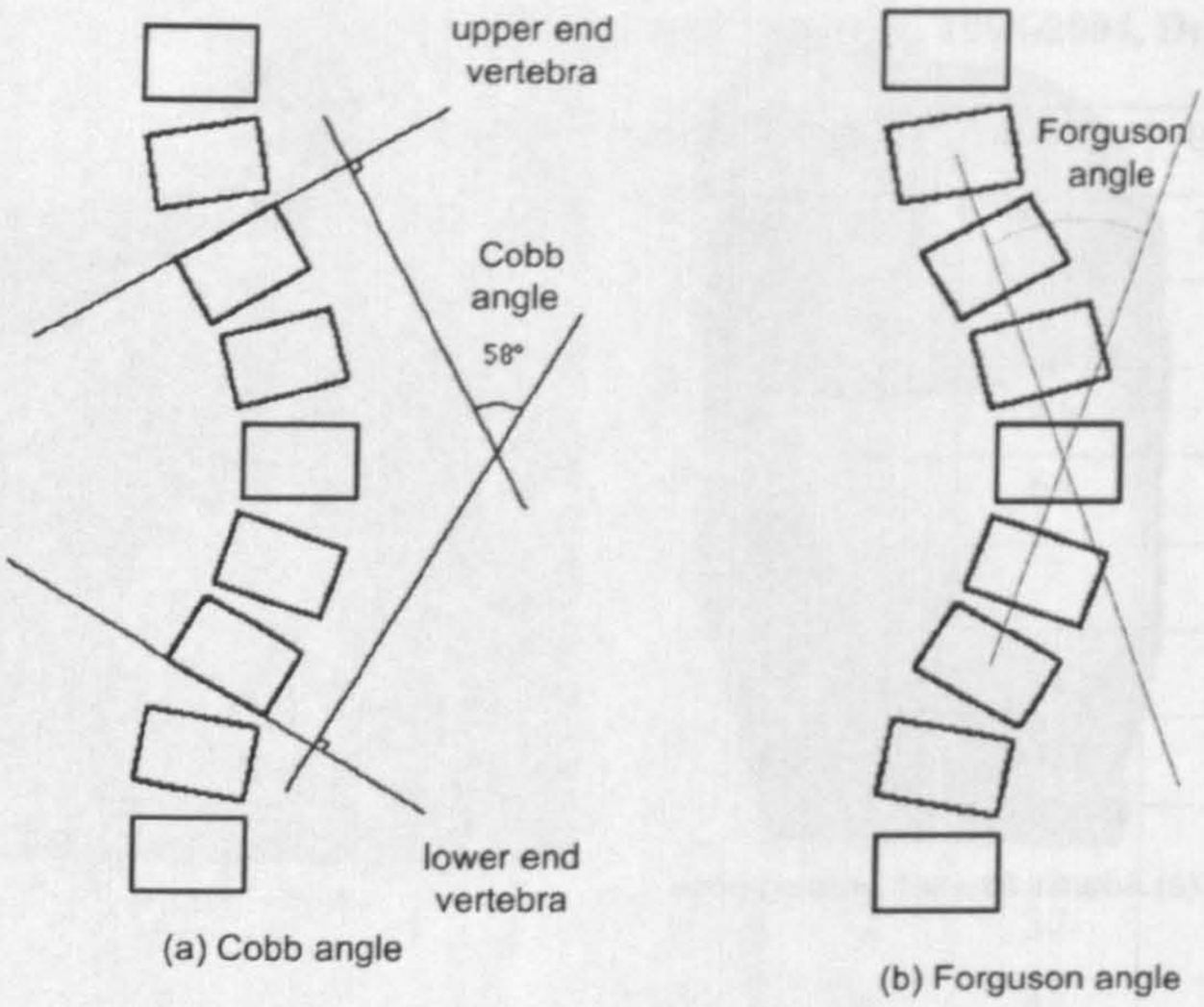


Figure 2.11: Curvature angle measurement on a radiograph (a) Cobb method (b) Ferguson method

To examine the geometric relationship between the Cobb and Ferguson angles, we approximate the spine by a smooth curve passing through the centroids of the vertebra bodies. For simplicity, let this approximating curve be a simple circular curve. From Figure 2.12, it is trivial to show that the Cobb angle (α) is twice the magnitude of the Ferguson angle (β), that is, $\alpha = 2\beta$.

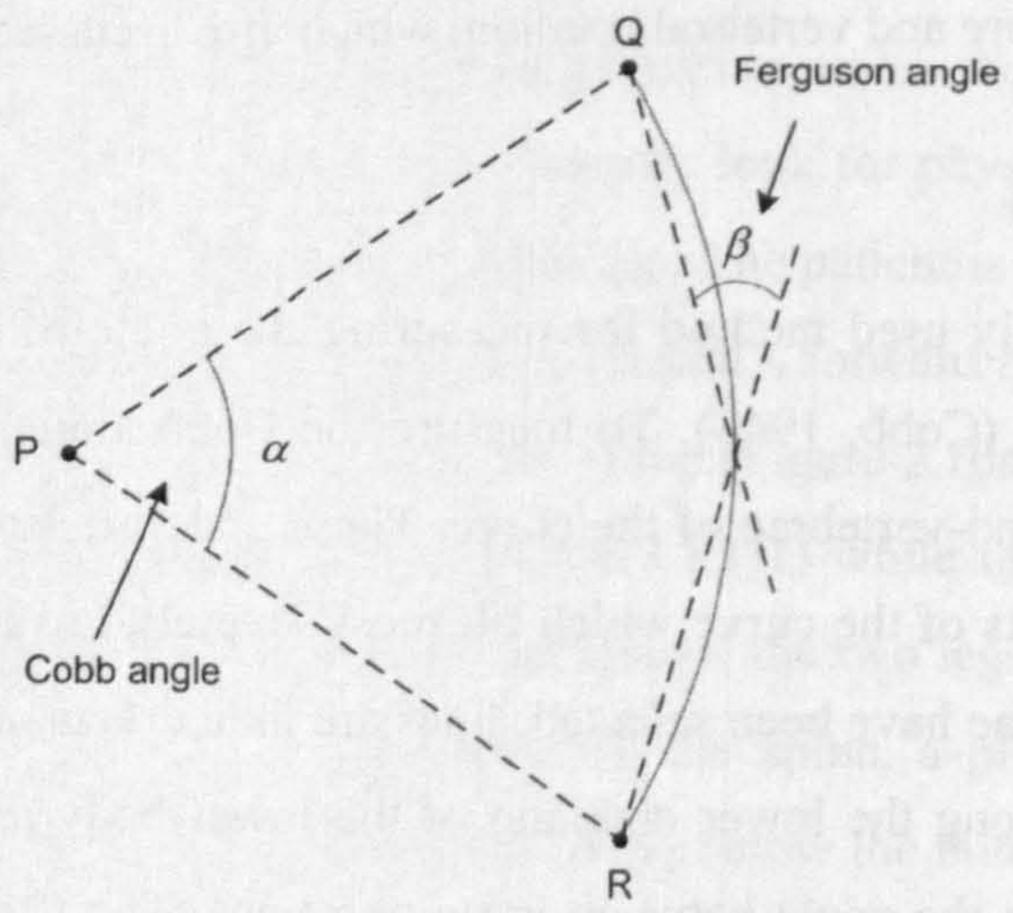


Figure 2.12: Relationship between Cobb and Ferguson angles

However, this relationship only holds if the spine can be approximated by a circular arc. Previous studies of actual spinal curves found the ratio between Cobb and Ferguson angles to be approximately 1.35—1.38 (Robinson et al, 1983; Stokes et al, 1993). This confirms that a

simple circular curve is inadequate approximation of the spinal curve. In Section 6.6, we use constrained polynomials to derive improved approximation of the back midline.

Vertebral rotation

As mentioned earlier, structural scoliosis is characterised by not only lateral curvature of the spine but also rotation of the vertebrae. The measurement of vertebral rotation thus provides a complementary means for assessing scoliotic deformity. However, since the radiograph is a two-dimensional projection of the spine with limited intensity, the measurement of vertebra rotation is generally difficult. One of the earliest methods for estimating vertebral rotation was developed by Nash and Moe (1969), who based their measurements on the position of the pedicles relative to the vertebral body line. The pedicles are two short, thick processes, which project backward, one on either side, from the upper part of the vertebral body, which can be seen on the radiographs (see Figure 2.2). When the spine is straight, the pedicles are projected symmetrically onto the radiographs. As the vertebra rotates, the pedicle on the convex side of the curve moves towards the concave side of the curve whiles the pedicle on the concave side of the curve disappears (Figure 2.13).

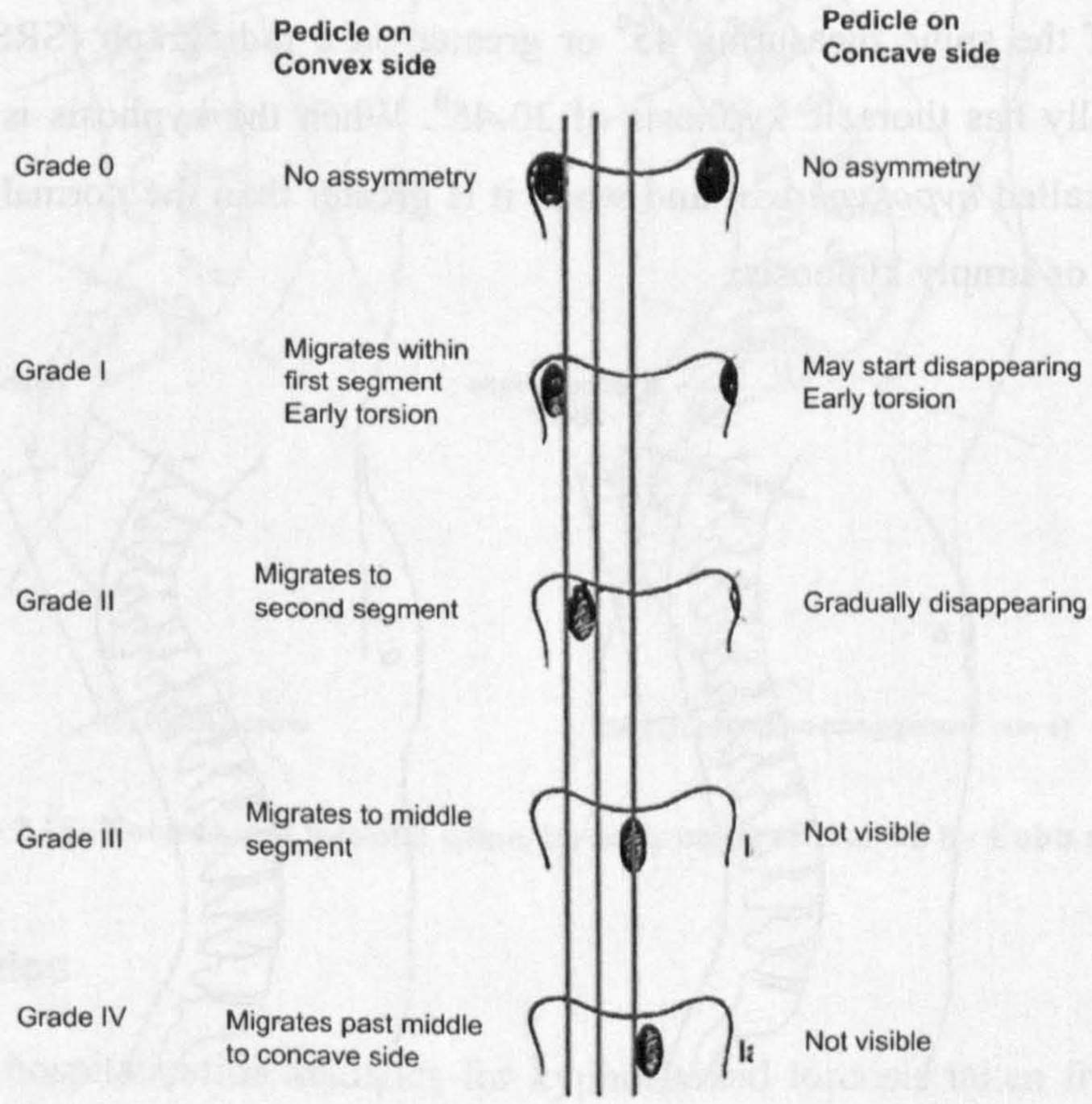


Figure 2.13: Nash and Moe method for measurement of vertebral rotation (Nash and Moe, 1969)

Nash and Moe used this movement of the pedicles to categorise the rotation of the vertebrae into five grades, namely, grade 0, grades I through IV, and assigned degrees of rotation to each category (Table 2.3).

Table 2.3: Rotation assigned to each category in Nash and Moe method

Grade	0	I	II	III	IV
Rotation	0°	5°-10°	15°-30°	35°-40°	>50°

Other methods for measurement of vertebral rotation from radiographs include those proposed by Cobb (1948) who graded the vertebral rotations based on the dislocation of the spinous processes in relation to the vertebral body. Modifications to the Nash and Moe's method have also been suggested (Drerup, 1984).

2.4.1.8. Kyphosis and Lordosis

Kyphosis refers to an outward curvature of the thoracic spine in the sagittal plane which can result in a noticeable round back or hunchback deformity (Figure 2.14). The term is usually used to describe an exaggerated curve. The Scoliosis Research Society defines kyphosis as a curvature of the spine measuring 45^0 or greater on a radiograph (SRS, 2000). The normal spine typically has thoracic kyphosis of $20\text{-}45^0$. When the kyphosis is less than the normal range, it is called *hypokyphosis*; and when it is greater than the normal range, it's called *hyperkyphosis* or simply kyphosis.

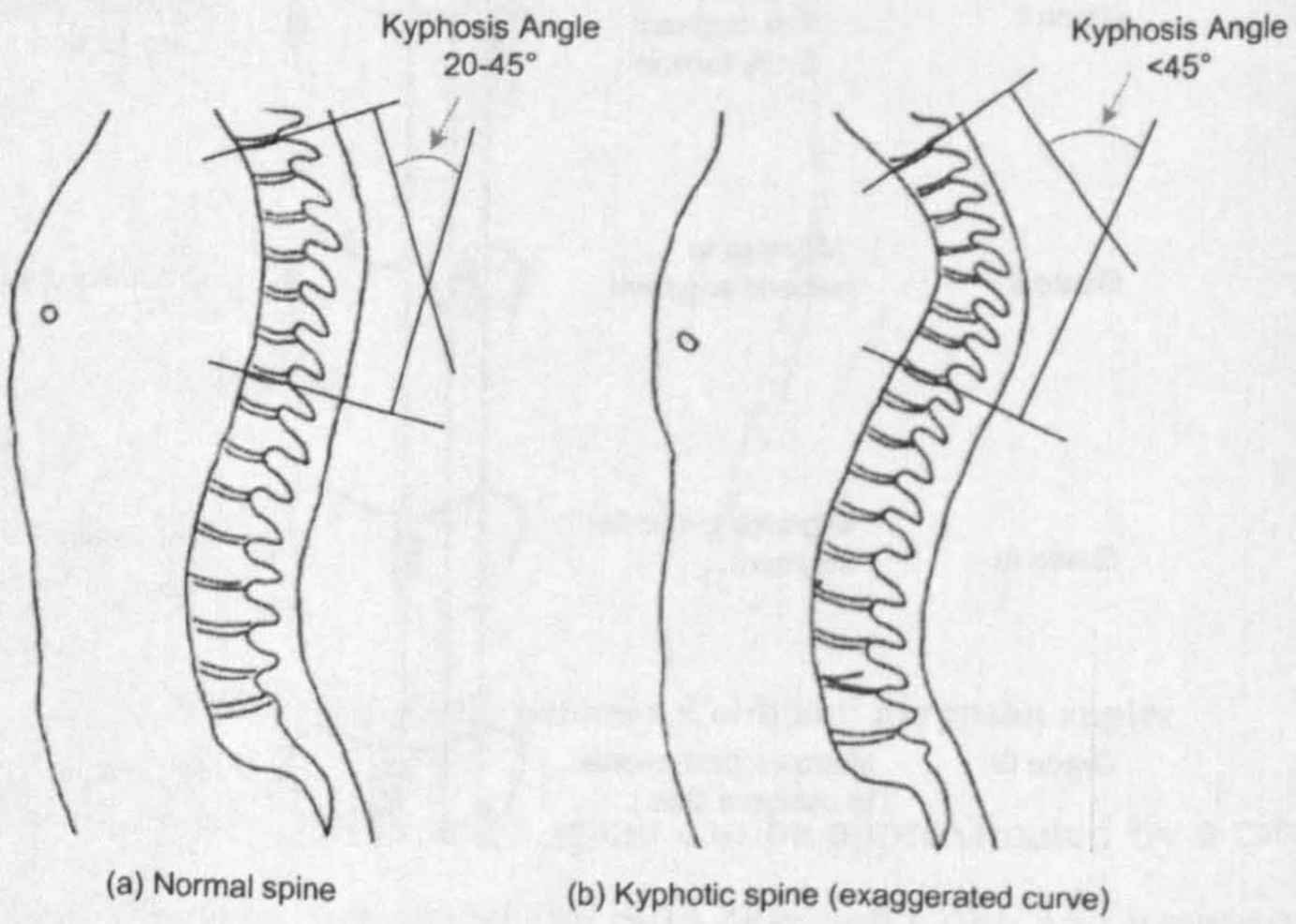


Figure 2.14: Normal and kyphotic spines, kyphosis angle measured by Cobb method

Postural kyphosis is the most common type and represents an exaggerated, but flexible, increase of the natural curve of the spine. This usually becomes noticeable during adolescence and it's more common among girls than boys and rarely causes pain. The condition is not associated with vertebral abnormality and the deformity is completely correctable by change in posture. *Congenital kyphosis* is secondary to congenital vertebral abnormality such as failure of formation, failure of segmentation or a combination of both. Structural kyphosis is associated with other conditions such as Scheuermann's disease, neuromuscular conditions, myelomeningocele, trauma, metabolic problems, skeletal dysplasias, etc. (McAllister, 1975).

Lordosis is the opposite of kyphosis and refers to an inward curvature of the lumbar spine in the sagittal plane producing what is known as the *swayback* (Figure 2.15). The term lordosis is used when the curve is exaggerated. *Hyperlordosis* is a lordosis greater than the normal range; and *hypolordosis* a lordosis less than the normal range. Lordosis tends to make the buttocks appear more prominent. Causes of lordosis include conditions such as congenital, neuromuscular, spondylolisthesis, skeletal diasplasia (McAllister, 1975).

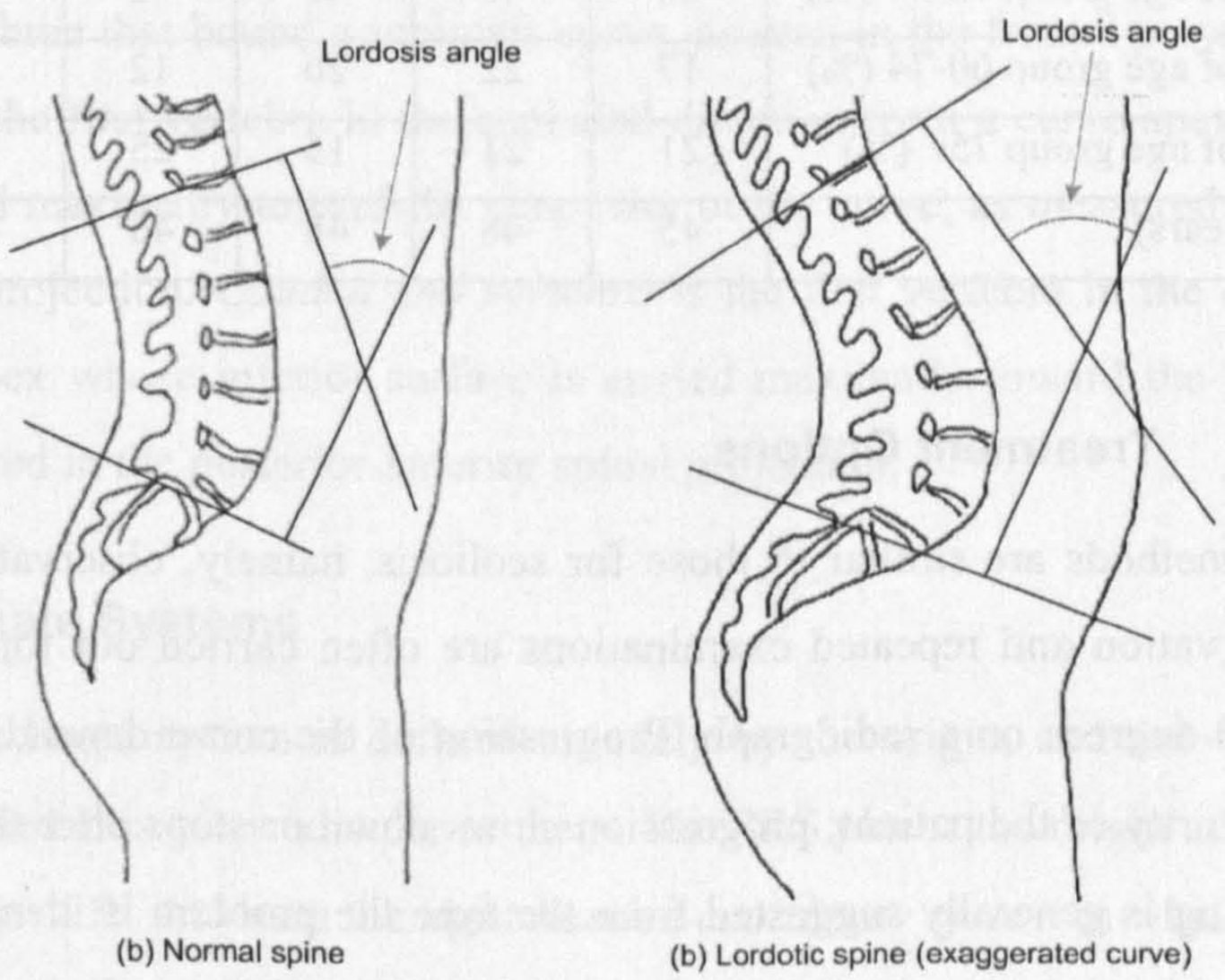


Figure 2.15: Normal and lordotic spine, lordosis angle measured by Cobb method

2.4.1.9. Statistics

Table 2.4 gives hospitalization statistics for kyphosis and lordosis taken from UK National Hospital Episode Statistics over the 1998-2004 periods. These statistics indicate kyphosis and lordosis affect only a small percentage of the population. Generally, an increase in thoracic

kyphosis and a decrease in lumbar lordosis occur with advancing age and are more pronounced in females. The mean age is just over 45 with 54% been females. Subjects with greater lumbar lordosis generally have greater thoracic kyphosis and vice versa.

Table 2.4: Kyphosis and lordosis statistics (Source: Hospital Episode Statistics, 1998-2004, Department of Health, UK)

Year	98-99	99-00	00-01	01-02	02-03	03-04	Mean
Hospital consultant episodes	248	257	219	229	256	217	238
Percentage of hospital adm. (%)	91	94	94	87	93	89	91
Percentage of Males (%)	46	48	45	43	47	46	46
Percentage of Females (%)	54	52	55	57	53	54	54
Percentage of emergency adm. (%)	17	14	12	16	18	16	16
Waiting list (%)	58	68	65	61	60	58	62
Mean waiting time (days)	175	136	159	137	159	134	155
Mean length of stay (days)	11	12	13	12	13	12	12
Percentage of age group 0-14 (%)	16	11	11	15	9	14	13
Percentage of age group 15-59 (%)	46	46	48	48	52	48	48
Percentage of age group 60-74 (%)	17	22	26	12	18	16	19
Percentage of age group 75+ (%)	21	21	15	25	20	22	21
Mean age (years)	45	48	48	46	48	47	47

2.4.1.10. Treatment Options

Treatment methods are similar to those for scoliosis, namely, observation, bracing and surgery. Observation and repeated examinations are often carried out for curves that measure less than 60 degrees on a radiograph. Progression of the curve depends upon the amount of skeletal maturity of the patient; progression slows down or stops after the patient reaches puberty. Bracing is generally suggested from the time the problem is identified until the end of spinal growth and maturation. For cases that cannot be managed non-surgically, spinal instrumentation and fusion are alternatives. Several different types of surgery are available to both straighten the spinal curve and relieve the patient's pain.

2.5. Three-dimensional Geometric Description of Spinal Deformities

The traditional definition of scoliosis as a lateral spinal curvature does not provide adequate description of the three-dimensional nature of the deformity. For the purpose of research and diagnosis, a more complete description is required. The SRS has developed three-dimensional geometric terminology of the spine which provides the basis for deriving surface measurements. In this thesis, we build on this terminology to derive various surface measurements for the back midline.

2.5.1. Basic Vertebra Definitions

A *vertebra centroid* is defined by SRS as the mid point of a vertebral body and is defined as the point half way between the centres of the two endplates of a vertebra. *Vertebra bodyline* is the 3-D curved line that passes through the centroids of the vertebral bodies. *Apical vertebra* or disc is the most laterally deviated vertebra or disc in a scoliosis curve i.e. the vertebra that has the greatest y coordinate in the global coordinate system. An *end vertebra* is the cephalad and caudal vertebrae that bound a scoliosis curve, as seen in the frontal projection. *Cephalad end vertebra* is the first vertebra in the cephalad direction from a curve apex whose superior surface is angled maximally toward the concavity of the curve, as measured in the posterior-anterior spinal projection. *Caudad end vertebra* is the first vertebra in the caudad direction from a curve apex whose inferior surface is angled maximally toward the concavity of the curve, as measured in the posterior-anterior spinal projection.

2.5.2. Coordinate Systems

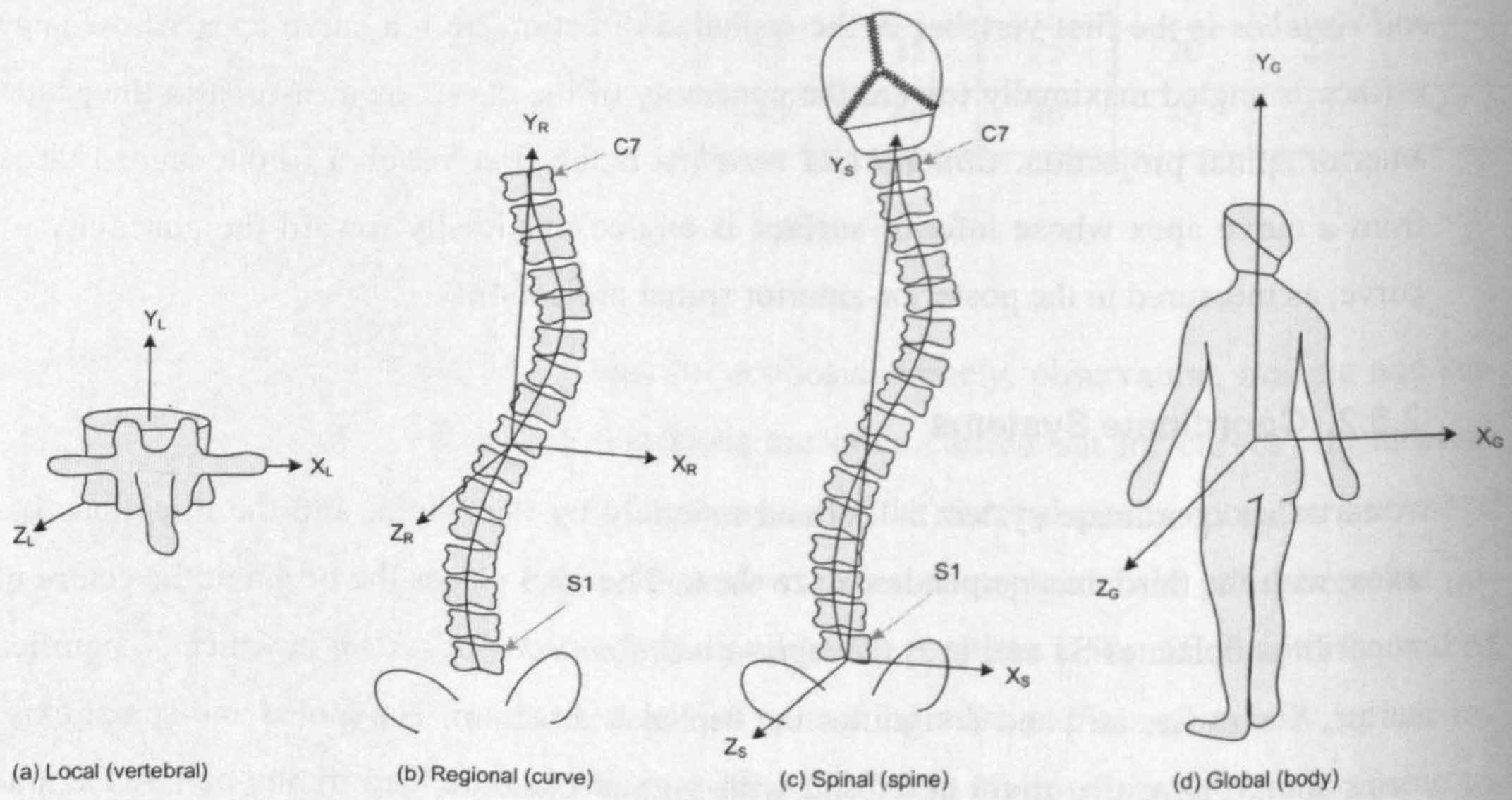
A cartesian coordinate system is defined uniquely by the origin, and the directions of two axes, with the third axis perpendicular to them. The SRS places the origin at the centre of the superior endplate of S1 and uses the *right-handed coordinate system* in which X signifies anterior, Y signifies left, and Z signifies the cephalad direction. The global and spinal axis systems should have the origin at S1 and with sagittal plane defined by the pelvis, the anterior superior iliac spines (ASIS) defining the transverse global (Y) direction. The other principal directions are aligned either with gravity (global system), or with spinal landmarks. In this investigation, we use the *left-handed coordinate system* and define four coordinate systems as follows (Figure 2.16):

- *Local coordinate system*. A vertebra-based coordinate system with the origin at the centroid of the vertebral body, Y-axis passing through the centres of the upper and lower endplates, and X-axis is parallel to a line joining similar landmarks on the bases of the right and left pedicles. This coordinate system is denoted in this thesis as $X_L Y_L Z_L$.

- *Regional coordinate system*. A curve based reference system with the Y-axis passing through the end vertebrae of a curve, denoted as $X_R Y_R Z_R$.

- *Spinal coordinate system*. A reference system for the entire spine which has its origin at S1, the Y-axis passing through C7, and the X-axis parallel to the line connecting the PSISs. This system is denoted as $X_S Y_S Z_S$.

- *Global coordinate system*. Conventional anatomic gravity-based axis system, with the origin at S1, the Y-axis is vertical (gravity line) and the XY-axis parallel to the line connecting the PSISs. This system is denoted as $X_G Y_G Z_G$.



(a) Local (vertebral) (b) Regional (curve) (c) Spinal (spine) (d) Global (body)

Figure 2.16: Coordinate systems used for scoliosis description

2.5.3. Regional and Global Planes

The plane which best accommodates the vertebrae in a specified region of the spine is known as the *best fit plane*. The vertical plane which shows the greatest spinal curvature by a specified method (e.g. Cobb) when a specified part of the spine is projected onto it is known as *plane of maximum curvature*. Similarly, the vertical plane with the minimum spinal curvature is known as *plane of minimum curvature*. The conventional global planes of the body are follows (Figure 2.17):

- Frontal or coronal plane (global XY-plane)
- Sagittal or lateral plane (global YZ-plane)
- Transverse or axial plane (global XZ-plane)

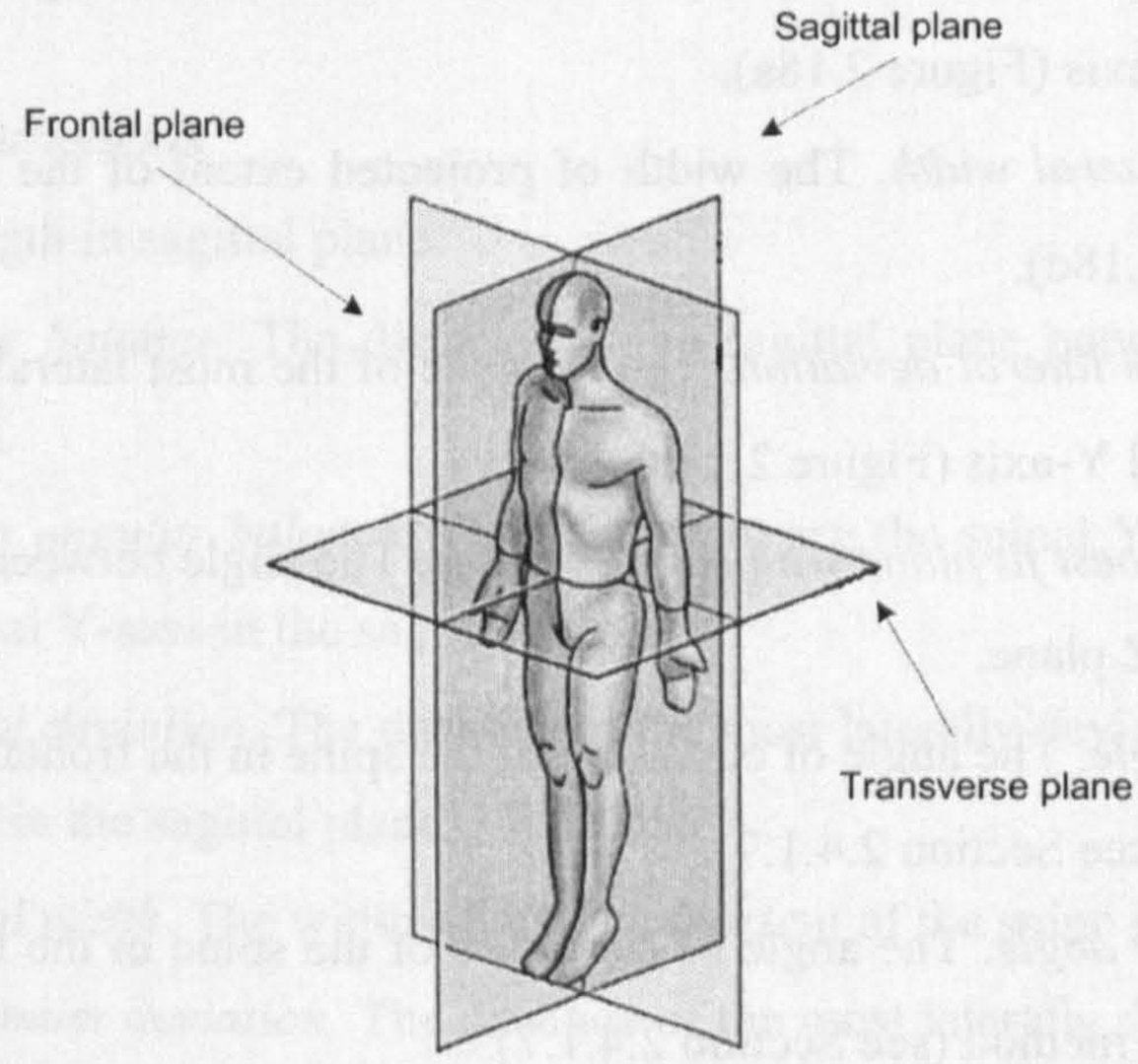


Figure 2.17: Global planes of the body (adapted from Bridwell, 2001a)

2.5.4. Spinal Measurements

The spinal measurements consist of linear, angular, area and volume measurements and have been grouped under two-dimensional (3D) and three-dimensional (2D) measurements. The 3D measurements are derived in 3D space, for example, 3D spinal length and volume of spinal bounding box. The 2D measurements are performed in 2D space defined by a specified plane of the body i.e. frontal, sagittal or transverse. The following is a summary of the relevant measurements derived in this investigation.

3D Spinal Measurements

- *3D spinal length*. The 3D length of the spine or a specified region of the spine.
- *Volume*. The volume of the 3D bounding box of the spine.

Frontal Plane Measurements

- Two-dimensional spinal length in frontal plane.
- *Frontal plane balance*. The distance in the frontal plane between C7 vertebra and the global Y-axis (Figure 2.18c).
- *Frontal plane angular balance*. The angle between spinal y-axis and global Y-axis in the frontal plane (Figure 2.18c).
- *Spinal lateral deviation*. The distance of the most laterally deviated vertebra from the spinal y-axis (Figure 2.18a).
- *Spinal lateral width*. The width of projected extent of the spine in the frontal plane (Figure 2.18d).
- *Maximum lateral deviation*. The distance of the most laterally deviated vertebra from the global Y-axis (Figure 2.18b).
- *Angle of best fit plane with sagittal plane*. The angle between the best fit plane and the global YZ plane.
- *Cobb angle*. The angle of curvature of the spine in the frontal plane measured by Cobb method (see Section 2.4.1.7).
- *Ferguson angle*. The angle of curvature of the spine in the frontal plane measured by Ferguson method (see Section 2.4.1.7).
- Projected area of the spine in the frontal plane (see Section 2.5.5).

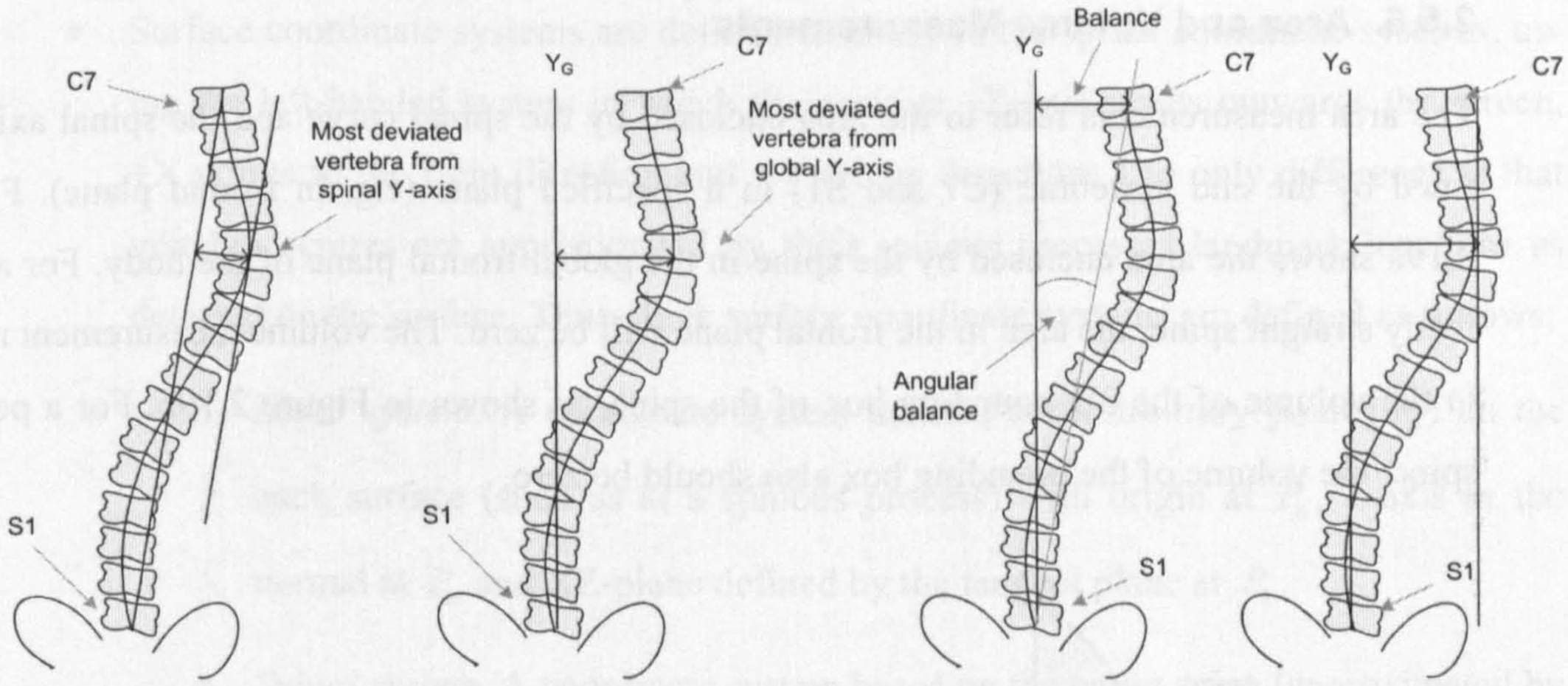


Figure 2.18: Spinal measurements in the frontal plane

Sagittal Plane Measurements

- 2D spinal length in sagittal plane.
- *Sagittal plane balance*. The distance in the sagittal plane between C7 vertebra and global Y-axis.
- *Sagittal plane angular balance*. The angle between the spinal Y-axis and the vertical from the global Y-axis in the sagittal plane.
- *Sagittal spinal deviation*. The distance of the most laterally deviated vertebra from the spinal Y-axis in the sagittal plane.
- *Sagittal spinal width*. The width of projected extent of the spine in the Sagittal plane.
- *Sagittal maximum deviation*. The distance of the most laterally deviated vertebra from the global Y-axis in the sagittal plane.
- *Angle of best fit line with frontal plane*. The angle between the best fit plane and the global XY plane.
- *Kyphosis angle* by Cobb or Ferguson method (Section 2.4.1.7).
- *Lordosis angle* by Cobb or Ferguson method (Section 2.4.1.7).
- Projected area of the spine in the sagittal plane (Section 2.5.5).

Transverse Plane Measurements

- 2D Spinal Length in transverse plane
- Projected area of the spine in the transverse plane

2.5.5. Area and Volume Measurements

The area measurements refer to the area enclosed by the spinal curve and the spinal axis defined by the end vertebrae (C7 and S1) in a specified plane (e.g. in frontal plane). Figure 2.19a shows the area enclosed by the spine in the global frontal plane of the body. For a perfectly straight spine, the area in the frontal plane will be zero. The volume measurement refers to the volume of the 3D bounding box of the spine, as shown in Figure 2.19b. For a perfect spine, the volume of the bounding box also should be zero.

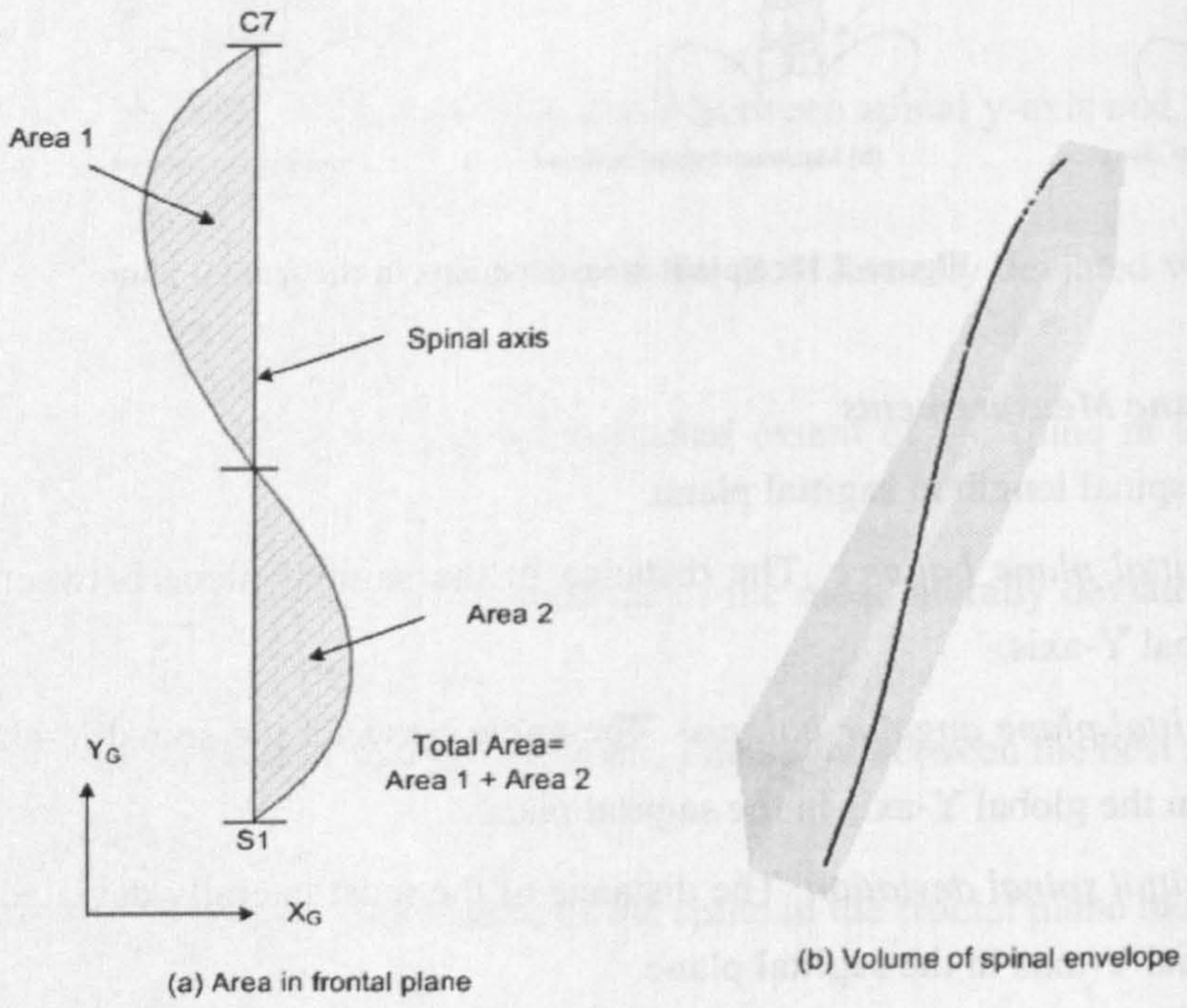


Figure 2.19: Area and volume measurements

2.5.6. Surface Measurements

In the computation of all surface measurements, the following general assumption is made—the location or centroid of each vertebra may be adequately approximated by the projection of its spinous process on the surface of the back. With this assumption, the following statements can be made:

- The vertebral body line can be approximated by the back midline connecting the spinous process landmarks detected on the surface of the back.
- The orientation of each vertebra can be approximated by the normal vector to the surface at the spinous process landmark of that vertebra.

- Surface coordinate systems are defined similarly to the spinal coordinate systems, using the left-handed system in which the positive +Z-axis points outwards the screen, +X points to the right direction and +Y the up direction. The only difference is that vertebra centres are approximated by their spinous processes landmark locations as detected on the surface. Thus, three surface coordinate systems are defined as follows:
    - *Local system*. A coordinate system defined at an arbitrary point (P_0) on the back surface (such as at a spinous process) with origin at P_0, Z-axis as the normal at P_0 and XZ-plane defined by the tangent plane at P_0.
    - *Spinal system*. A coordinate system based on the entire spine (approximated by the midline). Origin at S1 spinous process landmark, Y axis passing through C7 landmark, and X axis parallel to the line connecting the left and right PSIS landmarks.
    - *Global system*. A reference system formed by the three global planes of the body, same as defined in Section 2.5.2.

2.5.7. Pelvis Orientations

The relevant measurements with respect to the pelvis are the three orientation angles, namely, pelvic obliquity (rotation about global Z-axis), pelvic tilt or inclination (rotation about global Y-axis) and pelvic rotation (rotation about global X-axis). These are illustrated in Figure 2.20, in which $X_G Y_G Z_G$ represent the global axes of the body (see Sections 2.5.2 and 2.5.3).

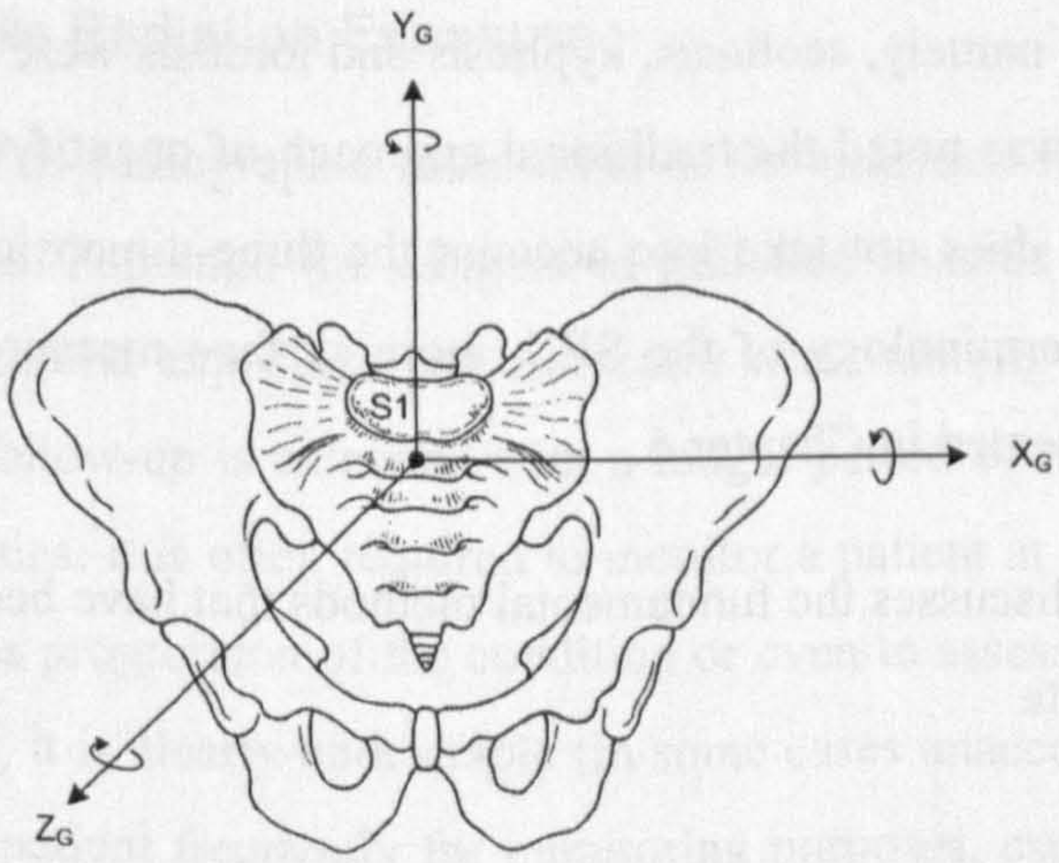
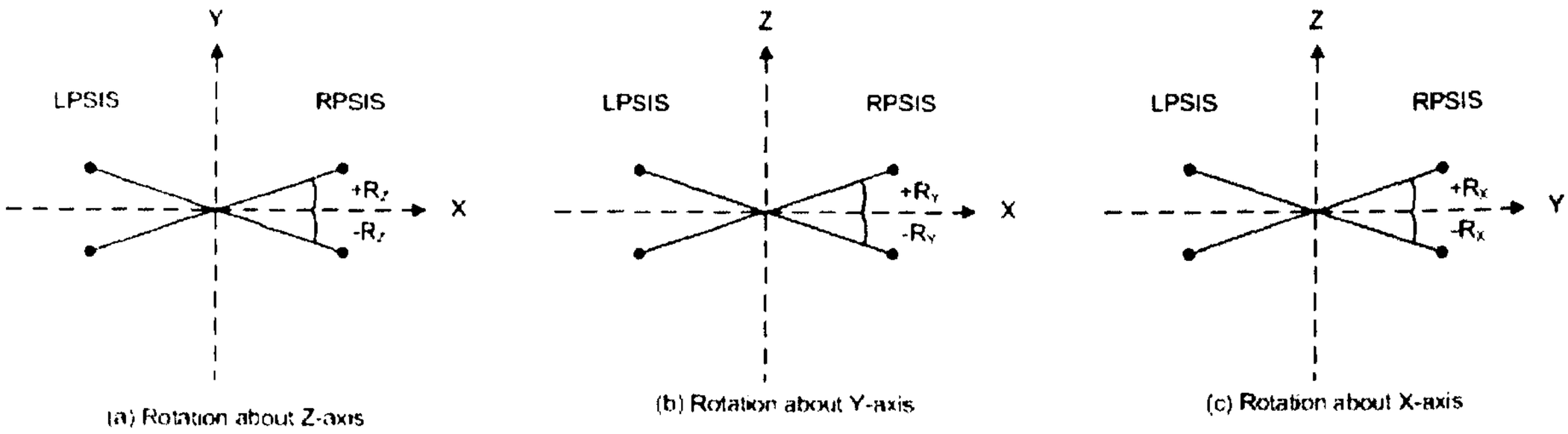


Figure 2.20: Orientation angles of the pelvis

Approximate orientations for the pelvis may be computed based on locations of the left and right PSIS landmarks. In Figure 2.21, XYZ represent the axes of the global coordinate system. The rotation angles can be computed by using the set of Equations 2.1.

$$R_Z = \tan^{-1}\left(\frac{dy}{dx}\right) \qquad R_Y = \tan^{-1}\left(\frac{dz}{dx}\right) \qquad R_X = \tan^{-1}\left(\frac{dz}{dy}\right) \tag{2.1}$$

where R_X, R_Y, R_Z are the rotation angles about the three global axes and dx, dy, dz are the differences in the coordinates of the left and right PSIS landmarks.



(a) Rotation about Z-axis (b) Rotation about Y-axis (c) Rotation about X-axis

Figure 2.21: Computation of pelvis orientations based on PSIS landmarks

2.6. Chapter Summary

This chapter has described relevant background to spinal and back deformities. The basic anatomy of the spine was discussed, including the visual and palpable landmarks of the back were discussed. These landmarks are of particular interest in this investigation since the success of any surface analysis techniques to a large extent depends on them. The three curvatures of the spine, namely, scoliosis, kyphosis and lordosis were discussed. In the last section of the chapter, it was noted the traditional approach of quantifying spinal deformities by the Cobb angle alone does not take into account the three-dimensional nature of the deformity. Building on the terminology of the SRS, more surface measurements were derived, which have been implemented in Chapter 6.

The next chapter discusses the fundamental methods that have been developed for acquisition of back surface data

.

3. Surface Measurement Methods for the Assessment of Spinal and Back Deformities

3.1. Introduction

This chapter discusses the fundamental surface measurement methods that have been developed for assessment of back deformities. In the past few decades, various surface measurement system and imaging modalities have been proposed. These range from tactile devices (such as the Scoliometer), non-contact structured light methods (such as moiré fringe topography, phase measuring methods, rasterstereography) to three-dimensional scanning systems (such as the ISIS). For each method, the underlying theory is briefly outlined and its application to back shape measurement is discussed including the main limitations. Recent advances in hardware and software technology have now enabled fast and accurate acquisition of both image and coordinates. The last section discusses three-dimensional laser scanners and their application to back shape measurement. The Minolta laser scanner, which has been used for data collection in this investigation, is also discussed.

3.2. Why Surface Measurement?

There are three primary reasons for assessing spinal deformities from surface measurements; these are discussed in this section.

3.2.1. Risk of Multiple Radiation Exposure

The obvious drawback of radiographic assessment is the radiation hazards associated with exposure to radiographs. Although the dangers of radiation hazards have been reduced by modern techniques, repeated exposures can still carry accumulating risk over the patient's lifetime, especially if follow-up is extended over a longer period of time. During the assessment of spinal deformities, it is often required to monitor a patient at regular intervals over a period of time, to assess progression of the condition or even to assess the effect of treatment methods. In such cases, it is clearly undesirable (in some cases unacceptable) to take full spinal radiographs of the patient frequently for monitoring purposes, especially in teenage girls in whom scoliosis has a high incidence. It is therefore desirable to explore surface methods for the assessment and monitoring of spinal deformities.

3.2.2. Cosmetic Deformity of the Back Surface

Structural deformity of the spine such as scoliosis will almost certainly result in corresponding deformity of the back or trunk surface (see Section 2.4.1.3). Studies indicate that it is the cosmetic deformity of the trunk surface which is more important to the patient and family, and is what frequently motivates them to seek treatment (Mahood et al, 1995). The patient is most aware of the visible deformity of their body such as the rib hump and may seek treatment for these reasons or because they are uncertain about the nature and prognosis of their condition. Clinically, the back or trunk surface is important not only because of its cosmetic relevance, but also because it is the interface for some treatment methods such as bracing. Assessment methods should therefore allow the quantification of the surface deformity in addition to the spinal deformity.

3.2.3. Need for Three-Dimensional Measurements

One problem with the measurement of the Cobb angle is that it is obtained from a two-dimensional, low resolution radiographic image of the spine. Consequently, it is usually difficult to measure the Cobb angle in a consistent manner. In fact, studies have shown significant variations in measurement of the Cobb angle—inter-observer variation is approximately $7\text{-}9^0$ at 95% confidence interval while intra-observer variation is about $4\text{-}5^0$ (Goldberg et al, 1988; Morrissy et al, 1990). Clearly, there is the need for a more objective and automated measurement techniques. Second, the Cobb angle alone does not provide a complete description of the three-dimensional nature of the deformity. More complete geometric descriptors (such as those discussed in Section 2.5) are needed to fully quantify the entire deformity. Radiography can offer three-dimensional information only if at least two radiographs are acquired. However, this option will only worsen the radiation problem. Three-dimensional surface measurements methods will allow the derivation of alternative or complementary geometric descriptors.

3.3. Surface Measurements Methods

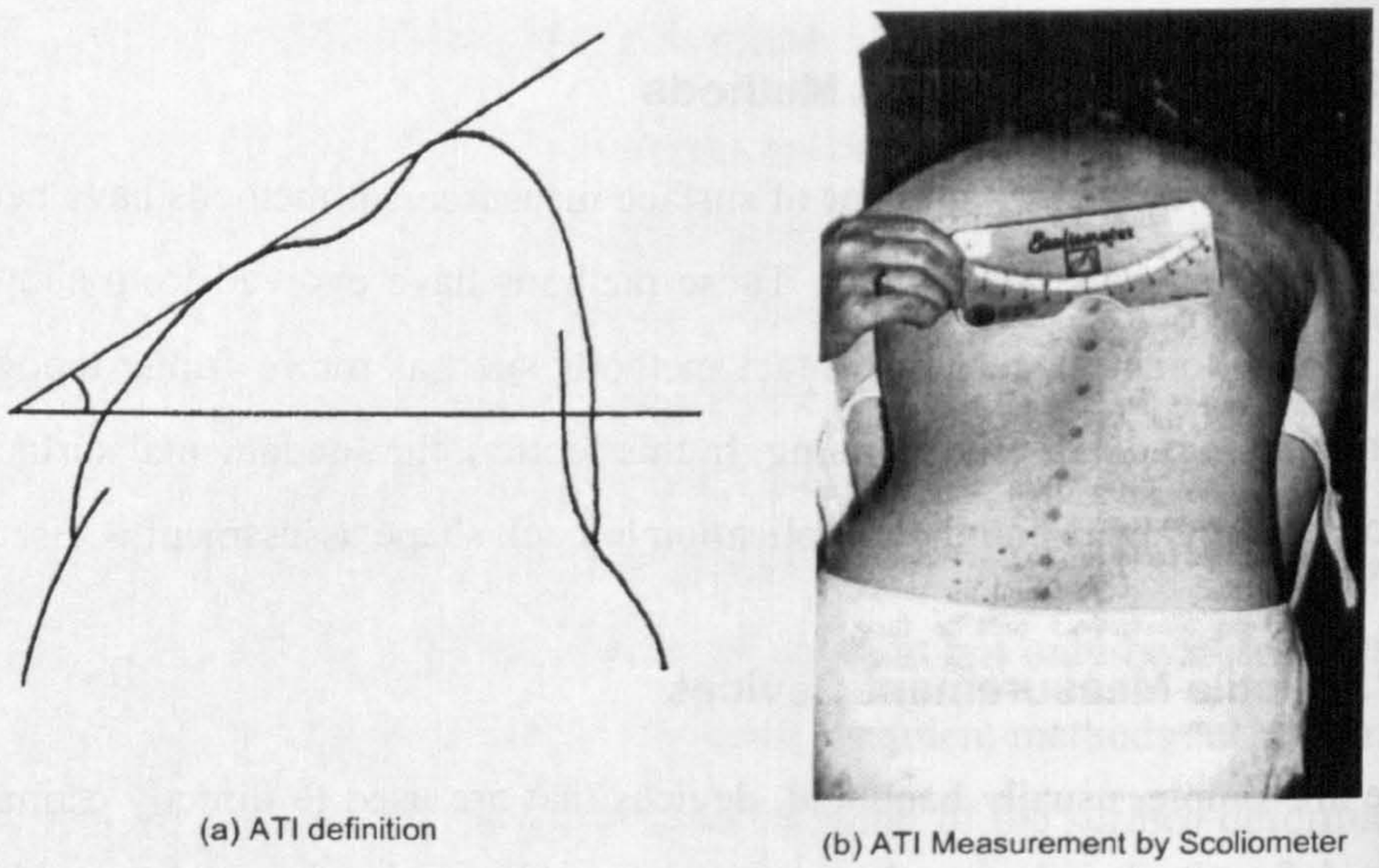
In the past few decades, a number of surface measurement methods have been developed and applied to back shape assessment. These methods have evolved from simple tactile devices (such as Scoliometer), to non-contact methods such as moiré fringe topography, structured light methods, and 3D laser scanning. In this section, the fundamental surface data acquisition methods are reviewed and their application to back shape assessment is discussed.

3.3.1. Tactile Measurement Devices

These are simple, usually handheld, devices that are used to directly estimate the magnitude of the deformity by placing the instrument on the back of the subject. Various instruments have been developed for measuring various aspects of the deformity. For example, the Scoliometer and the Formulator Body Contour Tracer (FBCT) are simple devices used to estimate directly the asymmetry of the back; while the Debrunner Kyphometer is used to measure the kyphosis angle (Debrunner, 1972). The Scoliometer and the FBCT are discussed further in the following sub-sections.

3.3.1.1. Scoliometer

The Scoliometer (also known as inclinometer) is a simple tactile device developed to directly measure what is termed the Angle of Trunk Inclination (ATI). The ATI is defined by the SRS as the angle between the horizontal and the plane across the back at the greatest elevation of a rib prominence or lumbar prominence, as measured by an inclinometer or Scoliometer (SRS, 1980). The ATI is usually measured with the patient in Adam's forward bending position (Figure 3.1).



(a) ATI definition

(b) ATI Measurement by Scoliometer

Figure 3.1: ATI definition and measurement by the Scoliometer

Dangerfield (1992a) extended this simple concept and recorded the ATI at each vertebral position and plotted the results as a graph of ATI against vertebral position. The graph represents the magnitude of scoliosis along the length of the spine. ATI measurements were taken with the patient in standing and prone positions as well, allowing a multi-level ATI map to be constructed and stored in patient records. Figure 3.2 shows a typical ATI map for a scoliosis patient.

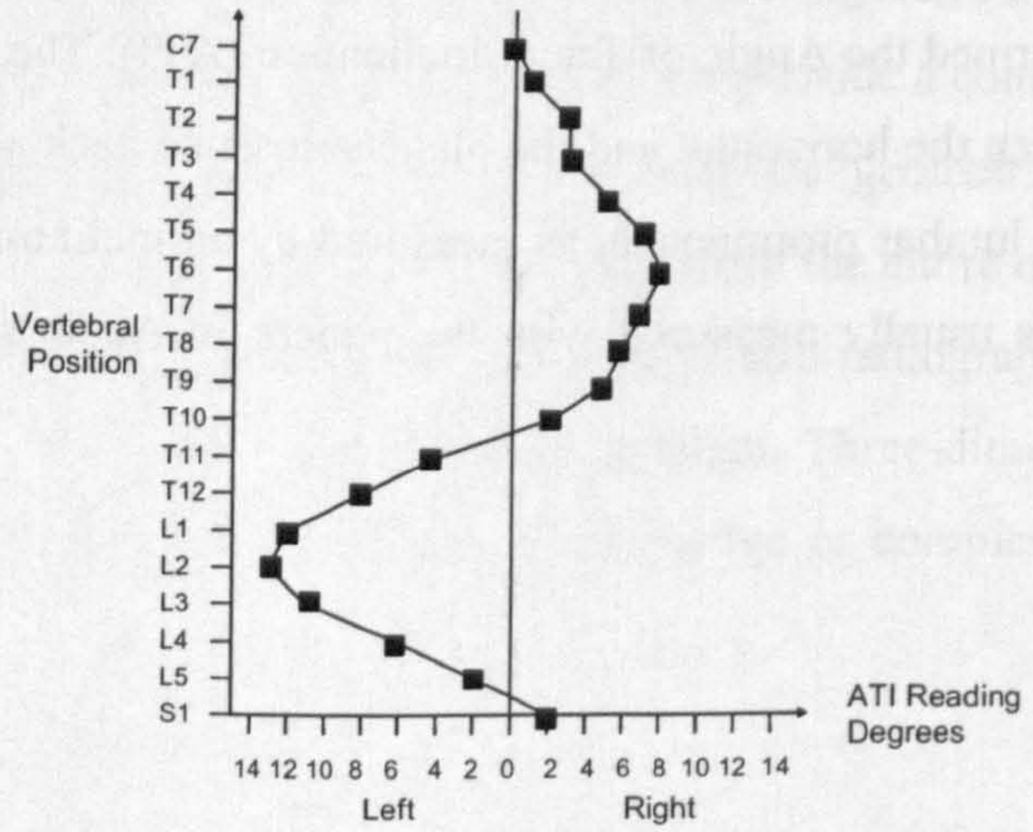


Figure 3.2: A typical ATI map for a scoliosis patient (Dangerfield, 1992a)

3.3.1.2. Formulator Body Contour Tracer

Another tactile device developed by Burwell et al (1983) known as Formulator Body Contour Tracer (FBCT) allowed a more accurate recording of the cross-sectional profiles of the back surface. The FBCT consists of a simple matrix of rods that are placed parallel to the coronal plane across the surface of the back (usually in the forward bending position). The rods assume the profile of the back as shown in Figure 3.3. To use the FBCT, the spinous processes are first palpated and marked with markers. The FGCT is then placed on the back of the subject in Adams's forward bending pose with its central rod coinciding with the spinous process marker. The rods are then allowed to fall under gravity to assume the contour of the back and then locked by tightening a winged nut on their housing and removed from the patient.

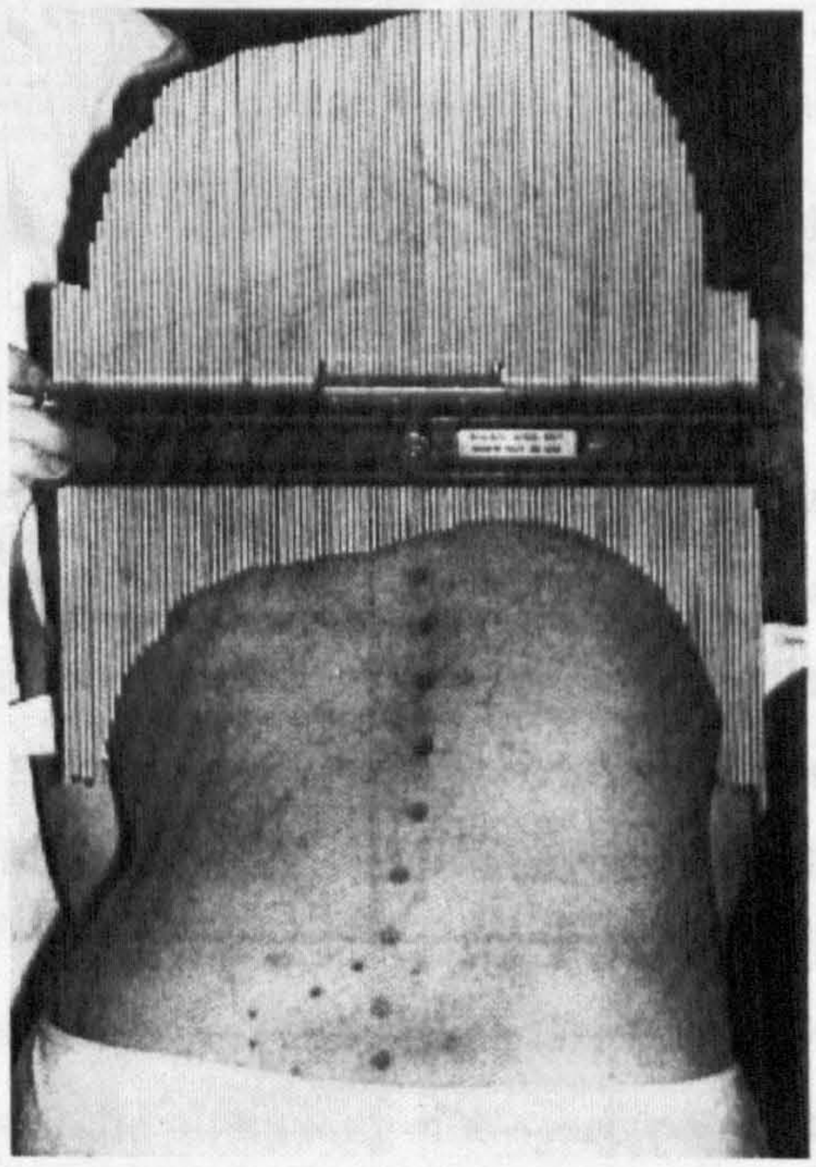


Figure 3.3: Use of FBCT to measure the back profile of a subject (Burwell et al, 1983)

The resulting cross-sectional profile of the FBCT is then transferred onto paper by tracing out the edges of the rods (Figure 3.4). A linear measurement known as Trunk Asymmetry Score (TAS) is calculated by sub-diving each half of the contour profile into five sections and measuring the distance from the contour to a datum line. The differences between corresponding left and right distances are summed, termed Crude Trunk Asymmetry Scores (CTAS). For comparison of different subjects, the CTAS values are standardised to a mean trunk diameter of the study group, giving Standardised Trunk Asymmetry Score (STAS).

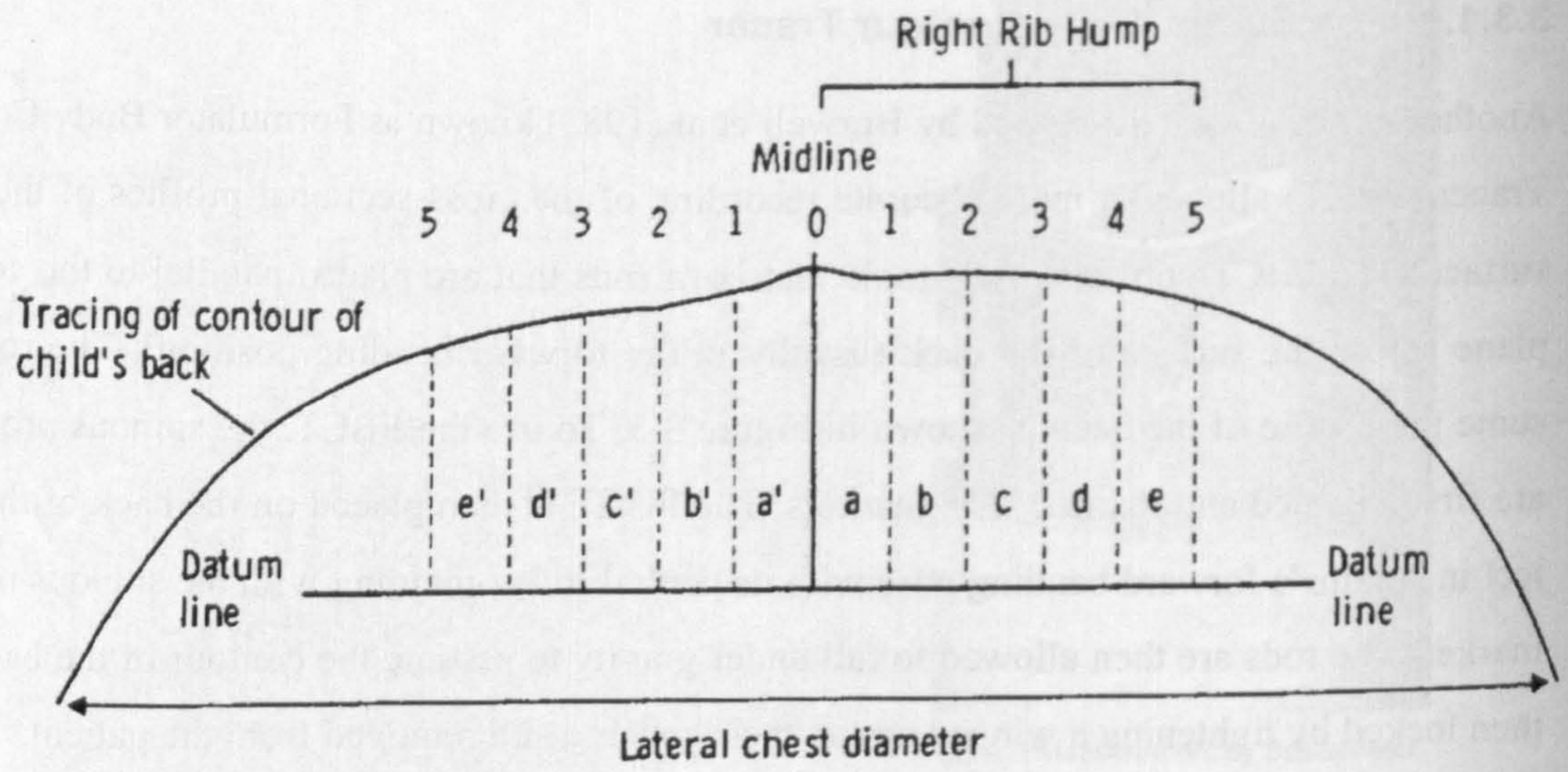


Figure 3.4: Measurement of Trunk Asymmetry Score (TAS) from FBCT profile (Burwell et al, 1983)

The main problem with tactile measurement devices is obvious: discomfort to the subject. If the subject is a patient under treatment, tactile measurement may even cause pain. Moreover, contact of the device with the back may cause the subject to become tense or flinch, resulting in a change of body posture which can affect the measurement (Pearson, 1996). There is also the issue of hygiene and safety which may have to be addressed in the case of contact measurement. In general, tactile devices are less accurate, laborious and time-consuming. For these reasons, non-contact measurements methods are preferred. In the following sections, non-contact optical methods are discussed, namely, moiré topography, structured light approaches, phase-measuring methods and 3D scanning systems.

3.3.2. Moiré Topography

Moiré topography is one of the earliest non-contact optical techniques which gained widespread clinical use in the assessment of the back shape. For example, researchers at Salisbury District Hospital have used moiré topography in routine clinical analysis of the back shape for over fifteen years. A moiré topograph (or topogram) provides a contour map which can be easily understood by a human observer. Figure 3.5 shows typical examples of moiré topographs of scoliosis patients. The asymmetry of the shape of the back can be easily observed from these fringes.

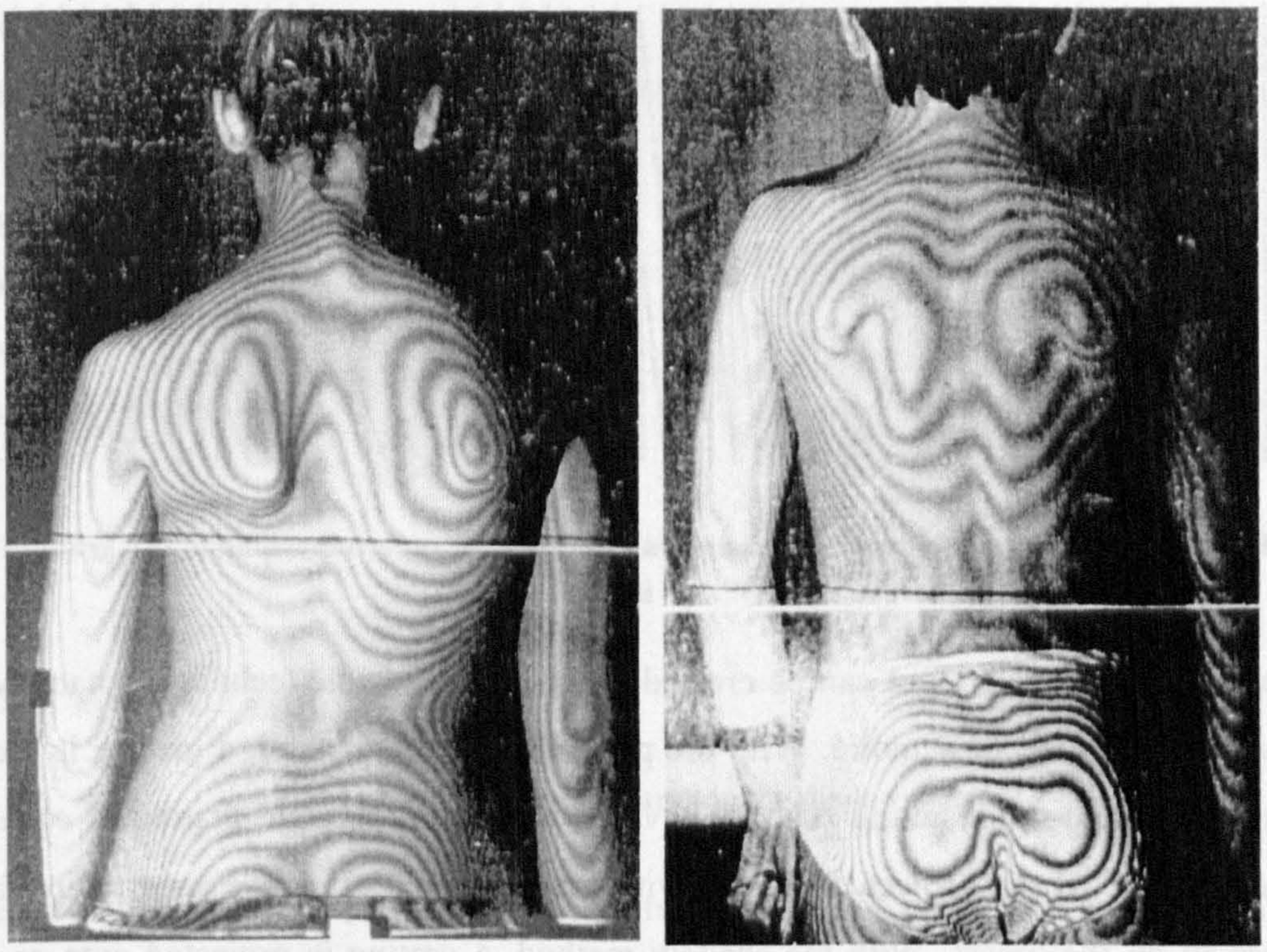


Figure 3.5: Examples of Moiré topograms of scoliosis patients

3.3.2.1. Principle of Moiré

The term moiré is used to describe interference fringes created by the superimposition of two or more geometric patterns. When two gratings are superimposed, moiré fringes are created as a result of the patterns having a difference in line spacing either in pitch or in line orientation. Depending on what type of geometric patterns is used and how they are combined, the moiré fringes may assume various meanings. This review is restricted to moiré fringes produced by the superposition of two periodic alternating line-space patterns, i.e. linear gratings. Figure 3.6 shows examples of the formation of moiré fringes by superimposing two identical gratings, with one at an angle to the vertical. Moiré fringes are observed occurring perpendicularly to the normal. As the angular displacement of the second grating is increased, more fringes are created.

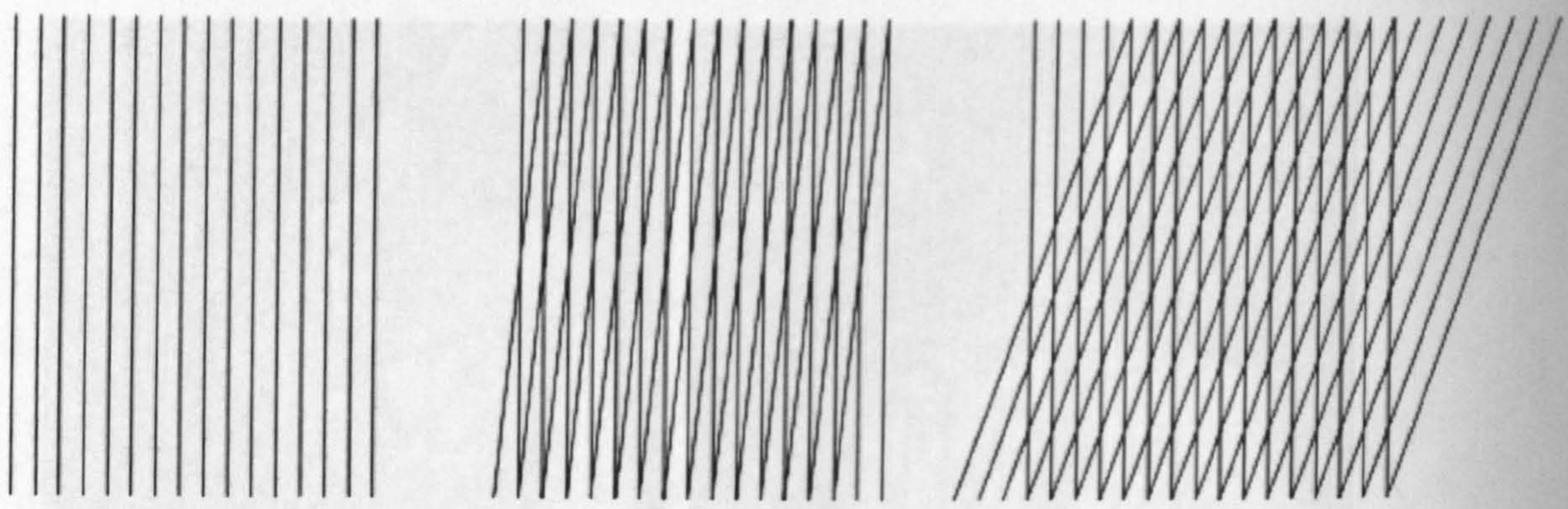


Figure 3.6: Formation of moiré fringes showing the effect of superimposition of two gratings with increasing orientation

In practice, moiré fringes can be created by one of two basic techniques, namely, projection moiré and the shadow moiré. With the projection moiré method, a grating is projected onto the object and the resulting shadow is viewed through a second grating of equal pitch. The interference of the projected grating with the reference grating causes moiré fringes to be formed on the object. With the shadow method, a grating is projected onto an object by a point light source and observed through the grating. The interference of the grating with its shadow produces moiré effects. The shadow moiré techniques is the most common approach used, and thus is discussed further.

3.3.2.2. Shadow Moiré Principle

Figure 3.7 illustrates the basic principle of the shadow moiré method. An equally spaced plane grating is projected onto an object by a point light source and observed through the grating by a camera. The resulting moiré pattern represents a contour line system showing equal depth from the grating if the light source and the observing point lie on a plane parallel to the grating.

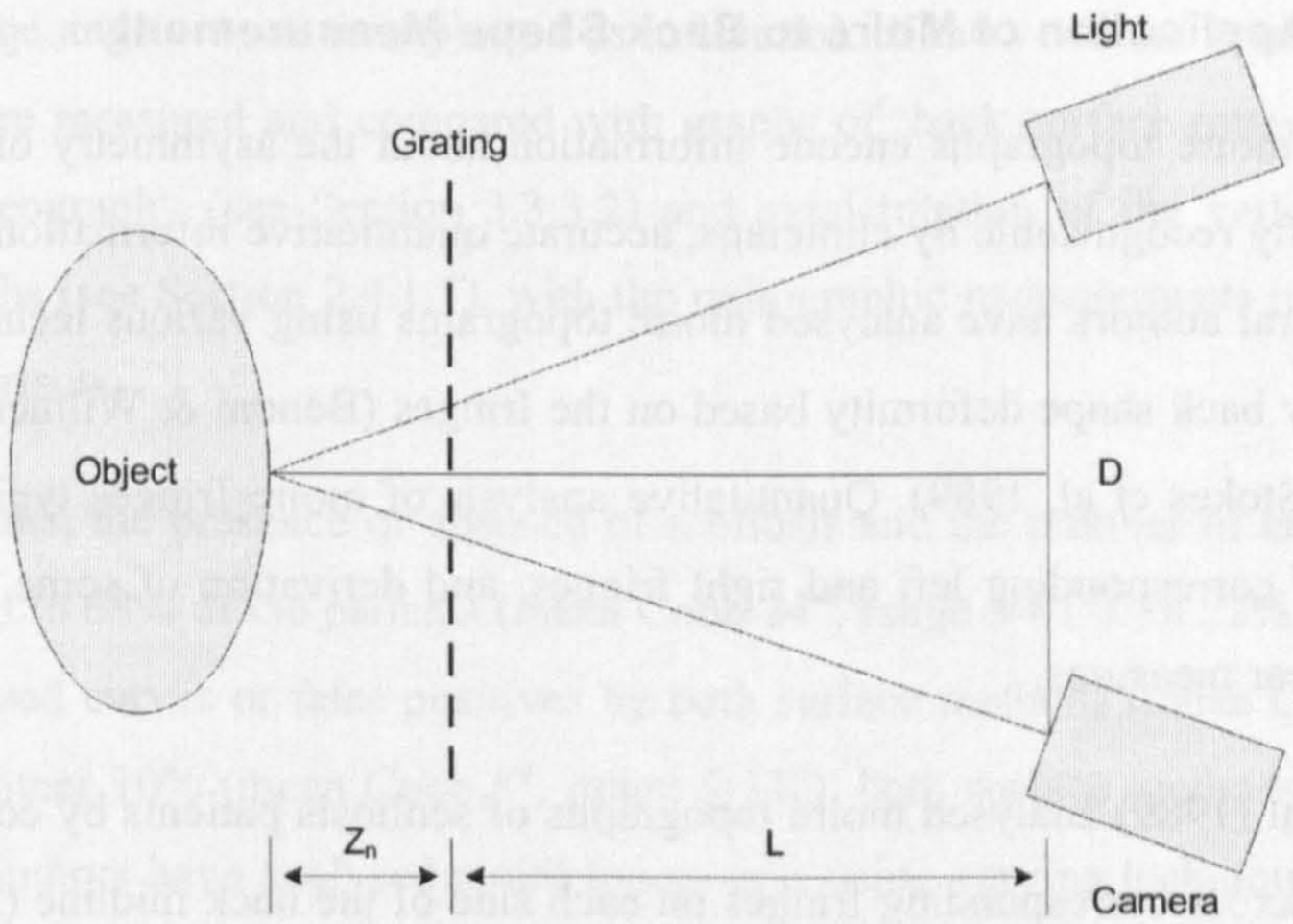


Figure 3.7: Typical configuration for shadow moiré

The height of the n^{th} fringe (Z_n) from the reference grating is given by Equation 3.1, which has been derived elsewhere (Takasaki, 1970; Takasaki, 1981).

$$Z_n = \frac{nL}{(D/p - n)} \tag{3.1}$$

where:

n is the fringe order

p is the pitch of the grating

L is the distance of the light source and the camera from the grating plane

D is distance between the camera and the light source

The separation between the n and $n-1$ fringes is given by Equation 3.2:

$$\Delta Z = \frac{npL}{D - np} - \frac{(n-1)pL}{D-(n-1)p},\ n = 0,1,2... \tag{3.2}$$

It is obvious from Equation 3.2 that the distance between two neighbouring fringes is not constant but gradually increases with the distance to the screen. This rate of increase is different for various instruments, and this gives problems for comparing results from various systems.

3.3.2.3. Application of Moiré to Back Shape Measurement

Although moiré topographs encode information about the asymmetry of the back in a form that is easily recognizable by clinicians, accurate quantitative information is not readily available. Several authors have analysed moiré topograms using various techniques in an attempt to quantify back shape deformity based on the fringes (Benoni & Willner, 1983; Moreland et al, 1983; Stokes et al, 1989). Quantitative analysis of moiré fringes typically involves comparison of corresponding left and right fringes, and derivation of some quantitative angular and/or linear measures.

Stokes et al (1989) analysed moiré topographs of scoliosis patients by constructing lines tangentially across corresponding fringes on each side of the back midline (Figure 3.8). The objective of this study was to determine the extent to which moiré topographs and rasterstereography (Section 3.3.3.2) could be used to detect the presence, magnitude, side and anatomic level of spinal deformity. The angle between each line and the horizontal was measured and termed 'moiré fringe angles'. Each angle was interpreted as a measure of both side-to-side asymmetry and sagittal curvature. The vertical height position of each fringe angle was measured relative to skin marks attached to each patient at levels T1 and L5 spinous processes.

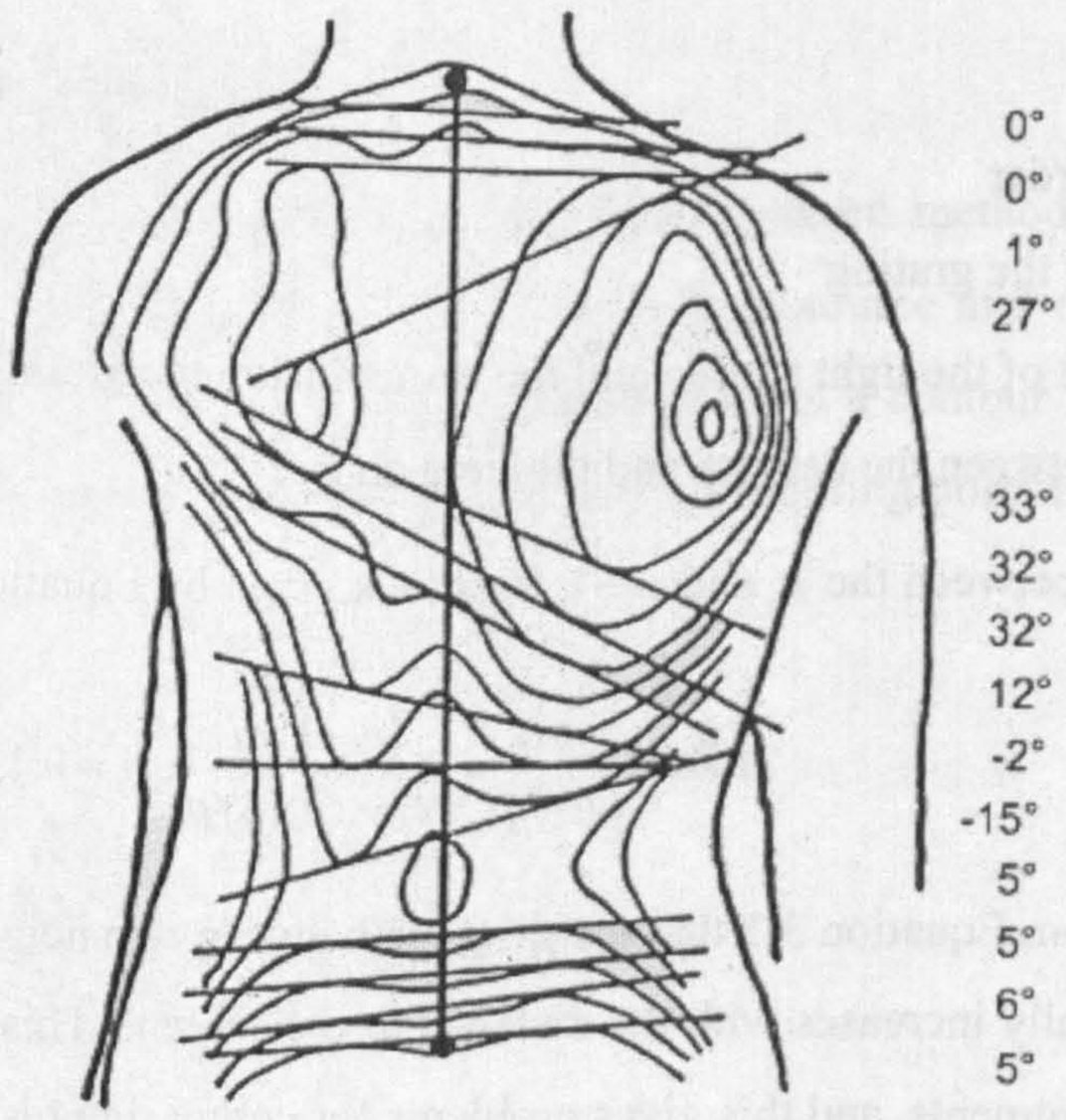


Figure 3.8: Analysis of moiré topograms by Stoke et al (1989). Double tangent lines are drawn for corresponding contours and the angles between these contours and the horizontal measured.

be restricted to those regions by applying a two-level thresholding to isolate regions of high positive and negative curvatures. We define high curvature regions as those regions with curvature outside user-defined standard deviations from the mean or median (Figure 6.2).

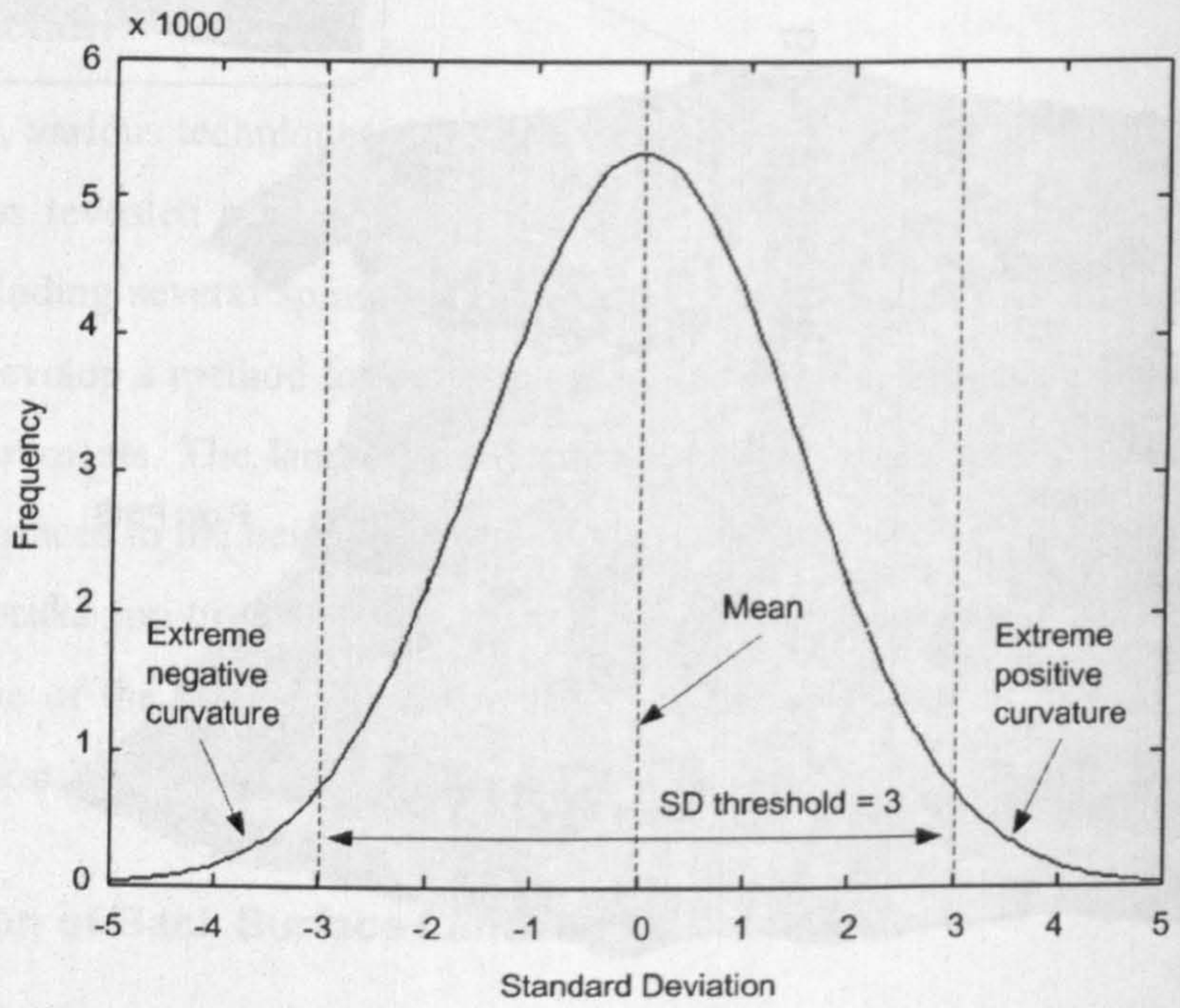


Figure 6.2: Search regions defined by 3 standard deviations from the median

The standard deviation threshold can be adjusted to detect landmarks of specific regions of the back. Figure 6.3 shows application of the two-level thresholding discussed above. The threshold is set to 0.002 standard deviations from the median in (a) and 0.005 standard deviations from the median (b). The median is used instead of the mean because it is less susceptible to outliers. In (a) the threshold is narrow, and thus more regions are detected whiles the larger threshold in (b) detects a few regions.

The moiré fringe angles were then plotted as a function of the vertebral level of the back at which they were measured and compared with graphs of 'back surface rotation' determined from rasterstereography (see Section 3.3.3.2) and axial rotation of the vertebrae measured from radiographs (see Section 2.4.1.7), with the radiographic measurements used as the standard for comparison.

From their results, the presence or absence of scoliosis and the number of curves were correctly identified in 68% of the patients (mean Cobb 24°, range 5-61°). In 22% of the patients, there were missed curves or false positives by both surface methods (mean Cobb 14°, range 0-42°). In the other 10% (mean Cobb 8°, range 5-11°), both surface methods gave incorrect results. Other authors have analysed moiré topographs using varying techniques and have reported similar large numbers of false positive findings (Suzuki et al, 1981; Armstrong et al, 1982; Willner, 1982; Moreland, 1983).

3.3.2.4. Problems with Moiré Topography

Moiré topography is a simple, low cost data acquisition method that encodes asymmetry information in a form directly recognisable to clinicians. Asymmetries in back shape, for example, are often obvious from the fringe pattern and direct observation of the moiré topograms has been used widely for assessing back asymmetry (Swain et al, 1992). However, the moiré technique has a number of problems which makes it unsuitable for objective back shape measurements.

The major problem with moiré topography is that the formation of the moiré fringes not only depends on shape of the body, but also on the patient positioning. The fringes are highly susceptible to the patient posture such that minor posture changes can alter the appearance of the fringes. Thus a direct inspection of moiré fringes may be misleading, as has been confirmed in several studies. In practice, to control patient posture, patients are often brought into standardised positioning pads or other fixing elements. However, in the case of the measurement of deformities, this procedure may actually result in further modification of the natural body posture (Frobin et al, 1983).

A second major problem with the moiré method is the difficulty in automating the detection of the fringes. Several authors have attempted to automate the technique and have reported a number of problems that make automation difficult (Hierholzer et al, 1983; Kawamura, 1983;

Moran et al, 1998; Windischbauer, 1983; Yatagai and Idesawa, 1982). Pearson (1996) summarised the problems associated with automated processing of moiré topograms as follows:

- Directional ambiguity. Difficulty in determining the sign of the contour interval without a priori knowledge of the surface. This directional ambiguity is sometimes known as the 'hills or valleys' problem, i.e. there are many surface shapes which give the same fringe pattern.

- Poor contrast. The presence of primary fringes on the object, and the grating through which the object is viewed cause low contrast in the fringe patterns in a moiré topograph, particularly at the shoulders, which limits automatic fringe detection.

- Systematic noise. Automated processing of moiré topographs involves the processing of the two-dimensional image to extract edges or intensity peaks. However, the grating itself presents local intensity peaks and sharp intensity edges which must be filtered out to enable accurate and reliable detection of fringe centres, edges, or phase. These problems limit the method as an affective method for back shape analysis.

3.3.3. Structured Light Projection Methods

Structured light methods for surface measurements are based on the projection of light pattern onto the patient's back and imaging the resulting pattern. Various structured light patterns are possible, however, transverse line pattern appear to have been used more successfully. Figure 3.9 shows a simple grating that might be projected to produce a line pattern. When this grating is projected unto the surface of an object and imaged, the deformity of the object's surface causes the raster profiles to be distorted. The resulting image can then be processed to derive shape information of the object measured.

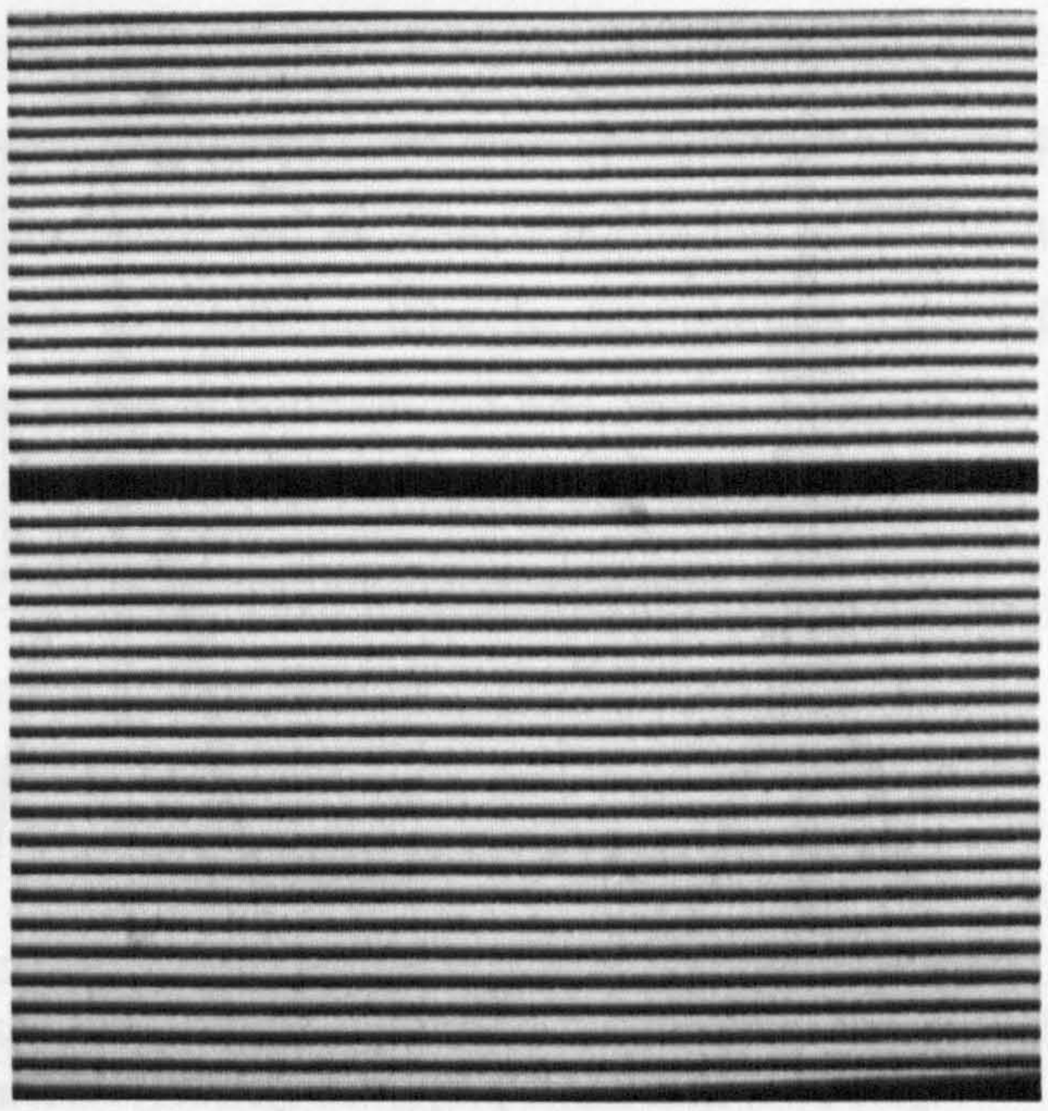


Figure 3.9: A simple grating for line raster projection

To demonstrate the principles of structured light projection systems, two systems are discussed, namely, the systems developed by Frobin et al (1981, 1983, 1988) and Pearson (1996). These systems were deployed in routine clinical measurements, and have been the basis for various studies, thus it is appropriate to discuss them further.

3.3.3.1. Principle of Line Raster Projection: Pearson System

Figure 3.10 shows the raster projection system developed by Pearson (1996). The patient stands on a foot-plate in a vertical frame. The raster pattern is projected onto the surface to be measured by a light source and an image is acquired by a camera angularly offset from the projector. The camera and projector are mounted with the optical centres of their lenses aligned vertically. The image acquired by the camera is transferred via a frame grabber into the memory of the computer

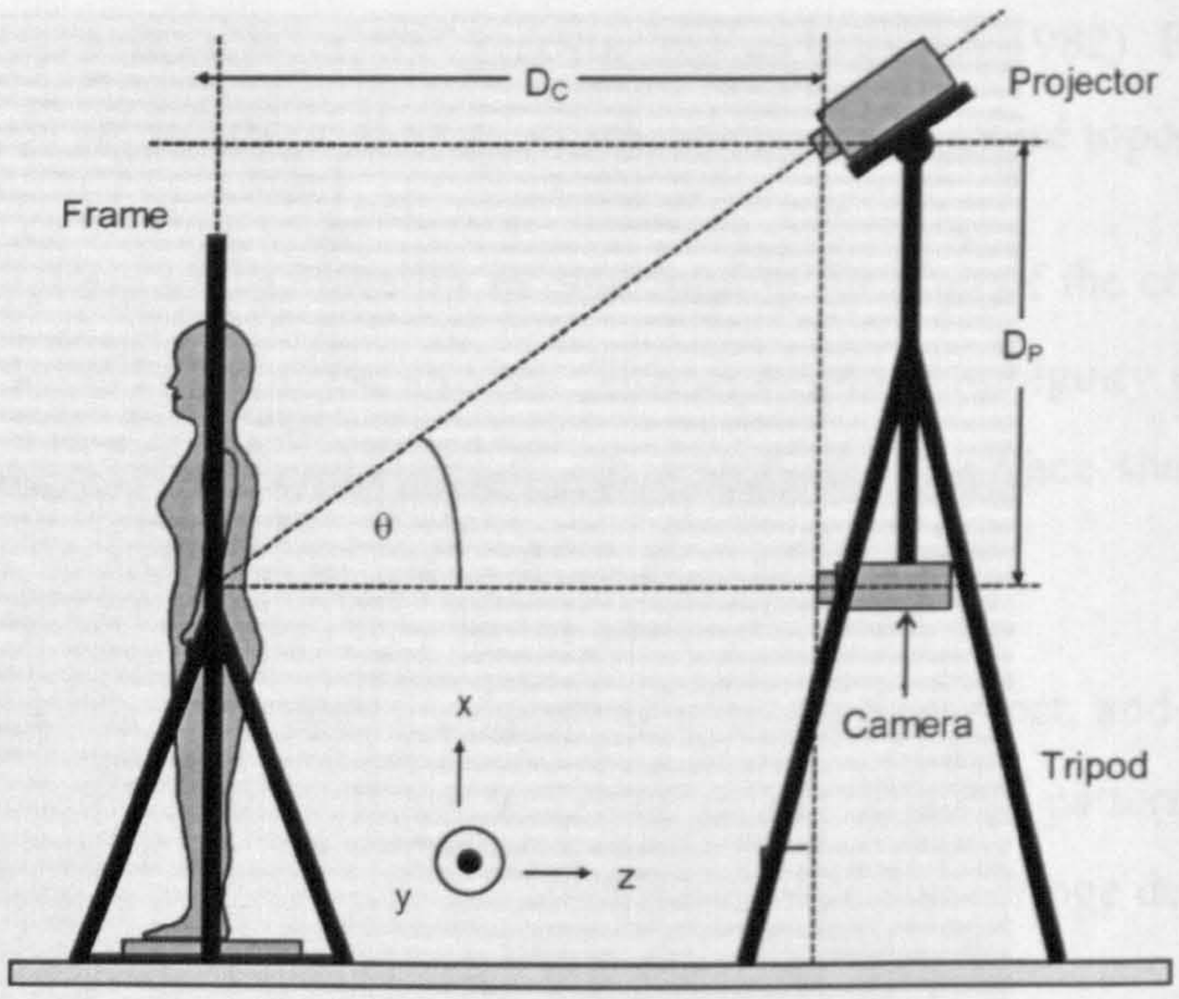


Figure 3.10: Line raster projection system (Pearson, 1996)

The raster pattern will be deformed according to the deformation of the surface, and this can be processed to derive three-dimensional coordinates of the back surface. The surface construction methods often rely on measuring the position of structured light markers on the object and comparing them with the position that a marker would have on some physical or conceptual reference plane. Triangulation based upon these two positions yields three-dimensional coordinates for the structured light markers on the object.

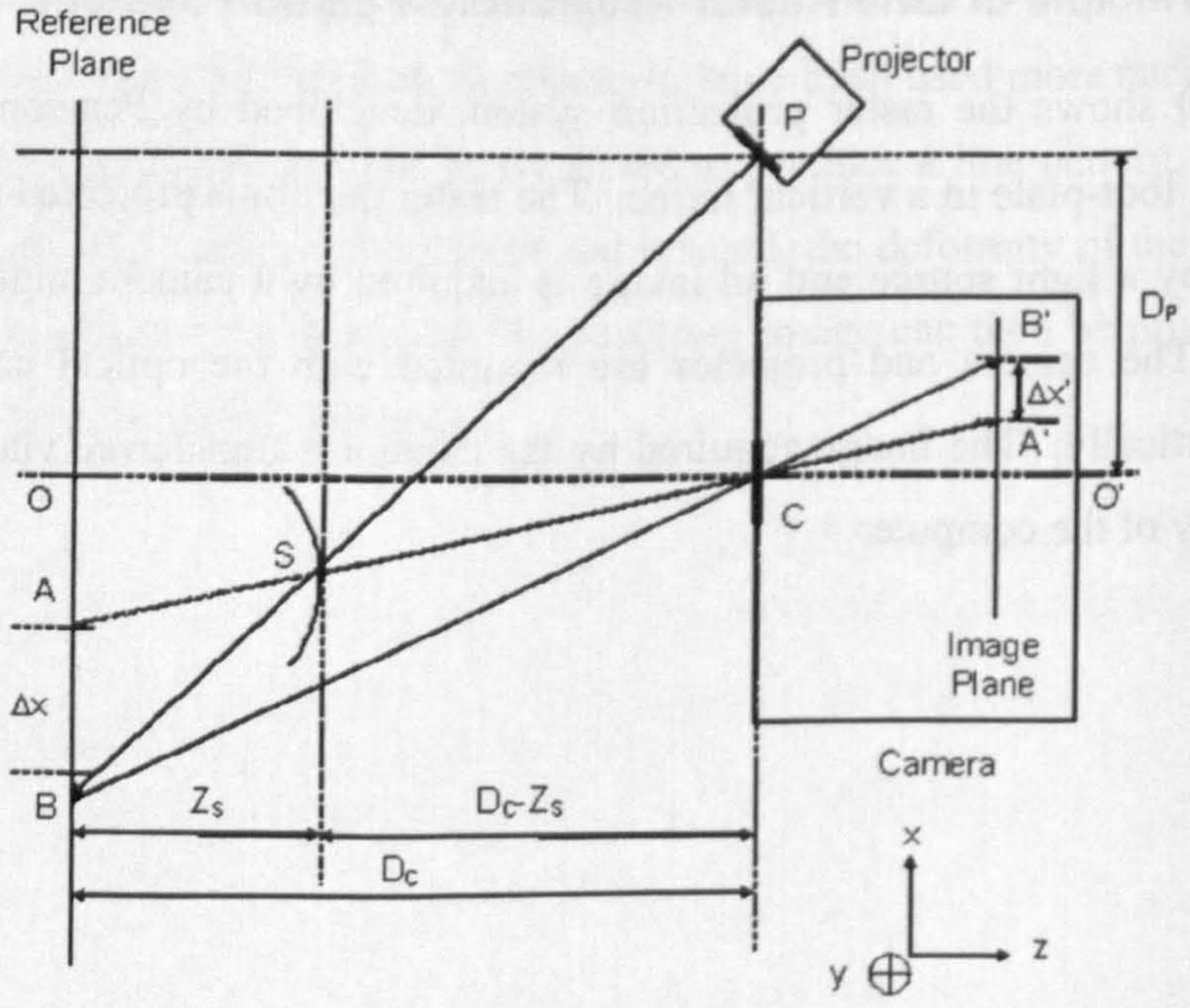
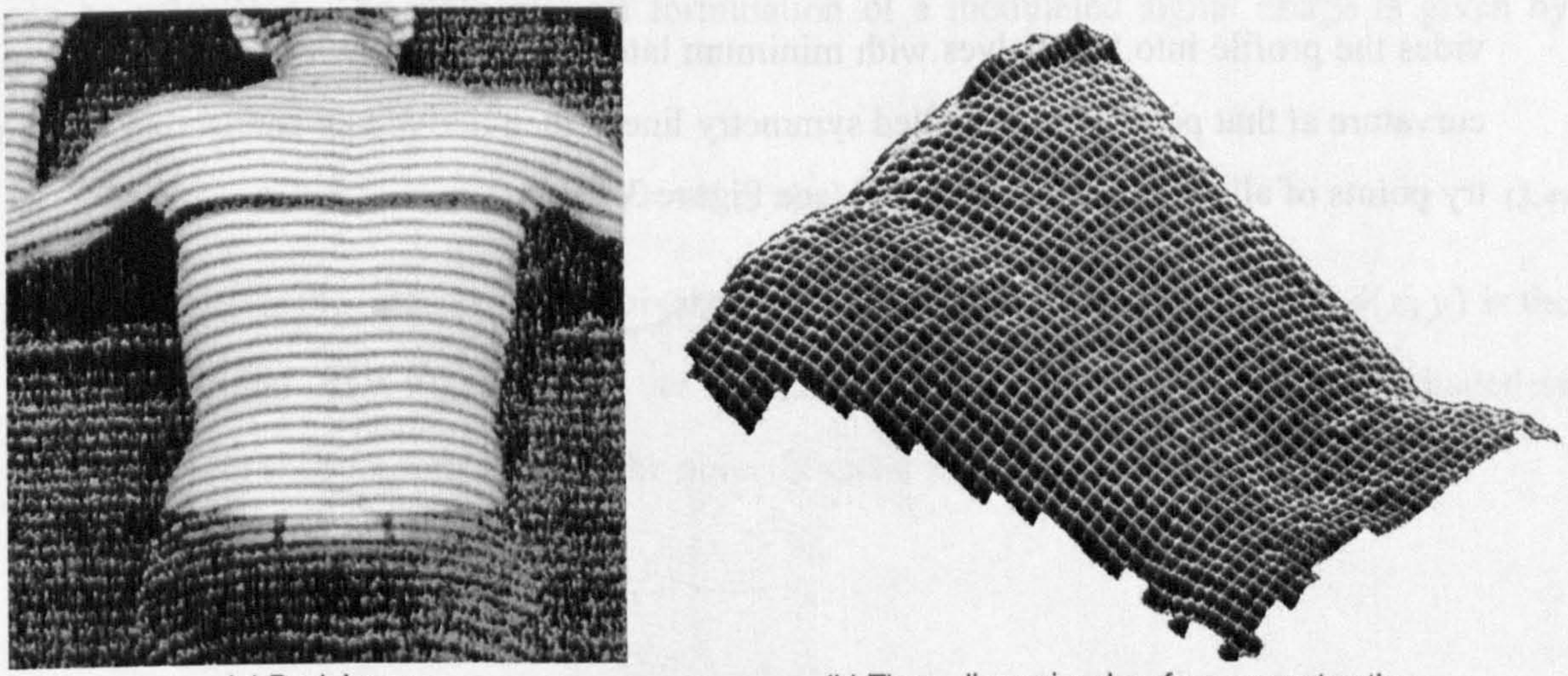


Figure 3.11: Optical geometry of Pearson system (Pearson, 1996)

The optical geometry of Pearson system is shown in Figure 3.11. A fringe from the projector at P strikes an element of the back surface at S and is imaged by the camera at A'. The XY plane is the plane of the image and the Z-axis points out of the image towards the camera. The position of the fringe as it falls on the reference plane is recorded in the image plane of the camera at B'. The distance between the two fringe locations A' and B' is measured in terms of pixels to be $\Delta X'$ and can be converted into absolute distance (ΔX) using scaling information at the reference plane. Then from similar triangles, ABS and CPS, the Z coordinate (Z_S) of the surface element at S is given by Equation 3.3.

$$\frac{Z_s}{\Delta X} = \frac{D_c - Z_s}{D_p} \Rightarrow Z_s = \frac{D_c \Delta X}{D_p + \Delta X} \tag{3.3}$$

Figure 3.12 shows actual back image acquired with Pearson system with the reconstructed surface.



(a) Back image

(b) Three-dimensional surface reconstruction

Figure 3.12: Back image and 3D reconstruction from line raster projection (Pearson, 1996)

The accuracy with which the system measured depth (z coordinate) was determined by making a series of measurements to planes positioned at different depths from the scanner, and was reported to be 0.87mm. The reference frame was positioned at a mean distance of 2340mm from the camera and projector. From the reconstructed back surface, the locations of the vertebrae are estimated manually and two measurements are determined, namely, angle of trunk inclination (ATI) and trunk asymmetry score (TAS).

3.3.3.2. Rasterstereography

Rasterstereography is a structured light projection method developed by Frobin and Hierholzer (1981, 1983, 1988). The method was originally based on photographic film which was digitised by scanning. It was later adapted for automated imaging using a CCD camera and produced as a clinical system. The optical configuration is essentially similar to the Pearson system and thus will not be repeated here. Rather, we focus on the data analysis techniques developed by the authors to derive relevant information for asymmetry analysis of the back surface, based on shape analysis of the horizontal profiles produced from rasterstereography data.

In Hierholzer (1985) and Drerup & Hierholzer (1985, 1987), the authors discuss a method of back shape analysis based on asymmetry of the horizontal profiles produced from rasterstereography. The symmetry point of each horizontal profile is defined by a point which divides the profile into two halves with minimum lateral asymmetry with respect to the surface curvature at that point. The so-called symmetry line is then derived by connecting the symmetry points of all the horizontal profiles (see Figure 3.13).

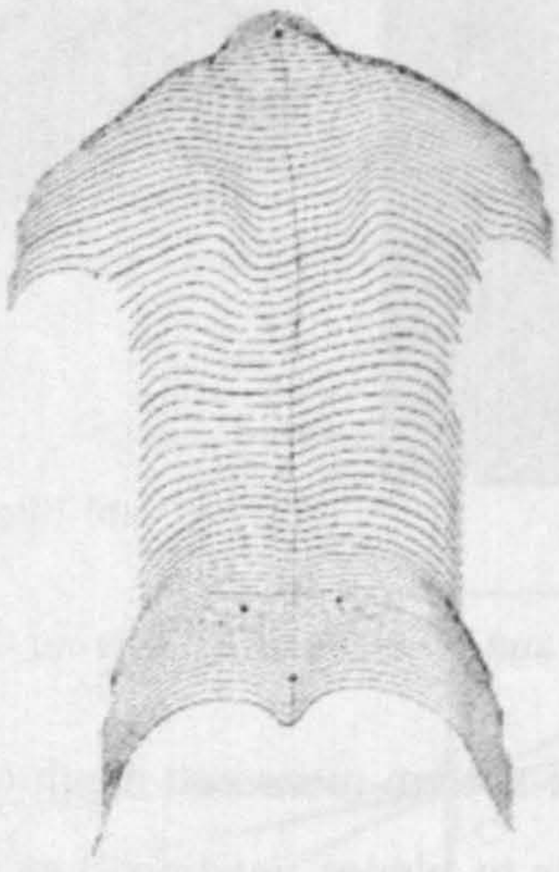


Figure 3.13: Horizontal profiles and the symmetry line (Drerup & Hierholzer, 1994)

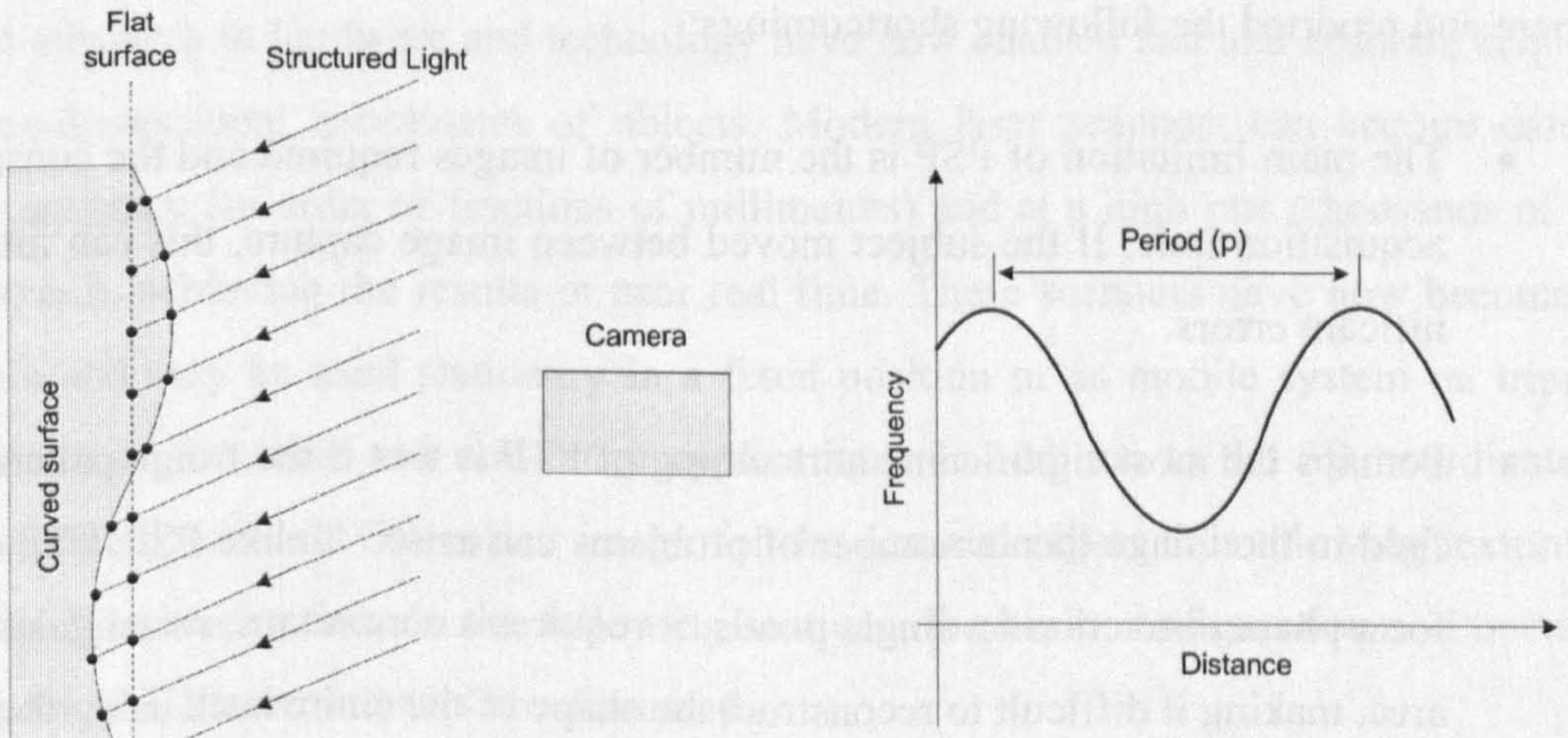
Detailed discussion of the shape analysis method used by these authors is postponed until the introduction of curvature analysis. In Chapter 6, the shape analysis method of Drerup & Hierholzer (1985, 1987) is compared with the method developed in this investigation.

3.3.4. Phase Measuring Methods

Structured light pattern projected on the surface of an object like the human body encodes information about the shape of the object which can be extracted by geometrically analysing the light pattern. An alternative method for extracting shape information is to regard the light pattern as a periodic sinusoidal signal which is modulated by the surface of the object. Figure 3.14 shows how a structured pattern is modified as it falls on a curved surface. The amount and direction of the modification depends on the roughness of the surface. In signal processing terms, the light signal has been *phase modulated* by the surface. This modulated signal encodes shape information of the surface which can be extracted by recovering the original signal, a process known as *demodulation*. It can be shown that if the incident light on the surface is periodic as shown in Figure 3.14(b), then there will be a phase shift proportional to the height change at the point of interest. If the phase shift can be measured, the height change can be calculated. The mathematical formulation of a modulated signal image is given by Equation 3.4 (Halioua, 1989):

$$I(x,y) = A(x,y) + B(x,y)\phi(x,y) \tag{3.4}$$

where $I(x,y)$ is the intensity of the pixel, $A(x,y)$ is the background intensity, $\phi(x,y)$ is the phase signal and $B(x,y)/A(x,y)$ is the fringe contrast. Equation 3.4 must be evaluated in order to obtain shape information of the object's surface.



(a) Modulation of structured light by curved surface

(b) Light intensity on flat surface

Figure 3.14: Phase measurement methods: modulation of structured light by a curved surface

Two main methods are used for phase extraction, namely, phase stepping profilometry (PSP) and Fourier transform profilometry (FTP). Phase stepping solves for the three unknowns in Equation 3.4 by taking at least three images, with the projection grating moved between image capture by a fraction of its period, known as the step size. The step size will typically have four values (0, 90, 180, 270), giving four images. The phase of each pixel in the image can then be calculated using pixel values from all four images. The resulting phase values are wrapped, i.e. in the range $\pm\pi$, so a further processing is required for unwrapping which is probably the most difficult part of the algorithm.

Fourier transform analysis requires only one image and relies on the calculation of the Fourier transform of the modulated signal image, which can be obtained by the Fast Fourier Transform algorithm (FFT) (Sorensen et al, 1987). A series of filters are applied to the FT image to remove negative frequency components, dc component (average signal) and any noise and harmonics in the positive spectrum (Pearson, 1996 for detail description of algorithm). After filtering, the inverse Fourier transform is then applied to the image, yielding both real and imaginary parts. The phase of the light pattern for each pixel is then calculated to be the angle between the real and imaginary parts, giving phase map, or phase image. Again unwrapping is required which can be followed by conversion to depth coordinates for surface reconstruction.

The literature reveals little about the application of phase measuring profilometry to the measurement of the back shape. Pearson (1996) applied the method to back shape measurement and reported the following shortcomings:

- The main limitation of PSP is the number of images required and the consequent data acquisition time. If the subject moved between image capture, this can introduce significant errors.

- Perhaps the most significant shortcoming of FTP is that if the fringe pattern is not full field in the image then a number of problems can arise. Unlike PSP, FTP cannot perform phase extraction on single pixels: it requires a continuous, rectangular sub-image area, making it difficult to reconstruct the shape of the entire back using the method.

- The FTP requires a considerable level of expertise and experience in filtering in the Fourier domain, making it very difficult to design filters that will work across a broad range of surfaces.

- Lastly, FTP requires the computation of both a forward and inverse Fourier transform which can be computationally expensive for low spec computers.

3.3.5. Three-dimensional Scanning Systems

Optical scanning systems are devices which use controlled angular deflection of an incoming beam, plane or other pattern of structured light. Scanning systems typically consist of a camera and a structured light. Mirrors mounted on galvanometers are used to control the deflection of the beam. The light source used for optical scanners is often a laser which produces a horizontal beam or spot on the subject. The reflected light from the subject is then imaged by a camera mounted orthogonally or angularly offset from the incoming beam. Consideration of the triangle formed by camera, projector and object yields three-dimensional coordinates for the structured light markers on the object.

An early and widely used optical scanning system for measuring human back shape in a clinical environment is the ISIS (Integrated Shape Imaging System) (Turner-Smith, 1988). The ISIS system used structured white light projector and a low cost 625-line television camera. The configuration of ISIS system was different from that used in modern scanners. The camera and projector were mounted in a fixed relationship within a unit and the entire unit was then made to swing about a horizontal axis in order to complete a scan. The scanning was completed within 2 seconds with approximately 3mm accuracy in all three axes.

Recent advances in hardware and technology have now enabled fast and accurate acquisition of three-dimensional coordinates of objects. Modern laser scanners can acquire data with higher accuracy (in order of fractions of millimetres) and at a high rate (thousands of points per second), achieving the results in near real time. These scanners have now become more portable and may be used stationary in a fixed position or as mobile system on tripods or similar stands for close and mid-range applications. In addition to the 3D coordinate data, most scanners deliver RGB colour image for the scanned object thus allowing texture mapping for 3D visualization. In the following sections, the basic configuration and operational principles of 3D laser scanners are discussed.

- Lastly, FTP requires the computation of both a forward and inverse Fourier transform which can be computationally expensive for low spec computers.

3.3.5. Three-dimensional Scanning Systems

Optical scanning systems are devices which use controlled angular deflection of an incoming beam, plane or other pattern of structured light. Scanning systems typically consist of a camera and a structured light. Mirrors mounted on galvanometers are used to control the deflection of the beam. The light source used for optical scanners is often a laser which produces a horizontal beam or spot on the subject. The reflected light from the subject is then imaged by a camera mounted orthogonally or angularly offset from the incoming beam. Consideration of the triangle formed by camera, projector and object yields three-dimensional coordinates for the structured light markers on the object.

An early and widely used optical scanning system for measuring human back shape in a clinical environment is the ISIS (Integrated Shape Imaging System) (Turner-Smith, 1988). The ISIS system used structured white light projector and a low cost 625-line television camera. The configuration of ISIS system was different from that used in modern scanners. The camera and projector were mounted in a fixed relationship within a unit and the entire unit was then made to swing about a horizontal axis in order to complete a scan. The scanning was completed within 2 seconds with approximately 3mm accuracy in all three axes.

Recent advances in hardware and technology have now enabled fast and accurate acquisition of three-dimensional coordinates of objects. Modern laser scanners can acquire data with higher accuracy (in order of fractions of millimetres) and at a high rate (thousands of points per second), achieving the results in near real time. These scanners have now become more portable and may be used stationary in a fixed position or as mobile system on tripods or similar stands for close and mid-range applications. In addition to the 3D coordinate data, most scanners deliver RGB colour image for the scanned object thus allowing texture mapping for 3D visualization. In the following sections, the basic configuration and operational principles of 3D laser scanners are discussed.

3.3.5.1. Basic Principle of 3D Laser Scanners

In general, 3D laser systems can be divided into two groups, namely, *monocular* and *binocular* systems (Wang et al, 2002). The monocular approach is based on the propagation of light (also called *ranging scanners*); while the binocular systems are based on the triangulation methods.

3.3.5.2. Ranging Scanners

There are two approaches used by ranging scanners, namely, *time of flight* and *phase comparison* (see Figure 3.15). With the time of flight approach, a laser pulse is sent to the object and the time it takes for the signal to travel to the object and back is recorded by the receiver sensor. This time interval multiplied by the speed of light and divided by 2 gives the distance from the laser head to the object. The 3D coordinates of the scanned point can then be computed if the position of the laser head is known in coordinated space, together with the scan angle of the laser mirror. Scanners use small rotating device for the angular reflection of the beam. The resolution and accuracy of time-of-flight scanners is quite limited: typical standard deviation is of the order of a few millimetres.

With the phase comparison approach, the transmitted laser beam is modulated by a harmonic wave and the distance is derived using the phase difference between the transmitted and received wave (see Section 3.3.4). Since a well defined returning signal is needed, scanners using the phase comparison method may also have a reduced range and tend to produce more wrong or dropped points (Wang et al, 2002).

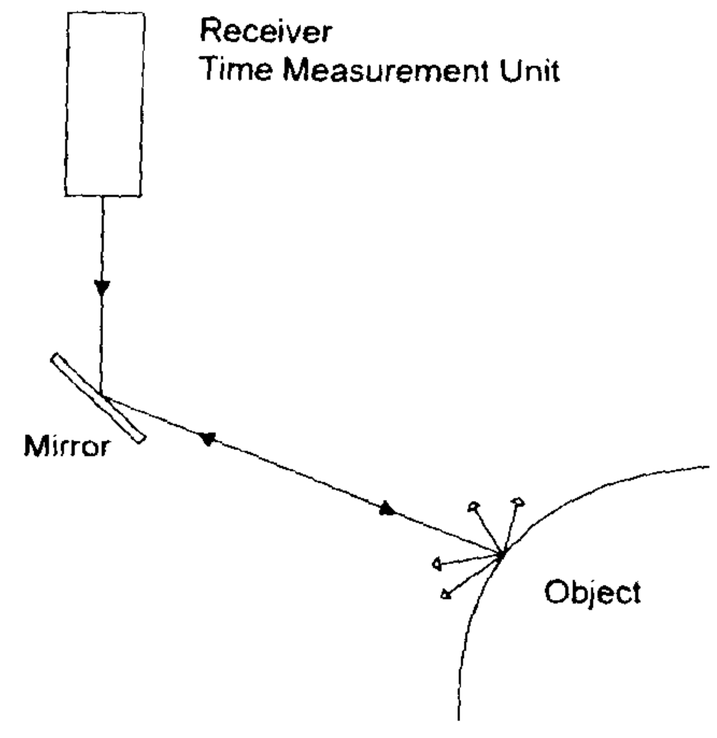


Figure 3.15: Time of flight laser scanner (adapted from Wang et al, 2002)

3.3.5.3. Triangulation Scanners

There are two types of triangulation scanners, namely, those that use *single camera solution* and those that use *double cameras solution* (Figure 3.16). The single camera scanner consists of a transmitting device, which sends a laser beam from one end of a mechanical base onto the object, and a CCD camera at the other end of the base which detects the laser beam from the object. Consideration of the resulting triangle formed by the laser projector, the camera and the object yields the 3D coordinates of the scanned surface. The double camera solution is a variation of the triangulation principle which uses two CCD cameras at each end of the measurement base. A separate projector transmits the laser spot or beam onto the object to be scanned. The projection may consist of a moving light spot or line, of moving stripe patterns, or of a static arbitrary pattern. The geometric solution is the same as with the one camera principle, thus resulting in the similar accuracy characteristics. Triangulating laser scanners typically have a higher resolution and accuracy than ranging scanners: typical accuracies are of the order of fraction of millimetres, thus making them ideal for scanning small objects or over short distances.

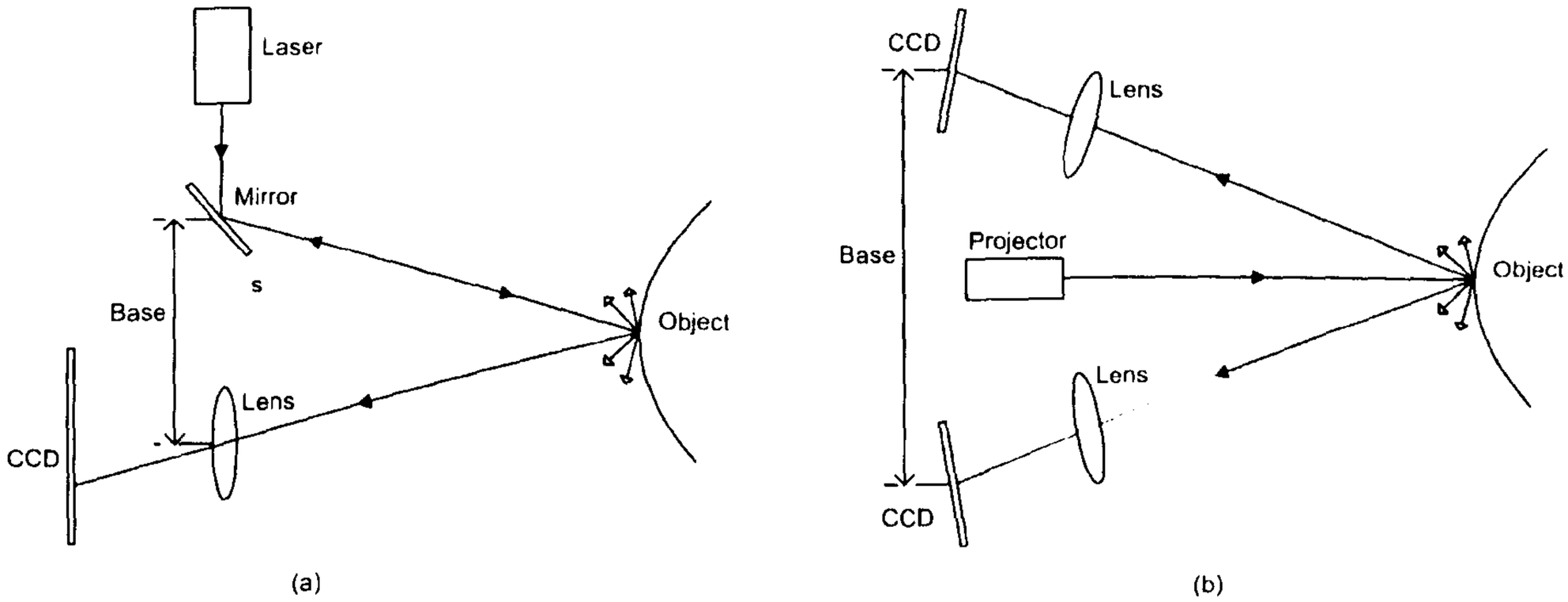


Figure 3.16: Triangulation scanners (adapted from Wang et al, 2002)

3.3.5.4. Application of 3D Laser Scanning to Back Shape Analysis

As discussed in the preceding sections, 3D laser scanners allow accurate and fast acquisition of three-dimensional data of objects. Thus the problem of data acquisition has been solved to a satisfactory level. However, methods to extract relevant information for objective analysis of the back shape have not been widely developed. From the literature, it appears that a few

studies have attempted to derive objective measurements from laser data for back shape analysis.

An interesting application of 3D laser scanning to back shape analysis, in the context of this study, is presented in Jaremko et al (2002b). In this paper, the authors acquired 3D coordinates of the entire torso and generated contours at 10mm vertical separation for the torso model. Various indices that described asymmetry of the torso were then computed based on these contours and compared with radiographic measurements. Indices computed for each cross-sectional torso contour are summarised as follows (Figure 3.17):

- Back surface rotation (BSR). The BSR in each contour was computed as the orientation of the dual tangent line joining the most posterior points (T_L and T_R) left and right of the contour centroid (BSR is θ_2 in Figure 3.17a).

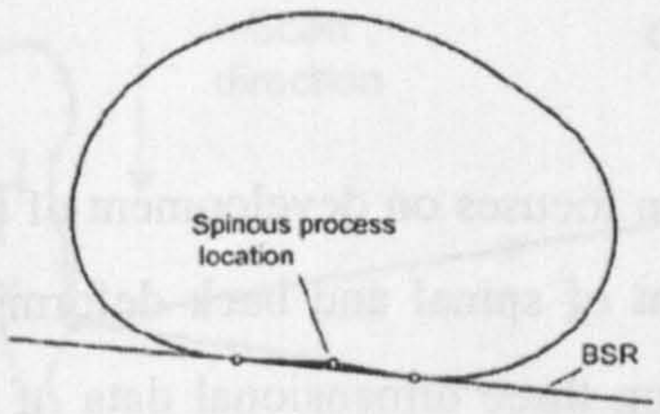
- Principal axis orientation (PAX). The orientation of the principal axes of inertia of each contour with respect to the PSIS reference axis (θ_1 in Figure 3.17b).

- Rib hump was computed as the difference in distance from left and right tangent points to PSIS reference axis ($d_L - d_R$ in Figure 3.17c).

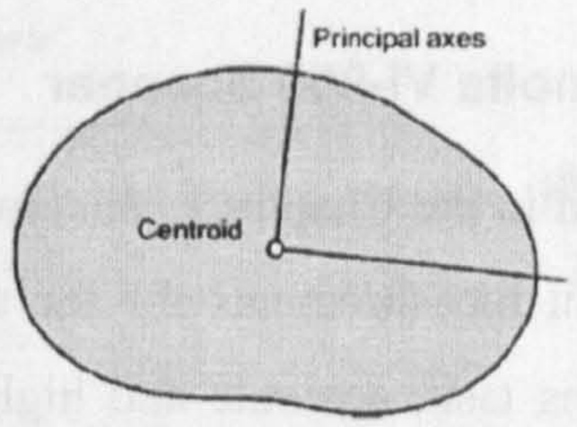
- Envelop indices. For each contour, the left-right difference between the lateral distance from the centroid to the outer envelop was computed. The left-right difference in aspect ratio was also computed as $dAP/dLat$ for each half-area.

- Half-area indices. Differences in geometric properties of the left and right halves of each cross-section were assessed. These measurements included differences in left and right half-areas; antero-posterior and lateral asymmetry of half-centroid locations (dXC and $Z_{CL} - Z_{CR}$ in Figure 3.17d); orientation of the line joining the half-centroids and left-right asymmetry of moments of inertia measured antero-posteriory and laterally.

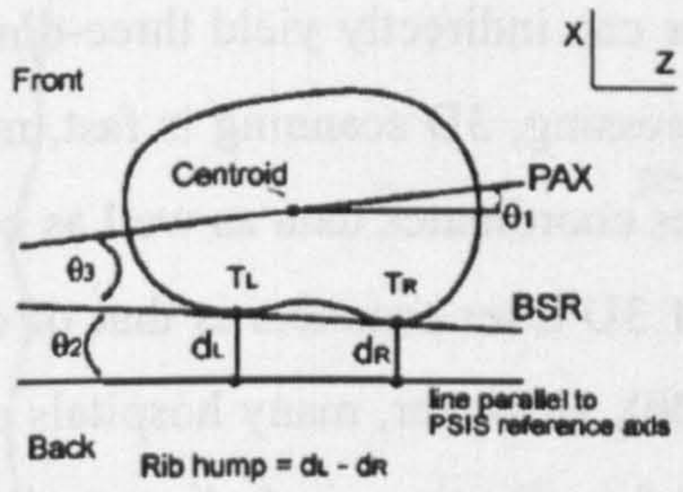
- Torso centroid line. The line connecting the centroid of the contours.



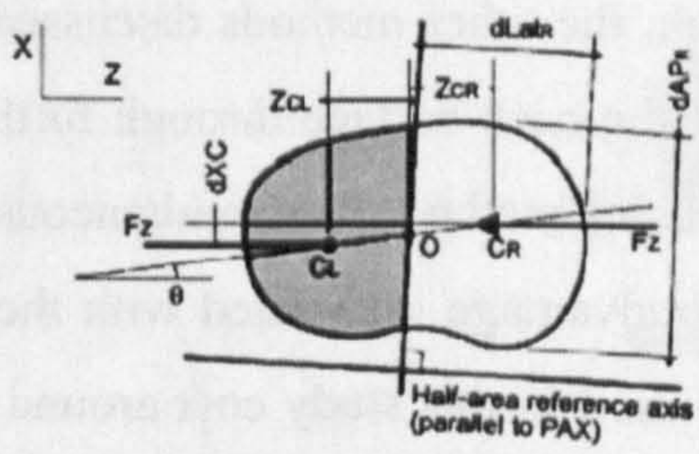
(a) Back Surface Rotation



(b) Principal axes of contour



(c) Index calculation



(d) Half area index calculation

Figure 3.17: Index calculation by Jaremko et al (2002b). (a) θ_1 is PAX orientation; θ_2 is the BSR; θ_3 is the difference between BSR and PAX; Rib hump is computed as $d_L - d_R$ (b) Aantero-posterior (dXC) and lateral (Z_{CL} vs. Z_{CR}) asymmetry; θ is the angle of rotation of line joining the half-centroids; dAP/dLat is the aspect ratio for each half-area, F_Z is the hypothetical unit-force applied inward at each half-centroid.

Correlation of surface each index with the Cobb angle was computed and stepwise regression was used to estimate the Cobb angle from selected torso indices. From their results, some of the indices developed correlated quite well to the Cobb angle (correlation up to 0.8). The Cobb angle was best estimated by patient age group, rib hump and left-right variation in unbraced patients and by centroid lateral deviation in braced patients. A regression model applied estimated the Cobb angle from the torso indices within 5^0 in 65% of patients and 10^0 in 88% ($r=0.91\pm6.1^0$).

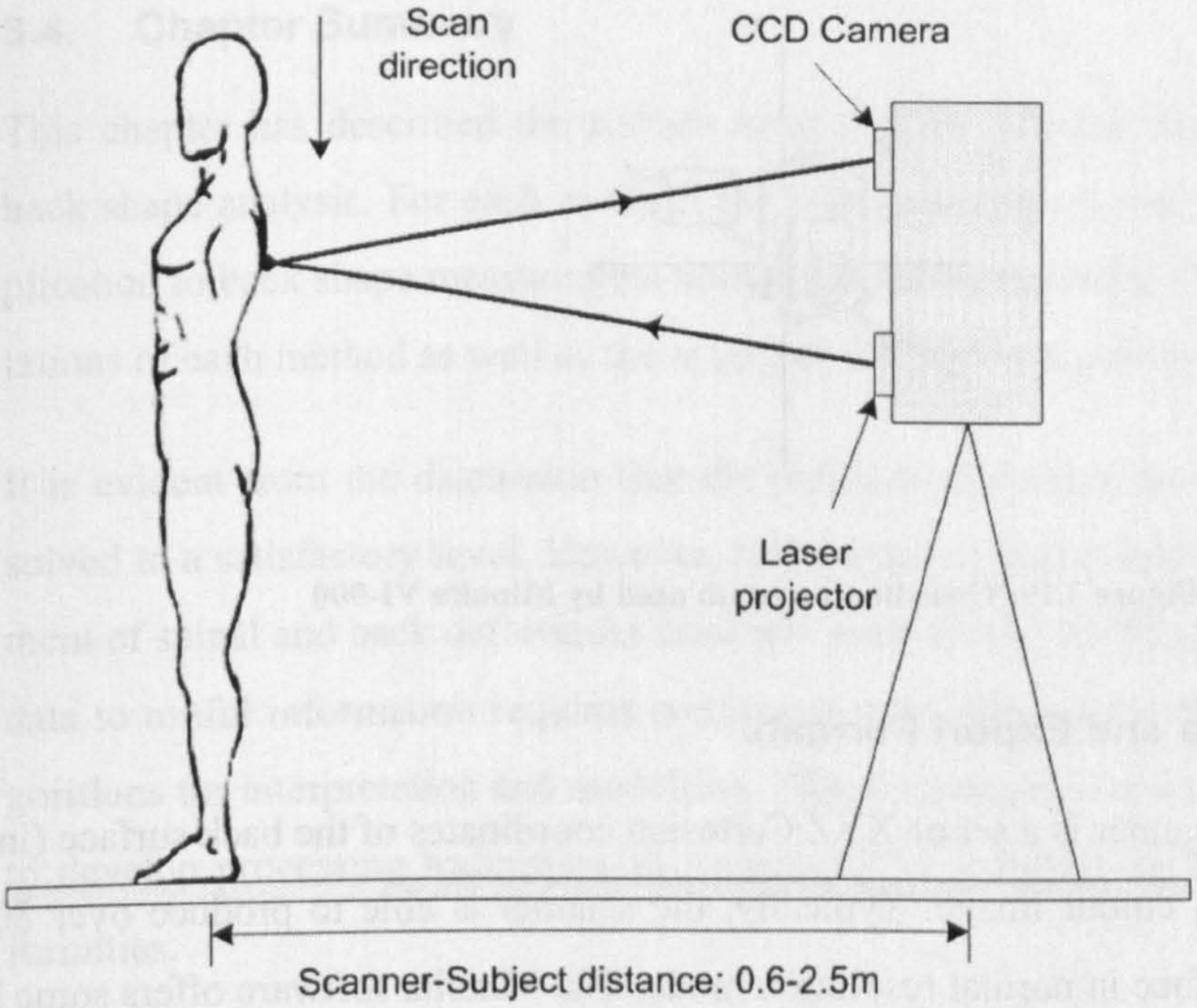
In other studies by the same authors (Jeremko et al, 2001, 2002c), genetic and neural network algorithms were used to estimate the Cobb angle from the indices computed from contours of the torso of 48 scoliosis patients. The Cobb angle was estimated to within 5^0 in two-thirds of the patients and to within 10^0 in six-sevenths of 48 scoliosis patients.

3.3.6. Minolta VI-900 Scanner

As discussed in the Chapter 1, this investigation focuses on development of techniques to extract relevant measurements for the assessment of spinal and back deformities. In order to carry out this task, accurate and high resolution three dimensional data of the back are required. These requirements make the 3D laser scanner the obvious choice for data acquisition. Although, the other methods discussed earlier can indirectly yield three-dimensional coordinates of the back surface through further processing, 3D scanning is fast, more accurate and non-invasive method that simultaneously gives coordinates data as well as colour image. The main disadvantage associated with the use of 3D laser scanners is that of cost (the Minolta scanner used in this study cost around £25,000). However, many hospitals and organisations do purchase laser scanners for a wide range of applications including medical, architectural, and manufacturing purposes. For this investigation, Minolta VI-900 laser scanner was used for data capture due to its availability and the fact that it fulfilled the accuracy requirements. This section discusses the principal features of the VI-900 scanner.

3.3.6.1. Operating Principle

The operation of VI-900 is based on laser triangulation principle and consists of a transmitting device which sends a horizontal laser beam onto the subject and a CCD camera which detects the reflected laser beam from the subject (Figure 3.18). The object is scanned by a plane of laser light coming from the VI-900's source aperture. The laser beam is swept across the field of view by a mirror, rotated by a galvanometer. The reflected light from the object is captured by the CCD camera. The three-dimensional coordinates of the object are then derived from the resulting triangulation. The entire object is completed in 0.3 seconds in fast mode, and 0.5 seconds in colour mode and 2.5 seconds in fine mode. Complete 3D digitization of objects can be accomplished by connecting a synchronised rotary platform, taking multiple scans by rotating the platform using a user-definable angular step (e.g. 45° or 90°), and registering the views to form a complete 3D model.



Minolta VI-900

Figure 3.18: Operating principle of Minolta VI-900 laser scanner

The scanner is equipped with three interchangeable lenses for variable scanning volumes, namely, tele lens (focal length 25mm), medium lens (focal length 14mm) and wide lens (focal length 8mm). The accuracy is thus dependent on the scanned area—Table 3.1 summarises the accuracy associated with each lens type in fine mode.

Table 3.1: Accuracy of VI-900 for each lens type in millimetres (fine mode)

Lens Type	X	Y	Z
Tele	0.22	0.16	0.10
Middle	0.38	0.31	0.20
Wide	1.40	1.04	0.40

3.3.6.2. Coordinate System

With the VI-900 scanner, three dimensional coordinates are defined using right-handed Cartesian coordinates system, the origin of which is set at the centre of the light-receiving lens of the CCD camera. The +Z direction refers to the direction facing the front side along the optical axis, +Y direction refers to the upward direction and +X direction refers to the direction facing the right, when viewed from the back of the light-receiving section (see Figure 3.19). The Z-coordinates of the scanned object are therefore always negative.

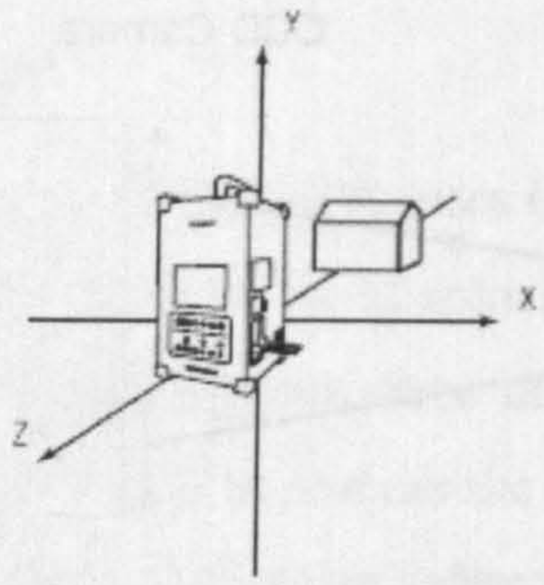


Figure 3.19: Coordinate system used by Minolta VI-900

3.3.6.3. Output Data and Export Formats

The output from the scanner is a set of XYZ Cartesian coordinates of the back surface (in millimetres) and an RGB colour image. Typically, the scanner is able to produce over 80,000 points on the back surface in normal resolution mode. The Minolta software offers some basic functions to visualize and edit the output data to a satisfactory level. For example, the actual number of points measured can be sub-sampled to the required size. Functions such as image registration and filling of holes in the surface are also available. The point data are visualized as a polygonal mesh, and can be exported to supported file formats such as ASCII, TXT or DXF for further processing.

3.3.6.4. Laser Safety Issues

The VI-900 scanner uses Class-2 laser with a maximum wavelength of 690nm. According to the Health and Safety Manual produced by the US National Institute of Environmental Health Sciences, Laser Class 2 is safe in operation:

> Class II - denotes low power visible lasers or laser systems which because of the normal human aversion responses (<0.25s), do not normally present a hazard but may present some potential for hazard if viewed directly for extended periods of time (like many conventional light sources) (Health and Safety Manual, Chapter 9).

Minolta VI-900 has been used for scanning various parts of the human body, including the breasts (Langmack et al, 2000) and the face and head (Boehnen & Flynn, 2005).

3.4. Chapter Summary

This chapter has described the surface measurement systems that have been developed for back shape analysis. For each method, the underlying theory was briefly outlined and its application to back shape measurement was discussed by reporting the work of others. The limitations of each method as well as the main advantages were pointed where appropriate.

It is evident from the discussion that the problem of surface data acquisition has now been solved to a satisfactory level. However, techniques to extract relevant information for assessment of spinal and back deformities have not been widely developed. To convert the 3D point data to useful information requires continuous research and development of methods and algorithms for interpretation and modelling. This investigation addresses this need and attempts to develop processing techniques to extract useful information for the assessment back deformities.

The next chapter provides the mathematical foundation for the shape analysis techniques that have been applied to back shape analysis in this investigation.

4. Mathematical Basis For Back Shape Analysis

4.1. Introduction

This chapter presents the mathematical basis for the shape analysis techniques that have been developed in this research for back shape assessment. Unlike many industrial and real world objects, the human back is a relatively smooth surface, that is, there are no crisp features on the surface of the back. This precludes the application of conventional feature detection techniques to back shape analysis. The approach adopted in this study is to apply mathematical analysis of surface shape or curvature. This comprises fitting a smooth mathematical surface to each point and its local neighbourhood and analysing the approximating surface to obtain knowledge about the surface shape at that point. The concept of the curvature of two-dimensional curves is first introduced and generalised into the third dimension. Various measures of surface curvature that have been proposed in various applications of differential geometry are then discussed, including their computational methods. The method of Least Squares has been used as the basis for various tasks, including surface fitting for curvature computation and landmarks extraction; polynomial fitting for derivation of the spinal curves; and coordinate transformation for comparing back surfaces. Consequently, the basic principles of Least Squares are discussed, including the practical applications to which the method has been committed.

4.2. Curvature Analysis

4.2.1. Curvature of Curves in the Plane

Curvature is a two-dimensional property of a curve and intuitively measures the deviation of the curve from a straight line at the point of interest (Thomas, 1972; Gray, 1997; O'Neill, 1996). Figure 4.1 shows a curve passing through a point P; N is the normal (or perpendicular) vector to the curve at P. The curvature of the curve at P represents the rate of change of direction of the curve at P. A circle passing through P with radius R is called the *osculating circle* or circle of curvature if the circle lies on the concave side of the curve and shares a common tangent line with the curve at point P. This circle makes the greatest possible contact with the curve at P. The radius of the osculating circle is called the *radius of curvature* and is inversely proportional to the curve's curvature:

$$K = \frac{\partial \omega}{dS} = \frac{2\pi}{2\pi R} = \frac{1}{R} \tag{4.1}$$

Thus, the smaller the radius of curvature, the more the curve deviates from a straight line and the greater its curvature. As the radius of curvature approaches infinity, the circle approaches a straight line and its curvature approaches zero.

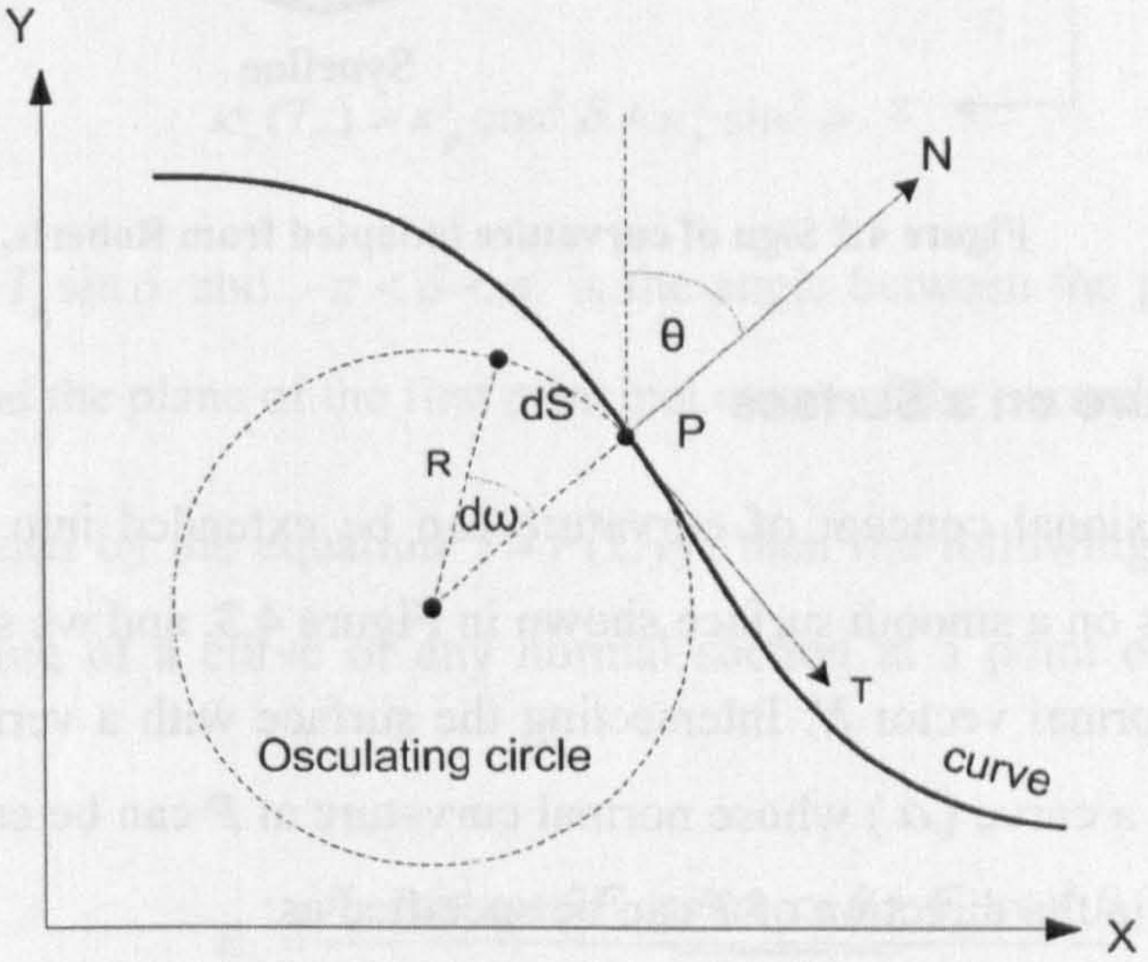


Figure 4.1: Curvature of a curve

Curvature is closely related to the second derivative and can be expressed in terms of derivatives as in Equation 4.2 (Thomas, 1972). Consequently, the second derivative is sometimes used as approximation of the curvature although this assumption is only valid when the normal coincides with the Z-axis.

$$\kappa = \frac{d^2y/dx^2}{(1+(dy/dx)^2)} \tag{4.2}$$

Sign of Curvature

Figure 4.2 describes curvature sign convention used in this thesis. The normal vectors to the curve are shown at regular intervals along the curve. Where the curve is flat, the normals are all parallel to each other, and the curvature of the curve is zero at these locations. Where the curve is convex, the normals diverge and the curvature is marked as positive; and where the curve is concave, the vectors converge and the curvature is negative.

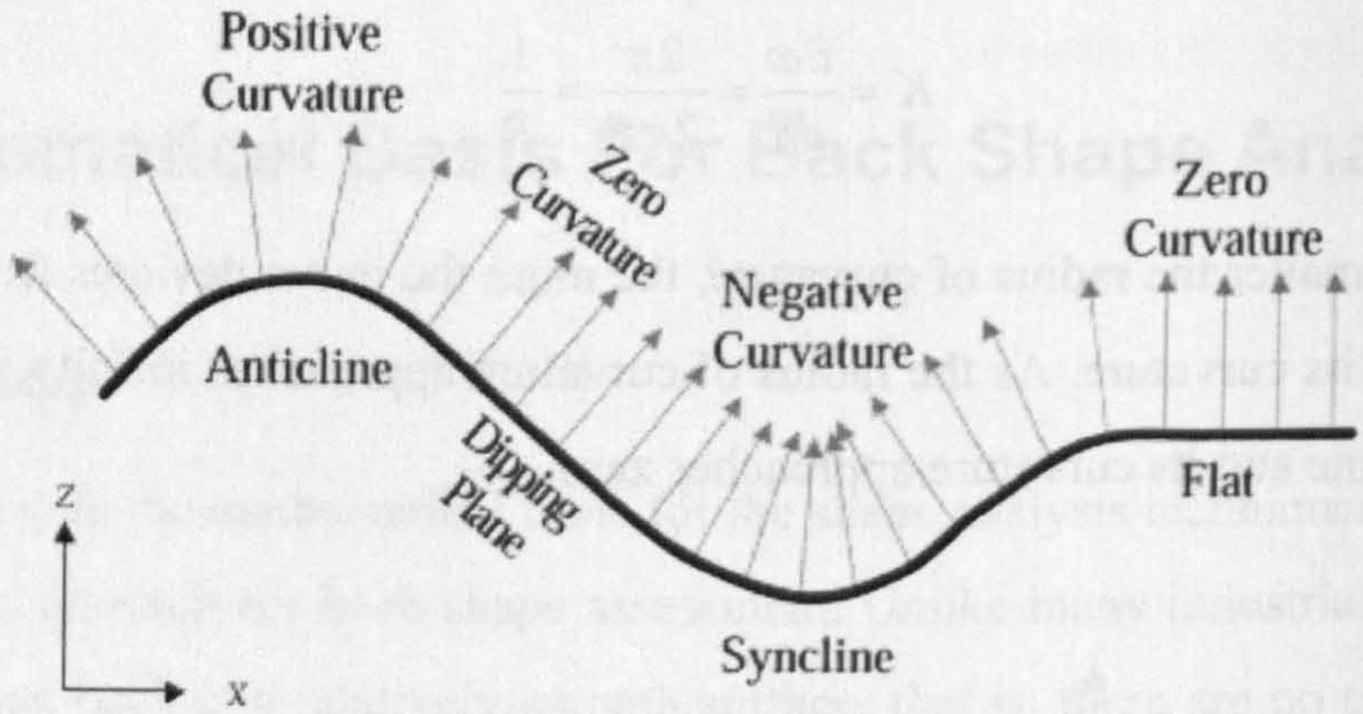


Figure 4.2 Sign of curvature (adapted from Roberts, 2001)

4.2.2. Curvature on a Surface

The two dimensional concept of curvature can be extended into the third dimension. The point P now lies on a smooth surface shown in Figure 4.3, and we specify the orientation of S at P with the normal vector N. Intersecting the surface with a vertical plane that contains P and N produces a curve (α) whose normal curvature at P can be calculated. The normal curvature of S at P in the direction of T can be specified as:

$$\alpha^{\cdot}(0) = \kappa_p(T)N \tag{4.3}$$

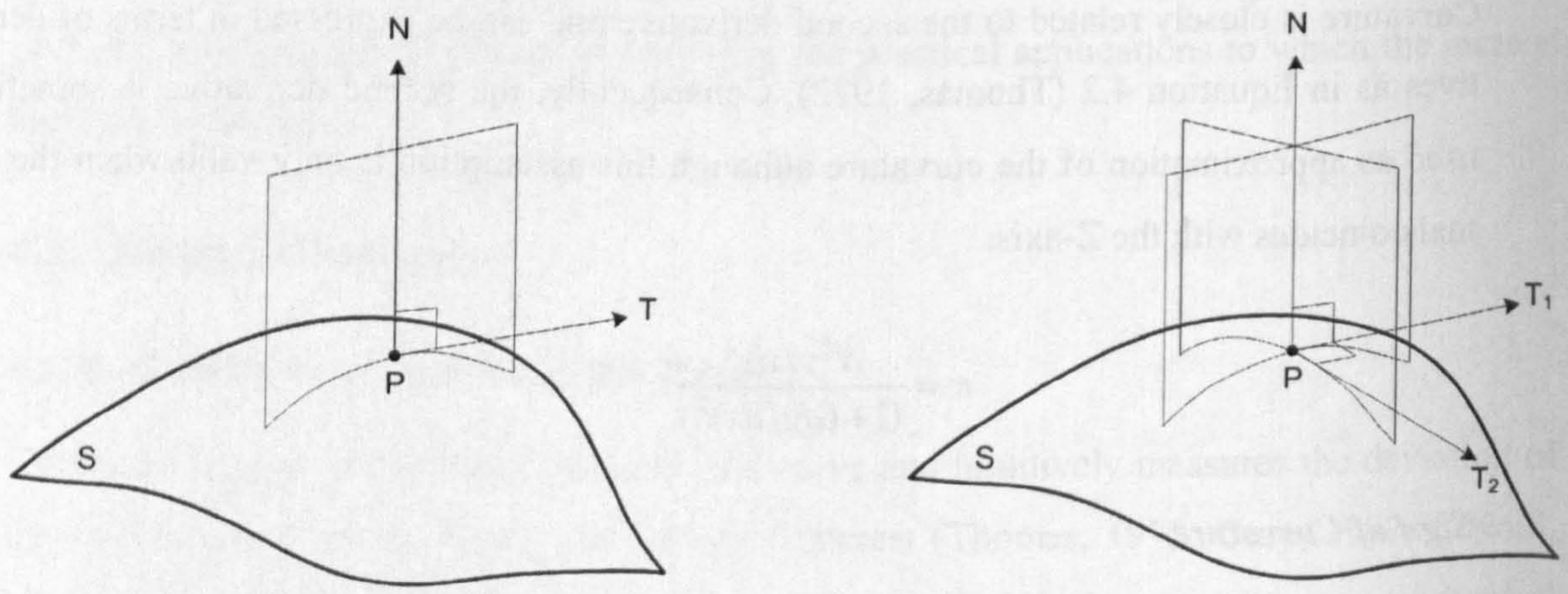


Figure 4.3: Curvature at a point P on a surface (a) shows a normal curvature on the surface S at P (b) shows the principal curvatures and directions T_1 and T_2

This curvature $\kappa_p(T)$ is only for a single curve on S and does not completely define the curvature of the surface at P. by rotating the plane section about normal N, we can construct an in-

finite number of planes which intersect the surface resulting in an infinite number of curvatures. From this infinite set, the maximum and minimum curvatures are called *principal curvatures* and their corresponding tangent vectors are called *principal directions*. The principal curvatures are the most useful subset of normal curvatures. Let the principal curvatures and directions be denoted by κ_1, κ_2 and T_1, T_2 respectively, then the following relationship, known as Euler's Theorem, holds for any normal curvature:

$$\kappa_p(T_\delta) = \kappa_p^1 \cos^2 \delta + \kappa_p^2 \sin^2 \delta \tag{4.4}$$

where $T_\delta = T_1 \cos\delta + T_2 \sin\delta$ and $-\pi < \delta < \pi$ is the angle between the plane of the curve of the normal section and the plane of the first principal curve of the normal section.

If the surface is defined by the equation $z = F(x, y)$, then the following general relationship holds for the curvature of a curve of any normal section at a point on the surface (Kepr, 1994):

$$\kappa_p = \frac{F_{xx}\cos^2\alpha + 2F_{xy}\cos\alpha\cos\beta + F_{yy}\cos^2\beta}{\cos\delta\sqrt{F_x^2 + F_y^2 + 1}} \tag{4.5}$$

where δ is the angle between the normal to the surface at the given point and the sectional plane; α and β are the angles which the tangent of the curve of the normal section at the point makes with the X and Y axes respectively. Equation 4.5 allows the determination of the curvature of the curve of arbitrary normal section at a point on the surface.

The principal curvatures can be combined in different ways to define various shape properties of the surface. First, the average of the two principal curvatures at a point on a surface is known as the *mean curvature* (κ_m) (Equation 4.6). The *Gaussian curvature* (κ_g) is defined as the product of the principal curvatures (Equation 4.7).

$$\kappa_m = \frac{1}{2}(\kappa_p^1 + \kappa_p^2) \tag{4.6}$$

$$\kappa_g = \kappa_p^1 \kappa_p^2 \tag{4.7}$$

The mean curvature has the sign of predominant curvature and thus can be used to distinguish more concave regions from more convex regions. Gaussian curvature highlights negative sur-

face curvature that occurs at hyperbolic patches since these patches occur where only one principal curvature is negative. Together, the mean and Gaussian curvatures can be used to distinguish different shape types as summarised in Figure 4.4.

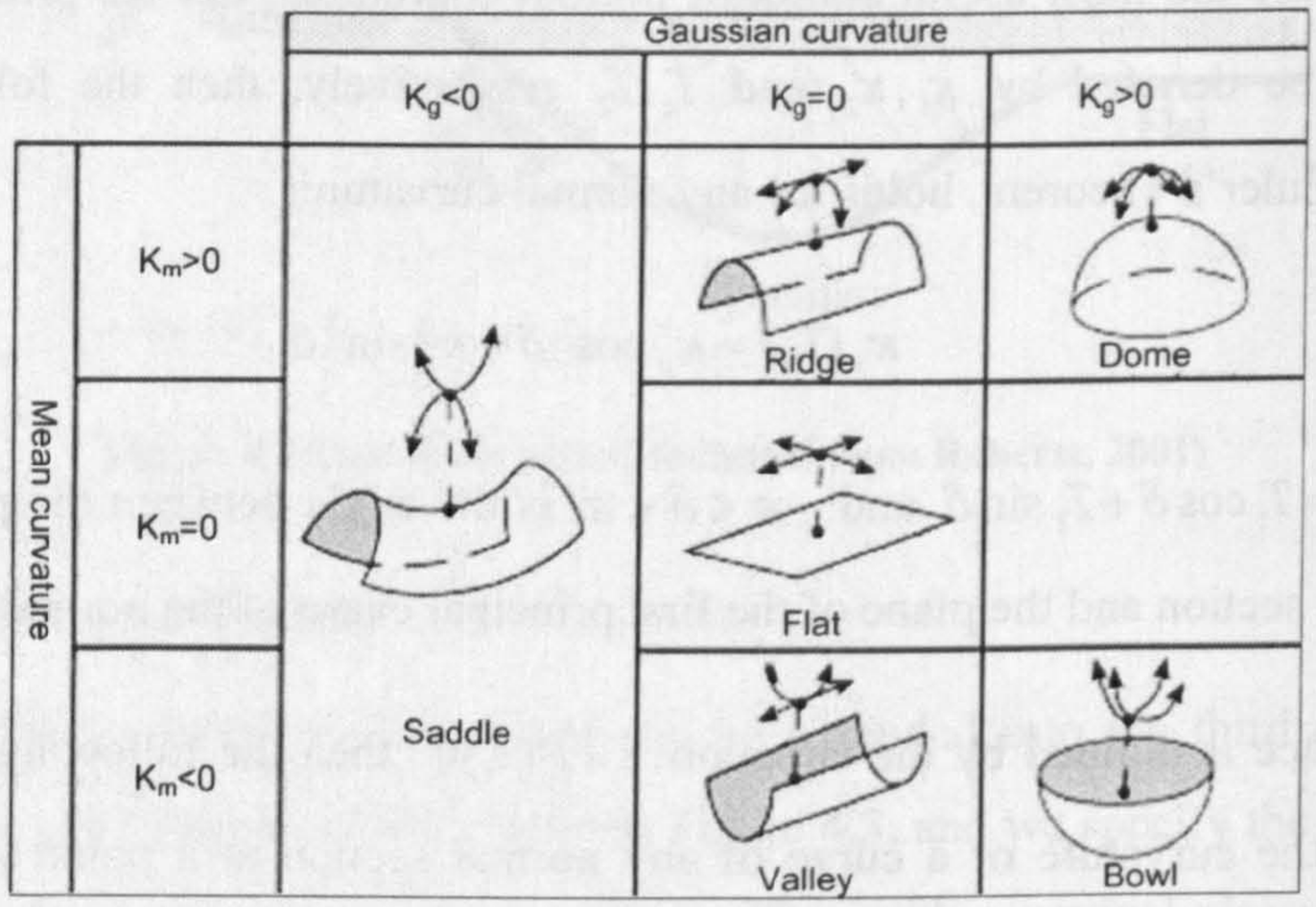


Figure 4.4: Shape classification based on curvature (adapted from Roberts, 2001).

4.2.3. Curvature Computation Methods

Various methods have been proposed in the literature for computing surface curvature depending on the data representation: point clouds, surface meshes or voxel data. A practical comparison of some of these methods has been given by Flynn & Jain (1989), Garland & Heckbert (1997), Krsek et al (1998) and Surazhsky et al (2003). Without loss of generality, methods for calculating surface curvature may be categorised as follows:

- Surface fitting methods
- Curve fitting methods
- Discrete curvature methods

4.2.3.1. Surface Fitting Methods

Surface fitting methods fit some analytic surface to the data and compute curvature from the approximating surface. Quadrics are usually used for approximation as most of the local surfaces to be recovered can be adequately approximated with quadrics. The simplest approach to obtain an approximating surface is by solving an over-determined system of linear equations (Hamann, 1993; Krsek et al, 1998; Stokely & Wu, 1992). Non-linear methods generally fit higher order surfaces, usually quadrics. Sinha & Besl (1990) also present another non-linear method based on B-splines. The most common approach is to fit a second order quadric surface of the form $z = F(x, y)$ to the data and estimate the coefficients by the Least Squares method. A more general case of the quadric fitting is described by Douros & Buxton (2002) in which a generic quadric of the form $F(x, y, z) = 0$ is fitted locally to the data, with all three dimensions given equal importance. However, this increases the number of coefficients to be estimated and the number of points required to solve the Least Squares problem. An alternative approach is the use of parametric representation: $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$ which is often preferred in differential geometry; but in practice, it introduces additional problem of finding the surface parameters, which can be non-trivial in some cases.

4.2.3.2. Curve Fitting Approach

These methods fit a family of curves lying in planar sections around each point and deduce the curvature of the surface from the curves. Martin (1998) developed a circle fitting method which selects point triplets, fits circles to these triplets, compute the curvature of each circle and estimate the normal curvature from corresponding curvature of the circles. The principal curvatures and directions are then estimated using Euler's Theorem in a Least Squares solution. Tookey and Ball (1997) extend this method using five instead of three points; but their method is only applicable to data on a regular grid. Tang and Medioni (1999) developed a method to determine the sign of Gaussian curvature and principal directions from noisy data based on tensor voting theory of Medioni. Taubin (1995) developed a method that uses a symmetric matrix with the eigenvectors as the principal directions and eigenvalues that are related by a homogenous linear transformation to the principal directions. This matrix is estimated for each vertex pairs that share a common edge in a triangular mesh. An interesting paper by Page et al (2002) extends the methods of Tang, Medioni and Taubin by employing a geodesic neighbourhood to improve curvature estimation and a voting scheme similar to the

tensor voting theory. Their method, known as normal vector voting, also results in robust classification of surface orientation to account for curvature singularities at creases and corners, and a robust estimation of both principal curvatures and directions.

4.2.3.3. Discrete Curvature Methods

These methods are only applicable to triangulated surfaces and are based on polyhedral metrics. The discrete Gaussian curvature of a vertex in a triangulation is defined as the area-weighted angular deficit with respect to the sum of angles of the neighbouring triangles (Borrelli et al, 2003).

$$\kappa_g = \frac{2\pi - \sum_{i=1}^{k} \phi_i}{1/3\, A} \tag{4.8}$$

where:

$A = \sum_{i=1}^{k} f_i$ is the sum of the areas of adjacent triangles or faces

ϕ_i is the angle deficit at a vertex calculated as the sum of the angles

Similarly, the discrete mean curvature of an edge e in a triangulation is half the angle between the normals of the faces adjacent to the edge e. The mean curvature of a portion (κ_m) of the triangulated surface is given by Equation 4.9.

$$\kappa_m = \frac{\sum m(e_i)}{1/3\, A} \tag{4.9}$$

where:

$$m(e) = \begin{cases} \gamma & \text{if } e \text{ is convex} \\ 0 & \text{if } e \text{ is plane} \\ -\gamma & \text{if } e \text{ is convex} \end{cases}$$

e_i is an edge of a vertex

$m(e)$ is the curvature function of an edge

γ is the angle between the surface normals of the adjacent faces sharing the edge e_i

$A = \sum_{i=1}^{k} f_i$ is sum of the area of adjacent triangles or faces

From the mean and Gaussian curvatures estimates, the principal curvatures can be estimated, but not principal directions.

4.2.3.4. Selected Method for Curvature Computation

A comparison of representatives of these methods has been carried out by Krsek et al (1998) and Surazhsky et al (2003). Their results indicate that quadric surface fitting method is the optimal method for curvature computation from point data. The circle fitting method is the fastest but more susceptible to noise and the discrete method is a quick method when the data is already triangulated. In our implementation, we have therefore adopted quadric surface fitting by Least Squares solution for curvature computation, the details of which are discussed in the next section.

4.2.4. Curvature Computation by Quadric Surface Fitting

The basic idea is to fit a quadric of the form $z = F(x, y)$ locally to the points. The method usually consists of the following steps:

- For each point (P_0) a neighbourhood of points is selected.
- The normal to the surface is then estimated by locally approximating the data by a plane using the Least Squares method.
- The points in the neighbourhood are transformed onto an orthogonal basis with the origin as P_0 and +z along the normal.
- A second order quadratic surface of the form is then fitted onto the points.
- Curvatures are then computed from the parameters of the fitted surface.

The second order quadric fitted is of the form shown in Equation (4.10):

$$z = F(x, y) = ax^2 + by^2 + cxy + dx + ey + f = 0 \qquad \textbf{(4.10)}$$

where (a, b, c, d, e, f) are the coefficients of the quadric surface and (x, y, z) are the coordinates of the set of points. At least 6 equations are required to solve for the 6 unknown coeffi-

cients. We define neighbourhood for each point using a 3x3 window of pixels which covers a region of $36mm^2$ square millimetres on the back surface (see Section 5.4.3 for justification of this). Using a 3x3 window gives 9 equations, the solution of which is obtained by the Least Squares method discussed in Section 4.3. With the coefficients obtained, various shape parameters can be calculated from the approximating surface using standard formulae from differential geometry.

Mean and Gaussian Curvatures

As discussed in Kepr (1994), if $z = F(x, y)$ is a surface in three-dimensional space, then the Gaussian (κ_g) and mean (κ_m) curvatures may be calculated by Equations 4.11 and 4.12 respectively:

$$\kappa_g = \frac{F_{xx}F_{yy} - F_{xy}^2}{(F_x^2 + F_y^2 + 1)^2} \tag{4.11}$$

$$\kappa_m = \frac{1}{2}\left[\frac{F_{xx}(1+F_y^2) + (1+F_x^2)F_{yy} - 2F_xF_yF_{xy}}{\left(F_x^2 + F_y^2 + 1\right)^{3/2}}\right] \tag{4.12}$$

where $F_x, F_y, F_{xx}, F_{yy}, F_{xy}$ are the partial derivatives which from Equation 4.10 are given by:

$$\begin{aligned} F_x &= 2ax + cy + d \\ F_y &= 2by + cx + e \\ F_{xx} &= 2a \\ F_{xx} &= 2b \\ F_{xy} &= c \end{aligned} \tag{4.13}$$

Principal Curvatures and Directions

The maximum and minimum principal curvatures can be obtained from Equations 4.14 and 4.15 as follows (Gray, 1997):

$$\kappa_{\max} = \kappa_m + \sqrt{\kappa_m^2 - \kappa_g} \tag{4.14}$$

$$\kappa_{\min} = \kappa_m - \sqrt{\kappa_m^2 - \kappa_g} \tag{4.15}$$

Absolute and Difference Curvatures

Other measures of a surface curvature are absolute curvature and difference between the principal curvatures. The difference curvature can highlight regions with large difference between the maximum and minimum curvatures.

$$\kappa_{abs} = |\kappa_{\max}| + |\kappa_{\min}| \tag{4.16}$$

$$\kappa_{diff} = |\kappa_{\max}| - |\kappa_{\min}| \tag{4.17}$$

Curvedness and Shape Index

Koenderink et al (1992) proposed alternative measures for surface curvature, namely, shape index which describes the type of surface and curvedness which measures the degree of the curvature. These can be computed from Equations 4.18 and 4.19 respectively.

$$\kappa_i = \frac{2}{\pi} \tan^{-1}\left(\frac{\kappa_{\min} + \kappa_{\max}}{\kappa_{\min} - \kappa_{\max}}\right) \tag{4.18}$$

$$\kappa_c = \sqrt{\frac{\kappa_{\min}^2 + \kappa_{\max}^2}{2}} \tag{4.19}$$

Profile Curvature

Intersecting the surface with a normal plane in the direction of the steepest gradient defines a curvature measure known as profile curvature (Equation 4.20). This curvature is a measure of the rate of change of gradient in the maximum gradient direction and has been used for delineating faults in terrain analysis (Wood, 1996).

$$\kappa_{prof} = \frac{F_{xx}F_x^2 + 2F_{xy}F_xF_y + F_{yy}F_y^2}{(F_x^2 + F_y^2)\sqrt{(F_x^2 + F_y^2 + 1)^3}} \tag{4.20}$$

Tangential curvature

The tangential curvature is obtained by intersecting the surface with a normal plane section in the direction perpendicular to the gradient (Wood, 1996).

$$\kappa_t = \frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{(F_x^2 + F_y^2)\sqrt{(F_x^2 + F_y^2 + 1)}} \tag{4.21}$$

Plan Curvature

The plan curvature represents the curvature of the contour created by intersecting the surface horizontally with the XY plane. This curvature measure however is not normal, and although it is used in some disciplines (e.g. Terrain Analysis), the plan curvature values are not well constrained and consequently very large values can occur at the culmination of ridges and valleys (Roberts, 2001).

$$\kappa_{plan} = \frac{F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2}{\sqrt{(F_x^2 + F_y^2)^3}} \tag{4.22}$$

4.2.5. Normal Vector and Gradient

A normal vector to a surface at the point of interest is useful parameter that measures of the orientation of the surface at that point. In Section 6.7.2, the normal vector is used to approximate orientations of spinous process landmarks in order to compute back surface rotation. The normal vector at a point (x_0, y_0) on a surface $z = F(x, y)$ is given by:

$$N = \begin{bmatrix} F_x(x_0, y_0) \\ F_y(x_0, y_0) \\ -1 \end{bmatrix} \tag{4.23}$$

where F_x and F_y are the partial derivatives of F evaluated at the point (x_0, y_0).

If the surface is defined implicitly in the form $F(x, y, z) = 0$, then the unit normal vector is given by Equation 4.24 (∇F denotes the gradient):

$$\hat{N} = \frac{\nabla F}{\sqrt{F_x^2 + F_y^2 + F_z^2}} \tag{4.24}$$

The gradient is a first-order differential operator that maps functions to vector fields. It is a generalization of the ordinary derivative, and as such conveys information about the rate of change of a function relative to small variations in its independent variables. We consider an

n-dimensional Euclidean space with orthogonal coordinates $x^1, x^2, ..., x^n$ and corresponding unit vectors $u_1, u_2, ..., u_n$. The gradient (∇F) of the function $F(x^1, x^2, ..., x^n)$ is defined to be the vector field given by the partial first order derivatives:

$$\nabla F = \sum_{i=1}^{n} \frac{\partial F}{\partial x^i} u_i$$

For a surface $z = F(x, y)$ in 3D space, the gradient is defined by a plane tangent to a surface at any given point. The direction of the gradient ∇F (known as aspect) is the direction of steepest ascent, while the magnitude is the rate of change (known as slope) in that direction. Geometrically, the gradient magnitude, given by Equation 4.25 describes the steepness of the hill $z = F(x, y)$ at the point on the hill located at $(x, y, F(x, y))$; while the aspect (Equation 4.26) is the direction of the projection of the surface normal at the point of interest onto the XY plane.

$$G = \|\nabla F\| = \sqrt{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2} \qquad \textbf{(4.25)}$$

$$A = \tan^{-1}\left[\frac{\partial F/\partial y}{\partial F/\partial x}\right] \qquad \textbf{(4.26)}$$

4.3. Least Squares Method

No measurement is ever exact, that is, every measurement contains errors. This is a fundamental and universally accepted fact. To illustrate this, consider the process of obtaining the distance between two points A and B on the ground with a steel tape. If this measurement process is repeated several times, it is very unlikely one will obtain exactly the same results. The variability of repeated measurements under similar conditions is an inherent quality of any measurement process (Wolf and Ghilani, 1997). This shows that the true value of a measured quantity is never known; hence, the exact amounts of the errors present in a set of measurements are unknown. In other words, every measurement process is a *random process (or stochastic process)* and the results of the measurements are themselves random variables. Consequently, in statistics, the general approach adopted is to collect more data than is necessary (i.e. redundant data) and find a means to arrive at a unique solution. One such method is

Least Squares which, in this investigation, has been used as the basis for various tasks, namely:

- Surface fitting for curvature computation and landmarks extraction
- Polynomial fitting for derivation of the back midline
- Coordinate transformation for the purposes of surface comparison

This section discusses the principles of the Least Squares method and the various applications to which it has been committed in this investigation.

4.3.1. Functional and Stochastic Models

In the measurement sciences, direct observations are rarely used as the required information. Instead, they are used to derive or compute other useful information such as lengths, angles, areas and volumes. The mathematical model used in the computational effort is typically composed of two parts: a *functional model* and a *stochastic model* (Anderson and Mikhail, 1998). The functional model describes the geometric or physical characteristics of the problem while the stochastic model describes the statistical properties of the elements involved in the functional model.

For example, suppose we are interested in the shape of a triangle. By measuring two of its angles, the triangle's shape will be uniquely determined. However, if one decides to measure all three internal angles, any attempt to construct a unique triangle will reveal inconsistencies in the measured angles. In this instance, the functional model is that the sum of the three angles must be 180. Since every measurement contains random errors, if the all three angles are used in the functional model, it is very unlikely that their sum will equal 180. Consequently, to obtain a unique triangle, we need both the functional model—sum of internal angles equals 180; and the stochastic model—statistical properties of the measured angles. In many cases, the measurement process is carried out by the same device, applying the same principle and under similar environmental conditions; so the measurements can be treated as having equal precision. In other cases, the measurements will be correlated and of unequal precision. Such statistical information is important and should be taken into account when deriving the required information. This introduces us to the concepts of *data fitting* or *adjustment* and the method of Least Squares.

4.3.2. Least Squares Criterion

We have emphasised the need for redundant measurements to account for the existence of random errors. However, when redundant measurements are acquired, the functional model will rarely be satisfied. Therefore, a means must be found to arrive at a set of values for the unknown parameters that are unique and consistent with the mathematical model. This process is known as data adjustment or estimation. The central idea is that of adjusting the measurements to make them consistent with the model, leading to the uniqueness of the estimated unknown parameters. Several techniques could be applied, but the Least Squares method is the simplest and can be conveniently constrained to follow the mathematical model. The method dates back to the German mathematician Gauss who stated: "but of all principles ours is the most simple; by others we would be led into the most complicated calculations" (Gauss, 1809, cited in Rousseeuw, 2003).

To derive the basic Least Squares criterion, let L be the vector of measurements and V be the vector of residuals which when added to L yields a set of new estimates $\hat{L}$ which is consistent with the functional model. Thus, we have:

$$\hat{L} = L + V$$

The statistical or stochastic properties of the measurements are expressed by weight matrix (W). Using these variables, the general form of the Least Squares criterion is given by (Mikhail, 1976; Anderson and Mikhail, 1998):

$$\phi = V^T W V \rightarrow \text{minimum} \tag{4.27}$$

The function ϕ is a scalar whose minimum is obtained by equating its partial derivatives to zero. If the measurements are uncorrelated, W will be a diagonal matrix, and Equation 4.27 will simplify to:

$$\phi = \sum_{i=1}^{n} w_i v_i^2 = w_1 v_1^2 + w_2 v_2^2 + \cdots + w_n v_n^2 \rightarrow \text{minimum} \tag{4.28}$$

Equation 4.28 states that the sum of the weighted squares of the residuals is a minimum. If the measurements are both uncorrelated and equal precision (i.e. $\mathbf{W} = \mathbf{I}$), ϕ becomes:

$$\phi = \sum_{i=1}^{n} v_i^2 = v_1^2 + v_2^2 + \cdots + v_n^2 \rightarrow \text{minimum} \tag{4.29}$$

Equation 4.29 seeks the *least sum of squares* of the residuals, hence the name '*least squares*'.

4.3.3. Least Squares Solution

There are various techniques for solving Least Squares problems. In this section, we discuss one of the most common methods, namely, solution by *parametric equations*, which is characterised by the following properties (Anderson and Mikhail, 1998):

- Each equation includes both the measurements and parameters (derived quantities)
- The number of equations equals the number of measurements
- Each equation contains only one measurement with a coefficient of 1

With these properties, the system of equations takes the following general form:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_1 &= l_1 + v_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_1 &= l_2 + v_2 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{np}x_1 &= l_n + v_n
\end{aligned} \tag{4.30}$$

where

$v_1, v_2, \cdots, v_n$ are the residuals for the n measurements

$l_1, l_2, \cdots, l_n$ are the constant terms for n equation which will usually contain the measurements

$x_1, x_2, \cdots, x_p$ are the p unknown parameters

$a_{11}, a_{12}, \cdots, a_{np}$ are numerical coefficients of the parameters

In a matrix form, we have:

$$\underset{(n\times p)}{A}\ \underset{(p\times 1)}{X} = \underset{(n\times 1)}{L} + \underset{(n\times 1)}{V} \tag{4.31}$$

where A is known as the design matrix, X is the vector of parameters and L is the vector of the measurements and V is the vector of residuals associated with the measurements.

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{21} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{nn} \end{bmatrix}, \quad X=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad L=\begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_m \end{bmatrix}, \quad V=\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

The Least Squares criterion is:

$$\phi = V^T \mathbf{W} V \rightarrow \text{minimum}$$

Substituting V from Equation 4.31, we have:

$$\begin{aligned} \phi &= (AX-L)^T W (AX-L) \\ &= A^T X^T WAX - A^T X^T WL - L^T WAX + L^T WL \\ &= A^T X^T WAX - 2L^T WAX + L^T WL \end{aligned} \tag{4.32}$$

For ϕ to be minimum, $\partial\phi/\partial X$ must be zero, thus:

$$\frac{\partial \phi}{\partial X} = 2X^T A^T WA - 2L^T WL = 0$$

$$(A^T WA)X = A^T WL \tag{4.33}$$

The set of equations (4.33) is known as *normal equations* from which the solution vector is obtained as:

$$X = (A^T WA)^{-1} A^T WL \tag{4.34}$$

With X computed, the residual vector may be computed as:

$$V = AX - L \tag{4.35}$$

and the Least Squares estimates of the measurements are computed as:

$$\hat{L} = L + V \tag{4.36}$$

4.3.4. Precision of the Estimated Parameters

The precision of the estimated parameters is given by the cofactor matrix (Q_{XX}) which can be obtained by applying a basic principle of error propagation (see Appendix E):

$$Q_{XX} = J_{XL} Q_{LL} J_{XL}^T = (A^T W A)^{-1} A^T W Q_{LL} W A (A^T W A)^{-1} \tag{4.37}$$

Noting that $W = Q^{-1}$, we have:

$$Q_{XX} = (A^T W A)^{-1} A^T W W^{-1} W A (A^T W A)^{-1}$$

$$Q_{XX} = (A^T W A)^{-1} \tag{4.38}$$

To obtain the covariance matrix for the estimated parameters, an estimate of the variance of the random error (known as reference variance) is required. This can be computed from Equation 4.39:

$$\hat{\sigma}_0^2 = \frac{V^T W V}{n-p} = \frac{V^T W V}{r} \tag{4.39}$$

in which r is the degrees of freedom obtained as the difference between the number of measurements (n) and number of parameters (p); and V is the vector of residuals obtained from Equation 4.35. The covariance matrix of the estimated parameters is given by Equation 4.40:

$$\sum_{XX} = \sigma_0^2 Q_{XX} \tag{4.40}$$

Similarly, precision information for the estimated measurements can be computed from Equation 4.41:

$$Q_{\hat{L}\hat{L}} = A(A^T W A)^{-1} A^T = A Q_{XX} A^T \text{ and } \sum_{\hat{L}\hat{L}} = \sigma_0^2 Q_{\hat{L}\hat{L}} \tag{4.41}$$

4.3.5. Goodness of Fit

Analysis of the goodness of fit is an important step in the Least Squares solution. The goodness of fit describes how well the measurements conform to the mathematical model; in other words, how best the model fit the original data. The most common techniques used to assess the goodness of fit are discussed in this section.

4.3.5.1. Residual Analysis

The first step in evaluating the goodness of the fit is to analyse the distribution of the residuals. Assuming the model fitted to the data is correct, the residuals approximate the random errors. If the residuals appear to have a random distribution, it generally suggests that the model fits the data well. On the other hand, if the residuals show a systematic pattern, it is a good sign that the model fits the data poorly.

4.3.5.2. Goodness of Fit Statistics

Goodness-of-fit statistics are numerical measures that describe how well the measurements conform to the mathematical model. Let y be the measurements, $\bar{y}$ be the mean and $\hat{y}$ be the Least Squares estimates which are consistent with mathematical model. The following goodness-of-fit statistics are commonly used:

- Sum of squared errors (SSE) measures the combined deviation of the independent variables from the fit, and is computed by simply summing the individual deviations.

$$SSE = \sum_{i=1}^{n} w_i (y_i - \bar{y})^2 \qquad \textbf{(4.42)}$$

  Due to cancellation effect of positive and negative deviations, the SSE is generally not a good measure of the goodness-of-fit.

- R-square statistic measures how successful the fit is in explaining the variation of the data. It is also called the coefficient of determination. R-square is defined as the ratio of the sum of squares of the regression (SSR) and the total sum of squares (SSE):

$$\text{R-Squared} = 1 - \frac{SSE}{SSR} \qquad \textbf{(4.43)}$$

  where $SSR = \sum_{i=1}^{n} w_i (\hat{y}_i - \bar{y})^2$ and $SSE = \sum_{i=1}^{n} w_i (y_i - \bar{y})^2$

  R-square takes a value between 0 and 1, with a value closer to 1 indicating a better fit. If the number of parameters in the mathematical model is increased, R-square might increase although the fit may not improve. To avoid this situation, the adjusted R-square statistic is used.

- Adjusted R-square uses the R-square statistic and adjusts it based on the degrees of freedom. The adjusted R-square statistic takes a value less than or equal to 1, with a value closer to 1 indicating a better fit.

$$\text{Adjusted R-Square} = 1 - \frac{SSE(n-1)}{SSR(p-1)} \tag{4.44}$$

- Root mean squared error, also known as standard error of the fit is defined as:

$$RMSE = \sqrt{\frac{SSE}{n-p}} \tag{4.45}$$

A RMSE value closer to 0 indicates a better fit.

4.3.5.3. Confidence and Prediction Intervals

In addition to residuals analysis and goodness-of-fit statistics, confidence intervals can be calculated for the estimated parameters, reference variance and the fitted function. Confidence intervals are useful in statistical analysis as they give the range within which the unknown true value of the population is expected to lie at a certain confidence level.

In Equation 4.39, an estimate of the reference variance (variance of the random error) is used in order to compute the covariance matrix. We can now construct a confidence interval for the reference variance using the χ^2 distribution. The two-tailed $(1-\alpha)100\%$ confidence interval for σ_0^2 is given by:

$$\frac{r\hat{\sigma}_0^2}{\chi^2_{r,\alpha/2}} < \sigma_0^2 < \frac{r\hat{\sigma}_0^2}{\chi^2_{r,1-\alpha/2}} \tag{4.46}$$

If the reference variance is incorrect or the mathematical model is inadequate, then σ_0^2 will typically fall outside this interval. Several reasons may account for the failure of σ_0^2, such as inadequacy or incorrect formulation of the mathematical model; failure to linearise non-linear system of equations; failure to eliminate any significant systematic or gross errors from measurements. In such cases, the model and the computation process will have to be revisited.

Using the t-distribution, we can also construct confidence intervals for estimated parameters of the fitted function. The $(1-\alpha)100\%$ confidence interval for the parameters is given by Equation 4.47:

$$X \pm t_{r,\alpha/2}\sqrt{\sigma_{XX}} \tag{4.47}$$

where X is the vector of estimated parameters, σ_{XX} is a vector of standard deviations of the estimated parameters, obtained from the diagonal elements from the covariance matrix of the parameters. The width of these confidence intervals is a measure of the overall quality of the fit.

Similarly, the confidence interval for the fitted function itself at a particular value x_0 is given by Equation 4.48:

$$\hat{L}_{x_0} \pm t_{r,\alpha/2}\sqrt{J_{x_0}\sigma_{XX}J_{x_0}^T} \tag{4.48}$$

where $\hat{L}_{x_0}$ is the vector of the Least Squares estimates of the measurements evaluated at x_0 and J_{x_0} is defined as a row vector of the Jacobian of the fitted function with respect to the parameters and evaluated at a specified predictor x_0.

We can also calculate the so-called prediction interval for a future measurements Y_0 given a particular value x_0. The $(1-\alpha)100\%$ prediction interval for a future observation Y_0 at the value x_0 is given by Equation 4.49:

$$\hat{L}_{x_0} \pm t_{r,\alpha/2}\sqrt{\hat{\sigma}_0^2 + J_{x_0}\sigma_{XY}J_{x_0}^T} \tag{4.49}$$

where all variables have their usual meanings.

4.3.6. Least Squares Solution of Non-Linear Systems

When the system of equations being dealt with is non-linear, the equations may be linearised by approximating with a first-order Taylor series, after which linear Least Squares methods can be applied. Let f be a non-linear function involving a set of independent variables x, y, then the Taylor series approximation of f is given by (Abramowitz and Stegun, 1972):

$$\begin{aligned} f = f(x,y) = f(x_0,y_0) &+ \frac{(\partial f/\partial x)_0}{1!}dx + \frac{(\partial^2 f/\partial x)_0}{2!}dx + \cdots + \frac{(\partial^n f/\partial x^n)_0}{n!}dx \\ &+ \frac{(\partial f/\partial y)_0}{1!}dy + \frac{(\partial^2 f/\partial y)_0}{2!}dy + \cdots + \frac{(\partial^n f/\partial y^n)_0}{n!}dy + R_n \end{aligned} \tag{4.50}$$

where x_0 and y_0 are approximations of x and y; $f(x_0, y_0)$ is the function evaluated at x_0, y_0; R_n is the reminder term, known as Lagrange remainder; and dx and dy are the corrections to the initial approximations. As more terms are included, a more exact approximation is obtained. However, as the order of each successive term increases, its contribution to the overall expression diminishes. In practice, only the first-order terms are often used, and Equation 4.50 reduces to:

$$f = f(x, y) = f(x_0, y_0) + \left(\frac{\partial f}{\partial x}\right)_0 dx + \left(\frac{\partial f}{\partial y}\right)_0 dy \quad \textbf{(4.51)}$$

The only unknowns in the Taylor's approximation are the corrections dx and dy which must be applied to the initial approximations to obtain the computed quantities. Thus, the Least Squares solution must be iterated to yield accurate results. The process of the Least Squares solution of a non-linear system can be summarised as follows:

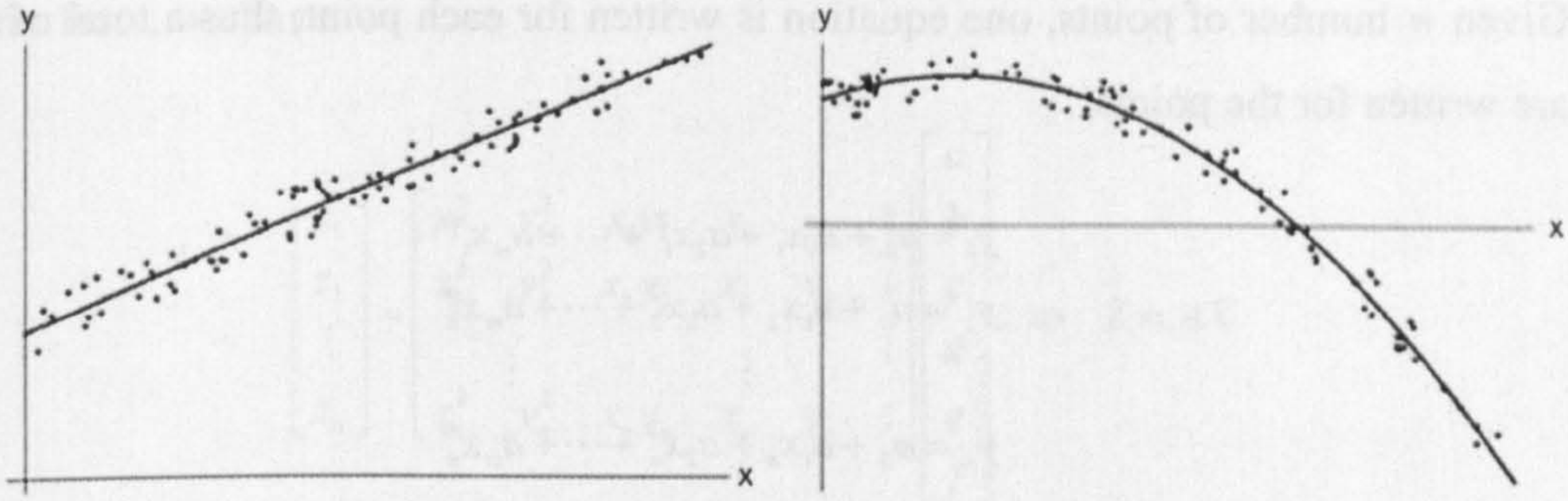
- Determine the Taylor series approximation for the equations
- Determine the initial approximations for the unknowns
- Use the Least Squares method to solve for the corrections to the unknowns
- Apply the corrections to the initial approximations
- Repeat steps 1-4 until the corrections become negligible

4.3.7. Applications of Least Squares

This section discusses the various applications to which the Least Squares method has been committed in this study, namely, curve fitting, surface fitting and coordinate transformation.

4.3.7.1. Polynomial Fitting

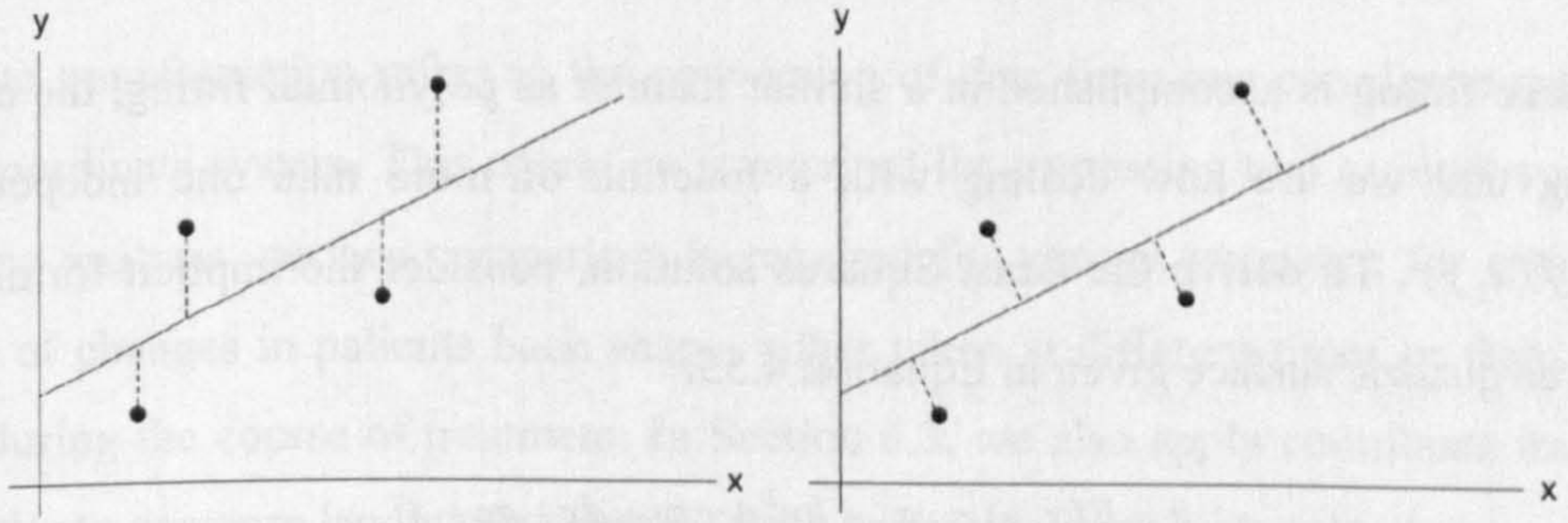
Polynomial curve fitting is a common problem in statistical data analysis. The problem consists of finding the best-fit curve to a given set of points by minimizing the sum of the squares of the residuals of the points from the curve (Figure 4.5).



(a) First degree polynomial (b) Second degree polynomial

Figure 4.5: Fitting a polynomial to a set of points (adapted from Weisstein, 2002)

The actual residuals are the perpendicular deviations from the fitted curve (Figure 4.6). However, in practice, the vertical offsets from the polynomial (or surface) are often minimised instead of the perpendicular offsets. There are several reasons for this choice. First, the fitting function is conveniently expressed in terms of the independent variable x which estimates y for a given x. This allows uncertainties of the data points along the X and Y axes to be incorporated in a simple manner (Weisstein, 2002). Moreover, using vertical deviations in a much simply fitting function for the fitting parameters than would be obtained using a fit based on perpendicular deviations. Generally, the difference between vertical and perpendicular fits is quite negligible for a reasonable number of noisy data points.



(a) Vertical offsets (b) Perpendicular offsets

Figure 4.6: Vertical versus perpendicular offsets (adapted from Weisstein, 2002)

To derive the Least Squares solution, consider the general form of a polynomial of degree m:

$$y = f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \tag{4.52}$$

where $a_0, a_1 \cdots a_m$ are the coefficients.

Given n number of points, one equation is written for each point; thus a total of n equations are written for the points:

$$\begin{aligned} y_1 &= a_0 + a_1 x_1 + a_2 x_1^2 + \cdots + a_m x_1^m \\ y_2 &= a_0 + a_1 x_2 + a_2 x_2^2 + \cdots + a_m x_2^m \\ &\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ y_n &= a_0 + a_1 x_n + a_2 x_n^2 + \cdots + a_m x_n^m \end{aligned} \tag{4.53}$$

Equation 4.53 can be represented concisely in a matrix form:

$$\begin{bmatrix} y_1 \\ y_2 \\ \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ & & & & \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} \Rightarrow Y = AX \tag{4.54}$$

where A is the design matrix, Y is the vector of measurements and X the vector of parameters. Adding residuals, Equation 4.54 can be written in a form for Least Squares solution, from which the solution vector can be obtained as in Equation 4.34. If the covariance (or weight) matrix is known, it can be incorporated into the solution. Precision quantities for the fitting parameters can be evaluated from the co-factor matrix discussed in Section 4.3.4.

4.3.7.2. Surface Fitting

Surface fitting is accomplished in a similar manner as polynomial fitting, the only difference being that we are now dealing with a function of more than one independent variable $z = f(x, y)$. To derive the Least Squares solution, consider the implicit form of the second degree quadric surface given in Equation 4.55:

$$z = F(x, y) = ax^2 + by^2 + cxy + dx + ey + f \tag{4.55}$$

where a, b, c, d, e, f are the coefficients.

For n number of points, we have the following equations:

$$\begin{aligned} z_1 &= ax_1^2 + by_1^2 + cx_1y_1 + dx_1 + ey_1 + f \\ z_2 &= ax_2^2 + by_2^2 + cx_2y_2 + dx_2 + ey_2 + f \\ &\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ z_n &= ax_n^2 + by_n^2 + cx_ny_n + dx_n + ey_n + f \end{aligned} \tag{4.56}$$

In matrix form, we have:

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} x_1^2 & y_1^2 & x_1y_1 & x_1 & y_1 & 1 \\ x_2^2 & y_2^2 & x_2y_2 & x_2 & y_2 & 1 \\ & & & & & 1 \\ x_n^2 & y_n^2 & x_ny_n & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} \Rightarrow \quad Z = AX \tag{4.57}$$

Adding residuals, Equation 4.57 can be written the general form for the Least Squares solution:

$$Z + V = AX$$

The solution vector and precision information are obtained as usual.

In Section 4.2.4, a plane is fitted to the local neighbourhood of points as part of the curvature computation process. The Least Squares solution is the same as discussed above, except that the equation of a plane (given by Equation 4.58) is used in formulating the linear system.

$$z = ax + by + c \tag{4.58}$$

4.3.7.3. Coordinate Transformation

Coordinate transformation refers to the conversion of data from one coordinate system into another coordinate system. This operation is required for comparing two or more surfaces. In back shape analysis, surface comparison is required for various purposes, for example, the detection of changes in patients back shape, either taken at different times or from different patients during the course of treatment. In Section 6.3, we also apply coordinate transformation in order to compare landmarks detected from curvature maps with palpation.

In this section, we discuss a mathematical model for Cartesian coordinate conversion based on a 7-parameter similarity transformation, which preserves angles and changes all distances in the same ratio (Anderson and Mikhail, 1998). This transformation model requires at least three points whose coordinates are known in both coordinate systems.

To develop the transformation model, let $S_1(X_1, Y_1, Z_1)$ and $S_2(X_2, Y_2, Z_2)$ be two surfaces defined in two independent coordinate systems. Transforming S_2 into S_1 system involves the following steps:

- Creating a common origin for the two systems (translation)
- Making the reference axes of the two systems parallel (rotation)
- Making equal dimensions in the two coordinate systems (scaling)

In other words, the transformation involves 7 parameters—three rotations to align the coordinate axes, three translations to make a common origin, and one scale factor to ensure equal dimensions in both reference systems (see Figure 4.7).

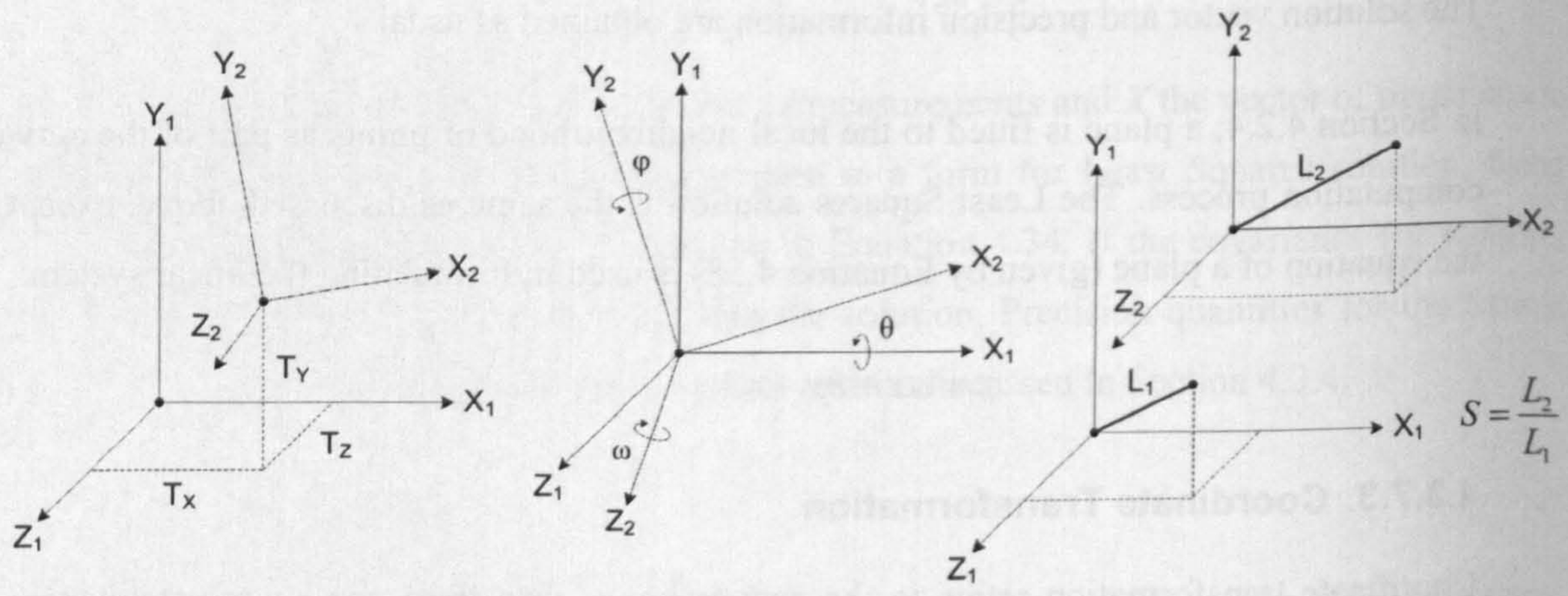


Figure 4.7: 7-parameter 3D similarity coordinate transformation

Thus, the transformation can be expressed by the following model:

$$\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} = SR \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} + \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix} \tag{4.59}$$

In Equation 4.59, S is the scale factor, T_X, T_Y, T_Z are the translations about X_1, Y_1 and Z_1 axes respectively, and R is the rotation matrix which is built from individual rotations ω, ϕ and κ about X_2, Y_2 and Z_2 axes respectively. The rotation matrix R represents the combined individual rotations about their respective axes. When rotating about the X-axis, the X coordinate remains constant while the Y and Z coordinates are transformed. Similarly, when rotating

about the X and Y axes, the respective X and Y coordinates remain constant. Thus the three rotation matrices are as follows:

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega) & -\sin(\omega) \\ 0 & \sin(\omega) & \cos(\omega) \end{bmatrix}, R_y = \begin{bmatrix} \cos(\phi) & 0 & \sin(\phi) \\ 0 & 1 & 0 \\ -\sin(\phi) & 0 & \cos(\phi) \end{bmatrix}, R_z = \begin{bmatrix} \cos(\kappa) & -\sin(\kappa) & 0 \\ \sin(\kappa) & \cos(\kappa) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \textbf{(4.60)}$$

These individual rotation matrices about the axes can be combined to obtain a single rotation matrix. The elements of R, shown in Equation 4.61, depends on the order in which the rotations are combined

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad \textbf{(4.61)}$$

For rotation about X, Y and Z in that order (R_{zyx}), we have:

$$\begin{aligned} r_{11} &= \cos(\phi)\cos(\kappa) \\ r_{12} &= \sin(\omega)\sin(\phi)\cos(\kappa) - \cos(\omega)\sin(\kappa) \\ r_{13} &= \cos(\omega)\sin(\phi)\cos(\kappa) + \sin(\omega)\sin(\kappa) \\ r_{21} &= \cos(\phi)\sin(\kappa) \\ r_{22} &= \sin(\omega)\sin(\phi)\sin(\kappa) + \cos(\omega)\cos(\kappa) \\ r_{23} &= \cos(\omega)\sin(\phi)\sin(\kappa) - \sin(\omega)\cos(\kappa) \\ r_{31} &= -\sin(\omega) \\ r_{32} &= \sin(\omega)\cos(\phi) \\ r_{33} &= \cos(\omega)\cos(\phi) \end{aligned}$$

For rotation about Z, Y and X (R_{xyz}) we have:

$$\begin{aligned} r_{11} &= \cos(\phi)\cos(\kappa) \\ r_{12} &= -\cos(\phi)\sin(\kappa) \\ r_{13} &= \sin(\phi) \\ r_{21} &= \cos(\omega)\sin(\kappa) + \sin(\omega)\sin(\phi)\cos(\kappa) \\ r_{22} &= \cos(\omega)\cos(\kappa) - \sin(\omega)\sin(\phi)\sin(\kappa) \\ r_{23} &= -\sin(\omega)\cos(\phi) \\ r_{31} &= \sin(\omega)\sin(\kappa) - \cos(\omega)\sin(\phi)\cos(\kappa) \\ r_{32} &= \sin(\omega)\cos(\kappa) + \cos(\omega)\sin(\phi)\sin(\kappa) \\ r_{33} &= \cos(\omega)\cos(\phi) \end{aligned}$$

Equation 4.59 can now be expanded to obtain the following set of equations:

$$\begin{aligned} X_1 &= S(r_{11}X_2 + r_{12}Y_2 + r_{13}Z_2) + T_X \\ Y_1 &= S(r_{21}X_2 + r_{22}Y_2 + r_{23}Z_2) + T_Y \\ Z_1 &= S(r_{31}X_2 + r_{32}Y_2 + r_{33}Z_2) + T_Z \end{aligned} \tag{4.62}$$

Equations 4.62 involve seven unknown parameters ($S, \omega, \phi, \kappa, T_X, T_Y, T_Z$), thus seven equations are required for a unique solution. Each point yields three equations, hence a minimum of 3 points are required whose coordinates are known in both systems (called *common points* or *control points*). Since three points gives 9 equations (more than required), a least squares solution is sought which minimises the sum of the squared residuals.

Equations 4.62 are non-linear in the unknown parameters, and must therefore be linearised by using Taylor's approximation. Applying the method described in Section 4.3.6 leads to the following linearised equations:

$$\begin{aligned} X_1 &= (X_1)_0 + (X_S)_0\, dS + (X_\omega)_0\, d\omega + (X_\phi)_0\, d\phi + (X_\kappa)_0\, d\kappa + (X_{T_X})_0\, dT_X \\ Y_1 &= (Y_1)_0 + (Y_S)_0\, dS + (Y_\omega)_0\, d\omega + (Y_\phi)_0\, d\phi + (Y_\kappa)_0\, d\kappa + (Y_{T_Y})_0\, dT_Y \\ Z_1 &= (Z_1)_0 + (Z_S)_0\, dS + (Z_\omega)_0\, d\omega + (Z_\phi)_0\, d\phi + (Z_\kappa)_0\, d\kappa + (Z_{T_Z})_0\, dT_Z \end{aligned} \tag{4.63}$$

where subscript 0 means evaluated at the approximations and X_j, Y_j, Z_j are the partial derivatives with respective to the j^{th} transformation parameter $(S, \omega, \phi, \kappa, T_X, T_Y, T_Z)$.

The partial derivative terms are:

$$X_S = \frac{\partial X_1}{\partial S} = r_{11}X_2 + r_{12}Y_2 + r_{13}Z_2$$

$$X_\omega = \frac{\partial X_1}{\partial \omega} = 0$$

$$X_\phi = \frac{\partial X_1}{\partial \phi} = S\left[-\sin(\phi)\cos(\kappa)X_2 + \sin(\phi)\sin(\kappa)Y_2 + \cos(\phi)Z_2\right]$$

$$X_\kappa = \frac{\partial X_1}{\partial \kappa} = S[r_{12}X_2 - r_{11}Z_2]$$

$$X_{T_X} = \frac{\partial X_1}{\partial T_X} = 1$$

$$Y_S = \frac{\partial Y_1}{\partial S} = r_{21}X_2 + r_{22}Y_2 + r_{23}Z_2$$

$$Y_\omega = \frac{\partial Y_1}{\partial \omega} = -S[r_{31}X_2 + r_{32}Y_2 + r_{33}Z_2]$$

$$Y_\phi = \frac{\partial Y_1}{\partial \phi} = S\left[\sin(\omega)\cos(\phi)\cos(\kappa)x_2 - \sin(\omega)\cos(\phi)\sin(\kappa)y_2 + \sin(\omega)\sin(\phi)z_2\right]$$

$$Y_\kappa = \frac{\partial Y_1}{\partial \kappa} = S[r_{22}X_2 - r_{21}Y_2]$$

$$Y_{T_Y} = \frac{\partial Y_1}{\partial T_Y} = 1$$

$$Z_S = \frac{\partial Z_1}{\partial S} = r_{31}X_2 + r_{32}Y_2 + r_{33}Z_2$$

$$Z_\omega = \frac{\partial Z_1}{\partial \omega} = S[r_{21}X_2 + r_{22}Y_2 + r_{23}Z_2]$$

$$Z_\phi = \frac{\partial Z_1}{\partial \phi} = S\left[-\cos(\omega)\cos(\phi)\cos(\kappa)X_2 + \cos(\omega)\cos(\phi)\sin(\kappa)Y_2 - \cos(\omega)\sin(\phi)Z_2\right]$$

$$Z_\kappa = \frac{\partial Z_1}{\partial \kappa} = S[r_{32}X_2 - r_{31}Y_2]$$

$$Z_{T_Z} = \frac{\partial Z_1}{\partial T_Z} = 1$$

Adding residuals, simplifying and rearranging, Equations 4.63 can be represented in the general matrix notation for the Least Squares solution:

$$\begin{bmatrix} (X_S)_0 & (X_\omega)_0 & (X_\phi)_0 & (X_\kappa)_0 & 1 & 0 & 0 \\ (Y_S)_0 & (Y_\omega)_0 & (Y_\phi)_0 & (Y_\kappa)_0 & 0 & 1 & 0 \\ (Z_S)_0 & (Z_\omega)_0 & (Z_\phi)_0 & (Z_\kappa)_0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} dS \\ d\omega \\ d\phi \\ d\kappa \\ dT_X \\ dT_Y \\ dT_Z \end{bmatrix} = \begin{bmatrix} X_1 - (X_1)_0 \\ Y_1 - (Y_1)_0 \\ Z_1 - (Z_1)_0 \end{bmatrix} + \begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix} \quad \textbf{(4.64)}$$

Each common point contributes 3 equations corresponding to the three rows of the coefficients to matrices A, L and V. Thus, if there are n number of common points, the following matrix equation results:

$$\underset{(3n\times 7)}{A}\ \underset{(7\times 1)}{X} = \underset{(3n\times 1)}{L} + \underset{(3n\times 1)}{V} \quad \textbf{(4.65)}$$

Since the unknowns in the linearised equations are corrections to the parameters, the Least Squares solution must be iterated till the corrections are negligible. Once the solution has reached satisfactory convergence, the latest approximations for the unknowns are the values for the transformation parameters. In this study, we formulate a test for convergence based on the combined root mean squared error of the fit:

$$RMSE = \sqrt{\frac{SSE}{3n-7}} \tag{4.66}$$

If the RMSE is less than a predefined limit, then the solution has converged successfully (in this study, we use a limit of $1x10^{-7}$). The estimated transformation parameters can then be used to transform any coordinates in S_2 system into S_1 system using Equation 4.62.

4.4. Chapter Summary

This chapter presented the mathematical foundation of the shape analysis techniques applied in this study. The concept of surface curvature, the various curvature measures and the computational methods were discussed. The Least Squares method was also discussed because it underlies many of the data analysis methods used in subsequent chapters. The Least Squares model represents a simple but effective method for analysing measurements with normally distributed errors, especially if the solution is properly constrained. The practical applications to which the Least Squares method has been committed in this investigation were also discussed, including polynomial fitting, surface fitting and coordinate transformation. Having laid down the relevant mathematical foundation, the next chapters present the shape analysis and visualization methods developed in this investigation.

A crucial aspect of any form of data interpretation and analysis is how to deal with the presence of errors in the data. Every measurement process contains random errors, and if the measurements are to be useful, the methods of interpretation should properly account for these errors. The next chapter presents the techniques that have been used in this investigation to determine nature of these errors and the methods used to effectively reduce their effect.

5. Error Reduction, Curvature Computation and Visualization

5.1. Introduction

This chapter presents three aspects of this research, namely, error reduction, curvature visualization and implementation issues. In Chapter 4, it was noted that all measurements have inherent errors and that these errors will be propagated to derived quantities. Unlike many industrial objects, human subjects are mobile and flexible, leading to further complications. It is simply difficult, if not impossible, to ensure that the human body remains in a static position for the time required to scan (maximum of 2.5 seconds in fine mode, see Section 3.3.6). Moreover, curvature computation depends upon second derivatives which emphasise random errors in the original data. Further still, the curvature quantities are not the final results—surface measurements are to be derived from them. Consequently, effective reduction of the errors in the original measurements constitutes a crucial component in back shape analysis. In Section 5.2, this issue is addressed through scanner calibration and empirical noise smoothing operations to ensure stable curvature computation.

Section 5.3 focuses on development of methods for effective visualization of the back shape. The objective here is to provide diagnostic visual tools that will enable the exploration of back deformities without the need for physical manipulation of the subject. A number of visualization techniques are explored, including continuous colour scales, discrete colour scales, contour maps and three-dimensional views.

The chapter concludes with addressing the major implementation and design issues regarding data structures and conversion, choice of development tools and system design. The software system which has been developed is the only platform used for testing, visualizing and conducting all the experiments in this investigation.

5.2. Error Assessment and Reduction

In Section 4.3, it was discussed that every measurement contains errors, i.e., every measurement process is a random process and that the results of the measurements are themselves random variables. Due to the existence of errors, before the original measurements can be utilised, they must be processed to account for the errors. Accounting for existence of errors in measurements involves two main steps (Wolf and Ghilani, 1997), namely:

- Performing statistical analysis of the measurements to assess the amount and distribution of the errors to determine whether or not the errors are within acceptable limits. This is a crucial step since the errors will be propagated to any quantities derived from the original measurements, if proper steps are not taken to reduce their effects.
- If the errors are within acceptable limits, adjusting and redistributing them so that they conform to the required mathematical model.

In Section 4.3, the Least Squares method was presented as a useful method for adjusting the measurements to make them conform to the mathematical model. In this section, the errors themselves are addressed. First, the sources and nature of the errors are determined. Then various experiments are conducted in attempt to estimate the distribution of the errors and to find a method to minimise their effect.

5.2.1. Sources and Classification of Errors

Errors in any measurement process stem from four main sources, namely:

- *Instrumental.* These are caused by imperfections in the construction of the measuring instrument. Instrumental errors are usually systematic and can be eliminated from the measurements.
- *Environmental.* These are errors caused by the variations in environmental conditions during the measurement process which may have significant effect on the measurement. These include changing environmental conditions such as, variations in atmospheric pressure, temperature, humidity, etc.
- *Personal.* These errors are due to limitations in human senses in using the measuring device (such as ability to perfectly level the instrument) as well as errors caused by the carelessness or mistakes of the human operator.

- *Object movement*. These are errors caused by movement of a mobile and/or flexible object like the human body during data capture. Object movement can be rigid (i.e. whole body movement), non-rigid (e.g. stretching) or a combination of both. In rigid-body movement, the whole body moves equally in the form of rotation or translation or a combination of both. This will typically manifest in the measurements as statistically large error if the movement is significant. With non-rigid movement different parts of the body move differently. An example is movement of the head during data capture which can cause unequal skin movement in other parts of the body. With a mobile and flexible object like the human body, these random movements are difficult to control, and processing techniques must cope with them.

Without loss of generality, measurement errors, irrespective of their source can be classified under three headings, namely, gross, systematic and random errors.

- *Gross errors*. These are statistically large errors that must be detected and removed from the measurements. Errors caused by mistakes and carelessness of the human operator as well as significant whole-body movement will typically fall under this category. Since gross errors are largely due to the imperfections of the individual operating the instrument, one way to reduce them is to ensure great care when operating the instrument. This implies that the instrument should be operated according to the manufacturer's guidelines, such as, ensuring that the object-scanner distance does not exceed 2.6m.

- *Systematic errors*. These errors follow some known physical laws and thus are predictable. Errors caused by device imperfections and some known environmental conditions will normally fall under this category. These errors usually follow a systematic pattern, and if the measurements are repeated under the similar conditions, the same pattern will be repeated. If known, these errors must be eliminated from the measurements before any used can be made of them. We address systematic errors through a calibration process in which the scanned data of regular standard objects are compared with their known measurements (see Section 5.2.2). This also allows us to verify the nominal accuracy of the instrument specified by the manufacturer.

- *Random errors*. These are the errors that remain after all gross and systematic errors have been removed from the measurements. They are generally small in magnitude

and caused by factors that are impossible to control, such as unknown environmental variations and non-rigid skin movement during data capture. Random errors generally do not follow any exact physical laws and thus must be dealt with by the mathematical theory of probability. Since it is impossible to perfectly control all the elements involved both in the construction and operation of the instrument, there will always be random errors in the measurements. We deal with random errors in the software environment using error reduction methods such as *smoothing*. A further reason to reduce the effect of random errors is that the curvature quantities that are to be computed from the original data depend upon second derivatives (see Sections 5.2.1 and 5.2.4). Consequently the results are quite susceptible to noise in the original data; and yet, the curvature quantities are not the final results, further measurements such as lengths, angles and areas are to be derived from them (see Chapter 6).

5.2.2. Calibration of the Laser Scanner

The purpose of the calibration experiment is two-fold, namely:

- To determine whether there are any significant systematic errors or drift in the measurements obtained with the Minolta laser scanner (VI-900)

- To verify the accuracy with which point data are measured, as reported by the instrument manufacturer (Minolta Corporation).

To perform the above tasks, two standard geometric objects were designed and manufactured to within a precision of 0.1mm. Both objects were made of wood, with brownish colour, specifically chosen to imitate the skin colour of the subjects scanned. The scanner and the objects were mounted on a level platform at a height of 1.3m from a horizontal floor and with scanner-object distance set to 1.6metres, the distance typically used for scanning human subjects. Various phases of the objects were then scanned and the distances shown in Figure 5.1 were measured in each coordinate axis and compared with the known distances.

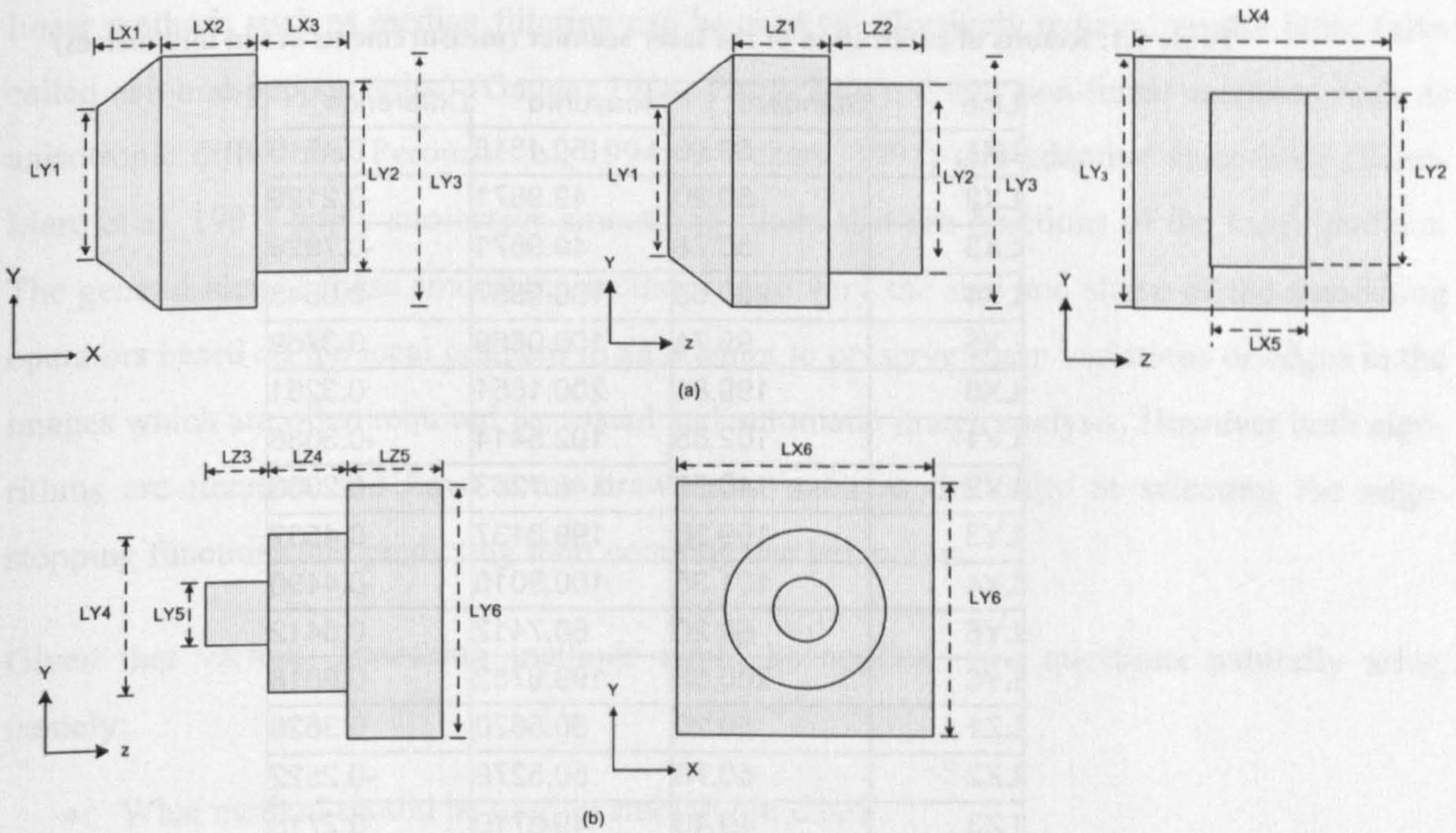


Figure 5.1: Geometric objects used for scanner calibration

Table 5.1 gives the differences between the measured and the known distances. The RMS values for the measurements in each axis were computed as 0.51mm in X, 0.45 in Y and 0.29 in Z. These RMS values represent the average deviation of the measured from the known distances and compare quiet well with the calibration data received from Minolta (see Table 3.1). These results indicate that there are no significant systematic errors or bias associated with the laser scanner if it is operated according to the manufacturer's guidelines.

In the above experiment, scanning was performed in fine mode which took up to 2.5 seconds to complete. The scan time can be reduced to 0.3 seconds in fast mode, but this will also degrade the accuracy obtainable. For scanning of human subjects, the highest accuracy is required for curvature analysis. On the other hand, a problem with human subjects is maintaining static posture for the duration of the scan: posture changes can introduce random errors in the measurement. However, as the next section demonstrates, the effects of these random errors can be reduced via smoothing operations in the software environment. Consequently, to maintain the highest achievable accuracy, the human subjects examined in this study were scanned in fine mode.

Table 5.1: Results of calibration of the laser scanner (measurements are in millimetres)

Line	Standard	Measured	Difference
LX1	50.00	50.4518	0.4518
LX2	50.20	49.9871	-0.2129
LX3	50.78	49.9871	-0.7929
LX4	200.65	199.9957	-0.6543
LX5	99.71	100.0869	0.3769
LX6	199.84	200.1651	0.3251
LY1	102.85	102.5414	-0.3086
LY2	140.52	140.7263	0.2063
LY3	199.36	199.8137	0.4537
LY4	101.35	100.9010	-0.4490
LY5	60.20	60.7412	0.5412
LY6	200.58	199.9782	-0.6018
LZ1	50.20	50.5620	0.3620
LZ2	50.78	50.5278	-0.2522
LZ3	49.40	49.6710	0.2710
LZ4	49.66	49.8910	0.2310
LZ5	50.62	50.9189	0.2989

5.2.3. Noise Smoothing

As already pointed out, random errors will still remain in the measurements even after known systematic errors have been eliminated. These random errors must now be dealt with in the software environment using smoothing methods. The problem of reducing the effect of errors (or *noise*) in image data is a fundamental problem in image processing. The difficulty lies in identifying which features in the data or image are genuine and which are due to errors or noise. The main idea behind many smoothing methods is to replace the erroneous pixels with some values derived from the local neighbourhood of that pixel. Thus, the general assumption is that local variations of pixel values in the image will be gradual so that pixels that are significantly different from their neighbours can often be attributed to noise. The drawback of smoothing, of course, is that sharp variations in pixel values (called *edges*) that are genuine are smoothed as well. Thus, minimal smoothing should be applied since over smoothing can actually modify the curvature being computed.

Various linear and non-linear methods have been proposed in the literature for reducing noise in image data. The effectiveness of any smoothing method however depends on the nature of the noise being dealt with. Linear smoothing methods such as simple averaging and Gaussian smoothing are found to perform well on images with uniform or Gaussian noise, while non-

linear methods such as median filtering can be used to effectively reduce impulse noise (also called salt-and-pepper noise) (Gauge, 1998; Pratt, 2001). Other non-linear methods such as anisotropic diffusions (Perona et al, 1994; Weickert, 1997) and adaptive smoothing (Saint-Marc et al, 1991) apply successive smoothing filters that are functions of the local gradient. The general idea of these smoothing methods is to vary the size and shape of the smoothing operators based on the local gradient in an attempt to preserve sharp variations or edges in the images which are often required for visual and automatic image analysis. However both algorithms are iterative and have some drawbacks, such as difficulty in selecting the edge-stopping functions and predicting their convergence behaviour.

Given that various smoothing methods could be applied, two questions naturally arise, namely:

- What method should be used to smooth our data?
- How much smoothing should be applied?

Since the effectiveness of any smoothing method depends on the nature of the noise, if the underlying distribution of the errors is known, an appropriate smoothing method may be devised. Thus, we address the first issue by performing empirical estimation of the distribution of errors or noise in our data. This is achieved by comparing actual laser data of a smooth machined cylinder with an ideal cylinder with the same dimensions. The second issue of the limit of the smoothing is important since over smoothing can result in a modification of the true curvature of the back surface. We address this issue by empirical approach—by comparing the curvature of the actual scanned data of the cylinder computed with increasing amount of smoothing with the true curvature of the cylinder.

5.2.4. Estimating the Error Distribution

To estimate the distribution of noise in our data, we compared the actual scanned data of a smooth machined cylinder of a known radius with its ideal cylinder data generated with the same radius, position and orientation. The difference between the actual and ideal datasets gives an image containing the noise data, the histogram of which approximates the distribution of noise in the data. Figure 5.2 shows the cylinder used for the experiment. A smooth white cylinder with 63mm radius and 200mm height was chosen. The distance between the scanner and cylinder was set to 1.6m, the average distance used for scanning human subjects.

The cylinder was mounted on a platform with equal height as that of the scanner's tripod and both were levelled using a handheld spirit level. Twenty scans of the cylinder were then taken without disturbing the experimental setup.

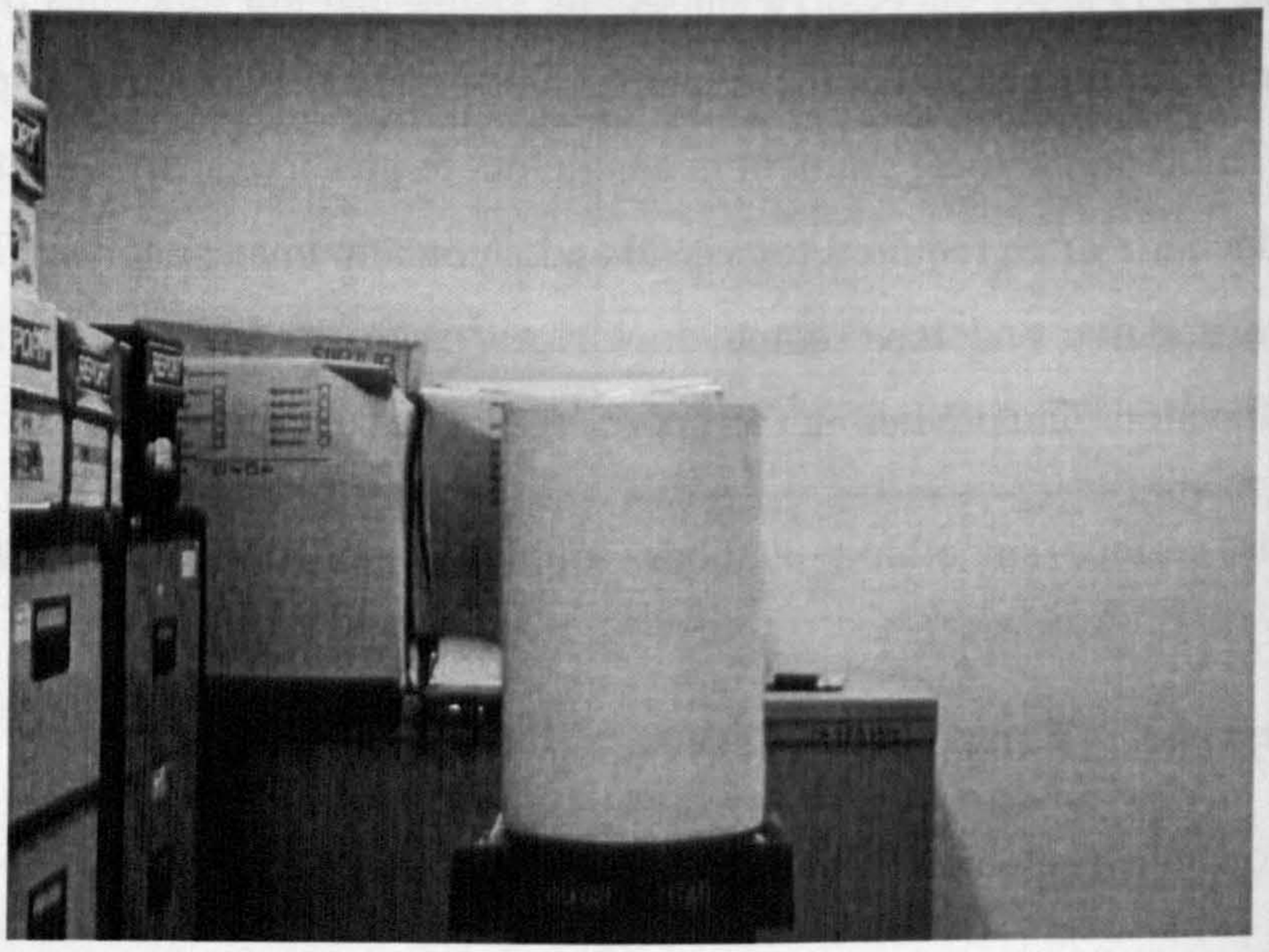


Figure 5.2: A cylinder used for the experiment

5.2.4.1. Analysis of Cylinder Data

The ideal data of the cylinder was generated using the same dimensions and orientation of the actual cylinder. The twenty datasets of the cylinder were interpolated onto a common grid with 0.5mm spacing and the mean dataset was computed by taking the arithmetic average on pixel by pixel basis. For comparison purposes, both the scanned and the ideal datasets were again interpolated onto a common grid with 0.5mm spacing. The error dataset was computed by subtracting the ideal dataset from the mean scanned dataset on pixel by pixel basis. To avoid the influence of boundary errors at the edges of the cylinder during data capture, a patch of 80 by 80 pixels of the noise data taken from the centre of the cylinder was analysed instead of the entire data (see Figure 5.3). The rationale behind this is that, near the edges of the cylinder, the geometric configurations of the points will be distorted due to the steep slope, and this could affect the accuracy of the derived coordinates. Thus, the errors at the edges of the cylinder would not be representative of the distribution of the random errors.

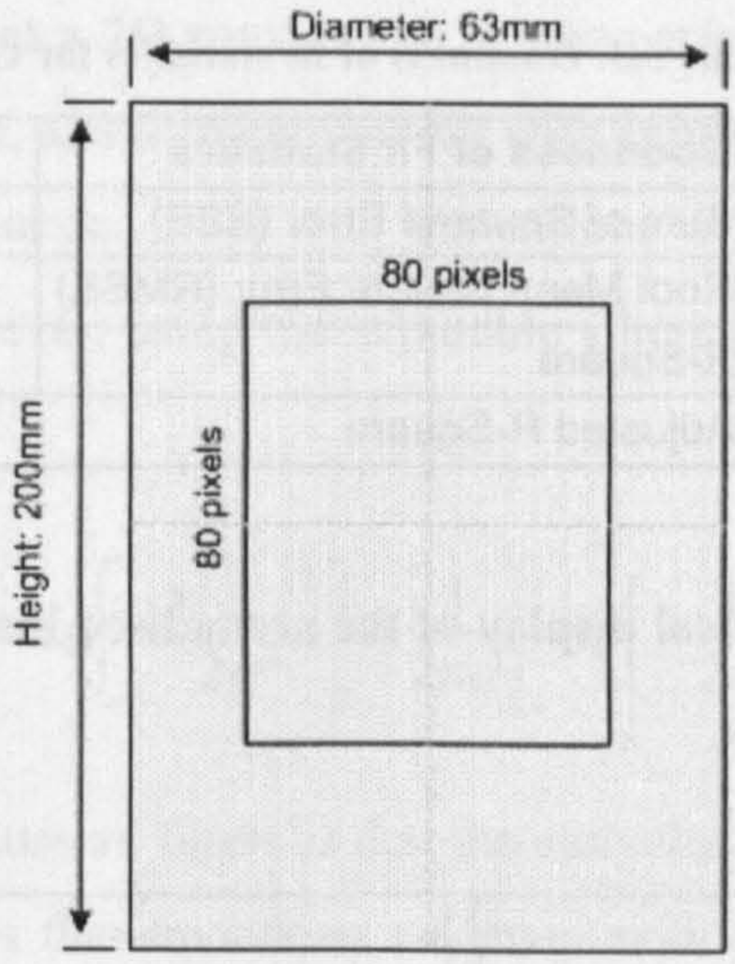


Figure 5.3: Patch of cylinder analysed: 80 by 80 pixels

The histogram of the noise data represents the distribution of the errors. The histogram was calculated by grouping the individual pixels of noise data into 500 bins, and counting the number of pixels falling within each bin. A Gaussian function was then fitted to the histogram using the method of Least Squares. The estimated parameters of the Gaussian function are shown in Table 5.2 together with their confidence intervals at 95% confidence level. The mean was estimated to be 0.0029±0.0046mm and the standard deviation was 0.4705±0.0079mm.

Table 5.2: Estimated parameters of Gaussian fitting

Coefficients (95% confidence interval)	Value	Lower Bound	Upper Bound
Mean	0.0029	0.0006	0.0052
Standard deviation (σ)	0.4705	0.4626	0.4784

The goodness-of-fit statistics presented in Table 5.3 show that the fitted Gaussian function describes the noise distribution quite well. The sum of squared errors (SSE) which represents the total deviation of the noise data and the fitted function is less than 3.0, while the root mean square error (RSME), which represents the standard error of the entire fit is less than 0.1. In conclusion, the results of this experiment show that the noise distribution is adequately approximated by a Gaussian distribution with standard deviation 0.47mm and mean 0.003mm, indicating that Gaussian smoothing will be appropriate for our data.

Table 5.3: Goodness of fit statistics for Gaussian fitting

Goodness of Fit Statistics	Value
Sum of Squared Error (SSE)	2.9817
Root Mean Square Error (RMSE)	0.0547
R-Square	0.9683
Adjusted R-Square	0.9682

Figure 5.4 shows a graphical display of the normalised histogram with the Gaussian fit superimposed.

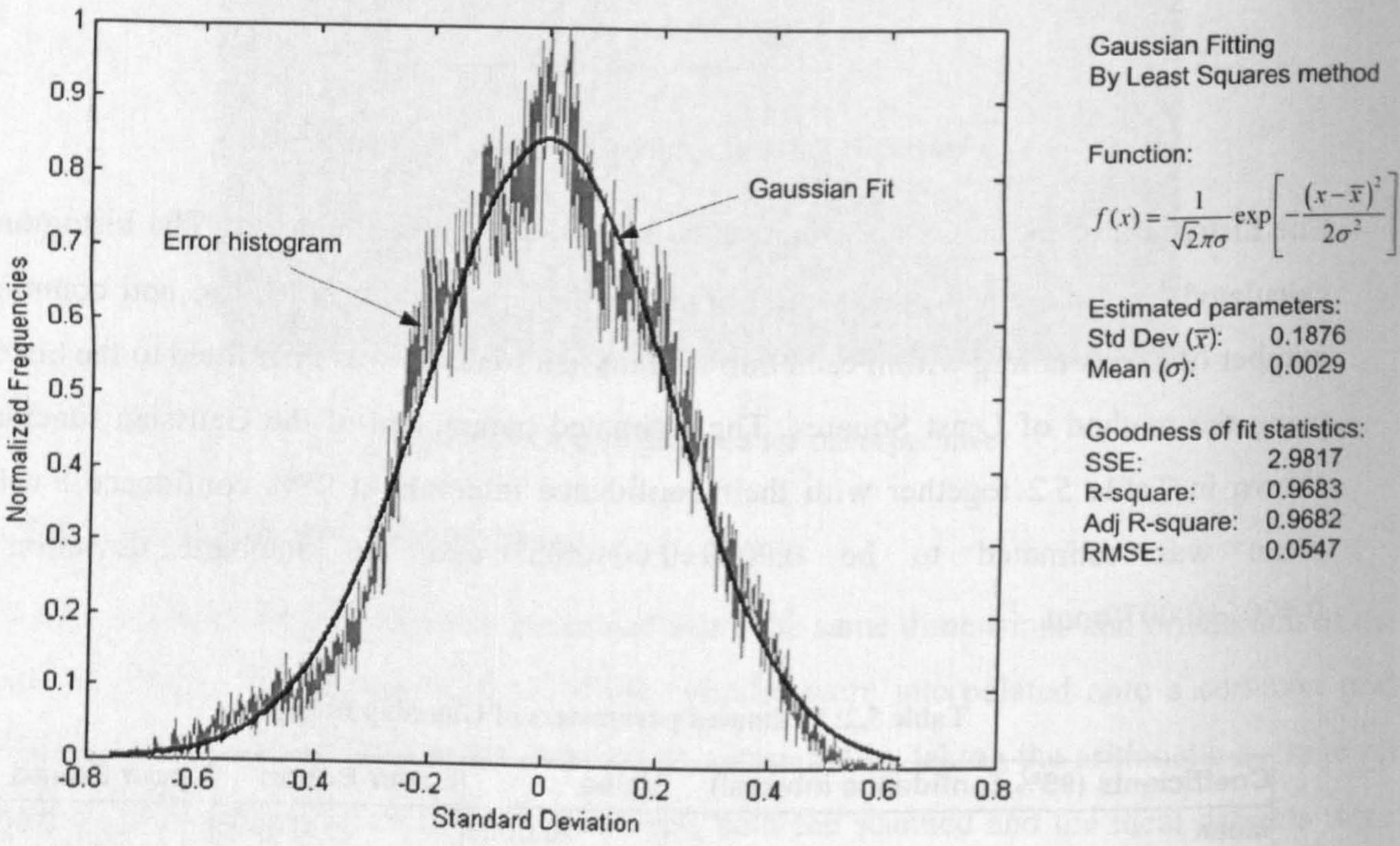


Figure 5.4: Noise distribution of cylinder data approximated by a Gaussian with standard deviation 0.19mm and mean 0.001mm

5.2.5. Gaussian Smoothing

The experimental results obtained in the previous section indicate that the errors in our data are approximately normally distributed and thus Gaussian smoothing will be optimal method for reducing the effect of the errors. In fact, Gaussian smoothing has a number of features that makes it a suitable choice for reducing normally distributed noise. The Gaussian function decreases monotonically from the central peak, thus giving higher weight to the central pixel. Moreover, the Gaussian is completely separable which makes its computationally efficient for

large images like our data. Thus a 2D convolution can be performed by first convolving with 1D Gaussian in the X direction, and then convolving with another 1D Gaussian in the Y direction (Equation 5.1). For a Gaussian filter with dimensions $m \times m$, 2D convolution requires m^2 operations per pixel. However, using the separable filters can reduce this to $2m$ operations per pixel.

$$G(x,y) = \frac{1}{2\pi\sigma}\exp\left[-\frac{x^2+y^2}{2\sigma^2}\right] = \frac{1}{2\pi\sigma}\left[\exp\left(-\frac{x^2}{2\sigma^2}\right) * \exp\left(-\frac{y^2}{2\sigma^2}\right)\right] \quad (5.1)$$

Another useful property of Gaussian filters is that the convolution of a Gaussian with itself is another Gaussian. This implies that smoothing an image with a small Gaussian filter several times is equivalent to smoothing the same image with a larger Gaussian. The Fourier transform of a Gaussian with standard deviation σ is another Gaussian with standard deviation $\alpha = 1/2\pi\sigma$ and amplitude $A = \sqrt{2\pi\sigma^2}$ (Gauge, 1998). Thus for large images, Gaussian filtering can be efficiently implemented using the Fast Fourier Transform algorithm.

5.2.6. Design of the Gaussian Filter

From Equation 5.1, the 2D continuous Gaussian function has one parameter that must be specified, namely, the standard deviation (σ). Due to the symmetric nature of the Gaussian, the standard deviation is the same in both X and Y directions, i.e. $\sigma = \sigma_X = \sigma_Y$. The standard deviation parameter controls how quickly the Gaussian approaches zero (see Figure 5.5). This implies that different amounts of smoothing can be obtained by varying the standard deviation—a small σ will produces a function which quickly approaches zero, and can be implemented using a small discrete filter; a large σ produces a function which slowly approaches zero. In other words, a higher standard deviation will result in greater amount of smoothing and vice versa. The issue of how much smoothing should be applied is addressed empirically in Section 5.2.7.

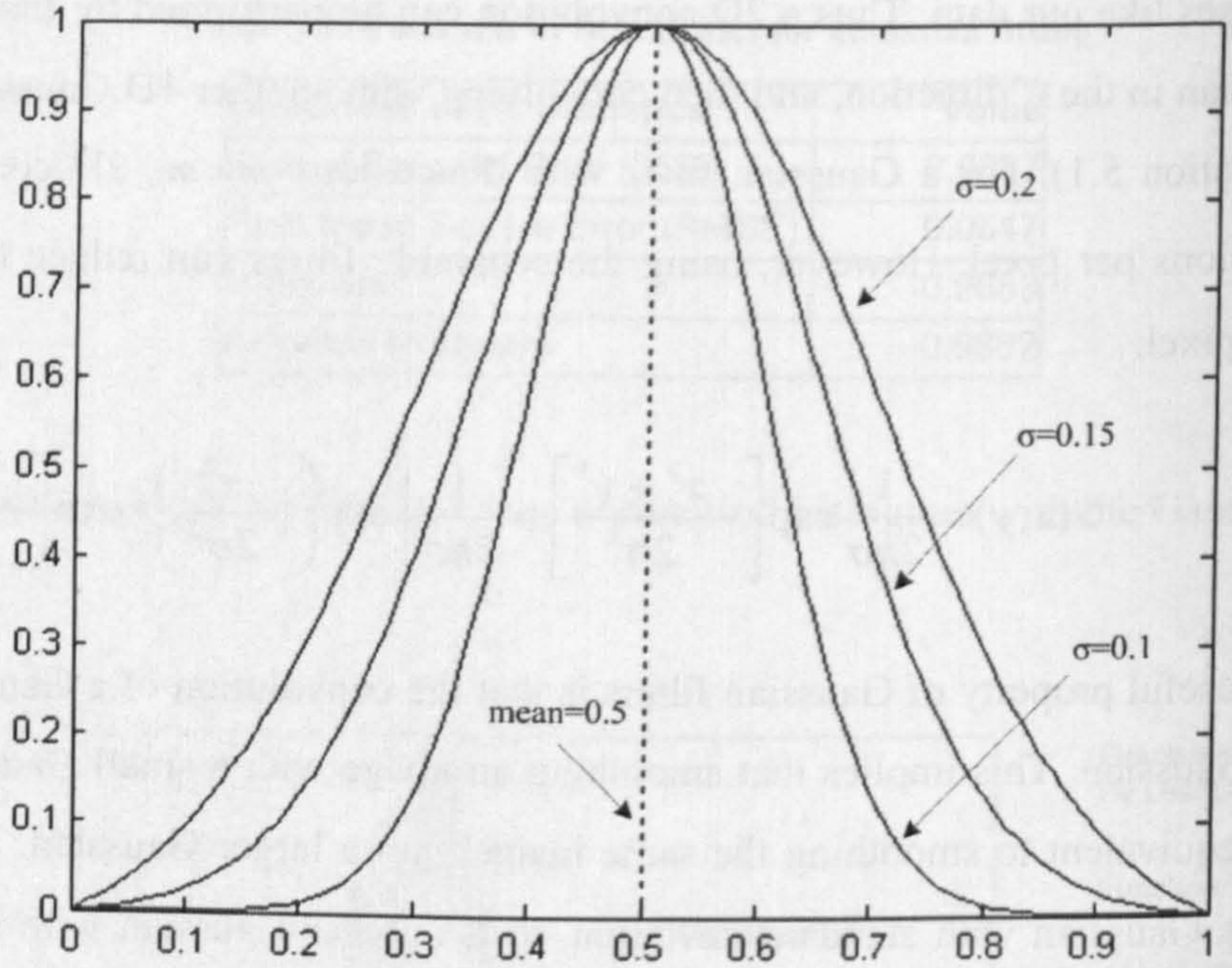


Figure 5.5: 1D Gaussian function with varying standard deviations

Once the appropriate standard deviation is chosen for the continuous Gaussian function, a discrete version must be derived in order to apply it to discrete images. We sample the continuous function by integrating it over the square pixels using the 2mm pixel size of the target data. Figure 5.6 shows an example of a 2D continuous Gaussian with 0.5 standard deviation sampled by a 5x5 discrete filter in Equation 5.2. This discrete filter can now be applied to the images using standard convolution methods (see Pratt, 2001).

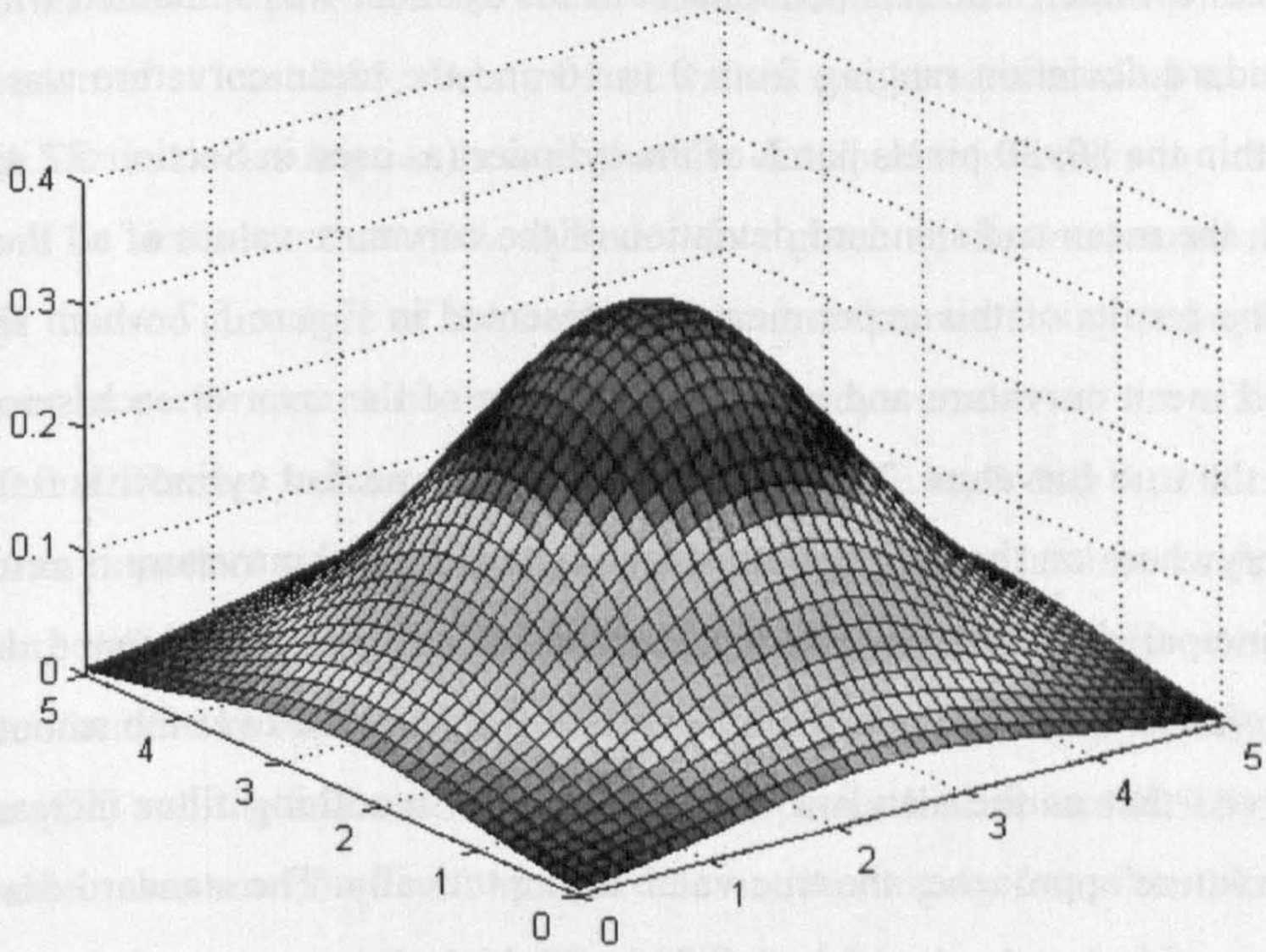


Figure 5.6: A continuous Gaussian with $\sigma = 0.5$, sampled by 5×5 discrete filter

$$G_{0.5} = \begin{bmatrix} 0 & 0 & 0.81 & 0 & 0 \\ 0 & 0.0113 & 0.0837 & 0.0113 & 0 \\ 0.0002 & 0.0837 & 0.6197 & 0.0837 & 0.0002 \\ 0 & 0.0113 & 0.0837 & 0.0113 & 0 \\ 0 & 0 & 0.81 & 0 & 0 \end{bmatrix} \quad (5.2)$$

5.2.7. Estimating Optimal Smoothing

This section addresses the second issue of the limit of smoothing. This is an important issue because smoothing has a basic drawback—true variations in the image, especially sharp variations (edges and creases) are also smoothed to some extent. Consequently, minimal smoothing should applied to avoid smoothing out true surface variations. Fortunately, the human back surface does not exhibit sharp variations in shape so the effect of carefully controlled smoothing should be minimal.

Since the amount of smoothing achievable with a Gaussian filter is directly proportional to the standard deviation parameter, the problem resolves into determining the standard deviation value for optimal smoothing. This issue is addressed by comparing the curvature computed from the scanned data of the cylinder with increasing smoothing with the true curvature

of the ideal cylinder. The scanned dataset of the cylinder was smoothed with a Gaussian filter with standard deviation ranging from 0 to 10 and the mean curvature was computed at each point within the 80x80 pixels patch of the cylinder (as used in Section 5.2.4). At each smoothing level, the mean and standard deviation of the curvature values of all the pixels were computed. The results of this experiment are presented in Figure 5.7 which shows a plot of the computed mean curvature and standard deviation of the error at each smoothing level compared to the true curvature. The true curvature of the perfect cylinder is 0.016mm and is constant everywhere on the cylinder—the minimum principal curvature is zero, while the maximum principal curvature is constant. Because of errors in the scanned data, the computed mean curvature of the cylinder is not constant, but variable over the smoothing range. It can be observed that as the standard deviation of the smoothing filter increases, the computed mean curvature approaches the true value asymptotically. The standard deviation of the error before smoothing commenced is 0.6894mm, which decreases towards zero with increasing smoothing. At a smoothing level of about $\sigma = 2.85$, the mean curvature has approached the true value to within 10% of the error, which gives a general indication of error limits for optimal smoothing.

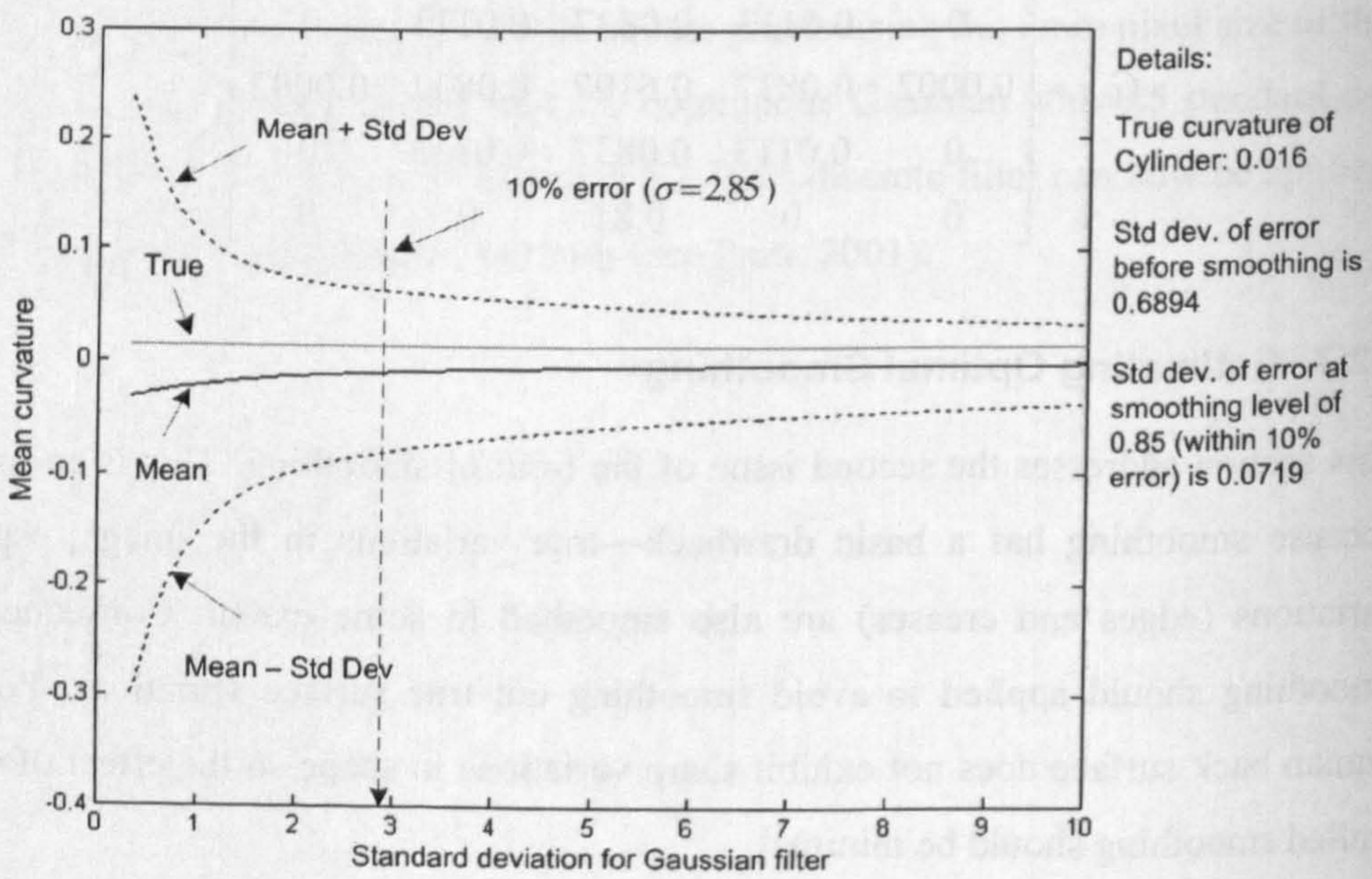


Figure 5.7: Comparison of computed curvature with true curvature of cylinder with increasing smoothing

The results of the above experiment show that if we smooth the back data with a Gaussian of standard deviation of at least 2.85, the curvature values can be estimated up to within 10% of the total error. The basic assumption here is that the distribution of errors in the scanned data of the back will be similar to that of the cylinder data given similar measurement conditions. Unfortunately, the true curvature of the back surface is unknown at any point, so similar experiment cannot be performed on back shape data. Instead, we assessed the effect of smoothing on curvature difference between certain landmarks of the back and their local neighbourhoods. This is an important step since successful detection of these landmarks depends on sufficient curvature differences between the landmarks and their local neighbourhood (see Section 6.2). Three landmarks of the back were chosen for this experiment, namely, seventh cervical vertebra (C7) and the left and right posterior superior iliac spines (PSIS). The scanned data of one subject was smoothed with standard deviation ranging from 0 to 10, and the mean and Gaussian curvatures were computed. The locations of the landmarks were extracted from the curvature maps. For each landmark, the absolute difference in curvature between the landmark and its 3x3 neighbourhood was computed at each smoothing level. The results of this experiment are presented Figure 5.8 and Figure 5.9. It can be observed that as the smoothing standard deviation increases, both the mean and Gaussian curvature differences decrease toward zero for all the landmarks. Beyond σ=5.5, the absolute curvature difference for C7 landmark effectively tends to zero ($\approx 1x10^{-5}$ for the mean curvature and $\approx 1x10^{-6}$ for the Gaussian curvature), thus making it difficult to accurately detect the landmark. It also appears from both figures that the curvature differences begin to stabilise after a smoothing level of about $\sigma = 2.0$. After further experiments with other curvature maps and landmarks, σ=1.5-3.5 was selected as optimal smoothing range for our data. All datasets analysed and visualized in this thesis are smoothed with σ=2.25, which gives good results both numerically and visually.

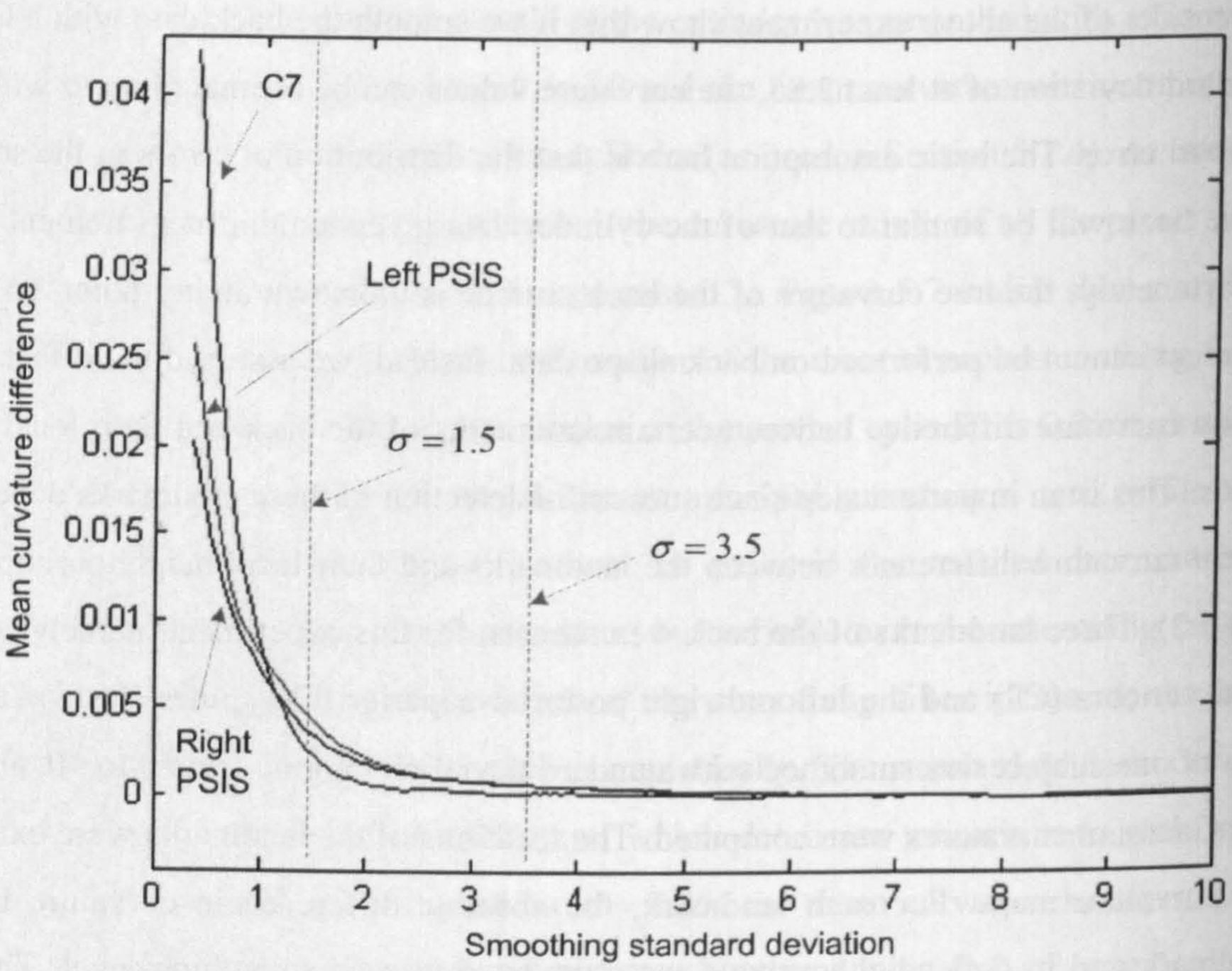


Figure 5.8: Effect of smoothing on Mean curvature difference for C7, LPSIS, and RPSIS

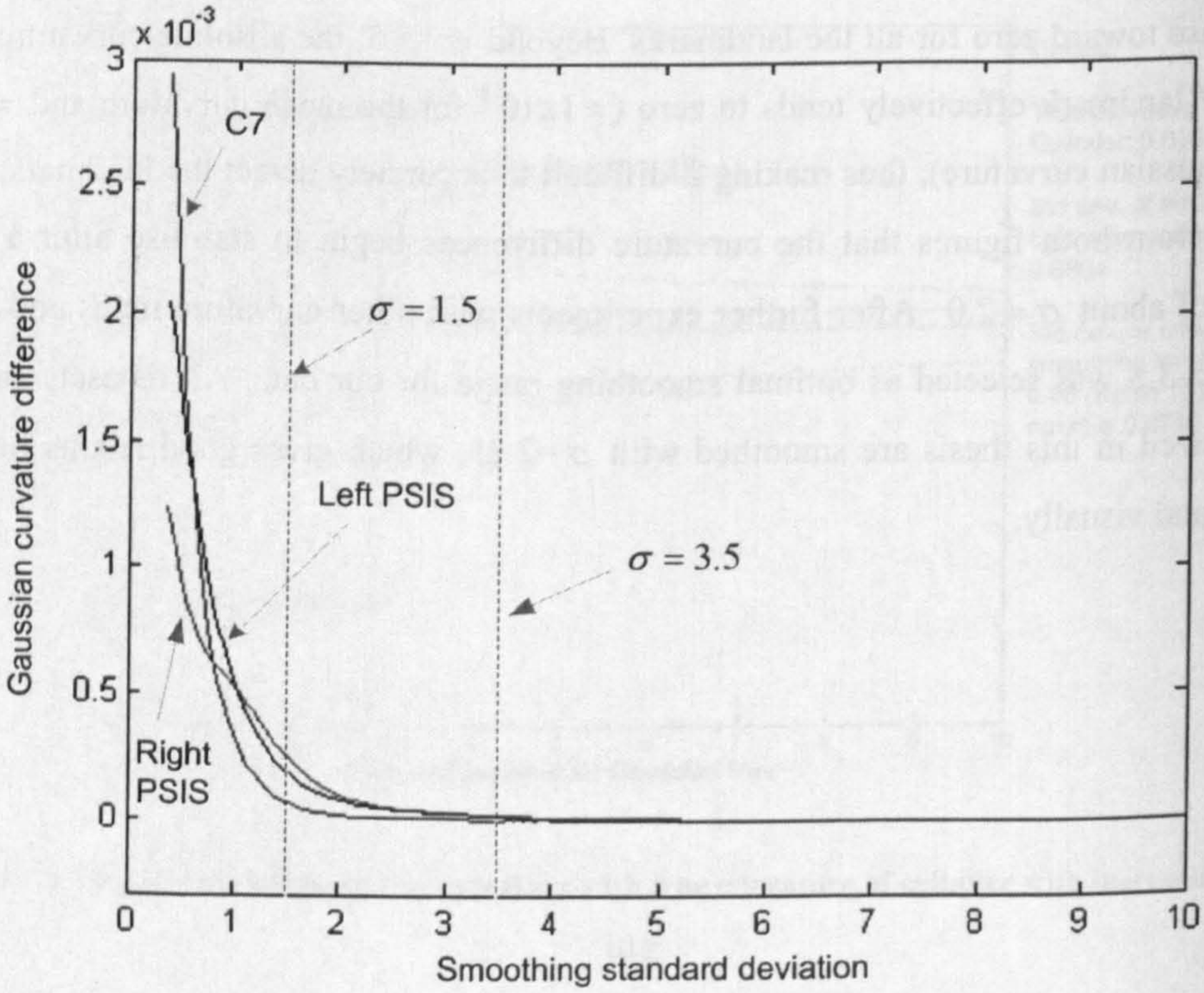


Figure 5.9: Effect of smoothing on Gaussian curvature difference for C7, LPSIS, and RPSIS

5.3. Visualization of the Back Shape

In this section, we explore various methods for visualizing the shape of the back. Visualization plays an important role in back shape analysis as it enables the exploration of the extent and changes of back deformities. The effect of treatment methods such as bracing or surgery can be explored using effective visualization tools without the need for physical manipulation of the patient with the problem. The main objective here is to provide diagnostic visual tools to aid the assessment of back deformities. To this end we explore a number of visualization methods, namely, continuous colour scales, discrete colour scales, contours and three-dimensional views.

5.3.1. Visualization Pipeline

The data visualization process is usually represented by a pipeline which consists of two main steps, namely, encoding and decoding, as depicted in Figure 5.10 (Wünsche, 2004). The data encoding step describes the process of converting the data into some visual attributes to be displayed on the output medium (computer screen or paper), and often involves operations such as transformations, mapping and rendering. The decoding part describes the perception and cognition of the visual information by the human subject.

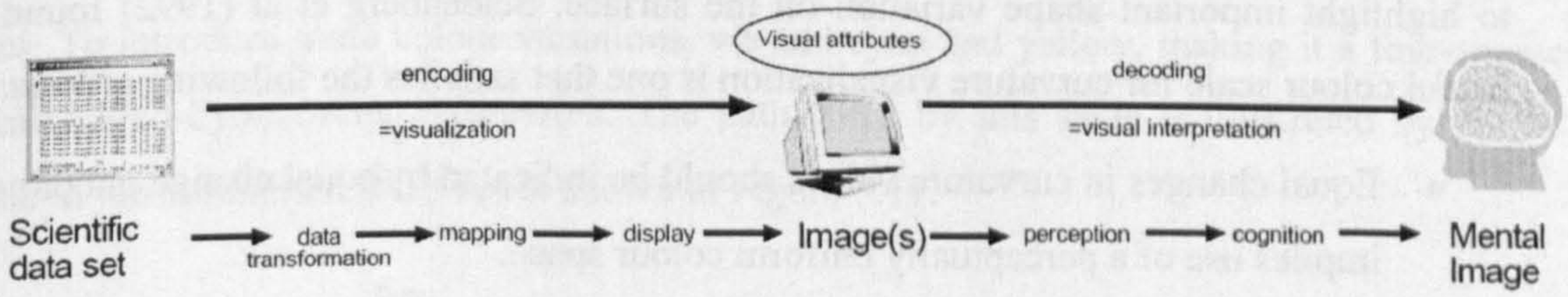


Figure 5.10: The visualization pipeline (adapted from Wünsche, 2004)

It is important to note that visualization is not an end in itself, but a step in the decision-making process. Consequently, visualization is only effective if the decoding process (perception and cognition) can be performed correctly and efficiently (Wünsche, 2004). Accurate perception implies that the data quantities and relationships conveyed by the visualization reflect the original data. To achieve this requirement, the visualization should preserve the order of the data values being visualized, preserve distances between values, group related values and separate unrelated ones. Efficiency of information perception implies that maximum

amount of information is perceived in minimal time, i.e. the visualization should maximise information absorption whiles minimizing the cost of interaction. To achieve this requirement, the visualization should be intuitive, easy to read and understandable. Perceptual and psychological aspects of colour should also be taken into account.

5.3.2. Colour Scales

Colour scales are effective and commonly used method for visualizing scalar data. Colour scales are well suited to give an overall impression of the distribution of the scalar field, although exact quantitative information cannot be perceived accurately. Moreover, the effectiveness of colour mapping depends on the colour scale used and on perceptual issues. Levkowitz (1997) summarises the following desirable properties of a colour scale:

- Colours should be perceived as preserving the order of the scalar values they represent.
- Colours should convey the distances between values they represent and should associate related values and separate unrelated values.
- Colours should be continuous for a continuous value range.

Colour scales are particularly effective for visualizing surface curvature where the objective is to highlight important shape variation on the surface. Seidenberg et al (1992) found that a useful colour scale for curvature visualization is one that satisfies the following criteria:

- Equal changes in curvature values should be indicated by equal change in colour. This implies use of a perceptually uniform colour space.
- There should be enough colours in the scale to detect important curvature variations.
- The middle and the two end colours must be unique since they represent flat, concave or convex areas which should be easily distinguishable.
- Colour scales should be used consistently. This will allow the frequent user to learn from experience and be able to associate certain colours with certain curvature values.

The first criterion of Seidenberg et al (1992) implies that a set of colours from a perceptually a uniform colour space is required. A colour space is perceptually uniform if a small change at one point in the space is approximately equally perceptible across the entire space. In other words, any two points with a certain distance apart in one part of the space should be percep-

tually the same as two other points with the same distance apart in another part of the space. The RGB (Red-Green-Blue) colour space, which is used by monitors and many display devices, is perceptually non-uniform. Interpolating a linear colour scale from a non-uniform colour space (such as the RGB space) may result in two adjacent colours in one part of the scale being perceptually different from two adjacent colours in other part of the scale. The problem of finding a uniform colour space, however, is a difficult one, and has been the subject of much research since the origin of colour science. A major advance was made by the CIE (Commission Internationale de l'Eclairage) in 1976 with the development of the CIELAB and CIELUV colour spaces, both of which were defined based on the non-uniform XYZ colour space defined by the CIE in 1931 (see Appendix D for description of these colour spaces). These two colour spaces are nearly uniform and have been accepted as the standard model for colour conversion (creating a perfectly uniform colour space is still problematic). For our colour scale, we perform the interpolation in the CIELAB space to ensure perceptual uniformity. Thus, the original RGB colours must be transformed into CIELAB for interpolation, and then transformed back into RGB for display.

5.3.3. Colour Scale Creation

To demonstrate the process of creating the colour scales, we use Blue-Green-Red scale in which blue represents low values, green represents middle values and red represents high values. To introduce more colour variations, we add cyan and yellow, making it a four-segment scale: *Blue-Cyan-Green-Yellow-Red.* The path taken by this scale is illustrated by the bold line in the normalised RGB space shown in Figure 5.11.

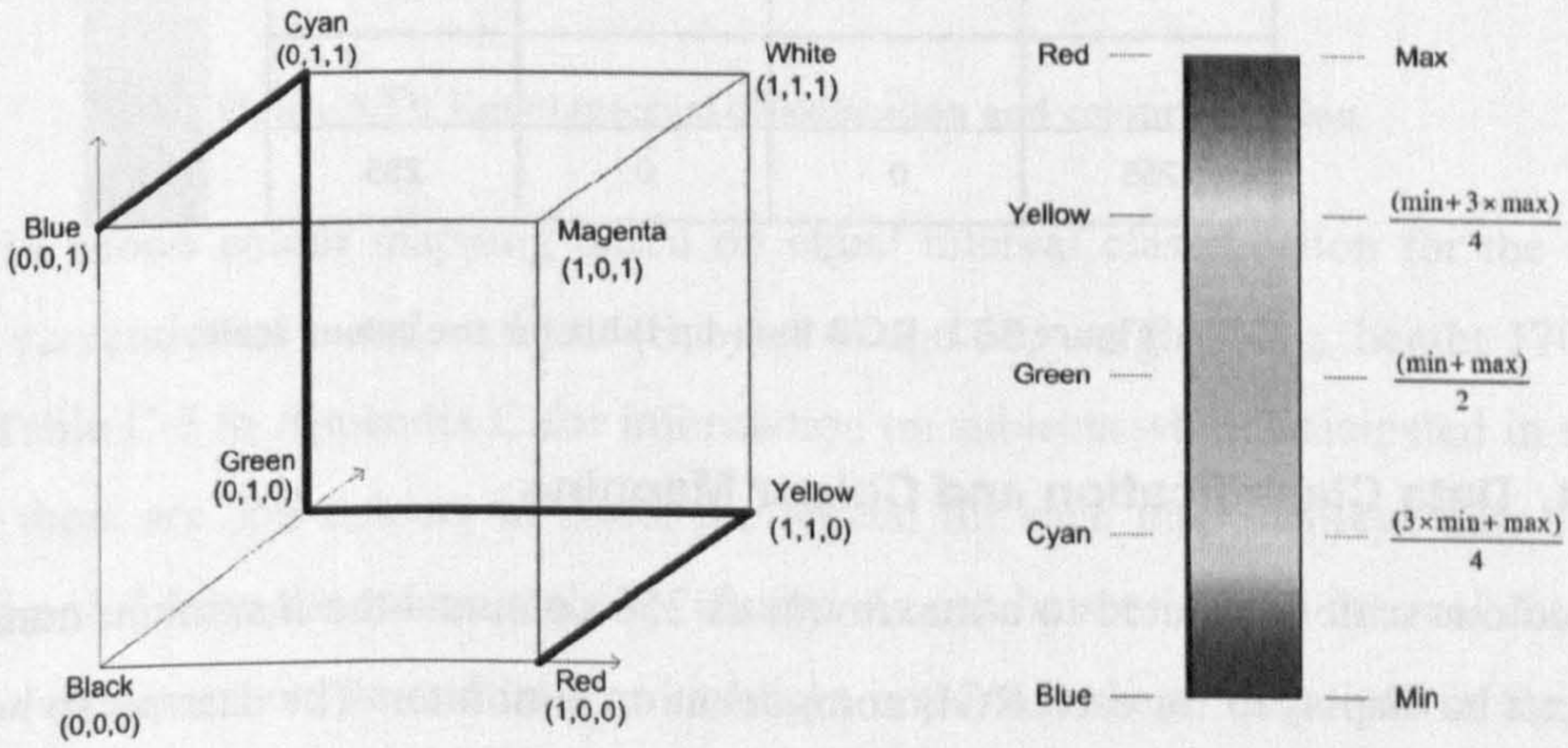


Figure 5.11: Blue-Cyan-Green-Yellow-Red colour scale

The process of creating the colour scale can be summarised as follows:

- Transform the RGB values for Red, Yellow, Green, Cyan and Blue from RGB space through XYZ space to CIELAB. The transformation is accomplished by using the set of equations given in Appendix D.
- Interpolate the required number of colours linearly in the CIELAB space. A total of 256 colours are interpolated for the entire scale, with equal number of colours in each segment (i.e. 64 for a four-segment scale).
- Transform the interpolated colours from CIELAB space through XYZ space back to RGB space. This transformation is accomplished by the set of equations given in Appendix D.
- A colour look-up table (LUT) is then created with all possible colours in the scale and an index associated with each RGB triplets (Figure 5.12). For each pixel in the image, the rendering algorithm calculates the colour index associated with the pixel's value, looks it up from the LUT and renders the pixel using this colour.

Index	Red	Green	Blue
0	245	0	0
1	245	0	0
2	245	0	0
3	247	0	0
4	247	8	0
5	247	21	0
...	...	...	...
255	0	0	255

Figure 5.12: RGB look-up table for the colour scale

5.3.4. Data Classification and Colour Mapping

The colour scale is limited to a maximum of 256 colours—the maximum number of colours that can be displayed for each RGB component on a monitor. The datasets to be visualized on the other hand typically contain over 36,000 pixels. To map the colour scale to the dataset, the data have to be classified or partitioned into the required number of available colours. In this

section we explore various classification methods and contrast enhancement techniques in order to maximise the visual contrast of the resulting images.

5.3.4.1. Equal Interval Classification

A simple method of mapping the colour scale to the curvature data is to partition the data using a fixed interval—each class occupies an equal interval along the number line (Figure 5.13). Thus, this method is identical to creating a grouped-frequency table or histogram of the data. Let $Z_{\min}$ and $Z_{\max}$ be the minimum and maximum data values, and n be the number of classes, then the class interval (I_{equal}) is calculated by Equation 5.3:

$$I_{equal} = \frac{Z_{\max} - Z_{\min}}{n} \tag{5.3}$$

To associate the colour map, the first and last colours in the scale are assigned to the first and last classes respectively, while all other classes are distributed linearly to the rest of the colours (Figure 5.13).

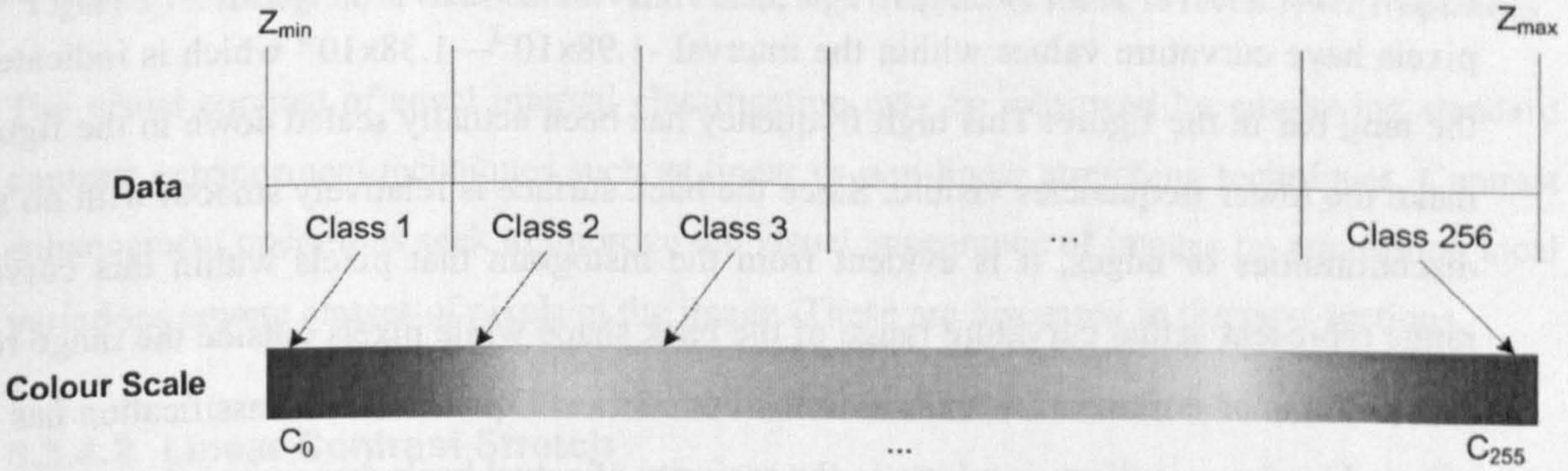
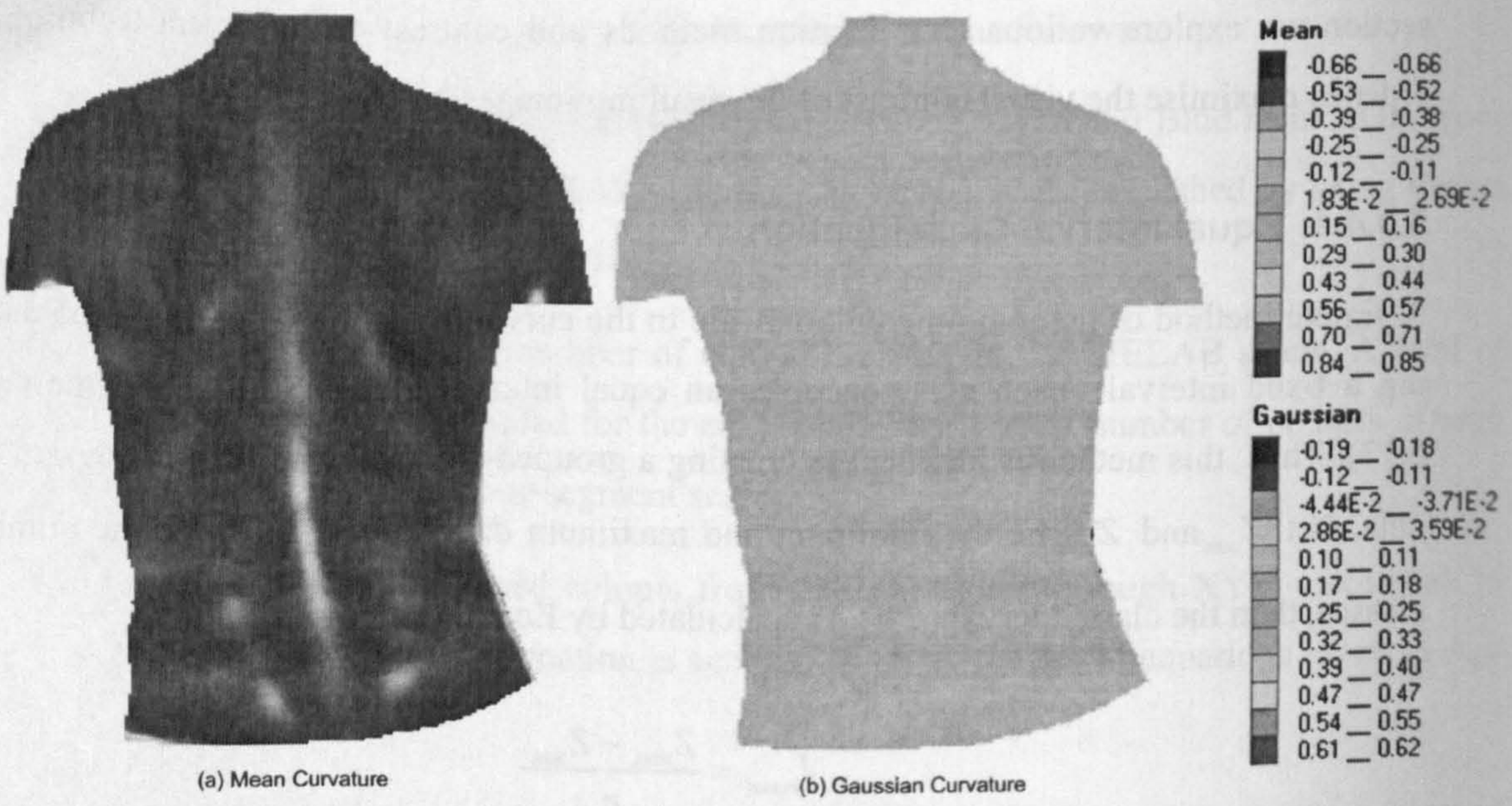


Figure 5.13: Equal interval classification and colour mapping

Figure 5.14 shows colour mapping based on equal interval classification for the mean and Gaussian curvatures of a male subject (Subject 1: age 30, weight 70kg, height 170cm, BMI 24). See Table C-7 in Appendix C for information on subjects who participated in this study. Although there are 256 colours in scale, the legend for each map displays only 12 colours linearly sampled from the colour scale for clarity. As can be seen, equal interval classification results in very poor contrast and fails to highlight important shape variations in the curvature maps. The Gaussian curvature map particularly has a very poor contrast.



(a) Mean Curvature

(b) Gaussian Curvature

Figure 5.14: Mapping of colour scale to curvature maps using equal interval classification (Subject 1)

The histogram of the Gaussian curvature map shown in Figure 5.15 explains why equal-interval classification results in a poor contrast image. As indicated by the histogram, the majority of the curvature values cluster around 0. To be precise, 35,835 out of the total of 36,868 pixels have curvature values within the interval $-1.98\text{x}10^{-3}$—$1.38\text{x}10^{-3}$ which is indicated by the long bar in the figure. This high frequency has been actually scaled down in the figure to make the lower frequencies visible. Since the back surface is relatively smooth with no sharp discontinuities or edges, it is evident from the histogram that pixels within this curvature range represent actual curvature range of the back shape while pixels outside the range represent extreme or erroneous curvatures at the boundaries. Equal interval classification has been affected by these outliers, resulting in the majority of actual back shape pixels being placed in only a few classes. Thus equal interval classification fails to respect the statistical distribution of the dataset.

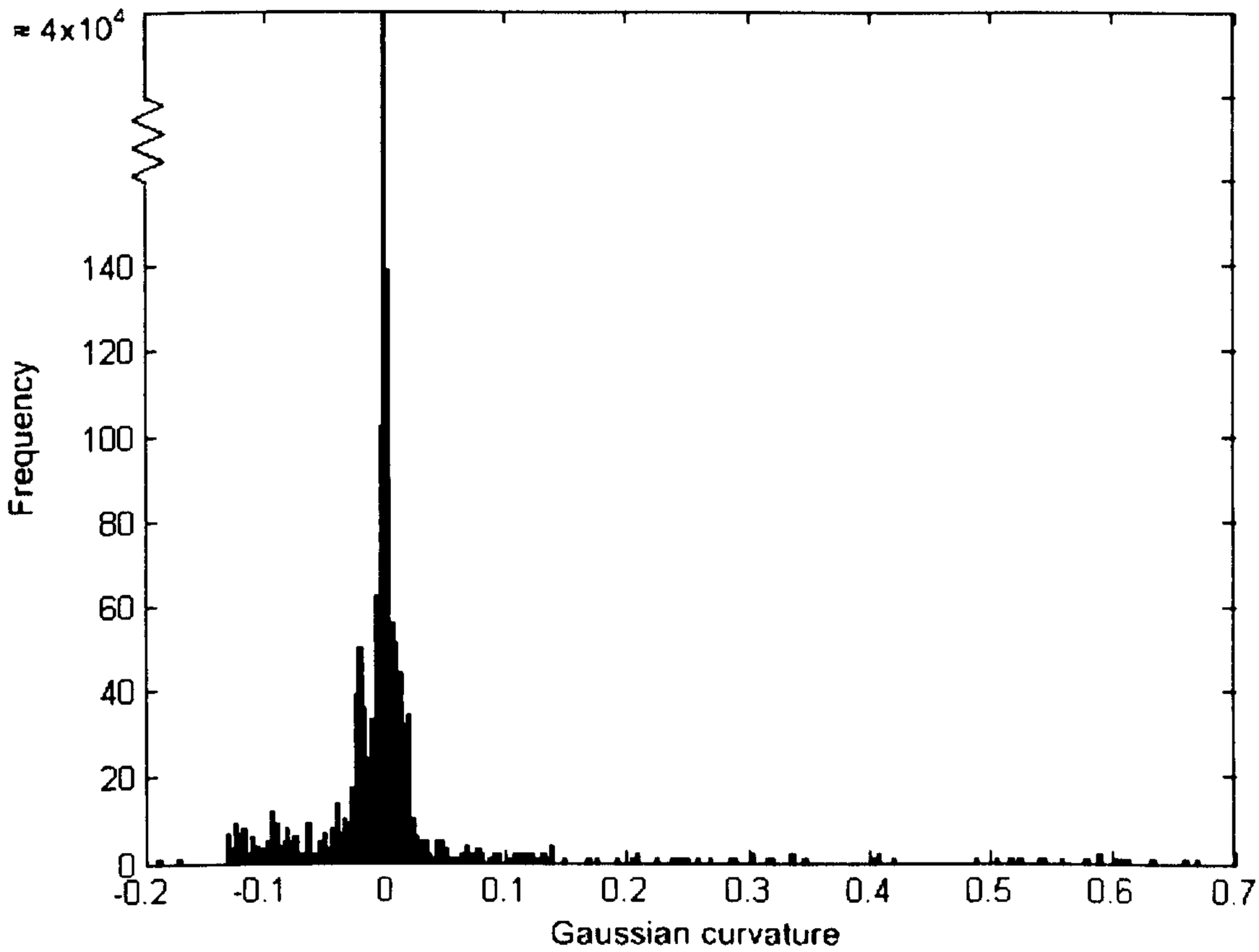


Figure 5.15: Histogram of Gaussian curvature map, high frequencies scaled to reveal lower frequencies

The visual contrast of equal interval classification may be improved by employing standard contrast enhancement techniques such as linear or non-linear stretching techniques. Contrast enhancement operations seek to improve the visual appearance of images by amplifying local variations among classes of pixels in the image. These are discussed in the next sections.

5.3.4.2. Linear Contrast Stretch

If the region of interest lies within a known range $Z_{low}...Z_{high}$, there is no need to display pixels lying outside this range. Pixels with values lower than the Z_{low} can be assigned to the beginning of the colour scale (C_0) and pixels with values above Z_{high} can be assigned to the end of the colour scale (C_{255}). All pixels with values within that range are then linearly distributed. This has the effect of amplifying the variations between adjacent pixels thereby enhancing contrast within the target region and effectively reducing contrast outside the region to zero. By adjusting Z_{low} and Z_{high} in Equation 5.4, low contrast details in different curvature regions of the image can be explored interactively.

$$C_i = C_0 + (i - Z_{\text{low}})\left(\frac{C_{255} - C_0}{Z_{\text{high}} - Z_{\text{low}}}\right) \tag{5.4}$$

Another method for defining the particular region to enhance is to cut off extreme pixels by considering the statistical distribution of the entire image. If we define extreme values as those pixels outside certain standard deviations from the mean or median (say n standard deviations), then all pixels outside the interval $\bar{Z} \pm n\sigma$ are mapped to the beginning and end of the colour scale while pixels within that range are stretched linearly between the end colours (see Figure 5.16). This allows the colour scale to avoid the extreme values. The standard deviation multiplier (n) can be adjusted to produce various degree of contrast enhancement.

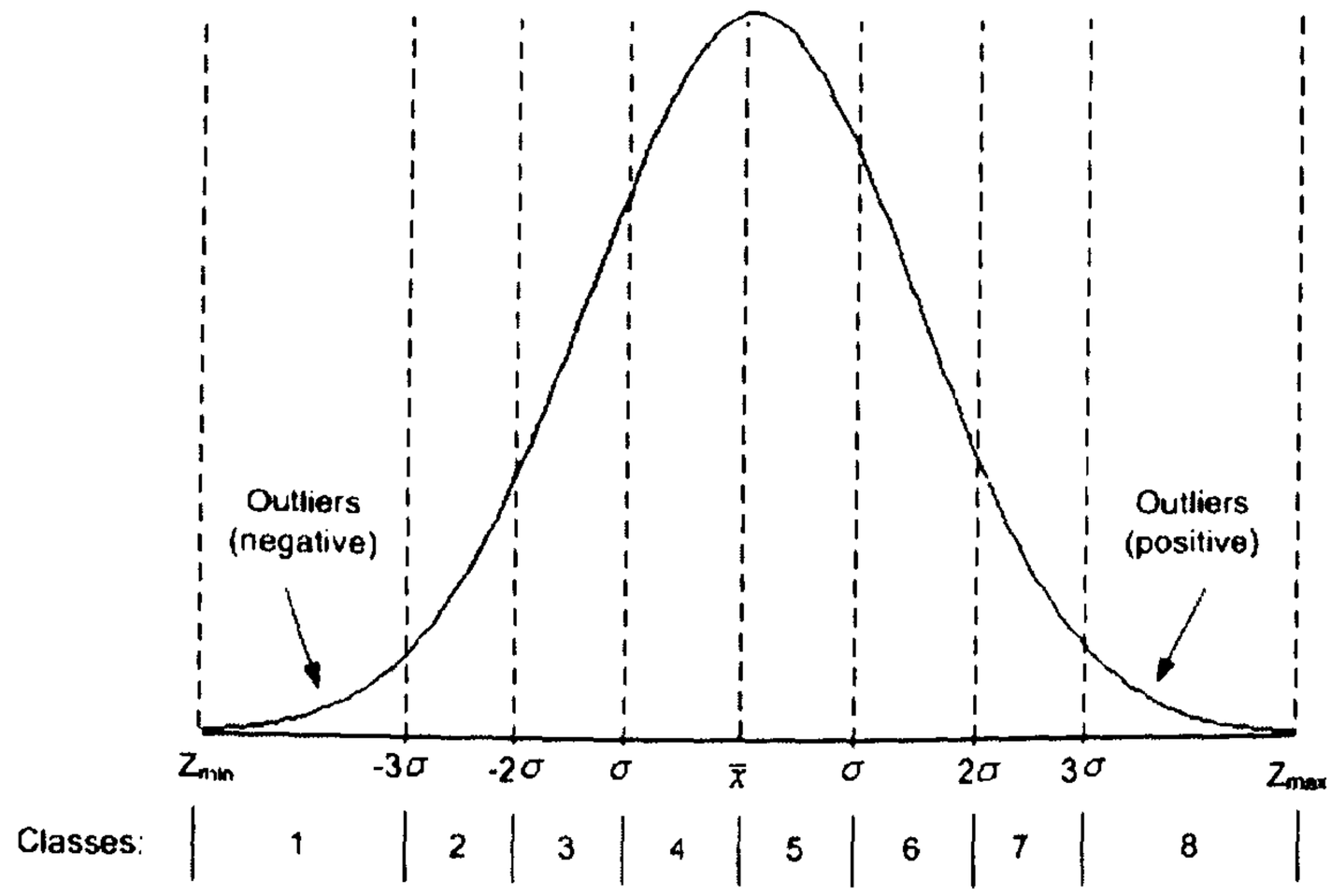
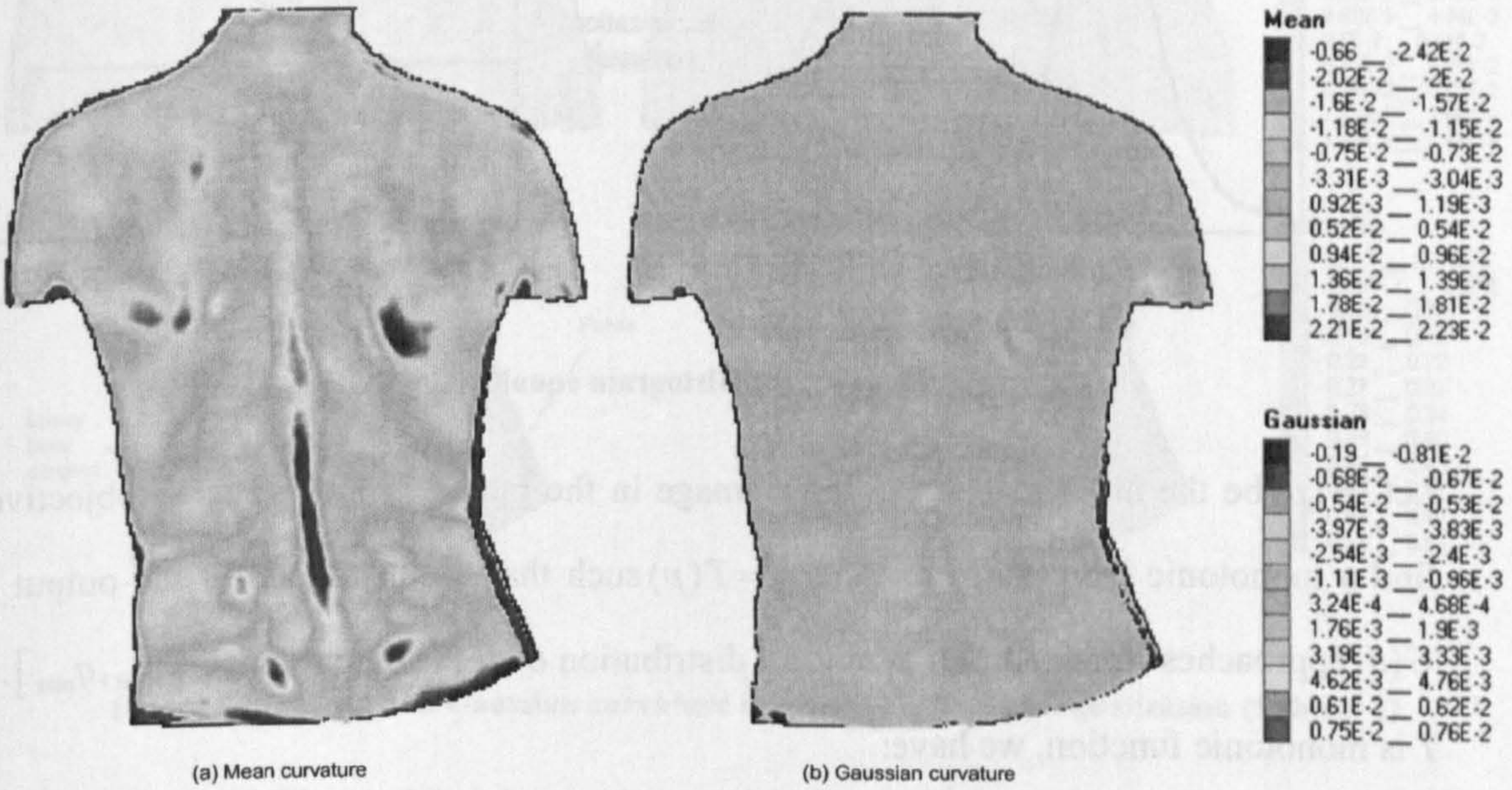


Figure 5.16: Standard deviation stretch, 8 classes

Figure 5.17 shows application of standard deviation stretch to the mean and Gaussian curvature maps after equal interval classification. A standard deviation threshold of $n = 0.5$ was used for the mean curvature while a threshold of $n = 0.1$ was used for the Gaussian. The mean curvature now shows a considerable visual contrast improvement whiles the Gaussian shows only a slight improvement. A higher threshold ($n > 0.5$) showed no contrast for the Gaussian map—highlighting the effect of outliers on equal interval classification. These results demonstrate that equal interval classification and standard deviation stretch are only well-suited to normally distributed data where the mean can serve as a useful dividing point, enabling contrast of values above and below it. However, since the mean is easily affected by

extreme values (outliers), this method will perform poorly with data that contain outliers or data with skewed distributions.



(a) Mean curvature

(b) Gaussian curvature

Figure 5.17: Standard deviation stretch applied to the mean and Gaussian curvature maps with equal interval classification (Subject 1)

5.3.4.3. Histogram Equalisation

Linear stretch techniques produce uniform contrast enhancement throughout the image, and thus can still be affected by outliers, as demonstrated in previous section. Non-linear, non-uniform enhancement such as histogram equalization is sometimes required to reveal subtle variations in the image. Histogram equalization allows the pixels values to be redistributed so that histogram of the enhanced image becomes uniform or follows some other distribution (Figure 5.18) (Pratt, 2001).

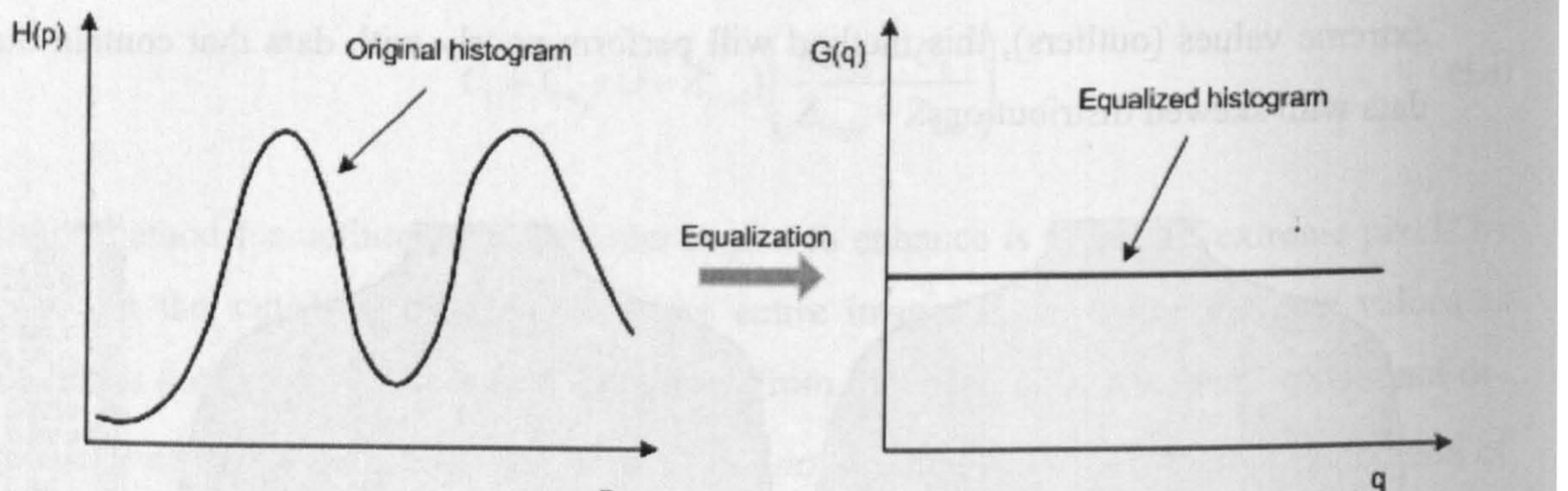


Figure 5.18: Histogram equalization

Let $H(p)$ be the histogram of the input image in the range $[p_{\min}, p_{\max}]$. The objective is to find a monotonic increasing function $q = T(p)$ such that the histogram of the output image $G(q)$ approaches (for example) a uniform distribution over the output range $[q_{\min}, q_{\max}]$. Since T is monotonic function, we have:

$$\sum_{i=0}^{j} G(q_i) = \sum_{i=0}^{j} H(p_i), \text{ where } q_i = T(p_i) \tag{5.5}$$

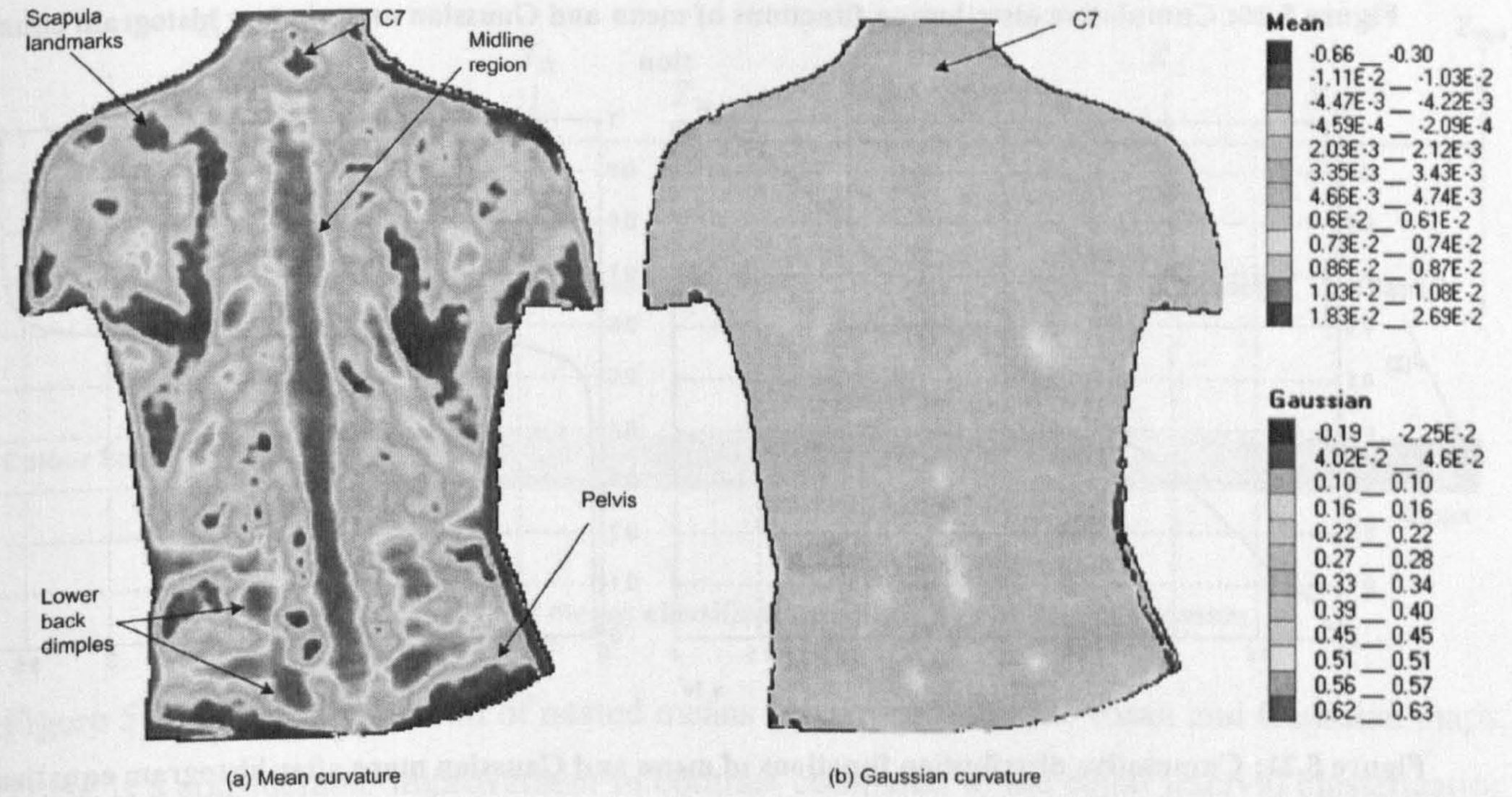
The summation on the right represents the cumulative histogram of the discrete input image. For a continuous distribution, the desired transformation is derived as:

$$q_i = T(p) = \frac{q_{\max} - q_{\min}}{nm} \int_{p_{\min}}^{p} H(s)ds + q_{\min} \tag{5.6}$$

The discrete approximation of Equation 5.6 is given by:

$$q_i = T(p_i) = \left(\frac{q_{\max} - q_{\min}}{n} \right) \sum_{i=0}^{j} H(p_i) + q_{\min} \tag{5.7}$$

Figure 5.19 shows application of histogram equalization to the mean and Gaussian curvature maps based on equal interval classification. The mean curvature now reveals more variations in the shape of the back: various landmarks are now visible—C7, midline region, lower back dimples, scapular areas, etc. The Gaussian curvature still suffers from poor contrast—global histogram equalization has failed to effectively improve contrast in the presence of extreme outliers.

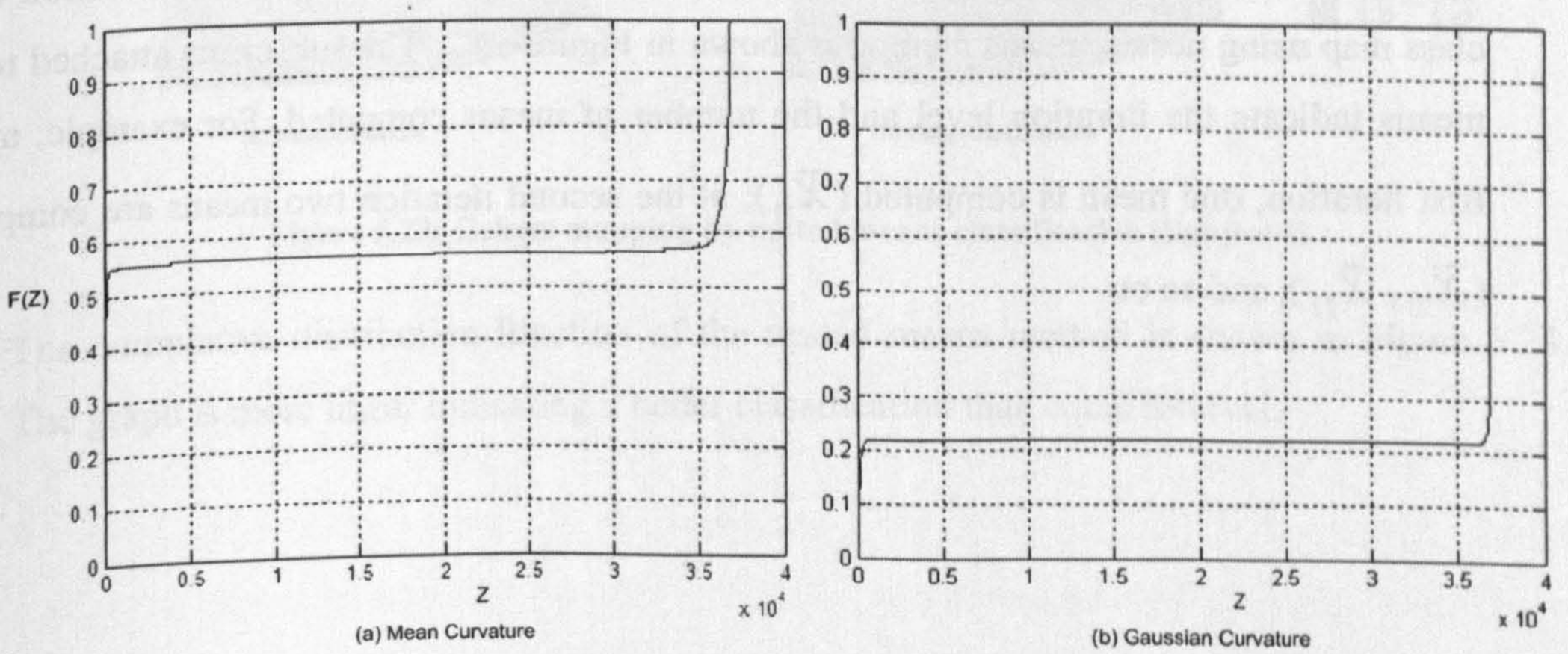


(a) Mean curvature

(b) Gaussian curvature

Figure 5.19: Mean and Gaussian curvature maps with histogram equalization (Subject 1)

Figure 5.20 and Figure 5.21 shows the empirical cumulative distribution functions of the mean and Gaussian curvature maps before and after histogram equalization. Global histogram equalization has been able to linearise the cumulative histogram for the mean curvature resulting in higher contrast. However, it failed to effectively linearise the cumulative distribution function of the Gaussian map, evidently due to presence of extreme values. Perhaps the Gaussian map could still be improved by localising the equalization process, an approach known as *adaptive histogram equalization* (Pratt, 2001). However, adaptive histogram equalization comes with high computational cost which makes it problematic for creating on-the-fly visualizations for large datasets; thus we investigate other methods for data classification.



(a) Mean Curvature

(b) Gaussian Curvature

Figure 5.20: Cumulative distribution functions of mean and Gaussian maps before histogram equalization

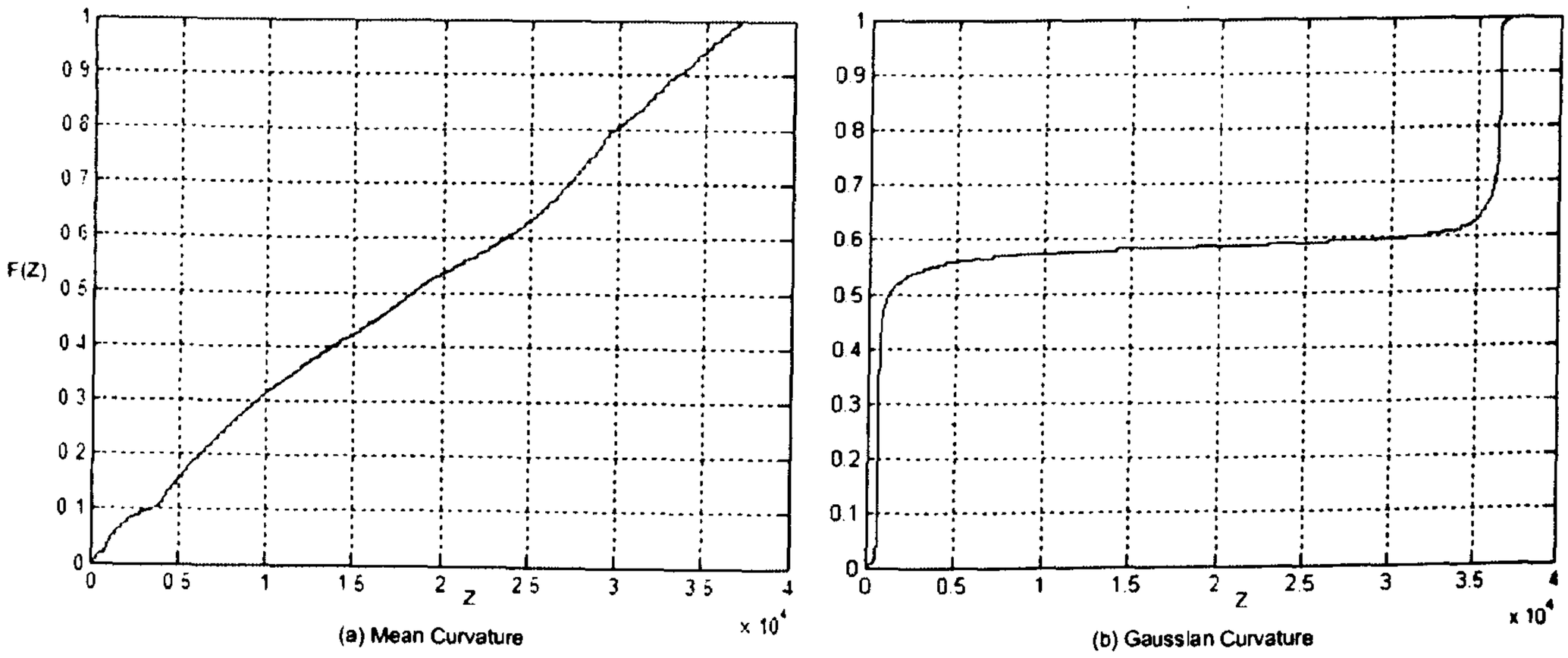


Figure 5.21: Cumulative distribution functions of mean and Gaussian maps after histogram equalization

5.3.4.4. Nested Means Classification

The nested means classification method, originally proposed by Scripter (1970), works by successively placing class breaks at the means of sub-sets of the data. First, the mean of the entire dataset is calculated and a class break is placed at this value. This separates the data into two classes—data values above and below the mean. Each class is further partitioned by calculating the means of the values within these two categories, and inserting class breaks at each of those two points—resulting in four classes. One more level of means will result in eight classes. The process can be repeated to create more classes—the number of resulting classes will always be 2^m, where m is the number of means used to divide the data. Thus the number of classes will be from the set {2, 4, 8, 16, 32, 64, 128, 256}. An illustration of 8-class map using nested means method is shown in Figure 5.22. The subscripts attached to the means indicate the iteration level and the number of means computed. For example, at the first iteration, one mean is computed ($\bar{X}_{11}$); at the second iteration two means are computed ($\bar{X}_{21}$, $\bar{X}_{22}$), and so on.

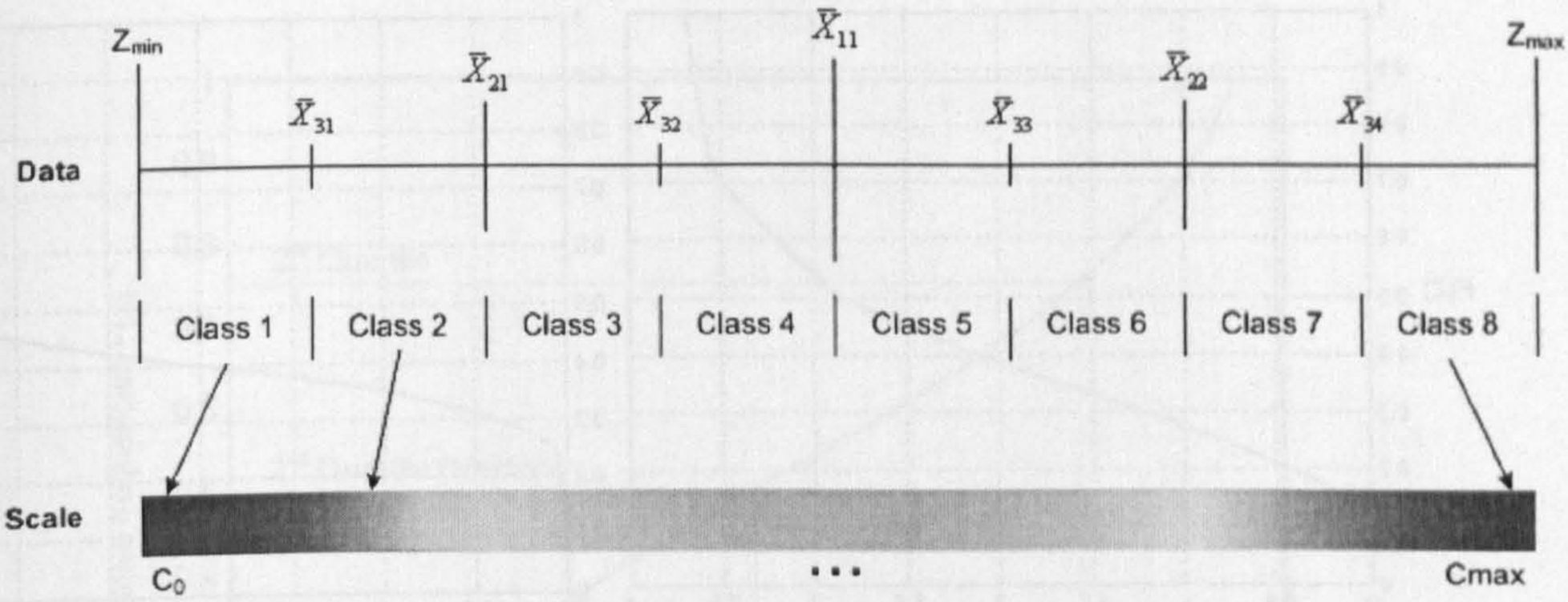
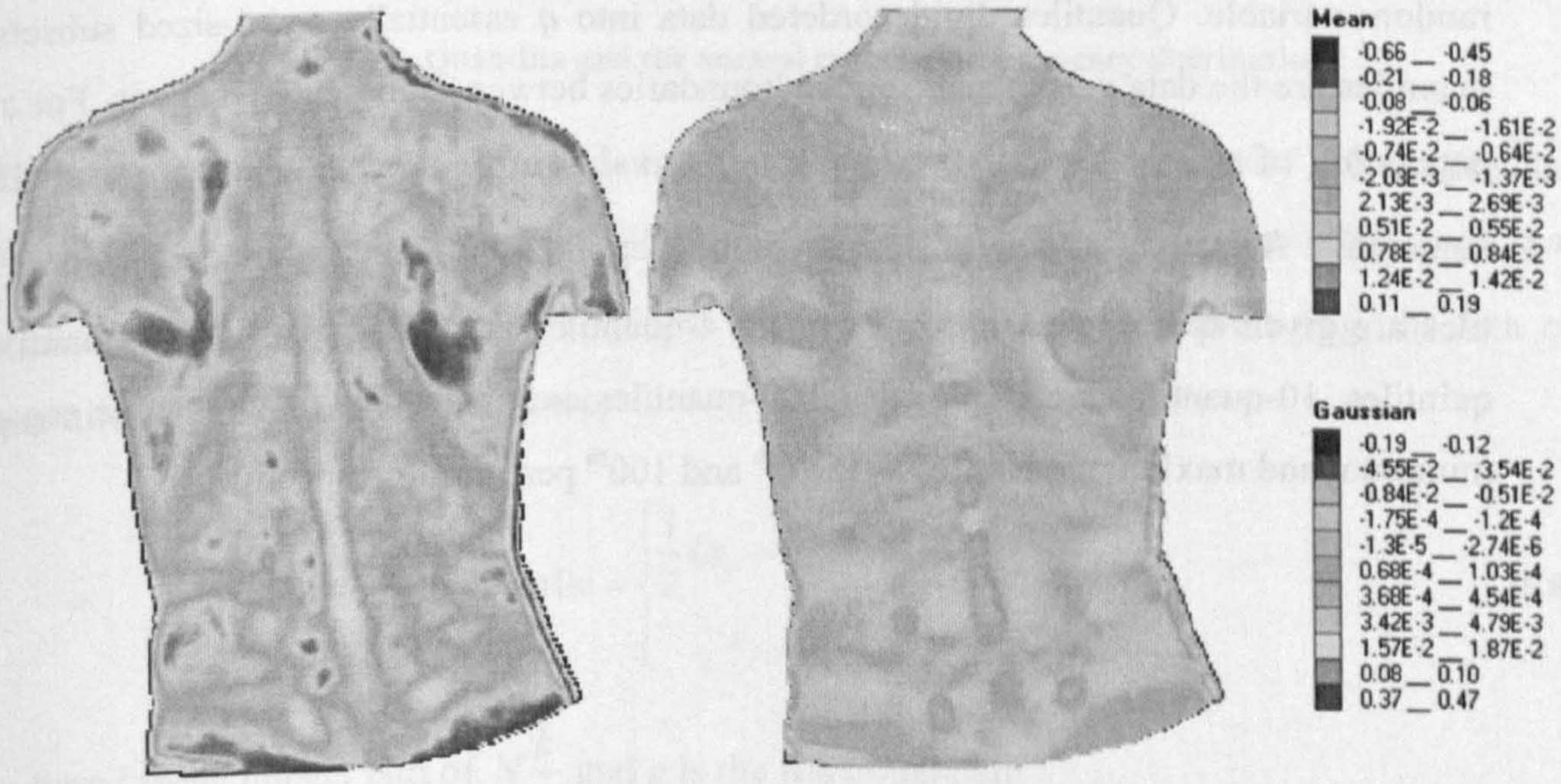


Figure 5.22: Nested means classification: 3 levels resulting in 8 classes

Figure 5.23 shows application of nested means classification to the mean and Gaussian maps. There is a considerable improvement in contrast compared to the equal interval classification method; thus this method could be a compromise for skewed data.

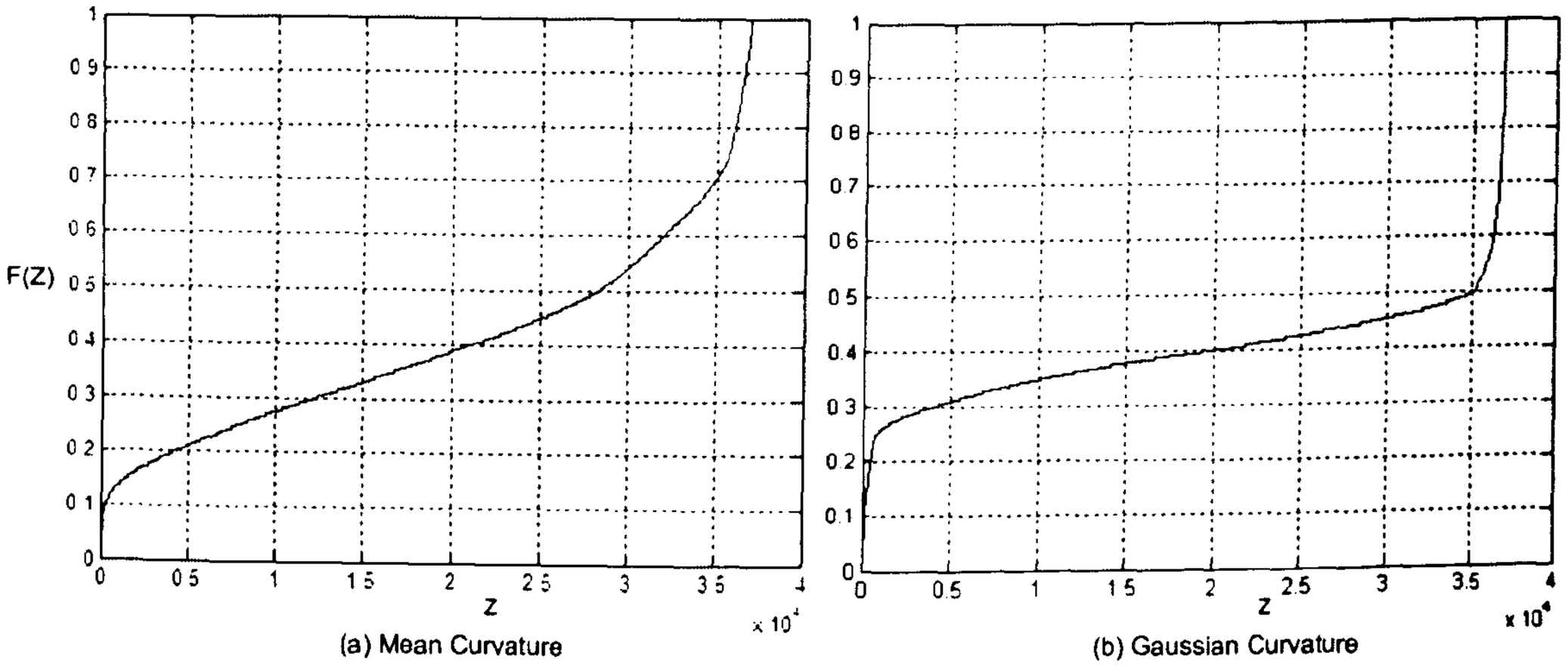


(a) Mean curvature

(b) Gaussian curvature

Figure 5.23: Colour mapping by nested means classification (Subject 1)

The cumulative distribution function of the nested means method is shown in Figure 5.24. The graph is more linear indicating a better classification than equal interval.



(a) Mean Curvature (b) Gaussian Curvature

Figure 5.24: Cumulative distribution functions for nested means classification

5.3.4.5. Quantile Classification

Quantiles are points taken at regular intervals from the cumulative distribution function of a random variable. Quantiles divide ordered data into q essentially equal-sized subsets—the quantiles are the data values marking the boundaries between consecutive classes. For a finite population of size N, the k-th q-quantile is that value of X, say x_k, which corresponds to a cumulative frequency of Nk/q (Kenney and Keeping, 1962) (see Figure 5.25). Some quantiles are given special names, for example: 4-quantiles are called quartiles, 5-quantiles are quintiles, 10-quantiles are deciles and 100-quantiles are percentiles (some people regard the minimum and maximum quantiles as the 0^{th} and 100^{th} percentiles respectively).

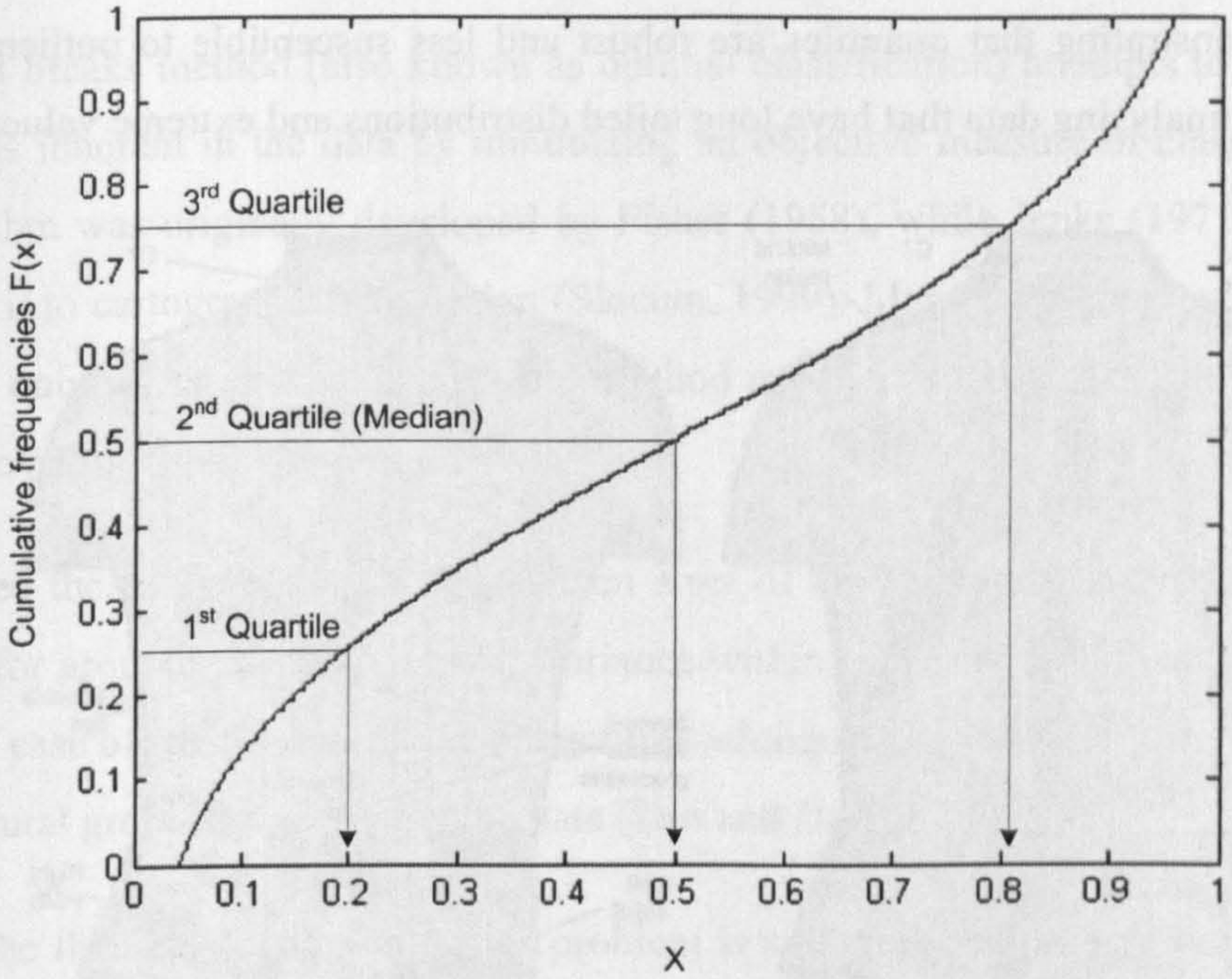


Figure 5.25: Quantiles and the normal cumulative frequency distribution

There are various methods for estimating the quantiles. Let N be the size of a finite population, and let $x_1, x_2, \cdots, x_n$ represent the ordered values of the sample population such that $x_1 \leq x_2 \cdots \leq x_n$. Then using the empirical distribution function with averaging, the k-th q-quantile can be estimated as follows:

$$k\text{-th quantile} = \begin{cases} \frac{1}{2}(x_j + x_{j+1}), & g = 0 \\ x_{j+1}, & g > 0 \end{cases} \tag{5.8}$$

where j is the integer part of $N\frac{k}{q}$ and g is the fractional part.

To implement quantile classification, the data is first sorted and the quantiles are computed according the total number of colours to be mapped (i.e. 256). To determine class membership for each data value, the sorted data are scanned, collecting class members until the number of members in each class is obtained. Figure 5.26 shows quantile classification applied to the mean and Gaussian curvature maps. Several landmarks are of the back are now clearly visible on both the mean and Gaussian curvature maps—including C7 and several spinous process landmarks, left and right PSIS landmarks, midline region, scapula landmarks and pel-

vis region. The Gaussian map which contains most extreme outliers is now properly classified, demonstrating that quantiles are robust and less susceptible to outliers, and thus very useful for analysing data that have long tailed distributions and extreme values.

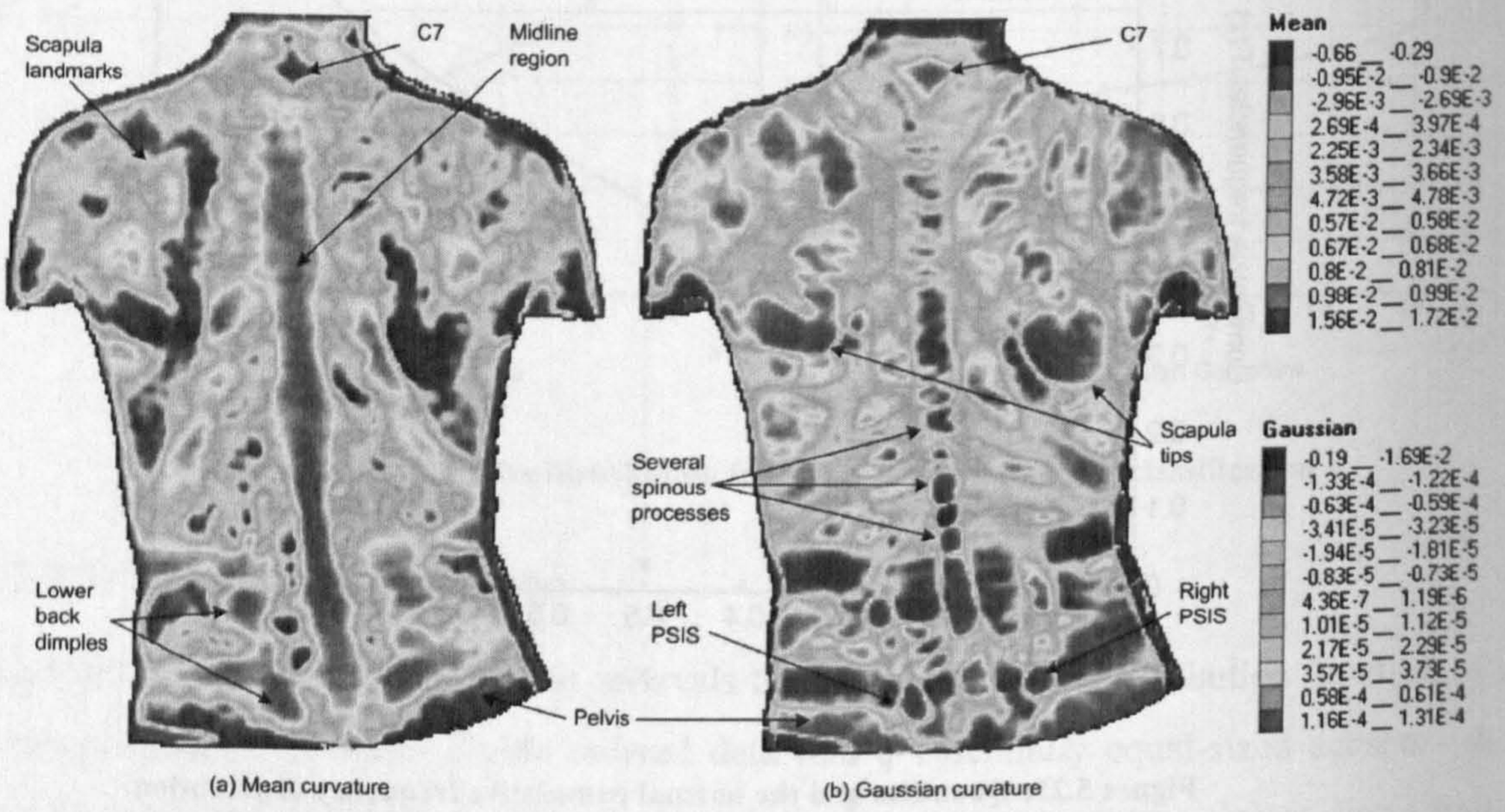


Figure 5.26: Quantile classification, 256 colours (Subject 1)

The empirical cumulative frequency distribution functions for the quantile classification are shown in Figure 5.27. The linear nature of the graphs shows that quantile classification is more robust to outliers and thus no further enhancement is necessary.

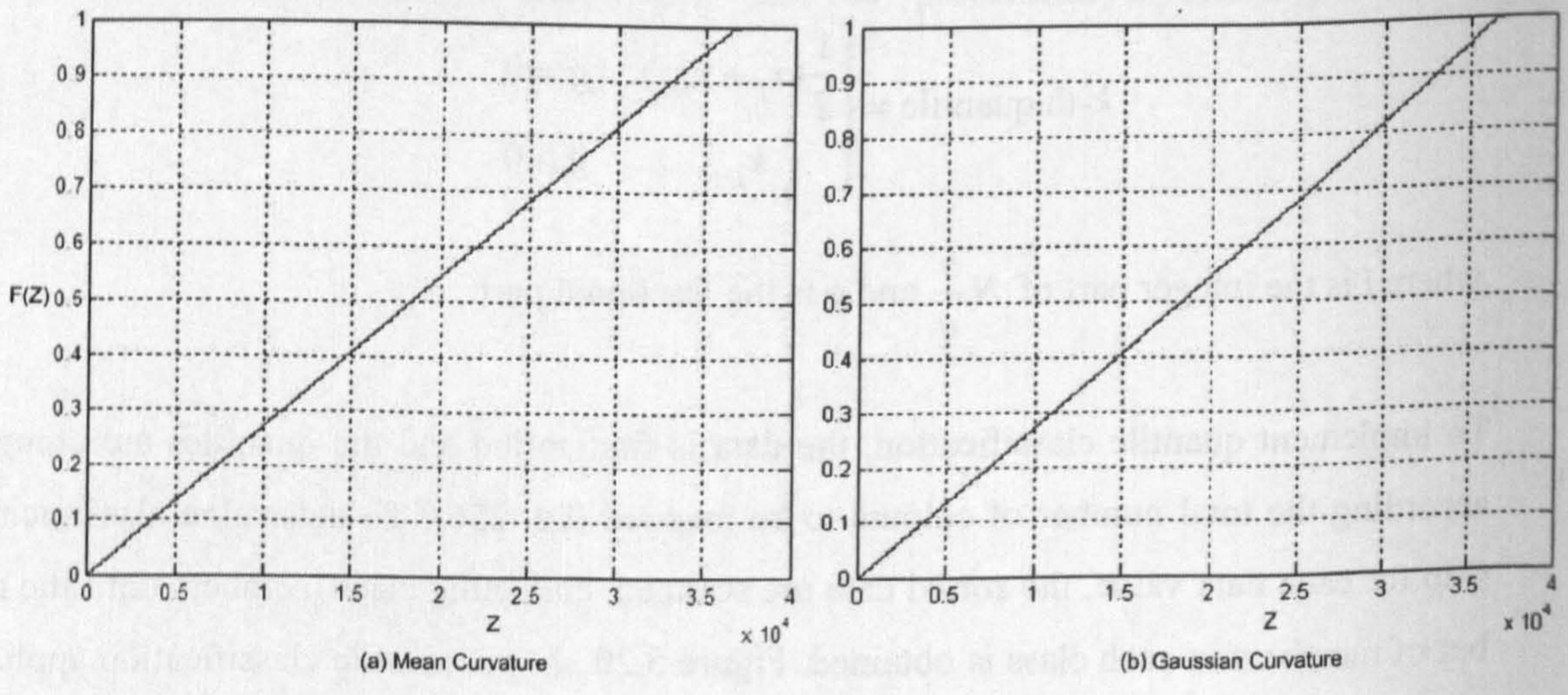


Figure 5.27: Cumulative distribution functions for Quantile classification

5.3.4.6. Natural Breaks Classification

The natural breaks method (also known as optimal classification) attempts to find groupings and patterns inherent in the data by minimizing an objective measure of classification error. The algorithm was originally developed by Fisher (1958), while Jenks (1971) modified and introduced it to cartographic map design (Slocum, 1999). Many cartographers have generally recognised only Jenks, and as a result, the method is often called 'Jenks optimal method' in much of the cartographic literature.

Fisher stated the problem as follows: "Given a set of arbitrary numbers, what is a practical procedure for grouping them so that the variance within groups is minimised?" This problem is a special case of multi-dimensional clustering techniques that partition data in an attempt to recover natural groupings present in the data (Jain and Dubes, 1988).

Although the theoretical solution to this problem is straightforward, there are difficulties associated with finding a practical solution. The first issue is how to select a classification criterion that translates one's intuitive notion of a natural 'cluster' into some mathematical formulation. Since there is no a precise definition of a cluster, classification criteria are highly application dependent. The most common classification criterion is based on the squared error criterion. With this criterion, the problem can be stated as follows. Let X be a finite set of size N with each element (x_i) assigned a weight (w_i), then the problem is to classify or partition the elements of X into n classes by minimizing the weighted sum of the squared deviation (D) in Equation 5.9:

$$D = \sum_{i=1}^{N} w_i \left(x_i - \bar{x}_i \right)^2 \tag{5.9}$$

where $\bar{x}_i$ is the weighted mean of those elements that are assigned to the class to which element i is assigned. In the special case when all elements have equal weights, D is simply the sum of the squared deviations. Other approaches use the median as a measure of central tendency and the associated absolute deviations about the class medians as a measure of the classification error (Slocum, 1999).

Unfortunately, the number of all possible classifications becomes prohibitively high for large datasets— $N = 100, n = 7$ will result in over 16 billion possible classifications to evaluate! Thus exhaustive evaluation of all possible partitions is not computationally feasible, even for

small datasets! Rather than consider all possible solutions, Fisher (1958) proved that any optimal partition is simply the sum of optimal partitions of subsets of the data. Consequently, not every possible solution needs to be individually computed, provided suitable records are kept. Although this optimization speeds up the computation process, it is still very slow for practically large datasets. Consequently, in practice, the following approach is often adopted to arrive at a solution. Starting with an initial classification, data are moved from one class to another in an effort to improve the value of the criterion function. Thus, successive iteration in the classification process is just a perturbation of the previous classification, resulting in evaluation of only a small number of classes. Unfortunately, algorithms based on this technique generally do not generally achieve a global minimum, but often converge to a local minimum, which may not be the optimal partition (Jain and Dubes, 1988).

Jenks (1971) modified the original algorithm of Fisher and developed a goodness fit statistic known as goodness of variance fit (GVF), which is calculated as:

$$GOF = 1 - \frac{SSE_C}{SSE_D} \tag{5.10}$$

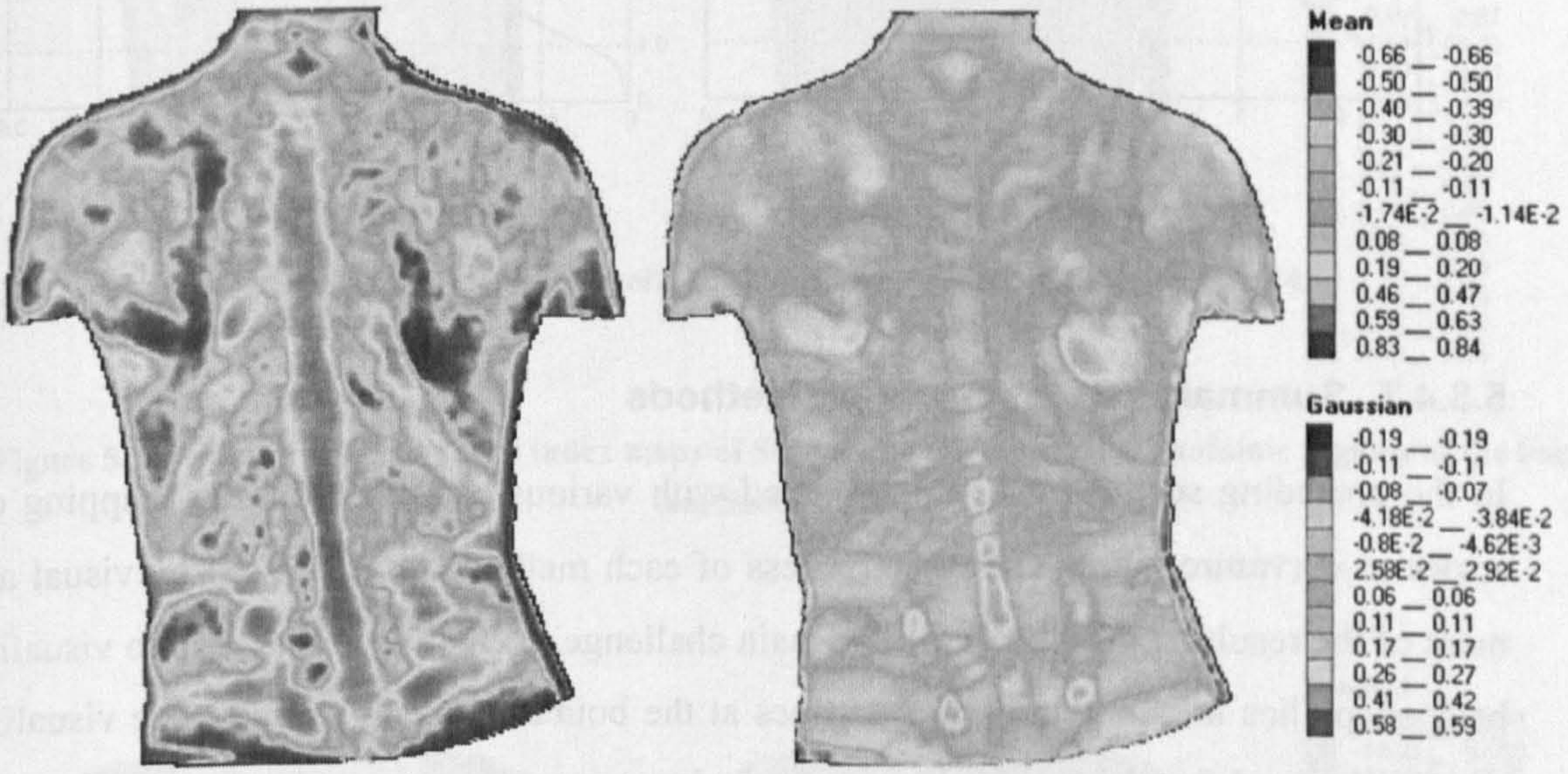
where SSD_D is the sum of the squared deviations of the data elements from the mean of the entire dataset, and SSD_C is the sum of the squares of the deviations of the class members from their class mean. Jenks also provided improved implementation of the algorithm (Groop, 1980; Robinson, 1995).

The steps in computing the optimal classification of any dataset can be summarised as follows:

- Sort the entire data.
- Compute the mean (or median) of the entire dataset and calculate the sum of the squares deviations (SSE_D) of the observations from the mean
- Develop class boundaries for the first iteration, calculate the class means (or median), and calculate the sum of the squares of the deviation of each class member from its class mean (denote this by SSE_C).
- Compute the good of fit statistic (GOF) which represents the sum of squared deviations between the classes.

- Repeat steps 2-4 until the GOF cannot be maximised further.

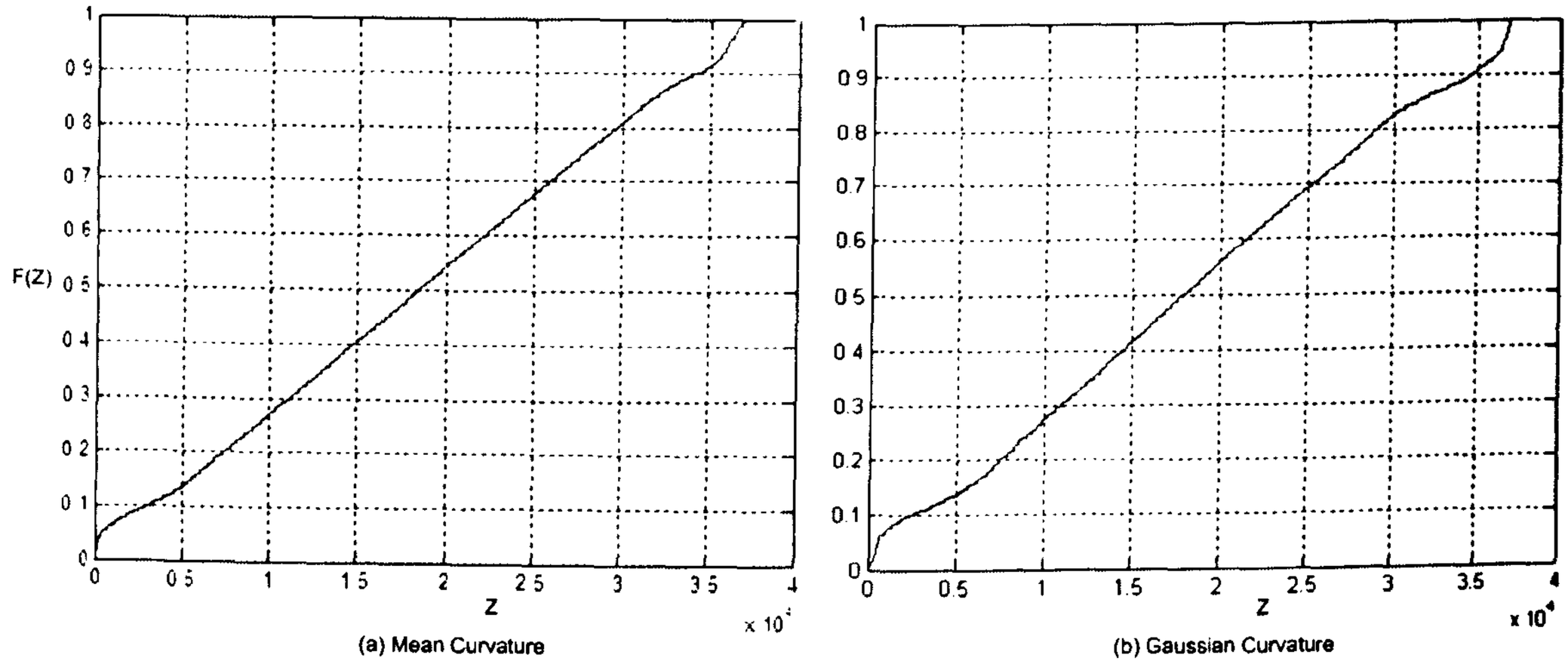
Figure 5.28 shows application of the natural breaks classification method to the mean and Gaussian maps. A total number of 256 classes were computed for both maps using the sum of squares about the class means as the measure of classification error. Both maps show considerable contrast, although the effect of extreme outliers is still obvious from the Gaussian map.



(a) Mean curvature (b) Gaussian curvature

Figure 5.28: Colour mapping by natural breaks classification (Subject 1)

The cumulative frequency distribution functions of the natural breaks method shown in Figure 5.29 for the mean and Gaussian maps. The graphs are almost linear, demonstrating that the method works quite well. The main problem with the method is the time taken for computation. While all of the other methods took just a few fractions of a second to compute, the natural breaks method took several minutes on a typical modern computer (Pentium 4, 2.4GHz, 1G RAM). This severely reduced the performance of the system, and hence this method is not feasible for generating visualizations on-the-fly.



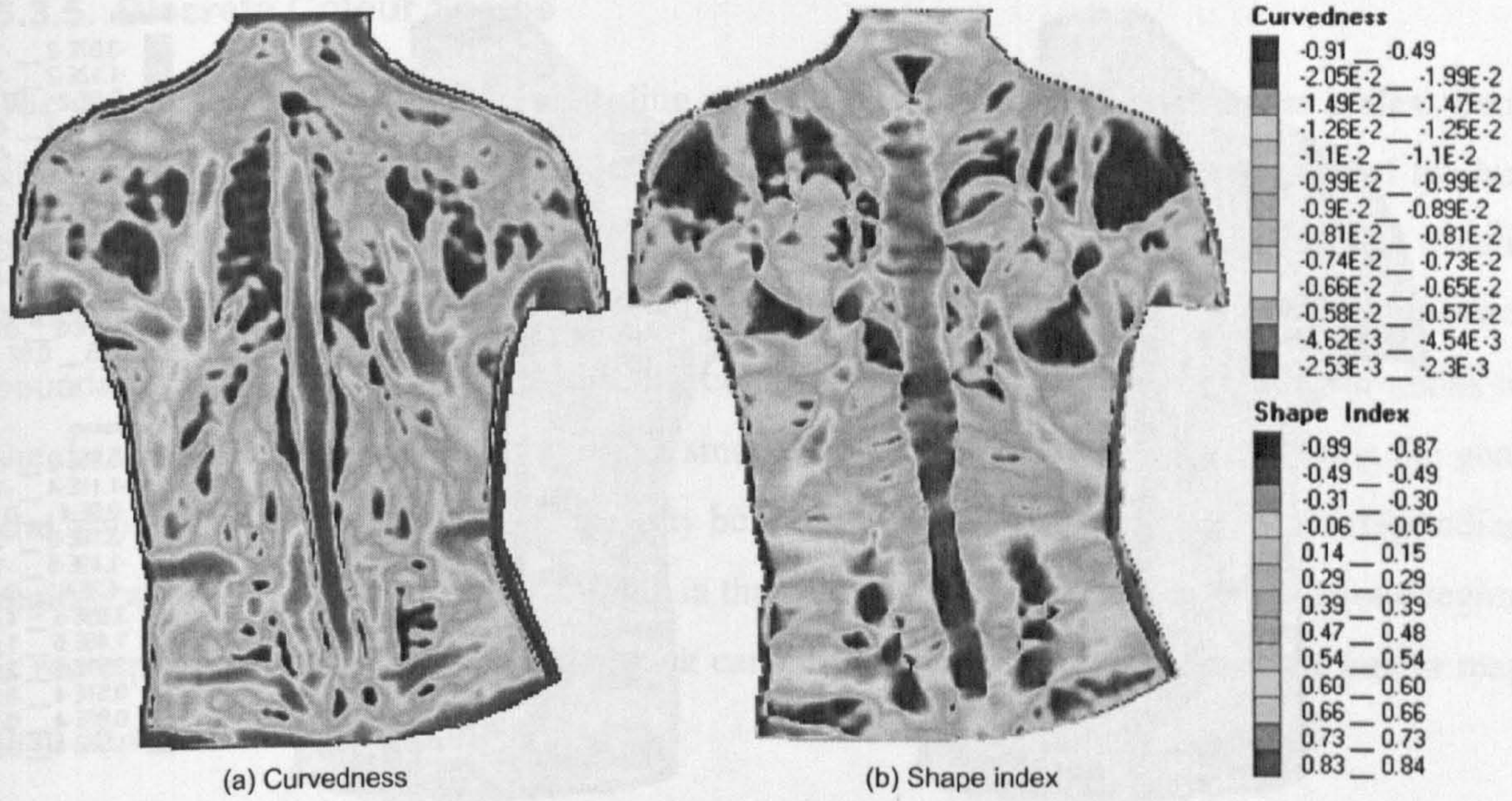
(a) Mean Curvature (b) Gaussian Curvature

Figure 5.29: Cumulative distribution functions for natural breaks classification

5.3.4.7. Summary of Classification Methods

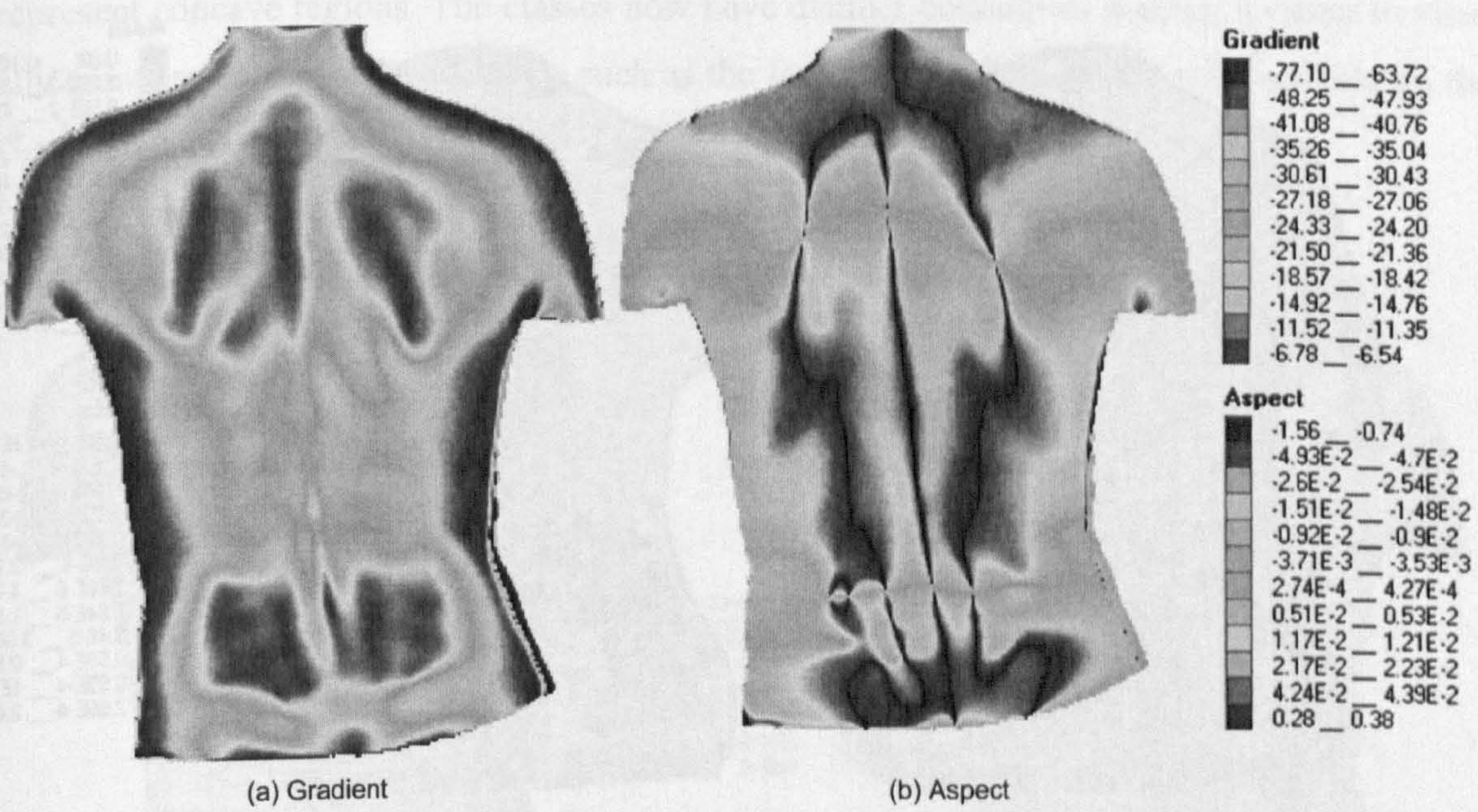
In the preceding sections, we experimented with various classifications for mapping colour scales to curvature maps. The effectiveness of each method was discussed by visual assessment of the resulting visualization. The main challenge in creating colour map to visualize the back shape lies in handling extreme values at the boundary of the dataset. The visualization results obtained from the preceding sections demonstrate that quantile classification is very robust to these extreme values. A further advantage of quantile classification is that the computation of quantiles is relatively simple and less computationally demanding task, compared with histogram equalization, nested means and natural break methods. For discrete data, quantile computation consists of sorting followed by equal internal classification of the sorted data. This implies that the visualizations can be generated on-the-fly and that various colour scales can be applied quickly to determine their effectiveness.

More examples of back visualizations of the same subject based on quantile classification are given in Figure 5.30 (curvedness and shape index) and Figure 5.31 (gradient and aspect). Figure 5.32 shows the mean and curvature maps of a second male subject (Subject 2: age 30, height 170cm, weight 76kg, BMI 24kg/m^2) based on quantile classification. And Figure 5.33 shows the mean and Gaussian curvature maps of a female subject (Subject 3: age 27, height 167cm, weight 58kg, BMI 21kg/m^2).



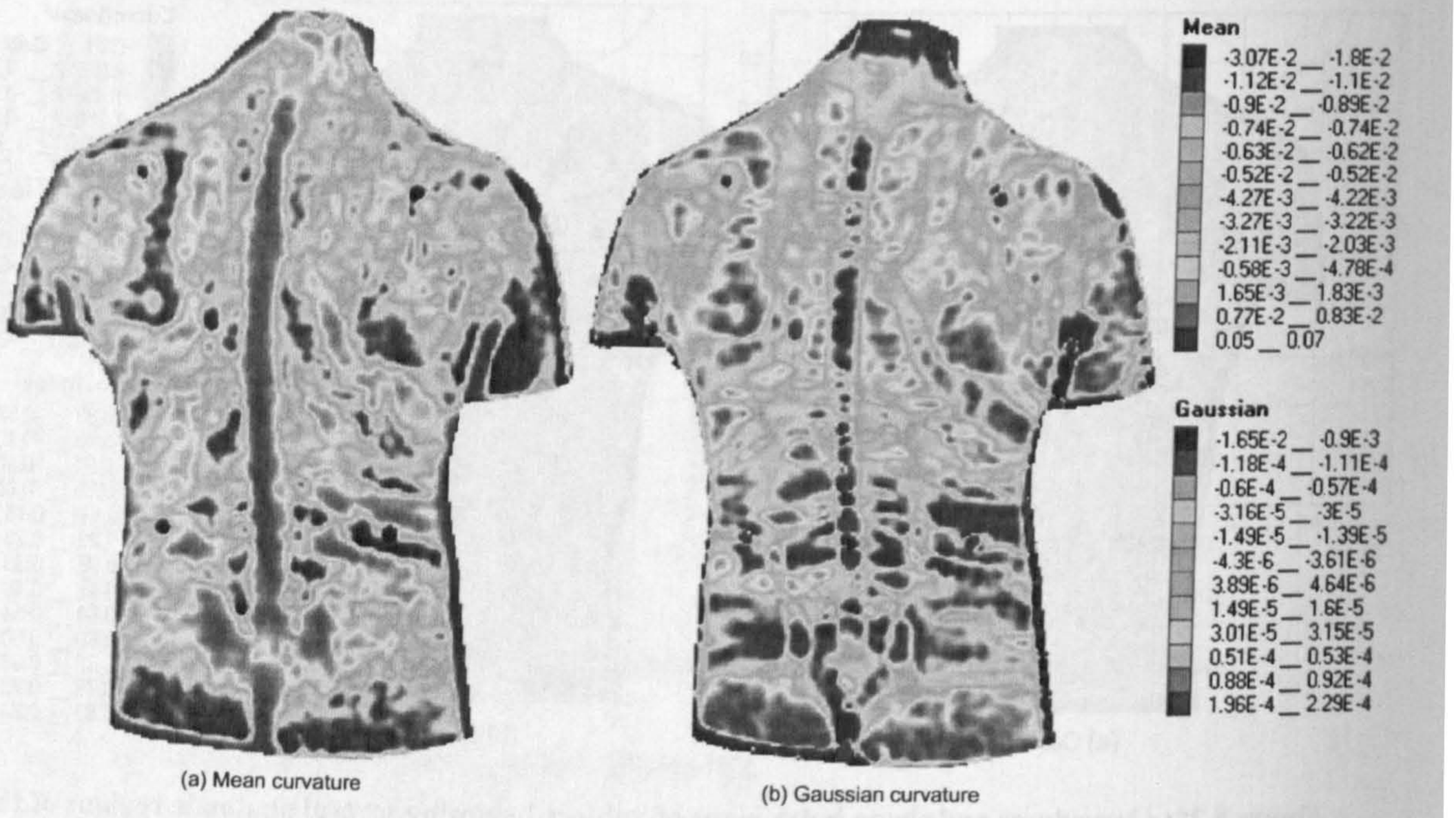
(a) Curvedness (b) Shape index

Figure 5.30: Curvedness and shape index maps of Subject 1 showing several anatomic regions of the back (see Section 4.2.4)



(a) Gradient (b) Aspect

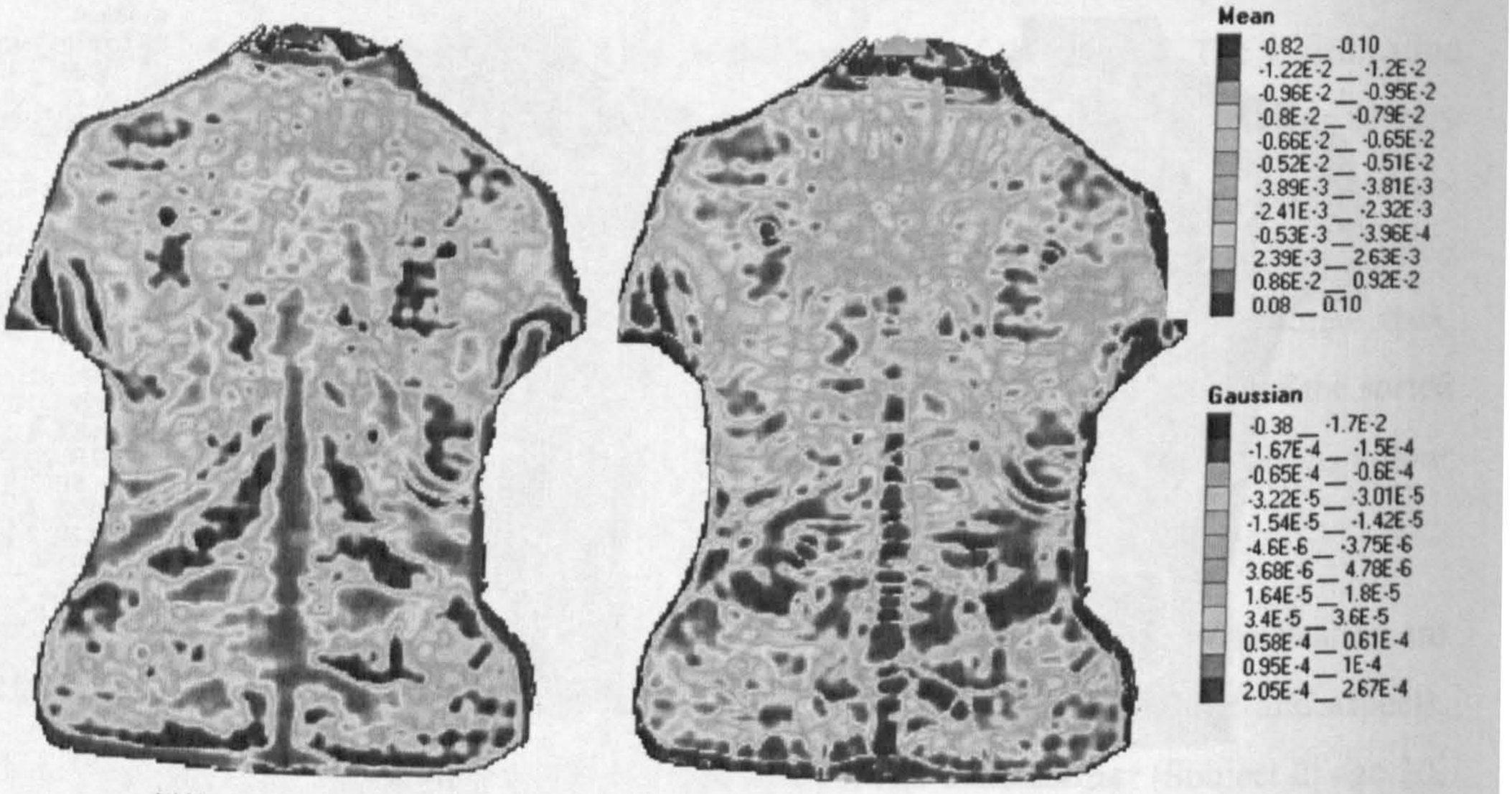
Figure 5.31: Gradient and aspect maps of Subject 1. Gradient describes steepest descent at each point on the surface while aspect describes the direction of this steepest descent (see Section 4.2.5)



(a) Mean curvature

(b) Gaussian curvature

Figure 5.32: Mean and Gaussian curvature maps by Quantile classification (Subject 9: male, age 30, height 175cm, weight 76kg, BMI 25kg/m^2)



(a) Mean curvature

(b) Gaussian curvature

Figure 5.33: Mean and Gaussian curvature maps by Quantile classification (Subject 25: female, age 27, height 167cm, weight 58kg, BMI 21kg/m^2)

5.3.5. Discrete Colour Scales

The colour scales shown in the preceding sections are perceptually continuous scales. Continuous scales are useful for mapping continuous variables (such as surface shape of the human back) as they preserve the continuity of the variable being mapped. Discrete colour scales limit the number of classes or colours to a few in order to create perceptually distinct boundaries between the classes. A discrete colour scale can sometimes highlight variations in the data which may not be visible from a smooth version. For example, in assessing the general asymmetry of the back shape, one may be interested in visually comparing corresponding regions at the left and right of the midline in the frontal plane to determine whether one region is more prominent than the other. This task can easily be carried out on a discrete colour map than on a smooth version.

Figure 5.34 shows the mean and Gaussian maps displayed with a discrete colour scale using the quantile classification method. The number of classes is restricted to 5 so that in the colour scale, Blue represents convex regions, Green represents relatively flat regions and Reds represent concave regions. The classes now have distinct boundaries making it easier to visually compare corresponding regions such as the lower tips of the left and right scapula on the Gaussian curvature map.

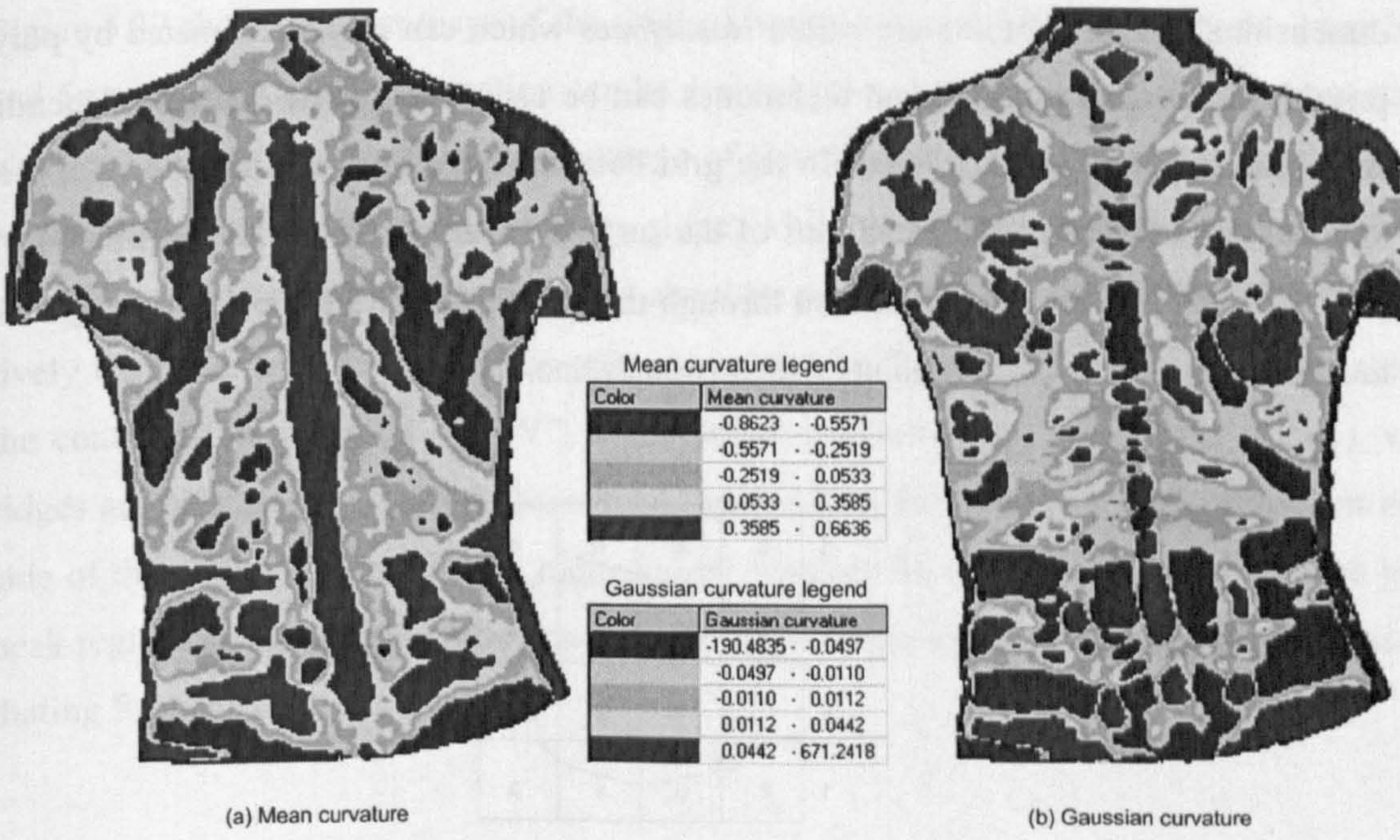


Figure 5.34: Discrete colour scale with quantile classification, 5 classes (Subject 1)

5.3.6. Contour Maps

Contours or isolines are lines of equal value. Contour maps are a natural extension to colour mapping. When visualising a colour-mapped surface, the eye often separates similarly coloured regions into distinct regions or classes. Contour maps effectively construct these boundaries. Contours are commonly used to visualize various data such as terrain elevation, meteorological and medical data; thus many people are familiar with them. Contours are useful in back shape analysis, as they encode information about asymmetry of the back in a form that is readily recognizable by clinicians, as is evident from the wide applications of moiré topography. It was discussed in Section 3.3.2.4 that one of the problems with moiré topographs is that the resulting contours are easily affected by rigid-body motion of the body during data capture. Since 3D laser scanning directly gives coordinates of the back surface, this is not a problem as contours can be interpolated relative to a reference plane such as the plane defined by the left and right PSIS landmarks. Another advantage is that contours with varying intervals can be generated on-the-fly to allow asymmetry analysis at varying levels of detail.

5.3.6.1. Contouring Algorithms

For a surface defined by $z = f(x, y)$, a 2D contour can be expressed as a curve $f(x, y) = c$ separating two regions $f(x, y) < c$ and $f(x, y) > c$, where c is the contour value. In three-dimensional space, contours are called *isosurfaces* which can be approximated by polygonal primitives. Various interpolation techniques can be used to compute contours depending on the particular data structure used. In the grid data structure shown in Figure 5.35, the pixel values shown represent the elevations of the surface at each point (i.e. z coordinate). A contour with value of 5 can be generated through the grid by linear interpolation along the edges and connecting the resulting points with a smooth line.

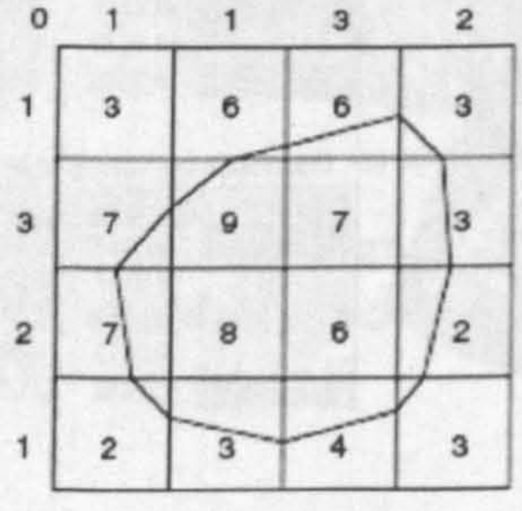


Figure 5.35: Contour generation by linear interpolation through a 2D grid data structure

One approach for practical contour generation is the matching square algorithm (matching cube in 3D), which relies on the fact that a contour can pass through a pixel in only a finite number of ways (see Lorensen & Cline, 1987). A look-up table (LUT) which contains all the possible intersection of a contour line and each pixel is constructed (see Figure 5.36). A vertex is considered inside a contour if its value is larger than the value of the contour line. Vertices with values less than the contour value are said to be outside the contour. For each pixel, the inside/outside state of each vertex is calculated. This data is used as input into the LUT to determine the topological state of the pixel, which is then used to calculate the contour location via interpolation.

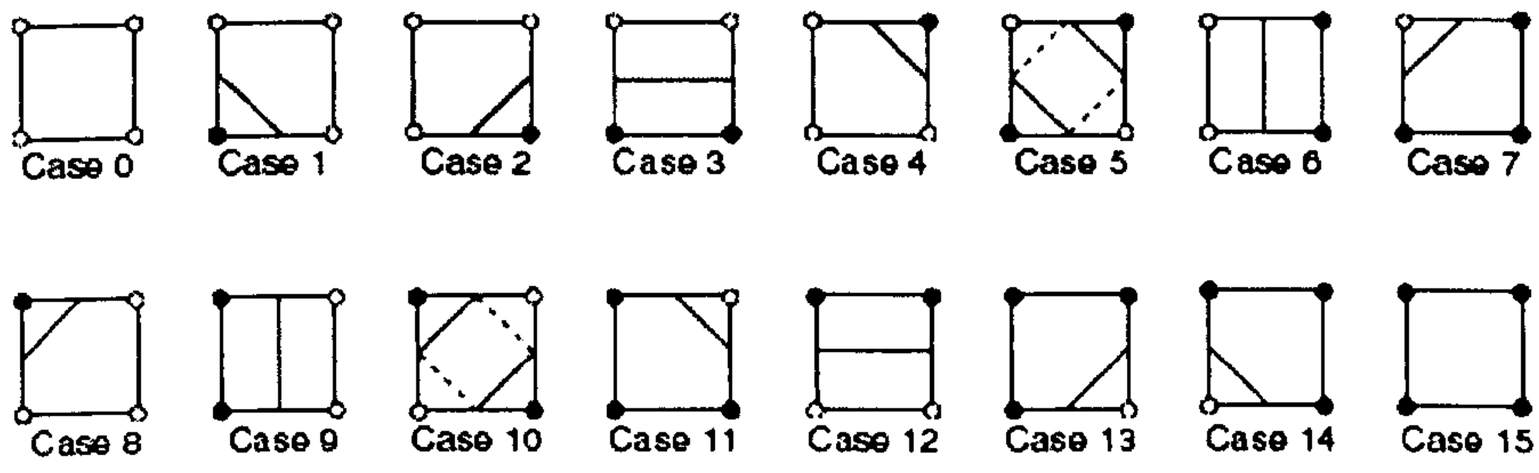
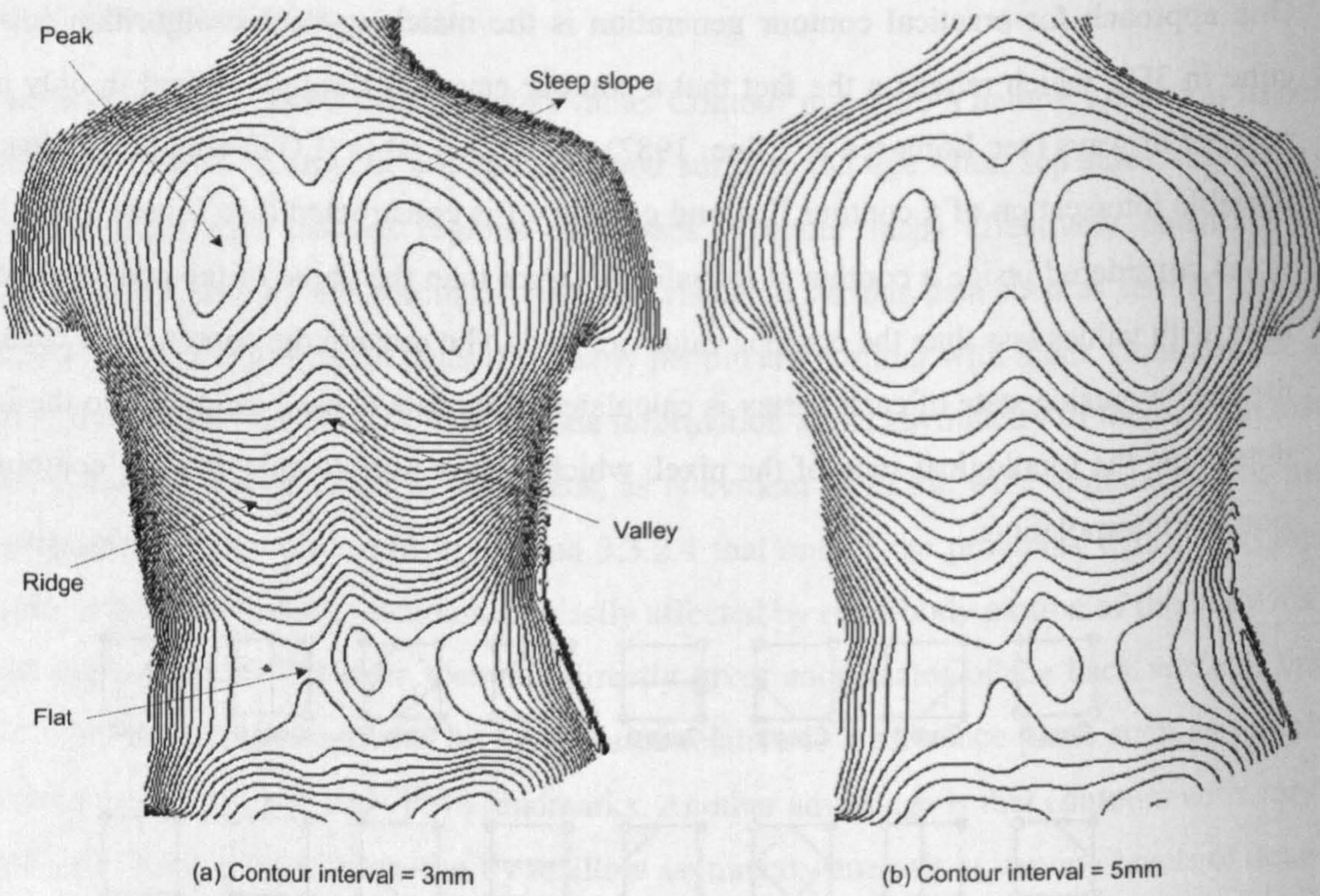


Figure 5.36: Matching square algorithm for contour generation in 2D

5.3.6.2. Interpretation of Contours

Figure 5.37 shows contour maps of the original back surface at different intervals: 3mm in (a) and 5mm in (b). Useful information can be derived from these contours. Moving perpendicular to the contours shows the rate of change of elevation of the back surface. Closely spaced contours indicate rapid change or steep slope, while widely spaced contours indicate slow change or gentle steep. For example, the shoulder regions display steep slope while the relatively flat lower back region display spaced contours in Figure 5.37. Valleys are indicated by the contours forming a 'U' (or 'V') shape pointing uphill (as in the midline region), while ridges are indicated by the 'U' shape contours pointing downhill (as on the ridges on either side of the midline). Summits are indicated by contours forming concentric curves (as in the peak regions of the scapula), and depressions are indicated by circular contours with lines radiating from the centre.

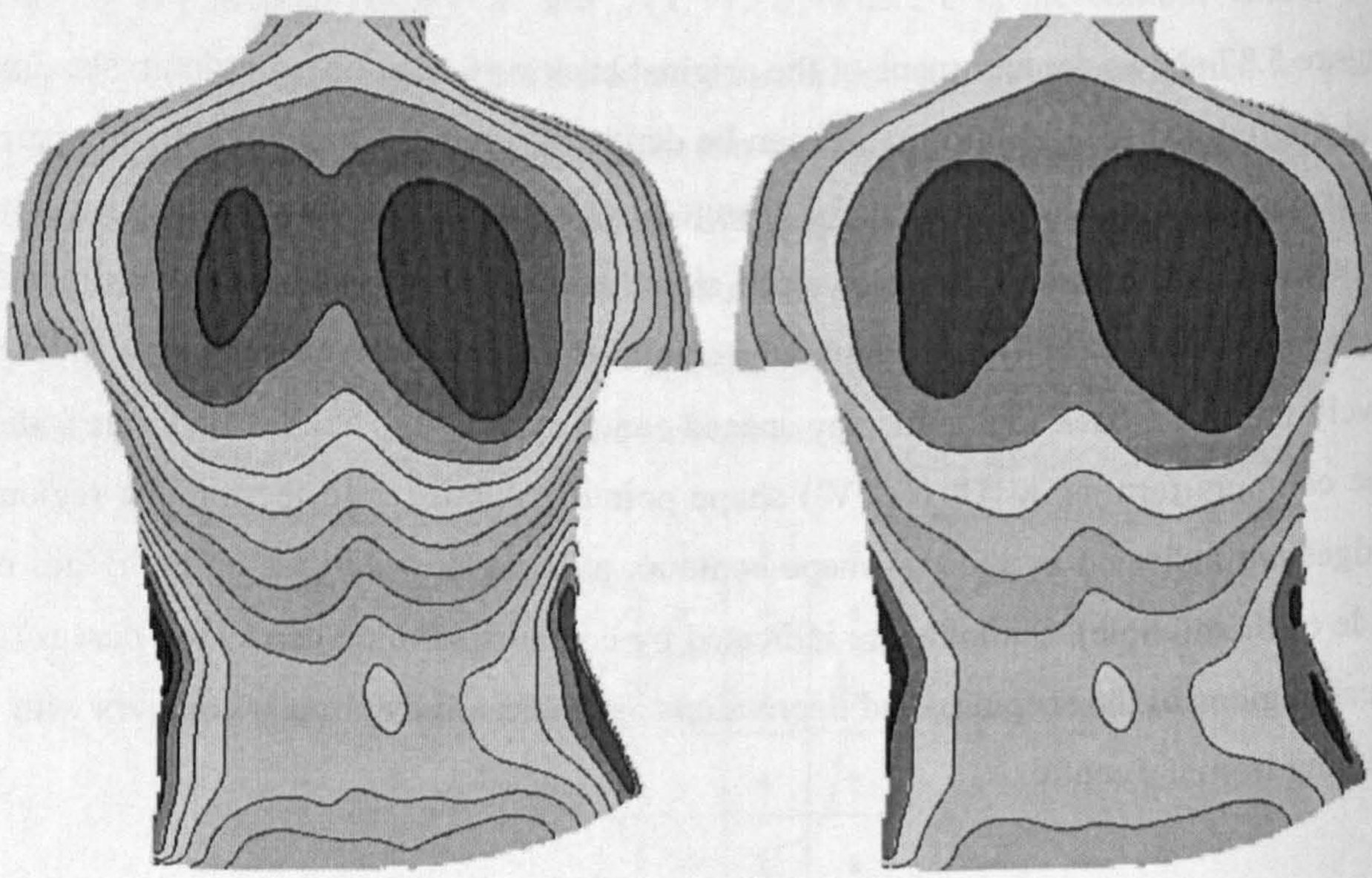


(a) Contour interval = 3mm

(b) Contour interval = 5mm

Figure 5.37: 2D contours of orginal back surface (Subject 1)

The space between contours can also be colour-filled to maximise contrast, as in Figure 5.38.



(a) Contour interval = 9mm

(a) Contour interval = 15mm

Figure 5.38: Contours with filled colours (Subject 1)

5.3.7. Three Dimensional Views

Three dimensional views allow real time exploration of the back shape, ability to zoom, rotate and focus on various aspects of the shape of the back in order to assess the extent of the deformity. Issues such as the effect of bracing or surgery can also be explored without the need to physically manipulate the patient with the problem. Figure 5.39 shows sample 3D views (a) colour map of original surface (b) colour map with contours superimposed in 3D.

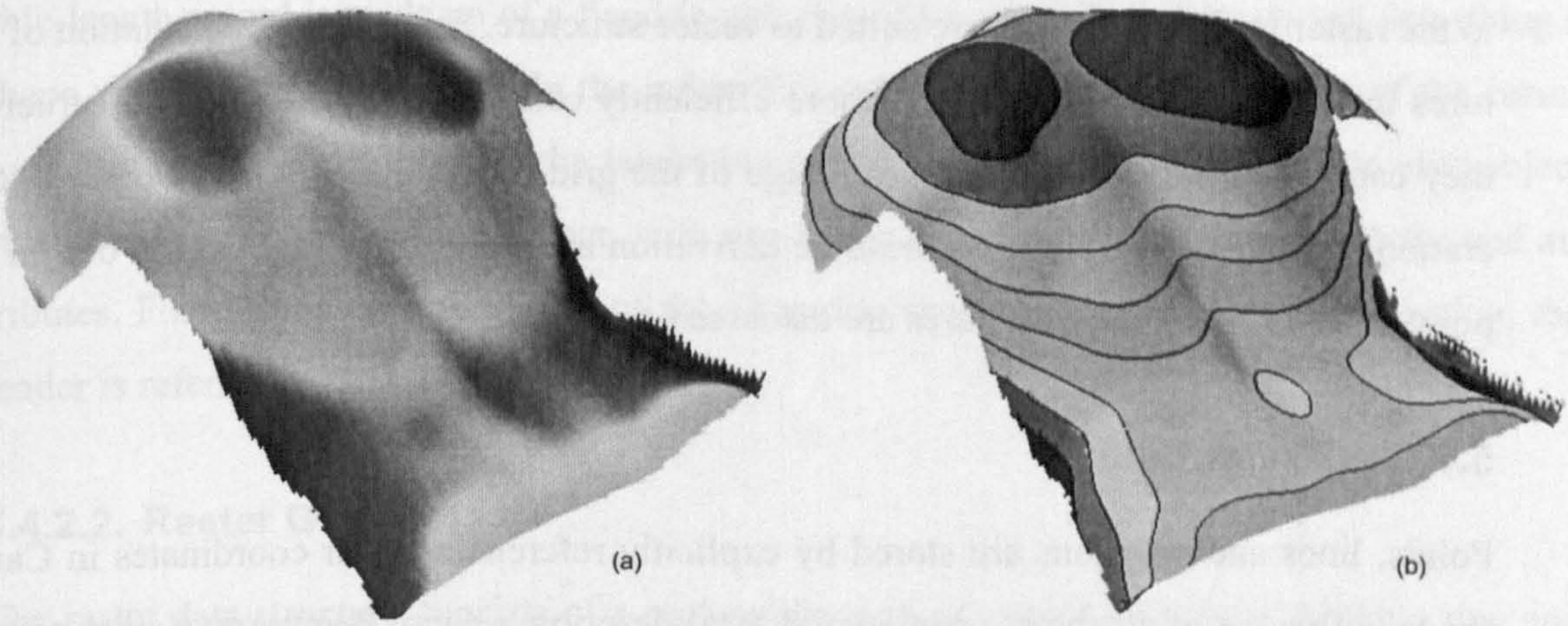


Figure 5.39: 3D views of original surface with contours superimposed (Subject 1)

5.4. Implementation Considerations

The complete software developed consisted of more than 50,000 lines of program code that were distributed in over 500 program files. Clearly, implementation of this size required careful planning and application of the essential principles of software engineering. This section briefly discusses the major design issues regarding data structures and conversion, system design and development tools.

5.4.1. Raw Data

The output from the 3D laser scanner is a set of XYZ Cartesian coordinates of the back surface (in millimetres) and an RGB colour image. Typically, the scanner is able to produce over 80,000 points on the back surface in normal resolution mode. The scanner comes with proprietor's software which offers some basic functions to visualize and edit the output data. For example, the actual number of points measured can be sub-sampled to the desired size. In our

implementation, the point data from the scanner is exported into ASCII, TXT or DXF format which is then imported into the data structures used in our implementation (discussed below).

5.4.2. Data Structures

A data structure refers to the organization of data in a form suitable for computer processing. Two types of data structures have been used in the implementation, namely, vector data and raster grid. The reason for using both data structures is that some operations are more suited to the raster grid while others are suited to vector structure. For example, calculation of curvatures by parabolic surface fitting is more efficiently carried out on the grid data structure, as they can be simplified by taking advantage of the grid configuration. On the other hand, operations such as curve fitting for midline derivation are more efficiently carried out on vector point data. The two data structures are discussed further below.

5.4.2.1. Vector Data

Points, lines and polygons are stored by explicitly referencing their coordinates in Cartesian space with a list of attributes (non-spatial data) describing them. Point arrays store each point as a row in the table, with positional and non-spatial attributes stored as fields. Lines are stored as strings of connected straight line segments defined by ordered sequence of their vertices. Polygons are stored as a closed loop of lines (Figure 5.40).

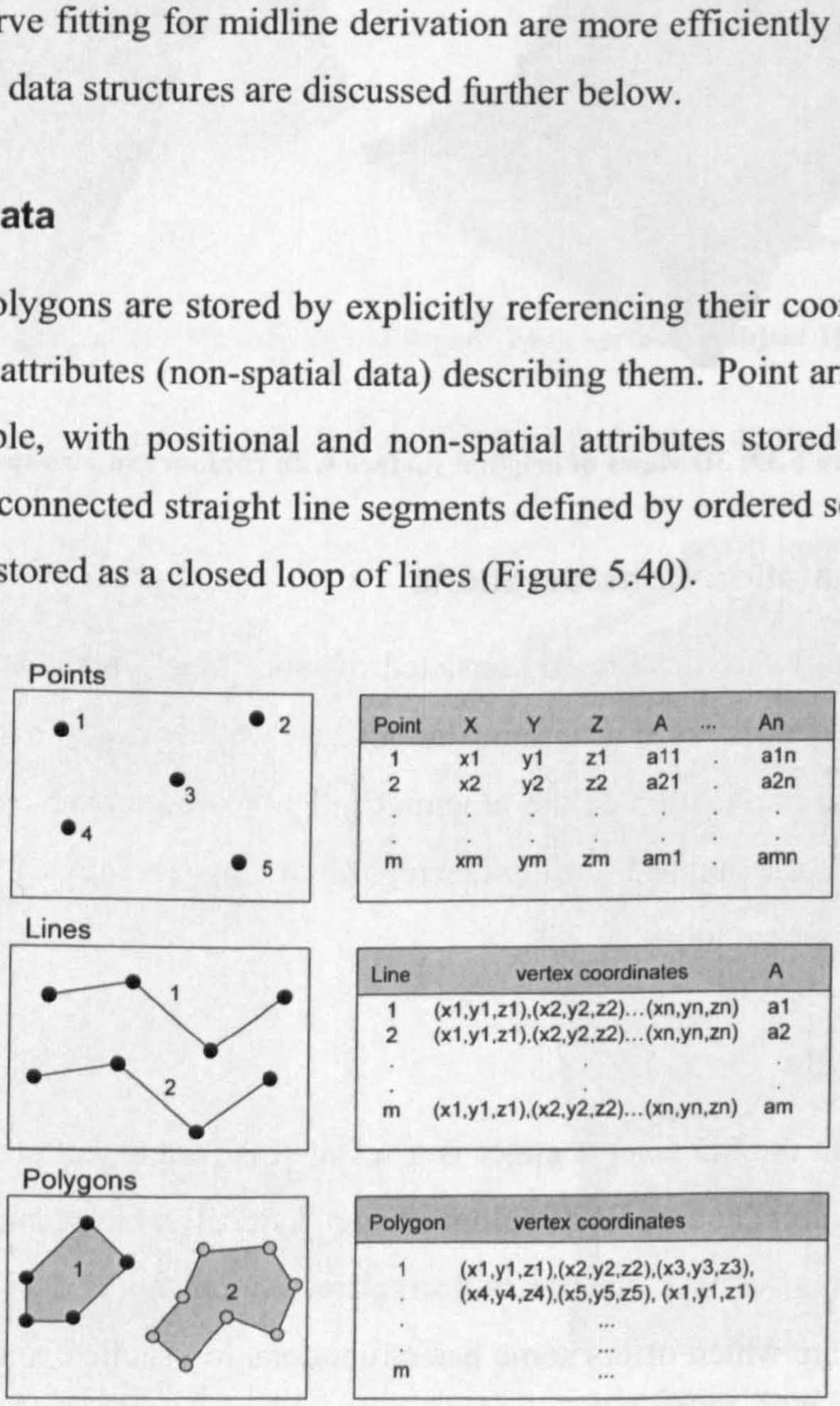


Figure 5.40: Vector data structure

A specific vector data structure template defined by the ESRI (Environmental Research Institute) and known as Shapefile has been used in our implementation. The Shapefile data structure has been proven to be efficient, robust and has been accepted as standard structure for spatial data in Geographic Information Systems (GIS) applications. Moreover, the author is completely familiar with this data structure. The Shapefile data structure consists of three separate files: a main file (.shp extension), an index file (.shx), and a database file (.dbf). The main file contains a fixed-length file header followed by variable-length records. Each variable-length record is made up of a fixed-length record header followed by record describing a shape with a list of its vertices. In the index file, each record contains the offset of the corresponding main file record from the beginning of the main file. The database contains object attributes with one record per object, with one-to-one relationship between geometry and attributes. For a complete description of the Shapefile specification and data organization, the reader is referred to (ESRI, 1998).

5.4.2.2. Raster Grid

The raster data structure consists of a rectangular grid or array of pixels in which a row and column coordinates reference a particular location in the raster (Figure 5.41). Each pixel is associated with a value (v_i) that determines what the pixel represents. For 2D images, v_i represents the colour value at that pixel location. For a general surface of the form $z = f(x, y)$, v_i represents the z-coordinate at the given (x, y) location. Given information about the size of the grid, and the scan order, the grid can be stored using two-dimensional arrays. The resolution of the raster data depends on its pixel size—the smaller the pixel size the higher the resolution and the larger the data size. Typically, in a raster data structure, the spatial coordinates are not explicitly stored, as the sequence of pixels provides an implicit spatial address, once the sequence of rows and columns (called *scan order*) is established. This configuration also allows a high amount of compression of raster data.

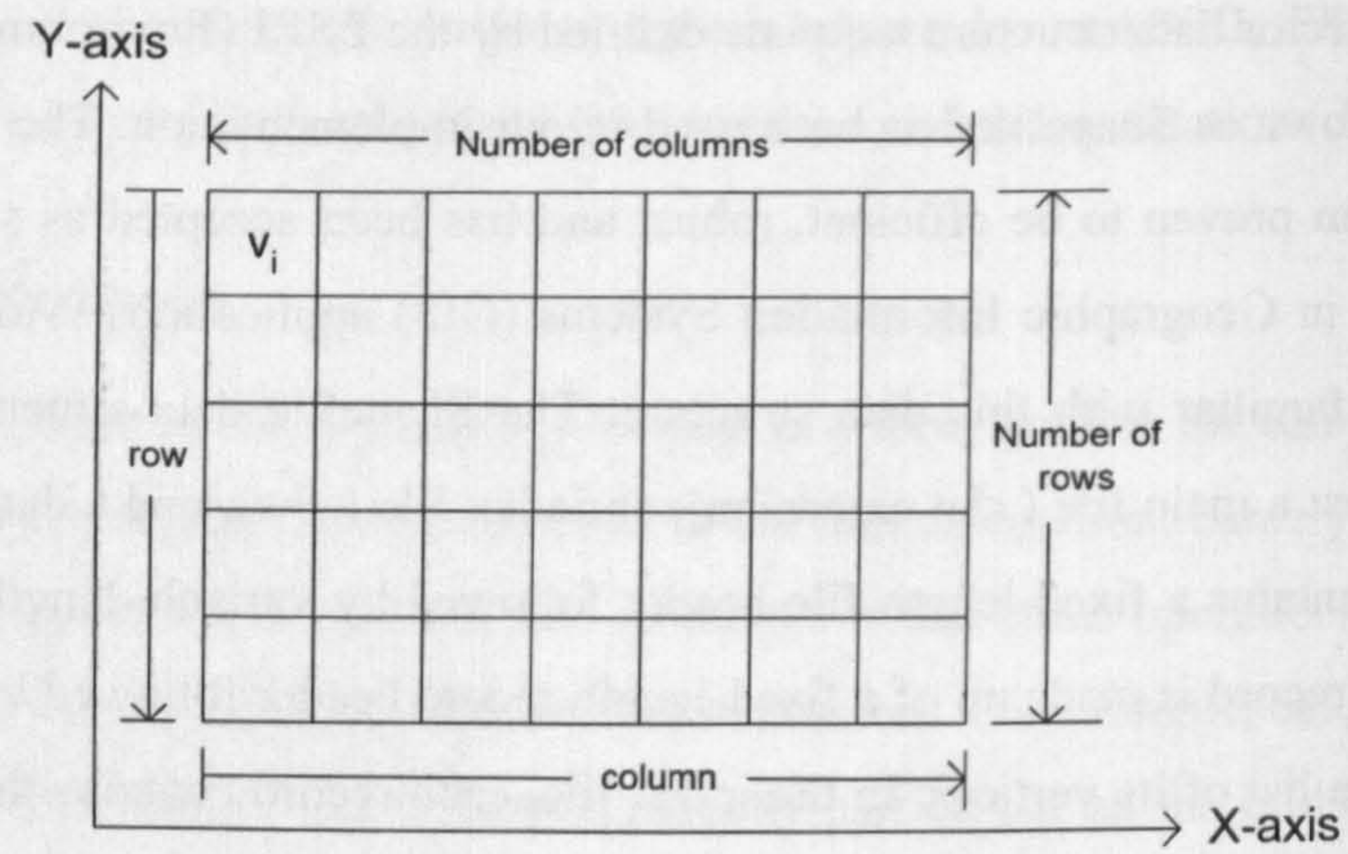


Figure 5.41 Raster grid: v_i represents grey level in 2D images; for general surfaces of the form z=f(x, y), v_i represents the z-coordinate at (x, y) location

5.4.3. Conversion of Point to Raster Data

Point vector data from the laser scanner is converted into raster data by simply superimposing a regular grid and interpolating the pixel value from nearby points. In Figure 5.42, each pixel's value is derived by interpolating over the points in the neighbourhood of that pixel. There are two considerations in designing the target raster grid, namely, interpolation method and pixel size (also known as raster period or resolution). These are described below.

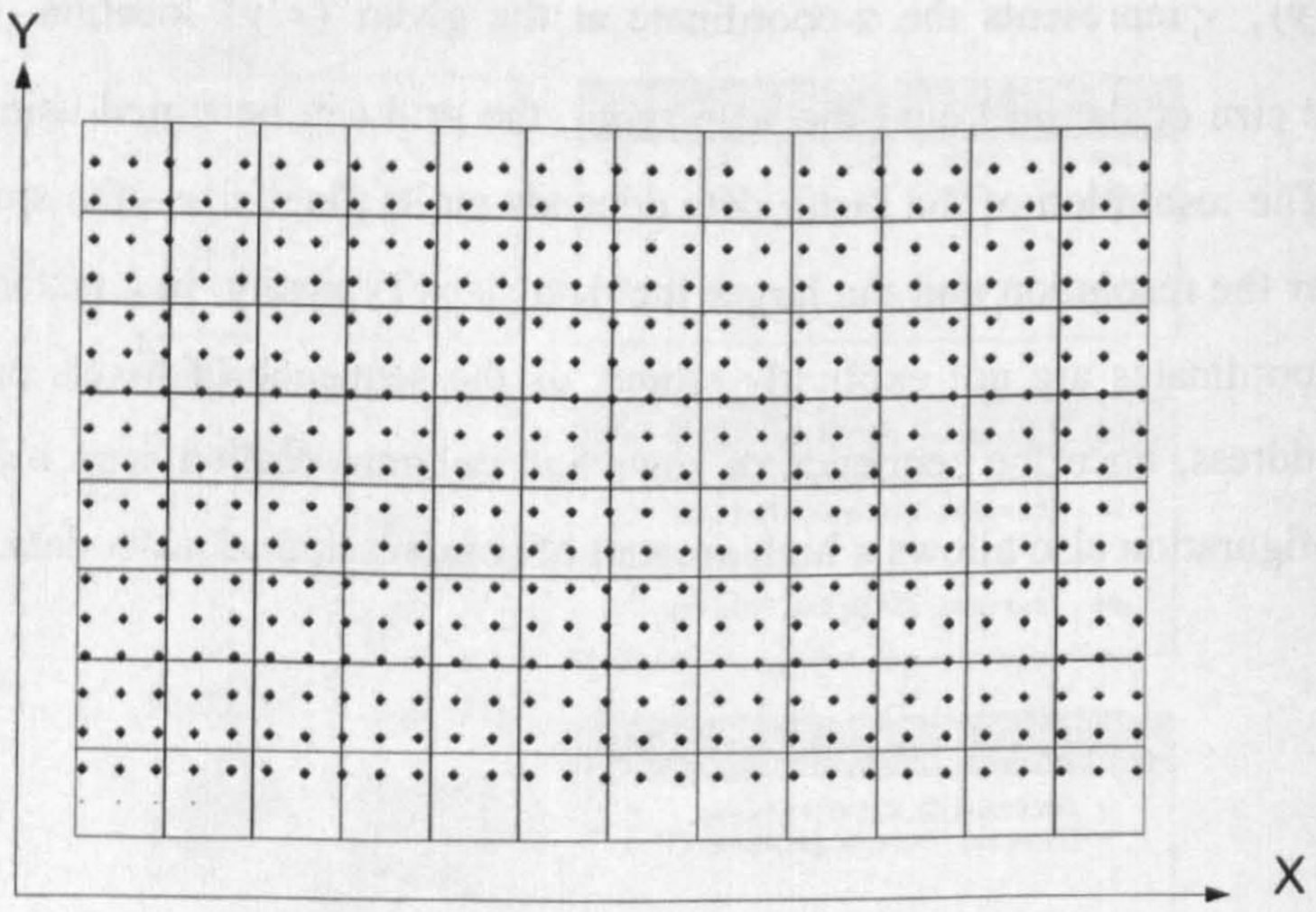


Figure 5.42: Conversion of point data into raster grid

There are various methods for interpolating the value of the target pixel including nearest neighbour, linear, bilinear and bicubic (Burrough, 1998). The nearest neighbour method simply selects the value of the nearest point, ignoring other neighbouring points. Nearest neighbour is thus computationally inexpensive, but not very accurate. Bilinear interpolation computes the average of the four points nearest to the target pixel, weighted by their distances. Bicubic interpolation goes further and uses 16 nearest points, and thus achieves a smoother surface approximation. Thus we use bicubic interpolation in the conversion of point data to raster data.

The pixel size or resolution of the raster was determined based on three considerations, namely, resolution of the original coordinate data, computational costs and the landmarks to be derived. The resolution given by the scanner was less than 1mm, producing over 80,000 pixels per back surface. Processing all these pixels is quite a computationally demanding task, especially since most of the shape analysis operations will have to be repeated for each pixel.

On the other hand, preliminary investigation of some of the landmarks that are to be detected revealed that the landmarks occupied an area of approximately 36 square millimetres or more on the back surface. Thus, using a very small pixel size may not necessarily improve detection of these landmarks. With these considerations, the pixel size of the raster grid was fixed at 2mm, resulting in approximately 40,000 pixels per back surface. This choice seems optimal for our purpose—the surface representation is fairly accurate compared with original resolution and the number of pixels to be processed is reduced by half. With a choice of 2mm pixel size, an area of 36 square millimetres will be partitioned into 9 pixels. This is desirable as most of the shape analysis operations involve solving over-determined systems of equations that require less than 9 equations to yield unique solution. For example, the surface fitting algorithm developed in Section 4.3.7.2 requires at least 6 equations for unique solution while the coordinate transformation model developed in Section 4.3.7.3 requires at least 7 equations, allowing a Least Squares solution to be sought.

5.4.4. System Design

The conceptual model of the system has been designed using in the unified modelling language (UML). The UML is a language for specifying, visualizing, constructing, and documenting the artifacts of software systems, as well as for business modelling (OMG, 1999). The UML was used because of its solid semantics, notation definitions and its wide accep-

tance as a standard model for designing software systems. The UML comprises a number of graphical elements that are combined to form models that provide multiple views of a system. Figure 5.43 shows a simplified UML model of the back shape analysis system. The UML model captures the main components of the system and their relationships. Each class or block in the model represents a specific component in the system that has a set of properties and operations to perform. For clarity, the model has been simplified to present only the most important components, without listing their attributes and properties. The system is being designed as a generic application for back shape analysis and visualization which can be deployed on any standard computer.

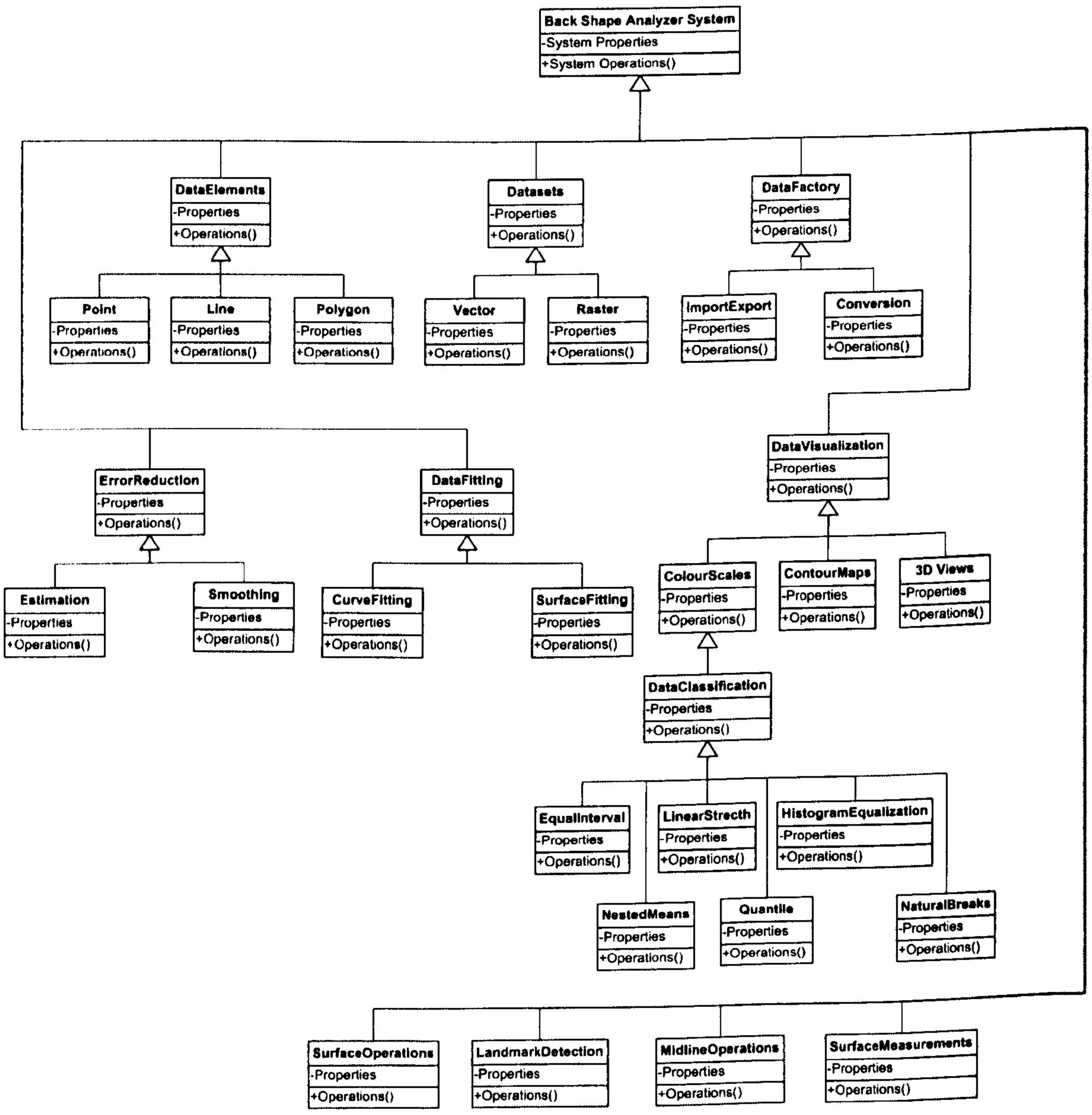


Figure 5.43: UML model of the Back Shape Analyzer system

5.4.5. Choice of Development Tools

It is one of the objectives of this investigation to implement the methods developed practically by developing a software system that will enable automation of back shape analysis and visualization. Such software can be adapted into a practical clinical system. The following factors were considered in the selection of appropriate development tools:

- The development tools should provide high level primitives and data structures so that the methods and algorithms could be quickly implemented.
- Familiarity with the development tools to enable fast learning curve. This allows one to concentrate on the core research issues rather implementation issues.
- Combination of various tools should be used where appropriate to enable rapid implementation of these algorithms.

After consideration of the above factors, the following tools were selected:

- Microsoft Visual Basic: This is as the main environment for developing the software. Visual Basic was designed to be easy to learn and use and allows rapid development of simple and fairly complex applications with graphical user interfaces (GUI).
- Microsoft Visual C++: One major problem with Visual Basic 6 is that it is slow in processing complex operations such as large matrix operations and solving complex systems of equations, as is involved in many image processing tasks and Least Squares solutions. Processing typical back shape data will involve manipulation of over 40,000 points which will cause noticeable performance issues for Visual Basic. A compromise solution was found by developing dynamic link libraries (DLL) in C++ (which is much more efficient) to handle the complex processing tasks. These DLLs were developed using Microsoft Component Object Model (COM) protocol which allowed them to be easily plugged-into Visual Basic.
- ArcGIS Engine: ArcGIS Engine consists of COM-compatible library developed by the ESRI for developing GIS applications. This library provides support for displaying ESRI's Shapefile and raster data structures, as well as set of geometric manipulation functions.

The system has all the functionalities of modern GUI applications, such as zooming, panning and rotating 3D model of the back. Figure 5.44 shows the 2D main interface of the system. More details of the software system and its full functionalities are given n in Appendix A.

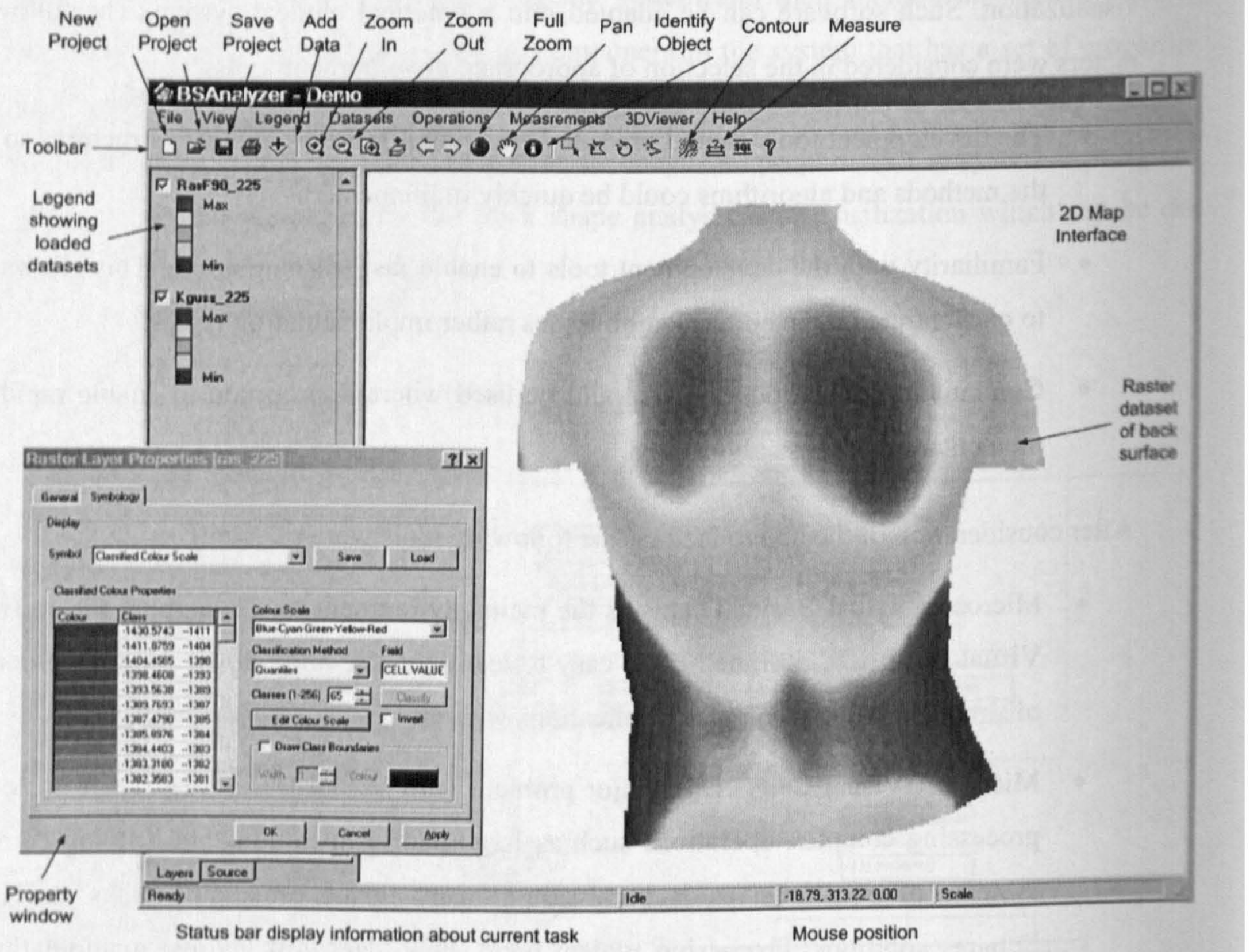


Figure 5.44: Main system interface 2D

5.5. Chapter Summary

This chapter has discussed three aspects of this study, namely, error reduction, curvature visualization and implementation issues. The need for effective error reduction was emphasised and addressed via scanner calibration and experimental analysis to estimate the nature and method for optimal error smoothing. With stable curvature maps computed, various techniques were developed for visualizing the shape of the back, including continuous and discrete colour maps, contour maps and 3D views. The visualizations revealed several anatomic regions of the back which highlighted the shape of the back, which will enable visual exploration of back deformities without the need for physical manipulation of the subject. Lastly, ma-

jor design and implementation issues regarding data structures, development tools and system design were addressed in the last part of the chapter.

6. Landmark Detection and Surface Measurements

6.1. Introduction

In Section 5.3, various techniques were developed for visualizing the shape of the back. The curvature maps revealed several landmarks of the back that were characterised by extreme curvature, including several spinous processes and the PSIS landmarks. The purpose of this chapter is to develop a method for detecting these landmarks, derive the midline and compute surface measurements. The landmark detection algorithm developed here consists of fitting constrained surfaces to the neighbourhood of the landmarks and analysing the surfaces to determine their peaks and troughs. The detected landmarks are then used as the basis for deriving the midline of the back and subsequent computation of the surface measurements discussed in Section 2.5.

6.2. Detection of Back Surface Landmarks

As demonstrated in Section 5.3, the curvature maps reveal several landmarks of the back represented as regions of high positive and negative curvature values. Figure 6.1 shows a 3D view of the Gaussian map with the curvature values scaled to emphasise high curvature regions, which appear visually as hills, valleys, ridges or dimples on the Gaussian map. The peaks or troughs of these regions represent local maxima or minima which can be estimated mathematically. The magnification of the region containing the C7 landmark on the Gaussian map shows that this region has a convex shape whose marked peak represents the C7 landmark. It is obvious that if the local surface containing the C7 landmark is isolated, the landmark represented by the peak point can be evaluated mathematically. Thus, a method to detect the landmarks consists of two main steps, namely:

- Searching for all convex and concave regions in the curvature maps
- Analysing each mathematically regions to estimate their peak or troughs

These two steps are discussed in detail.

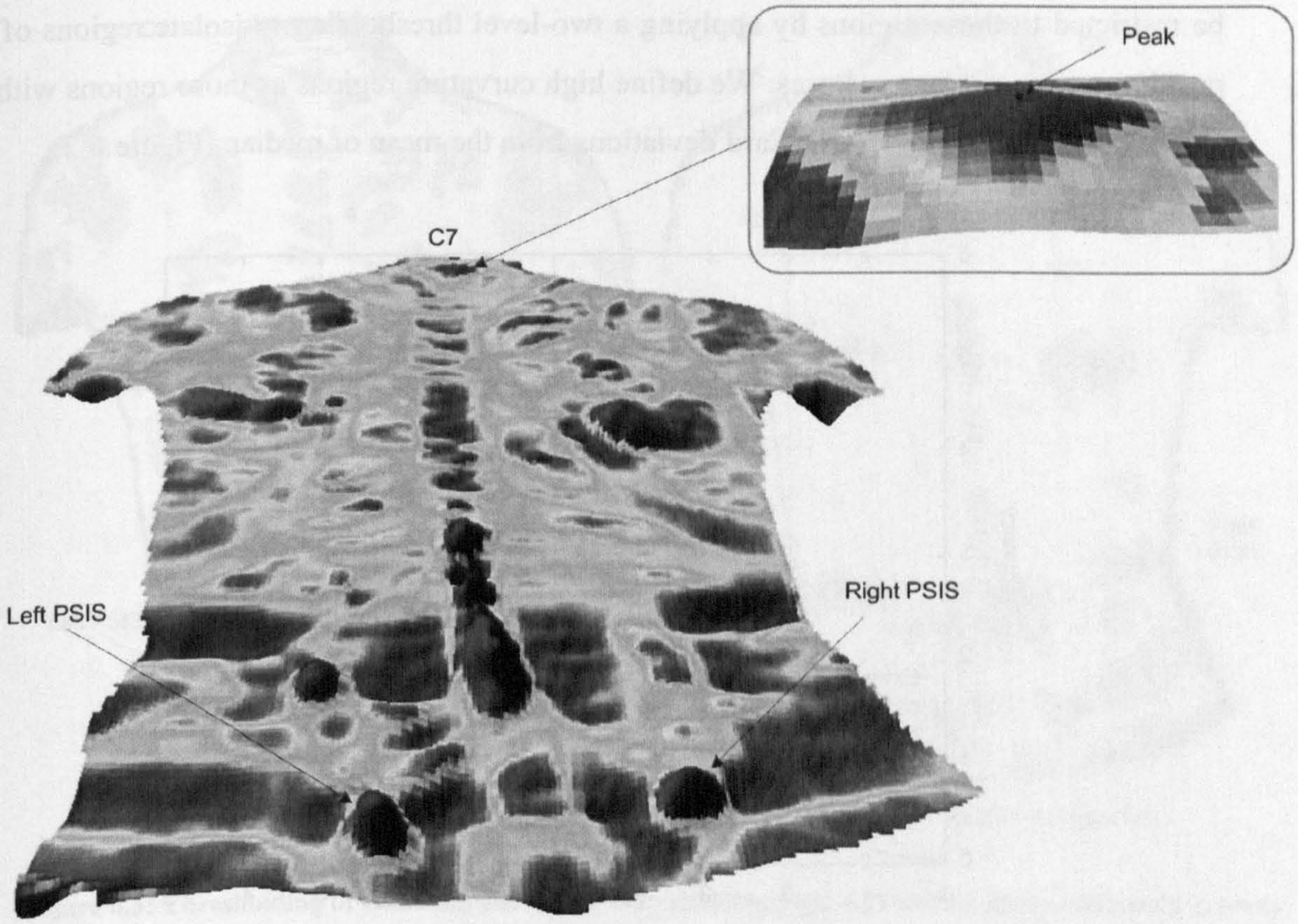


Figure 6.1: 3D View of Gaussian curvature map showing peak of region containing C7 landmark

6.2.1. Searching for Local Convex and Concave Surfaces

There are two basic surface types of interest, namely convex and concave regions. A local surface containing a landmark is defined using 3x3 pixels, which covers an area of 36 square millimetres on the back surface (see Section 5.4.3). To isolate convex and concave regions, the following simple rule is used in the search algorithm:

$$\begin{cases} \kappa_c > \kappa_i & \text{convex region} \\ \kappa_c < \kappa_i & \text{concave region} \end{cases} \tag{6.1}$$

where κ_c is the curvature of the central pixel and κ_i is the curvature of the other pixels in the neighbourhood. This rule simply states that a convex region (or concave) is a region in which the central pixel's curvature is greater than (less than for concave) the curvature of all surrounding pixels in the neighbourhood. This criterion will return all convex and concave regions of the back, irrespective of whether or not they contain landmarks of interest. However, since the anatomic landmarks are usually characterised by higher curvatures, the search can

be restricted to those regions by applying a two-level thresholding to isolate regions of high positive and negative curvatures. We define high curvature regions as those regions with curvature outside user-defined standard deviations from the mean or median (Figure 6.2).

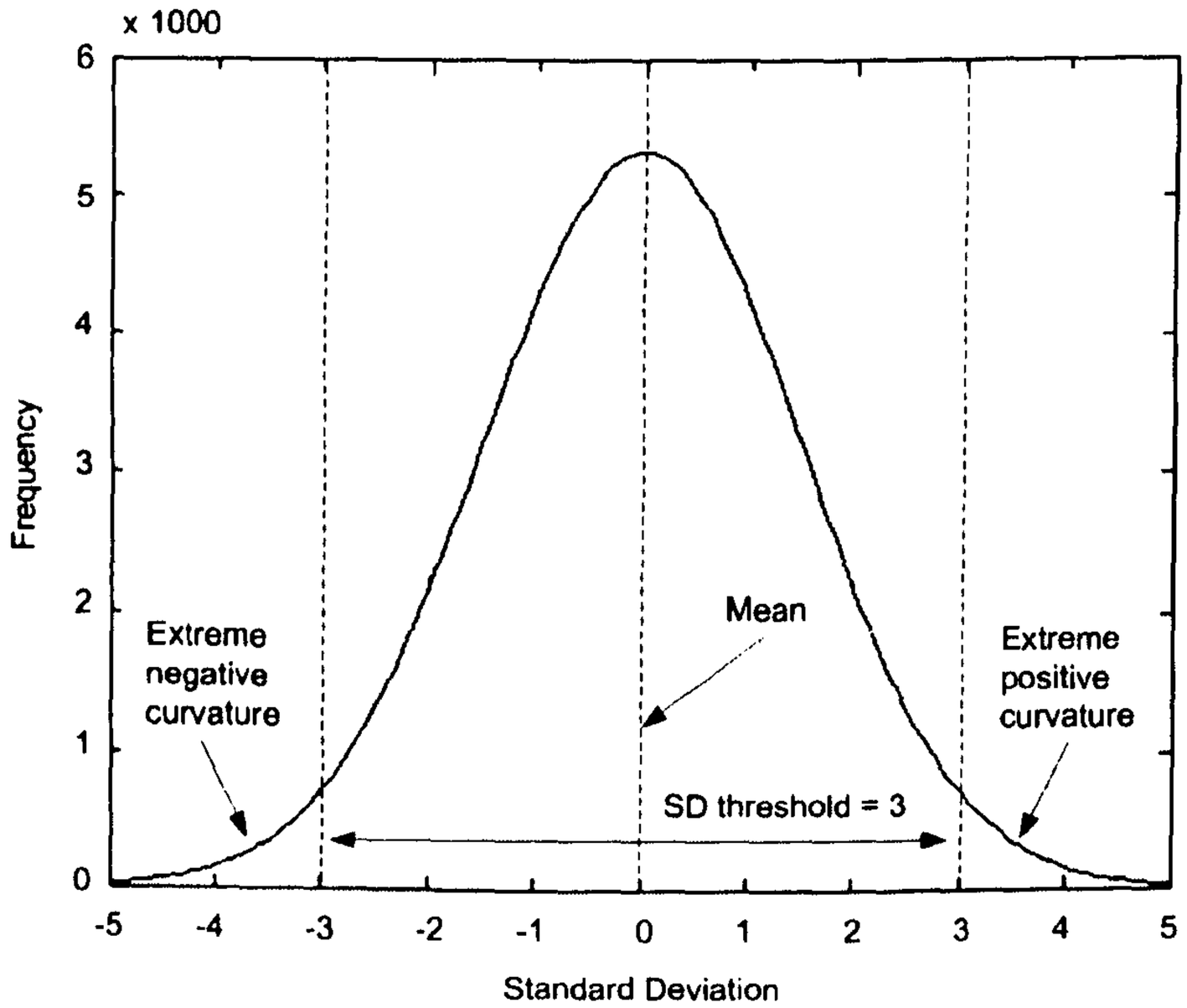


Figure 6.2: Search regions defined by 3 standard deviations from the median

The standard deviation threshold can be adjusted to detect landmarks of specific regions of the back. Figure 6.3 shows application of the two-level thresholding discussed above. The threshold is set to 0.002 standard deviations from the median in (a) and 0.005 standard deviations from the median (b). The median is used instead of the mean because it is less susceptible to outliers. In (a) the threshold is narrow, and thus more regions are detected whiles the larger threshold in (b) detects a few regions.

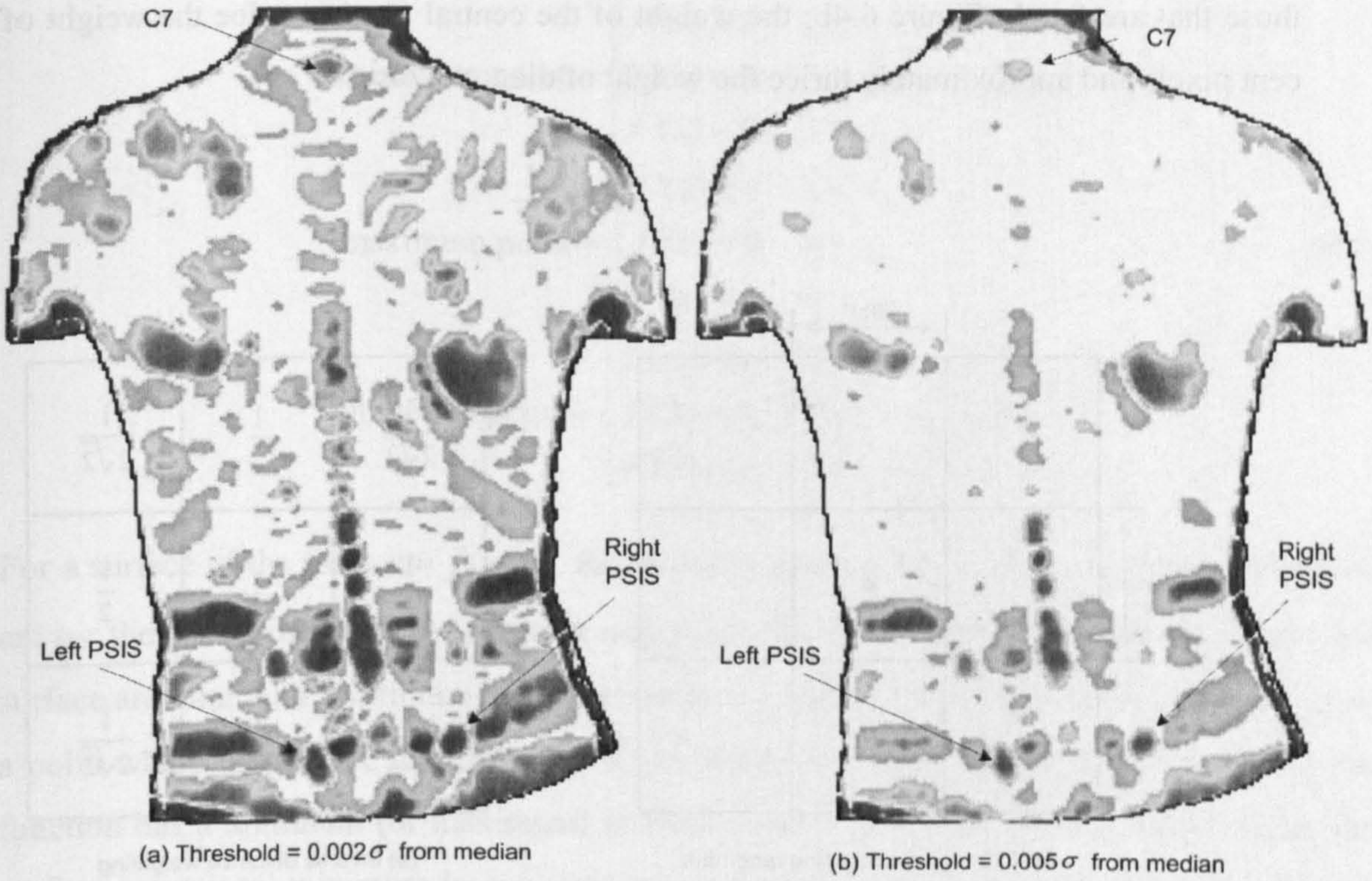


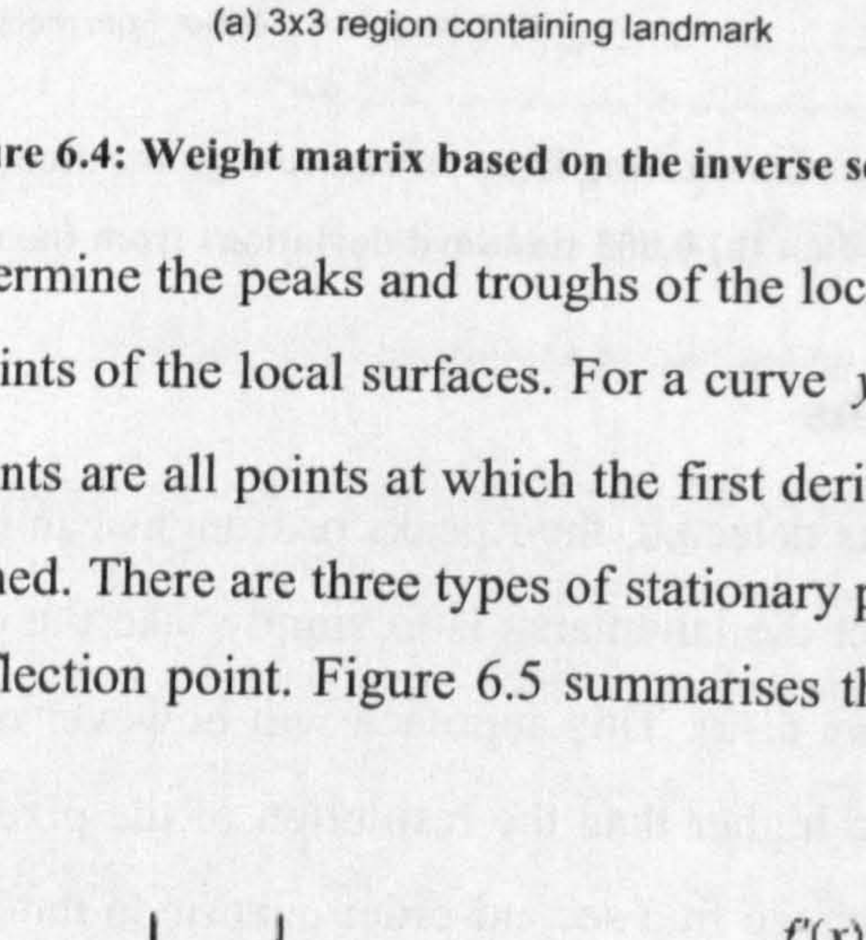
Figure 6.3: Thresholding of Gaussian curvature map showing high curvature regions, threshold is set to (a) 0.002 standard deviations from the median (b) 0.005 standard deviations from the median

6.2.2. Detection of Peaks and Troughs

With the local convex and concave regions detected, their peaks or troughs can be evaluated mathematically. A simple method to detect the landmarks is to simply take the central pixel with the highest curvature value (see Figure 6.4a). This approach will however restrict the X and Y accuracy of the landmarks to be no higher than the resolution of the pixel size of the data (i.e. 2mm). To obtain a better estimate, we fit a second-order quadric to the local convex and concave regions using Equation 4.10 and the Least Squares method. The peaks and troughs of the approximating local surfaces are then estimated by examining the stationary or critical points on the surface. It is also important to properly constrain the surface fitting process so that the resulting surface passes through the centre pixel. This can be achieved by assigning a higher weight to the central pixel in the Least Squares solution. We define a weight function based on the inverse square distance principle: the weight of each pixel (w_p) is inversely proportional to the square of the distance (d) of the pixel to the central pixel (Figure 6.4b). This implies that pixels that are close to the central pixel will have higher weights than

those that are far. In Figure 6.4b, the weight of the central pixel is twice the weight of adjacent pixels and approximately thrice the weight of diagonal pixels.

$$w_p = \frac{1}{d^2} \tag{6.2}$$



(a) 3x3 region containing landmark

$\frac{1}{2\sqrt{2}}$	$\frac{1}{2}$	$\frac{1}{2\sqrt{2}}$
$\frac{1}{2}$	1	$\frac{1}{2}$
$\frac{1}{2\sqrt{2}}$	$\frac{1}{2}$	$\frac{1}{2\sqrt{2}}$

(b) Inverse distance weighting

Figure 6.4: Weight matrix based on the inverse square distance for a 3x3 window, pixel size is 2mm

To determine the peaks and troughs of the local surfaces, we examine the stationary (or critical) points of the local surfaces. For a curve $y = f(x)$ in two-dimensional space, the stationary points are all points at which the first derivative of the function is zero (i.e. $f'(x) = 0$) or undefined. There are three types of stationary points on a curve, namely, minimum, maximum and inflection point. Figure 6.5 summarises the properties of these three types of stationary points.

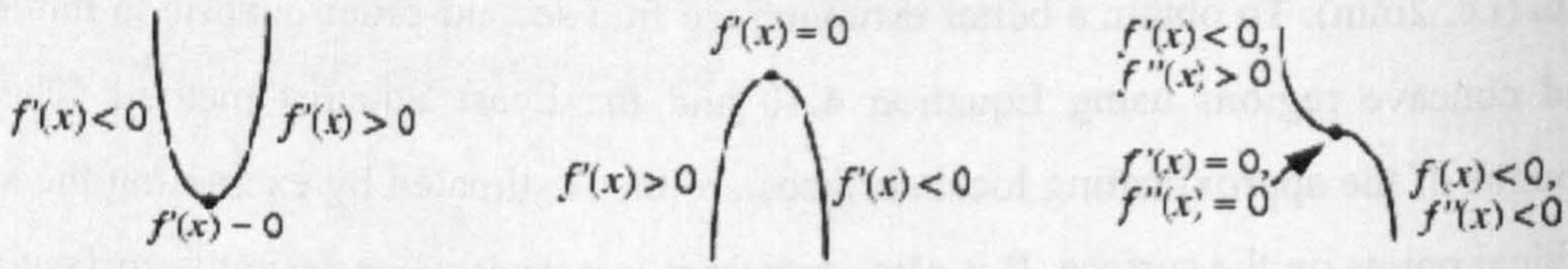


Figure 6.5: Stationary points on a curve (a) inflection point (b) local minimum (c) local maximum

A stationary point x_0 on a curve is a local minimum if $f'(x) < 0$ for all $x < x_0$ and $f'(x) > 0$ for all $x > x_0$. For a local maximum, $f'(x) > 0$ for all $x < x_0$ and $f'(x) < 0$ for all $x > x_0$. An inflection point is a point at which the sign of the curvature of the curve changes. These are summarised in Equation 6.3.

$$\text{minimum point} = \begin{cases} f'(x) < 0 & x < x_0 \\ f'(x) = 0 & x = x_0 \\ f'(x) > 0 & x > x_0 \end{cases}$$

$$\text{maximum point} = \begin{cases} f'(x) < 0 & x < x_0 \\ f'(x) = 0 & x = x_0 \\ f'(x) > 0 & x > x_0 \end{cases} \quad \textbf{(6.3)}$$

$$\text{inflection point} = \begin{cases} f'(x) < 0,\ f''(x) > 0 & x < x_0 \\ f'(x) = 0,\ f''(x) = 0 & x = x_0 \\ f'(x) > 0,\ f''(x) < 0 & x > x_0 \end{cases}$$

For a surface of the form $z = f(x, y)$, the critical points occurs at all points where the gradient (or the partial derivatives) is zero or undefined. The three types of the critical points on a surface are minimum, maximum and saddle points. A minimum (or maximum) point occurs at a point where the surface is locally below (or above) the height of the critical point, i.e. the function has a minimum (or maximum) in both x and y directions. With a saddle point, the surface has a local minimum in one direction and a local maximum in the other (see Figure 6.6).

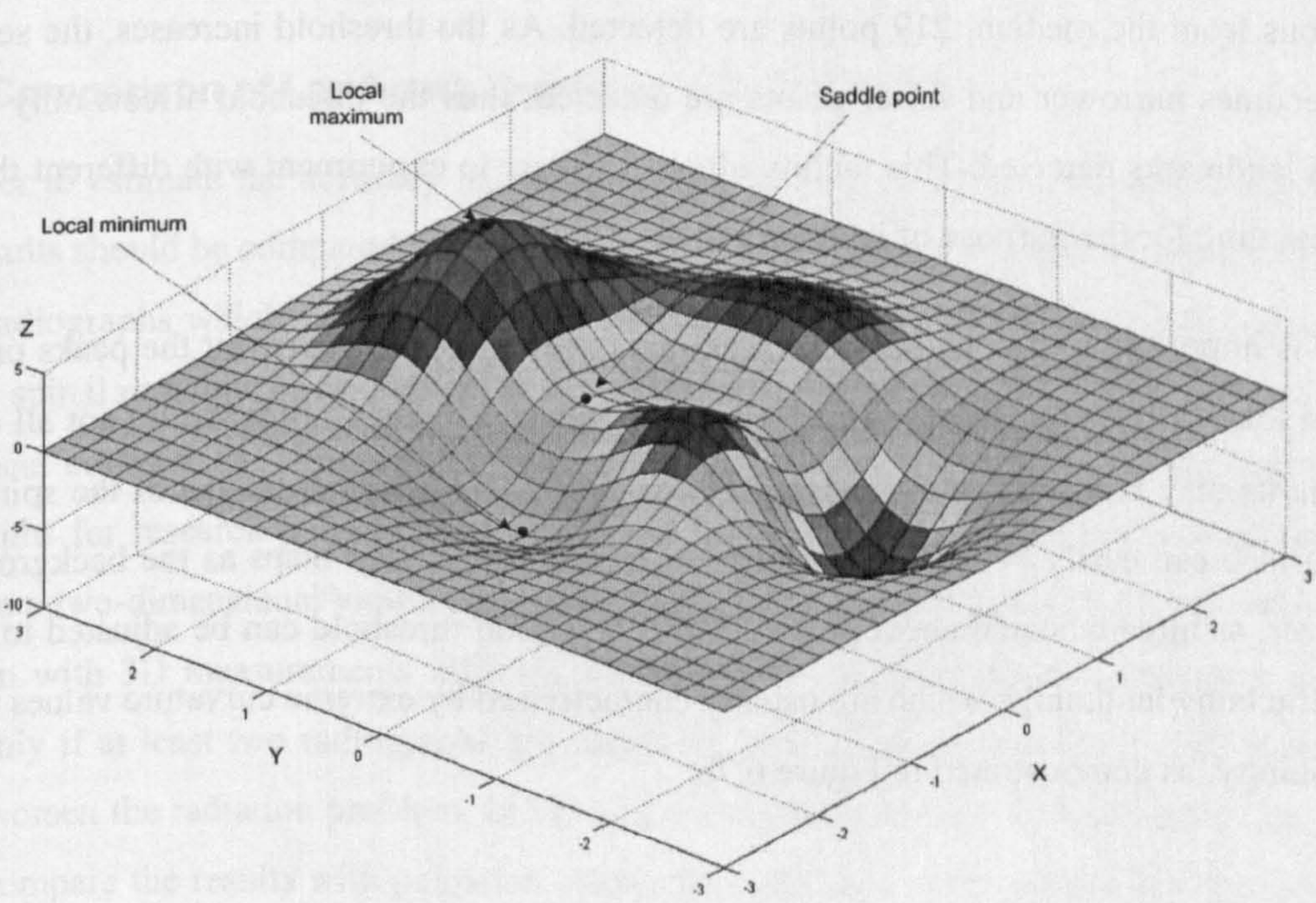


Figure 6.6: The critical points on a surface

Now equating the partial derivatives of Equation 4.10 to zero yields the following:

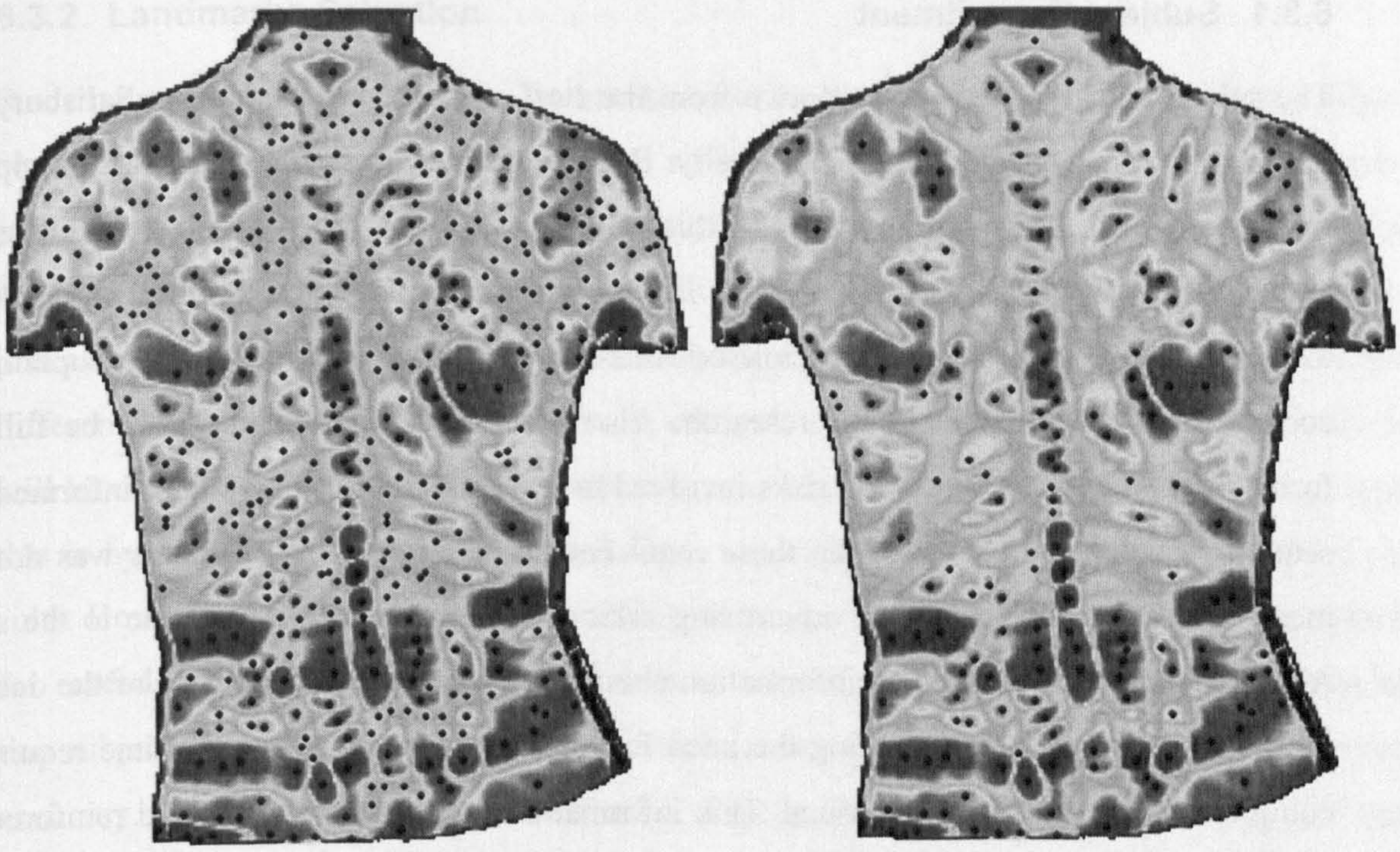
$$x_0 = \frac{2bd - ce}{c^2 - 4ab}, \quad y_0 = \frac{2ae - dc}{c^2 - 4ab} \tag{6.4}$$

The type of point can be determined using the second derivative test. Let the discriminant (D) be defined as $D = F_{xx}F_{yy} - F_{xy}F_{yx}$, then the following hold:

- If $D > 0$, $F_{XX} > 0$ and $F_{XX} + F_{YY} > 0$, the point is a local minimum
- If $D < 0$, $F_{XX} < 0$ and $F_{XX} + F_{YY} < 0$, the point is a local maximum
- If $D < 0$, the point is a saddle point
- If $D = 0$, then further tests needed

Figure 6.7 shows the result of application of the algorithm to detect landmarks from the Gaussian curvature map (Subject 1). Several landmarks are detected, including C7, the left and right PSIS and several spinous process landmarks. The number of points detected depends on the standard deviation threshold: high threshold will detect fewer landmarks and vice versa. With no threshold, a total of 524 points are detected and at threshold of 0.002 standard deviations from the median, 219 points are detected. As the threshold increases, the search region becomes narrower and fewer points are detected; thus the threshold affects only the number of landmarks detected. This facility allows the user to experiment with different thresholds in real time for the purpose of landmark identification.

It is important to note the detected landmarks in Figure 6.7 represent the peaks or troughs of all local convex and concave surfaces in the search region. Consequently, not all of the landmarks may represent bony landmarks of interest. The bony landmarks of the spine and back surface can easily be identified and labelled with the colour maps as the background. Moreover, as already stated above, the standard deviation threshold can be adjusted to detect specific bony landmarks which are usually characterised by extreme curvature values within their vicinity, as demonstrated in Figure 6.7b.



(a) No threshold, 524 landmarks detected

(b) Threshold = 0.003σ, 219 landmarks detected

Figure 6.7: Landmark detected from Gaussian curvature map at different standard deviation thresholds

6.3. Comparison of Landmark Detection with Palpation

In order to estimate the accuracy and precision of landmarks detected from curvature maps, the results should be compared with some 'ground truth'. Ideally, comparison should be made with radiographs which establish the locations of the actual vertebrae. This will require full-length spinal radiographs and laser scan data from a number of subjects in order to carry out statistical comparison. However, in many cases, asking people to take full-length spinal radiographs for research purposes is unacceptable or undesirable. Moreover, radiographs only encode a two-dimensional view of the spine with low intensity resolution, which makes comparison with 3D measurements difficult. Radiography can offer three-dimensional information only if at least two radiographs are acquired, as in CT scans. However, this option will only worsen the radiation problem. In view of these above problems, an alternative approach is to compare the results with palpation. Although palpation is error-prone and is expected to be less accurate than landmarks detected from the curvature maps, it can still give an indication of the error limits of landmark detection and demonstrate the usefulness of the developed method. This option was thus chosen for validation purposes.

6.3.1. Subject Recruitment

The subjects for the study were drawn from the staff and research students at Salisbury District Hospital and Bournemouth University. Prior to the data collection process, an approval was obtained from the NHS Research Ethics Committee (REC) (a copy of the ethics acceptance letter is attached in Appendix B-1). The basic principles of research ethics were adhered to during data collection. The principle of voluntary participation requires that people not be coerced into participating in the research. Also prospective participants must be fully informed about the procedures and risks involved in research and must give their informed consent to participate. In view of the these requirements, the recruitment process was done by means of poster and newsletter advertising asking for volunteers to contribute to the study. All volunteers were first given information about the study, intended purpose of the data and the scanning procedures including the need for undressing, positioning, the time required to complete the scan and safety issues. This information was given verbally and reinforced by printed material with questions invited at all stages. A copy of the information provided to participants is attached in Appendix B-2. At the conclusion of the information phase the subjects were asked if they wished to participate and their consent was recorded and signed by each party with the clear proviso that the subject could withdraw if they changed their mind (see Appendix B-3 for a copy of the consent form). To ensure anonymity of the participants, data collected for each person did not include any personal information such as name or address. In addition, all data files were saved under arbitrary names such as 'Subject 1'.

The data collection was carried out in three rounds due to difficulty in recruiting volunteer subjects: in the first round, 12 volunteers were scanned, in the second round, 8 volunteers were scanned and 5 were scanned in the last round. In all, 25 'normal' subjects participated, consisting of 23 males and 2 females, all volunteers. Table 6.1 summarises the relevant information about the subjects.

Table 6.1: Information on subjects used for evaluation of landmark detection

Statistic	Age	Height (cm)	Weight (kg)	BMI (kg/m^2)
Min	24	167	58	21
Max	48	202	107	28
Range	24	35	49	8
Mean	28	174	74	24
Median	27	172	72	24
St Deviation	5	7	9	2

6.3.2. Landmarks Selection

Since palpation is error-prone, the landmarks used for the comparison should be easily palpable with improved precision. The precision with which landmarks can be palpated on a subject depends on factors such as the body mass index (BMI) of the subject and experience of the one performing the palpation. In other words, palpation, to a larger extent, is a subjective process. Consequently, it is it difficult to compare results from different subjects or even for different landmarks of the same person. In order to compare results of different subjects, it will be useful to be able to quantify palpation precision for the subjects. One way to achieve this is by palpating each subject several times and calculating standard deviation of the results. However, this is unacceptable or undesirable as it would require subjects to be retained for extended time. To simplify these problems, a brief preliminary physical examination of one subject was conducted to determine which landmarks could be palpated with improved accuracy. From this preliminary examination, three landmarks were selected, namely, C7 and the left and right PSIS landmarks. The C7 is the most prominent spinous process in a large proportion of the population (≈ 60-70%, Cramer & Darby, 2005) while the PSIS landmarks are generally visible as a pair of dimples. S1 could not be included in the study due to the fact that some of the subjects were uncomfortable with palpating the S1 landmark. C7 has as a convex shape while the PSIS dimples have concave shapes, thus at least both shape types were represented in the study. For the subjects examined, the C7 was found to be the most prominent landmark.

6.3.3. Palpation and Data Acquisition

The three selected landmarks (C7 and the PSISs) were palpated on all the subjects and marked with white circular dots by Dr Steve Crook who is involved in routine assessment of back deformities at Salisbury District Hospital. The shape and size of the markers were determined after several experiments with various maker shapes and sizes. Circular markers were chosen because it is easier to extract the location of their centres with improved accuracy. The diameter of the dots used was 5mm—optimal diameter that could be accurately detected from the scanned data following preliminary investigations. The markers were glued onto the back surface by means of a medically approved gel (TAC GEL). To permit comparison of the markers and detected landmarks, four additional registration markers (denoted as G1, G2, G3, and G4) were also placed on selected landmarks for registration purpose, as shown in Figure 6.8. The transformation model developed in Section 4.3.7.3 requires at least

three common points with good geometric configuration for improved accuracy. Thus, the registration markers were chosen to be global to the back surface, with rectangular configuration.

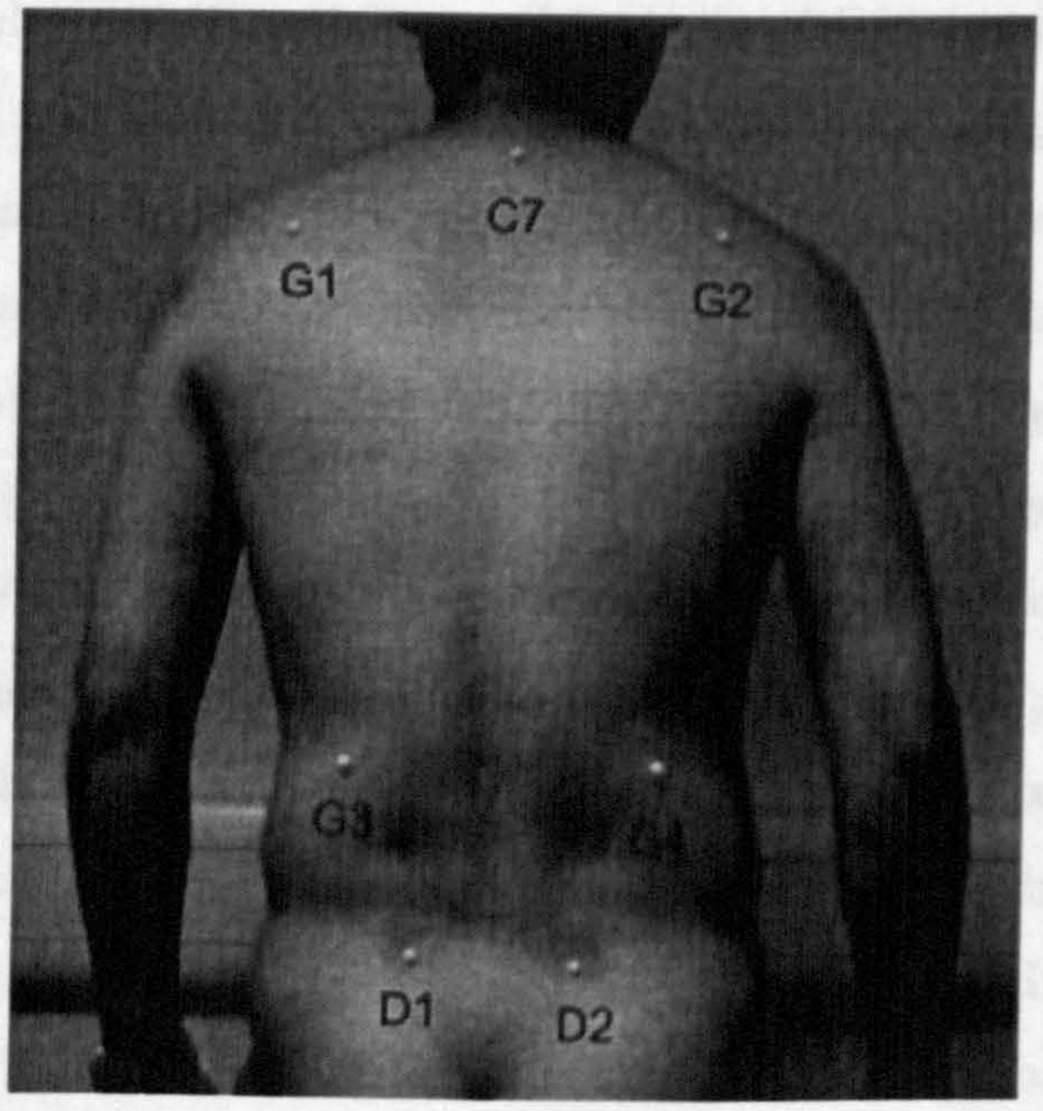


Figure 6.8: Landmarks used in experiment: C7, left and right PSIS; G1-G4 are registration markers

After subject preparation, two images were acquired for each subject, one with both the test and the registration markers and the other with only the registration markers. The first image was captured within 2.5 seconds after preparation and marker placement. The three test markers were then carefully removed without disturbing the registration markers and the second image was taken. This allowed accurate registration and comparison of the two images to be carried out.

6.3.4. Data Analysis

In order to make the comparison, the locations of the makers and the registration points had to be extracted from the datasets. Due to random errors, it was difficult to accurately extract the locations of the makers from the raw datasets so both datasets were first smoothed with a Gaussian filter at $\sigma = 2.25$. The X and Y coordinates of the centres of the markers were then extracted from the first image and the coordinates of the registration markers (G1-G4) were also extracted from both images. For each subject, the mean, Gaussian and curvedness curvature maps were computed, and the landmarks were detected using the algorithm developed in Section 6.2.

To compare the landmarks detected from the curvature maps with the marker positions, it was necessarily to transform them into the same coordinate system using the registration points. For each subject, the two corresponding datasets were registered by transforming the second image into the coordinate system of the first image using the 7-parameter transformation model discussed in Section 4.3.7.3. The transformation was accomplished with a mean RMSE of 1.33×10^{-8} for all 25 subjects, which was very accurate. The details of the transformation process for each subject are given in Appendix C.

The deviations between the marker positions and the transformed curvature landmarks were then computed by subtracting the later from the former. The results of this experiment are shown in Table 6.2 which gives the root mean squares (RMS) of the deviations between X and Y coordinates of the markers and their corresponding curvature positions. The RMS statistic is calculated by first squaring the deviations, finding the mean of the squared deviations and taking the square root. For a random variable that has both positive and negative values, the RMS statistic (known as quadratic mean) is a useful statistics than the arithmetic mean.

Table 6.2: Comparison of marker positions with curvature positions in the frontal plane, showing the RMS of deviations in X and Y (in millimetres)

	Gaussian		Mean		Curvedness	
	ΔX	ΔY	ΔX	ΔY	ΔX	ΔY
C7	0.8371	0.8823	0.8203	0.8802	0.8305	0.9810
Left PSIS	1.3252	1.3991	1.3328	1.3957	1.7434	1.8424
Right PSIS	0.5185	0.7196	0.7531	1.0922	1.7651	1.1576

6.3.5. Discussion of the Results

The RMS values given in Table 6.2 represent the average deviations between the markers and curvature positions, zero being a perfect match. Thus, an RMS value close to 0 indicates a better match. If one assumes the markers as the 'ground truth', then the RMS values give the accuracy of landmark detection compared to palpation.

These results indicate that both X and Y accuracies were slightly higher for C7 than the PSIS landmarks. One possible explanation for this is that usually C7 has a large spinous process which produces relatively prominent peak on the back and thus could be seen and palpated with improved accuracy. The PSIS landmarks appear as a pair of dimples which have relatively shallow depression, making palpation less accurate. Accuracy in X was also generally higher than accuracy in Y for all landmarks and curvature maps. Perhaps, this could be partly

attributed to factors such as a slight flexion of the neck which can induce skin movement in Y-axis more than in X. Breathing can also cause unequal skin movement in X-axis and Y-axis. Although patients were instructed to stand still and hold their breadth during data capture, it is impossible to perfectly control all the factors involved.

In comparing the accuracy obtained from the curvature maps, the mean RMS was for all three landmarks was 0.89 in X, 1.00 in Y for the Gaussian curvature map; 0.97 in X, 1.12 in Y for the mean curvature; and 1.45 in X, 1.33 in Y for the curvedness map. Thus, the accuracies of the curvature maps are comparable although the Gaussian curvature gave slightly more accurate match than the mean and curvedness maps. Moreover, there were slight but noticeable differences between RPSIS and LPSIS on the mean and Gaussian curvature maps compared to the curvedness map. However, no geometric explanation could be found for this, and more experiments with a larger sample size is required to make any conclusive inference. Combining the results of all curvature maps, C7 landmark was detected with accuracy 0.83mm in X, 0.91mm in Y; LPSIS with accuracy of 1.47 in X, 1.55 in Y and RPSIS with accuracy of 1.01 in X, 0.99 in Y. Combining the RMS values for all three landmarks and curvature maps, the overall mean accuracy for landmark detection was approximately 1.10mm in X and 1.15mm in Y.

These results were rather surprising given the reliability with which landmarks can be palpated. Preliminary investigation with one subject showed that palpation of the three landmarks (C7, LPSIS and RPSIS) was expected to be performed within a precision of approximately 4 millimetres. However, it is important to note that the reported accuracy of landmark detection depends on the accuracy with which the two datasets were registered. The registration model developed in Section 4.3.7.3 uses a scale factor to correct for unequal scaling in the coordinate systems of the two datasets. And since posture changes appear to be the major sources of random errors during data capture (see Section 6.4), accounting for scale changes is expected to improve registration. Moreover, the markers were extracted after smoothing operations which reduce random errors by Gaussian averaging of neighbouring pixels. Thus, the locations of the makers represent their average locations after smoothing, which could have contributed to the higher accuracy.

The main limitation with this experiment is that the sample size was rather small (i.e. less than 30 subjects), with badly balanced gender ratio (only 2 females out of 25). Moreover, the

subjects did not include adolescents (minimum age 24) in whom scoliosis has higher incidence rates. Due to the nature of the experiment, it was very difficult recruiting volunteers, particularly female subjects. The experiment required subjects to undress their backs for the scan and it was necessarily to accurately palpate the landmarks and mark them with the white dots. All these procedures took approximately 3 minutes to complete. Although, a gown was provided to cover the front part of the body, even some of the male subjects felt uncomfortable. This limited the experiment. Of the 25 subjects examined, 76% were within what is generally considered as the normal weight range (BMI 21-25) while the other 24% were overweight (BMI range 25-30). No significant differences in accuracy were observed between the two BMI groups, between the genders or between the three separate data collections on different days and locations. However, this could be due to the smaller sample size, and further studies involving more subjects may indicate otherwise. Ignoring the sample size, the result does indicate that the method could work well with subjects classed as normal weight and overweight.

6.4. Precison of Landmark Detection

There are two important aspects of the quality of every measurements process, namely, accuracy and precision. Accuracy refers to the closeness of a measured quantity to a known standard (assumed or true value). Precision on the other hand describes consistency among measurements, i.e. it is a measure of the closeness of a set of repeated measurements to each other and is usually measured by the standard deviation. In the previous section, the accuracy of landmark detection was assessed by comparison with palpation (which was used as benchmark). To determine the precision of our landmark detection algorithm, one subject was scanned twenty times (Subject 1: age 30, weight 70kg, height 170cm, BMI 24). The subject moved freely between scans to ensure some variability in posture. The twenty datasets were smoothed with a Gaussian filter ($\sigma = 2.25$), and the Gaussian curvature map was computed and the landmarks were detected according. Twenty landmarks consisting of several spinous processes including C7, S1 and the left and right PSIS landmarks were then selected. Then using the transformation model developed in Section 4.3.7.3, all the datasets were transformed into the coordinate system of the first dataset and the coordinates of the selected landmarks were compared. Table 6.3 gives the standard deviations of the transformed coordinates.

Table 6.3: Standard deviations of 20 landmarks detected from 20 Gaussian maps of Subject 1

Point	Std ΔX	Std ΔY
1(C7)	0.2024	0.2433
2	0.7285	0.8223
3	1.3401	1.2110
4	0.5820	1.2102
5	1.4166	0.8539
6	0.6019	0.4336
7	1.5527	1.3728
8	0.8391	1.2000
9	0.2035	0.6813
10	0.5616	1.9298
11	0.2032	0.4289
12	0.1241	0.2766
13	0.1939	0.2920
14	0.2131	0.3983
15	0.2790	0.6096
16	0.2713	0.7156
17	0.3763	0.7900
18 (S1)	0.7536	0.5147
19(LPSIS)	0.7540	1.3332
20(RPSIS)	0.3988	0.5304
Mean	0.5798	0.7924

The mean of the standard deviations in Table 6.3 was calculated to be 0.58mm in X and 0.79mm in Y. The highest precision was obtained for C7 i.e. 0.20mm and 0.24mm in X and Y respectively. Compared to these results, the accuracy obtained via comparison with palpation was quite high. This could be attributed to the fact that during the palpation experiment, the subjects remained in the same posture as much as possible between the two scans in order to maximise the registration accuracy. In the case of the precision environment, the subject moved freely between the twenty scans, introducing some variability in posture. Thus, it appears that posture changes during data capture are the major source of random errors.

6.5. Comparison with the Work of Drerup and Hierholzer

The closest work relating to landmark detection from curvature maps of the back is that of Drerup and Hierholzer (1985, 1987). In Drerup and Hierholzer (1985), the authors used a different method to detect the C7 landmark using the Gaussian curvature map computed from surface reconstruction of moiré topographs. With their method, the authors first estimate the C7 landmark by the pixel with the highest Gaussian curvature and improve this estimate by averaging all points within a defined neighbourhood of C7. The averaging is performed by calculating the centroid of all points with Gaussian curvature above a defined threshold, weighted by the Gaussian curvature. The authors compared their results with palpation involving twenty subjects and reported the mean and standard deviation of the differences between radiographic and surface measurements as X (1+2.5mm) and Y (1+6mm).

In order to compare their results with that obtained in this investigation, the root mean square deviation should be deduced. The arithmetic mean is not a good statistic for a random variable with positive and negative values, since they will cancel out yielding a value that is generally close to zero, which doesn't represent a perfect match. The RMS can be calculated from the mean and standard deviation as follows:

$$x_{rms} = \sqrt{\bar{x}^2 + \sigma_x^2} = \sqrt{1^2 + 2.5^2} = 2.69\text{mm}$$
$$y_{rms} = \sqrt{\bar{y}^2 + \sigma_y^2} = \sqrt{1^2 + 6.0^2} = 6.08\text{mm}$$

In another paper (Drerup & Hierholzer, 1987), the authors discuss an improved method for back shape analysis based on asymmetry of the horizontal profiles produced from raster-stereography (see Section 3.3.3.2). From shape analysis of these profiles, the so-called symmetry line is derived, which approximates the back midline. In order to determine the symmetry line, a function of lateral asymmetry is defined in terms of the local curvature at any point on the back surface. Consider a horizontal profile passing through a point P on the back surface such that P divides the profile into left and right parts. For each profile, a measure of asymmetry is calculated by comparing the curvature of points lying symmetrically to the left and right of the symmetry point P using Equation 6.5.

$$a = (H_L - H_R) + \left(G_1^2 - 2G_L^2 G_R^2 \cos 2\varepsilon + G_R^2\right)/2 \qquad \textbf{(6.5)}$$

where H_L and H_R are the mean curvatures of corresponding left and right points; G is the difference between the principal curvatures and ε is the angle difference between the corresponding principal directions.

The mean asymmetry for the entire profile is then calculated as in Equation 6.6.

$$A = \frac{1}{s}\int a dx \tag{6.6}$$

where s is the length of the integral which extends from P over the left part of the horizontal profile (excluding those parts not having a corresponding parts on the right). Figure 6.9 shows a typical example of the asymmetry function computed from the profiles of the back surface and its minima.

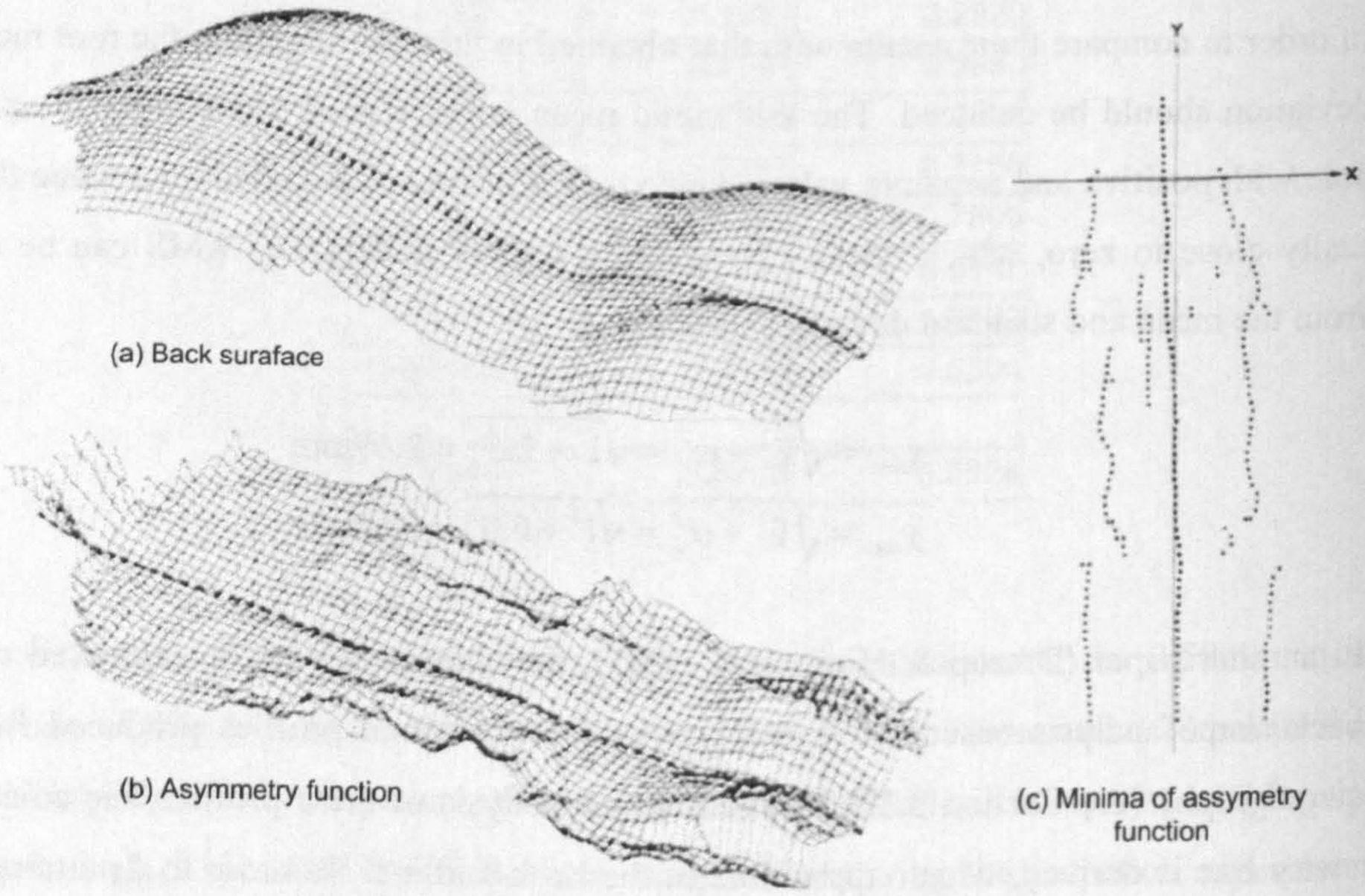


Figure 6.9: Typical back surface, the symmetry function and its minima (Drerup & Hierholzer, 1987)

The minima of the asymmetry profiles are determined and the symmetry line is constructed by searching for a sequence of minima using a set of geometric constraints. Figure 6.9 (c) shows derived minima of the asymmetry function. Using the derived symmetry line as the basis, four anatomic landmarks (C7, LPSIS, RPSIS and S1) were then estimated from the asymmetry function and compared with their corresponding radiographic positions. The accuracy of landmark detection was determined using the mean and standard deviation of length

differences between radiographic and surface measurements, resulting in X (LPSIS—RPSIS): -1.5±6.0 mm; Y (C7—Mid-point of PSIS) 2.3±9.9 mm. The precision of the landmark localization from the asymmetry function was determined by repetitive measurements of the same subjects as 1.1mm which is quite high. The accuracy results represent rather lower accuracy than that obtained in the previous study by the same authors. Perhaps, this could be attributed to the distances not being measured perfectly in X and Y axis, resulting in diagonal distances. In most cases, the line connecting the PSIS landmarks will not perfectly align with the X-axis of the global coordinate system. Consequently, without correction, the derived statistic may represent a diagonal distance which may contribute to the lower accuracy obtained. Moreover, no attempt was made to correct for unequal scaling in the two coordinate systems.

6.5.1. Concluding Remarks

In comparing the landmark detection algorithm of Drerup & Hierholzer (1985, 1987) with the algorithm developed in this investigation, the following concluding remarks can be made.

Although a 3D surface can be constructed from the horizontal profiles of rastestereographic data, the shape analysis method used by Drerup and Hierholzer (1987) still is tied to the nature of the underlying profiles, leading to a more complicated and computationally expensive analysis. The authors derive a function that expresses the idea of asymmetry of the corresponding points along the horizontal profiles and aggregate the results over the entire surface. The method developed here exploits the curvature maps directly to detect the landmarks, and thus is simple and less computationally demanding. Moreover, our method detects landmarks for the entire surface, rather than only for certain regions of the back as the asymmetry function of Drerup & Hierholzer (1987) reveal in Figure 6.9(c). This implies that not all detected points may be landmarks of interest in the case of our method. However, the standard deviation threshold can be altered to detect landmarks with most extreme curvature, which are typically the bony landmarks. Due to the simplicity of this approach, the operation can be performed on-the-fly, with different threshold values.

The improved accuracy reported in this investigation can be attributed to several factors, including the following:

- Effective error smoothing. As discussed in Section 5.2, error smoothing is a critical aspect of curvature computation since the results are susceptible to random errors in

the original data. And since all measurements contain random errors, an effective method must be found to reduce these errors. To determine optimal method for our data, a number of experiments were conducted to estimate error distribution and to estimate the amount of smoothing, leading to more stable results.

- Application of a rigorous registration algorithm. To compare surface data taken at different times, accurate registration is crucial. In the case of the human back, registration is further complicated by the flexible nature of the human body, which can cause unequal scaling. This implies that distances measured in one coordinate system will not necessarily be equal to corresponding distances in the other system. The registration algorithm developed in Section 4.3.7.3 attempts to correct this by incorporating a scaling factor (uniform in all three axes), thereby improving the accuracy of the registration. The algorithm can further be improved by incorporating non-uniform scaling, but this will introduce more parameters into the model.

- Evidently, the higher accuracy and resolution of the original back shape data will have contributed to the overall accuracy of the results. The resolution of our raster data was 2mm compared with 7.5mm resolution used by Drerup & Hierholzer (1987).

- Our landmark detection combines results from three curvature maps, namely, mean, Gaussian and curvedness maps. There are various measures of surface curvature and while each can be useful, combining the results from various curvature maps can significantly improve the accuracy of landmark detection.

- The algorithm developed here is simple, straight forward, is based on direct mathematical analysis of the curvature maps and application of a weight matrix to constrain the least square solution, thereby leading to improved estimates.

6.6. Derivation of the Back Midline

The back surface midline is defined by the projection of the spinous processes onto the back surface, and is thus the closest surface approximation of the vertebra body line. Accurate derivation of the midline therefore depends on accurate detection of the spinous process landmarks. In this study, we derive the midline by fitting a smooth curve through the detected spinous process landmarks via the usual method of constrained Least Squares. There are three things that require consideration:

- Selection of type of curve for approximating the midline
- Selection of spinous process landmarks to be used for derivation of the midline
- Formulation of the weight function for constraining the fitting process

6.6.1. Selection of Type of Fitting Curve

There are a number of curves that may be used to approximate the midline including polynomials and trigonometric functions. Since most common types of frontal spinal curves encountered are either a single, double or triple curves, it is proposed that a polynomial up to the fourth degree will be adequate to describe the midline. Figure 6.10 shows examples of polynomial functions of varying degrees—degree 1 corresponds to a straight line, degree 2 to a single curve, degree 3 to a double-curve and degree 4 to a triple curve. The required degree of polynomial to be fitted can be selected based on the general shape of the midline which can be obtained from the colour visualizations. The curve fitting process is first performed in the frontal plane to derive the 2D midline. To obtain a 3D curve, the resulting 2D curve is superimposed on the smooth version of the original surface, and the Z-coordinate is interpolated from this surface.

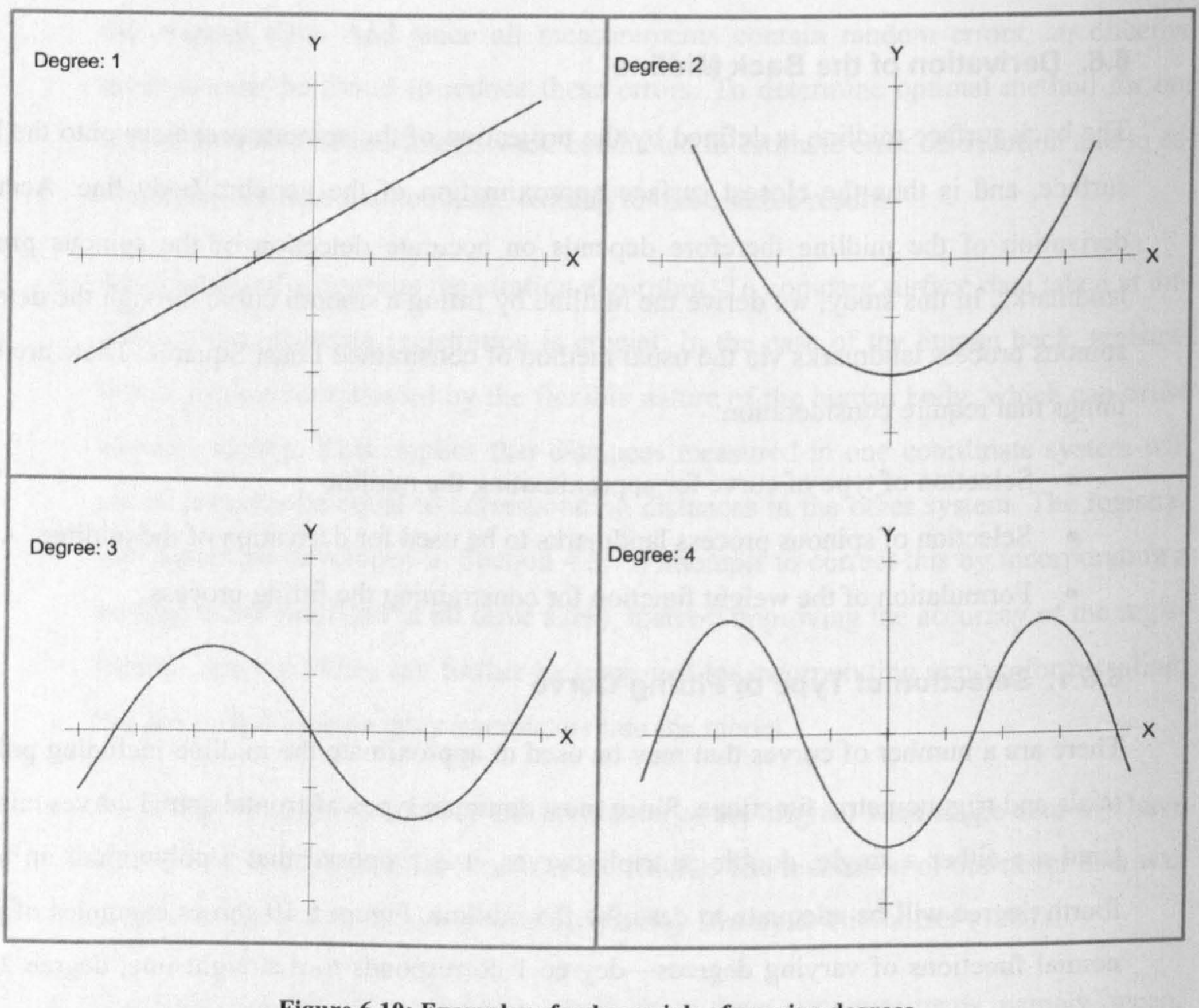


Figure 6.10: Examples of polynomials of varying degrees

6.6.2. Landmark Selection

To derive the midline, a number of spinous processes should be selected from the landmarks detected in Section 6.2. Since it is computationally expensive to select the landmarks automatically by searching, the current implementation relies on the human operator to select the spinous processes using the mouse. These landmarks can easily be identified using the colour visualizations of the curvature maps as the background. The operator selects the landmarks by drawing an irregular polygon using the mouse around the points (Figure 6.11). To allow the curve fitting to be properly constrained, the selection should include the C7 and S1 end points, which are usually detected. The PSIS landmarks are also required to compute the general alignment of the spine to the pelvis.

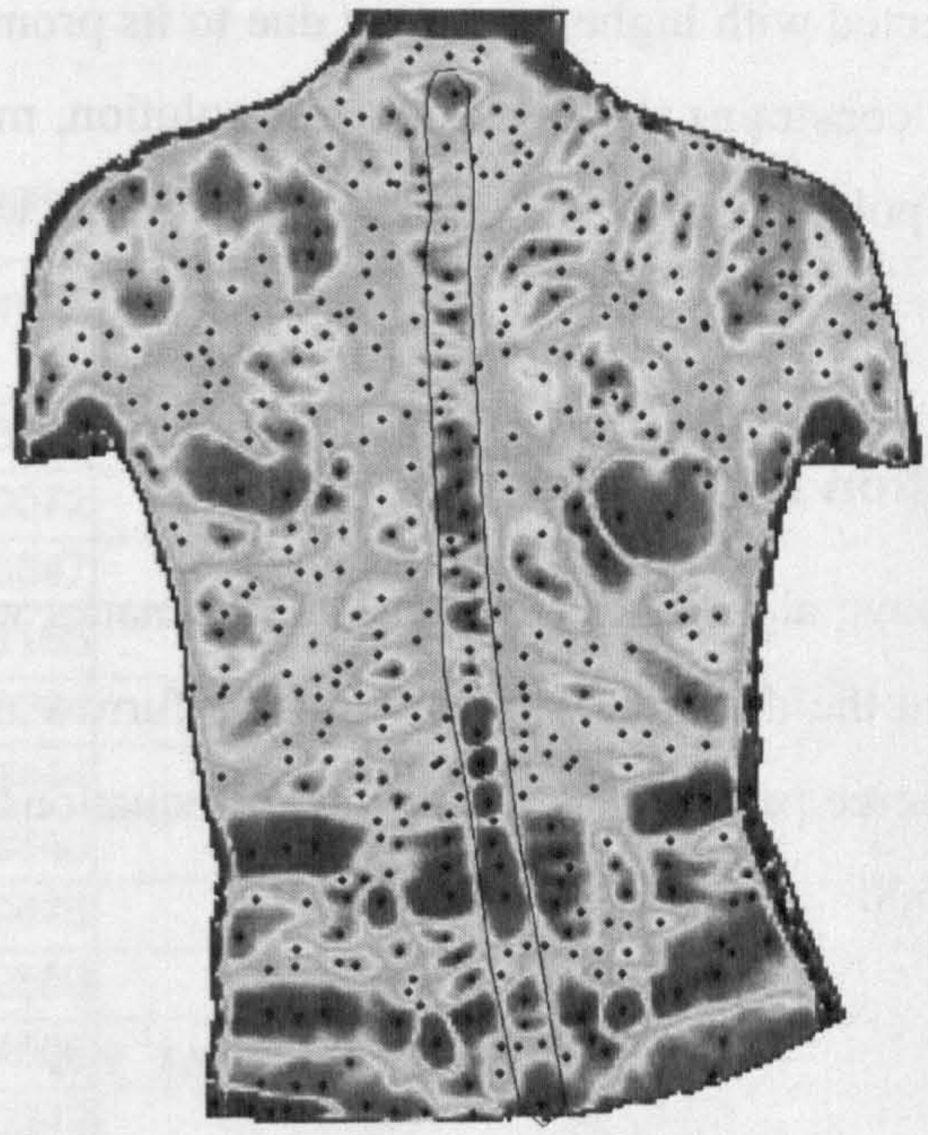


Figure 6.11: Landmark selection for midline derivation, polygon encloses selected landmarks

6.6.3. Weight Function

Since the Least Squares solution can be influenced by extreme or erroneous points that deviate considerably from the true curve, it is important to properly constrain the solution to ensure robust results. Two factors affect the behaviour of the fitted curve, namely, the degree of the polynomial and the weight function. The curve degree determines the nature of the resulting curve within the configuration of the selected points. The degree of the curve is determined using the general shape of the midline but can be changed by the user. The weight function is used to constrain the Least Squares solution so that the resulting midline passes through the C7 and S1 spinous processes which mark the spinal region of interest. This is achieved by assigning higher weights to these two end points and lower weights to all other participating points. We formulate the following weight function based on these considerations:

$$w_i = \begin{cases} 0.45 & \text{for C7 and S1} \\ 0.1/n & \text{for all other points} \end{cases} \tag{6.7}$$

The function in Equation 6.7 assigns much higher weights to C7 and S1 (90% of the weight) and the lower weights (remaining 10%) to the rest of the points. The motivation for assigning higher weight to C7 and S1 is that they mark the end points of the spine and, and that C7 in

particular can be detected with higher accuracy due to its prominent projection to the surface. This weight function constrains the Least Squares solution, making it less susceptible to the influence of extreme points, even if the selection polygon is widened to include points which are not on the spine.

6.6.4. Implementation and Results

To implement the above algorithms, a set of 30 landmarks were selected by drawing a polygonal region around the detected landmarks in the furrow of the back (as in Figure 6.11). Then using a third degree polynomial, the system of equations is developed using the general relation in Equation 6.8:

$$y = f(x) = a_3 x^3 + a_2 x^2 + a_1 x^1 + a_0 \tag{6.8}$$

With a total of 30 points and 4 coefficients to be estimated, there are 36 degrees of freedom. One equation is written for each pair of (x, y) data point and the weight matrix is formed. The solution is then obtained using the Least Squares method. Table 6.4 gives the estimated coefficients with their confidence intervals computed at 95% confidence level.

Table 6.4: Coefficients of polynomial fit with 95% confidence interval

Coefficients	Value	Lower Bound	Upper Bound
a3	6.50E-08	-1.504e-007	2.805e-007
a2	0.0002624	0.0002404	0.0002845
a1	-0.1012	-0.1127	-0.08978
a0	-19. 3662	-20.40	-18.329

The original data points, estimated data points, the residuals and the standard errors of the estimates are presented in Table 6.5. A graphical display of the original data points and the fitted polynomial is also shown in Figure 6.12.

Table 6.5: Results of polynomial fitting with residuals and standard errors for estimated function

Point ID	Dependent x	Independent y	Fitted ŷ	Residuals y- ŷ	Std. Error of Function
1(C7)	248.3095	-27.3155	-27.3289	-0.0134	0.3089
2	223.0072	-31.4910	-28.1714	3.3197	0.4905
3	213.3047	-25.4633	-28.3902	-2.9269	0.6036
4	185.1165	-31.1707	-28.7024	2.4683	0.8665
5	163.8284	-27.7761	-28.6232	-0.8471	0.9859
6	154.3444	-27.4559	-28.5018	-1.0459	1.0187
7	145.5340	-25.5327	-28.3418	-2.8091	1.0389
8	132.9475	-31.7454	-28.0350	3.7104	1.0523
9	123.2559	-29.2698	-27.7365	1.5333	1.0518
10	109.4590	-30.1864	-27.2188	2.9676	1.0375
11	94.0342	-27.8143	-26.5121	1.3022	1.0072
12	83.6663	-23.8020	-25.9619	-2.1599	0.9813
13	68.7304	-23.4799	-25.0640	-1.5841	0.9407
14	54.9722	-20.9829	-24.1280	-3.1452	0.9049
15	8.6371	-19.1399	-20.2211	-1.0811	0.8532
16	3.0370	-17.2374	-19.6713	-2.4338	0.8574
17	-8.7872	-18.3971	-18.4563	-0.0592	0.8734
18	-17.3884	-18.0269	-17.5267	0.5003	0.8904
19	-26.6432	-19.4077	-16.4836	2.9241	0.9124
20	-46.4601	-15.3812	-14.1023	1.2789	0.9663
21	-60.1738	-12.6086	-12.3377	0.2708	1.0028
22	-78.7045	-10.0200	-9.8037	0.2163	1.0416
23	-88.0221	-7.1401	-8.4653	-1.3251	1.0534
24	-97.5416	-8.2458	-7.0539	1.1919	1.0584
25	-108.5884	-4.7365	-5.3607	-0.6242	1.0537
26	-122.9544	-1.9416	-3.0709	-1.1293	1.0276
27	-149.5413	-0.8307	1.4257	2.2564	0.9087
28	-171.1173	6.3585	5.3175	-1.0410	0.7354
29	-182.7560	6.7317	7.5057	0.7740	0.6132
30(S1)	-218.2487	14.5617	14.5553	-0.0065	0.3092

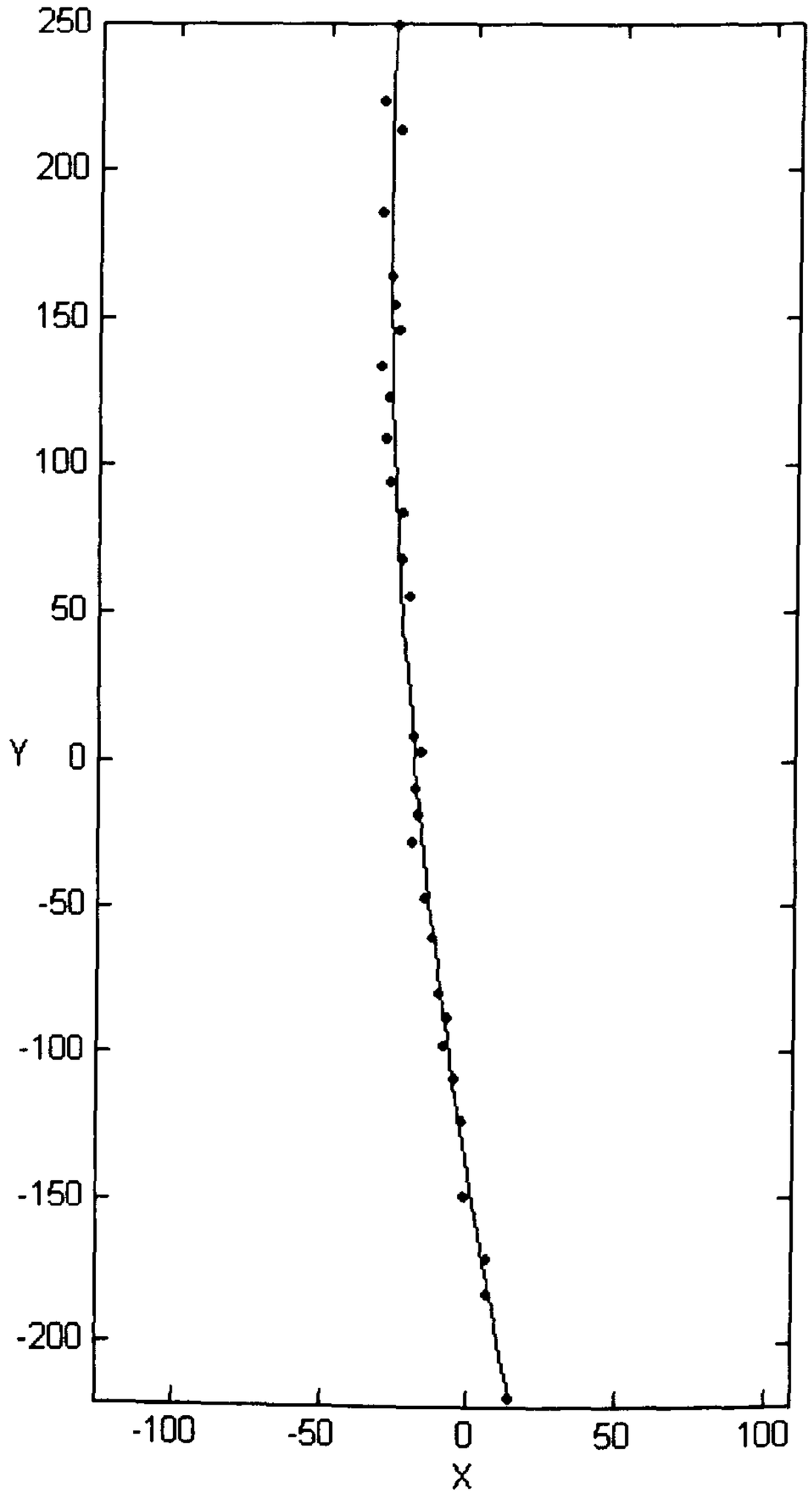


Figure 6.12: Results of the polynomial fit

6.6.5. Evaluating the Goodness of Fit

To evaluate how well the chosen model (i.e. constrained polynomial) fits the data (the spinous process landmarks), the goodness of fit should be analysed.

Residual Analysis

The first step in evaluating the goodness of fit is to analyse the residual data which are approximation of the random errors in the data. From Table 6.5, all the residuals lie within the range -3.4—3.7mm. The mean of the residuals is approximately 0.0827 and the standard deviation is less 2mm. This lies within the pixel size of the original data, indicating the function

fits the data quite well. A plot of the residuals relative to the fit (represented by the vertical zero line) is shown in Figure 6.13. The residuals are randomly scattered about the fit (i.e. they do not show any significant skewed distribution), which indicates that the model describes the data well.

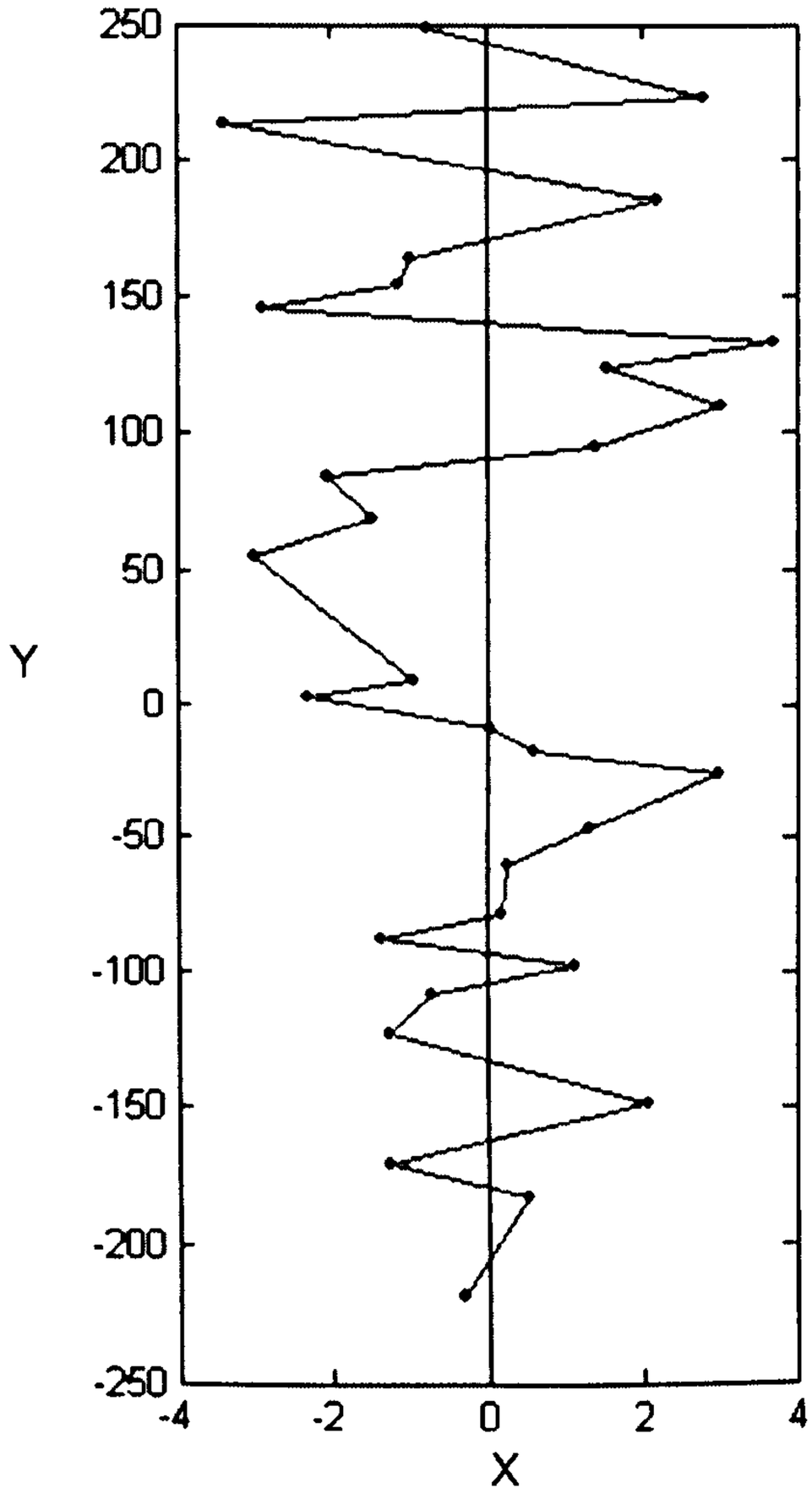


Figure 6.13: Residual plot, error is measured in X-axis

Goodness of Fit Statistics

The goodness-of-fit statistics have been computed and presented in Table 6.6. The sum of squared errors (SSE) which represents the total deviation of the original data from the fit is 0.39; and the root mean square error which represents the mean deviation of the original data values from the fit is only 0.12, indicating that the model fits the data well. Both R-square and adjusted R-square also indicate a good fit since they are approximately equal to 1. The reference variance or the variance of the random error is also estimated as 0.015.

Table 6.6: Goodness of fit statistics for polynomial fitting

Goodness of Fit Statistics	Value
Sum of Squared Error (SSE)	0.3870
Root Mean Square Error (RMSE)	0.1220
R-Square	0.9992
Adjusted R-Square	0.9991
Reference Variance	0.0149

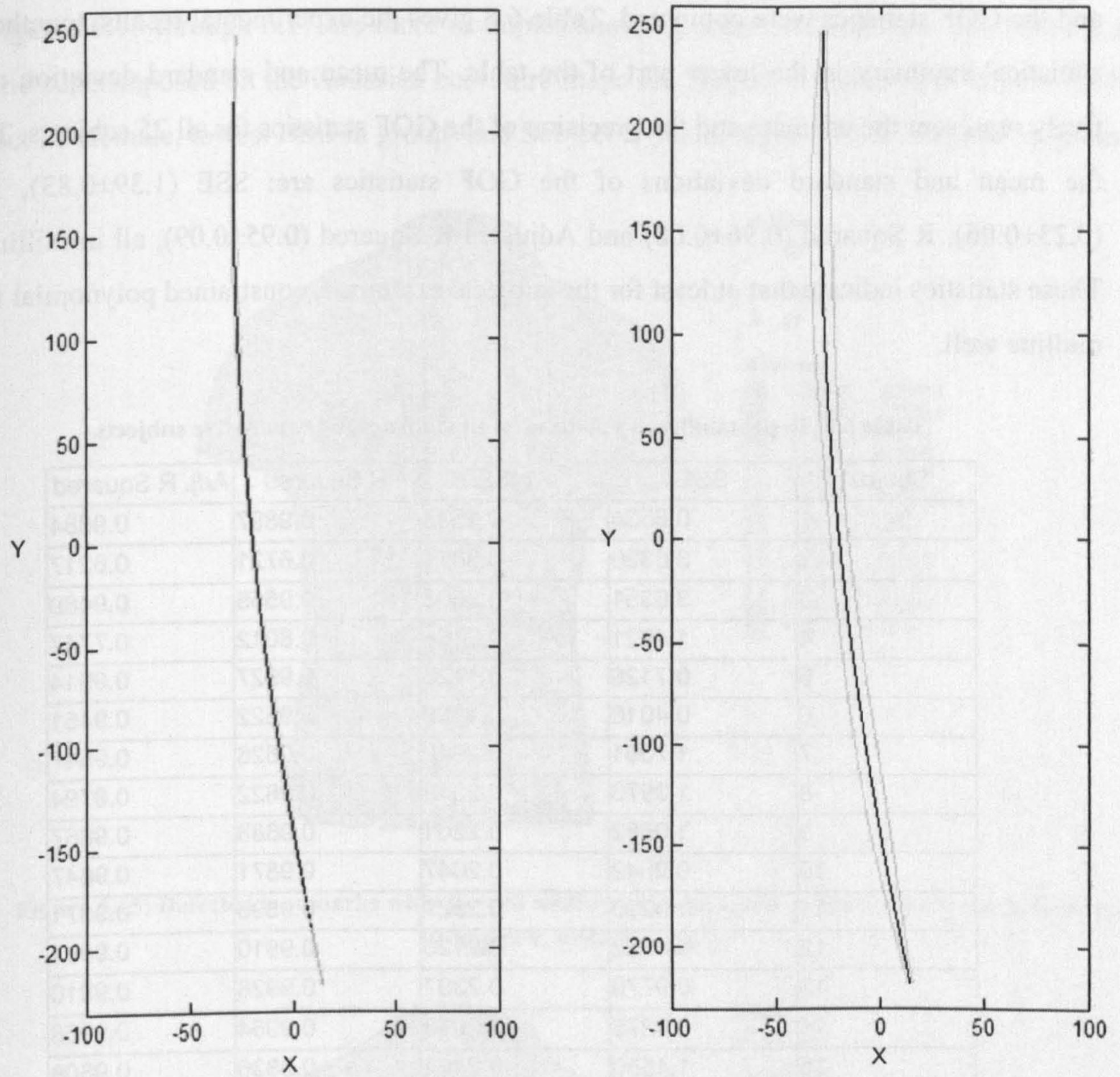
Confidence Bounds for the Fitted Function

The standard errors for the fitted function are listed in the sixth column of Table 6.5. These values can be used to construct confidence bounds for the fitted function at a specified confidence level, using Equation 4.43. The confidence bounds captures the region within which the true values of the fitted function are expected to lie had there been no errors in the original measurements and had the function fitted the measurements perfectly well. Table 6.7 gives statistical summary of the computed standard errors for the fitted function. The mean standard error of the fitted function is 0.88, estimated within ±0.21 precision (standard deviation). This represents the average error in the estimated function, propagated from the errors in the original measurements. This value is small which is indicative of the fact that the model fits the data well. This also implies the propagated errors will have minimal effect in any linear, angular and area measurements computed from the derived midline.

Table 6.7: Statistics of the residuals, and standard errors of fitted function

Statistic	Standard Error of fitted Function
Minimum	0.3089
Maximum	1.0584
Mean	0.8815
Median	0.9535
Std Dev	0.2124

Figure 6.14 shows a graphical display of the confidence bounds for the estimated function: (a) shows the confidence bounds constructed with the original standard error values and in (b) the standard errors used to construct the confidence intervals are scaled by a factor of 5 in order to make the graph visible. The confidence region is narrow at the end points (C7 and S1) due to the assignment of higher weights to these points.



(a) Confidence bounds – original values

(b) Confidence bounds – scaled by a factor of 5

Figure 6.14: Confidence bounds for fitted function (a) original values (b) scaled values

6.6.6. Repeatability of the Goodness of Fit Statistics

The goodness of fit statistics describes how well the model (i.e. constrained polynomials) fits the detected spinous processes. To assess the repeatability of the GOF statistics derived above and hence the suitability of constrained polynomials in describing the back midline, the process was applied to the same twenty five subjects used in the landmark validation experiment. The curvature maps of each subject were computed at a smoothing level of 2.25 and the landmarks were detected. For each subject, a number of spinous process landmarks were selected by drawing a polygonal region around the points. The selection always included the C7

and S1 landmarks. A polynomial of the third degree was then fitted to the selected landmarks, and the GOF statistics were computed. Table 6.8 gives the experimental results, together with statistical summary in the lower part of the table. The mean and standard deviation respectively represent the estimate and the precision of the GOF statistics for all 25 subjects. That is, the mean and standard deviations of the GOF statistics are: SSE (1.39±0.83), RMSE (0.23±0.06), R Squared (0.96±0.08) and Adjusted R Squared (0.95±0.09), all in millimetres. These statistics indicate that at least for the subjects examined, constrained polynomial fits the midline well.

Table 6.8: Repeatability of goodness of fit statistics for twenty five subjects

Subject	SSE	RMSE	R Squared	Adj. R Squared
1	0.8534	0.1633	0.9897	0.9884
2	3.6390	0.3671	0.6721	0.6217
3	3.6351	0.3603	0.9555	0.9489
4	1.4921	0.2194	0.8012	0.7747
5	0.7128	0.1723	0.9927	0.9914
6	0.4016	0.1198	0.9522	0.9451
7	1.7391	0.2949	0.9626	0.9547
8	1.3970	0.2318	0.9822	0.9794
9	1.0662	0.2201	0.9888	0.9867
10	0.9642	0.2047	0.9871	0.9847
11	1.0290	0.2327	0.9895	0.9871
12	0.8562	0.2123	0.9910	0.9891
13	0.9770	0.2397	0.9928	0.9910
14	0.9275	0.2053	0.9964	0.9958
15	1.4550	0.2462	0.9836	0.9808
16	1.4899	0.2602	0.9920	0.9905
17	1.2862	0.2475	0.9926	0.9911
18	1.1660	0.2118	0.9884	0.9865
19	1.5034	0.2405	0.9883	0.9865
20	1.1084	0.2065	0.9913	0.9899
21	1.8277	0.2704	0.9702	0.9652
22	0.3329	0.1203	0.9981	0.9978
23	0.6312	0.1734	0.9920	0.9904
24	0.5780	0.1901	0.9852	0.9812
25	0.8649	0.1898	0.9950	0.9941
Minimum	0.3329	0.1198	0.6721	0.6217
Maximum	3.6390	0.3671	0.9981	0.9978
Range	3.3061	0.2474	0.3260	0.3761
Mean	1.2774	0.2240	0.9652	0.9597
Std Deviation	0.8091	0.0590	0.0723	0.0829

Figures 6.15 through 6.17 are more examples showing detected landmarks and derived midline superimposed on the Gaussian curvature maps for Subject 9 (male, BMI 25km/m^2), Subject 25 (female, lowest BMI in group) and Subject 2 (male, highest BMI 28km/m^2 in group).

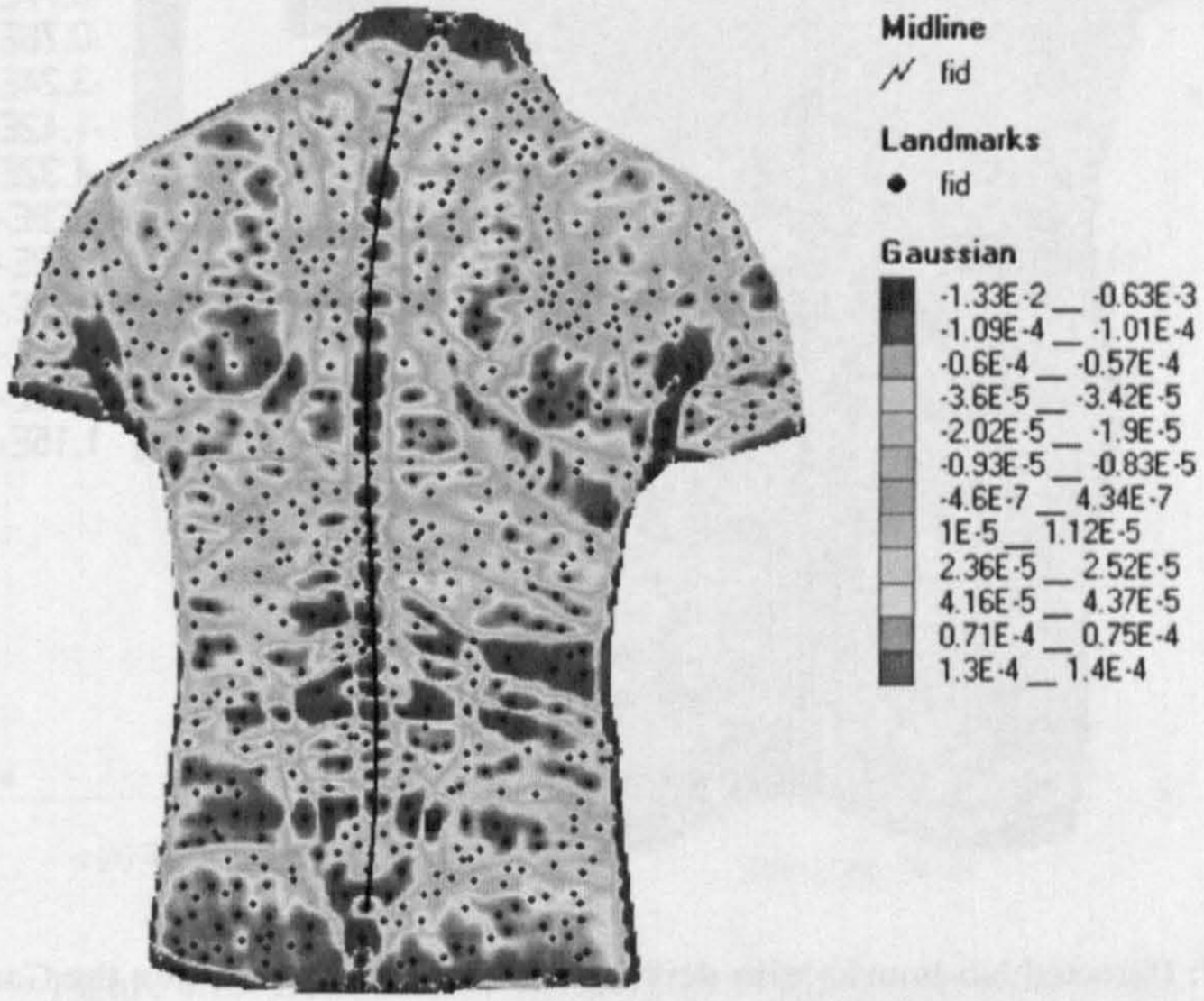


Figure 6.15: Detected landmarks with derived midline superimposed on the Gaussian curvature map (Subject 9, BMI 25 kg/m^2)

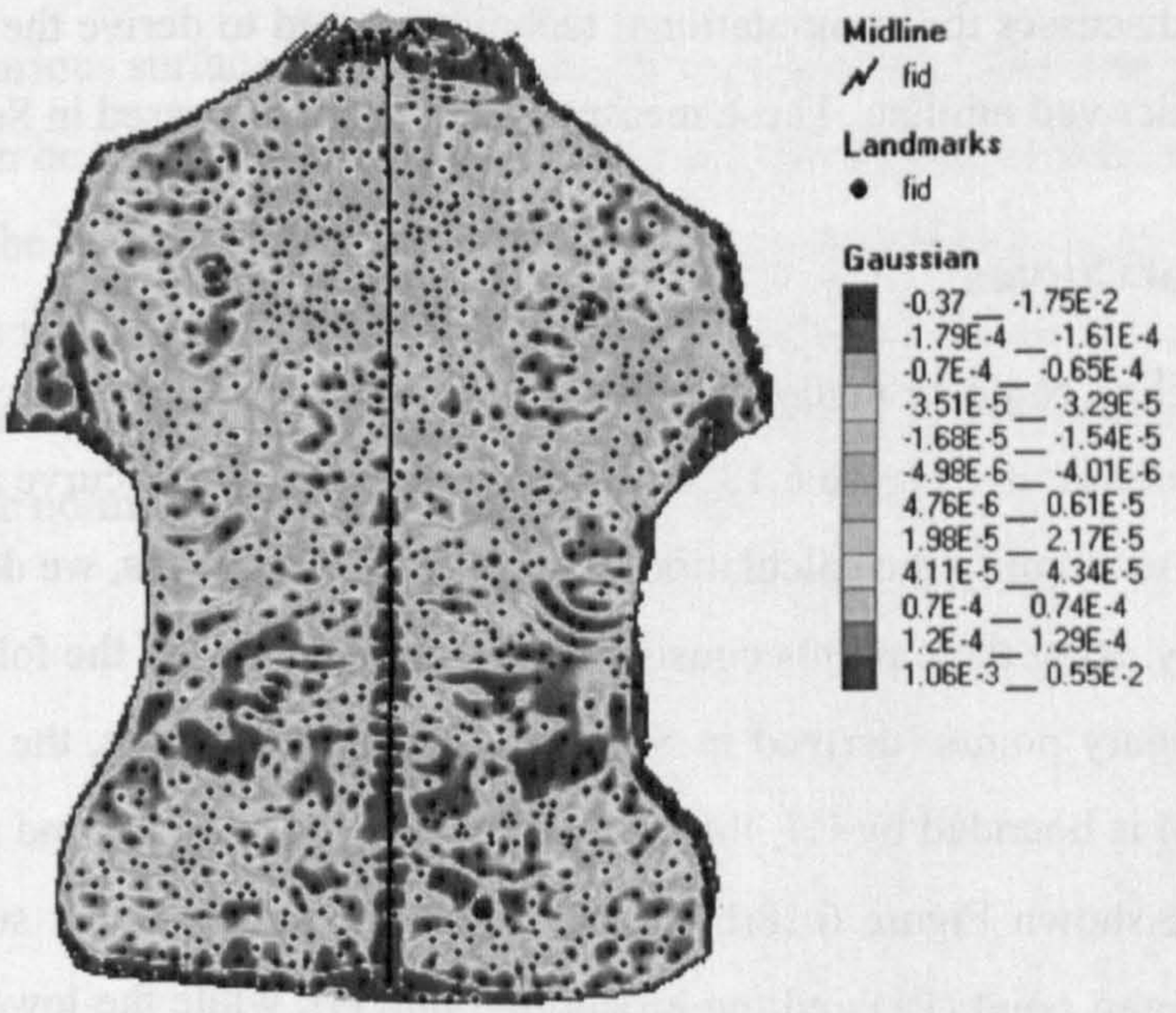


Figure 6.16: Detected landmarks with derived midline superimposed on the Gaussian curvature map (Subject 25, female, lowest BMI 22kg/m^2)

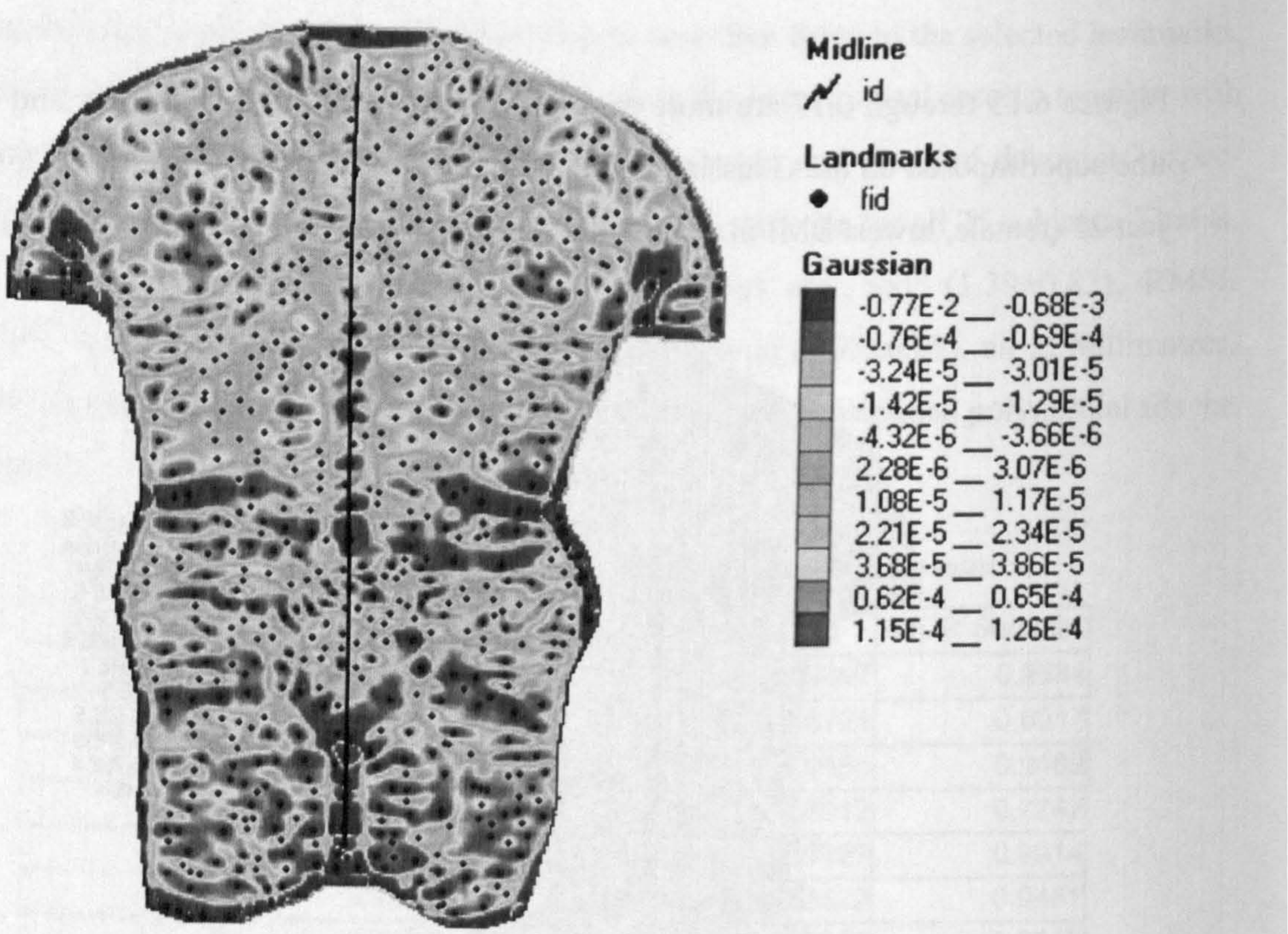


Figure 6.17: Detected landmarks with derived midline superimposed on the Gaussian curvature map (Subject 2, male, highest BMI 28kg/m^2)

6.7. Computation of Surface Measurements

This section discusses the computational techniques used to derive the surface measurements based on the derived midline. These measurements were discussed in Section 2.5.

6.7.1. Spinal Curves

The back midline could be made up of a single, double or even triple curve depending on the extent of the deformity. Figure 6.18 shows examples of a single-curve and double-curve midline. In order to simplify the calculation of surface measurements, we define a simple curve of the midline by using three points consisting of a combination of the following points: C7, S1, and the stationary points (derived in Section 6.2.2). For example, the single curve shown in Figure 6.18(a) is bounded by C7, the minimum point (P1) and the end point (S1). The double curve midline shown Figure 6.18(b) has two sub-curves: the upper sub-curve is defined by C7, the minimum point (P1) and the inflection point (I); while the lower sub-curve is defined by the inflection point (I), maximum point (P2) and S1. A simple curve must always have one

and only one local minimum or maximum. These definitions allow us to evaluate the properties of each individual curve as well as the entire midline.

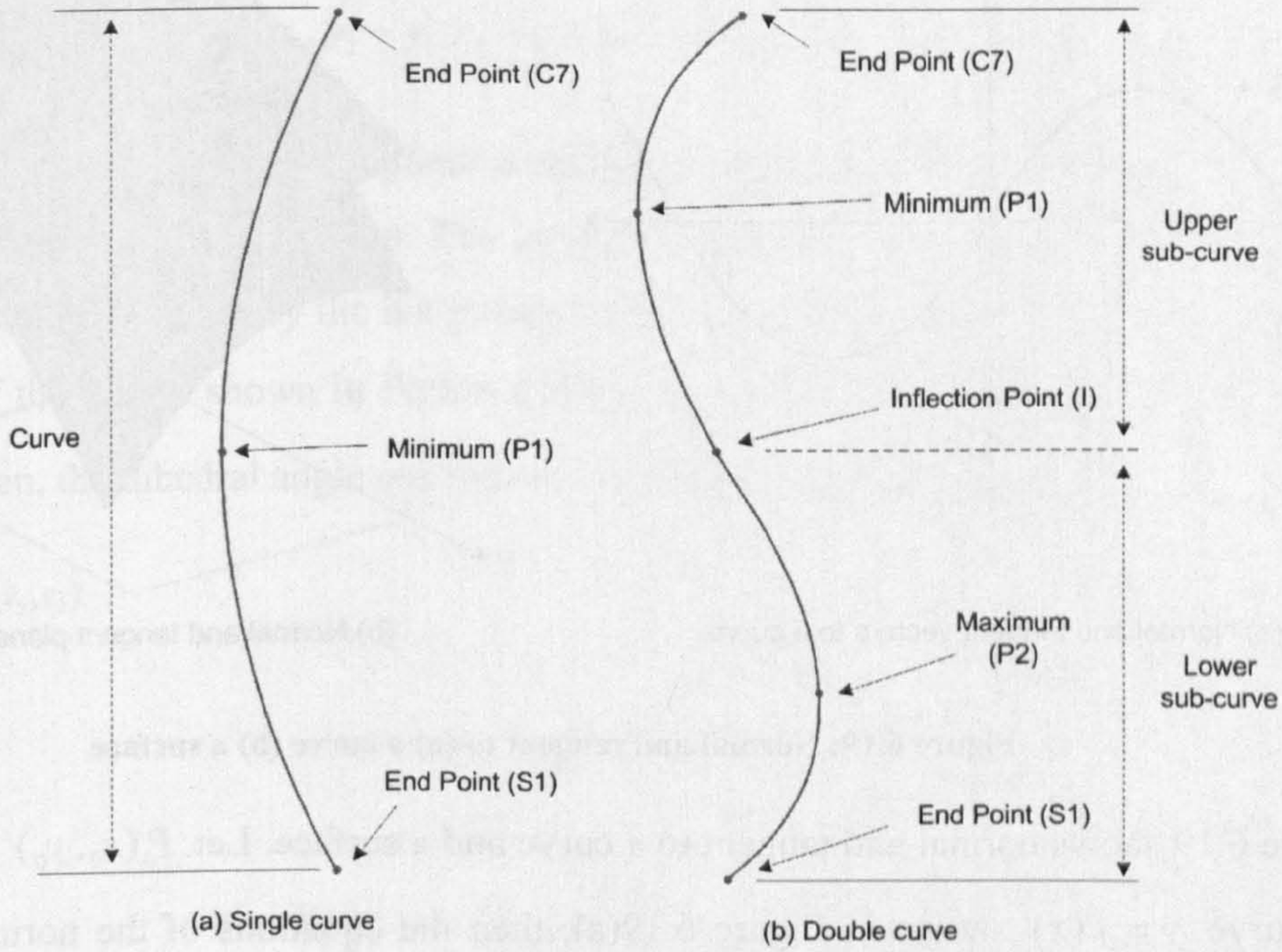
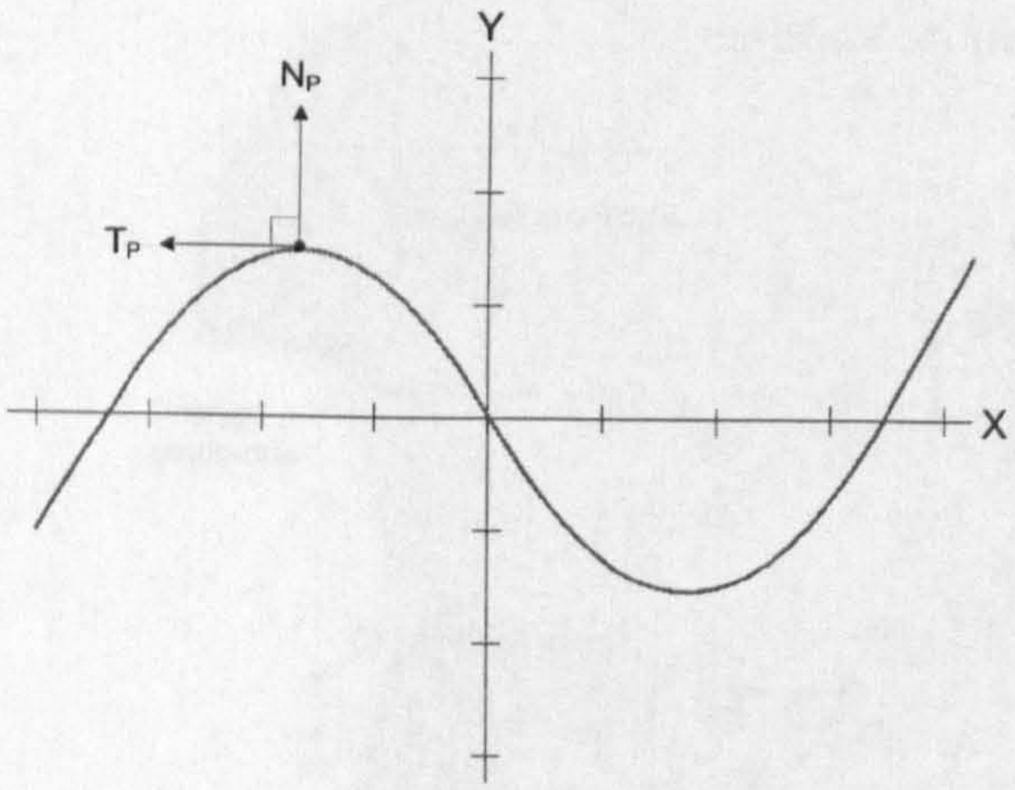


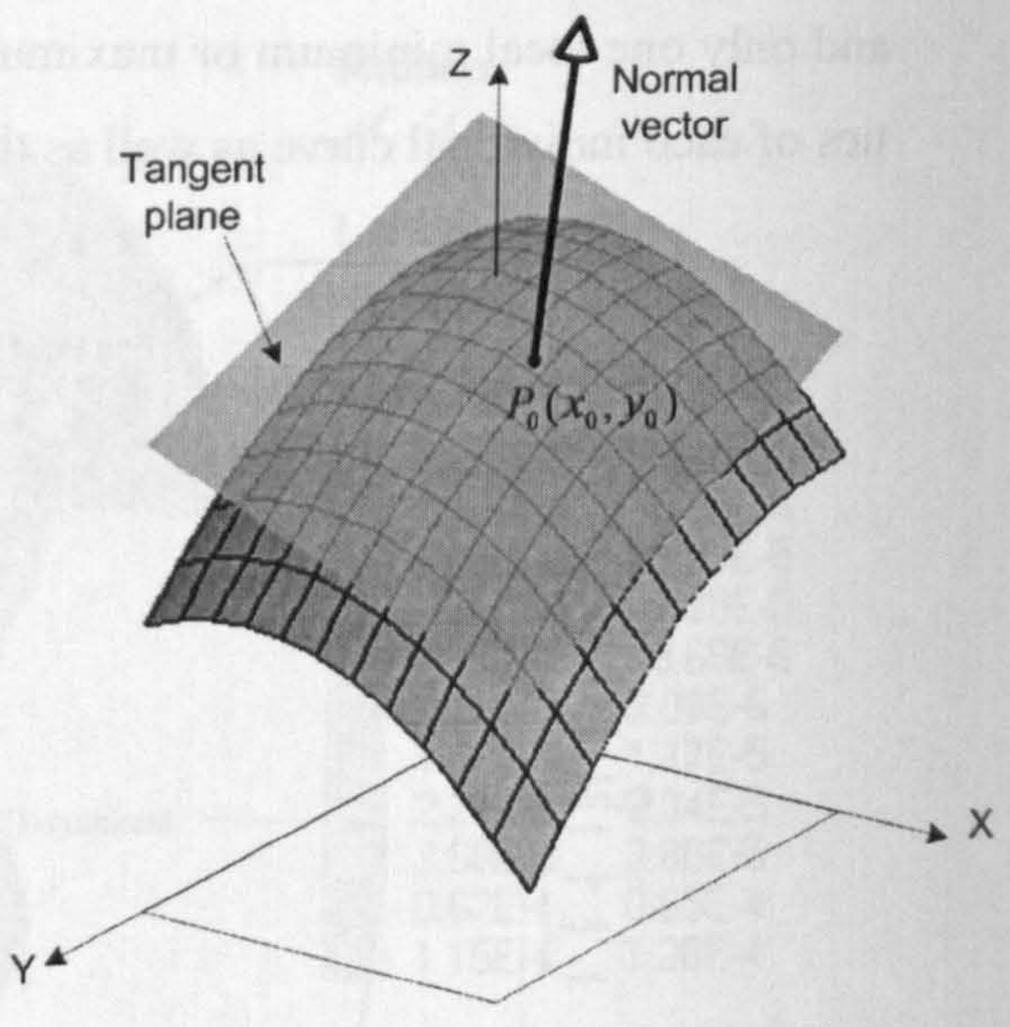
Figure 6.18: Definition of midline curves using stationary points

6.7.2. Normal and Tangent Vectors

Normal and tangent vectors are useful geometric shape descriptors which are required for the computation of various surface measurements. For example, the Cobb angle of a simple curve of the midline can be calculated as the angle between the normal vectors at the end points of the curve; while the Ferguson angle is the angle between the tangent vectors. The normal vector (or simply the normal) is a vector that is perpendicular to a curve or surface (i.e. meet the curve or surface at right angle). A tangent vector is a vector that passes through a point on the curve, meeting the normal at right angle (Figure 6.19).



(a) Normal and tangent vectors to a curve



(b) Normal and tangent plane to a surface

Figure 6.19: Normal and tangent to (a) a curve (b) a surface

Figure 6.19 shows normal and tangent to a curve and a surface. Let $P_0(x_0, y_0)$ be the point on the curve $y = f(x)$ shown in Figure 6.19(a), then the equations of the normal and tangent lines at P_0 are given by Equation 6.9.

$$\text{tangent line:} \quad y = f(a) + f'(a)(x-a)$$
$$\text{normal line:} \quad y = f(a) + \frac{1}{f'(a)}(x-a) \tag{6.9}$$

On a surface, the normal vector gives a measure of the orientation of the surface at that point. Let $P_0(x_0, y_0)$ be on the surface $z = F(x, y)$ shown in Figure 6.19(b); then the normal vector can be calculated using Equation 4.2.5. The surface has a non-vertical tangent plane at $P_0(x_0, y_0)$ with the following equation:

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x-x_0) + f_y(y_0, y_0)(y-y_0) \tag{6.10}$$

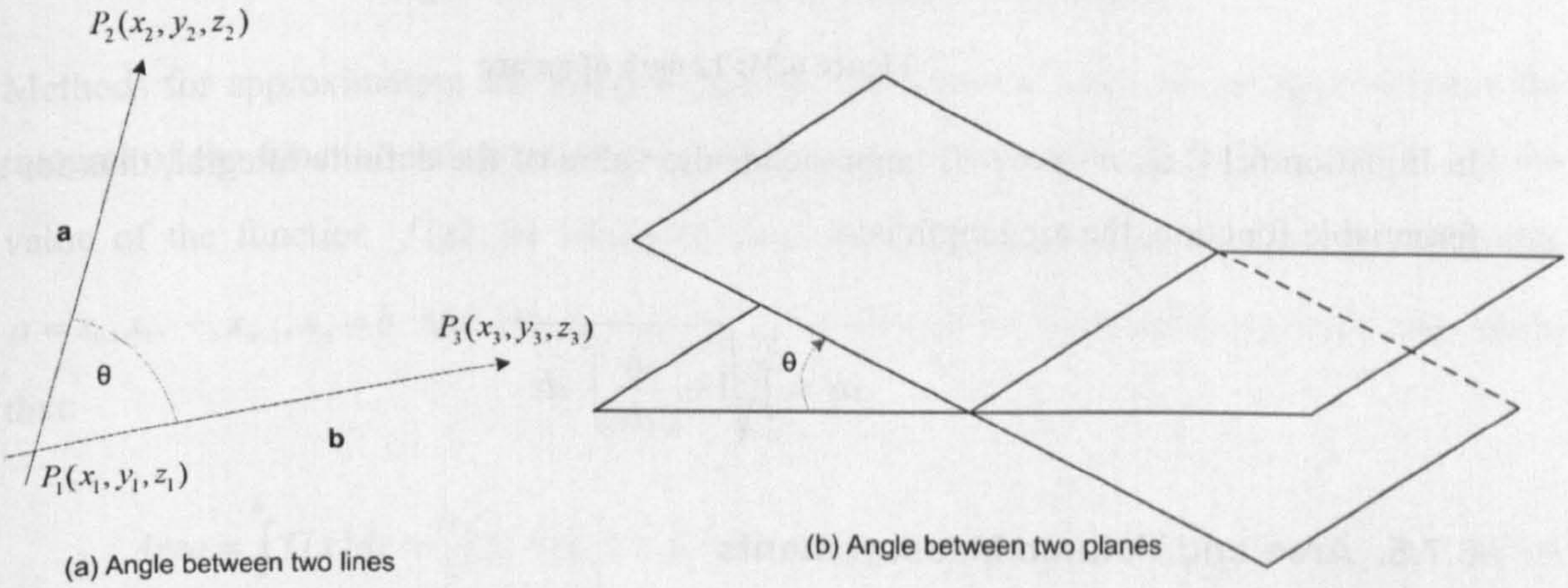
6.7.3. Angular Measurements

The following methods are used to compute the angular measurements discussed in Section 2.5. The angle between any two intersecting lines can be calculated using vector mathematics. In Figure 6.20(a), the angle between the lines $\overline{P_1P_2}$ and $\overline{P_1P_3}$ is given by the angle between their respective direction vectors **a** and **b** .

$$\cos\theta = \frac{a_1 a_2 + b_1 b_2 + c_1 c_2}{\sqrt{a_1^2 + b_1^2 + b_1^2}\sqrt{a_2^2 + b_2^2 + b_2^2}} \tag{6.11}$$

where $\mathbf{a}(a_1, a_2, a_3) = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$ and $\mathbf{b}(b_1, b_2, b_3) = (x_3 - x_1, y_3 - y_1, z_3 - z_1)$.

The angle between a line and a plane is defined by the orientation of the normal vector to the plane at the point of intersection. The angle between any two intersecting planes, known as 'dihedral angle' is given by the dot product of their respective normal vectors. Let the normal vectors of the planes shown in Figure 6.20(b), be $a_1 = (a_1, b_1, c_1)$ and $a_2 = (a_2, b_2, c_2)$ respectively. Then, the dihedral angle can then be calculated using Equation 6.11.



(a) Angle between two lines

(b) Angle between two planes

Figure 6.20: Angle between two intersecting (a) lines (b) planes

6.7.4. Linear Measurements

The linear distance between any two points $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ is calculated as the Euclidean distance between the points:

$$L = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \tag{6.12}$$

In two-dimension, the distance is obtained by simply setting the Z-dimension to zero. To obtain the length of the midline, the entire curve is subdivided into n small segments and the length of the curve is approximated as the sum of the lengths of the resulting segments. For example, suppose we want to obtain the 2D length of the curve $y = f(x)$ between $x = a$ and $x = b$. We parameterise the entire curve by the arc length and approximate the length of the arc by the length of the chord PG shown in Figure 6.21.

$$\delta s = \sqrt{\delta x^2 + \delta y^2} \Rightarrow \delta s = \sqrt{1 + \left(\frac{\delta y}{\delta x}\right)^2}\, \delta x \tag{6.13}$$

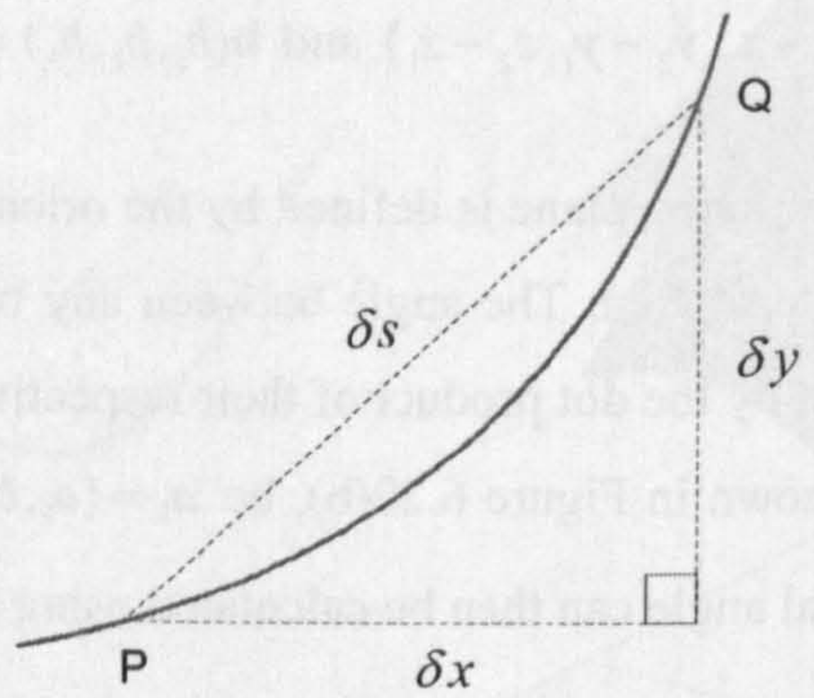


Figure 6.21: Length of an arc

In Equation 6.14, as $n \to \infty$, δs approaches the value of the definite integral, thus for a differentiable function, the arc length is:

$$ds = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \tag{6.14}$$

6.7.5. Area and Volume Measurements

The area enclosed by the midline about the spinal axis (C7-S1) is given by the sum of the areas of the sub-curves about the spinal axis. The area of a curve under the 'X-axis' over an interval [a, b] is defined as the integral of the curve over that interval. The total area of the double curve shown in Figure 6.22 is the sum of the areas of the sub-curves, i.e., $A_T = A_1 + A_2$. For the midline, the X-axis will be represented by the spinal axis, that is, the line connecting C7 and S1 landmarks.

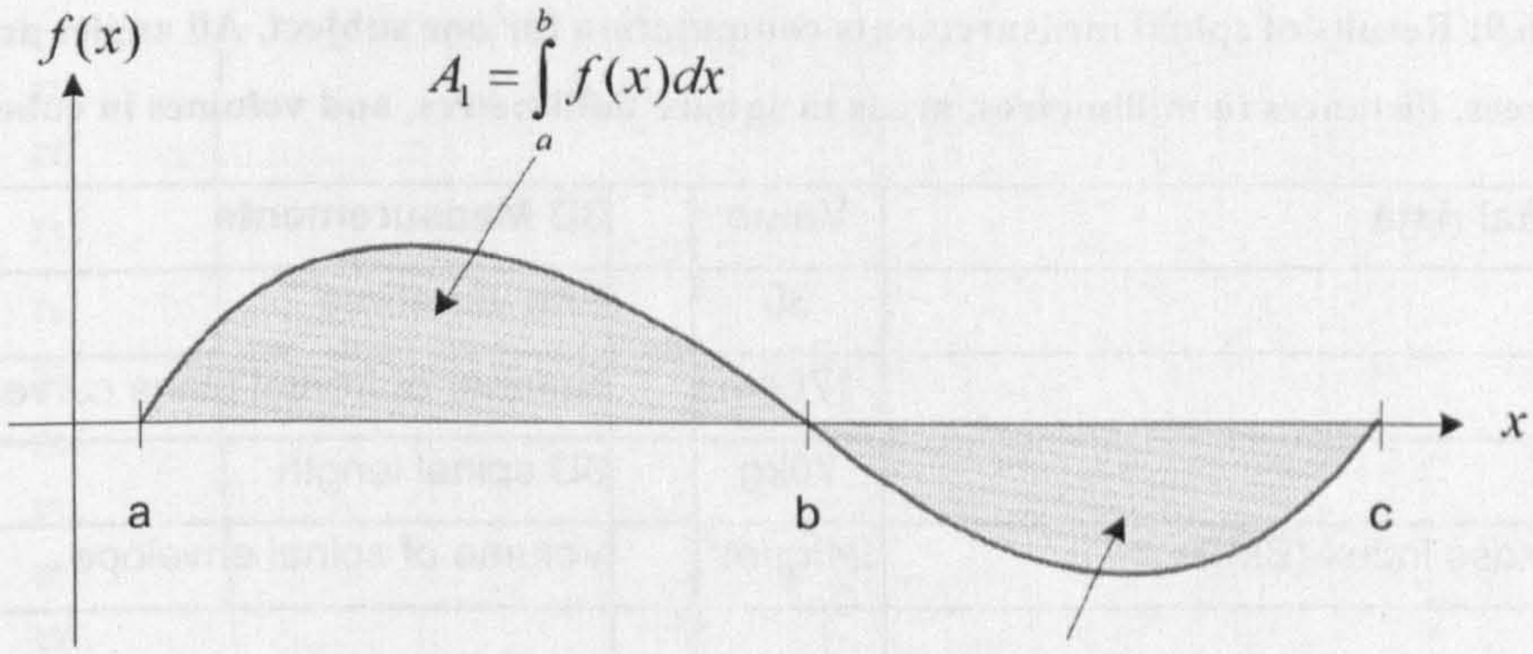


Figure 6.22: Area enclosed by a double curve midline

Methods for approximating the integral include the Simpson's rule which approximates the integral of the function using quadratic polynomials (Abramowitz & Stegun, 1972). Let the value of the function $f(x)$ be tabulated over the interval $[a,b]$ at equally spaced points $a = x_0, x_1, \cdots, x_{n-1}, x_n = b$ and let $f_n = f(x_n)$, $h = (b-a)/n$. Then the Simpson's rule states that:

$$Area = \int_a^b f(x)dx \approx \frac{h}{3}[f_0 + 4(f_1 + f_3 + \cdots + f_{2n-1}) + 2(f_2 + f_4 + \cdots + f_{2n-2}) + f_n] \quad \textbf{(6.15)}$$

The volume of the spinal bounding box is calculated using the straight-forward formula in Equation 6.16.

$$Volume = Length \times Width \times Depth \quad \textbf{(6.16)}$$

6.7.6. Implementation of the Surface Measurements

Using the techniques discussed in the preceding sections, the surface measurements have been implemented based on the derived midline. Table 6.9 gives an example of the results of surface measurements computed for one subject with a mild curve (age 30, weight 70kg, height 170cm, BMI 24).

Table 6.9: Results of spinal measurements computation for one subject. All angles are measured in degrees, distances in millimetres, areas in square millimetres, and volumes in cube millimetres.

Personal data	Value		3D Measurements	Value
Age	30		End vertebrae	C7-S1
Height	170cm		Number of frontal plane curves	1
Weight	70kg		3D spinal length	486
Body mass index (BMI)	24kg/m^2		Volume of spinal envelope	1524921
Frontal Plane Measurements	**Value**		**Sagittal Plane Measurements**	**Value**
2D spinal length in frontal plane	469.61		2D spinal length in sagittal plane	483.06
Frontal plane balance (C7)	41.88		Sagittal plane balance (C7)	42.61
Frontal plane angular balance (C7)	5.13°		Sagittal plane angular balance (C7)	5.22°
Spinal lateral deviation	14.4		Sagittal spinal deviation	25.44
Spinal lateral width	43.26		Sagittal spinal width	75.55
Maximum lateral deviation	43.26		Sagittal maximum deviation	60.00
Angle of best fit plane with sagittal plane	5.13°		Angle of best fit line with frontal plane	5.23°
Maximum analytical Cobb angle	13.99°		Kyphosis angle by Cobb method	44.44°
Maximum analytical Ferguson angle	7.04°		Kyphosis angle by Ferguson method	30.57°
Projected area in frontal plane	4492		Lordosis angle by Cobb method	43.76°
			Lordosis angle by Ferguson method	24.3°
			Projected area in sagittal plane	6202

Figure 6.23 shows a graphical display of the derived midline curve in frontal plane (a) and in sagittal plane (b). The positions of the vertebrae were estimated by their spinous processes landmarks detected from the curvature maps. The line connecting the left and right PSIS landmarks is also shown. The curve in Figure 6.23(c) represents the back surface rotation (BSR, see Section 3.3.5.4) estimated by the orientation of the normal vector at each vertebral level.

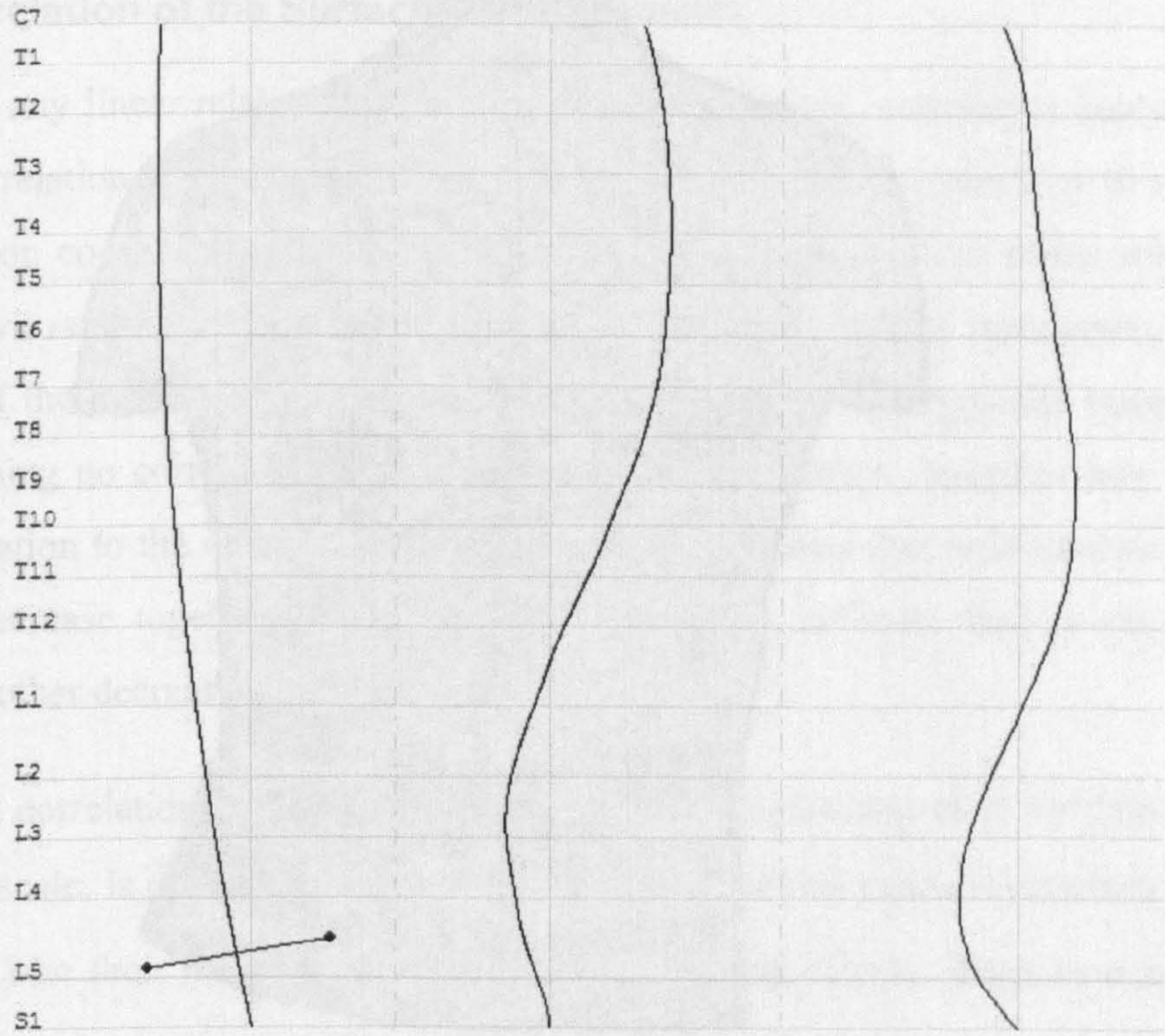


Figure 6.23: Midline curves for one subject (a) frontal plane (b) sagittal plane (c) back surface rotation (BSR)

Figure 6.24 shows a display of the Gaussian curvature map with the detected spinous process landmarks and the derived midline superimposed. In Figure 6.25, the 3D midline is superimposed on the original surface of the back.

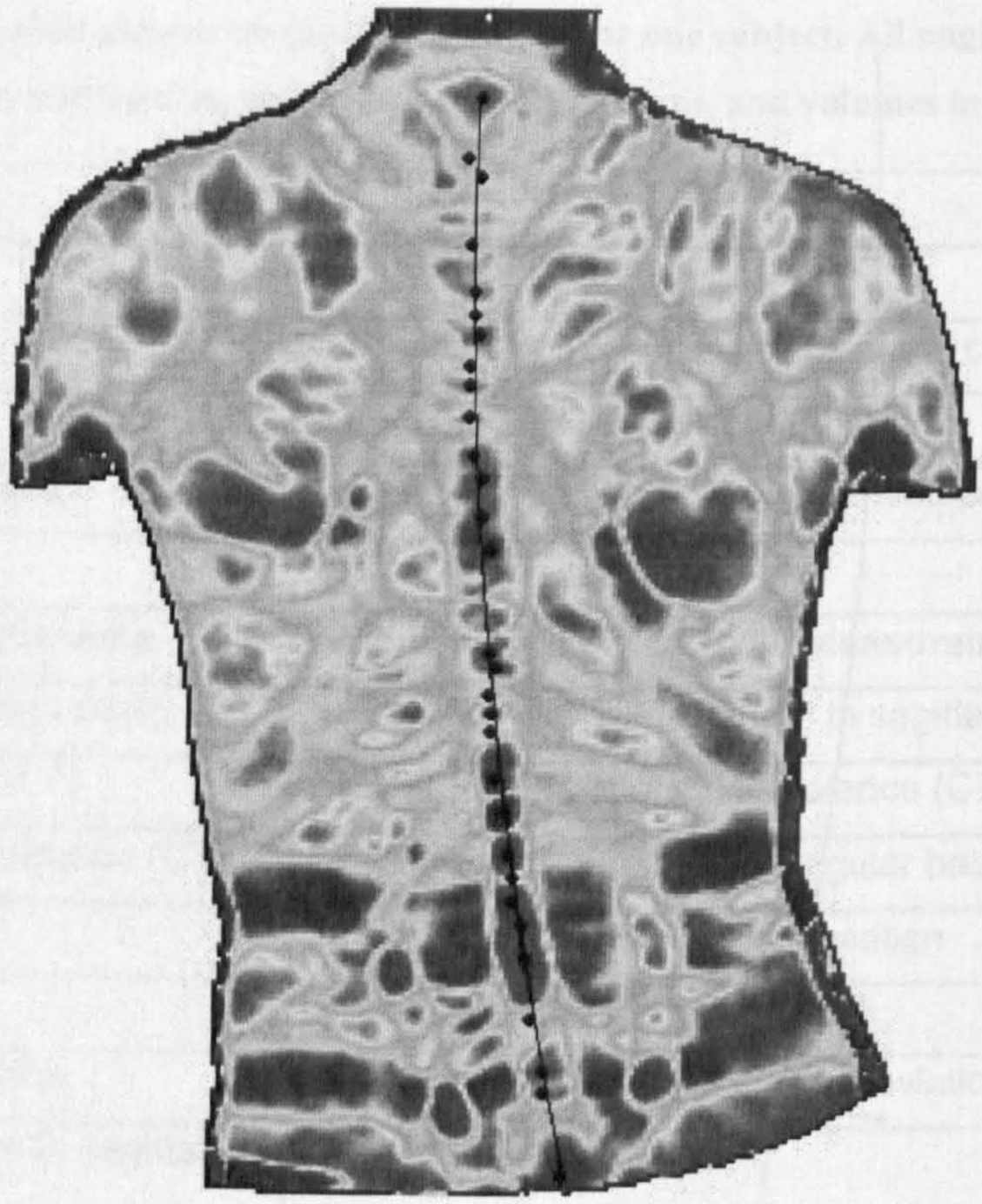


Figure 6.24: Gaussian curvature map with detected spinous process landmarks and the derived midline superimposed

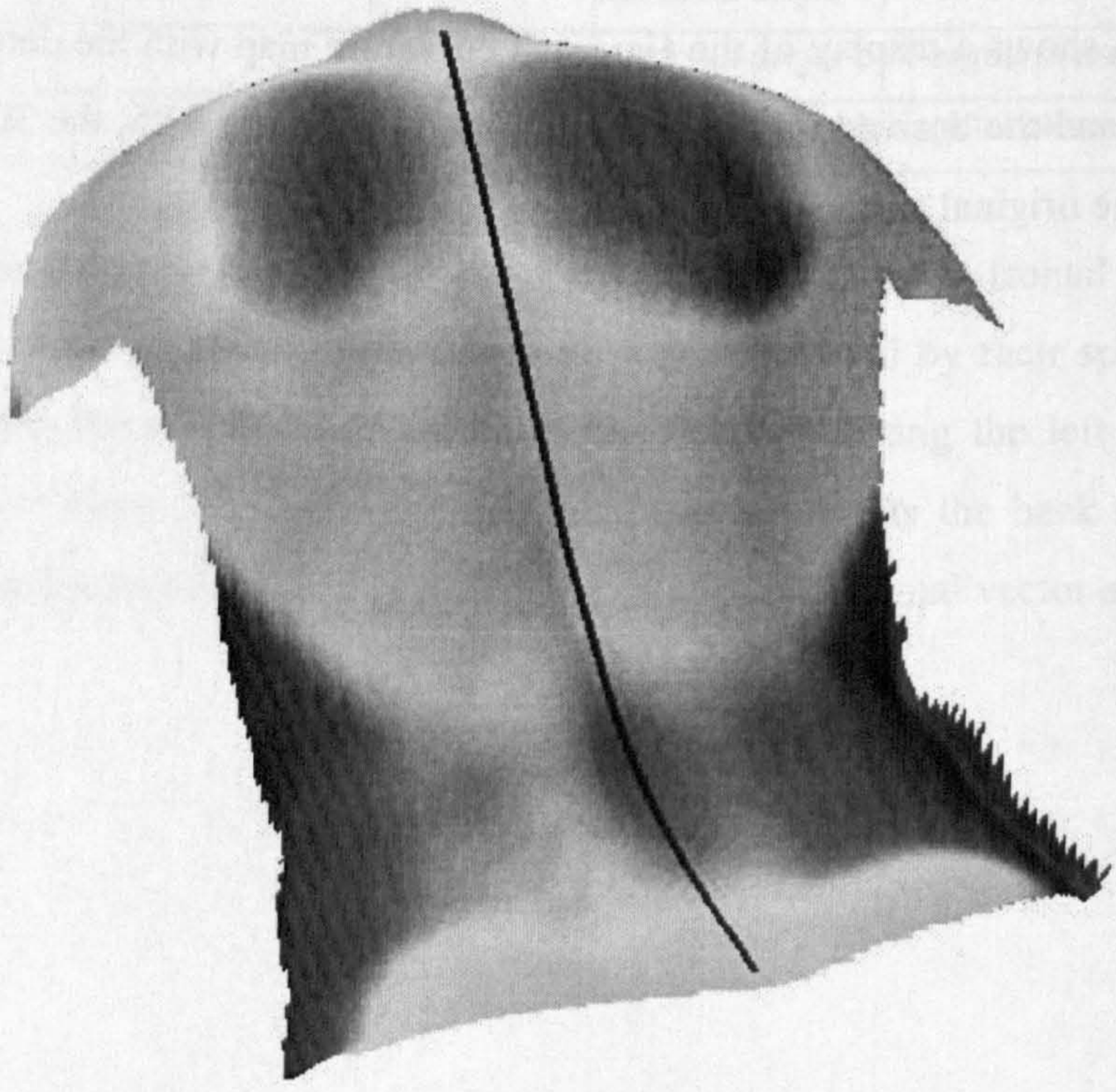


Figure 6.25: Midline superimposed on original surface of the back

6.7.7. Correlation of the Surface Measurements

To examine any linear relationships among the measurement, correlation analysis were performed. Correlation is a measure of the tendency of two random variables to vary together; the correlation coefficient indicates the strength and direction of the linear relationship between the two random variables. The strength of the relationship is represented by the absolute value of the magnitude of the correlation coefficient (which typically ranges from -1 to +1); 0 meaning no correlation. The direction of the correlation describes how one variable varies in relation to the other. A positive correlation indicates that both random variables increase or decrease together, whilst negative correlation indicates that as one variable increases, the other decreases, and vice versa.

The Pearson correlation coefficient, which is the most commonly used for data measured on the interval scale, is defined as follows. Let x and y be two random variables and ($\bar{x}$, σ_x) and ($\bar{y}$, σ_y) be their mean and standard deviations respectively. Then Pearson correlation coefficient (r) is defined as the product of the standardised scores of the two variables divided by the degrees of freedom (Equation 6.17).

$$r = \frac{1}{n-1}\left(\frac{(x-\bar{x})}{\sigma_x} \bullet \frac{(y-\bar{y})}{\sigma_y}\right) \tag{6.17}$$

The standardised scores are dimensionless quantities derived by subtracting the mean and diving by the standard deviation, making the correlation independent on the original unit of the measurements.

To carry out correlation analysis, the surface measurements were computed for the 25 subjects and Equation 6.17 was used to compute the Pearson correlation coefficient between each pair of surface measurement. Table 6.10 and Table 6.11 give the correlation matrix for the frontal plane and the sagittal plane measurements respectively, which list the correlation coefficients between all pairs of surface measurements. The correlation coefficient for any pair of surface measurements can be located by intersecting corresponding row and column of the tables. For example, the correction coefficient between frontal plane balance and spinal lateral width can be found by intersecting row 5 and column 8 in Table 6.10 to give $r = 0.780$. The diagonal elements of the correlation matrices are all 1 because these are the correlations be-

tween each variable and itself. The lower and upper triangular portions of are mirror images of each other.

To provide qualitative description of the correlation results, we use the following classification for both positive and negative correlation: weak ± (0—0.39), moderate ± (0.40—0.69), and strong ± (0.70—1.0). With this classification, it can be observed that the surface Cobb angle correlates strongly with the Ferguson angle ($r = 0.944$), the lateral deviation ($r = 0.934$), projected area in the frontal plane ($r = 0.733$); and moderately with the spinal lateral width ($r = 0.630$). All other measurements correlate weakly with the Cobb angle. In the frontal plane, both linear and angular balance correlate strongly with lateral width, maximum lateral deviation and the angle of the best fit plane ($r > 0.70$ in all cases). Similar observation can be made for corresponding measurements in the sagittal plane. The strongest correlations were observed between balance and the angle of best fit plane ($r > 0.99$ in all cases).

It is important to note that the correlation between any pair of the measurements does not represent the accuracy of the measurements, but rather the strength and direction of the relationship between them. Some of the measurements relate directly to each other and thus have strong correlation with each other; while other measurements have indirect or no relationship with each other, resulting in weak correlation. For example, there is a direct relationship between linear and angular balance in both frontal and sagittal planes, which is indicated a strong correlation between (i.e. $r = 0.997$ in frontal plane and $r = 0.994$ in the sagittal plane). There is also a strong correlation between the surface Cobb angle and the spinal lateral deviation: as the lateral deviation increases, the Cobb angle increases as well. On the other hand, weak negative correlation exists between 3D spinal length and many measurements in the frontal and sagittal plane.

6.7.8. Concluding Remarks on Surface Measurements

From a clinical perspective, the radiographic Cobb angle has been accepted as the 'gold standard' for scoliosis assessment. However, as already pointed out in Sections 1.2 and 3.2, there are problems with the use of Cobb angle as the 'gold standard', namely, the wide variations in its measurements and the fact that it is obtained from a two-dimensional, low resolution radiograph which doesn't account for vertebral rotation. These issues make comparison with radiographs difficult. To facilitate radiographic comparison, an objective method should be use for the measurement of the Cobb angle, such as fitting a smooth curve to the spine and

measuring the Cobb angle based on the inflexion points of the curve, rather than the end vertebrae.

Overall, this study has demonstrated that surface measurements can serve as alternative or complementary measurements for the assessment of back and spinal deformities. However, due to smaller sample size examined, conclusive inferences cannot be made. Further studies involving several subjects and body types are required to validate the surface measurements.

Table 6.10: Correlations matrix for Frontal plane measurements

	3D Len	Vol	2D Len	Balance	AngBal	LatDev	LatWidth	MaxLatDev	BestFitAng	Cobb	Ferg	Area
Len3D	1	0.0392	0.9881	-0.4263	-0.4748	0.1827	-0.1651	-0.3961	-0.4748	0.2574	0.1624	0.2505
Vol	0.0392	1	-0.0424	0.6006	0.5958	0.6247	0.8323	0.5675	0.5958	0.5413	0.6174	0.464
2D Len	0.9881	-0.0424	1	-0.4387	-0.4904	0.1781	-0.1818	-0.4135	-0.4904	0.2821	0.1727	0.2525
Balance	-0.4263	0.6006	-0.4387	1	0.9973	0.2093	0.7977	0.9108	0.9973	0.1866	0.2331	0.1463
AngBal	-0.4748	0.5958	-0.4904	0.9973	1	0.1766	0.7837	0.9143	1	0.148	0.2017	0.1053
LatDev	0.1827	0.6247	0.1781	0.2093	0.1766	1	0.6836	0.1458	0.1766	0.9384	0.9574	0.7973
LatWidth	-0.1651	0.8323	-0.1818	0.7977	0.7837	0.6836	1	0.759	0.7837	0.6297	0.6846	0.5588
MaxLatDev	-0.3961	0.5675	-0.4135	0.9108	0.9143	0.1458	0.759	1	0.9143	0.1228	0.1941	0.174
BestFitAng	-0.4748	0.5958	-0.4904	0.9973	1	0.1766	0.7837	0.9143	1	0.148	0.2017	0.1053
Cobb	0.2574	0.5413	0.2821	0.1866	0.148	0.9384	0.6297	0.1228	0.148	1	0.944	0.7334
Ferg	0.1624	0.6174	0.1727	0.2331	0.2017	0.9574	0.6846	0.1941	0.2017	0.944	1	0.7511
Area	0.2505	0.464	0.2525	0.1463	0.1053	0.7973	0.5588	0.174	0.1053	0.7334	0.7511	1

Table 6.11: Correlation matrix for Sagittal plane measurements

	3D Len	Vol	LenSP	Balance	AngBal	LatDev	LatWidth	MaxLatDev	BestFitAng	KyphCobb	KyphFerg	LordCobb	LordFerg	Area
3D Len	1	0.0392	0.9996	0.1802	0.093	0.4438	0.2078	0.129	0.0924	0.0345	-0.0434	-0.3484	-0.339	0.6375
Vol	0.0392	1	0.0214	0.7725	0.8023	0.4453	0.8351	0.854	0.8034	0.4904	0.612	-0.1391	-0.1008	0.2195
2D Len	0.9996	0.0214	1	0.1724	0.0848	0.4451	0.2007	0.1207	0.0842	0.0358	-0.0491	-0.347	-0.3373	0.6377
Balance	0.1802	0.7725	0.1724	1	0.9937	0.5861	0.8123	0.748	0.9935	0.4903	0.4973	-0.2976	-0.2469	0.5544
AngBal	0.093	0.8023	0.0848	0.9937	1	0.5505	0.8204	0.76	1	0.4989	0.5176	-0.2579	-0.2078	0.4849
LatDev	0.4438	0.4453	0.4451	0.5861	0.5505	1	0.6818	0.5897	0.5506	0.8212	0.7761	-0.4757	-0.39	0.7179
LatWidth	0.2078	0.8351	0.2007	0.8123	0.8204	0.6818	1	0.9523	0.8216	0.7138	0.705	-0.0887	-0.0196	0.5128
MaxLatDev	0.129	0.854	0.1207	0.748	0.76	0.5897	0.9523	1	0.7617	0.599	0.6437	-0.1584	-0.1028	0.38
BestFitAng	0.0924	0.8034	0.0842	0.9935	1	0.5506	0.8216	0.7617	1	0.4997	0.5187	-0.2574	-0.2072	0.4843
KyphCobb	0.0345	0.4904	0.0358	0.4903	0.4989	0.8212	0.7138	0.599	0.4997	1	0.8853	0.0321	0.1273	0.4188
KyphFerg	-0.0434	0.612	-0.0491	0.4973	0.5176	0.7761	0.705	0.6437	0.5187	0.8853	1	-0.177	-0.0941	0.3228
LordCobb	-0.3484	-0.1391	-0.347	-0.2976	-0.2579	-0.4757	-0.0887	-0.1584	-0.2574	0.0321	-0.177	1	0.9948	-0.385
LordFerg	-0.339	-0.1008	-0.3373	-0.2469	-0.2078	-0.39	-0.0196	-0.1028	-0.2072	0.1273	-0.0941	0.9948	1	-0.3256
Area	0.6375	0.2195	0.6377	0.5544	0.4849	0.7179	0.5128	0.38	0.4843	0.4188	0.3228	-0.385	-0.3256	1

6.8. Chapter Summary

This chapter has presented three aspects of this research, namely, landmark detection, midline derivation and computation of surface measurements.

The landmark detection algorithm developed is based on surface analysis of the curvature maps using constrained surface fitting. For validation, the results were compared with palpation and an accuracy of 1.10 mm in X and 1.15mm in Y was reported for three landmarks, assuming palpation results as the 'gold standard'. The precision (or repeatability) of landmark detection was determined by multiple scans of the same subject, and reported as 0.58mm in X and 0.79mm in Y.

Based on the detected landmarks, the midline of the back was derived using constrained polynomial fitting. The mean standard error of the fit was estimated as 0.23±0.06mm, demonstrating suitability of constrained polynomial for the subjects examined.

Using the derived midline, various surface measurements were computed and the linear relationships among them were examined via correlation analysis. Several surface measurements correlated strongly with the Cobb angle (e.g. the lateral deviation, projected area in the frontal plane) while others correlated weakly. Comparison with radiographic data is still required to validate the usefulness of the surface measurements.

7. Conclusions and Further Work

7.1. Introduction

This thesis has presented the content and progression of research into automated shape analysis and visualization of the human back. This chapter summarises the conclusions drawn from the research, as well as recommendations for further research.

7.2. Summary and Conclusions

The study was organised into four phases, each with its specific objective as outlined in Chapter 1. Each of these objectives is revisited and the extent to which it has been met is discussed.

7.2.1. Curvature Computation, Error Smoothing and Landmark Detection

All surface measurements and analysis techniques pertaining to spinal deformities rely on the assumption that deformity of the spine will generally result in corresponding deformity of the back surface. The success of any surface measurements method, to a large extent, depends on the detectability of the bony landmarks that protrude to the surface of the back. In this research, methods have been developed to detect these landmarks using curvature analysis of three-dimensional laser scans of the back.

A critical aspect of curvature computation is that the curvature values depend on second derivatives; consequently, they are susceptible to random errors in the original data. In the case of back shape analysis, the problem was further complicated by the mobile and flexible nature of human subjects which introduced subtle random errors during data capture, presenting a major challenge. The problem of errors was addressed through scanner calibration and noise smoothing operations. First, the scanner was calibrated to ensure there were no systematic errors or drifts in the scanned data. The results of the calibration closely matched the accuracy characteristics given by the manufacturer. Second, smoothing techniques were developed to reduce the effect of the remaining random. Two issues with error smoothing were identified, namely, selection of smoothing method and the amount of smoothing to apply. Since the effectiveness of any smoothing method depends on the nature of the noise, a smoothing method was selected based on experimental estimate of the distribution of the random errors. This

was achieved by comparing actual laser data of a smooth machined cylinder with an ideal cylinder with the same dimension and orientation, resulting in a distribution which was well approximated by a Gaussian with mean 0.003mm and standard deviation 0.47mm. The amount of smoothing was determined by comparing the curvature of the actual scanned data of the cylinder computed with increasing amount of smoothing with the true curvature of the cylinder. From these experiments, an optimal smoothing range was found for the data that permitted stable curvature computation and landmark detection.

The landmark detection method developed consists of searching for all local convex and concave regions containing the landmarks, fitting surfaces to these regions and analysing the resulting surfaces to determine their peaks and troughs. A two-level thresholding defined by standard deviation units from the median was used to restrict the search to higher curvature regions. The surface fitting was constrained using inverse square distance weightings for the points within the neighbourhood of the landmark. The peaks of the resulting local surfaces were then determined by analysing their critical points. The landmarks were detected from three curvature maps, namely, mean, Gaussian *and curvedness and the* results were averaged.

Accuracy and precision are important aspects of any measurement process. Accuracy describes the closeness of a measured quantity with its true value while precision describes the repeatability of a set of measurements. The accuracy of the detected landmarks was assessed by comparison with palpation, due to the unavailability of radiographs. Assuming palpation as the standard, accuracies of approximately 1.10mm in X and 1.15mm in Y were obtained. Precision was assessed by scanning one subject twenty times, and comparing several landmarks detected from the datasets, producing 0.58mm in X and 0.79mm in Y. These results demonstrate that back surface landmarks can be detected with sufficient accuracy to allow assessment of spinal and back deformities. The improved accuracy reported in this study can be attributed to several factors, including effective error smoothing; rigorous registration algorithm for surface comparison; direct analysis of multiple curvature measures and constrained least square solution.

7.2.2. Visualization of the Back Shape

The second objective of this research was to develop methods for visualizing the shape of the back. The aim was to develop diagnostic tools for the visual assessment of back deformities. A number of visualization techniques were explored, namely, continuous and discrete colour scales, contour maps and three-dimensional views. The main challenge with the use of colour scales was how to map the colour scale to the curvature data to maximise visual contrast. Mapping or classification methods that were investigated included equal interval, linear and non-linear contrast stretching (i.e. histogram equalization), nested means, quantiles and natural breaks classification. Due to the extreme values in the data (especially the Gaussian curvature map), histogram equalization failed to yield satisfactory results. Quantile classification was found to be more robust and less susceptible to outliers, resulting in visualizations with higher contrast. The visualizations revealed several landmarks and regions of the back which could be very useful for assessment of the back.

Contour maps encode information about asymmetry of the back in a form that is readily recognizable by clinicians, as is evident from the widespread use of moiré topography. With the automated system, contour maps can be generated at varying intervals, allowing the user to compare corresponding contours at different levels of detail to assess asymmetry. The area and volume enclosed by corresponding contours can be calculated and displayed in a simple form. The three-dimensional views allow the user to explore the 3D shape of the back by rotating the surface.

7.2.3. Midline Derivation and Surface Measurements

A third objective of this research was to derive the back midline, which is the closest surface approximation to the spine, and to compute additional surface measurements to quantify the deformity. From the detected spinous process landmarks, the midline was derived using constrained polynomial fitting via the method of Least Squares. Polynomials were used because of their simplicity and the fact that they can easily be constrained to follow the spinal model. To ensure the derived midline passed through the spinal end points (i.e. C7 and S1), the curve fitting was constrained by assigning higher weights to these points. The suitability of constrained polynomials was assessed by analysing the goodness of fit graphically and numerically. The mean square error of the fit for twenty five subjects examined was computed as 0.22±0.05mm, demonstrating the suitability of constrained polynomials for the subjects ex-

amined. With the midline derived, various surface measurements were computed, consisting of linear, angular, area measurements in frontal and sagittal planes, as well as 3D length and volume of the spinal bounding box. Back surface rotation was estimated by the orientations of the surface normals at the spinous process locations. Graphical displays of the midline were also given.

7.2.4. Automated System

The final objective of this research was to develop a software system to automate the tasks involved in analysing, visualization and monitoring back deformities. A software platform has been developed using various tools and the methods developed in this study have been realised by practical implementation. Tools have been developed for importing scanner data, performing smoothing operations, computing curvature maps, creating various visualizations, detecting the landmarks, deriving the back midline and computing the surface measurements. The system has all the functionalities of modern GUI applications, such as zooming, panning and rotating the 3D model of the back. Graphical displays of the derived spine, results of the polynomial fitting including goodness of fit statistics and the various surface measurements are all given by the system. The software has been designed as a generic application for back shape analysis and visualization that can be installed on any standard computer. Most of the algorithms developed in this research have been implemented. The system can readily be deployed at the centre of efficient management of patient data, where patient images, radiographs, back shape measurements, together with other relevant clinical data could be combined and used for decision making.

7.2.5. Concluding Remarks

The novel aspects of this investigation can be summarised as follows: development of effective smoothing methods for reduction of random errors for curvature computation; improved shape analysis and landmark detection algorithms; effective shape visualization techniques; derivation of the back midline using constrained polynomials and derivation of three-dimensional surface measurements. The shape analysis and visualization techniques developed in this study, together with the software platform should facilitate the assessment and monitoring of back deformities, enable the effectiveness of treatment methods to be assessed, and comparison of back shape data to track changes and improve knowledge of skeletal-surface relationships.

7.3. Recommendations for Further Work

7.3.1. Comparison with Radiographs

One shortcoming of this study was the unavailability of radiographs for comparison. Since radiographs establish the true locations of the bony landmarks, it is important to compare the results obtained in this study with radiographic measurements. There are two key aspects that require further validation by radiographic measurements, namely, landmark detection and surface measurements. The landmark detection algorithm developed in this investigation was compared with palpation and an accuracy of less than 1.2mm was reported in both X and Y. However, since palpation is error prone, comparison with radiographic data will provide a useful validation. However, as already pointed out in Section 6.7.8, due to the low resolution and two-dimensional nature of radiographs, this is a difficult task. This was confirmed by Drerup and Hierholzer (1987) who compared their results with radiographs and reported lower accuracies than that obtained from previous comparison with palpation (see Section 6.5). However, the experiment of Drerup and Hierholzer (1987) had one major limitation; namely, no attempt was made to correct for unequal scaling in both systems, which was shown to be a crucial factor in comparing data in different frame of reference. The transformation model developed in Section 4.3.7.3 can be used to improve the registration process in a future radiographic comparison. The second aspect of the study that requires further validation is the derived surface measurements. Due to the low turn-up of the volunteers, only 25 subjects were examined, with badly-balanced gender ration. A longitudinal study involving more subjects is required to validate the methods developed in this study. Such a longitudinal study will also confirm the correlations observed among these measurements.

7.3.2. Effect of Body Type on the Accuracy of Landmark Detection

This investigation has demonstrated that back surface landmarks can be detected with an accuracy of approximately 1.2mm compared with palpation. In general, the detectability of back surface landmarks depends to a large extent, on the characteristics of the body types being examined. Body fat in overweight people can substantially conceal many of the landmarks, making their detectability more difficult. Consequently, it is important to assess landmark detectability using various body types and characteristics such as age, gender, height, weight and BMI. This will require a longitudinal study involving several subjects with different body

characteristics, which could not be carried out in this investigation due to lack of enough data and difficulty in recruiting subjects.

7.3.3. Automated System

As described earlier, the system developed in this research has the functionalities of modern GUI applications, and most of the algorithms developed have been implemented. The main areas of further work relating to the software system are as follows:

- Automated communication between the scanner and the software system. The current implementation relies on the user to export back shape data from the scanner in the form of XYZ coordinates into a text file. The text file is then imported into the system using the tools that have been developed. To facilitate automated processing, the system can be programmed to communicate directly with the scanner. Minolta exposes a set of software development interfaces that can be used to automate the communication between VI-900 and the developed software. Consultation with Minolta also revealed that they are willing to corporate in the development of such a system. It is therefore recommended that this aspect of the system be carried, as it will facilitate further experimental studies such as those mentioned above.

- Patient database and decision support components. The system can be linked to a database system that stores patient data (e.g. history, radiographs, surface measurements, etc). A decision-support component can then be built to evaluate the different data of a subject to provide relevant information for clinical decision making.

7.3.4. Other Application Areas

The shape analysis and visualization techniques developed and implemented in this research should be applicable to other biomedical measurement problems, in particular, the following:

- Effect of treatment methods and reconstructive surgeries on various parts of the body can be explored using the analysis and visualization techniques developed in this study without the need for physical examination of the patient. Breast reconstruction offers restoration of breast symmetry to women whose breasts have been removed due to cancer or other disease by creating breast mounds that is similar in size, shape, contour, volume and position. The shape analysis and visualization techniques developed

in this study will permit models of breast moulds to be created, analysed, visualized and applied to a torso model of the subject before the actual implants. Another example is wound reconstruction where the measurement and visualization of irregularly shaped wounds are required. Indeed, the methods developed here could benefit the entire cosmetic surgery industry.

- Facial landmarks detection for face recognition. Face recognition systems attempt to identify a person's face by matching the features of the current image with those of a known image. The curvature analysis and landmark detection methods developed in this study can be used to locate the facial landmarks to facilitate recognition.

- Seated posture analysis, such as design of wheelchair seats for the disabled where the objective will be to provide measurements required for comfort design. The midline and the surface measurements can be used to derive information required for the design.

- This methods developed here system could also be beneficial to the healthcare and sports therapy sectors, for example, in the measurement of muscle symmetry.

Appendices

Appendix A: System Functionalities and Screen Shots

The developed software has standard functionalities that are found in modern GUI applications, including the following:

- Basic visualization tools: zooming in and out, panning, etc.
- Tools to display attributes of points or pixels
- Tools for onscreen measurements of distances, angles, areas, etc
- 3D visualization: rotating, scaling, and transforming surfaces with mouse events
- Importing scanner data from TXT file into vector and raster data structures
- Displaying the attribute table associated with each data layer
- Performing SQL query on the database associated with geometric objects
- Tools for converting data between different formats
- Various tools for creating the various visualization discussed in Chapter 5
- Tools for automatically detecting landmarks from curvature maps, deriving the midline and computing the various surface measurements discussed in Chapter 6

The following screen shots demonstrate the main interfaces and functionalities of the system.

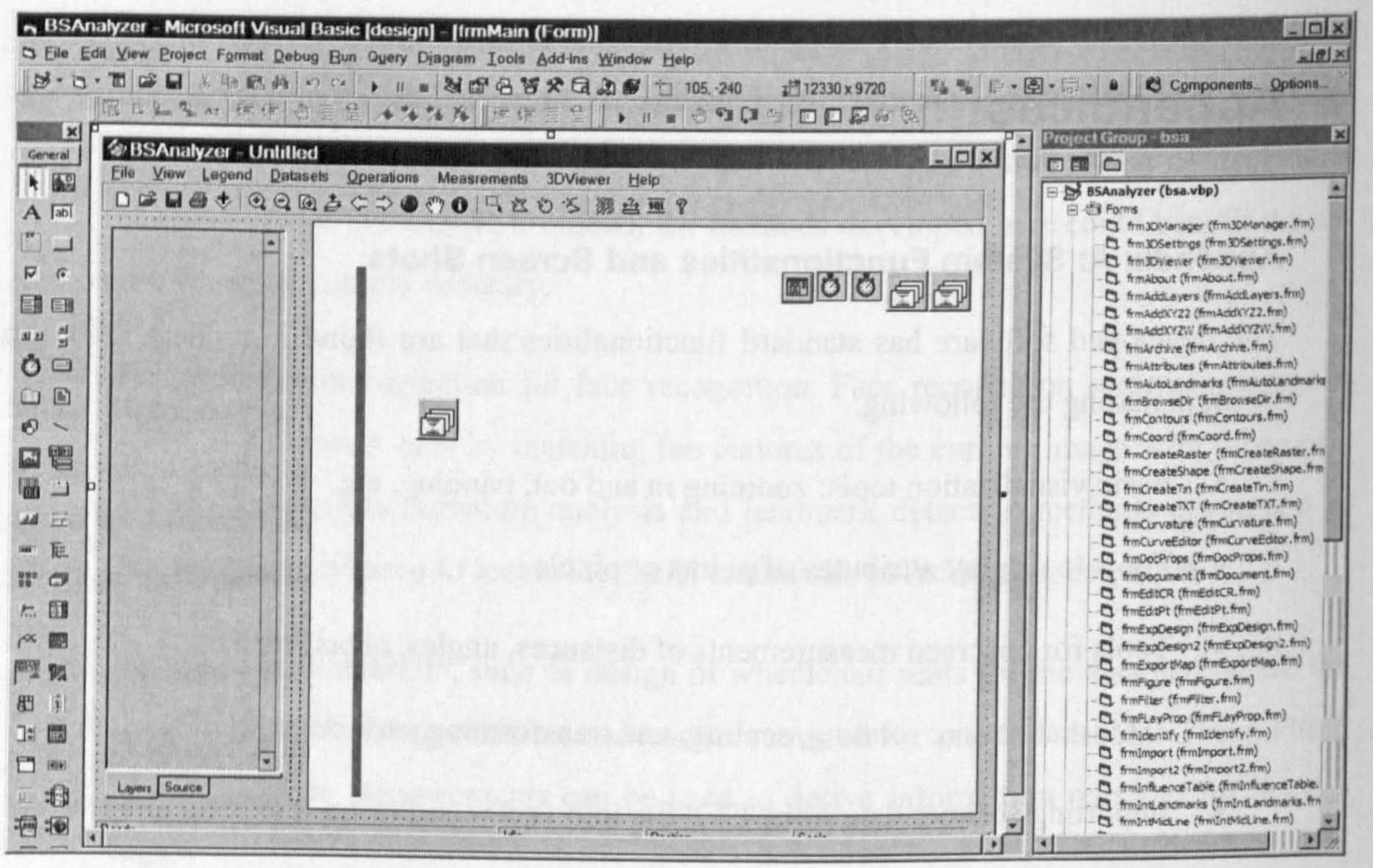


Figure A-2: Main system interface at design time in Visual Basic

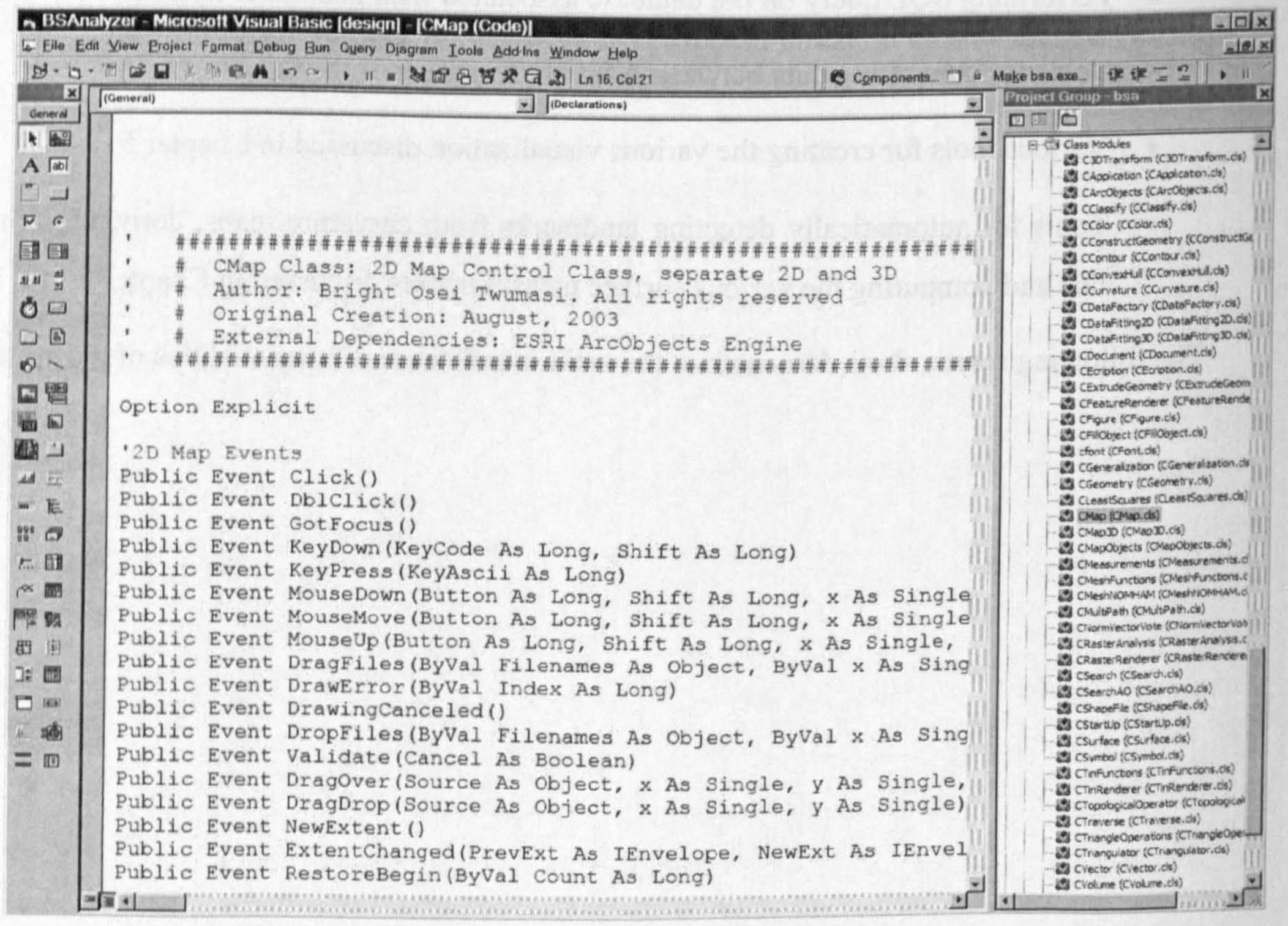


Figure A-3: Program development in Visual Basic – Code Snapshot

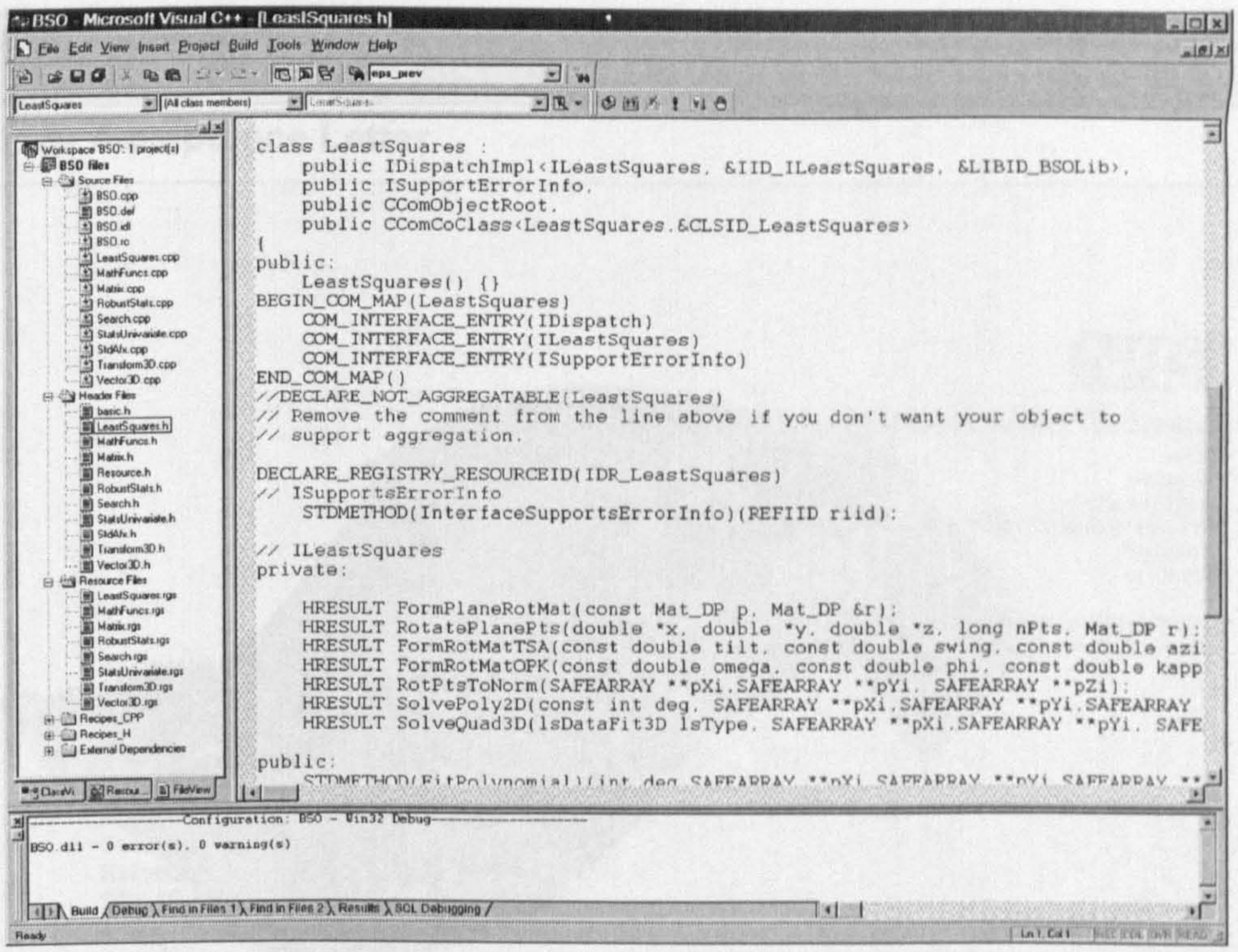


Figure A-4: Development of COM components in Microsoft Visual C++

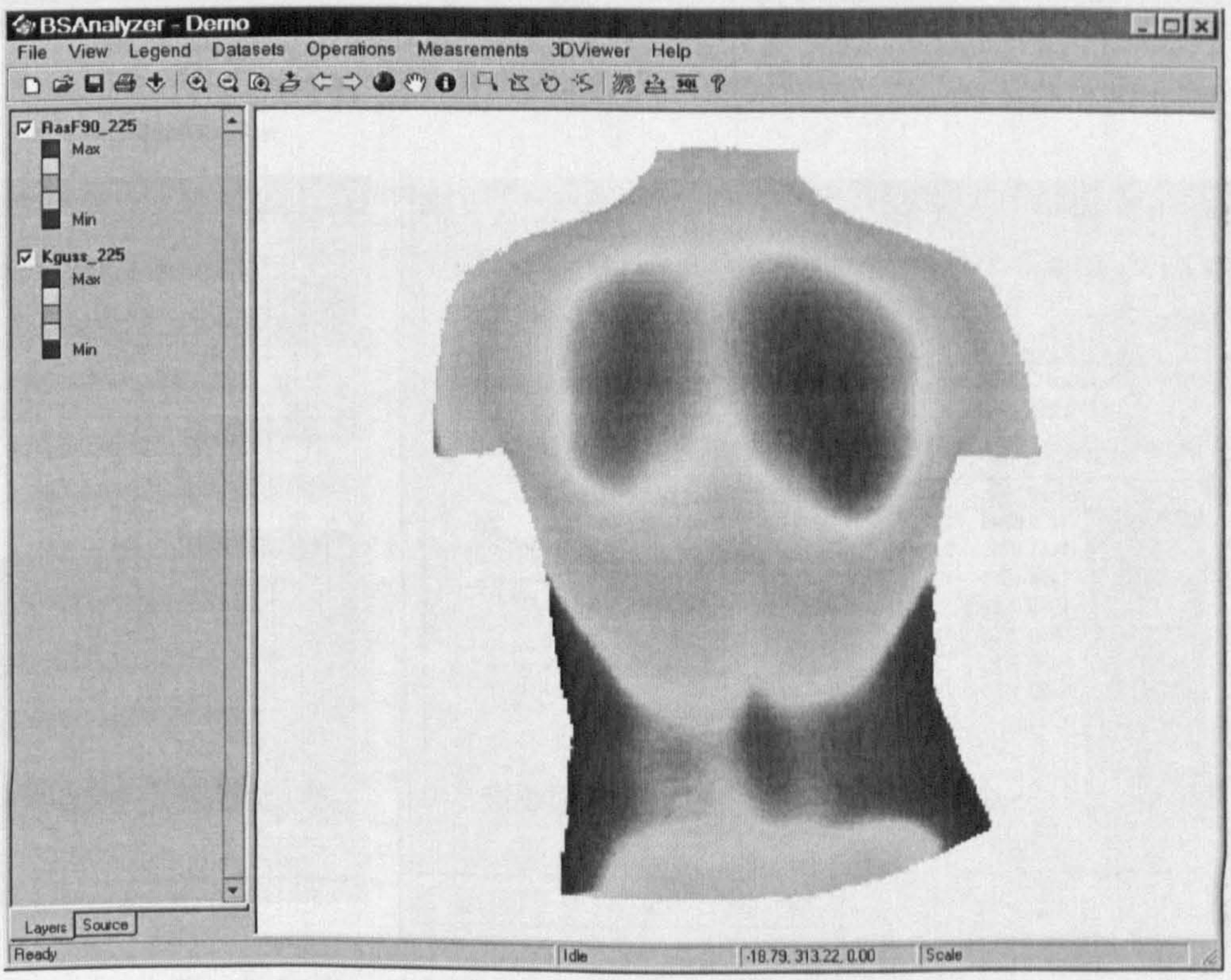


Figure A-5: Main program interface at run time

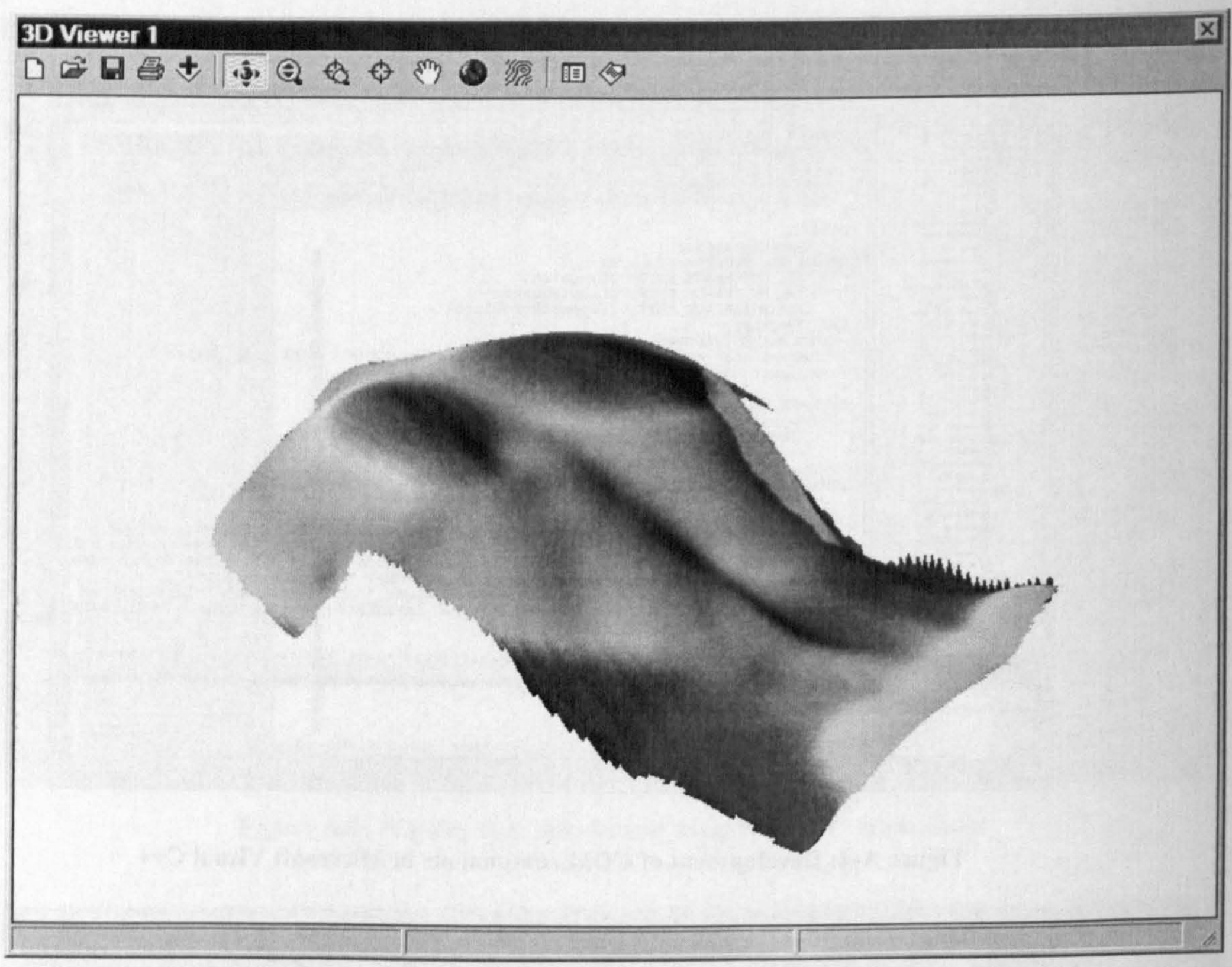


Figure A-6: 3D interface of the system

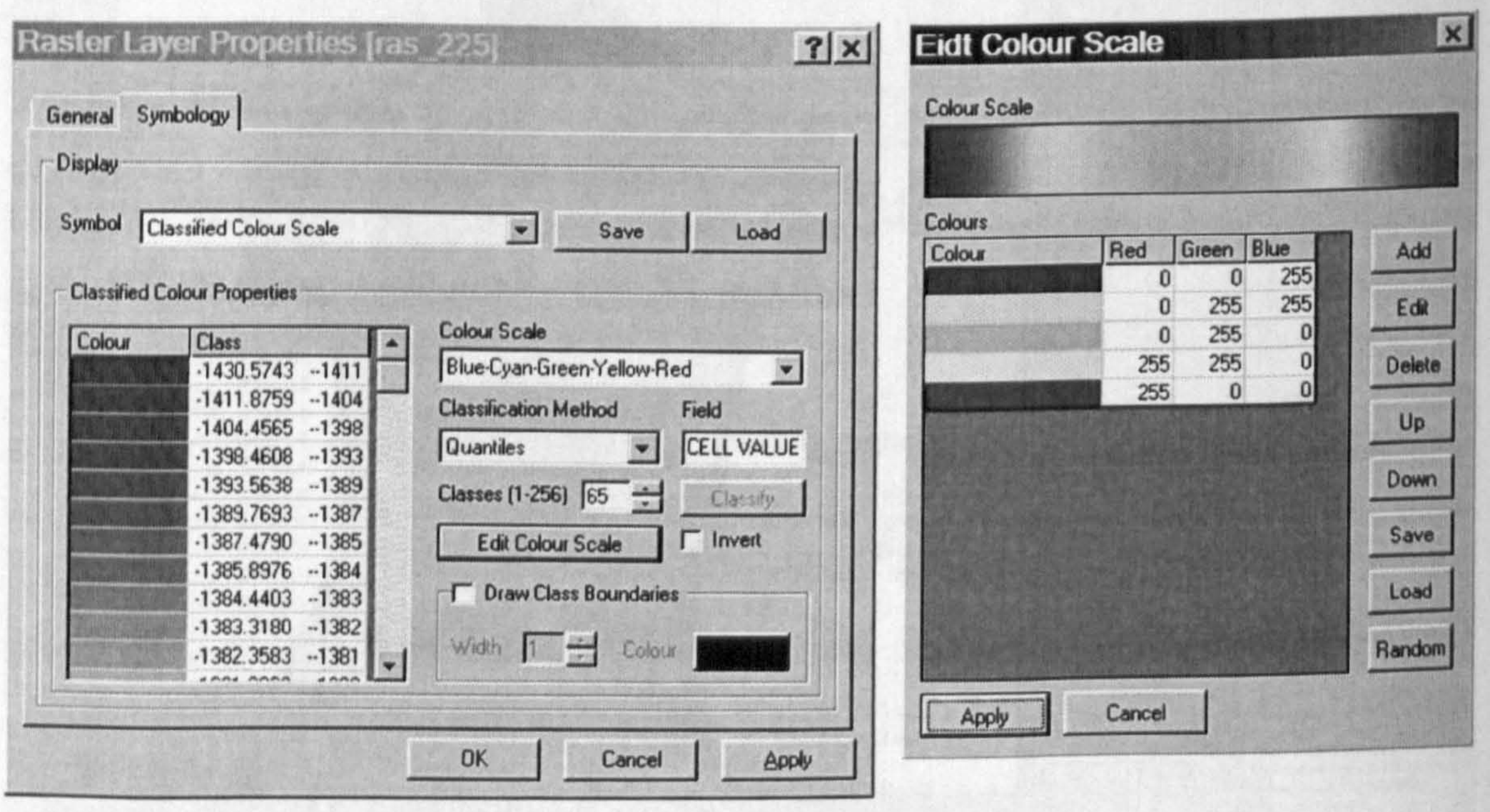


Figure A-7: Dataset properties window and colour editor

Appendix B: Ethics Approval, Subject Information and Consent Forms

B-1: Ethics Acceptance Letter

NHS

Salisbury and South Wiltshire Research Ethics Committee
Unit B
Valentines
Epsom Square
White Horse Business Park
Trowbridge
BA14 0XG

Telephone: 01225 756752
Facsimile: 01225 754648

11 August 2005

Dr Steven E Crook
Clinical Scientist and Deputy Head of Department
Department of Medical Physics, Salisbury Healthcare NHS Trust
Department of Medical Physics
Salisbury District Hospital
Salisbury
SP2 8BJ

Dear Dr Crook

Full title of study: **The automated extraction of topographic features from the human dorsal view to determine extent of postural abnormality and suggest correction.**

REC reference number: **05/Q2008/46**

The Research Ethics Committee reviewed the above application at the meeting held on 27 July 2005.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation. The Committee would like to commend you for a very well presented project.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Application		14 July 2005
Application	3	14 July 2005
Investigator CV		14 July 2005
Letter from Sponsor		11 July 2005
Copies of Advertisements	1	14 July 2005
Participant Information Sheet	3	14 July 2005
Participant Consent Form	3	14 July 2005
Other	1	14 July 2005
Other	1	14 July 2005

Management approval

The study should not commence at any NHS site until the local Principal Investigator has obtained final management approval from the R&D Department for the relevant NHS care organisation.

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

REC reference number 05/Q2008/46 please quote this number on all correspondence

With the Committee's best wishes for the success of this project,

Yours sincerely

Kirsten Peck

Kirsten Peck
Coordinator
Salisbury and South Wiltshire Research Ethics Committee

Enclosures:

Attendance at Committee meeting on 27 July 2005
Standard approval conditions
SF1

An advisory committee to Avon, Gloucestershire and Wiltshire Strategic Health Authority

B-2: Volunteer Information Regarding Data Collection

Data collection for surface measurement system for analysis and monitoring spinal and back deformities

Dear Volunteer

Thank you for your interest in this project. The information that you are reading is designed to explain the purpose of this project and why we need your help. In addition we hope to address any questions that you may have and reassure you as to the safety and confidentiality of your contribution.

Introduction to the Problem

Back pain and related symptoms has a major impact on the UK workforce. People with a curve in their back have an increased risk of back pain. This may start with personal discomfort, but can lead to increased back pain, disrupting productivity; often requiring expensive and prolonged treatment. There are also cosmetic issues when the curve, called a *scoliosis*, causes distortion of the whole trunk (see Figure 1). The assessment technique presently used relies on a skilled operator to interpret and assess images of the back and legs that have been enhanced by a shadow pattern on the skin. This procedure is very subjective. X-rays are often prescribed for people identified as having spinal curves but they pose a small health risk which is increased if the deformity is to be monitored over time.

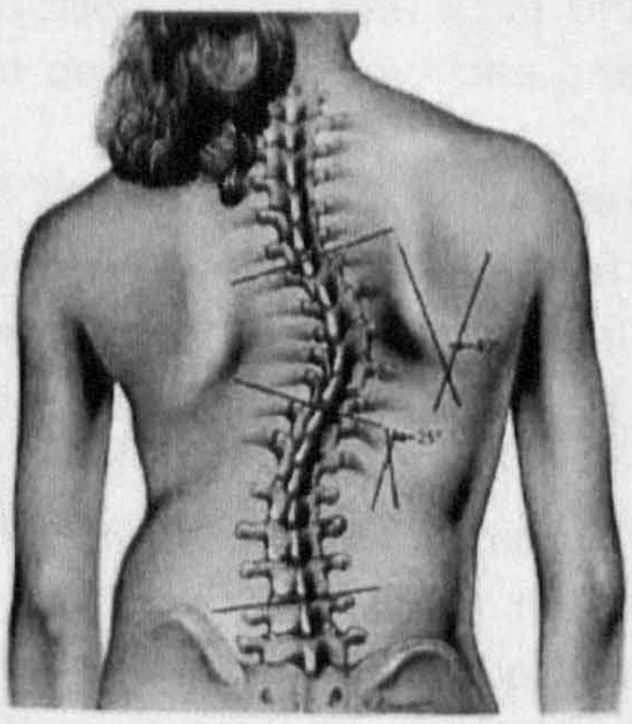


Figure 1: A picture of a girl with Scoliosis

This project was setup to investigate whether surface measurements methods like 3D Laser imaging could be use to diagnose spinal and back shape deformity. 3D laser scanning is fast, accurate, safe, non-invasive and does not carry with it the hazard associated with X-rays. The data obtained from laser scanner will be processed using computer algorithms to extract relevant information for clinical decision making.

Need for Data Collection

Preliminary work has determined that certain key features of the *back and upper buttocks* can be characterised on the limited data set available to the researcher. It is apparent that fine tuning of the feature detection can improve its performance on one image but may not offer the same improvement on another. The human subjects that the system will have to cope with will come in all shapes and sizes. It is therefore essential that a wider range of data be collected in order to ensure that the final system can be applied to the general population who will benefit.

Data Required

- Person information—The following information will be collected for each person will include age, sex, height, weight, occupation, work type, back pain symptoms if any.
- Laser scan data—Five scans will be taken for each person, 3 scans without markers and 2 scans with markers. These markers are 5 small sticky tags placed on the back by Dr Steve Crook for reference purposes. New markers will be used for each volunteer.

Data Collection Procedure

The scanning process is simple: the individual stands about 1-2metres from the scanner which will acquire the image within 3 seconds. The scanner cannot work through clothing so volunteers will have to undress their back for the scan. However, we will provide a hospital gown that will be worn to cover the front part of the body. There will be two members of the research team (male) present plus a chaperone if it is a lady being scanned. The researchers will be Dr. Steven Crook who is a medical expert from Salisbury District Hospital (NHS) and Bright Osei Twumasi (PhD student). *The parts of the body being scanned are the back and upper buttocks regions.* Each scan takes only 3 seconds and we hope to keep each volunteer for no more than 5 minutes.

Technology and Safety issues

The scanning will be performed with a Minolta non-contact Laser Scanner (VI-900, CLASS-2, 690nm). According to Health and Safety Manual produced by the US National Institute of Environmental Health Sciences, Laser Class 2 is safe in operation:

> Class II - denotes low power visible lasers or laser systems which because of the normal human aversion responses (<0.25s), do not normally present a hazard but may present some potential for hazard if viewed directly for extended periods of time (like many conventional light sources)[1].

Minolta VI-900 has been used for scanning various parts of the human body. Examples include: scanning the human breasts (Langmack et al, 2000), face and head scanning (see Minolta website for example medial applications [3]).

Anonymity of Volunteers and Data Usage

The anonymity of all volunteers will be completely ensured. To achieve this, data collected for each person <u>WILL NOT</u> include any personal information such as name or address. In addition, all data files will be saved under arbitrary names such as 'File 1', etc. The collected data will <u>ONLY</u> be used for the purposes of this research.

Thank you for your contribution,

The Project Team
Bright Osei Twumasi
Dr Steve Crook (Co-Supervisor)
Dr Martin Lefley (Co-Supervisor)

References

[1] US National Institute of Environmental Health Sciences, Health and Safety Manual, http://www.niehs.nih.gov/odhsb/manual/home.htm
[2] Langmack, K.A. et al (2000), Practical intensity modulated radiotherapy for improving dose homogeneity in tangent breast irradiation. IOS, Birmingham, May 2000, Web: http://www.minolta-3d.com/applications/eng/breast.html
[3] Minolta website: http://www.konicaminolta-3d.com/applications/

B-3: Subject Consent Form

CONSENT FORM

I confirm that I have received and read the details of the research experiment being undertaken on 3 November 2005 and understand the purpose for which this data is being collected.

I am a willing participant in the experiment and volunteer for an image of my back to be made using the Minolta VI-900 Laser Scanner.

My consent is given for the data collected to be used for research purposes only in connection with the PhD project currently entitled: "A surface measurement system for analysis and monitoring of spinal and back deformities".

Please print your name and sign below:

Name:.....

Subject's Signature:.....

Researcher's Signature:.....

Appendix C: Experimental Results

Table C-1: Comparison of marker with curvature coordinates: C7 Landmark

	C7 Landmark							
	Marker		Mean		Gaussian		Curvedness	
Subject	X	Y	X	Y	X	Y	X	Y
1	117.2161	277.4766	104.1252	295.8676	104.4328	296.2114	105.1059	296.9769
2	-16.0672	350.7410	-13.0621	353.1176	-13.3660	352.9869	-13.8340	352.5286
3	3.9486	374.5987	5.4788	379.9731	5.5560	380.0766	5.6048	380.2145
4	49.1898	359.0524	49.3176	361.3703	49.2017	361.4492	48.8292	362.8847
5	62.3831	240.0596	62.3573	242.7546	62.4628	242.5665	62.5380	242.4431
6	72.8813	297.9698	76.3015	301.1226	76.0914	301.0238	75.9903	300.8780
7	59.6311	263.3015	69.2781	261.7227	69.2841	261.7200	69.2996	261.7309
8	-94.9169	282.3878	-77.7349	284.2188	-77.7203	284.1900	-77.7114	284.1744
9	-31.5186	310.1729	-34.6812	314.6903	-34.6976	314.6777	-34.6892	314.6778
10	-1.7271	254.5057	-1.8529	259.2171	-1.8755	259.2082	-1.9023	259.1481
11	1.6706	270.6169	-7.4190	273.9843	-7.3873	273.8082	-7.3733	273.5465
12	-27.3529	266.8883	-25.1020	268.5803	-25.0414	268.6031	-25.0056	268.6162
13	-1.2310	267.0826	0.5740	269.5772	0.6085	269.5163	0.6256	269.4370
14	-14.6983	271.4948	-8.3844	267.9566	-8.3433	267.8359	-8.3125	267.6732
15	-14.7840	269.8635	-25.2748	268.8143	-25.2632	268.7529	-25.2436	268.6553
16	-15.3687	258.8470	-16.5699	266.3303	-16.6035	266.3515	-16.6209	266.3475
17	-2.3025	269.2527	-27.6716	258.0495	-27.6637	257.9910	-27.6975	258.0974
18	11.6914	264.3965	14.6477	270.5806	14.6751	270.5749	14.6771	270.5669
19	17.8108	268.2011	22.5621	265.7024	22.5938	265.6094	22.6086	265.4936
20	24.0377	266.6423	26.3733	264.0218	26.3842	263.8803	26.4122	263.6743
21	27.0314	143.3286	29.7186	145.9696	29.6740	145.8347	29.6539	145.6416
22	32.7897	151.0389	34.4580	151.7286	34.4533	151.7514	34.4480	151.8007
23	34.7865	231.0443	43.5101	236.8885	43.3883	236.7073	43.2993	236.3594
24	58.4540	24.8035	59.0471	28.1679	59.0558	28.1220	59.0656	28.0355
25	113.0531	149.8119	126.1696	146.1624	126.2444	146.241	126.3156	146.3214

Table C-2: Comparison of marker with curvature coordinates: Left PSIS landmark

	Left PSIS Landmark							
	Marker		Mean		Gaussian		Curvedness	
Subject	X	Y	X	Y	X	Y	X	Y
1	67.3343	-166.5801	57.4961	-163.4530	57.5804	-163.6244	58.1918	-164.3334
2	-71.3149	-132.7183	-75.3118	-131.2625	-75.8641	-130.7535	-76.9509	-130.5034
3	-45.0249	-118.3190	-46.9125	-122.2182	-46.6069	-121.9943	-46.2232	-121.7850
4	-13.2377	-114.2452	-14.6187	-119.0072	-14.6064	-119.0867	-14.4579	-119.1363
5	16.8254	-175.5168	14.5925	-176.5323	14.3956	-176.4777	14.1961	-176.4205
6	14.6993	-162.4589	22.8212	-159.8046	23.0141	-159.7345	23.1680	-159.7174
7	-15.7908	-143.6774	-8.7947	-146.4558	-8.3943	-146.2417	-7.7935	-146.2183
8	-153.8745	-190.2485	-137.9146	-189.2326	-136.7516	-189.3578	-136.8179	-190.2942
9	-102.7379	-123.8262	-105.9316	-120.6510	-106.6162	-120.6634	-108.7441	-120.7756
10	-46.8649	-231.4667	-44.7346	-232.2378	-44.5752	-232.5283	-43.8736	-232.6314
11	-62.5731	-202.8060	-61.3191	-201.3493	-60.5217	-202.0168	-55.8872	-205.5017
12	-76.9640	-197.1423	-81.2278	-199.3960	-80.4777	-199.3469	-78.6786	-200.5377
13	-66.3003	-207.6578	-61.8828	-200.1144	-61.8828	-200.1144	-61.8828	-200.1144
14	-66.0990	-213.3382	-59.7568	-217.8105	-59.5207	-217.7816	-59.1691	-217.7628
15	-53.3523	-207.4049	-68.3184	-220.5379	-68.3184	-220.5379	-68.3184	-220.5379
16	-66.1997	-211.1948	-68.0095	-213.1664	-67.4794	-212.0935	-67.4794	-212.0935
17	-51.9613	-218.0078	-77.5086	-212.0466	-77.0655	-211.4329	-77.0655	-211.4329
18	-29.4813	-208.1473	-33.4931	-209.2525	-32.7166	-208.5044	-32.7166	-208.5044
19	-22.9876	-215.3139	-27.0831	-213.9421	-26.8071	-213.1638	-26.9260	-212.1488
20	-21.3475	-218.6510	-25.8678	-216.4383	-25.8209	-215.8584	-25.9201	-214.5933
21	4.4979	-339.3927	7.0301	-339.8115	6.8549	-340.1189	6.6680	-340.4133
22	15.0538	-320.6202	14.9859	-323.1441	15.0057	-323.1547	14.8384	-323.3810
23	10.0572	-228.6803	11.0895	-228.3671	11.0615	-228.3692	11.0773	-228.3629
24	40.2966	-374.7460	41.9642	-376.6701	41.2995	-376.3719	40.7426	-376.1744
25	36.7598	-305.0357	43.3693	-302.9504	43.2820	-302.7518	43.0103	-302.3633

Table C-3: Comparison of marker with curvature coordinates: Right PSIS landmark

	Right PSIS Landmark							
	Marker		Mean		Gaussian		Curvedness	
Subject	X	Y	X	Y	X	Y	X	Y
1	147.3282	-175.3103	139.9582	-171.3010	139.9939	-171.2365	140.0198	-171.1970
2	42.4853	-128.6121	39.8861	-128.7873	40.3369	-128.9066	41.9935	-129.1151
3	46.5450	-124.6779	47.3540	-128.9808	47.5331	-129.0686	48.0886	-128.5115
4	100.3903	-113.5572	99.6857	-118.2228	99.2144	-118.2669	98.7639	-118.2535
5	112.5570	-175.1901	111.2536	-176.5743	111.2997	-176.3913	111.3488	-176.2270
6	138.2655	-163.8205	145.9701	-160.1093	145.8220	-160.4081	145.2140	-160.5856
7	121.0145	-145.0352	128.5905	-148.5756	128.6725	-148.6128	128.8050	-148.7849
8	-50.9892	-197.6490	-37.0321	-197.8450	-38.4569	-194.9021	-38.0940	-195.3401
9	9.5664	-122.0055	6.8996	-121.5690	7.4490	-122.6797	3.1354	-121.7226
10	24.0751	-233.8416	26.7898	-233.8366	27.0499	-233.8977	27.3626	-233.9867
11	23.9458	-205.7038	25.7839	-202.0427	24.4710	-203.0571	21.1221	-204.7833
12	7.5915	-198.0414	7.2873	-197.2581	7.9472	-197.0128	8.3517	-196.4370
13	23.1572	-210.1955	25.0698	-203.0717	24.8753	-203.2942	24.4460	-203.5255
14	3.6924	-219.5874	10.3091	-224.0471	10.1199	-223.9962	9.8535	-223.9136
15	14.9837	-213.1182	-1.6567	-229.7181	-1.4711	-229.8333	-1.2590	-229.8992
16	-1.6412	-216.8443	-1.9720	-218.8038	-2.1437	-218.8127	-2.4705	-218.7608
17	16.4653	-229.8292	-12.3135	-223.3574	-12.5013	-222.7920	-13.1377	-221.9041
18	37.2503	-215.0671	34.4443	-215.5880	34.3640	-215.6055	34.2495	-215.6083
19	60.9863	-211.4174	56.1153	-210.5411	55.0978	-210.7615	52.7178	-211.2141
20	51.7288	-222.3694	46.7961	-220.9223	46.7241	-220.9470	46.6513	-220.9565
21	84.09582	-331.275	87.14055	-331.486	86.69519	-330.53	86.53885	-330.106
22	101.3199	-327.277	101.5641	-325.199	101.6196	-325.53	99.1292	-325.448
23	101.3505	-228.45	102.7844	-228.968	102.8525	-228.969	102.9234	-228.976
24	100.518	-376.275	100.2004	-378.294	100.2197	-377.412	100.3754	-376.457
25	155.5925	-306.23	160.8852	-305.986	160.8981	-305.989	160.8469	-306.042

Table C-4: Coordinates of registration points (G1-G4) for the two datasets

		Dataset 1 (Landmark + Registration Pts)				Dataset 2 (Only Registration Pts)			
Subject		G1	G2	G3	G4	G1	G2	G3	G4
Subject 1	X	-37.12537	219.4589	-8.88462	186.7085	-22.7183	229.4091	5.617006	193.2408
	Y	247.1235	236.176	-95.3976	-87.2106	233.4544	218.4557	-100.535	-95.1582
Subject 2	X	-166.3304	140.1374	-116.154	66.14103	-167.781	137.9188	-114.363	67.40671
	Y	256.2174	255.3299	-15.7661	-11.5357	252.1105	255.2401	-18.181	-11.4669
Subject 3	X	-94.6839	83.70362	-94.4499	106.1563	-93.8236	81.00121	-92.6822	104.2927
	Y	291.8804	281.0836	-8.52961	-10.8232	287.7138	277.6917	-6.98899	-8.51656
Subject 4	X	-96.38848	153.4684	-29.1325	104.5694	-93.8056	152.2885	-27.5905	104.2865
	Y	301.3449	299.859	26.73748	3.278807	300.4957	298.6463	29.51576	6.25844
Subject 5	X	-62.25797	170.8501	-17.73	137.346	-60.7153	170.371	-15.0192	137.8675
	Y	205.9717	190.5424	-102.934	-98.6724	202.7024	188.2333	-102.455	-97.2813
Subject 6	X	-91.82621	233.9404	-6.95713	161.2374	-95.8453	229.3457	-13.6994	154.0859
	Y	229.6909	231.1156	-59.2867	-55.4109	228.1798	226.5907	-61.3597	-59.2535
Subject 7	X	-77.21065	206.8432	-32.7671	168.7464	-85.8261	196.9384	-40.0207	160.8764
	Y	209.7824	193.5929	-68.1369	-68.2321	210.7739	196.2922	-66.5645	-64.863
Subject 8	X	-209.1253	43.18889	-174.876	7.412749	-225.453	25.95366	-189.142	-7.35924
	Y	238.286	233.7038	-76.2782	-76.5999	235.709	232.8534	-77.585	-76.6813
Subject 9	X	-165.5099	106.5335	-149.952	52.63502	-161.894	109.4146	-146.388	55.3513
	Y	242.1706	222.6229	-41.5689	-48.9636	237.9052	218.1752	-44.6499	-52.4948
Subject 10	X	-125.0038	108.8591	-98.2335	92.76987	-124.176	107.1744	-99.362	89.81152
	Y	160.979	153.1701	-139.04	-139.811	157.8912	148.9954	-139.278	-141.067
Subject 11	X	-128.3253	106.5387	-93.4729	55.30323	-120.439	113.125	-92.7134	55.87132
	Y	211.4203	177.1546	-112.025	-108.785	210.5888	171.4207	-112.719	-112.917
Subject 12	X	-138.4315	93.08801	-112.192	35.74161	-139.701	91.43764	-112.237	35.73747
	Y	221.1797	175.1378	-116.327	-110.613	220.334	174.0411	-118.201	-111.432
Subject 13	X	-121.5094	115.6726	-94.0389	56.7343	-124.562	114.9806	-96.702	55.43461
	Y	207.0664	170.5311	-113.39	-110.24	203.9415	167.0655	-119.547	-116.632
Subject 14	X	-128.2773	96.54027	-102.171	47.44449	-134.579	90.04663	-108.602	40.94533
	Y	207.9045	167.0316	-123.311	-117.635	211.543	170.6787	-119.07	-113.456
Subject 15	X	-143.3184	88.76761	-111.644	35.2437	-128.205	97.44795	-95.1317	49.72784
	Y	206.4048	162.3289	-124.121	-119.109	207.3596	168.162	-108.174	-105.871
Subject 16	X	-138.9832	91.99573	-110.634	39.03134	-135.427	91.22737	-108.171	38.87591
	Y	208.4096	166.1069	-115.996	-111.155	202.2479	160.4945	-115.97	-111.414
Subject 17	X	-147.8769	67.24769	-121.129	29.05956	-126.435	96.66809	-97.5211	58.29395
	Y	196.708	163.6056	-118.912	-113.621	205.2219	171.8982	-121.776	-115.524
Subject 18	X	-103.4046	118.5982	-74.1699	75.24016	-104.319	115.5858	-71.1433	76.22273
	Y	216.0216	175.2866	-112.931	-108.611	209.2395	171.7289	-114.949	-108.994
Subject 19	X	-95.92059	126.5473	-65.8254	82.24883	-100.467	124.7792	-64.2408	85.00691
	Y	202.851	157.6524	-126.729	-123.971	202.6952	161.2513	-128.9	-123.53
Subject 20	X	-96.15496	136.4105	-65.8383	83.12543	-98.774	136.4552	-63.9932	86.84823
	Y	203.4553	166.3173	-127.502	-123.777	203.7776	169.7868	-129.615	-123.758
Subject 21	X	-80.8916	142.3046	-39.2341	120.9677	-82.8455	139.2609	-41.4192	117.823
	Y	61.40132	56.56577	-247.113	-242.336	59.10379	54.49698	-247.183	-242.465

		Dataset 1 (Landmark + Registration Pts)				Dataset 2 (Only Registration Pts)			
Subject		G1	G2	G3	G4	G1	G2	G3	G4
Subject 22	X	-50.99592	152.1527	-17.6208	136.7157	-52.6193	150.7175	-18.6736	135.5231
	Y	62.07431	56.6194	-253.605	-249.874	61.10012	56.12783	-254.809	-250.572
Subject 23	X	-73.36669	143.7562	-30.0434	127.6296	-78.1425	138.8739	-31.782	126.0097
	Y	174.2541	168.481	-139.125	-149.525	173.4972	165.7111	-139.734	-148.615
Subject 24	X	-50.3374	165.5925	-18.894	140.8434	-48.733	164.1367	-17.032	140.2889
	Y	-89.93621	-93.8899	-296.434	-305.504	-92.1604	-95.2366	-295.897	-304.234
Subject 25	X	13.82722	220.792	40.25789	168.7727	1.061533	210.5025	32.24497	161.5173
	Y	48.42283	29.8657	-185.699	-190.573	49.35711	33.90919	-186.695	-189.261

Table C-5: Transformation parameters and RME errors (mean RMSE: 1.33×10^{-08})

Subject	s	ω	Φ	κ	Tx	Ty	Tz	RMSE
1	0.97314	3.2E-25	1.05E-25	-0.00858	13.14214	-8.34542	-1.4E-23	1.01E-10
2	0.99677	1.87E-25	7.38E-25	0.01279	1.327136	-0.99189	-3.8E-23	2.7E-08
3	0.980915	1.35E-27	7.37E-27	0.003667	0.016806	1.713946	-1.9E-25	4.24E-12
4	0.98598	1.43E-28	-1.7E-28	-0.00034	1.077229	3.147466	-2.8E-26	5.22E-15
5	0.988217	-8.9E-26	3.82E-25	0.004059	1.942143	-0.58146	2.58E-23	5.64E-08
6	0.99893	2.62E-26	-3.7E-25	-0.00933	-6.35037	-2.2013	-3E-23	4.53E-09
7	0.996859	-1.9E-28	5.21E-28	0.006557	-7.76464	1.935077	4.72E-26	1.94E-15
8	0.996704	-2.5E-28	6.38E-26	0.007098	-15.3626	-0.34931	-5.3E-24	2.01E-08
9	0.996511	-3.5E-27	-1.2E-26	-0.00097	2.967051	-3.54268	7.79E-25	2.95E-09
10	0.990074	7.14E-27	-2.3E-26	-0.00528	-1.33581	-2.12977	6.15E-26	1.13E-08
11	0.997969	-1.5E-26	3.74E-26	-0.02171	3.007275	-3.07782	9.54E-26	4.4E-08
12	1.000467	1.4E-27	-2.4E-27	0.00362	-0.57508	-1.06843	1.35E-26	7.2E-12
13	1.009833	-4.9E-29	2.12E-28	-2.4E-05	-1.82169	-5.16368	-4E-28	5.26E-09
14	0.998482	-3.5E-28	-3.5E-27	-0.00026	-6.47311	3.971408	8.78E-26	1.36E-08
15	0.966251	4.89E-26	-3.8E-25	0.009455	12.87803	10.35245	1.08E-23	1.85E-10
16	0.981143	-6.2E-27	-3.1E-26	-0.00147	0.661696	-2.34947	1.14E-24	5.58E-08
17	1.036456	5.99E-27	9.28E-26	0.004088	27.635	2.028592	-4.4E-24	3.12E-08
18	0.987122	3.02E-26	5.22E-25	0.012642	0.602911	-2.686	5.5E-25	4.24E-09
19	1.008355	1.32E-28	-2.9E-28	0.017667	-0.10039	-0.00609	-7E-27	8.46E-15
20	1.009417	8E-28	-1.5E-27	0.013674	1.023143	-0.05025	-4.5E-26	1.77E-13
21	0.993593	3.92E-29	-1.1E-28	-6.8E-05	-2.34629	-1.73404	-4.6E-28	4E-14
22	1.000615	6.48E-29	4.98E-29	0.001648	-1.51831	-0.87357	8.99E-27	1.43E-14
23	0.995861	4.02E-25	2.54E-25	0.006811	-2.98764	-1.03496	5.09E-24	7.71E-10
24	0.986369	-9.3E-26	3.03E-25	0.003383	0.517135	-3.31758	-8.9E-26	5.44E-08
25	1.009335	-5E-28	5.18E-29	0.0171	-11.8855	0.093677	-3.1E-26	2E-13

Table C-6: Information on subjects used in the study for comparing landmarks detected from curvature maps with palpation. All subjects are males, unless indicated otherwise.

Subject	Age	Height (cm)	Weight (kg)	BMI (kg/m^2)
1	30	170	70	24.22
2	48	170	80	27.68
3	31	202	107	26.22
4	26	182	78	23.55
5	27	176	75	24.21
6	27	172	65	21.97
7	30	182	80	24.15
8	24	168	80	28.34
9	30	175	76	24.82
10	27	168	72	25.51
11	26	168	70	24.80
12	24	170	75	25.95
13	27	169	74	25.91
14	24	175	72	23.51
15	30	176	78	25.18
16	29	174	70	23.12
17	25	175	70	22.86
18	26	176	80	25.83
19	27	167	65	23.31
20	25	172	68	22.99
21	24	168	70	24.80
22	24	174	75	24.77
23	27	170	70	24.22
24 (female)	30	170	64	22.15
25 (female)	27	167	58	20.80

Appendix D: Colour Conversion Formulae

D-1: Conversion from RGB space to XYZ space

Given an RGB colour whose components are in the nominal range [0.0, 1.0] and whose gamma is γ, the equations for converting from RGB space to XYZ space are:

$$[X \quad Y \quad Z] = [r \quad g \quad b][M] \tag{D-1}$$

where if the RGB system is sRGB, then

$$r = \begin{cases} R/12.92 & R \le 0.04045 \\ ((R+0.55)/1.055)^{2.4} & R > 0.04045 \end{cases}$$

$$g = \begin{cases} G/12.92 & G \le 0.04045 \\ ((G+0.55)/1.055)^{2.4} & G > 0.04045 \end{cases}$$

$$b = \begin{cases} B/12.92 & B \le 0.04045 \\ ((B+0.55)/1.055)^{2.4} & B > 0.04045 \end{cases}$$

and it the RGB system is not sRGB, then

$$r = R^{\gamma}$$
$$g = G^{\gamma}$$
$$b = B^{\gamma}$$

The transformation matrix (M) is calculated from the chromaticity coordinates of the RGB system (xr, yr), (xg, yg) and (xb, yb) and its reference white (XW, YW, ZW) using the following formula:

$$M = \begin{bmatrix} S_r X_r & S_r Y_r & S_r Z_r \\ S_g X_g & S_g Y_g & S_g Z_g \\ S_b X_b & S_b Y_b & S_b Z_b \end{bmatrix}$$

where

$$\begin{bmatrix} X_p \\ Y_p \\ Z_p \end{bmatrix} = \begin{bmatrix} x_p / y_p \\ 1 \\ (1 - x_p - y_p)/y_p \end{bmatrix} \text{ and } p = [r \quad g \quad b]$$

Table D-1: Results of matrix computation for commonly used RGB working spaces

RGB Working Space	Ref. White	RGB to XYZ [M]			XYZ to RGB $[M]^{-1}$		
Adobe RGB (1998)	D65	0.576700	0.297361	0.0270328	2.04148	-0.969258	0.0134455
		0.185556	0.627355	0.0706879	-0.564977	1.87599	-0.118373
		0.188212	0.0752847	0.991248	-0.344713	0.0415557	1.01527
Apple RGB	D65	0.449695	0.244634	0.0251829	2.95176	-1.0851	0.0854804
		0.316251	0.672034	0.141184	-1.28951	1.99084	-0.269456
		0.18452	0.0833318	0.922602	-0.47388	0.0372023	1.09113
PAL / SE-CAM	D65	0.430587	0.222021	0.0201837	3.06313	-0.969258	0.0678674
		0.341545	0.706645	0.129551	-1.39328	1.87599	-0.228821
		0.178336	0.0713342	0.939234	-0.475788	0.0415557	1.06919
sRGB	D65	0.412424	0.212656	0.0193324	3.24071	-0.969258	0.0556352
		0.357579	0.715158	0.119193	-1.53726	1.87599	-0.203996
		0.180464	0.0721856	0.950444	-0.498571	0.0415557	1.05707

D-2: Conversion from XYZ space to CIELAB space

The conversion from XYZ space to CIELab space is accomplished by the set of equations D-2.

$$L = 116 f_y - 16$$
$$a^* = 500\left(f_x - f_y\right) \qquad \textbf{(D-2)}$$
$$b^* = 200\left(f_y - f_z\right)$$

where:

$$f_x = \begin{cases} \sqrt[3]{x_r} & x_r > \varepsilon \\ \dfrac{\kappa y_r + 16}{116} & x_r \le \varepsilon \end{cases}$$

$$f_y = \begin{cases} \sqrt[3]{y_r} & y_r > \varepsilon \\ \dfrac{\kappa y_r + 16}{116} & y_r \le \varepsilon \end{cases}$$

$$f_z = \begin{cases} \sqrt[3]{z_r} & z_r > \varepsilon \\ \dfrac{\kappa z_r + 16}{116} & z_r \le \varepsilon \end{cases}$$

$$x_r = \frac{X}{X_r}$$

$$y_r = \frac{Y}{Y_r}$$

$$z_r = \frac{Z}{Z_r}$$

$$\varepsilon = \begin{cases} 0.008856 & \text{Actual CIE Standard} \\ 216/24389 & \text{Intent of the CIE Standard} \end{cases}$$

$$\kappa = \begin{cases} 903.3 & \text{Actual CIE Standard} \\ 24389/27 & \text{Intent of the CIE Standard} \end{cases}$$

It is sometimes convenient to express CIELAB colours in cylindrical coordinates system, which can be computed as follows:

$$C^*_{ab} = \left[a^{*2} + b^{*2}\right]^{1/2}, \quad h_{ab} = \tan^{-1}\left(\frac{b^*}{a^*}\right)\left(\frac{180}{\pi}\right) \qquad \textbf{(D-3)}$$

The polar coordinates are useful since difference in the chroma term C^*_{ab} can be correlated with to perceived colour difference and the difference in hue term h_{ab} can be correlated with perceived hue difference.

D-3: Conversion from XYZ space to RGB space

Given an XYZ colour whose components are in the nominal range [0.0, 1.0] and whose reference white is the same as that of the RGB system, the conversion to RGB is as follows is accomplished by the following equations:

$$\begin{bmatrix} r & g & b \end{bmatrix} = \begin{bmatrix} X & Y & Z \end{bmatrix} \begin{bmatrix} M \end{bmatrix}^{-1} \qquad \textbf{(D-4)}$$

where if the RGB system is sRGB, then

$$R = \begin{cases} 12.92r & r \le 0.0031308 \\ 1.055r^{1/2.4} - 0.055 & r > 0.0031308 \end{cases}$$

$$G = \begin{cases} 12.92r & g \le 0.0031308 \\ 1.055g^{1/2.4} - 0.055 & g > 0.0031308 \end{cases}$$

$$B = \begin{cases} 12.92r & b \le 0.0031308 \\ 1.055b^{1/2.4} - 0.055 & b > 0.0031308 \end{cases}$$

and if the RGB is not sRGB, then

$$R = r^{1/\gamma}$$
$$G = g^{1/\gamma}$$
$$B = b^{1/\gamma}$$

The inverse transformation matrix M^{-1} is obtained from Table D-1 and γ is the gamma value of the RGB colour system used.

D-4: Conversion from CIELAB space to XYZ space

The conversion from CIELAB to XYZ space requires a reference white point $[X_r \quad Y_r \quad Z_r]$.

$$\begin{aligned} X &= x_r X_r \\ Y &= y_r Y_r \\ Z &= z_r Z_r \end{aligned} \qquad \textbf{(D-5)}$$

where:

$$x_r = \begin{cases} f_x^3 & f_x^3 > \varepsilon \\ (116 f_x - 16/\kappa) & f_x^3 \le \varepsilon \end{cases}$$

$$y_r = \begin{cases} ((L+16)/116)^3 & L > \kappa\varepsilon \\ L/\kappa & L \le \kappa\varepsilon \end{cases}$$

$$z_r = \begin{cases} f_z^3 & f_z^3 > \varepsilon \\ (116 f_z - 16/\kappa) & f_z^3 \le \varepsilon \end{cases}$$

$$f_x = \frac{a^*}{500} + f_y$$

$$f_z = f_y - \frac{b^*}{200}$$

$$f_y = \begin{cases} (L+16)/116 & y_r > \varepsilon \\ (\kappa y_r + 16)/116 & y_r \le \varepsilon \end{cases}$$

$$\varepsilon = \begin{cases} 0.008856 & \text{Actual CIE Standard} \\ 216/24389 & \text{Intent of the CIE Standard} \end{cases}$$

$$\kappa = \begin{cases} 903.3 & \text{Actual CIE Standard} \\ 24389/27 & \text{Intent of the CIE Standard} \end{cases}$$

D-5: Computing CIELAB Colour Difference

The CIELAB colour difference is computed as the Euclidean distance between two points in CIELAB space:

$$\Delta E^*_{ab} = \left[\left(\Delta L^*\right)^2 + \left(\Delta a^*\right)^2 + \left(\Delta b^*\right)^2 \right]^{1/2} \quad \textbf{(D-6)}$$

where:

$$\Delta L^* = L^*_T - L^*_S$$
$$\Delta a^* = a^*_T - a^*_S$$
$$\Delta b^* = b^*_T - b^*_S$$

The subscript S and T refers to refers to the standard and the trial respectively—in industrial application of colour difference, it is typical to regard one of the colours as a standard and the other as a trial (or sample) that is supposed to be a visual match to the standard.

Other optimal colour difference formulae have been proposed with varying advantages and disadvantages, such as CMC, CIE94, CIEDE2000 (Westland & Ripamonti, 2004). These formulae are more complex and supposed to be optimal than ΔE^*_{ab}. The CIE recently recommended CIEDE2000 formula for evaluating small colour difference. However, ΔE^*_{ab} is still the current CIE recommendation for evaluating large colour differences (i.e. ΔE^*_{ab} >5).

Appendix E: Basic Statistical Concepts

E-1: Basic Terminologies

The following statistical terminologies were assumed as the basis for most of the analysis methods applied in this Thesis.

- *Standard error*. The standard error of a sample of sample size n is defined as the standard deviation of the sample divided by $\sqrt{n}$. The standard error of a statistical estimate is also defined as the square root of the estimated error variance of the quantity, i.e., $se = \sqrt{\hat{\sigma}^2}$.

- *Root mean square*. The root mean square (RSM), also known as the quadratic mean, is a measure of the magnitude of a random variable, calculated as in Equation E-1.

$$x_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \tag{E-1}$$

The relationship between x_{rms} and $\overline{x}$ is expressed by Equation E-2, which shows that the RMS is always greater than or equal to the mean, in that the RMS includes the square of the standard deviation. The RMS is a useful statistics than the mean if the random variable has both positive and negative values.

$$x_{rms} = \overline{x}^2 + \sigma_x^2 \tag{E-2}$$

- *True value*. The true value is a variable's theoretically correct and exact value. This is simply the mean of the population and is indeterminate since the population of measurements is infinite.

- *Error*. An error is the difference between any individual measurement of a variable and its true value. This is also indeterminate since the true value is unknown. Errors are expressed by:

$$\varepsilon_i = x_i - \mu \tag{E-3}$$

where x_i is an individual measurement associated with error ε_i and true value μ.

- *Residual.* This is the difference between any individual measurement and the most probable value or the mean. The terms error and residual are often used interchangeably; and although they have similar behaviour, they are theoretically different.

- *Degrees of freedom.* This refers to the number of measurements that are in excess of the number required to solve for the unknown parameters, i.e., the number of *redundant measurements*. The degree of freedom (r) is computed as the number of observations (n) minus the number of parameters (p): $r = n - p$.

- *Accuracy and precision.* Although accuracy and precision are often used interchangeably, the two concepts are statistically different. Accuracy refers to the closeness of a measured quantity to its true value or the most probable. Since the true value is indeterminate, absolute accuracy is also unknown. Precision on the other hand describes consistency among measurements, i.e. it is a measure of the closeness of a set of repeated measurements to each other and is usually measured by the standard deviation. A set of measurements may be precise but not accurate and vice versa. An example often used to illustrate the difference between precision and accuracy is the groupings of riffle shots (Mikhail, 1976; Wolf and Ghilani, 1997). Figure E-1 shows four possible groupings: (a) is both accurate and precise because the measurements are *close to each othe*r and the mean is *close to the centre*; (b) is precise but not accurate; (c) is accurate but not precise; (d) is both inaccurate and imprecise. In general, when making measurements, the results in (a) are preferred while (c) and (d) are undesirable. The results in (b) can be accepted if proper steps are taken to correct for the presence of the systematic errors that caused the entire set of measurements to deviate from the true value.

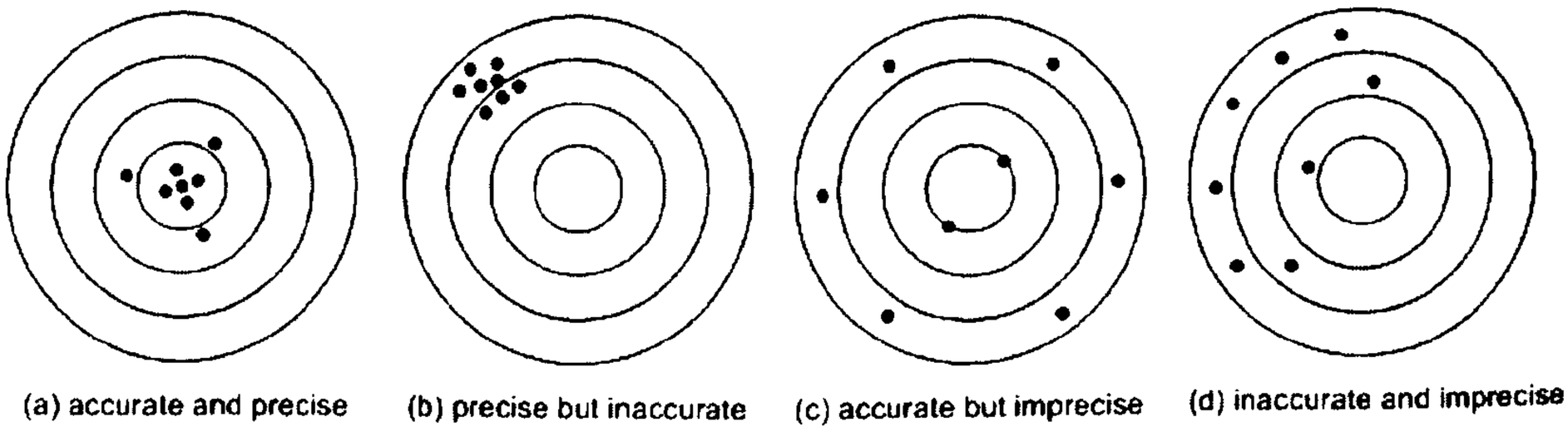


Figure E-1: Difference between accuracy and precision (adapted from Wolf & Ghilani, 1997)

- *Relative precision*. This refers to the ratio of a measure of precision (standard deviation) to the quantity measured or estimated.

E-2: Propagation of Random Errors

Measurements may be direct or indirect. Direct measurements are made by applying a device directly to the unknown quantity and observing its value. Indirect measurements are derived or computed from direct measurements using their mathematical relationships to the direct measurements. During this process, the errors inherent in the direct measurements are propagated by the computational process into the computed quantities. Thus, the computed quantities contain errors that are functions of the original errors. This concept is known as error propagation and is discussed in this section.

Let y be a quantity computed from several independent measurements $x_1, x_2, \cdots, x_n$ which have standard deviations $\sigma_1, \sigma_2, \cdots, \sigma_n$ respectively. Then using a general non-linear function $y = f(x_1, x_2, \cdots, x_n)$, the combined propagated error in σ_y can be computed from Equation E-4 (Anderson and Mikhail, 1998):

$$\sigma_y^2 = \left(\frac{\partial y}{\partial x_1}\right)\sigma_1^2 + \left(\frac{\partial y}{\partial x_2}\right)\sigma_2^2 + \cdots + \left(\frac{\partial y}{\partial x_n}\right)\sigma_n^2 \qquad \textbf{(E-4)}$$

If y is related to x by a linear function $y = a_1x_1 + a_2x_2 + \cdots + a_nx_n$, Equation E-4 simplifies to:

$$\sigma_y^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2 \qquad \textbf{(E-5)}$$

The one-dimensional case contains only one random variable x with mean μ_x and variance σ_x^2. The two-dimensional case has two random variables x and y, with means μ_x and μ_y, and variances σ_x^2 and σ_y^2 respectively, and covariance σ_{xy}. These three parameters can be collected into a square symmetric matrix with the variances along the main diagonal and the covariance along the off-diagonal. This matrix is called the *variance-covariance matrix* or simply the *covariance matrix*.

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \qquad \textbf{(E-6)}$$

Equation E-6 can be extended to multidimensional case for n random variables $x_1, x_2, \cdots, x_n$:

$$\sum_{xx} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_n^2 \end{bmatrix} \quad \textbf{(E-7)}$$

which is $n \times n$ square symmetric matrix.

Often, in practice, the variances and covariances of the random errors are not known in absolute terms but only to a scale factor termed the reference variance (denoted σ_0^2). The square root of the reference variance (σ_0^2) is called the reference standard deviation or standard error of unit weight. The relative variances and covariances are called cofactors, and are given by:

$$q_{ii} = \frac{\sigma_i^2}{\sigma_0^2},\ q_{ij} = \frac{\sigma_{ij}}{\sigma_0^2} \quad \textbf{(E-8)}$$

Collecting the cofactors into a square matrix produces the cofactor matrix denoted by Q. If Q is non-singular, its inverse is called the weight matrix denoted by W, and the relationship in Equation E-9 is deduced.

$$W = Q^{-1} = \sigma_0^2 \sum^{-1} \quad \textbf{(E-9)}$$

This can be generalised to correlated measurements and multiple functions. Let y be a set of n measurements, each of which is a function of another set x of n random variables. Given the covariance matrix $\sum_{xx}$ (or the cofactor matrix Q_{xx}) for the variables x, the covariance matrix $\sum_{yy}$ (or the cofactor matrix Q_{yy}) for the computed quantities y is given by:

$$\sum_{yy} = J_{yx} \sum_{xx} J_{yx}^T \text{ or } Q_{yy} = J_{yx} Q_{xx} J_{yx}^T \quad \textbf{(E-10)}$$

where J_{yx} is $m \times n$ and called the Jacobian matrix which contains the partial derivatives of y with respect to x and is given by:

$$J_{yx} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} \quad \textbf{(E-11)}$$

References

[1]. Abramowitz, M. and Stegun, I. A., 1972. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. New York: Dover.

[2]. Allrefer.com, 2006. Forward bend test. Web address: http://health.allrefer.com/pictures-images/forward-bend-test.html. Accessed: 23 September, 2006.

[3]. Anderson, J.M. and Mikhail, E.M., 1998. Surveying: Theory and Practice. USA: McGraw-Hill.

[4]. Arai, B. and Read, C., 2005. Web address: www.conleyread.net/en_us/docs/231_ConleyRead_ProjectFinalPresentation_BenjaminArai.ppt. Accessed: 14 October 2006.

[5]. Armstrong, G.W.D. et al, 1982. Non-standard vertebral rotation in scoliosis screening patients. Its prevalence and relation to the clinical deformity. Spine, 7:50-54.

[6]. Aubin, C.E. et al, 1997. Morphometric evaluations of personalised 3D reconstructions and geometric models of the human spine. Medical and Biological Engineering and Computing, 35(6):611-18.

[7]. Baylis, W.J. and Rzonca, E.C., 1988. Functional and structural limb length discrepancies: evaluation and treatment. Clinics Podiatric Medicine and Surgery, 5(3):509-20.

[8]. Benoni, G. and Willner, S., 1983. Asymmetry of moiré pattern in scoliosis at different levels of the spine. In: B. Drerup et al, eds. Moiré Fringe Topography and Spinal Deformity, Publisher, 93-98.

[9]. Boehler, W. and Marbs, A., 2002. 3D scanning instruments. In: Boehler W. and Patias P. eds. International Workshop on Scanning for Cultural Heritage Recording, CIPA WG6 and ISPRS Comm. V, Corfu, Greece, 9-12.

[10]. Boehnen, C. and Flynn, P., 2005. Accuracy of 3D Scanning Technologies in a Face Scanning Scenario. Proceedings of the Fifth International Conference on 3-D Digital Imaging and Modelling, 310 – 317.

[11]. Borrelli, V. et al, 2003. On the angular defect of triangulations and the pointwise approximation of curvatures. Computer Aided Geometric Design, 20:319–341.

[12]. Bowen, R.M., 1995. Respiratory management in scoliosis. In: J. Lonstein et al, eds. Moe's Textbook of Scoliosis and Other Spinal Deformities. Philadelphia: W.B. Saunders Company, 572-580.

[13]. Branthwaite, M.A., 1986. Cardiorespiratory consequences of unfused idiopathic scoliosis. British Journal of Diseases of the Chest, 80:360-9.

[14]. Bridwell, K., 2001a. Anatomical Planes of the Body. Web address: http://www.spineuniverse.com/displayarticle.php/article1023.html. Accessed: 2nd November, 2005.

[15]. Bridwell, K., 2001b. Vertebral Column, Web Address: http://www.spineuniverse.com/displayarticle.php /article1286.html, Accessed on: 2nd November, 2005.

[16]. Bunnell, W.P., 1984. An objective criterion for scoliosis screening. Journal of Bone and Joint Surgery, 66A(9):1381-1387.

[17]. Burrough, P.A. and McDonnel, R.A., 1998. Principles of Geographical Information Systems. Oxford: Oxford University Press, UK.

[18]. Burwell, R.G. et al, 1983. Standardised trunk asymmetry scores. In: B. Drerup et al, eds. Moiré Fringe Topography and Spinal Deformity, 27-32.

[19]. Burwell, R.G., 2003. Aetiology of idiopathic scoliosis: current Concepts. Paediatric Rehabilitation, 6(3-4):137-170.

[20]. Cassar-Pullicino, V.N. and Eisenstein, S.M., 2002. Imaging in scoliosis: What, why and how? Clinical Radiology, 57:543-562.

[21]. Cobb, J.R., 1948. Outline for the study of scoliosis. In: C.V. Mosby ed. American Academy of Orthopaedic Surgeons, Instructional Course Lectures, 261-275.

[22]. Cockerill, W. et al, 2000. Does location of vertebral deformity within the spine influence back pain and disability? Annals of the Rheumatic Diseases, 59:368-371.

[23]. Cordover, A.M., 1997. Natural history of adolescent thoracolumbar and lumbar idiopathic scoliosis into adulthood. Journal of Spinal Disorders, 10:193-6.

[24]. Croft, P. et al, 1997. Low back pain. In: A. Stevens and J. Raftery eds. Health Care Needs Assessment: The epidemiologically based needs assessment reviews, Second series. Abingdon, UK: Radcliffe Publishing, 129-179.

[25]. Cruickshank, J.L. et al, 1989. Curve patterns in idiopathic scoliosis: A clinical and radiographic study. Journal of Bone and Joint Surgery, 71B:259-263.

[26]. Dangerfield, P.H. et al, 1992a. Multiple spinal level measurement of body surface deformity in scoliosis. In: A. Alberti et al eds. Surface Topography and Spinal Deformity. Stuttgart: Gustav Fischer Verlag, 212-216.

[27]. Dangerfield, P.H., 2003. The classification of spinal deformities. Paediatric Rehabilitation, 6(3-4):133-136.

[28]. Dawson, E.G. et al, 1993. Optoelectronic evaluation of trunk deformity in scoliosis. Spine, 18 (3):326-331.

[29]. Demers, M.N., 2000. Fundamentals of Geographic Information Systems, Second Edition. John Willey & Sons.

[30]. Dennis, A.L., 2003. Scoliosis: A straightforward approach to the curved spine. JAAPA, 16(9):17-8, 21-3.

[31]. Dent, B.D., 1999. Cartography: Thematic Map Design. MCGraw-Hill Inc.

[32]. Diab, K.M., et al, 1995. Accuracy and applicability of measurement of the scoliotic angle at the frontal plane by Cobb's method, by Ferguson's method and by a new method. European Spine, 4(5):291-5

[33]. Douros, I. and Buxton, B., 2002. Three-dimensional surface curvature estimation using quadric surface patches. Proceedings of Numerisation 3D-Scanning.

[34]. Drerup, B. and Hierholzer, E., 1985. Objective determination of anatomical landmarks on body surface: measurement of the vertebra prominens from surface curvature. Journal of Biomechanics, 18(6):467-474.

[35]. Drerup, B. and Hierholzer, E., 1996. Assessment of scoliotic deformity from back shape asymmetry using improved mathematical model. Clinical Biomechanics, 11(7):376-383.

[36]. Drerup, B., 1984. Principles of measurement of vertebral rotation from frontal projections of the pedicles. Journal of Biomechanics, 17(12):923—935.

[37]. ESRI, 1998. ESRI Shapefile Technical Description, Environmental Systems Research Institute. Web: http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf. Accessed: August, 2nd, 2003.

[38]. Ettinger, B. et al, 1992. Contribution of vertebral deformities to chronic back pain and disability. The Study of Osteoporotic Fractures Research Group. Journal of Bone and Mineral Research, 7(4):449-56.

[39]. Ferguson, A.B., 1930. The study and treatment of scoliosis. South Medical Journal, 23:116-120.

[40]. Fisher, W.D., 1958. On grouping for maximum homogeneity. American Statistical Association Journal, 53:789-798.

[41]. Flynn, P.J. and Jain, A.K., 1989. On reliable curvature estimation. Proceedings of the International Conference on Computer Vision and Pattern Recognition, 110-116.

[42]. Friberg O., 1983. Clinical Symptoms and Biomechanics of Lumbar Spine and Hip Joint in Leg Length Inequality. Spine, 8(6):643-651.

[43]. Frobin W. and Hierholzer E., 1981. Rasterstereography: a photogrammetric method for measurement of body surfaces. Photogrammetric Engineering and Remote Sensing, 47(12):1717-1724.

[44]. Frobin, W. and Hierholzer, E., 1982. Analysis of human back shape using surface curvatures. Journal of Biomechanics, 15(5):379-390.

[45]. Frobin W. and Hierholzer E., 1983. Automatic Measurement of body surfaces using rasterstereography. Photogrammetric Engineering and Remote Sensing, 49(3):377-384.

[46]. Frobin W. and Hierholzer E., 1988. Real-time rasterstereography using a solid state camera. Proceedings of SPIE 1030, 28-34.

[47]. Frobin, W. et al, 1983. Shape analysis of surfaces: extraction of shape from coordinate data. In: B. Drerrup *et al, eds. Moire* Fringe Topography and Spinal Deformity, 71-82.

[48]. Garland, M. and Heckbert, P.S., 1997. Surface simplification using quadric error metrics. In: Computer Graphics Proceedings (SIGGRAPH'97), 209-216.

[49]. Gauch, J.M., 1998. Noise Removal and Contrast Enhancement. In: S. J. Sangwine et al, eds. The Colour Image Processing Handbook. London, England: Chapman and Hall.

[50]. Goldberg, M.S. et al, 1988. Observer variation in assessing spinal curvature and skeletal development in adolescent idiopathic scoliosis. Spine, 13:1372-1377.

[51]. Gomes, A.S. et al, 1995. Automated 360 degree profilometry of human trunk for spinal deformity analysis. In: D. D'Amico and D.E. Santambrogio eds. Three Dimensional Analysis of Spinal Deformities. Amsterdam: IOS Press, 423-329.

[52]. Gray, A., 1997. Modern Differential Geometry of Curves and Surfaces with Mathematica, 2nd Edition. Boca Raton, FL: CRC Press.

[53]. Gray, H., 1918. Anatomy of the human body. Online Edition, Published May 2000 by Bartleby (http://www.bartleby.com/107). Accessed: 04/01/2005.

[54]. Groop, R.E., 1980. JENKS: An Optimal Data Classification Program for Choropleth Mapping. Technical Report 3. Department of Geography, Michigan State University, 1980.

[55]. Grossman, T.W. et al, 1995. An evaluation of the Adams forward bend test and the Scoliometer in a scoliosis school screening setting. Journal of Pediatric Orthopaedics, 15:535-538.

[56]. Halioua, M. and Liu, H., 1989. Optical three-dimensional sensing by phase measuring profilometry. Optics and Lasers in Engineering, 11:185-215.

[57]. Hamann, B., 1993. Curvature approximation for triangulated surfaces. In: G. Farin et al, eds. Geometric Modelling, Computing Supplement 8, Springer-Verlag, 139-153.

[58]. Haris, J.D. and Turner-Smith, A.R., 1983. Clinical back shape analysis with anatomically defined datum. In: B. Drerup ed. Moire Fringe Topography and Spinal Deformity, 99-106.

[59]. Hartigan, J.A., 1975. Clustering Algorithms. New York: John Wiley.

[60]. Hierholzer, E. and Drerup, B., 1995. High resolution rasterstereography. In: M. D'Amico et al, eds. Three Dimensional Analysis of Spinal Deformities. Amsterdam: IOPress, 435-439.

[61]. Hierholzer, E. et al, 1983. Computerised data acquisition and evaluation of moiré topograms and rasterstereography. In: B. Drerup et al, eds. Moire Fringe Topography and Spinal Deformity, 93-98 or 233-240.

[62]. Howard, A. et al, 1998. A comparative study of TLSO, Charleston, and Milwaukee braces for idiopathic scoliosis. Spine, 23(22):2404-2411.

[63]. Jähne, B., 1995. Digital image processing: concepts, algorithms and scientific applications. Berlin, Heidelberg: Springer-Verlag.

[64]. Jaremko, J.L. et al, 2001. Estimation of spinal deformity in scoliosis from torso surface cross sections. Spine, 26(14):1583-91.

[65]. Jaremko, J.L. et al, 2002a. Comparison of Cobb angles measured manually, calculated from 3-D spinal reconstruction, and estimated from torso asymmetry. Computer Methods in Biomechanics and Biomedical Engineering, 5(4):277-281.

[66]. Jaremko, J.L. et al, 2002b. Indices of torso asymmetry related to spinal deformity in scoliosis. Clinical Biomechanics, 17:559–568.

[67]. Jaremko, J.L. et al, 2002c. Genetic algorithm-neural network estimation of Cobb angle from torso asymmetry in scoliosis. Journal of Biomechanical Engineering, 124:496-503.

[68]. Jenks, G.F. and Fred G.C., 1971. Error on choropleth maps: definition, measurement, reduction. Annals of Association of American Geographers, 61:217-44.

[69]. Jenks, G.F., 1977. Optimal data classification for choropleth maps. Occasional Paper No. 2, Department of Geography, *University of Kansas.*

[70]. Kawamura, T. et al, 1983. Computerised moire topography and anterior chest wall deformity. In: B. Drerup et al, eds. Moiré Fringe Topography and Spinal Deformity, 241-248.

[71]. Kenney, J. F. and Keeping, E. S., 1964. Quantiles. In: J.F. Kenney ed. Mathematics of Statistics, Part 1, Princeton, NJ: Van Nostrand, 37-38.

[72]. Keper, B., 1994. Differential geometry. In: K. Roktorys ed. Survey of Applicable Mathematics, Kluwer Academic Publishers, 260-334.

[73]. Kerpt, J. et al, 1998. Shadow moiré topography: theory and application. Acta Universitatis Palackianae Olomucensis, Facultas Rerum Naturalium, Physica, 37:137-159.

[74]. Klaber Moffett, J. et al, 1995. Back pain: Its management and cost to society. York, UK: Centre for Health Economics, University of York.

[75]. Koenderink, J.J. et al, 1992. Surface shape and curvature scales. Image and Vision Computing, 10(8): 557-565.

[76]. Krsek, P. et al, 1998. Algorithms for computing curvatures from range data. In: A. Ball et al, eds. The Mathematics of Surfaces VIII, Information Geometers, 1-16.

[77]. Lawrence, J.S., 1969. Disc degeneration: its frequency and relationship to symptoms. Annals of the Rheumatic Diseases, 28:121-37.

[78]. Seidenberg, L.R. et al, 1992. Surface curvature analysis using colour. Proceedings of the IEEE Conference on Visualization, Visualization '92, 260-267.

[79]. Lenke, L.G. et al, 2001. Adolescent idiopathic scoliosis: a new classification to determine the extent of spinal arthrodesis. Journal of Bone and Joint Surgery, 83A:1169-1181.

[80]. Levkowitz, H., 1997. Colour theory and modelling for computer graphics, visualization and multimedia applications. Massachuset, USA: Kluwer Academic Services.

[81]. Lorensen, W. E. and Cline, H. E., 1987. Marching cubes: a high-resolution 3D surface construction algorithm. Computer Graphics, 21(3):163-169.

[82]. Mahood, J.K. et al, 1995. Perceptions of cosmetic deformity in scoliosis. In: M. D'Amico et al, eds. Three Dimensional Analysis of Spinal Deformities, Amsterdam: IOS Press, 239-242.

[83]. Maniadakisa, N. and Gray, A., 2000. The economic burden of back pain in the UK. Pain, 84: 95-103.

[84]. Martin, R.R., 1998. Estimation of principal curvatures from range data. International Journal of Shape Modelling, 4:99-109.

[85]. Mayo, N.E. et al, 1994. The Ste-Justine Adolescent Idiopathic Scoliosis Cohort Study. Part III: Back pain. Spine, 19(14):1573-81.

[86]. McAlister, W.H. and Shackelford, M.G.D., 1975. Classification of spinal curvatures. Radiologic Clinics of North America, 13(1):93-111.

[87]. Medioni, G. et al, 2000. A Computational Framework for Segmentation and Grouping, *Elsevier, Amsterdam.*

[88]. Mikhail, E.M., 1976. Observations and Least-squares. New York: *Thomas Y. Crowell.*

[89]. Morais, T. et al, 1995. Age- and sex-specific prevalence of scoliosis and the value of school screening programs. American Journal of Public Health, 75:1377-80.

[90]. Moran, A.H., 1998. Automatic digitization and analysis of moire topograms for clinical use, Thesis (PhD), University of Bath, UK.

[91]. Moreland, M.S. et al, 1983. The clinical use of moiré topography for spinal deformity: a perspective of challenge. In: B. Drerup et al, eds. Moiré Fringe Topography and Spinal Deformity, 129-140.

[92]. Morrissy, R.T. et al, 1990. Measurement of the Cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error. Journal of Bone Joint Surgery Am, 72:320-327.

[93]. Nash C. and Moe J., 1969. A study of vertebral rotation. Journal of Bone and Joint Surgery, 51(2):223-229.

[94]. NIAMS, 2001. Questions and answers about Scoliosis in Children and Adolescents. U.S. Department of Health and Human Services, NIH Publication No. 01-4862.

[95]. NIEHS Health and Safety Manual, Chapter 9. Web address: http://www.niehs.nih.gov/odhsb/manual/man9b.htm. Accesses: 22 June 2005.

[96]. O'Neill, O., 1996. Elementary Differential Geometry, Academic Press.

[97]. OMG, 2001. Unified Modelling Language Specification, Version 1.4, September 2001. Website: http://www.omg.org/technology/documents/formal/uml.htm. Accessed on November 3, 2003.

[98]. Page, D.L. et al, 2002. Normal vector voting: crease detection and curvature estimation on large, noisy meshes. Graphical Models, Special Issue on Large Triangle Mesh Models, 64(3-4):199-229.

[99]. Pearson, J.J., 1996. Automated body shape measurement in scoliosis. Thesis (PhD), Liverpool John Moores University, UK.

[100]. Perdriolle, R. and Vidal, J., 1987. Morphology of scoliosis: the three dimensional evolution. Orthopedics, 10:909-915.

[101]. Perona, P. Et al, 1994. Anisotropic diffusion. In: B. M. ter Haar Romeny, ed. Geometry-driven diffusion in computer vision, Dordrecht: Kluwer Academic Publishers, 72-92.

[102]. Poncet, P. et al, 2001. Geometric torsion in idiopathic scoliosis: three-dimensional analysis and proposal for a new classification. Spine, 26:2235-2243.

[103]. Pratt, W.K., 2001. Digital Image Processing. New York: John Willey and Sons.

[104]. Pruijs, J.E. et al, 1994. Variation in Cobb angle measurements in scoliosis. Skeletal Radiology, 23:517-520.

[105]. Richardson, M.L., 2000. Approaches to differential diagnosis in musculoskeletal imaging, Web Address: http://www.rad.washington.edu/mskbook/scoliosis.html, Accessed: 2nd November, 2005.

[106]. Roberts, A., 2001. Curvature attributes and their application to 3D interpreted horizons. First Break, 19(2):85-100.

[107]. Robinson, A.H. et al, 1995. Elements of Cartography. USA: John Wiley & Songs.

[108]. Robinson, E.F., 1983. Statistical assessment of two methods of measuring scoliosis before treatment. Canadian Medical Association Journal, 129:839-841.

[109]. Ross, J. et al, 1987. Back pain and spinal deformity in cystic fibrosis. American Journal of Diseases of Children, 141(12):1313-6.

[110]. Rousseeuw, P.J. and Leroy, A.M. 2003. Robust Regression and Outlier Detection. Willey and Sons.

[111]. Russ, J.C., 2002. The image processing handbook, Fourth Edition. Florida: CRC Press LLC.

[112]. Saint-Marc, P. et al, 1991. Adaptive smoothing: a general tool for early vision. IEEE Transaction on PAMI, 13(6):514-529.

[113]. SAUK, 2000. The Scoliosis Association, UK. Web address: http://www.sauk.org.uk/spinal_braces.htm. Accessed: 07 June, 2007.

[114]. Scharr, H. et al, 2003. Image Statistics and Anisotropic Diffusion. Ninth IEEE International Conference on Computer Vision (ICCV'03), 2:840

[115]. Scoliosis Research Society (SRS), 2000. In depth review of scoliosis. Internet Address: http://www.srs.org/patients/review/default.asp, Accessed on 10, October 2004.

[116]. Scripter, M., 1970. Nested-means map classes for statistical maps. Annals of the Association of American Geographers, 60:385-393.

[117]. Sinha, S.S. and Besl, P.J., 1990. Principal patches: a viewpoint-invariant surface description. IEEE Transaction on Pattern Analysis and Machine Intelligence, 226-231.

[118]. Slocum, T.A., 1999. Thematic cartography and visualization. Upper Saddle River, New Jersey: Prentice Hall.

[119]. Soong, T.T., 2004. Fundamentals of probability and statistics for Engineers. England: John Willey and Sons.

[120]. Sorensen, H.V. et al, 1987. Real-valued Fast Fourier Transform algorithms. IEEE Transactions in Acoustics, Speech and Signal Processing, ASSP-35:849–863.

[121]. Stokely, E.M. and Wu, S.Y., 1992. Surface parameterization and curvature measurement of arbitrary 3D objects: Five practical methods. IEEE Transaction on Pattern Analysis and Machine Intelligence, 14(8):833-840.

[122]. Stokes I.A.F. and Moreland M.S., 1989. Concordance of shape of the back asymmetry and spine shape in idiopathic scoliosis. Spine, 14(1):73-78.

[123]. Stokes, I.A. et al, 1993. Reexamination of the Cobb and Ferguson angles: bigger is not always better. Journal of Spinal Disorders, 6(4):333-8.

[124]. Stokes, I.A.F., 1994. Three-dimensional terminology of spinal deformity. Spine, 19(2):236-248.

[125]. Surazhsky, T. et al, 2003. A comparison of Gaussian and mean curvatures estimation methods on triangular meshes. Proceedings of the 2003 IEEE International Conference on Robotics and Automation, ICRA, 1021-1026.

[126]. Suziki, N. et al, 1981. Application of moiré topography to spinal deformity. In: M.S. Moreland et al, eds. Moiré Fringe Topography and Spinal Deformity, 66-75.