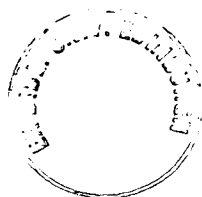


# Statistical methods for segmenting X-ray CT images of sheep

*Caroline D. Robinson*

Doctor of Philosophy  
University of Edinburgh

2000



# Abstract

X-ray computed tomography (CT) is a non-invasive imaging technique widely used in medical diagnosis to detect physiological abnormalities. Recently it has been adopted for estimating tissue proportions in live sheep. This thesis is concerned with the development of statistical methods for automating the estimation of tissue proportions from CT images.

The first stage in the estimation process is to segment sectional images into the internal organs, the carcass and the area external to the sheep. This is currently achieved by manually extracting boundaries which encircle the internal organs of the sheep, and is undesirable because it is a very subjective and tedious process. We explore the use of deformable templates to automate this stage, by means of a parametrised stochastic template which describes the shape and variability of these boundaries. The manually segmented boundaries from 24 lumbar images are parametrised using Fourier coefficients, which are reduced in dimensionality using principal components in order to estimate a distribution on the parameters of the template. Templates are fitted to further images using a criterion which combines the local pixel gradient and closeness to the estimated template distribution.

Having isolated the carcass region, we estimate the proportions of fat and muscle by modelling the probability density function of the pixel values in the segmented image, taking into account that many pixel values are generated from a mixture of two or more tissues. The spatial response of the CT machine is investigated by examining a sharp boundary in the image. Modelling this response as an isotropic bivariate normal density leads to a new probability density function for the values of the mixed pixels in the image, and hence to a combined distribution with the remaining pixels.

In a simulation study, this proposed density function is fitted to histograms of pixel values by maximum likelihood, and is shown to estimate the tissue proportions more accurately than the threshold-based method currently in use. The effect of taking some account of a pixel's spatial context by examining the greyscale values of its neighbours is also investigated.

# Acknowledgements

I would like to thank my supervisors, Dr Chris Glasbey and Dr Chris Theobald for their continued and invaluable advice and support throughout this thesis. I am also very grateful to Professor Tom Leonard for his supervision in the early months of my research and to Neil Hamilton-Smith for his patience to assist with my Fortran queries.

In addition, many thanks go to Mark Young, Kirsty McLean and Neil Robson at the SAC-BioSS CT unit for not only providing me with the data, but taking the time to demonstrate the workings of the scanner and their image analysis system.

Finally, and certainly not least, I would like to thank the Engineering and Physical Sciences Research Council and the Scottish Agricultural College for funding this project.

# Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.



# Table of Contents

<b>Chapter 1</b>	<b>Introduction</b>	<b>5</b>
1.1	Why use X-ray CT for sheep breeding? . . . . .	5
1.2	Digital images and Hounsfield units . . . . .	6
1.3	The nature of the data . . . . .	8
1.4	Current approach to tissue estimation . . . . .	10
1.5	Outline of the thesis . . . . .	11
<b>Chapter 2</b>	<b>X-ray computed tomography</b>	<b>13</b>
2.1	Introduction to X-ray CT . . . . .	13
2.2	X-ray CT and its history . . . . .	15
2.3	The scanning system of an X-ray CT machine . . . . .	16
2.3.1	Modes of scanning . . . . .	18
2.4	X-ray CT scanning of sheep using the SAC-BioSS scanner . . . . .	20
2.4.1	The scanning strategies used for sheep . . . . .	20
<b>Chapter 3</b>	<b>Deformable templates</b>	<b>22</b>
3.1	Introduction . . . . .	22
3.2	Low-level segmentation . . . . .	22
3.3	Deformable templates . . . . .	24
3.3.1	Energy minimising approach . . . . .	25
3.3.2	Bayesian formulation . . . . .	26
3.3.3	Optimisation of the energy and Bayesian approaches . . . . .	27

3.4	Free-form models . . . . .	27
3.5	Parametric deformable templates . . . . .	29
3.5.1	Hand-crafted models . . . . .	30
3.5.2	Prototype based templates . . . . .	31
3.6	Summary and conclusions . . . . .	36
<b>Chapter 4 Segmentation of sheep images</b>		<b>37</b>
4.1	Introduction . . . . .	37
4.2	Manual segmentation of the sheep images . . . . .	38
4.3	Modelling the carcass boundary using deformable templates . . . . .	40
4.4	Reduced-rank approximation and principal components . . . . .	45
4.4.1	Assessing results and principal scores from the reduced-rank approximation . . . . .	48
4.5	Fitting boundaries to the training images using the reduced-rank approximation . . . . .	49
4.6	Image filtering . . . . .	52
4.6.1	Spatial filtering . . . . .	53
4.7	Matching criterion . . . . .	56
4.8	Nelder-Mead optimisation algorithm . . . . .	57
4.8.1	Results from optimising the objective function using the Nelder-Mead algorithm . . . . .	58
4.9	Redefinition of the template . . . . .	60
4.9.1	Smoothing the image before optimisation . . . . .	64
4.9.2	Results after redefining the template . . . . .	67
4.10	Applying the redefined model to validation data . . . . .	71
4.11	Effect of adding mirror images to the training set . . . . .	72
4.12	Summary and conclusions . . . . .	78

<b>Chapter 5</b>	<b>Probability density function of pixel values</b>	<b>80</b>
5.1	Introduction . . . . .	80
5.2	Review of recent literature on estimating the probability distribution of pixel values . . . . .	81
5.3	Approximating the probability density function of the pixel values	83
5.3.1	Review of literature on estimating the point spread function, (PSF) . . . . .	83
5.3.2	Estimation of the PSF of the SAC-BioSS CT scanner . . . . .	86
5.3.3	Observed image for a circular normal PSF and a linear boundary . . . . .	91
5.3.4	Using the estimated PSF to approximate the probability density function of pixel values . . . . .	93
5.4	The distribution of perpendicular distances of pixels from a random line . . . . .	98
5.5	Summary and conclusions . . . . .	100
<b>Chapter 6</b>	<b>Estimation of tissue proportions</b>	<b>101</b>
6.1	Introduction . . . . .	101
6.2	Random processes of geometrical objects . . . . .	102
6.2.1	Representation space for directed lines in the plane . . . . .	102
6.3	Simulation study . . . . .	104
6.4	Methods used for estimating tissue proportions . . . . .	109
6.4.1	Threshold methods . . . . .	109
6.4.2	Mixed pixel methods . . . . .	111
6.5	Results and discussion . . . . .	113
6.6	Modelling thin layers of tissue . . . . .	119
6.6.1	Assessing the accuracy of the original assumptions . . . . .	120
6.6.2	Modifying the simulations . . . . .	122
6.7	Incorporating spatial information of pixels to estimate the tissue proportions . . . . .	124

6.7.1	Selecting thresholds for the variances of pure tissues . . . .	128
6.7.2	Results from using spatial information . . . . .	130
6.8	Summary and conclusions . . . . .	133
<b>Chapter 7 Conclusions and further work</b>		<b>135</b>
7.1	Conclusions . . . . .	136
7.2	Further work . . . . .	139
<b>Bibliography</b>		<b>141</b>

# Chapter 1

## Introduction

The aim of this thesis is to develop statistical methods for automated estimation of the amount of fat, muscle and bone in X-ray computed tomography images of sheep.

X-ray computed tomography (CT) is a non-invasive imaging technique whose primary application is in medical diagnosis. However, in a very different application, CT scanning is being used increasingly in sheep breeding programmes, and offers potential for estimating tissue proportions in live sheep to aid selection.

### 1.1 Why use X-ray CT for sheep breeding?

In the past, ultrasound imaging has proved a very useful tool for breeders who wish to improve the carcass composition of their animals (see Simm, 1998). However, X-ray CT offers a more accurate imaging technology, allowing genetic improvement of carcass composition to be accelerated and more characteristics to be assessed, e.g. the ratio of muscle to fat, studies of fat distribution or development of muscle. The quality of the carcass is becoming of greater importance for breeders, but with ultrasound imaging the most useful and critical characteristics, whose manipulation causes the greatest impact, are difficult to assess objectively and accurately in live sheep.

Therefore in 1997, with this understanding the Scottish Agricultural College (SAC) collaborated with Biomathematics and Statistics Scotland (BioSS), aiming to benefit from the use of CT scanning in meat sheep breeding programmes. Meat breeds, such as Charollais, Suffolks and Texels, produce about 70% of the lambs reared for meat production in Britain, so genetic improvement in this sector will have a major impact in the whole sheep industry. It is believed by breeders that

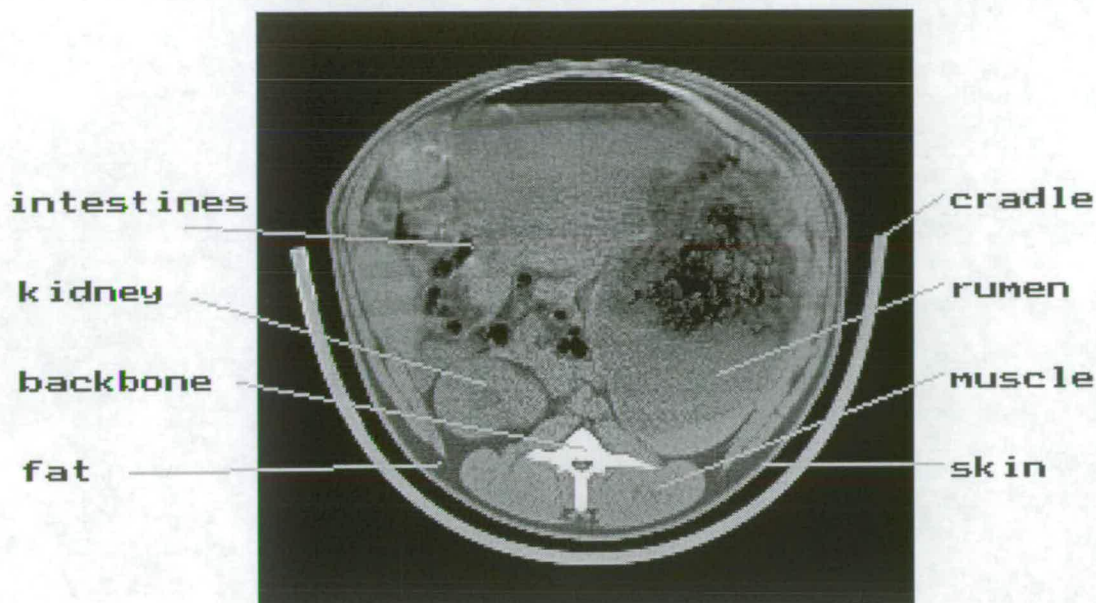


Figure 1.1: X-ray CT image showing a cross-section of the abdomen of a sheep with the main regions identified.

the economic benefits to the industry could be increased by up to 50% by using CT measurements as opposed to ultrasound.

CT scanning provides more accurate information on body tissues than ultrasound; this results in fewer animals requiring to be studied. Detailed measurements can now be made on tissues of live animals, whereas previously these calculations could only be made on dissection of the animal. In addition, some entities such as the distribution of fat in the body or the shape of individual muscles can now be visualised in a way which was previously unavailable. Figure 1.1 shows a typical X-ray CT image through the abdomen of a sheep; the U-shaped cradle in which the sheep is lying can also be seen.

## 1.2 Digital images and Hounsfield units

X-ray CT measures tissue density, and each CT picture is a digital image which is made up of a grid of small squares, which are called pixels. Often the grid is square and the number of rows and columns is either 256 or 512, with the labelling of pixels starting in the top left-hand corner.

Each pixel is assigned a value which corresponds to the average density of the tissue around that point in the image. Air is least dense and tissues such as bone are most dense. The pixel values are given Hounsfield units (HU), which measure



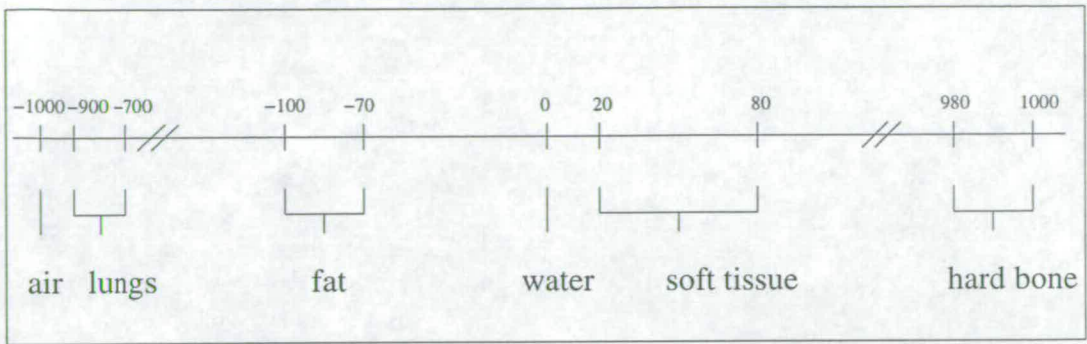


Figure 1.2: *The Hounsfield scale and values for various body tissues.*

how opaque different tissues are to the passage of X-rays and range between  $-1024$  for air to in excess of  $+1000$  for dense bone. Figure 1.2 shows the full Hounsfield scale and the ranges of values for various tissues.

However, the human eye can distinguish only a small number of shades of grey, around 10-15. Therefore, if the full Hounsfield scale (approximately 2000 HU) is displayed in an image then small differences in density appear homogeneous and are masked. The computer displays different density values as different shades of grey and these are known as grey-levels. Usually the computer discretizes the density values as 256 shades of grey, where 0 is displayed as black (representing air) and 255 as white (representing bone). The choice of how to display pixel values is arbitrary and the range of densities displayed can be altered to suit the particular investigation by selecting upper and lower limits on the density range required. For example, we can choose to highlight just dense tissues or have a large range of densities displayed in the image. Pixels which have a Hounsfield number larger than the upper limit and smaller than the lower limit can be assigned to appear as white and black in the image.

For the purposes of this thesis, we re-configure the sheep images using a scale which emphasizes the differences between fat and muscle tissues by selecting the Hounsfield units in the range  $[-256, +256]$ . A linear transformation on this range is used to assign greyscale values with  $-256$  and  $+256$  represented by greyscale values 0 and 255 respectively. Hounsfield values less than  $-256$  are represented by greyscale value 0 and the Hounsfield values greater than  $+256$  are represented by greyscale value 255. Therefore, in a typical image of a sheep, visually we see air as black, fat as dark grey, muscle as light grey and bone as white; see Figure 1.1. As in the convention with X-ray plates, see Chapter 2, light areas in the image denote regions which transmitted less X-rays (usually bone). Muscles and internal organs appear slightly lighter than fat tissue because they are slightly



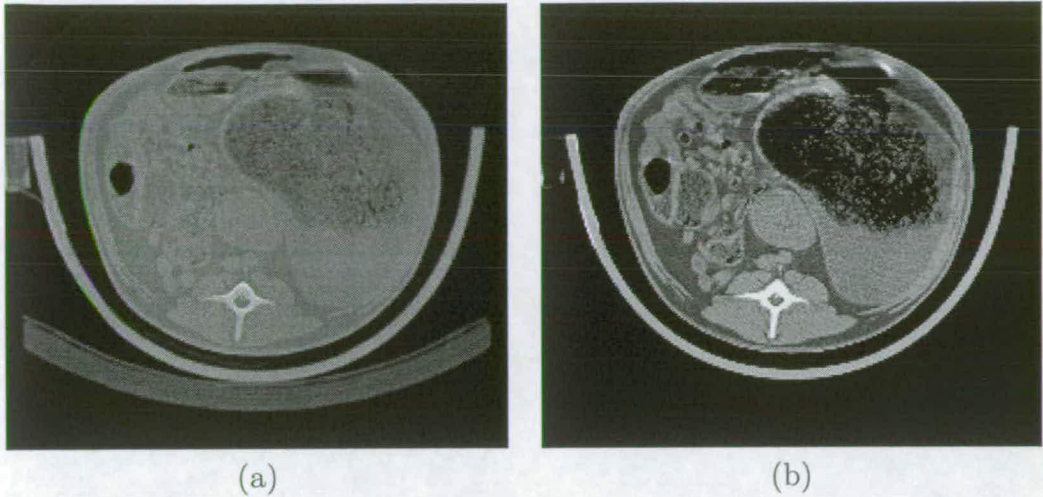


Figure 1.3: *X-ray CT image through the abdomen of a sheep where (a) the full Hounsfield range is displayed (b) a subset of the Hounsfield range is displayed to highlight the differences between fat and muscle tissue.*

more opaque to X-rays. Figure 1.3 (a) shows the full Hounsfield range displayed in a typical cross-section through the abdomen of a sheep and (b) shows the same image with Hounsfield units in the range  $[-256, +256]$  to highlight the fat and muscle tissues.

### 1.3 The nature of the data

The SAC-BioSS CT scanner is a Siemens Somatom CR system, which is described in more detail in Chapter 2. The dataset consists of 24 Suffolk lambs, comprising both males and females, scanned at the age of 22 weeks.

Experts at the SAC-BioSS CT unit identified several CT scan positions in the whole body which were most informative for the estimation of body composition. Therefore, the convention now established at the CT unit is to obtain a conventional X-ray image (in the vertical plane through the animal), together with three anatomically located CT cross-sectional images for each sheep. The process of obtaining these images is described in Chapter 2. These three locations are the ischium (upper legs), lumbar (lower back) and the thorax (chest region) and they comprise one each in the main carcass regions. Hence, this is intended to allow accurate estimation of tissue proportions and distributions from only three scans. The three CT scans are taken parallel to each other and are perpendicular to the long axis of the sheep. Each image contains a  $256 \times 256$  array of pixels with the sheep scanned lying on its back, and the left side of the sheep is displayed on the



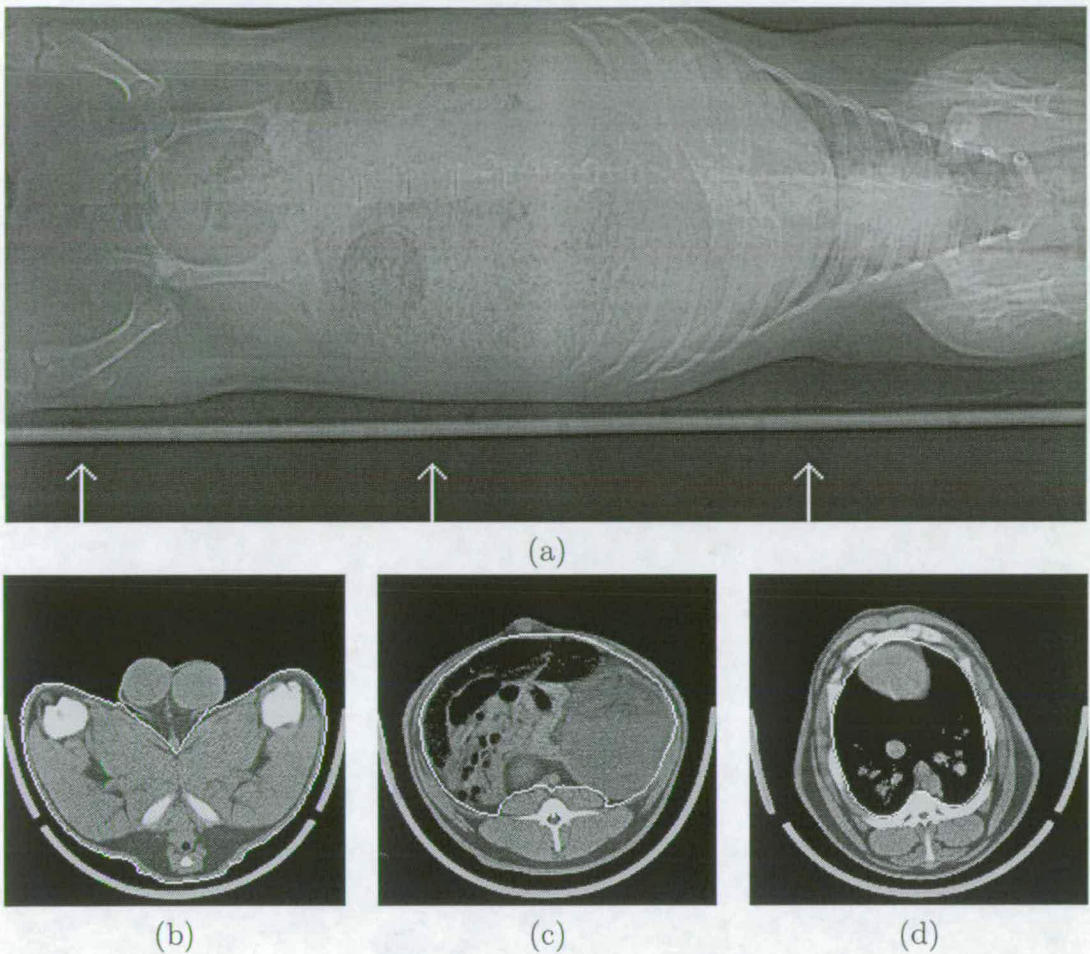


Figure 1.4: (a) Conventional X-ray image of a male sheep lying on its back in a cradle. The cross-sectional CT images, (b) ischium (c) lumbar (d) thorax are displayed, (with the manually drawn boundaries) and are located at the positions marked by arrows.

right to the viewer. Figure 1.4 shows a typical example of each of these images, where (a) shows a conventional X-ray image of a sheep lying on its back in a cradle. The position of the three cross-sectional images relative to this conventional image are marked with arrows. These cross-sections are (b) ischium, (c) lumbar and (d) thorax images, which have been trimmed to become  $191 \times 191$  arrays of pixels. The manually drawn boundaries, which are explained in Section 1.4, for each cross-sectional image are shown in white. It is attempted, as much as possible, to restrain the sheep to lie vertically in the cradle. However, due to movement of some of the animals after restraint, the corresponding images show slight rotation from the central axis of the body from the vertical. Examples of this are shown in Figures 1.3 and 1.4.

## 1.4 Current approach to tissue estimation

The approach which is currently in use at the SAC-BioSS CT unit for estimation of tissue proportions in sheep takes place in three stages, which are outlined briefly below. For a more detailed summary of each stage, see Glasbey and Robinson (2000).

The first stage involves segmenting the relevant tissue areas of fat, muscle and bone from images similar to Figures 1.4 (b), (c) and (d), by excluding the internal and external organs, the cradle and the air. This is currently achieved by tracing round the digestive tract, internal organs and internal fat using a computer mouse. These manual boundaries for each anatomical position are shown in white in Figures 1.4 (b), (c) and (d). Another boundary around the outside of the sheep is required in Figures 1.4 (c) and (d), but can be found comparatively easily: see Chapter 4. The greyscale values of the pixels contained within the inner boundary and outside the outer one are set to zero, so the remaining non-zero pixels correspond to fat, muscle and bone. These images are referred to as 'segmented images' and are used by the operator in stage two to predict the relative body composition.

In the second stage, the animal's fat and muscle tissue volumes are estimated from the segmented images. At present, this is achieved by examining the histograms of greyscale values of these pixels and using three threshold values to segment the image into air, fat, muscle and bone. These threshold values are  $-196\text{HU}$ ,  $-24\text{HU}$  and  $176\text{HU}$ .

Finally, in the third stage, the overall amounts of fat, muscle and bone in the carcass are predicted using these tissue measurements from the three CT images of each sheep, together with the weight of the animal at the time of scanning. This is known as the 'liveweight'. These predictions are made using multiple regression procedures. Initially, the prediction equations used by SAC-BIOSS CT unit were breed specific, e.g. for Charollais, Suffolks or Texels. However, it was found that when the breed specific equations were used to predict carcass tissue weights in other breeds, statistically significant differences between breed equations were found. Therefore, a generalised prediction equation has been derived by fitting the tissue areas from the three CT images as covariates and breed as a fixed effect in a generalised linear model.

The research in this thesis comprises the improvement of the first two stages of this current process, in terms of efficiency, accuracy of estimation results and

computation time. Although not discussed further, stage three is mentioned for completeness.

## 1.5 Outline of the thesis

In Chapter 2, a concise history and the process of X-ray CT are presented, and the differences to conventional X-ray images are highlighted. The hardware and software of the scanning system of a typical X-ray CT machine, and in particular the SAC-BioSS CT scanner, is discussed, together with the different modes of scanning available.

We propose to use deformable templates to automate the segmentation stage of the process, which is currently performed manually. In Chapter 3 the problems with low-level segmentation techniques for these X-ray CT images of sheep are presented, hence stressing the requirement for an alternative approach. The advantages of using deformable templates are considered and a concise review of the literature in this field is provided. The review is separated into free-form models and parametric deformable templates, although the latter are predominantly discussed as this approach is employed in Chapter 4.

In Chapter 4, a parametrised template model is constructed for the distribution of boundaries which encircle a sheep's internal organs in order to identify the carcass. This is achieved by modelling a training set of 24 lumbar images similar to Figure 1.4 (c). Fourier coefficients are used to parametrise the boundaries and the matrix of coefficients is reduced in dimensionality using principal components. As a measure of fit to further images we use a criterion which combines the local pixel gradient with the estimated distribution of boundaries. This criterion is optimised using the Nelder-Mead algorithm, and the model is validated on an independent set of images from sheep of the same breed and age. To complete Chapter 4, we investigate the effect of adding the mirror images of each of the 24 lumbar regions to the training set.

Chapter 5 is concerned with estimation of the relevant tissue proportions after automating the segmentation of these images. In approaching this it is appreciated that many of the pixel values are generated by a mixture of two or more tissues, due to the finite resolution of the X-ray CT machine. The spatial response of the CT machine is investigated by examining a sharp boundary in the image. This is used to derive a new probability density function for the mixed pixels and hence a distribution of greyscale values of all the pixels in the image.

In Chapter 6, the proposed probability density function is fitted to simulated histograms of pixel values containing two tissue types, using maximum likelihood. The results are compared with those obtained using the threshold method currently in use at the SAC-BioSS CT unit. To complete this chapter, a pixel's spatial context is considered by examining the greyscale values of its neighbours, in an investigation which is shown to produce more efficient estimation results than the maximum likelihood approach.

# Chapter 2

## X-ray computed tomography

### 2.1 Introduction to X-ray CT

When X-rays were discovered in 1895, they provided a non-invasive, internal visualisation of the body's structures without surgery. The potential of this as a new medical diagnostic technique was immediately recognised. Since then the technique has been improved through the development of more sophisticated and effective instruments. Due to the ionising effect of the X-rays, the patient is liable to some risks, but these are generally accepted because of the advantages of the direct visualisation of structures: see Shrimpton and Wall (1995).

In the ordinary medical use of X-rays, the image is something like a shadow. All points in the path of any X-ray beam are projected onto the same point on the detector film, resulting in structures being superimposed on the photographic image, rendering detailed visualisation of these separate structures difficult. Figure 2.1 shows an example of a sheep which has been X-rayed in this way. The positioning of the bones can be fairly clearly seen, but other organs and soft tissues are not so easily detectable.

This form of conventional X-ray imaging technique has several important limitations:

- Due to the superposition of a three-dimensional structure onto a two dimensional detector, much of the detail on features is lost. This is very evident from Figure 2.1.
- Small differences in the attenuation values of different tissues are not detectable in the values that the X-ray film records.
- A large percentage of the transmitted X-rays are scattered from the patient



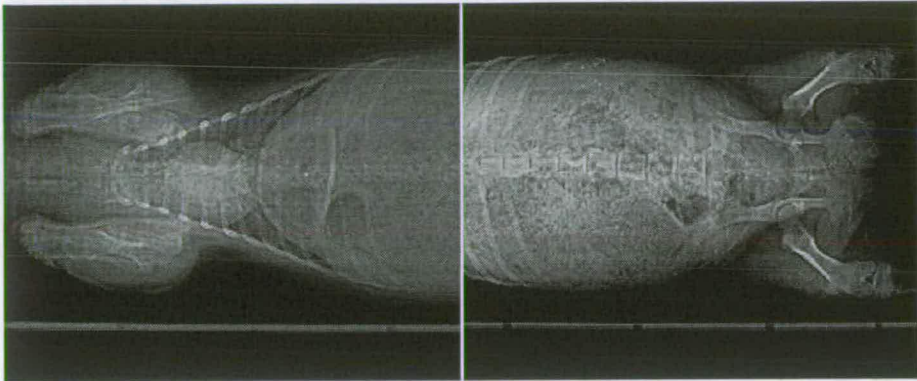


Figure 2.1: An example of a conventional X-ray scan (known as a topograph) of a typical sheep. The sheep is scanned in two images as the table can only be moved by the computer by a certain amount, which is shorter than the full length of the sheep. From left to right, the head, ribs, heart and legs can be seen.

and are not recorded by the detector. This therefore reduces the signal to noise ratio of the recorded information.

When X-ray CT (computed tomography) was developed in the 1970s, it provided a more precise method of viewing the location of structures inside the human body than had previously been available. It is termed *X-ray CT* because it makes use of X-ray attenuation (reduction in intensity of X-rays due to tissue absorption and scatter). 'Tomography' is derived from a Greek word meaning 'section' where a computer (rather than an X-ray film) is used to produce an image of the subject. Unlike conventional X-ray imaging, X-ray CT provides a picture of a single slice, (or cross-section) termed a *tomograph*, through the body, without structures being superimposed. Tomographs are constructed from thousands of X-ray beams that lie in the plane of the cross-section. As well as enhanced image quality, the development of X-ray CT has led to reduced examination times. Therefore, with the introduction of X-ray CT, many of the problems associated with conventional X-ray imaging were either eliminated or greatly diminished, see Shepp and Kruskal (1978).

Today tomography is used in a wide variety of fields, such as medical imaging, seismology, and underwater acoustic imaging. Although the main use of X-ray CT is in medical imaging to detect tissue abnormalities, it has also been used to study the growth and development of animals.

## 2.2 X-ray CT and its history

The fundamental purpose of CT is to reconstruct an image of a cross-section of an object using data collected from many individual beams of X-rays that are passed through the cross-section. The data from the beams is processed by a computer, which uses a mathematical algorithm to convert the X-ray attenuation measurements into a two-dimensional cross-sectional image, which is displayed on a video screen. This process is known as *image reconstruction*.

In 1917, an Austrian mathematician, Radon (1917), first investigated the mathematical principles which still have such a central role in tomography. He proved that a two or three-dimensional object could be reconstructed uniquely from the infinite set of all its projections. Radon's formulation later became known as the projection formula.

In the 1930s, conventional tomography (also known as longitudinal tomography) was developed in order to overcome the problem of structures being superimposed, which existed with traditional X-ray imaging. Conventional tomography differs from traditional X-ray imaging in that the X-ray source and detector move in opposite directions to each other. This is performed in such a way that only points in a plane parallel to the film are in focus on the film. This method does not completely remove the superposition problem, but results in sharper focus of the objects in the plane than those not in the plane.

Recognising the restrictions, Kuhl and Edwards (1968) introduced transverse section scanning in an attempt to avoid them. This technique is a form of the basic reconstruction method, known as *summation* or *simple back projection*. However, it also has limitations, in that due to the reconstruction, blurring of sharp features occurs.

Bracewell and Riddle (1967) proposed the mathematical approximation of Radon's integral formula in radioastronomy, which came to be known as the *convolution method*. This method was modified and refined by several investigators, including Cormack (1963) who proposed it for medical use, and was to become the reconstruction method used by most X-ray CT scanners developed in the latter half of the 1970s.

In the late 1960s, Hounsfield (1972) was independently developing the ideas that mathematical techniques could be used to reconstruct the internal structure of the body from a number of different X-ray measurements. As a result of his work, he developed a numerical, executable, mathematical solution to the reconstruc-

tion problem. Hounsfield (1972) also showed that he could produce considerably more accurate measurements of the absolute value of X-ray attenuation coefficients within the body with a quantitative tomographic technique rather than the conventional X-ray imaging. In 1971, when working for EMI Ltd, he invented the first clinical X-ray CT machine to give an image accurate enough to be of value in medical diagnosis. Mathematical analysis using Fourier transforms has led to algorithms which are much more accurate and efficient than the algorithm used in the first commercial tomography machines.

The most widely used technique for comparing the three reconstruction algorithms mentioned (summation, convolution method and the Fourier transform method) has been to compare the reconstructed images from each method when applied to data taken from real human objects. Another technique is to use *phantoms*; this means taking data from a man-made object of known structure instead of a human subject. The most well known phantom in reconstruction is a head phantom: see Shepp and Logan (1974). The mentioned image reconstruction methods are not described in this thesis, but full details may be found in Robb (1982), Rosenfeld and Kak (1982) and Schalkoff (1989).

## 2.3 The scanning system of an X-ray CT machine

The scanning system consists of 5 main components, which can be seen in Figure 2.2. These are

- An operator's console from which the rest of the system is controlled.
- A bed on which the subject lies (the position, relative to the gantry, is under the control of the computer).
- The gantry which contains the X-ray source and detectors (both of which can spin around the subject).
- The computer, which is responsible for co-ordinating all system functions and calculating the image reconstruction.
- High voltage generator.

An X-ray CT image, or tomograph, is obtained with the subject placed in the aperture of the gantry. An X-ray source spins round the subject sending a thin focused beam of X-rays through the subject. On the opposite side to the X-ray source an array of X-ray detectors measures the degree to which the X-ray beam is attenuated. The amount of absorption of the X-rays depends on the physical



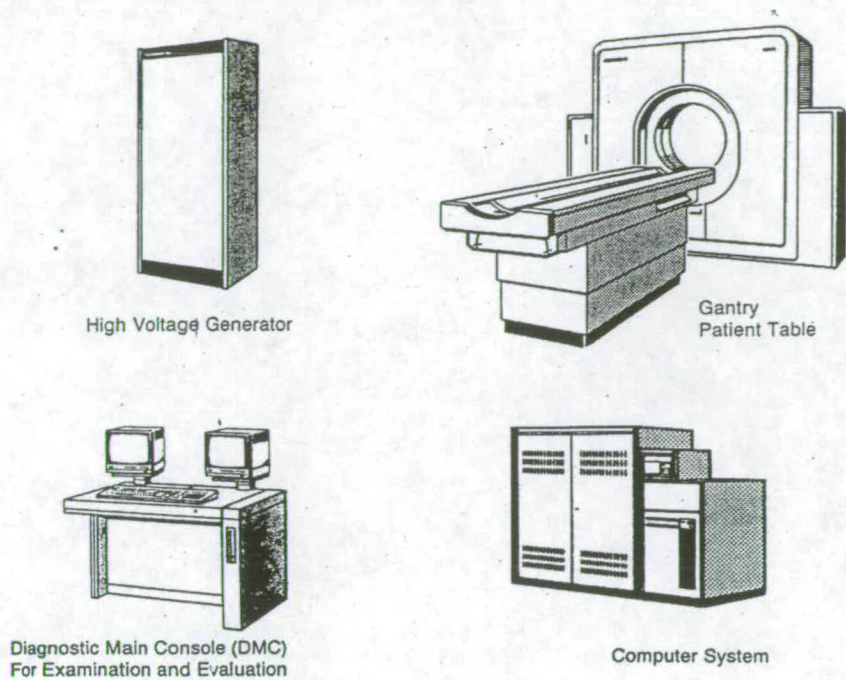


Figure 2.2: *The main components of a typical X-ray CT machine. This image was taken from the operating manual of the Siemens Somatom scanner.*

density and atomic composition of the structures which they pass through and on the energy of the X-ray beam. Although two spatially separated X-ray beams of equal energy may be recorded by the detector as having the same total reduction in intensity, they may have passed through entirely different materials. This is because attenuation of the beam is also dependent on the length of the path through the object, as well as the composition of the path. The computer records the degree of attenuation at many points around the subject. It then builds up a picture of a two-dimensional slice through the subject, describing the density of each point in that picture. This density is based on the degree to which the X-ray beam is attenuated from the different positions around the subject. The finite number of attenuation values corresponding with the scanned object are organised into a matrix. The translation of these numbers into Hounsfield units, (see Chapter 1) and hence corresponding grey-levels, then creates a visual image of the scanned cross-sectional area. The number of pixels in the reconstructed image is dependent on the number of individual projections through the subject and therefore also influences the quality of the image resolution.

Obtaining a tomograph at a precise anatomical location is aided by the use of *topographs* or *scout images*: a topograph for a typical sheep is shown in Figure 2.1. These look like conventional X-ray plates and are obtained by holding the X-ray source in one position and slowly moving the subject past it. The operator

of the scanner can then locate the position of a desired slice from this topograph and the computer will move the motorised bed so a tomograph image is taken at that exact location. Technically, X-ray CT scanners can produce whole body scans, but in practice this is not realistic as more than 100 scans per animal would have to be taken and analysed.

### **2.3.1 Modes of scanning**

There are two possible modes used for scanning the cross-section. These are the *parallel* and the *fan-beam* mode. Since the development of CT scanners, the designs of the scanners have changed in order to become more efficient. Many texts give detailed accounts of the various types of scanners, such as Herman (1980), Robb (1982), Anton and Rorres (1991) and Hiriyannaiah (1997). The current engineering emphasis is on speed of scanning and the reduction of the number of moving parts, which is achieved by the use of many detectors in combination with a fan-shaped beam of X-rays.

#### **Parallel mode**

In the parallel mode, a single X-ray source and the X-ray detector pair are moved in parallel in a direction perpendicular to the line connecting the source to the detector. Many measurements of the parallel beams are recorded. Then the source and detector pair are rotated through a small angle (typically one degree), and another set of measurements is taken. This is repeated until the desired number of beam measurements is completed. Advantages of this method are that there is little noise due to scatter and the detector can be calibrated at the beginning of each parallel beam scan. The disadvantages are that it is very time consuming and inappropriate for imaging organs which cannot stay stationary for more than a few seconds. Also, undue exposure to ionising radiation for extended periods are involved. For example, in the original 1971 machine, 160 parallel measurements were taken through 180 angles spaced 1 degree apart, a total of 28,000 beam measurements. Each scan took approximately five and a half minutes.

#### **Fan-beam mode**

This type of scanner was introduced to speed up the scanning process without losing most of the desirable features of the parallel mode. In fan beam tomogra-



phy, the X-ray beam is collimated so that a thin, planar fan beam of rays diverges from the X-ray source and passes through the object before being collected by a detector array (containing between 200 and 1000 detectors) on the other side of the field of view. This method involves only one motion, where the X-ray tube and detector array are rotated around the patient through many angles (approximately 500) and a set of measurements is taken at each angle. In each source position, the fan beam completely covers the object and each complete scan is performed within several seconds.

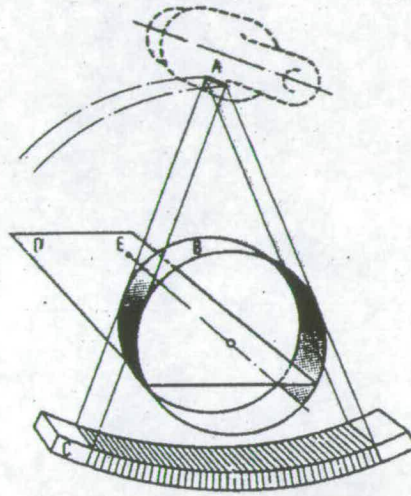


Figure 2.3: *This shows how the equiangular fan beam system appears. The X-rays are emitted from A, and C is the circular arc containing the detector array. This was taken from the Siemens Somatom scanner operating manual.*

There are two major detector configurations in the fan beam mode, (for more details see Hiriyannaiah, 1997).

**Equiangular fan-beam:** The detector array lies on a circular arc (or a complete circle in some present day scanners) such that the angle between two detectors is constant as well as the distance between the detectors. If the detectors were collinear then the spacing between the detectors would obviously be unequal. Figure 2.3 shows an equiangular fan beam system with rotating X-ray tube and a circular arc containing the detector array. The enclosed circle in this image is the gantry in which the subject lies.

**Equidistant collinear detector fan beam:** The detectors lie on a straight line and are equidistant from each other, but the angular interval is not equal.

## 2.4 X-ray CT scanning of sheep using the SAC-BioSS scanner

The SAC-BioSS scanner is a Siemens Somatom CR system, and the main components of this machine are shown in Figure 2.2. This is a fan-beam scanner, and the gantry contains the rotating X-ray tube-detector system. The sheep lies on a longitudinally movable table top, normally in the supine position with its head pointing towards the gantry. Reconstruction and storage of the image is performed using a  $256 \times 256$  matrix, with a choice of attenuation measurements taken at either 360 or 720 angles in order to produce a single scan: refer to Subsection 2.3.1 for further details.

Different durations of X-ray pulses are available (2 and 4 ms), as are different slice thicknesses (2, 4 and 8mm): the 2mm option is used for all the tomograph images presented throughout this thesis. The measured attenuation values are converted into the international Hounsfield unit (HU) range: see Chapter 1.

### 2.4.1 The scanning strategies used for sheep

There are two strategies commonly used in CT scanning of sheep. One is used to predict the relative body composition of the animal from a few (2-7) images, taken at specific anatomical locations. These specific locations are found using the topograph, see Section 2.3. The other method is used to estimate whole body tissue size from 15-20 images.

The first strategy is used when knowledge of the relative tissue size is sufficient. The few anatomical sites are chosen according to their likelihood of having a strong relationship with body composition or to allow assessment of a number of anatomical entities from only a few images. This approach is best suited to comparisons of large numbers of animals within or between breeds. However, it may overestimate or underestimate differences between individual animals with different body shapes. Body and carcass composition can then be predicted from tissue areas in each slice using equations derived from calibration studies. At the SAC-BioSS CT unit, three anatomical locations from each sheep are used in the reference slice approach, namely the ischium (upper legs), lumbar (lower back region) and thorax (chest) images: see Chapter 1 for examples of scans at each of these positions. These particular positions are chosen based on expert knowledge of positions which provide the most accurate analysis of the full animal composition. Comparison of different sets of predictors was carried out using stepwise

regression techniques. It was found that these three CT scans consistently topped the list of the best predictors. Also, these three predictors include one each from the three main regions of the carcass: leg, loin and chest/shoulder.

The second strategy is known as the Cavalieri approach, and is named such as it is implemented using Cavalieri's Theorem, a 14th century Italian mathematician (see Sugakkai, 1987). The volume of each tissue can be estimated accurately by taking 12-15 images randomly positioned but equally spaced, and measuring the area of each tissue in the image and then multiplying the sum of these areas by the inter-image distance (for the SAC-BioSS machine this is 2, 4, or 8mm). This approach is used when the absolute tissue size is required to be known, such as in studies of nutrition.

# Chapter 3

## Deformable templates

### 3.1 Introduction

Deformable templates are a model-based approach to extracting structures from images, which are able to accommodate the significant variability of biological structures across individuals and over time. They were originally developed for application to problems associated to computer vision and computer graphics, to segment, visualize, track and quantify a diversity of structures. In particular, in segmentation, the aim is primarily to extract boundaries belonging to the structure and to use this information to build a consistent model of the object of interest.

This chapter explains how segmentation was traditionally performed, (see Section 3.2) and goes on to discuss the disadvantages of these methods. Deformable templates were introduced and developed in order to overcome these deficiencies. In addition, a review is presented of the various types of template that have been employed by authors since they were first introduced. The review mainly focuses on the parametric deformable templates in Section 3.5, since they incorporate prior information from the images into the model. This has been the most commonly used approach in recent years.

### 3.2 Low-level segmentation

In many medical and biological image analysis tasks, the first stage is to segment the image into subsets of pixels corresponding to the different tissues and structures present. In order to be able to identify objects of interest, it is a common approach to extract features such as boundaries. Using segmentation methods

based on these features, the image is separated into its constituent regions and can then be classified by identifying which object corresponds to each region. In medical and biological images, there is natural variability both between and within subjects. That is, almost every object of interest in the body can vary in size, shape, orientation, location and appearance. This makes the task of identifying and segmenting the image into the structures of interest very difficult.

In many applications, segmentation is performed manually, where a skilled operator manually traces round the region of interest in each image, using a computer mouse. This approach has serious weaknesses in that there will be operator bias or fatigue, since it is a very tedious, subjective and time-consuming process. It also raises the question of whether the operator would be able to produce a similar segmentation if carried out on different occasions. Therefore, for manual segmentation to be reliable, consistent and reproducible, expert knowledge of the features and underlying image is required.

Other approaches involve using low-level automatic segmentation techniques, for example grey-level thresholding or the application of edge operators. However, these methods also have many well-known problems. Thresholding is generally successful only if there is little overlap in the greyscale pixel values in the different regions of the image. Edge detectors for boundary fitting have the problem that they completely rely on the local neighbourhood of pixels in the original image and can therefore generate infeasible object boundaries, spurious edges and gaps in the objects of interest. In addition, the edges found do not necessarily correspond to the actual boundaries of the objects. Furthermore, they are of little use in noisy images, e.g. ultrasound images, and they ignore model-based information and the higher-order organisation of the image. Therefore, these model-free techniques also require expert intervention as they are unlikely to work well due to the under-constrained nature of the problem.

These weaknesses for both manual and low-level processing show there is a need for quantitative data to be efficiently extracted using a more accurate, repeatable, objective and automatic method. That is, in general, segmentation cannot be addressed adequately without 'high-level' prior knowledge of the size, location, intensity of tissues, the number of objects the image contains and how they interact with each other.

The consequence of these disadvantages is that powerful models are needed which are not only robust against noisy data but are also capable of accurately representing complex shapes of many anatomical structures. In order to be fully effective,

these models must allow for expected variation in size, shape and appearance of the objects in the images. However, segmenting structures and reconstructing a compact geometric representation of the structures is difficult because of the size of datasets and the complexity and variability of the anatomic shapes of interest. A solution to this problem is the elegant and robust method of deformable templates (with image preprocessing), which will now be described in detail.

### 3.3 Deformable templates

The use of deformable templates is a model-based approach to extracting meaningful structures from the data, which accommodates the significant variability of biological structures across subjects. They may be used to overcome the difficulties and drawbacks of manual interpretation and traditional image processing techniques, as mentioned in Section 3.2. They are also popular because they have the ability to combine low-level knowledge from the image with high-level knowledge about the location, size and shape of features such as the anatomic structures.

By using parametrised geometric models of objects that are likely to be present in the image, results from imperfect data can be improved. The amount of shape information incorporated into the model can vary from local and general to global and very specific. For example, the model can incorporate only smoothness or continuity constraints or may specify the exact shape of the object using e.g. a ‘hand-crafted’ parametric form, see Subsection 3.5.1. In general, the overall template can consist of more than one object template, and when it is overlaid onto the image, it should segment it into the regions of interest (see Phillips and Smith, 1994). Most authors address the problem of object identification as a process of boundary finding and incorporate the global shape information into the model. In order to exploit model-based information, the knowledge of the shape should be included as explicitly and specifically as possible. Obviously there will be differences between the objects present between individuals due to natural variation, varying imaging conditions and noise. Deformable templates are adaptable and flexible in the modelling of the object of interest and can cope with variations while still maintaining the certain structure of the object. These models of shapes and appearance of flexible objects may be used when searching the images for new examples of the objects. Such models usually have a number of parameters to control the shape and size of all the parts of the model. Structures in the image can therefore be considered as a deformation of a given template.



Early attempts at template matching were undertaken by using rigid templates, which are restricted to transformations of rotation, translation and scaling. As they are very simple models because of the limited degrees of freedom, their use is very restricted, since biological applications undergo non-rigid transformations. However, a deformable template can undergo both non-rigid transformations and the transformations mentioned above. They can adapt to fit any given image data and they have the ability to impose geometrical constraints on the shape. Also, it is possible to include local information from the image.

The concept of template matching was introduced by Fischler and Elschlager (1973) to model facial features using a set of basic rigid features connected by a mathematical representation of springs. The individual features (eyes, mouth, nose, hair) do not deform and each feature has a local measure of fit to the image. The springs joining the rigid features act not only to constrain the relative movement, i.e. ensure that the spatial relationships between the features remain reasonable, but to measure the 'cost' of the movement by how much they are stretched from the equilibrium position. During the matching process the entire structure is deformed until all the features have a good local fit in an image and the spring forces are balanced. Although it has the benefit of being a very general approach, the main weakness is the simplicity of the local fitness measure that Fischler and Elschlager (1973) proposed to use. It is strongly scale dependent and may fail with certain noise levels in the image.

The deformable templates used in previous literature fall into two categories:

- Free-form models
- Parametric deformable templates.

These will be discussed in Sections 3.4 and 3.5 respectively. However, before they are discussed it should be pointed out that there are two ways in which to quantify the success of a particular instance of a deformable template and distinguish which template provides the best segmentation for a given image. This can be achieved by either using an energy minimising approach, or the more recent approach, which uses a Bayesian framework.

### 3.3.1 Energy minimising approach

Here an energy function is defined which consists of two separate parts, namely the internal and external energy functions. Each of these individual functions may consist of one or more terms themselves, (see for example Yuille et al.,

1992). The internal energy is independent of the input image and is related only to the geometric shape, which is a fundamental property of the template. The external energy function is related to the image data in some way, depending on the object of interest, e.g. edges, lines, valleys or peaks. However, in most papers the authors specify the interaction between the image and the template in terms of the image greyscale values and/or edge information from the image. In principle it is possible to combine other types of information such as texture or colour. Combining the external energy with the internal energy allows the template to interact with the image and it is therefore attracted to the desired prominent features. In Section 3.4 the internal and external functions used by Kass et al. (1988) are discussed. The two energy functions are combined generally in such a way that the overall energy, also called the objective function, is minimised by the optimal template.

### 3.3.2 Bayesian formulation

A convenient approach is also to view the problem in a probabilistic framework, hence allowing the incorporation of very specific prior knowledge about features and the spatial relationships between them into the template model. The probabilistic framework also provides a measure of uncertainty of the estimated shape parameters after the model is fitted to the data and is used to make inferences about the parameters of interest. The prior model usually represents the initial knowledge of the distribution of the template parameters. We let the vector of parameters of the model be denoted by  $\theta$ , with prior probability density  $p(\theta)$ . The prior density must accurately reflect the variation between the parameters if the general template is to be capable of adapting to closely represent a sufficiently wide range of particular cases of the modelled object or objects. The likelihood function (also known as the observation or imaging model) measures the ‘goodness of fit’ of a given template with the observed image, denoted by  $p(I|\theta)$ , where  $I$  represents a certain image, (see Fisker and Carstensen, 1998). In many cases the likelihood is thought of as specifying how well the existing templates agree with the boundaries of the images, i.e. a measure of the edge strength agreement. For examples of Bayesian formulations, see Phillips and Smith (1994), Rueckert and Burger (1997), Glasbey (1998). These papers will be discussed in Subsection 3.5.2. The likelihood function is chosen according to the specific problem in hand. Using Bayes theorem, the posterior density,  $p(\theta|I)$ , of the parameters given the image is obtained, where  $p(\theta|I) \propto p(I|\theta)p(\theta)$ . The posterior density,  $p(\theta|I)$ , (also denoted the objective function) can then be maximised to find the parameters of

the best fitting template for that given image.

### 3.3.3 Optimisation of the energy and Bayesian approaches

The internal energy function can be interpreted as a prior distribution over the expected shapes, using the Boltzmann or Gibbs distribution. Similarly, the external energy can be shown how to be related to the likelihood, therefore showing that the approaches are equivalent. For full details on this relationship between the two methods see McNerney and Terzopoulos (1996) and Jain et al. (1998). Several authors have assigned weights to the two different components of the given objective function, both in the Bayesian and energy formulation, see Glasbey (1998). The choice of these weights and other parameters of the template affect the overall success of the model.

Various methods are used in the literature for the optimisation of the objective function (from either the energy or Bayesian approaches), using either deterministic or stochastic techniques. The choice of optimisation technique is not discussed here, but various approaches are mentioned in Chapter 4. As part of the optimisation stage, some authors have proposed to use a multi-resolutional approach proceeding from a coarse to fine resolution in a number of steps: see Jain et al. (1996).

## 3.4 Free-form models

McNerney and Terzopoulos (1996) and Jain et al. (1998) provide very comprehensive reviews of these free-form methods. They are used because they can represent any arbitrary shape as long as some given constraints such as continuity and smoothness are satisfied. They are generally free to follow almost any smooth boundary as they contain only a few constraints on the overall shapes. In general, they are used to approximate the locations and shapes of boundaries, based on the assumption that the boundaries are piecewise continuous and smooth. Among the first and primary uses of deformable models was the application of deformable contour models, also known as active contour models (see Kass et al., 1988). They refer to their model as a ‘snake’, which they describe to be

*an energy minimising spline guided by external constraint forces and influenced by image forces that pull it towards features such as lines*

*and edges. Snakes are active contour models. They lock onto nearby edges, localising them accurately.*

Kass et al. (1988) describe the model to be active because it is always minimising the energy function and can adapt to any given shape. Snakes do not try to solve the global problem of finding prominent image contours as they rely on other mechanisms to place them somewhere near the desired contour. However, from any starting point, the snake deforms itself to resemble and fit the nearest prominent contour. In most cases where deformable models are used, they have associated deformable energy functions, which increase monotonically as the model deforms away from the specified natural position. Kass et al. (1988) optimise in the image space and this makes it difficult to incorporate global shape information. Other authors, including Cootes et al. (1994) and Glasbey (1998), optimise in the parameter space. Kass et al. (1988) propose to minimise the energy function which is a weighted combination of three parts:

1. An internal contour energy which characterises the deformation of a stretchy, flexible contour and which serves to impose a piecewise smoothness constraint,
2. An image energy which attracts the contour to prominent features of the image like lines and edges,
3. An external constraint energy, which is responsible for putting the contour near a desired local minimum.

To apply snakes to images, image energies are designed whose local minimum coincide with e.g. the maximum and minimum greyscale values, edges and other features of interest. If the snake is designed to be attracted to edges, then it is fairly common to convolve the edge image with a Gaussian smoothing filter, see Subsection 4.9.1. This has the effect of controlling the spatial extent of the local minimum of the energy function.

At the time snakes were introduced, they offered a unique semi-automatic way of locating edges, lines and contours using essentially a single method. However, they have limitations in that they only use local information from the image and therefore do not try and solve the global problem of finding prominent image contours. They are very sensitive to their initial position and the amount of noise in the image. After the initial position is specified, the snake deforms itself to resemble and fit a nearby edge or line by minimising the energy function. However, it is very likely to be attracted to a local minimum of the energy function rather than the global minimum.

For full automation of this approach, the models need to describe the size, shape, location and orientation of the object, while also allowing expected variations in these characteristics. In many medical and biological images, this information is generally known and therefore as much information as possible should be incorporated into the model. As a result, many authors use parametric deformable templates.

### 3.5 Parametric deformable templates

Jain et al. (1998) provide a good description of the parametric deformable models. A parametric deformable template is a parametrised geometric model of the shape or the object boundary being modelled, usually using a small number of parameters. These types of deformable models are preferable to free-form models because prior information of the global structure of the object is included in the template. An advantage of this type of global model is that gaps in the boundaries which occur from noise or occluding objects are bridged. As stated earlier, in biological and medical images the location, shape, size etc. is generally known and this information may be incorporated into the template model in the form of initial conditions, data constraints, constraints on the model shape parameters or directly into the model fitting procedure. Parametrisation of the template allows for deformations in its shape and by specifying a probability distribution over the parameters makes it possible to quantify the relative probability of the likely types of deformation.

Parametric deformable templates are similar to snakes as they also address the problem of matching objects with known shapes. However, they differ in that the parametric form comprises more prior knowledge about the object being detected. Also, the forces on the parametric templates are global rather than local, therefore reducing the number of local minima in the objective function. Since they also involve only a small number of parameters, they are fairly easy to implement.

There are two ways in which to parametrise the shape and variation in these types of templates:

1. Using a collection of parameterized curves, known as ‘hand-crafted models’.
2. Using a prototype template.

In both of these methods, the prior shape preferences are clearly expressed by the parameters of the model. The deformable template interacts with the features

of the image by changing the parameters of the template to alter the measure of fit with the image. Each value of the objective function quantifies how well the template matches the object in that given image. By searching the possible parameter values using an optimisation technique, the set of parameters which optimises the objective function may be chosen.

### 3.5.1 Hand-crafted models

These deformable models are built from simple subcomponents such as circles, ellipses, lines or arcs. The components are allowed some degree of freedom to move around relative to each other and have the possibility to change scale and orientation, for example see Yuille (1991) and Yuille et al. (1992). For this type of approach it is necessary that the shapes used for the templates are of known structure. The variations in the shapes of the template are determined by a probability distribution placed on the allowed parameters. Yuille (1991) and Yuille et al. (1992) have used this approach to extract facial features and determine the spatial relationship between them. In particular, they have designed models for the eye and mouth using circles and parabolic curves. They use an energy-minimisation approach where the external energy function is defined in terms of edges, peaks and valleys of the greyscale values and also the original image greyscale values. They find that for these types of templates, a good starting point for the contour is required in order to obtain meaningful results. They also assume that the orientation and scale of the individual templates are known. Lipson et al. (1990) use a similar scheme to map elliptical models of vertebrae onto CT images of spines. Hill and Taylor (1992) use these hand-crafted models for identifying the left ventricle of the heart in ultrasound images. They first select control points on each image and use parametric cubic splines which interpolate these control points. To find the set of parameter values which best position the template in the image, they search through the parameter space, projecting instances of the template back into the image until the most consistent with the observed is found. They again use an objective function based on edge evidence within the image in order to evaluate the degree to which the image supports a particular instance of the model.

Grenander et al. (1991) represent shapes as a set of boundary points which are connected by arcs, with a statistical model of the relationships between neighbouring arcs. They also show that by using the model of the outline of a hand that it can be deformed to fit degraded images of hands. They are able to achieve this by considering sections of the boundary and determining the most probable

positions given the rest of the boundary and through incorporating local image data.

The advantage of this type of modelling lies in the simplicity of the template because of the limited degrees of freedom generally used. Also, models of this type are capable of capturing detailed knowledge of the expected shapes. However, this approach has distinct disadvantages in that it lacks generality, since the templates have to be hand-crafted for specific applications and they are not easily applicable to complex objects. Also, the shape being modelled has to be well-defined in order for it to be represented by a set of curves with (preferably) a small number of parameters. These templates are generally restricted to scaling, rotation and translation whereas in most cases, anatomical features also undergo non-rigid deformations, due to their own deformation as well as deformation from adjacent structures.

### 3.5.2 Prototype based templates

This is a more flexible approach to designing a deformable template and is the most commonly used method in the recent literature. In these types of templates the object is defined around a ‘standard or prototype’ template which normally describes the most likely or characteristic shape of the objects to be identified, where the objects have a global structure but deviate from this structure from individual to individual. The prototype template is selected by using prior knowledge of the objects of interest or else it is obtained empirically from training samples, e.g. see Baldock (1992), Cootes et al. (1995) and Glasbey (1998). A parametric mapping of the prototype is used in order to obtain different instances of the object. As with the ‘hand-crafted’ templates, (in Subsection 3.5.1) different parameter values give different instances of the object. The pattern theory proposed by Grenander (1993) describes a framework to represent classes of shapes that show a substantial amount of variation.

Success in using this type of template is dependent on the accuracy of the description of the object and the variations allowed within the template parameters.

The discussion of the literature using these types of templates will be divided into the two methods mentioned in Subsections 3.3.1 and 3.3.2, although the majority of the papers use the Bayesian approach. Another method, using a landmark approach, is also described.

## **Prototype templates using an energy minimising approach**

Baldock (1992) models the boundaries of the heart chambers in two-dimensional ultrasound images and represents these boundaries as a number of curved lines. An operator manually draws round the required structures by using line segments. Dynamic programming is used to refine the initial estimates of a structure using the information encoded in the model. An energy function is designed containing four terms, with weights applied to each (similar to Yuille et al., 1992). The model used is entirely empirical and he describes the model as ‘trainable’ because it uses information from many images to update the parameters of the model. The model then uses the knowledge from the manual templates to establish an initial guess of the parameters in the model. An advantage of building in this training mechanism is that during the routine use of the system, the model can be continuously updated by new examples.

Rueckert and Burger (1997) are also interested in segmenting the left and right ventricles of the heart, and their template consists of two ellipsoids. They describe each object as an ordered set of vertices, where they think of the vertices as landmarks which describe the outline of the shapes. They interpolate all the vertices using interpolating splines so that the boundary is smooth between these vertices. They initially approach the problem in an energy minimising framework with an internal and external energy functions, but later describe how this may be viewed in a probabilistic framework also. They describe how they can use the prior information on the shapes of these object boundaries to increase the robustness of the algorithm in situations where part of the boundary is occluded in some way.

## **Prototype templates using a Bayesian approach**

Staib and Duncan (1992) use a deformable contour model mainly on two dimensional echocardiograms to extract the left ventricle of the heart. They represent open and closed boundary templates for various other objects which are smooth and continuously deformable by using elliptic Fourier descriptors. The Fourier coefficients are the parameters of the deformable template and the probability distribution on these coefficients is selected in order to bias the template towards some particular shapes with different amount of variability among the objects of interest. They derive distributions for each of the parameters over a training set and the likelihood function which they use is based on the correlation between the template and the boundary strength in the original image. They then pro-



ceed using a Bayesian approach by locating the maximum of the posterior density (*MAP*), which determines the most likely set of parameter values of the template for each given image. However, the relationships between the variations in shape and variations in the parameters of the trigonometric expansion of Fourier descriptors is not straightforward because the parameters (the Fourier coefficients) are correlated. Like others, they place separate limits on the model parameters on the basis of the distributions determined directly from the training set. Since, as usually is the case, the parameters are correlated over the training set, this approach does not effectively restrict the shape which can be generated to ones similar to those found in the original training set.

Jain et al. (1996) deform the prototype shape of many objects using, like most others, boundary information of the objects of interest. Their template consists of representative edges and a set of probabilistic deformation transformations on the template. Their approach differs from others in that the prior shape information is specified as a binary image. The prototype template is not parametrised but contains edge information in the form of a bitmap, and the deformed templates are obtained by applying transformations to the prototype. The prototype is deformed by locally stretching, squeezing and twisting, and the variability in the shape is obtained by imposing a probability distribution on the possible mappings. Using a Bayesian framework and including a weighting on the prior and likelihood components, they find the *MAP*. Their likelihood is similar to the external energy function described by Kass et al. (1988), but different from more recent approaches in that they use the gradient directions combined with the edge positions in the image. They use a multi-resolutional approach, from a coarse to fine resolution, to initiate matchings at finer resolutions.

At the coarsest stage, a smoothed version of the image is used which has fewer local optima, and this helps to roughly find the global optimum without great consideration to the accuracy of its position. At the finer stages, the amount of smoothing applied to the image is less since more accurate localisation of the template is desired. The template from the previous resolution stage is used as an initial guess. Their approach is very general because they can match curved, (open or closed) polygonal objects using the boundary and gradient information of the image only. They show their method can perform good matching independent of the location, size, orientation and number of objects, present in the image.

Phillips and Smith (1994) develop a hierarchical model to organise prior information to segment many facial features in images. The hierarchical models allow the prior information to be organised in a structure so that at one extreme it considers

local variation in the individual templates and in the other extreme it considers global variation. They make the assumption that the pixel values within each region after segmentation are independent in order to form the likelihood, which is unlikely to be true. However, they find this assumption seems to work quite well, assuming that there is an obvious difference between the means of the segments. They use the hierarchical models to try and reflect the relative importance of the individual features. They construct the boundaries for the head and face using a small number of key points and a closed contour is achieved by fitting a smooth curve through these points. The face is parametrised in terms of proportions of the head template. They also make an assumption of symmetry, that the right side of the face mirrors the left side, when fitting the individual features such as the eyes, mouth and nose. They also use a smoothed version of the original image, similar to Jain et al. (1996), for the initial search of parameters.

### **Prototype templates using a landmark approach**

A number of researchers have built deformable templates by using the distribution of sets of 'landmark' points which mark significant positions on an object. Recent work using this type of template has focused on using training samples to actively learn about the shape models.

The approach which Cootes et al. (1995) adopt is to find a model for the shape representation in which estimates of the shape parameters are uncorrelated over the training set, using principal component analysis (PCA). Having performed PCA, simple limits on each parameter constrain the model to generate shapes similar to those in the training set.

The work by Cootes et al. (1994) and Cootes et al. (1995) on prototype template matching is based on deformations which are learned from correctly segmented training images. They call this method 'Active Shape Models'. They represent objects by sets of labelled points, which are placed in the same way on the object boundary found in each image. Then, the sets of points over the training images are aligned with respect to a set of axes in order to establish similarities between the landmarks of training samples of the same objects. By examining the statistics of the labelled points and using PCA, the point distribution model (PDM) is derived. The PDM method works by modelling how different labelled points tend to move together as the shape varies. They use the largest eigenvectors of the covariance matrix of the labelled points of the shape to describe the average positions of the points and describe the main modes of variation found from the

training set. They propose to use the mean shape as their prototype template. The object boundaries are then segmented using this PDM by combining the current example of the PDM with the image at each iteration. At this point the pixels in the neighbourhood of each landmark point are examined to see if an edge is present there. It should be noted that the landmarks can come from several different objects in the image, (see Cootes et al., 1994, 1995), or from just a single object, (see Cootes et al., 1992). Although in many of their papers in this area, they are interested in locating the left ventricle in a two-dimensional image of the heart, the later papers also include the nearby edge of the right ventricle and the left cavity of the heart in their model. This is to improve the proficiency of the template at locating the desired object.

In an early paper on their work, Cootes et al. (1992) use a method defined as the Chord Length Distribution (CLD), to build models for the left ventricle only. Here, they represent the shapes using polygons with the landmark points. For each polygon in the training set, the chord lengths between pairs of points are calculated. The correlation between pairs of chords is modelled by calculating their means and variances and applying PCA to choose a smaller set of chord vectors which can explain most of the variation in the shape. Although the CLD is better than the PDM at representing objects which can bend, the reconstruction shape from the distances between points is iterative and slow.

The main contribution of this work is that the active shape model is capable of learning the typical pattern of the shape and uses this information to deform the template in ways that reflect the variation from the training set, therefore giving a more powerful description than the active contour models proposed by Kass et al. (1988). However, a weakness of this approach is that they found it is sensitive to partially occluded objects and it is not capable of handling large changes in scale or orientation.

Glasbey (1998) used a similar approach to Cootes et al. (1995), by incorporating a PDM into a Bayesian framework to segment ultrasound images of sheep. Glasbey (1998) also makes the assumption similar to Phillips and Smith (1994) that the pixels in each segment are independent and identically distributed in order to form the likelihood. A weighting is used on the likelihood, similar to Jain et al. (1996).

## 3.6 Summary and conclusions

It has been shown how deformable templates offer an attractive approach to modelling anatomic structures because they are able to represent complex shapes and shape variability. They overcome many of the limitations of the traditional low-level image processing techniques by providing compact and analytical representations of the object shape.

The ideas and methods described in this chapter will be used when a deformable template is designed in Chapter 4 to segment the X-ray CT images of the sheep.

# Chapter 4

## Segmentation of sheep images

### 4.1 Introduction

In this chapter deformable templates are used to segment the sheep images described in Chapter 1 to identify the region of interest, namely the area of the sheep that contains the muscle, fat and bone. Currently, considerable human intervention is required to achieve this. We explore the use of deformable templates to automate the segmentation. The deformable template is constructed in Section 4.3, from a training set of 24 manually segmented images. Fourier coefficients are used to parametrise the template boundaries and the coefficients are reduced in dimensionality using principal components to estimate a distribution on the parameters of the template: see Section 4.4. In order to define a matching criterion between a given template and the boundaries of an image, we examine local edge information in the image in Section 4.7. The Nelder-Mead algorithm is used in Section 4.8 to optimise the objective function, which is a combination of the matching criterion and the distribution of the template parameters. The amount of prior information incorporated from the training set into the template is then updated to improve the performance of this deformable template approach. The results have been validated on an independent set of images, Section 4.10. Finally, it has been decided to add the mirror images of each image into the training set and the same analysis performed (see Section 4.11), to examine the effect on the results by doubling the training set and allowing for symmetry.

## 4.2 Manual segmentation of the sheep images

The low-level methods of segmentation, described in Section 3.2, are unsatisfactory for the segmentation of these sheep images. The desired goal is to extract the region of the sheep which contains the muscle, fat and bone and separate this region from the areas that contain the cradle and the internal organs. The area of interest will be referred to as the ‘carcass region’. Thresholding is unsuitable for the segmentation because, as can be seen from Figure 4.2, the internal organs have similar density to muscle and therefore appear in the image as having similar greyscale values. In addition, the edge detection filters are unsuitable due to the reasons mentioned in Section 3.2. An example of an edge filter applied to a typical sheep image is shown in Figure 4.12 (page 53). It can be seen that the boundaries are discontinuous in this image, and the filter also highlights many boundaries which are not of direct interest. Therefore, expert human intervention is required in order to select the correct boundary to follow from all the possibilities. In many other segmentation problems the object of interest is well defined and the edge detector highlights and identifies the boundary very easily. However, due to the complexity of the shape of the boundary to be extracted in these X-ray images, edge filters used on their own are seen to very ineffective.

Therefore, the segmentation is performed manually at present. As stated in Section 1.4, two boundaries are required to segment these lumbar images. The boundary around the outside of the sheep is relatively easy to find and the method used to select this boundary is described in Section 4.10. In order to identify the boundary which encircles the internal organs, a skilled operator manually traces the region of interest in each tomograph using a computer mouse. In attempting to reduce bias between operators, each skilled operator uses a list of criteria when deciding on the best position of the hand-drawn inner carcass boundary. These criteria are described below. Figure 4.1 labels the appropriate regions of a typical sheep lumbar scan to assist interpretation of these criteria.

- There is no air allowed in the carcass area. Therefore, any air pockets that appear in an image are automatically assumed to be contained in the internal organs region.
- If a large air pocket exists at the top of the rumen, the inner boundary would generally run along the outer edge of the pocket. Typical air pockets in the rumen are visible in Figure 4.2.
- When moving away from the backbone, in either a clockwise or anti-clockwise



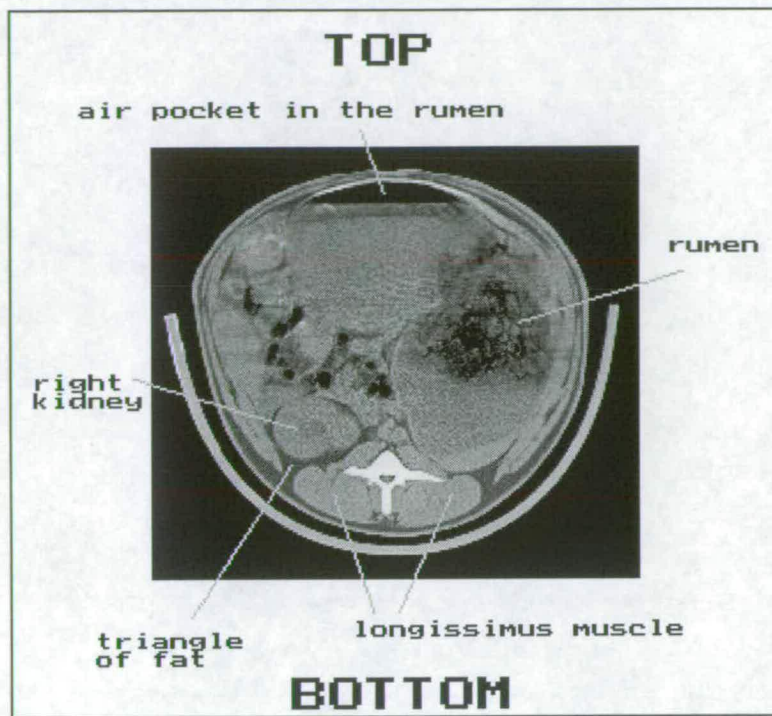


Figure 4.1: A typical lumbar scan of a sheep, with various organs and muscles labelled. These landmarks are used to manually select the inner boundary which encircles the internal organs.

direction towards the top of the image, around the abdominal wall area, choose the inner-most layer of muscle before drawing the boundary.

- Around the abdominal wall contained in the carcass, starting from the back-bone region there is a transition from three layers of muscle, to two and then finally only one layer at the top of the image.
- The triangle of fat at the lateral edge of the longissimus muscle and below the right kidney (seen on the left hand side of the image as an approximately circular light grey object) should remain in the carcass area.
- If one side of the image around the abdominal wall has unclear boundary definition between the fat and muscle and the internal organs, then the boundary is drawn symmetrically with the visible side of the image. (assuming that the outer boundary is symmetric with respect to the axis running through the spinal column.)

## 4.3 Modelling the carcass boundary using deformable templates

As stated in Chapter 1, three anatomical positions are scanned for each of the sheep in the training set. It has been decided that the lumbar image is the most difficult to segment automatically due to the internal organs lying beside the region of interest. Therefore, a deformable template approach is used for segmentation at this anatomical position. Although it has not been attempted here, a similar approach could be employed to build a model to segment the remaining two positions (see Glasbey and Young, 2000).

It is already known from Section 3.3 that designing a deformable template to detect a boundary or object falls into two parts.

1. Providing a parametric model for the template. Once the form of the model is specified, the parameters for each image can be evaluated and the probability distribution of these parameters may be estimated by statistical methods, assuming that the size of the training set is large enough.
2. Given a model for the template, it is necessary to quantify how the template will appear in an image by specifying a measure of agreement, or ‘matching criterion’, in order to determine how well the template fits an image. It is this part that is generally harder.

1. and 2. are combined to form what is denoted as the ‘optimisation function’. The inner boundaries from the training set of 24 manually segmented lumbar images, described in Section 1.3, are used in the construction of a stochastic model for the distribution of these inner boundaries which segment the images into the region of the internal organs and the carcass region. Images (a) and (b) in Figure 4.2 show typical X-ray images before segmentation and (c) and (d) show the same images with the manually drawn boundaries superimposed. Figure 4.3 shows the extracted carcass region from Figure 4.2 (c), with all other pixel values set to zero.

The knowledge of all the realised shapes, sizes and locations of the inner boundaries within the training data is used to help build a probability distribution for the parameters of the template, based on the parametric model. Figure 4.4 shows some inner boundaries taken from the set of 24 training images. The variability in size and slight differences in the orientation at which the sheep are lying can be seen.

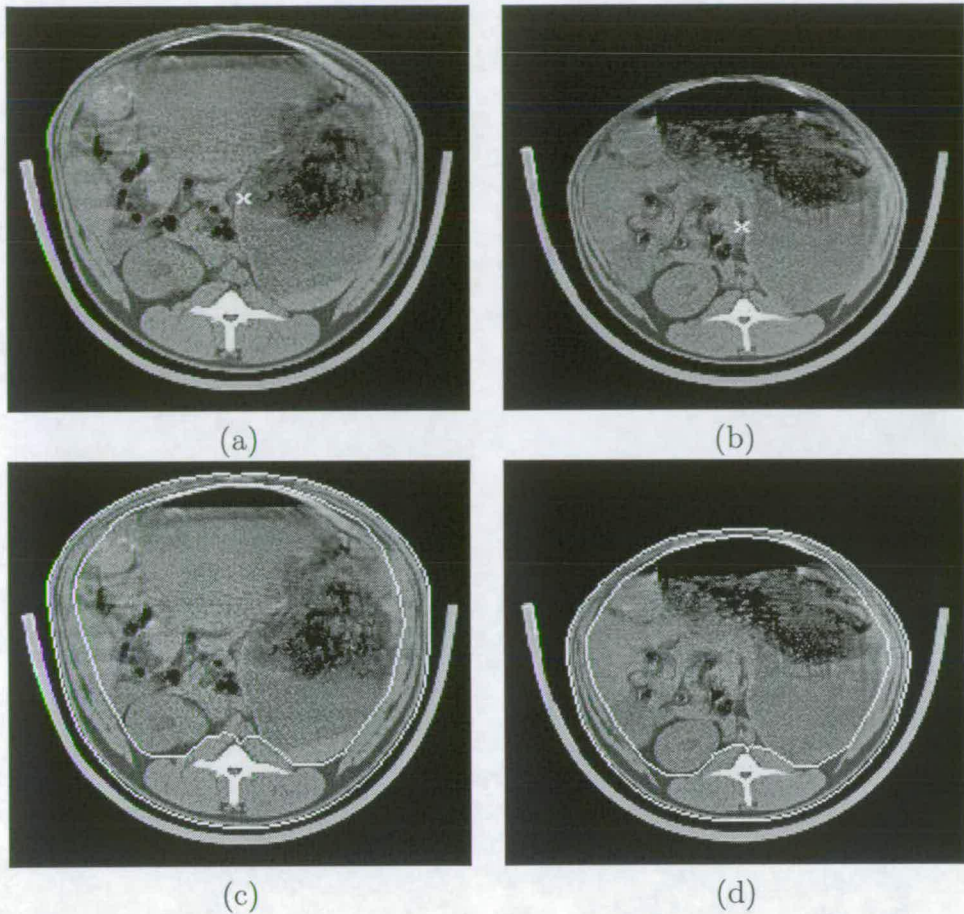


Figure 4.2: *Examples of X-ray CT images before ((a) and (b)) and after ((c) and (d)) manual segmentation. The inner and outer hand-drawn boundaries are superimposed in white on the original image. The white 'x' in (a) and (b) corresponds to the centroid of the outer boundary.*

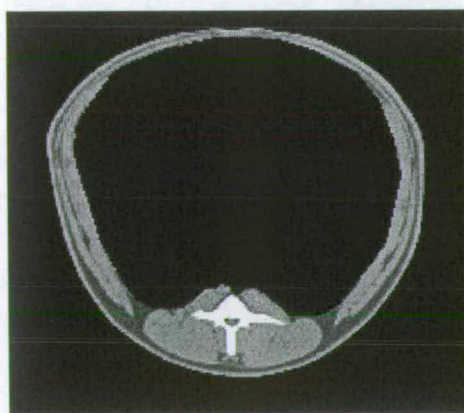


Figure 4.3: *The carcass region of the sheep in Figure 4.2 (c): the fat, muscle and bone can be seen to lie between the manually drawn boundaries which have been used to 'blank out' the internal organs.*



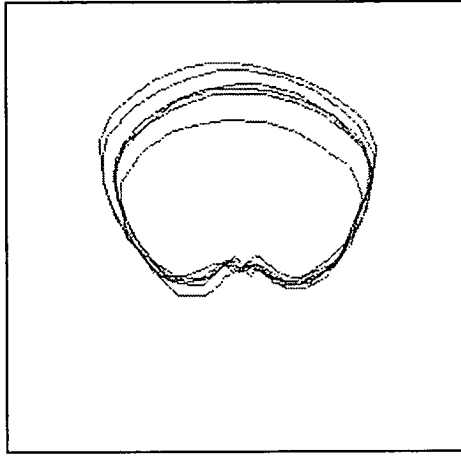


Figure 4.4: *Some realisations of shapes, sizes and locations of the inner boundary taken from the sample of known inner boundaries in the training set.*

There are many ways in which to parametrise the template, e.g. splines (Rueckert and Burger, 1997), landmark points (Cootes et al., 1995), polygons (Cootes et al., 1992) and using hand-crafted models, (Yuille, 1991). It has been decided to use a Fourier parametrisation, similar to Staib and Duncan (1992). Fourier descriptors are initially used to parametrise the inner boundaries using a polar representation, with co-ordinates  $r, \theta$ . Fourier representations of the boundary express the curve as a weighted sum of a set of orthogonal known functions, i.e. cosines and sines. A Fourier parametrisation is suitable for these images because it is concise, periodic and is very general in that it is not restricted to a particular type of object, unlike the hand-crafted models used by Yuille et al. (1992), see Subsection 3.5.1. Also, the  $r, \theta$  descriptor has the characteristic that the number of terms used is the same regardless of the size of the object to be identified. Staib and Duncan (1992) use a basis function of

$$(1/(2\pi), \cos \theta/\pi, \sin \theta/\pi, \cos 2\theta/\pi, \sin 2\theta/\pi, \dots, ).$$

They also state that the number of harmonics,  $h$ , that should be chosen to parametrise the boundaries, is a trade-off between the accuracy and the conciseness of the fitted boundary (in comparison to the true boundary) and the degree of smoothing of the fitted boundary.

It was decided that it was only necessary to model the inner carcass boundary since the outer boundary for each image can be easily identified using low-level image analysis techniques, such as thresholding and then forming a labelled image of connected components. The  $x, y$  coordinates of all the pixels lying on the manually segmented inner boundary are extracted, centred with respect to a 'central pixel' in each image and converted to the usual polar representation. The central

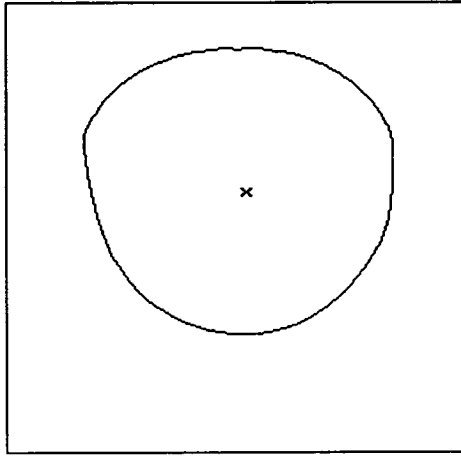


Figure 4.5: *An outer boundary region which is used to calculate the co-ordinates of the centroid. The position of the centroid for this image is marked by an 'x'.*

pixel chosen is the centroid (geometric centre) of each binary outer boundary. This is shown in Figure 4.5, where the centroid is marked by an 'x'. The centroid of an object has the useful property that it is approximately invariant with respect to the orientation that the object is scanned, if a discrete representation of the object is used. For a set of  $n_i$  pixels lying on the  $i$ th binary outer boundary, (as shown in Figure 4.5), at pixel positions  $(x_1, y_1), (x_2, y_2), \dots, (x_{n_i}, y_{n_i})$  in the image grid, the centroid of the  $i$ th image  $(\bar{x}_i, \bar{y}_i)$  is given by

$$\bar{x}_i = \frac{1}{n_i} \sum_{t=1}^{n_i} x_{t_i} \quad \bar{y}_i = \frac{1}{n_i} \sum_{t=1}^{n_i} y_{t_i}. \quad (4.1)$$

Examples of the position of the centroid are also shown in Figure 4.2 (a) and (b). Having extracted the centroid and the  $x, y$  coordinates of the pixels on the inner boundary, the corresponding radius  $r$  and angle  $\theta$  for each point is calculated, where  $\theta$  is calculated in the range  $(-\pi, +\pi]$  with zero taken parallel to the vertical axis.

For each of the  $N$  training images ( $i = 1, \dots, N$ ), we have an  $m_i$ -vector  $\mathbf{r}_i$  of distances from the centroid to the inner carcass boundary. Here,  $m_i$  is the number of pixels lying on the manually drawn inner boundary of image  $i$ , and the  $m_i$  from the training set vary from 330 to 450 pixels. This is due to the natural variability in sizes of the sheep at the same age and of the same breed. We treat  $\mathbf{r}_i$  as the vector of responses in a linear regression model

$$\mathbf{E}(\mathbf{r}_i | \mathbf{X}_i) = \mathbf{X}_i \boldsymbol{\beta}_i, \quad \text{Var}(\mathbf{r}_i | \mathbf{X}_i) = \sigma^2 \mathbf{I}_{m_i}. \quad (4.2)$$

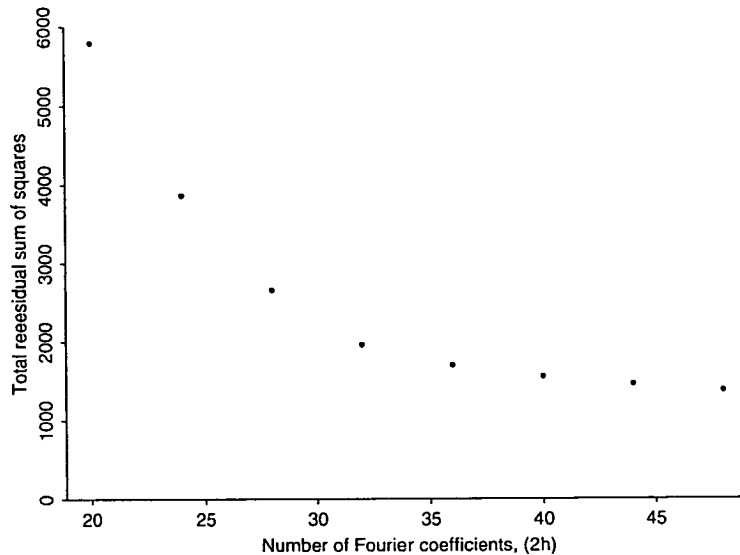


Figure 4.6: Total RSS from fitting the 24 training images vs number of Fourier coefficients,  $2h$ .

For the inner boundaries,  $\mathbf{X}_i$  is chosen to be an  $m_i \times p$  matrix whose  $j$ th row comprises  $1, \cos \theta_j, \cos 2\theta_j, \dots, \cos h\theta_j, \sin \theta_j, \sin 2\theta_j, \dots, \sin h\theta_j$ , and  $\beta_i$  is a  $p$ -vector of Fourier coefficients. The higher indexed basis functions, i.e.  $\cos h\theta$  and  $\sin h\theta$ , represent the higher spatial variation. Therefore,  $h$  Fourier harmonics are used, giving a total of  $p = (2h + 1)$  Fourier coefficients. Let  $\mathbf{b}_i$  denote the least-squares estimate of  $\beta_i$  and  $\mathbf{B}$  the  $N \times p$  matrix with rows  $\mathbf{b}_1^T, \dots, \mathbf{b}_N^T$ . Various numbers of harmonics are tried in order to parametrise each image, and the total residual sum of squares (RSS) over the 24 training images is plotted against the total number of harmonics,  $2h$ , which equals the number of Fourier coefficients (not counting the constant term): see Figure 4.6. On examination of this graph it is decided to parametrise each boundary using  $h = 20$  harmonics. Staib and Duncan (1992) choose the number of harmonics that reconstruct the boundary within a predetermined fixed error bound. For each of the 24 sheep images, 41 regression coefficients are estimated using least squares. The total variation can not be completely explained (i.e. the RSS does not fall to zero) using 25 harmonics (i.e.  $p = 51$ ). This is in part due to a relatively small number of harmonics in the model and to the fact that these inner boundaries are not all star-shaped with respect to the centroid of the outer boundary. The area which causes some boundaries not to be star-shaped is the region of the hand-drawn boundary around the backbone. It is possible to find rays from the centroid which intersected the boundary at more than one point.

Figure 4.7 shows the pixels (displayed using the radii and angles) of the manually



drawn inner boundaries (displayed as ‘•’) for images (a) and (b) shown in Figure 4.2. The continuous boundaries fitted to these data points, using the Fourier representation, have been superimposed onto the original data.

## 4.4 Reduced-rank approximation and principal components

To generalise from the 24 training images to future images without hand-drawn boundaries, we might suppose that the vector of coefficients,  $\beta_i$  in (4.2), are taken at random from a  $p$ -variate distribution, where  $p = 2h+1$ , and use the mean vector and variance matrix of the  $\mathbf{b}_i$  to estimate the moments of this distribution. With only 24 images, the sample variance is clearly a poor estimate of the population variance. To reduce its sampling variability, we use a reduced-rank approximation to the matrix of estimated coefficients. This approximation is closely related to the principal component analysis of the sample variance matrix of the  $\mathbf{b}_i$ .

Before the theory of the reduced rank approximation is presented, some definitions are given.

**Definition 4.1** *An  $m \times m$  matrix  $\mathbf{A}$  is idempotent if  $\mathbf{A}\mathbf{A} = \mathbf{A}^2 = \mathbf{A}$ .*

Consequences of this, (Harville, 1997) are

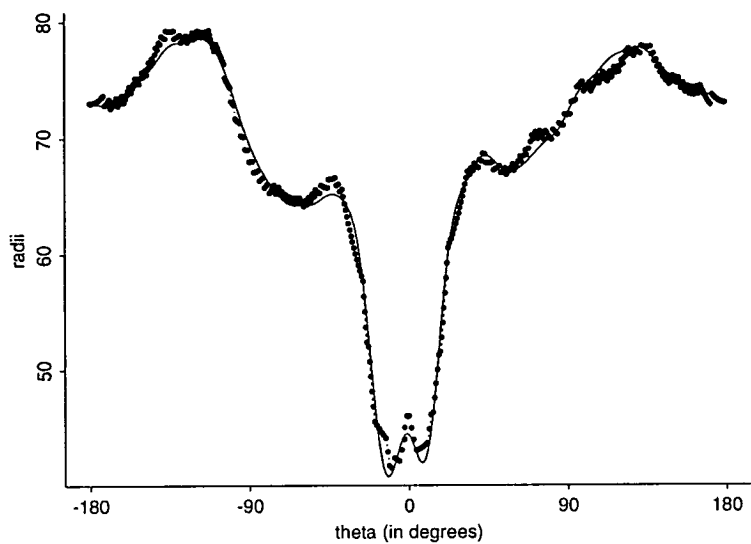
- If  $\mathbf{A}$  is symmetric and idempotent then the rank of  $\mathbf{A} = \text{trace of } \mathbf{A}$ .
- If  $\mathbf{A}$  is idempotent, then  $\mathbf{I}_m - \mathbf{A}$  is idempotent.

**Definition 4.2** *A Frobenius norm is the norm on matrices that arises by treating a matrix as a vector and using the Euclidean norm of that vector. If  $\mathbf{B}$  is an  $m \times n$  matrix, with elements  $b_{ij}$ , then*

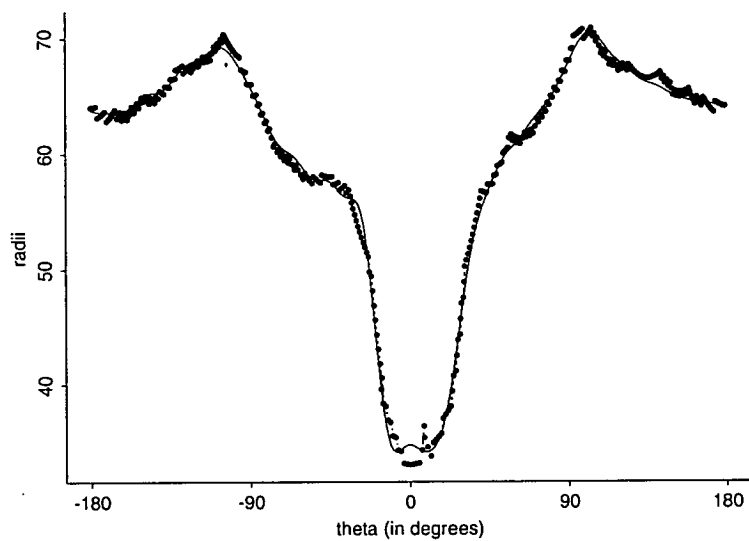
$$\|\mathbf{B}\|_F^2 = \sum_{i=1}^m \sum_{j=1}^n |b_{ij}|^2.$$

Let  $\mathbf{J}_N$  and  $\mathbf{H}_N$ , (a centering matrix), denote the symmetric, idempotent  $N \times N$  matrices  $N^{-1}\mathbf{1}_N\mathbf{1}_N^T$  and  $\mathbf{I}_N - N^{-1}\mathbf{1}_N\mathbf{1}_N^T$ , where  $\mathbf{1}_N$  denotes an  $N$ -vector of 1's, and write  $\mathbf{B}$ , the  $N \times p$  matrix, as

$$\begin{aligned} \mathbf{B} &= \mathbf{J}_N\mathbf{B} + \mathbf{H}_N\mathbf{B} \\ &= \mathbf{1}_N\bar{\mathbf{b}}^T + \mathbf{H}_N\mathbf{B} \end{aligned} \tag{4.3}$$



(a)



(b)

Figure 4.7: Radii vs angles for the original hand-drawn inner boundary pixels (shown as '•') for the X-ray images in Figure 4.2. The estimated boundaries for these images using  $p = 41$  Fourier coefficients have been superimposed as a continuous line.

where  $\bar{\mathbf{b}}$  is the mean of the  $\mathbf{b}_i$  and the rows of  $\mathbf{H}_N \mathbf{B}$  equal the deviations  $(\mathbf{b}_i - \bar{\mathbf{b}})^\top$  from the mean. Then  $r = \text{rank}(\mathbf{H}_N \mathbf{B}) \leq \min(N - 1, p)$ . To reduce the sampling variation in  $\mathbf{B}$ , we replace  $\mathbf{H}_N \mathbf{B}$  by an approximation with rank  $q < r$ , (see Basilevsky, 1983).

Let  $\mathbf{D}$  be a diagonal  $r \times r$  matrix whose diagonal elements are the  $r$  positive eigenvalues,  $d_1 \geq d_2 \geq \dots \geq d_r$ , of the symmetric, positive semi-definite matrix  $(\mathbf{H}_N \mathbf{B})^\top \mathbf{H}_N \mathbf{B} = \mathbf{B}^\top \mathbf{H}_N \mathbf{B}$ . This is the matrix of sums of squares and products of the  $\mathbf{b}_i$  about their mean, with entries in the  $j$ th row and  $k$ th column being

$$\sum_{i=1}^N (b_{ij} - \bar{b}_j)(b_{ik} - \bar{b}_k), \quad (1 \leq j, k \leq p).$$

If  $\mathbf{P}$  denotes the  $p \times r$  matrix of the corresponding eigenvectors then  $\mathbf{B}^\top \mathbf{H}_N \mathbf{B}$  has the decomposition  $\mathbf{P} \mathbf{D} \mathbf{P}^\top$  with  $\mathbf{P}^\top \mathbf{P} = \mathbf{I}_r$ .

The matrix  $\mathbf{H}_N \mathbf{B}$  has singular value decomposition

$$\mathbf{H}_N \mathbf{B} = \mathbf{Q} \mathbf{D}^{1/2} \mathbf{P}^\top, \quad (4.4)$$

where

$$\mathbf{Q} = \mathbf{H}_N \mathbf{B} \mathbf{P} \mathbf{D}^{-1/2} \quad (4.5)$$

and is an  $N \times r$  matrix. Then  $\mathbf{Q}$  has orthonormal columns  $\mathbf{q}_1, \dots, \mathbf{q}_r$ , since

$$\begin{aligned} \mathbf{Q}^\top \mathbf{Q} &= \mathbf{D}^{-1/2} \mathbf{P}^\top \mathbf{B}^\top \mathbf{H}_N \mathbf{B} \mathbf{P} \mathbf{D}^{-1/2} \\ &= \mathbf{D}^{-1/2} \mathbf{D} \mathbf{D}^{-1/2} \\ &= \mathbf{I}_r. \end{aligned}$$

For some specified integer  $q < r$ , let  $\mathbf{P}_1$  and  $\mathbf{Q}_1$  denote the submatrices of  $\mathbf{P}$  and  $\mathbf{Q}$  comprising their first  $q$  columns and let  $\mathbf{D}_1$  be diagonal  $(d_1, \dots, d_q)$ . Then, it is possible to approximate  $\mathbf{H}_N \mathbf{B}$  by

$$\begin{aligned} \mathbf{C}_q &= \mathbf{Q}_1 \mathbf{D}_1^{1/2} \mathbf{P}_1^\top \\ &= \mathbf{H}_N \mathbf{B} \mathbf{P}_1 \mathbf{P}_1^\top, \end{aligned} \quad (4.6)$$

where  $\mathbf{C}_q$  is the best approximation among  $N \times p$  matrices of rank  $q$  in the sense of minimising  $\|\mathbf{H}_N \mathbf{B} - \mathbf{C}_q\|_F$ , (see Stewart, 1973, page 322). The  $N \times q$  matrix  $\mathbf{H}_N \mathbf{B} \mathbf{P}_1$  comprises the principal component scores for the first  $q$  principal

components based on  $\mathbf{B}^T \mathbf{H}_N \mathbf{B}$ . This  $N \times q$  matrix has sums of squares and products matrix  $\mathbf{D}_1$ , since

$$\begin{aligned} (\mathbf{H}_N \mathbf{B} \mathbf{P}_1)^T \mathbf{H}_N \mathbf{B} \mathbf{P}_1 &= \mathbf{P}_1^T \mathbf{P} \mathbf{D} \mathbf{P}^T \mathbf{P}_1 \\ &= \begin{bmatrix} \mathbf{I}_q & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0}^T & \mathbf{D}_2 \end{bmatrix} \begin{bmatrix} \mathbf{I}_q \\ \mathbf{0}^T \end{bmatrix} \\ &= \mathbf{D}_1, \end{aligned}$$

where  $\mathbf{0}$  is  $q \times (r - q)$  and  $\mathbf{D}_2$  is  $(r - q) \times (r - q)$ .

Having found an approximation to  $\mathbf{H}_N \mathbf{B}$ , it can be seen from (4.3) that  $\mathbf{B}$  is approximated by

$$\mathbf{1}_N \bar{\mathbf{b}}^T + \mathbf{H}_N \mathbf{B} \mathbf{P}_1 \mathbf{P}_1^T. \quad (4.7)$$

The  $m_i$ -vector of fitted values for the  $i$ th regression is  $\mathbf{X}_i \mathbf{b}_i$ , and this is replaced by the approximation  $\mathbf{X}_i \hat{\mathbf{b}}_i$ , where  $\hat{\mathbf{b}}_i^T$  is the  $i$ th row of  $\mathbf{C}_q$ .

#### 4.4.1 Assessing results and principal scores from the reduced-rank approximation

The vectors  $\mathbf{b}_i$ , of estimates of the Fourier coefficients for each of the 24 training images are used to form the  $24 \times 41$  matrix  $\mathbf{B}$ . Performing the reduced rank approximation on  $\mathbf{H}_N \mathbf{B}$  reveals that the first six principal components accounted for 96.3% of the variation, as measured by the trace of  $(\mathbf{B}^T \mathbf{H}_N \mathbf{B})$ , between the original 41 variables. Figure 4.8 shows the cumulative proportions explained as the number of principal components is increased.

The first six principal components are assessed for multivariate normality. Plots of the principal component scores can reveal suspect observations in the data as well as provide checks on the assumption of normality. Since the principal components are linear combinations of the original  $p$  variables, it is not unreasonable to expect them to be normal. Figure 4.9 shows scatter diagrams for pairs of the first few principal component scores centred about their means. It is already known that the principal component scores are uncorrelated, but the scatter plots can be used to reveal any unusual observations. Figure 4.9 (a) has an outlier at  $(10.77, -0.64)$  which corresponds to image 8 and similarly Figure 4.9 (b) has an obvious outlier at  $(2.72, 0.74)$  which corresponds to image 20. These results could indicate that the fitted templates for images 8 and 20 would not fit as accurately to the hand-drawn template as for the other images in the training set. It can be seen from

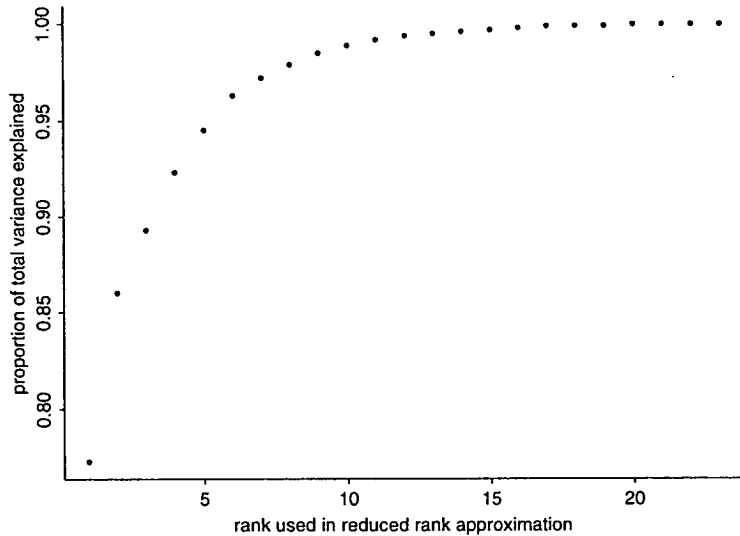


Figure 4.8: *Proportions of the total variation in the regression coefficients explained as the number of principal components varies.*

the probability plots Figure 4.10 that the first six principal component scores do not contradict this normality hypothesis.

The reduced-rank model represents the  $i$ th vector of regression coefficients,  $\hat{\mathbf{b}}_i$ , as  $\bar{\mathbf{b}} + \mathbf{P}_1 \mathbf{d}_i$  with  $\mathbf{d}_i^T$ , the  $i$ th row of  $\mathbf{H}_N \mathbf{B} \mathbf{P}_1$ , so that the  $\hat{\mathbf{b}}_i$  lie in a  $q$ -dimensional hyperplane defined by  $\bar{\mathbf{b}}$  and  $\mathbf{P}_1$ . For a future image from the same population, we take the vector of coefficients  $\mathbf{b}_f$  to equal  $\bar{\mathbf{b}} + \mathbf{P}_1 \mathbf{d}_f$ , where

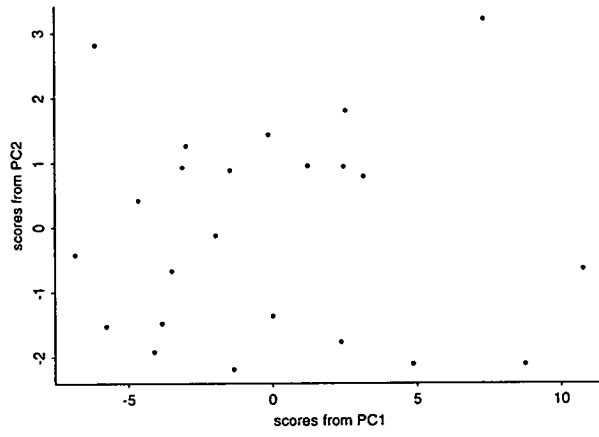
$$\mathbf{d}_f \sim N_q(\mathbf{0}, \mathbf{D}_1). \quad (4.8)$$

If a common set of  $m$  angles is used then the matrices  $\mathbf{X}_i$  reduce to a common  $m \times p$  matrix  $\mathbf{X}$  and the vector of radii is fitted by a model of the form

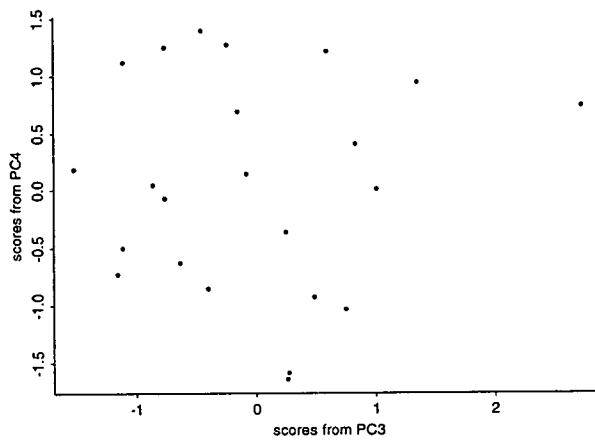
$$\hat{\mathbf{r}} = \mathbf{X} \bar{\mathbf{b}} + \mathbf{X} \mathbf{P}_1 \mathbf{d}_f. \quad (4.9)$$

## 4.5 Fitting boundaries to the training images using the reduced-rank approximation

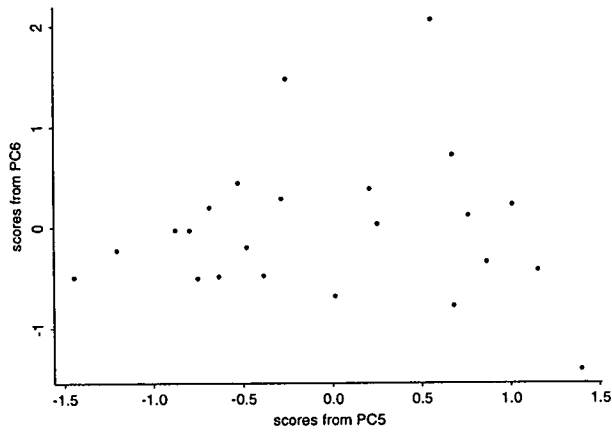
The images in the training set have varying numbers of pixels lying on the hand-drawn inner boundaries (330 to 450 approximately). Therefore, to standardise so that each image is reconstructed in the same way, a standard set of  $m = 360$  equally spaced angles is used to construct a common  $\mathbf{X}$ , and the corresponding radii estimated. Therefore, for each image,  $\mathbf{X}$  is now a  $m \times p$  matrix whose



(a)



(b)



(c)

Figure 4.9: Scatter plots of principal component scores for the 24 training images, (centred about their means). (a) first PC vs second (b) third vs fourth (c) fifth vs sixth.



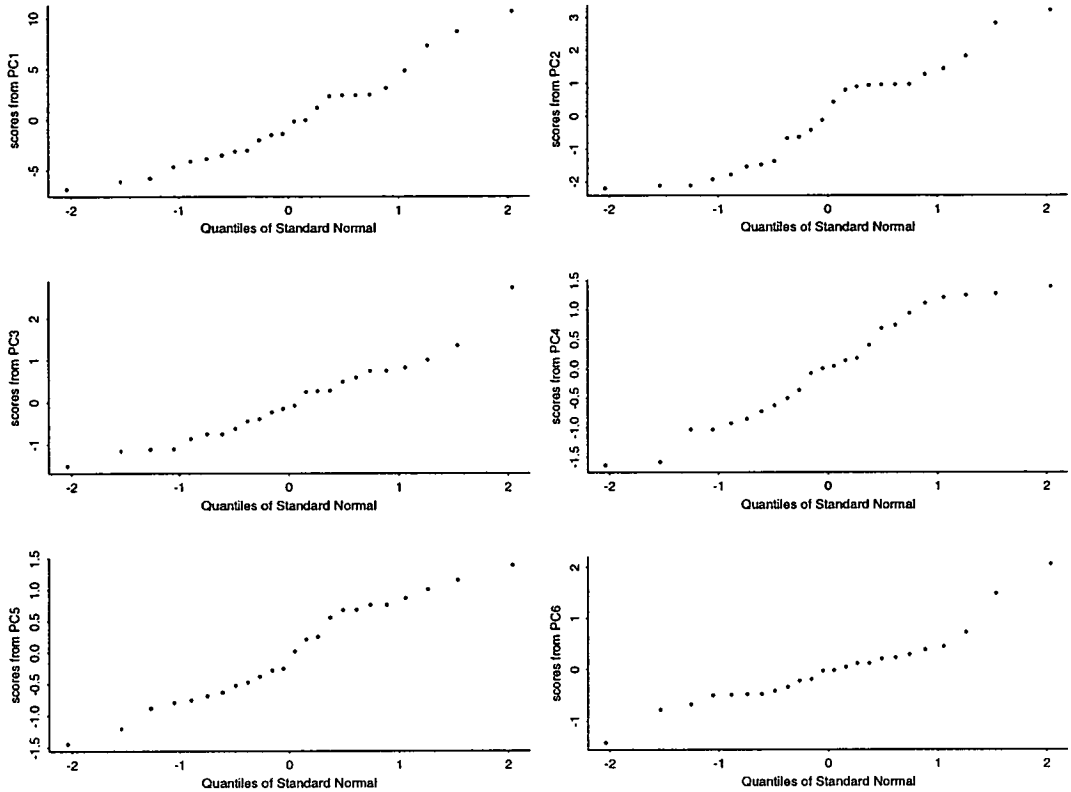


Figure 4.10: *Normal probability plots of the first 6 principal component scores, (centred about their means), for 24 training images.*

columns are  $1, \cos \theta, \dots, \cos h\theta, \sin \theta, \dots, \sin h\theta$ , where  $\theta$  is a 360-vector of equally spaced angles, making the columns of  $\mathbf{X}$  orthogonal. Taking  $q$  to be 6, the fitted coefficients in the appropriate row of  $\mathbf{C}_6$  for each image are used to reconstruct the inner boundaries for the 24 images one at a time.

Using standard Fourier theory gives

$$\mathbf{X}^T \mathbf{X} = m \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \frac{1}{2} \mathbf{I}_{2h} \end{bmatrix},$$

so that

$$\text{Var}(\mathbf{b}_i) = \frac{\sigma^2}{m} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & 2\mathbf{I}_{2h} \end{bmatrix}.$$

This shows that all the elements of  $\mathbf{b}_i$ , except the first have the same variance, (given  $\beta_i$ ), justifying calculating the principal components from the covariance matrix rather than the correlation matrix.

The effects on the template by varying the first few principal components was examined, although the results are not displayed here. It was found that the first principal component affects the orientation of the template with respect to the vertical axis. This indicates that more prior information from the hand-drawn

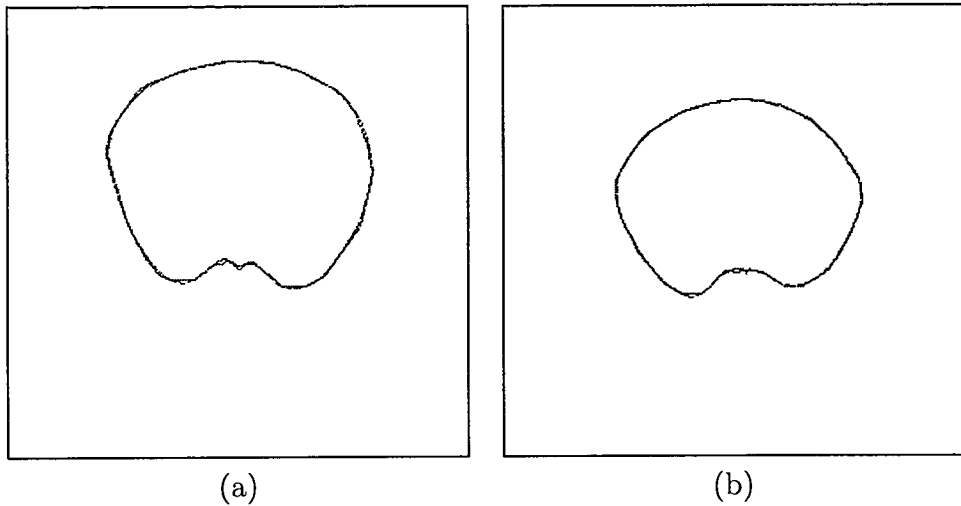


Figure 4.11: *Manually drawn (lighter) and fitted boundaries (darker) for the images in Figure 4.2 for a reduced-rank approximation with rank 6.*

boundaries could be included in the modelling process. This has been examined further in Section 4.9.

Figure 4.11 shows the fitted boundaries for a reduced rank approximation of rank 6, for the images in Figure 4.2, superimposed onto the hand-drawn boundaries. It can be seen that these fitted boundaries agree with the hand-drawn boundaries very accurately, and the main region of difference between the two boundaries is situated around the backbone. This is partly because when approximating the boundary initially by the 41 Fourier coefficients, a star-shaped boundary is fitted and also because the hand-drawn boundary is less smooth at this region in comparison to other parts, e.g. around the rumen.

For future images, the estimated probability distribution of the parameters of the template (which can be derived from (4.8)) can be combined with the matching criterion to form the optimisation function, which is then fitted to each image.

## 4.6 Image filtering

To extract information about the position of the boundaries, it has been decided to enhance the view of the inner boundaries in the images from the training set using an edge detector filter. Boundaries of objects tend to show up as intensity discontinuities in an image. The main objective of enhancement techniques is to transform an image so the result is more suitable than the original image for a specific application. There are two methods that may be used, see Gonzalez and Woods (1993).

**Spatial domain methods:** These methods refer to all the pixels forming an image, and the procedures operate directly on these pixels. It is these methods that have been used to transform the X-ray CT sheep images.

**Frequency domain methods:** These methods are based on transforming the spatial information in an image into frequency information by using the Fourier Transform. This uses a series of sine and cosine waves to fully represent the image.

### 4.6.1 Spatial filtering

As mentioned in Section 4.6, this is a pixel-by-pixel transformation of an image. Enhancement techniques based on this approach are referred to as filtering techniques, which are used to reduce noise by smoothing and/or enhance certain features of an image. Let the original greyscale pixel values of the  $256 \times 256$  sheep images be denoted by  $f_{u,v}$ , for  $u, v = 0, \dots, 255$ ; the row and column coordinates respectively. Similarly, let  $g_{u,v}$  be the transformed pixel values after filtering. The main approach is to use a submatrix of the area centred around each pixel, and this submatrix of pixels is generally square. The centre of the square is moved from pixel to pixel to obtain a new value  $g_{u,v}$  at each location by weighting the greyscale values of the pixel being processed,  $f_{u,v}$ , and also those of its neighbours. The values of the weights in this square determine the nature of the transformation and are often known as the ‘mask’ or the ‘filter’. If the weights of the original greyscale values of the pixels in this square are all positive then the filter will smooth the image. An example of this is the Gaussian filter, which will be discussed in Subsection 4.9.1.

Generally, the size of the filter used is a square of  $3 \times 3$  pixels, but it is possible to use  $5 \times 5$  or  $7 \times 7$  also, but the number of calculations per pixel increases rapidly as larger matrices are used. For pixels located on the border of the original image,  $g_{u,v}$  is computed using partial neighbourhoods. There are two types of spatial filters:

1. Linear filters: These use linear combinations of the greyscale values of the pixels in the original image, which depend on the weights of the filter. They are unable to smooth without simultaneously blurring the edges.
2. Non-linear filters: These can smooth without blurring edges and can detect edges at all orientations simultaneously. This type of filter does not use the

weights in the same way as linear filters, and is dependent on the type of filter how the weights are combined.

For linear filters, the basic approach is to sum products between the filter weights and the greyscale values of the pixels at a specific location of the submatrix in the image. Therefore, using a  $3 \times 3$  linear filter with specified weights  $w_{k,l}$ , for  $k, l = -1, 0, 1$  (with central pixel being  $w_{0,0}$ ) on the original greyscale values,  $f_{u,v}$ , will produce a transformed image with pixel values  $g_{u,v}$  where

$$g_{u,v} = \sum_{k=-1}^1 \sum_{l=-1}^1 w_{k,l} f_{u+k,v+l} \quad (4.10)$$

for  $u, v = 1, \dots, 254$  in the CT images. The transformed values for the border pixels are found using partial neighbourhoods.

Edge detection is by far the most common approach for detecting meaningful discontinuities in greyscale values. An edge is a boundary between two regions where the greyscale levels change rapidly. It is assumed that the regions in question are sufficiently homogeneous that the transition between the two regions can be determined on the basis of greyscale discontinuities alone. A region of rapid change is one where the first difference of the greyscale value, i.e. the gradient of greyscales, is at a maximum. If a directional derivative is used as a measure of edge strength, its response would vary with orientation of the edge. To avoid this, the magnitude of the gradient which gives the rate of change in the direction of greatest steepness is used. The first derivative at any point is obtained by using the magnitude of the gradient at that point. The gradient vector of the original image  $f$  at coordinates  $(x, y)$  is defined to be the vector  $\mathbf{f}$ , where

$$\mathbf{f} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix},$$

and the magnitude is given by

$$\text{mag} = \left[ \left( \frac{\partial f}{\partial x} \right)^2 + \left( \frac{\partial f}{\partial y} \right)^2 \right]^{1/2}. \quad (4.11)$$

One of the simplest non-linear edge detectors used for approximating the magnitude of the gradient, using (4.11), is Prewitt's filter: see Glasbey and Horgan (1995) for details. This uses a  $3 \times 3$  submatrix to approximate  $\frac{\partial f}{\partial x}$  and  $\frac{\partial f}{\partial y}$ . So, for a  $3 \times 3$  area of the image, with greyscale values  $f_{u+k,v+l}$ , the weights  $w_{k,l}$  for

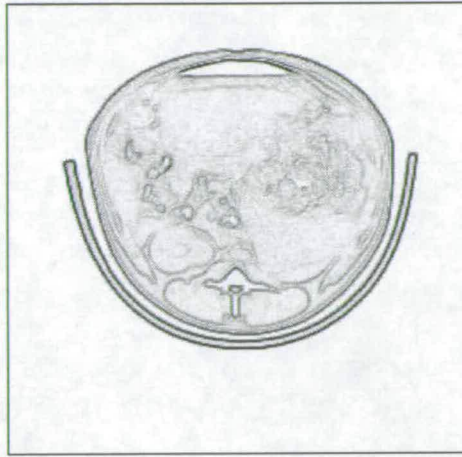


Figure 4.12: *Example of Prewitt's gradient filter applied to the original X-ray CT image shown in Figure 4.2 (a), with the 0 greyscale values shown in white and large greyscale values shown by dark pixels.*

$k, l = -1, 0, 1$ , used to approximate  $\frac{\partial f}{\partial x}$  are given by

$$\frac{1}{6} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

and the weights to approximate  $\frac{\partial f}{\partial y}$  are given by

$$\frac{1}{6} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}.$$

The filter coefficients sum to zero, indicating a response of zero in areas of constant greyscale values, as expected of a derivative operator. Although  $\frac{\partial f}{\partial x}$  and  $\frac{\partial f}{\partial y}$  use linear operations on the weights, the magnitude calculation does not, therefore the overall gradient filter is non linear. For further information on these filters see Ballard and Brown (1982) and Joyce-Loebl (1985).

Figure 4.12 shows the effect of applying Prewitt's gradient filter to an X-ray CT image. In the original images, black pixels have a greyscale value 0 and white pixels have a greyscale value 255. However, in Figure 4.12 the black pixels correspond to a greyscale value of 255 and the white pixels correspond to a greyscale value 0. This has been done for ease of viewing the defined edges from the gradient filter, and will be referred to as the 'inverse gradient filtered image' in future references. Larger pixel values (i.e. darker pixels) in Figure 4.12 correspond to a large change in the greyscale values around a point  $(x, y)$  in the image. Therefore, the outer boundary is seen as very dark because the filter detected a very strong edge as it passed from the inside to the outside of the sheep carcass. However, in

the middle of the image, most of the pixels are approximately the same greyscale value because the different internal organs all have approximately the same density in the original image. Therefore, as the filter passes from one organ to another it does not detect a large change in greyscale values. Hence, the filtered image has a smaller gradient at a pixel in this area than at a pixel lying on the outer boundary.

## 4.7 Matching criterion

As mentioned in Section 4.3, the deformable template model includes some measure of how well the template agrees with the true inner boundary in an image. Therefore, in order to form this matching criterion it is necessary to define the measure of fit between the template and the image. To do this, we examine the greyscale values from the gradient filtered image. It has been decided that averaging over the edge gradients of the pixels corresponding to the fitted boundary, defined by the template parameters, could be used as a measure of fit with the image. For images in the training set, the radii of the 360 pixels lying on the fitted boundary are found using  $\mathbf{X}\bar{\mathbf{b}} + \mathbf{X}\mathbf{P}_1\mathbf{d}_i$ , where  $\mathbf{d}_i$  is the  $i$ th row of  $\mathbf{C}_q$ . For future images, the 360 radii are expressed as  $\mathbf{X}\bar{\mathbf{b}} + \mathbf{X}\mathbf{P}_1\mathbf{d}_f$  where  $\mathbf{d}_f$  has distribution in (4.8). Obviously the stronger edges have a larger gradient value and hence a larger greyscale value in the gradient filtered images. Therefore, this suggests that the best fitting template maximises this average. Given the parameters of the template, the 360 radii and hence the corresponding Cartesian coordinates may be calculated. If these coordinates are integer valued, then using their corresponding pixel values in the gradient filtered image,  $g_{u,v}$ , we define the matching criterion to be

$$I(\mathbf{d}_f) = \exp\left(\frac{1}{360} \sum_u \sum_v g_{u,v}\right), \quad (4.12)$$

where the summation is over the given 360 greyscale values. It should be noted that  $I$  is a function of  $\mathbf{d}_f$  since the 360 gradient values are selected based on the template obtained given a particular instance of  $\mathbf{d}_f$ . However, these Cartesian coordinates are not integer valued. Therefore, bilinear interpolation, based on the four nearest pixels in the image is used to obtain a weighted sum of gradient values for each of the 360 radii. See Glasbey and Horgan (1995, page 42), for details on bilinear interpolation.

As stated in Section 4.3, the objective function is made up of two parts and can be thought of in either an energy minimisation framework or from a Bayesian



viewpoint. Therefore, similar to Glasbey (1998) and Phillips and Smith (1994), the estimated probability density of  $\mathbf{d}_f$ , (found from (4.8)) can be viewed as analogous to a prior density and the matching criterion as analogous to a likelihood. Then, summing the logs of these two components gives the objective function to be maximised, with a weighting  $c$  on the matching criterion. This is given by

$$F(\mathbf{d}_f) = \frac{1}{360} \sum_u \sum_v g_{u,v} - c \frac{1}{2} \mathbf{d}_f^T \mathbf{D}_1^{-1} \mathbf{d}_f. \quad (4.13)$$

The effect of varying  $c$  changes the emphasis on each of the two components of the objective function. These results are described later and are shown in Table 4.1. Many authors have similarly used a weighting between the two parts of the objective function, e.g. see Yuille et al. (1992), Phillips and Smith (1994) and Glasbey (1998). As in Glasbey (1998), it was decided to optimise this function  $F$  using the Nelder-Mead optimisation algorithm (Nelder and Mead, 1965). However, due to this being a minimisation technique,  $-F$  is used as the function to be optimised.

## 4.8 Nelder-Mead optimisation algorithm

Many numerical optimisation methods require the gradient of the objective function to be evaluated because it gives the direction of greatest increase in the function value. Therefore, the gradient provides the best local direction in which to move in the parameter space in order to locally maximise the objective function. However, an analytic or numerical form of the gradient must be found in order to make use of these methods, which in many cases can be very difficult. Methods which do not incorporate gradient information have an advantage of allowing greater flexibility in forming the objective function since they are not restricted by differentiability. Therefore, the Nelder-Mead algorithm is selected for the purposes of this optimisation as derivatives are not required to be calculated.

The Nelder-Mead algorithm, (see Press et al., 1996), is a downhill simplex method used for the minimisation of a function of  $M$  variables. No assumptions are made about the function except that it is continuous and has a unique minimum in the area of search. It is not very efficient in terms of the number of function evaluations that it requires and Powell's method, (see Press et al., 1996), is almost always faster in all applications. However, the downhill simplex method will converge even when the initial simplex straddles two or more valleys, a property not shared by Powell's method; the optimisation method used by Staib and Duncan (1992).

A simplex is a geometrical figure consisting, in  $M$  variables, of  $M + 1$  vertices, and all their interconnecting line segments. It is necessary to give the algorithm a starting guess in  $M$  dimensions with  $M + 1$  points, defined by an initial simplex. The method depends on the comparison of the function values at the  $M + 1$  vertices of the simplex, followed by the replacement of the vertex of the highest function value by another point. The simplex adapts itself to the local values of the function, and contracts on to the final minimum.

A general problem occurring in many deterministic minimisation methods is that of false convergence at a point other than the global minimum. Therefore, it is generally a good idea to restart a multi-dimensional routine at a simplex near where it claims to have found a minimum.

### 4.8.1 Results from optimising the objective function using the Nelder-Mead algorithm

The Nelder-Mead algorithm is used to minimise the objective function shown in (4.13) with  $c$  initially taken to be 1. This is carried out using 100 randomly chosen starting vectors (of length 6) from the estimated distribution on the template parameters, shown in (4.8). For each of the training images, the smallest of the minima (over the 100 starting positions) of the objective function is selected and the corresponding template is superimposed onto the gradient filtered image. Figure 4.13 shows examples of some of the different templates which minimised (4.16), using these 100 random starts. We find from this optimisation that in many cases the template giving the apparent minimum value of the objective function from each of the different starting vectors is actually recovering parts of the outer boundary. In some cases parts of the fitted template lie beyond the outer boundary: see Figure 4.13 (b).

It can be seen from Figure 4.13 that some of these templates provide a very poor solution to the segmentation, in particular the innermost boundary in Figure 4.13 (a) is not even close to the desired boundary.

From these 100 starts, 100 minimum values of  $-F$  are found and the smallest of these 100 values is taken to be the overall optimal best fitting template. Figure 4.14 shows these best fitting templates (lighter boundary) together with the original hand-drawn boundaries (darker boundary) for the two images from Figure 4.2. It may be seen that the manual template is very poorly recovered in both of these cases. The radii of the templates are over-estimated so that the area within the carcass region is vastly under-estimated in each image.

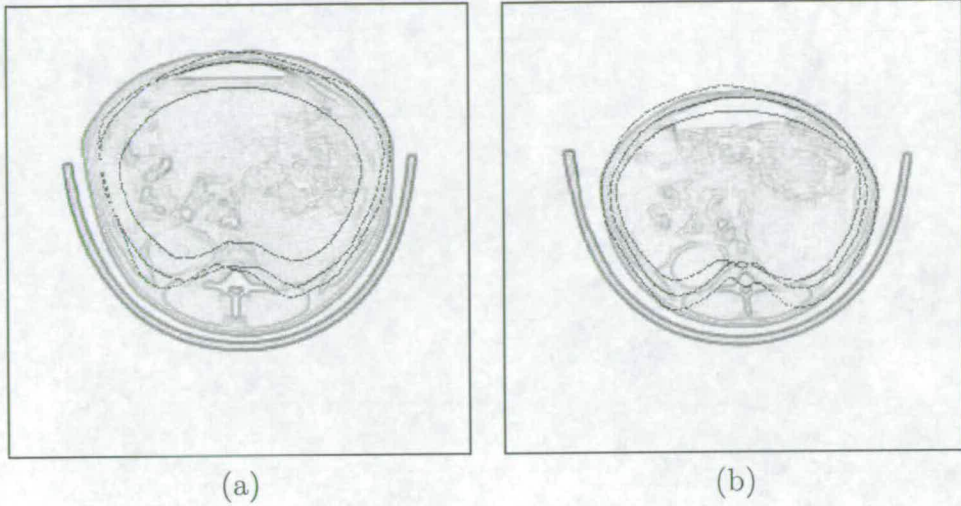


Figure 4.13: *Examples of the 'best' template from the 100 starting vectors, after optimisation using the Nelder-Mead algorithm. These templates have been superimposed onto the inverse gradient filtered X-ray images from Figure 4.2.*

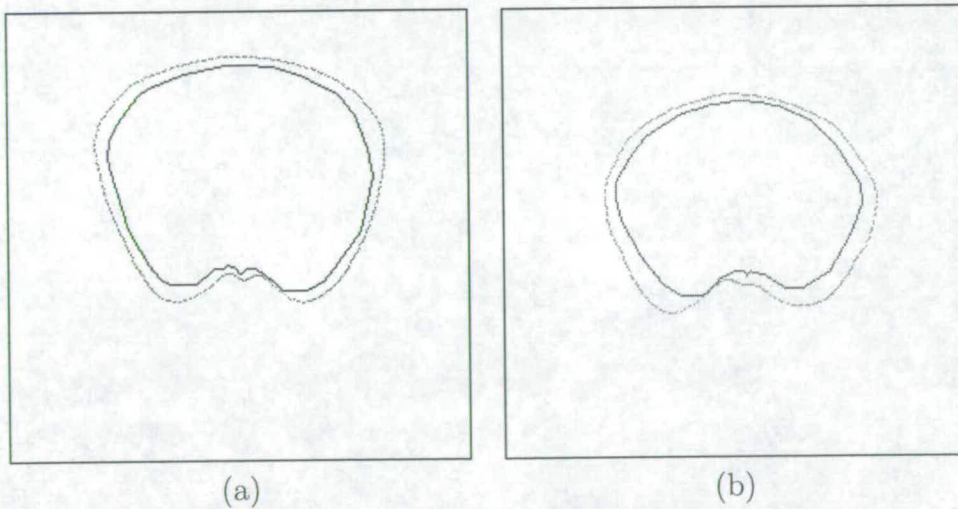


Figure 4.14: *Results after applying Nelder-Mead algorithm from 100 different starting positions. The darker boundary is the manual boundary and the lighter boundary is the best fitting template (obtained by the set of parameters which minimised the objective function). Again, these correspond to the original images in Figure 4.2.*

It can be seen from Figures 4.13 and 4.14 that the general shape of the inner boundary is correctly recovered using this parametrised template approach, but not the correct size. This suggests that the amount of prior knowledge used to form the template which reflects the size of the boundary needs to be increased.

## 4.9 Redefinition of the template

The results in Subsection 4.8.1 are clearly unsatisfactory since the model does not constrain the template to remain in the appropriate region of each image, i.e. inside the outer boundary of the sheep. Therefore, it has been decided that more prior information reflecting the size of the boundaries i.e. height and width should be incorporated into the model. The inner carcass boundaries will still be modelled using a polar representation, but now information from the outer boundary pixels will be included within the polar form. This is required so the model can be trained to ensure that the inner boundary lies inside the outer boundary. Therefore, we represent the radii of the inner boundary pixels as proportions of the distance to the outer boundary pixels situated at the same orientation. This is similar to Phillips and Smith (1994) who constrain the face template to lie inside the head template using proportions: see Subsection 3.5.2 for more details on this paper. Using this idea constrains the radii to lie in the range  $[0,1]$ : the closer to 1 implying the closer to the outer boundary that the template will lie. Incorporating this prior information will prevent the template from falling outside the outer boundary in the optimisation routine.

Previously the orientation of each sheep has not considered. For the redefinition of the templates this is included in the parametrisation. This depends on how much the sheep had rotated from the vertical axis after being restrained in the cradle. This was achieved by changing the reference axis of the angles from the vertical to the line joining the centroid of the outer boundary (marked with a white 'x' in the images in Figure 4.2) and the centre of the spinal column, which is easily located in all images either manually or automatically. This approach is similar to Cootes et al. (1995), who align their landmark points over the training set with respect to a set of axes: see Subsection 3.5.2.

The pixels lying on the inner boundary of each training image are then extracted and the corresponding  $r, \theta$  for each pixel found. This process is repeated for the outer boundary. From Section 4.3 it is known that the hand-drawn inner boundaries contain 330-450 pixels, whereas the corresponding hand-drawn outer boundaries contain more pixels than on the inner boundaries, approximately 380-

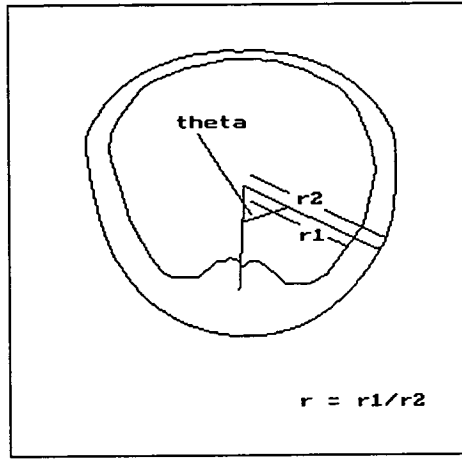


Figure 4.15: Angle,  $\theta$ , and proportional radius used for the redefined template where  $r_1$  is  $r_{\text{inner}}$  and  $r_2$  is  $r_{\text{outer}}$ .  $\theta$  is measured from the line joining the centroid and the centre of the spinal column.

490 pixels. This indicates that only a subset of the outer boundary pixels are necessary to calculate the appropriate proportional radii. Therefore, we choose the  $m_i$  pixels from each outer boundary ( $i = 1, \dots, N$ ) whose corresponding angles have the smallest absolute difference to the inner boundary angles. Having selected the  $m_i$  pixels, the corresponding outer boundary radii are found in order to calculate the  $m_i$ -vector of proportional radii ( $i = 1, \dots, N$ ). Figure 4.15 shows how the radius  $r$  for the corresponding angle on the inner boundary is calculated, where  $r = \frac{r_{\text{inner}}}{r_{\text{outer}}}$ .

When the manual outer boundary is viewed, it appeared to be very ‘jagged’ because it followed along the edges of the pixels. Therefore, it is decided to smooth the outer boundary by converting to polar coordinates (with the redefined axis) and using Fourier coefficients, as in Section 4.3. We decide to be consistent with Section 4.3 and parametrise the outer boundary using  $h_o = 20$  harmonics.

It is this new smoothed outer boundary that is used to extract the  $m_i$  radii to express the inner proportional radii as

$$r = \frac{r_{\text{inner}}}{r_{\text{fitted outer}}}. \quad (4.14)$$

Then, using these new  $m_i$  proportional radii, the regression analysis is repeated to estimate a new set of Fourier coefficients which models the inner boundary of each image. As in Section 4.3, we vary the number of Fourier harmonics,  $h$  in the parametrisation of the inner boundary. Figure 4.16 shows the total residual sum of squares (RSS) over the 24 images plotted against the total number of Fourier harmonics. From this, it has been decided to parametrise each boundary using 41 Fourier coefficients ( $2h + 1$ ) as before.

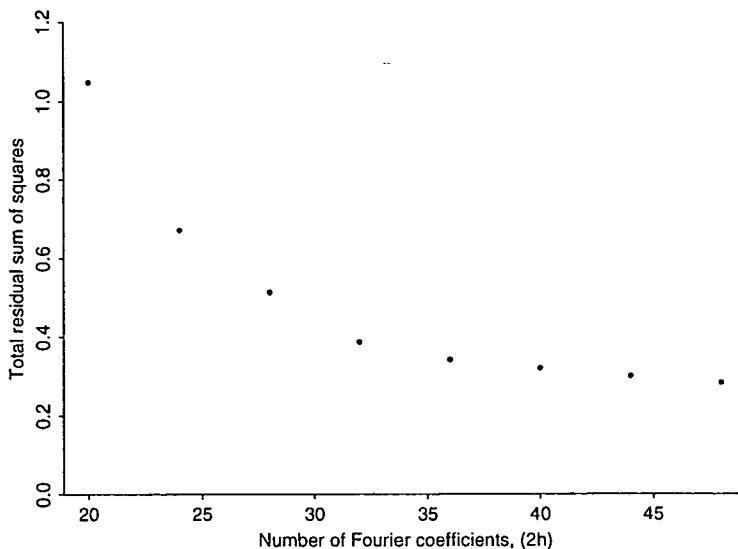
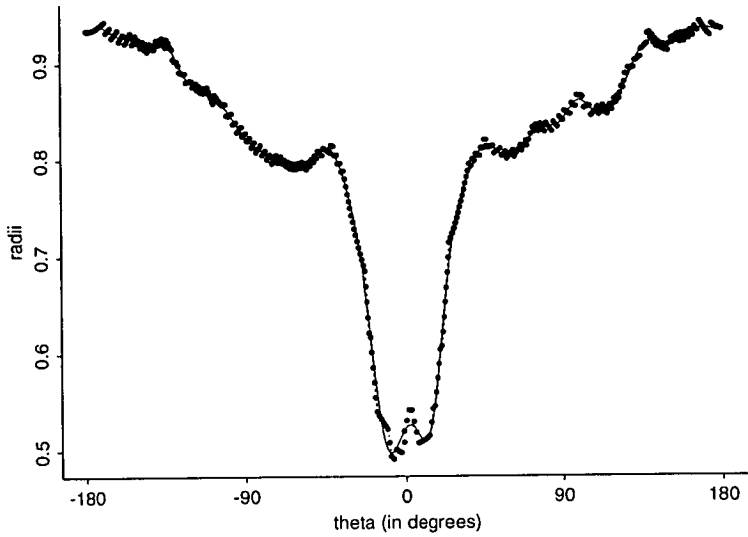


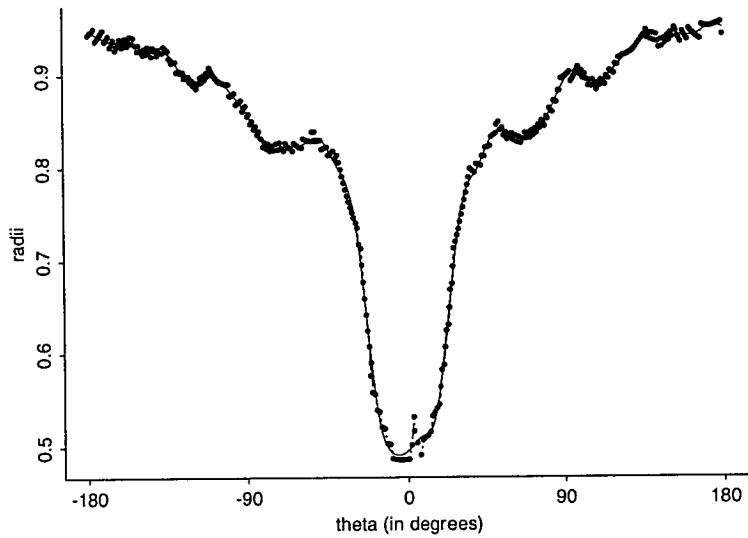
Figure 4.16: *Total RSS from fitting the proportional radii in the 24 training images vs total number of Fourier coefficients.*

Figure 4.17 shows the new fitted Fourier boundaries, again superimposed as a continuous line onto the original data points for the same images in Figure 4.2. It can be seen here that these images exhibit more symmetry than the corresponding images for the first parametrisation, shown in Figure 4.7. This is due to the redefinition of the template which helps to correct the asymmetry of the animal due to gravity when it is not upright in the cradle. The large ‘dip’ in the radii, centred around  $\theta = 0$ , corresponds to the pixels around the backbone region. This evidence of symmetry was one of the initial ideas that we mentioned in Section 4.2 when considering where to position the templates, since it is valuable to manual interpretation.

The reduced rank analysis was repeated on the matrix  $\mathbf{H}_N \mathbf{B}$  and Figure 4.19 shows scatter plots of the first few pairs of scores. In Figure 4.19 (a) the two outliers at  $(-0.036, -0.017)$  and  $(-0.017, 0.034)$  correspond to image 8 and image 20 respectively. In Figure 4.19 (b) the outlier at  $(-0.026, -0.005)$  corresponds to image 4. As before, this could indicate that the fitted templates for these images would not fit as accurately to the hand-drawn boundaries as for the remaining images. See Table 4.3 for further evidence of this. It appears the outliers in Figures 4.19 are more apparent than those in Figures 4.9. The results of this reduced rank approximation reveal that the first six components account for 86.7% of the variation between the original 41 variables and again the first six principal component scores support the normality hypothesis. Figure 4.18 shows the proportion of the cumulative variation explained as the number of principal components varies.



(a)



(b)

Figure 4.17: *Proportional radii vs angles for the original hand-drawn inner boundary pixels for the X-ray images in Figure 4.2. The estimated proportional radii for the images found using 41 Fourier coefficients have been superimposed as continuous lines.*



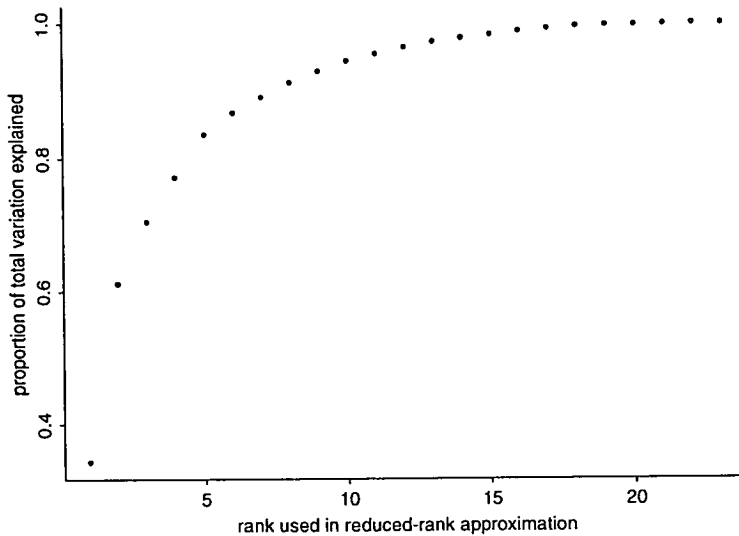
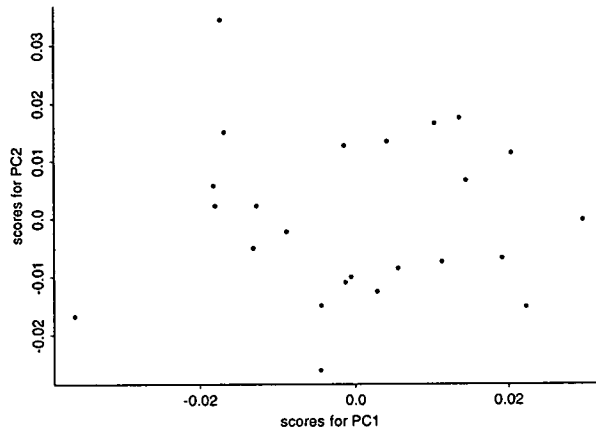


Figure 4.18: *Proportions of total variation in the regression coefficients explained as the number of principal components varies for the redefined template.*

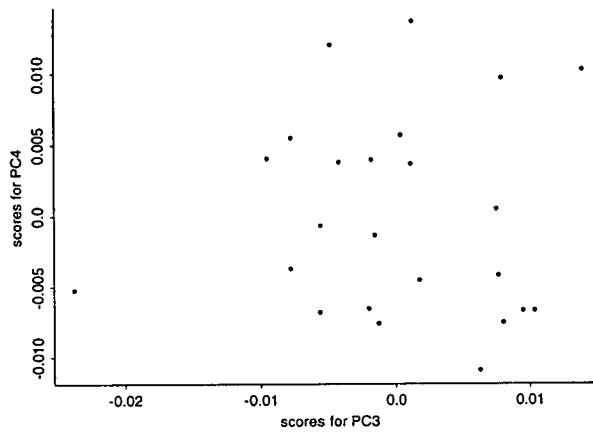
Figure 4.20 shows the effect on the template of varying each of the first three principal components. The darker lines show the average template increased by two standard deviations for the first three principal components in turn, and the lighter lines are for the average template decreased by two standard deviations. If the average template (i.e. no principal components used) had been added to these images, it would have been situated between the two boundaries already displayed. As can be seen, the first component mainly affects the amount of fat below the kidney that is included in the carcass: (refer to Figure 4.1). The second component specifies which layer of muscle to follow around the abdominal wall, while the third component affects the asymmetry of the image.

### 4.9.1 Smoothing the image before optimisation

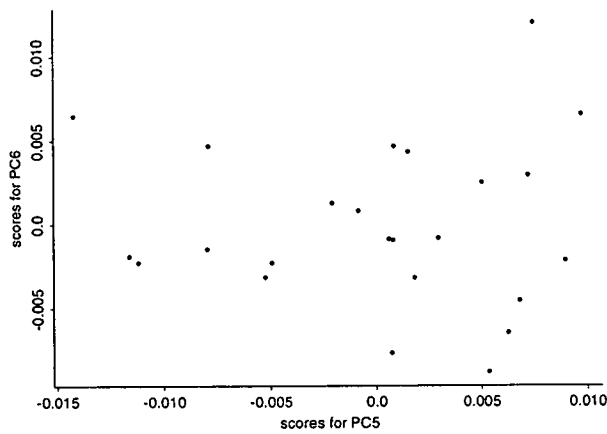
Before optimisation is carried out, the gradient filtered image is also modified slightly. Firstly, we decide to ‘blank out’ the outer boundary because of the problem discovered earlier (Subsection 4.8.1) that the optimisation routine found an optimum lying along parts of the outer boundary. This is easily achieved by setting all the pixel values on this boundary to zero, so they appear as white in the gradient filtered image shown in Figure 4.12. Next, a Gaussian filter, with a variance of  $\frac{8}{3}$ , was applied to simultaneously smooth and interpolate between pixels. The choice of variance used with the Gaussian filter is important, because as Staib and Duncan (1992) point out, it is possible to use too much smoothing



(a)



(b)



(c)

Figure 4.19: Scatter plots of principal component scores (centred about their mean) for the 24 training images (a) first PC vs second (b) third vs fourth (c) fifth vs sixth.

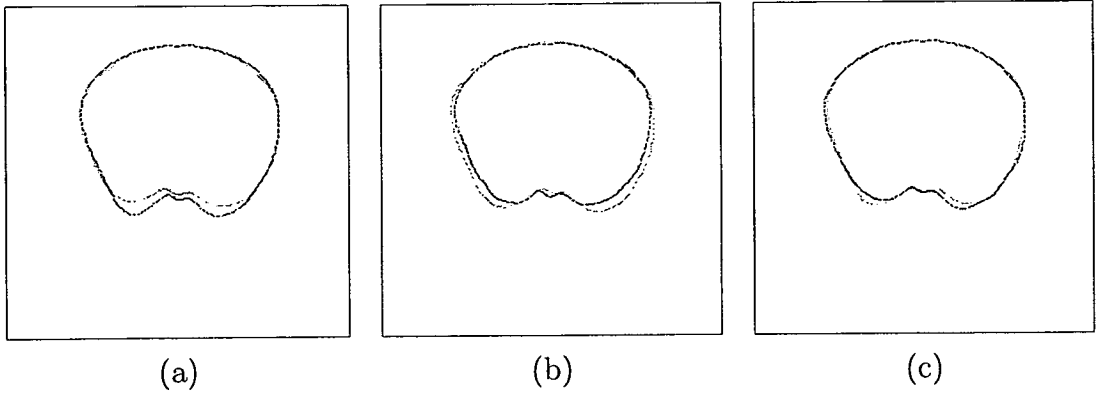


Figure 4.20: Average template with principal component scores increased (darker line) and decreased (lighter line) by two standard deviations for (a) first (b) second (c) third principal component.

(i.e. too large a variance) and this has the effect of going further than simplifying individual objects, so that they become no longer recognisable. Figure 4.21 shows the effect of applying this Gaussian filter to the gradient filtered image, with the outer boundary removed. It can be seen that this filter smoothed the image and caused a blurring effect, i.e. reduced the number of local optima in the objective function in (4.13). Many authors have used a similar approach for smoothing the image, e.g. see Lipson et al. (1990) and Staib and Duncan (1992). Table 4.2 shows evidence that these modifications produce more accurate positioning of the template in the image after optimisation when the proportional radii are fitted. The results in this table will be discussed later after the ‘goodness of fit criterion’ has been introduced, i.e. a measure quantifying how well the fitted template matches the manually drawn boundary.

The Gaussian filter is a linear filter, in which a weighted average of the neighbourhood pixels for each central pixel is formed, refer to Section 4.6. Gaussian filters have weights specified by the probability density function of a bivariate Normal distribution with variance  $\sigma^2$ . Therefore, the coefficients are defined by

$$w_{kl} = \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{(k^2 + l^2)}{2\sigma^2} \right\} \quad \text{for } k, l = -[3\sigma], -[2\sigma], \dots, +[3\sigma], \quad (4.15)$$

where  $[3\sigma]$  represents the integer part of  $3\sigma$ . The divisor of  $2\pi\sigma^2$  ensures the coefficients sum approximately to unity, which is a common convention with smoothing filters. For further details on this filter see Glasbey and Horgan (1995).

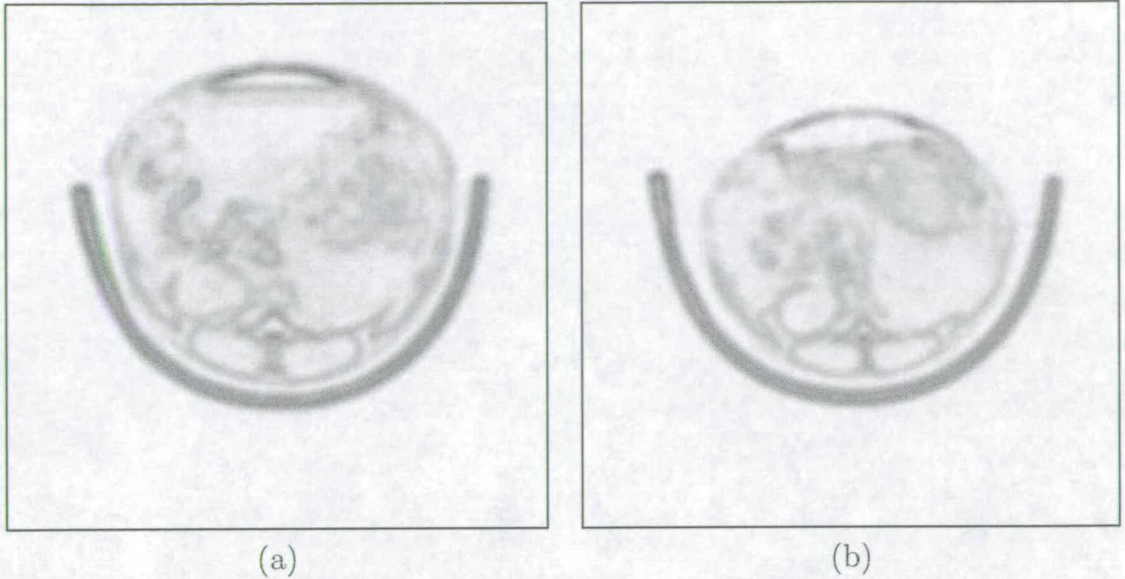


Figure 4.21: *Examples of the initial images in Figure 4.2 after Prewitt's filter has been applied, the outer boundaries removed and finally the Gaussian filter applied. As before, the inverse of these images are shown for ease of viewing.*

#### 4.9.2 Results after redefining the template

The objective function in (4.13) was minimised for a range of values of  $c$ , using 20 starting points for each value. It was found that this number was adequate as the variation between the best and worst fitted boundaries (i.e. corresponding to the parameter values which produced the smallest and largest value of the objective function using the Nelder-Mead routine) was quite small. This is due to incorporating more information from the manual boundaries into the template parametrisation and because the templates are standardised in size and orientation. Also, if the backbone is not vertically below the centroid, the inner and outer boundaries are distorted by gravity in a similar way. As before, the boundaries are reconstructed using only 360 radii, which are at equally spaced angles from the reference axis. The average sum of squares (SS) of the differences between the manual proportional radii and the fitted proportional radii (for the 24 images in the training set) for each value of  $c$  were found. These SS were found using

$$SS = \sum_{t=1}^{360} (r_t - \hat{r}_t)^2 \quad \text{where} \quad r = \frac{r_{\text{inner}}}{r_{\text{outer}}}. \quad (4.16)$$

These average sums of squares are shown in Table 4.1 for varying values of  $c$  and varying ranks. It can be seen that  $c = 1$  produces the lowest average sum of squares for each rank. Therefore, when validating this model on new data,  $c$  will be fixed at 1. This criterion in (4.16) is used to calculate the results in Table 4.2.

$c$	Rank used in reduced-rank approximation							
	0	1	2	3	4	5	6	7
9.0	0.154	0.148	0.145	0.145	0.145	0.145	0.145	0.144
3.0	0.154	0.139	0.132	0.133	0.133	0.132	0.132	0.132
1.0	0.154	0.126	<b>0.124</b>	0.132	0.130	0.128	0.128	0.128
0.3	0.154	0.130	0.166	0.227	0.219	0.222	0.199	0.275

Table 4.1: Average SS produced for various values of  $c$  in the objective function, (4.13), for images which have had both the gradient and Gaussian filters applied, and the outer boundaries ‘removed’ prior to optimisation.

Image type	Rank used in reduced-rank approximation							
	0	1	2	3	4	5	6	
A	0.154	0.180	0.216	0.199	0.170	0.220	0.215	
B	0.154	0.128	0.136	0.141	0.141	0.138	0.139	
C	0.154	0.126	0.124	0.132	0.130	0.128	0.128	

Table 4.2: Average SS over the 24 training images for various ranks used in the reduced-rank approximation and for different filters applied. Image types are A: Gradient filter, B: Gradient and Gaussian filters, C: Gradient and Gaussian filters with the outer boundaries removed. These values were found using  $c = 1$  in (4.13) and then using (4.16).

It can be seen from this table that by smoothing the edge image with a Gaussian filter (middle row of results) there is a significant reduction in the average SS. It is also evident that the removal of the outer boundary improves the results even further. In fact, it can be seen that using the gradient filter on its own produced worse fitting templates than if the average template had been used instead (i.e. zero principal components.)

The agreement between the best fitting template (darker), and the hand drawn boundaries (lighter), found using  $c = 1$ , for the images in Figure 4.2, may be seen in Figure 4.22. These results may be compared to those in Figure 4.14, and it is very evident that the manual template has been very accurately reconstructed. These best-fitting templates have been superimposed onto the inverse gradient filtered images (for ease of viewing) in Figure 4.23 to show the segmentation.

Although the previous choice uses rank 6 in the reduced-rank approximation, it can be seen from Table 4.1 that for  $c = 1$ , this is not the optimal rank. Using a reduced rank approximation with rank 2 gives a smaller average SS than rank 6 (and all other ranks), based on the total amount of variation from the true boundary for the training images. The first two principal components explain 34.4% and 61.1% of the total variability respectively.



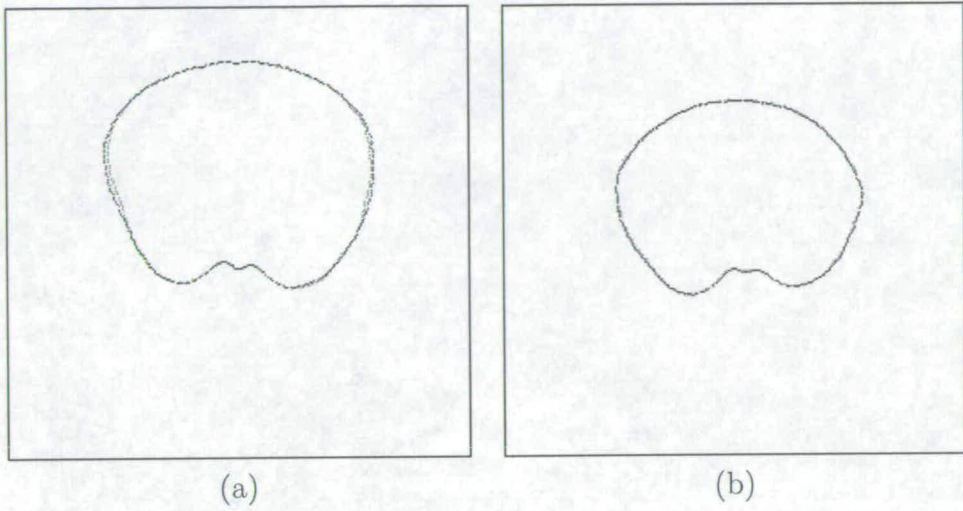


Figure 4.22: *Hand drawn boundaries (lighter) with the best fitting templates superimposed (darker), for the images in Figure 4.2.*

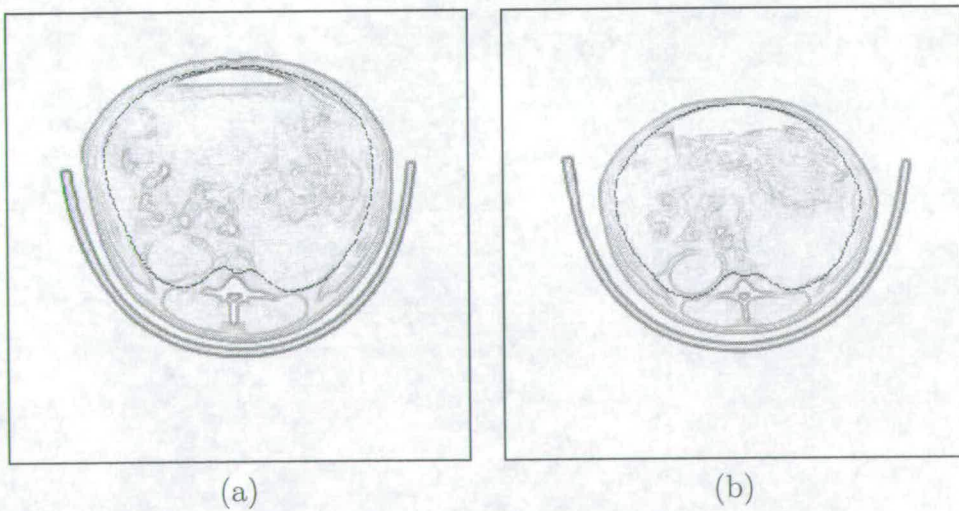


Figure 4.23: *X-ray images from Figure 4.2, with the template boundary superimposed onto the inverse gradient filtered image, based on optimisation of the redefined parametrisation of the template.*

Image	Rank used for reduced-rank approximation							
	0	1	2	3	4	5	6	7
1	0.117	0.092	0.071	0.110	0.108	0.092	0.093	0.094
2	0.225	0.095	0.103	0.110	0.106	0.108	0.107	0.106
3	0.065	0.060	0.064	0.059	0.065	0.068	0.070	0.069
4	0.230	0.242	0.088	0.084	0.076	0.077	0.073	0.075
5	0.068	0.068	0.041	0.045	0.043	0.043	0.045	0.045
6	0.185	0.206	0.103	0.118	0.119	0.114	0.118	0.117
7	0.093	0.076	0.068	0.100	0.099	0.097	0.095	0.096
8	0.314	0.236	0.179	0.170	0.166	0.166	0.160	0.161
9	0.235	0.219	0.274	0.288	0.289	0.278	0.272	0.276
10	0.138	0.078	0.083	0.097	0.096	0.097	0.092	0.090
11	0.240	0.130	0.125	0.143	0.137	0.147	0.142	0.143
12	0.095	0.095	0.081	0.074	0.056	0.055	0.055	0.055
13	0.074	0.091	0.081	0.074	0.074	0.072	0.074	0.074
14	0.064	0.062	0.036	0.047	0.047	0.045	0.044	0.042
15	0.190	0.152	0.152	0.169	0.179	0.181	0.178	0.182
16	0.122	0.071	0.084	0.062	0.063	0.063	0.064	0.066
17	0.168	0.095	0.121	0.114	0.116	0.109	0.112	0.112
18	0.104	0.080	0.068	0.077	0.070	0.069	0.068	0.073
19	0.091	0.091	0.095	0.091	0.094	0.092	0.091	0.092
20	0.442	0.405	0.554	0.603	0.604	0.600	0.617	0.619
21	0.092	0.097	0.054	0.063	0.056	0.058	0.055	0.055
22	0.114	0.117	0.133	0.126	0.129	0.116	0.118	0.120
23	0.078	0.079	0.124	0.135	0.134	0.127	0.128	0.124
24	0.159	0.095	0.217	0.206	0.197	0.195	0.197	0.194
Average	0.154	0.126	0.124	0.132	0.130	0.128	0.128	0.128

Table 4.3: Sums of squares of difference between the hand-drawn inner boundary and the fitted boundary (described in (4.16)) and the average sum of squares for the 24 training images, for the various ranks used to produce the fitted template using image type C.

Although the fitted boundaries are constructed with only 360 pixels, the majority of the original hand-drawn boundaries have many more pixels. Therefore, the 360 pixels that we extract from the hand-drawn boundary are those that lie closest in orientation to those on the fitted boundary. Table 4.3 displays the SS for each individual image and also the average SS over the full training set, for various ranks used in the fitting of the boundaries. It is also evident from Table 4.3 that image 20 is the worst, in terms of the criterion used in (4.16). Image 20 and image 8 were expected to do worse than the other images as they were viewed as outliers in Figure 4.19 (a): see Section 4.3 for previous comments. Figure 4.24 (a) displays the hand-drawn boundary, for image 20, with the poorly fitted boundary superimposed. Figure 4.24 (b) shows the original cross-section of image 20 with



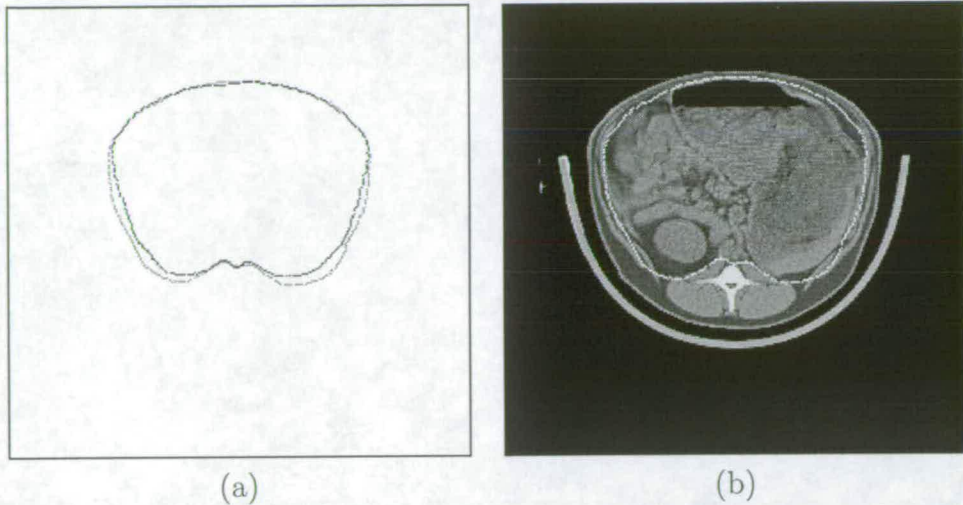


Figure 4.24: (a) Manual boundary (darker) with the fitted template (lighter) superimposed for image 20 (b) fitted template for image 20 superimposed onto the original image to show the poor segmentation.

the fitted boundary superimposed. This shows more clearly the areas of fat and muscle that have been incorrectly approximated. It can be seen from Figure 4.24 (b) that this is a fairly fat sheep and as a result of this, it ‘appears’ to the viewer that the continuous layers of muscle around the abdominal wall do not exist. The ‘islands’ of muscle within the carcass region make it difficult for the optimisation routine to find an accurate fitted template.

## 4.10 Applying the redefined model to validation data

Until now the algorithm had been built and tested only on the training data. However, it is necessary to validate the algorithm on an independent set of images. The new set of data consisted of 10 images, tomographs taken at the same lumbar vertebrae position and for the same breed and age of sheep as in the training data. The main problem encountered here is that the outer boundary has to be identified for the new images, whereas in the training set the pixels lying on the outer boundary are known and therefore can be easily modelled using the parametrisation. This is a relatively easy task using low level image analysis techniques involving several steps:

1. Threshold the image at a value of 0. This creates a binary image separating the background pixels (greyscale value 0) from the non-zero image pixels.
2. Using this binary image, a labelled image may be formed. This is where

the connected components in an image are labelled from 1 up to the total number of connected components in an image.

3. After labelling, it is fairly easy to extract the pixels which lie on the connected component corresponding to the outer boundary.

After the outer boundary is identified, the coordinates of the centroid are found using (4.1) and the outer boundary is smoothed using the Fourier parametrisation of Section 4.3. The coordinates of the centre of the spinal column are easily identified in all images.

The same procedure as in Section 4.4 was repeated, using proportional radii (measured from the line joining the centroid and backbone). Prewitt's and the Gaussian filters are applied and the outer boundary removed and the Nelder-Mead algorithm used to fit the boundary with  $d_f$  from (4.9) having length 2. Each boundary is reconstructed using 360 equally spaced angles to estimate the radii and bilinear interpolation of the corresponding real valued  $(x, y)$  coordinates is used to obtain 360 smoothed gradient values.

As would be expected, the fit of the template is not quite as good as it is for the original training data. However, the results are fairly accurate in comparison with those of the manual interpretation of the new data. The fit for the validation images is quantified using the sums of squares criterion in (4.16). The SS of differences of the true and fitted proportional radii for each of the 10 new images and their average SS are displayed in Table 4.4. An average SS of differences (rank 2) of 0.180 was found for the validation set, in comparison to 0.124 (rank 2) for the training data. Figure 4.25 shows the manual interpretation (lighter boundaries) of these new images with the fitted template assuming rank 2 (darker boundaries) superimposed. Figure 4.26 shows the 10 new images with the fitted template superimposed, again on the inverse of the gradient filtered image for ease of viewing.

## 4.11 Effect of adding mirror images to the training set

It has been decided to investigate how the results from the reduced-rank approximation would be affected if the mirror image of each training image is also used. This idea is motivated by the near-symmetry of the images and also because in the manual interpretation, symmetry is used to draw parts of the boundary if there is no clear boundary definition on one side of the image: see Section 4.2.

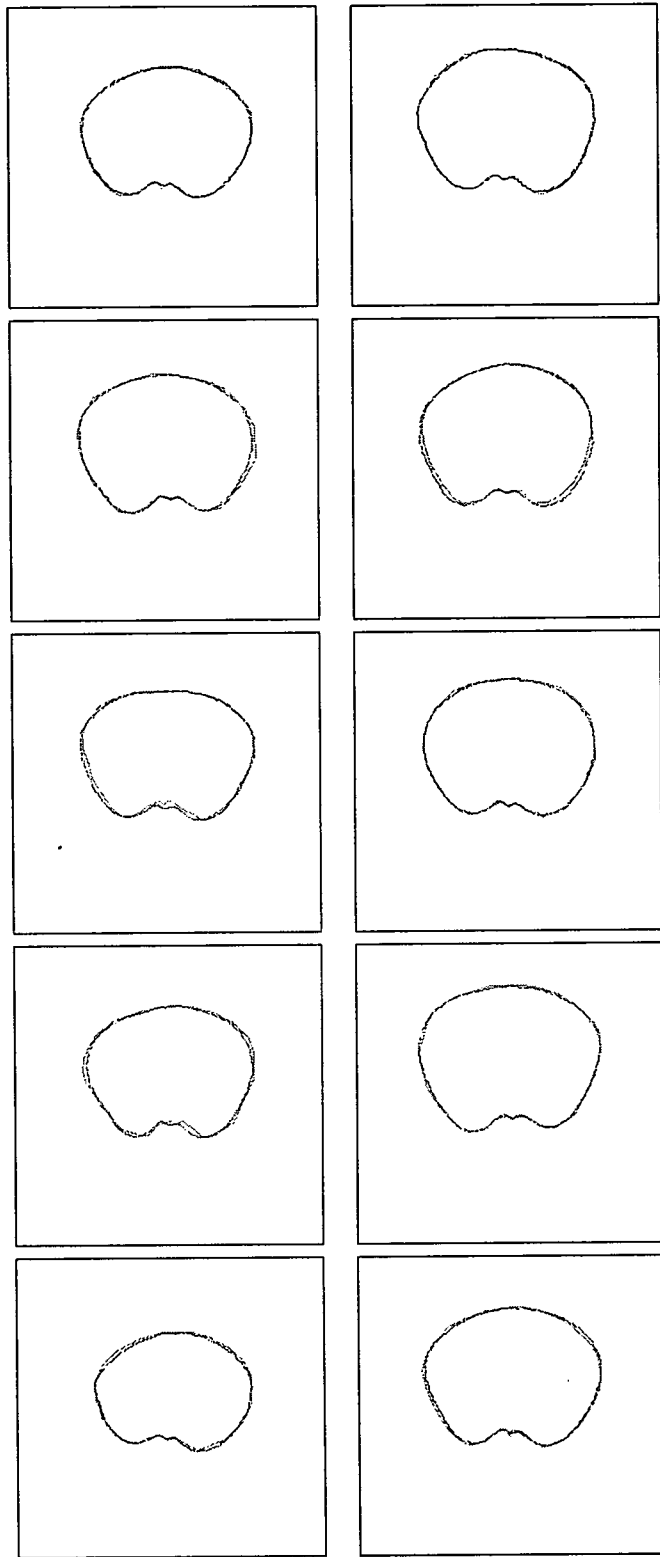


Figure 4.25: *The manual boundaries (lighter) for the validation data set with the fitted boundaries (darker) superimposed.*

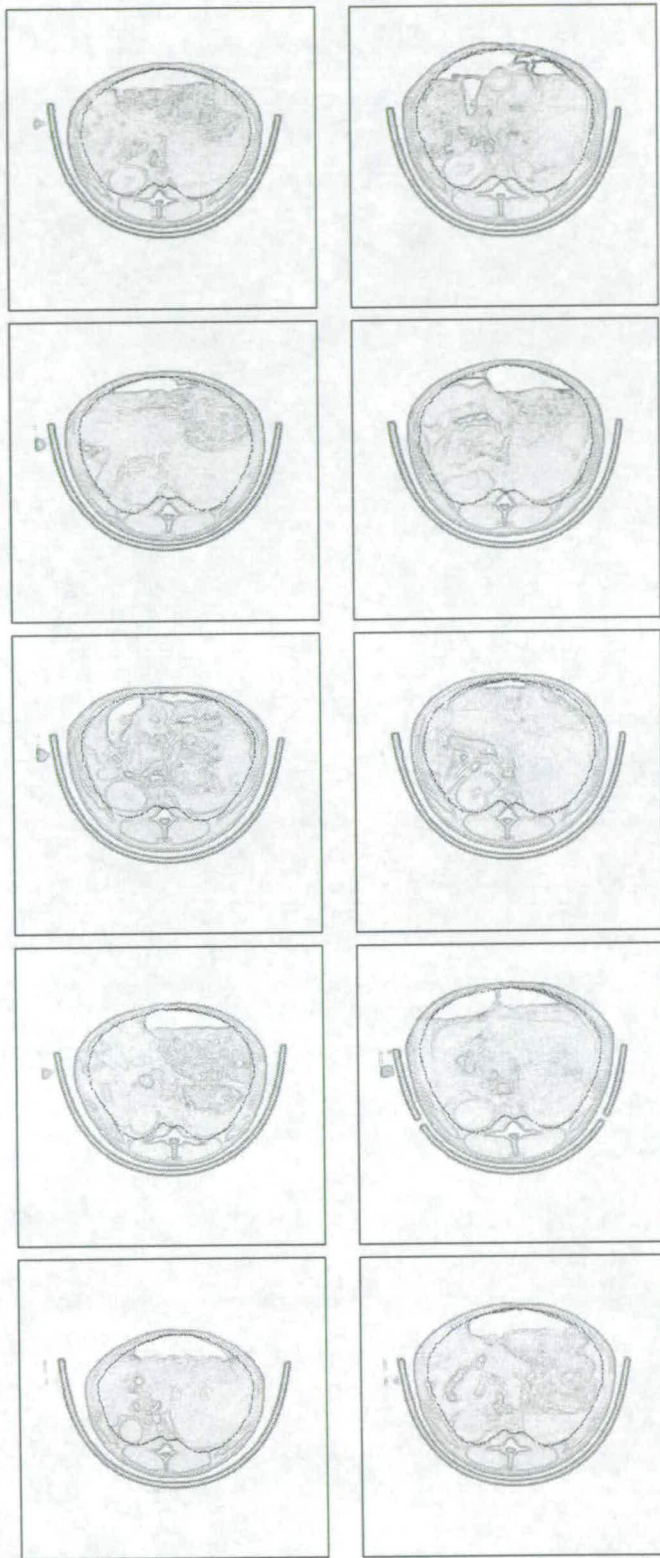


Figure 4.26: *Prewitt's filter applied to the validation data set, with the fitted templates (using rank 2) superimposed onto the inverse gradient filtered images.*

Image number	Sum of Squares of differences (rank 2)
1	0.108
2	0.063
3	0.156
4	0.245
5	0.350
6	0.077
7	0.328
8	0.080
9	0.248
10	0.146
average	0.180

Table 4.4: *Sums of squares of differences between the true proportional radii and the fitted proportional radii for each image in the validation set with rank 2.*

The polar representation of these mirror images is easily obtained by negating the vector of angles of the boundary pixels, so that a pixel previously with polar representation  $(r, \theta)$  now becomes  $(r, -\theta)$ . The polar form for the mirror images is based on the polar representation of the images in Section 4.3.

Using these mirror images with the original images has the effect of doubling the training set. To see how this affects the principal component scores and the corresponding eigenvectors and eigenvalues, suppose the  $i$ th vector of estimated Fourier coefficients in Section 4.3 is partitioned as

$$\mathbf{b}_i = \begin{bmatrix} \mathbf{b}_{ci} \\ \mathbf{b}_{si} \end{bmatrix} \quad i = 1, \dots, N,$$

where the two sub-vectors are the estimated coefficients for the cosines (and constant term) and the sines. Therefore, the corresponding partition of  $\mathbf{B}$  is given by

$$\mathbf{B} = [ \mathbf{B}_c \quad \mathbf{B}_s ],$$

where  $\mathbf{B}_c$  and  $\mathbf{B}_s$  are  $N \times (h + 1)$  and  $N \times h$  respectively. The corresponding matrix for the mirror images is

$$[ \mathbf{B}_c \quad -\mathbf{B}_s ].$$

Therefore, including the original and mirror images replaces  $\mathbf{B}$  by the  $2N \times p$  matrix

$$\begin{bmatrix} \mathbf{B}_c & \mathbf{B}_s \\ \mathbf{B}_c & -\mathbf{B}_s \end{bmatrix},$$

so that  $\bar{\mathbf{b}}$ , (which is  $p \times 1$ ), is replaced by  $\begin{bmatrix} \bar{\mathbf{b}}_c \\ \mathbf{0} \end{bmatrix}$ , where  $\bar{\mathbf{b}}_c$  is  $(h+1) \times 1$  and  $\mathbf{0}$  is  $h \times 1$ . Therefore, (4.3) becomes

$$\begin{aligned} \begin{bmatrix} \mathbf{B}_c & \mathbf{B}_s \\ \mathbf{B}_c & -\mathbf{B}_s \end{bmatrix} &= \mathbf{1}_{2N} \begin{bmatrix} \bar{\mathbf{b}}_c^T & \mathbf{0}^T \end{bmatrix} + \mathbf{H}_{2N} \begin{bmatrix} \mathbf{B}_c & \mathbf{B}_s \\ \mathbf{B}_c & -\mathbf{B}_s \end{bmatrix} \\ &= \mathbf{1}_{2N} \begin{bmatrix} \bar{\mathbf{b}}_c^T & \mathbf{0}^T \end{bmatrix} + \begin{bmatrix} \mathbf{B}_c - \mathbf{1}_N \bar{\mathbf{b}}_c^T & \mathbf{B}_s \\ \mathbf{B}_c - \mathbf{1}_N \bar{\mathbf{b}}_c^T & -\mathbf{B}_s \end{bmatrix} \end{aligned} \quad (4.17)$$

and  $\mathbf{B}^T \mathbf{H}_N \mathbf{B}$  is replaced by

$$\begin{bmatrix} \mathbf{B}_c^T & \mathbf{B}_c^T \\ \mathbf{B}_s^T & -\mathbf{B}_s^T \end{bmatrix} \begin{bmatrix} \mathbf{I}_N - \frac{1}{2} \mathbf{J}_N & -\frac{1}{2} \mathbf{J}_N \\ -\frac{1}{2} \mathbf{J}_N & \mathbf{I}_N - \frac{1}{2} \mathbf{J}_N \end{bmatrix} \begin{bmatrix} \mathbf{B}_c & \mathbf{B}_s \\ \mathbf{B}_c & -\mathbf{B}_s \end{bmatrix} = \begin{bmatrix} 2\mathbf{B}_c^T \mathbf{H}_N \mathbf{B}_c & \mathbf{0} \\ \mathbf{0} & 2\mathbf{B}_s^T \mathbf{B}_s \end{bmatrix}.$$

Thus, if  $\mathbf{p}_c$ ,  $\mathbf{p}_s$  are any eigenvectors of  $2\mathbf{B}_c^T \mathbf{H}_N \mathbf{B}_c$  and  $2\mathbf{B}_s^T \mathbf{B}_s$  respectively then  $\begin{bmatrix} \mathbf{p}_c \\ \mathbf{0} \end{bmatrix}$  and  $\begin{bmatrix} \mathbf{0} \\ \mathbf{p}_s \end{bmatrix}$  are eigenvectors of

$$\begin{bmatrix} 2\mathbf{B}_c^T \mathbf{H}_N \mathbf{B}_c & \mathbf{0} \\ \mathbf{0} & 2\mathbf{B}_s^T \mathbf{B}_s \end{bmatrix}.$$

The eigenvalues are twice those of  $\mathbf{B}_c^T \mathbf{H}_N \mathbf{B}_c$  and of  $\mathbf{B}_s^T \mathbf{B}_s$ . Since this is block-diagonal there are  $\min(N-1, \frac{1}{2}(p+1))$  non-zero eigenvalues for  $\mathbf{B}_c^T \mathbf{H}_N \mathbf{B}_c$  and  $\min(N, \frac{1}{2}(p-1))$  non-zero eigenvalues for  $\mathbf{B}_s^T \mathbf{B}_s$ .

It can be seen from Table 4.5 that when the mirror images are added to the original 24 images, producing a training set twice the previous size, seven principal components accounted for 85.2% of the total variation between the 41 Fourier coefficients. It can be seen that the eigenvalues are very similar for both analyses and the last two columns of Table 4.5 indicate whether the largest eigenvalues came from either  $\mathbf{B}_c^T \mathbf{H}_N \mathbf{B}_c$  or  $\mathbf{B}_s^T \mathbf{B}_s$ . It can be seen that the first, second, fourth, fifth and seventh largest eigenvalues are from  $\mathbf{B}_c^T \mathbf{H}_N \mathbf{B}_c$ . It is interesting to note that the third eigenvalue in the mirror image reduced rank analysis actually comes from  $2\mathbf{B}_s^T \mathbf{B}_s$ . It was shown in Figure 4.20, that the third principal component affected the asymmetry of the fitted boundary, showing that all the eigenvectors from  $2\mathbf{B}_s^T \mathbf{B}_s$  also affect the asymmetry to some extent in the images. This idea has not been investigated here but could be pursued further in future work, see Section 7.2.

Table 4.6 shows that three principal components minimised (4.16), where the average SS of differences (over the 24 images) between the true and fitted proportional radii is 0.132. The average SS from the previous analysis, not including mirror images in the training set, have been included to provide direct comparison. It can be seen that the sum of squares for ranks 0 to 7, is slightly higher



Previous analysis		Mirror image analysis			
eigenvalue $\times 10^{-5}$	% of variance	eigenvalue $\times 10^{-5}$	% variance	from matrices of	
				cosines	sines
25.4	34.4	24.5	32.0	×	
19.8	26.7	18.8	24.5	×	
6.83	9.3	6.82	8.9		×
5.01	6.7	5.82	7.6	×	
4.70	6.4	4.60	6.0	×	
2.39	3.2	2.84	3.7		×
1.72	2.3	1.89	2.5	×	

Table 4.5: *The largest seven eigenvalues for the reduced rank approximation, before and after the mirror images of the training set were included in the analysis. The last two columns indicate whether the eigenvalues in the mirror image analysis come from the matrices containing either the cosine or sine coefficients.*

Type of analysis	Rank used in reduced rank approximation							
	0	1	2	3	4	5	6	7
Mirror images	0.175	0.145	0.136	0.132	0.134	0.133	0.133	0.133
No mirror images	0.154	0.126	0.124	0.132	0.130	0.128	0.128	0.128

Table 4.6: *Average SS over 24 training images using various ranks in the reduced rank approximation when the mirror images were included in the analysis. The averages from Table 4.3, when mirror images are not included in the analysis, are shown for comparison.*

when the mirror images are included. Figure 4.27 shows the images in Figures 4.2 (a) and (b) with the fitted boundaries using the mirror image analysis, with rank 3, superimposed.

The increase in the average sum of squares for the various ranks by including the mirror images could be expected because it assumes symmetry in some average sense. However, the real test of whether the inclusion of these mirror images actually improves the boundary fitting results comes from applying this method to images not in the training set. Therefore, using the same ten validation images as in Section 4.10, boundaries are reconstructed using ranks from 0 to 7, in the reduced rank approximation which included mirror images. The average SS over the ten images are shown in Table 4.7. The corresponding results when the mirror images are not included in the analysis for these validation images are shown for comparison. It can be seen that for each of the ranks, (including the mean template with rank 0) the average SS over the ten images is actually larger when the mirror images of the 24 images are included in the training set. These results indicate that the inclusion of the mirror images of the training set actually lead to the inner boundaries being fitted less accurately than when the mirror images



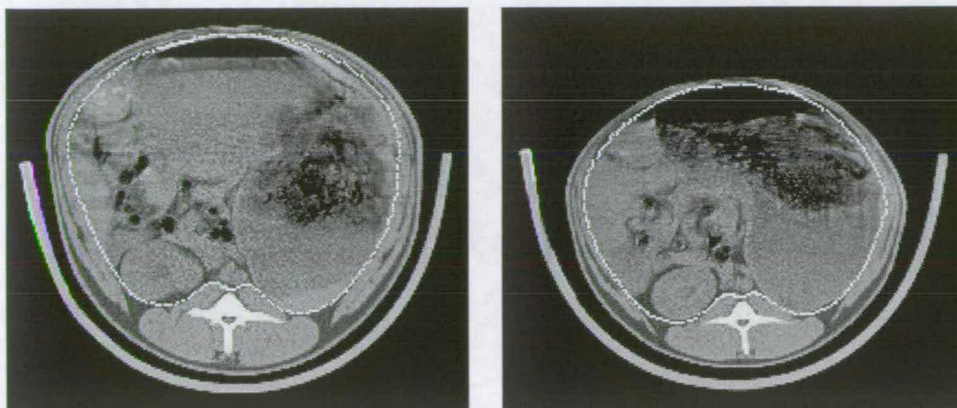


Figure 4.27: *Fitted boundaries, using the results from the reduced-rank approximation when the mirror images were included in the training set, superimposed onto the original image for the images displayed in Figures 4.2 (a) and (b) respectively.*

Type of analysis	Rank used in reduced rank approximation							
	0	1	2	3	4	5	6	7
Mirror images	0.229	0.191	0.181	0.178	0.180	0.181	0.182	0.183
No mirror images	0.220	0.186	0.180	0.178	0.179	0.175	0.174	0.174

Table 4.7: *Average SS over 10 validation images using various ranks in the reduced rank analysis approximation when the mirror images were included in the analysis. The averages when mirror images are not included in the analysis, are included for comparison.*

are not included.

## 4.12 Summary and conclusions

A deformable template approach has been used to perform the segmentation of the X-ray CT lumbar images, see also Glasbey et al. (1999). A training set of 24 images were used to build a parametric model of the template with Fourier coefficients. Principal component analysis was used to reduce the dimensionality and estimate a distribution on the parameters of the template. This distribution was combined with local gradient information from the image and various image enhancements were used on the images prior to optimisation. We found that the segmentation was greatly improved by including more prior information in the model concerning the size and orientation of the sheep. This was achieved by considering the inner boundary as a proportion of the outer boundary and taking into account the rotation of some sheep in the cradle after restraint.

We also investigated the inclusion of the mirror images of the 24 inner boundaries

into the training set. However, we were unable to improve upon the accuracy of the fitted boundaries prior to this investigation.

It can be seen from the results presented in this chapter, that deformable templates provide a very effective approach for the segmentation of the X-ray CT images. On comparison with the manual segmentation, deformable templates are capable of locating the required inner boundary extremely accurately.

We now consider the second stage of the process (see Section 1.4), which involves estimating tissue proportions in the segmented images. This is discussed in Chapter 5.

# Chapter 5

## Probability density function of pixel values

### 5.1 Introduction

Having automated the identification of the relevant tissue areas to produce images similar to Figure 4.3, the next step is to estimate the proportions of each tissue present in each image. One method for estimation is to consider each pixel separately, assign it to a tissue type and finally sum over all the pixels.

Classifying each pixel would be a relatively simple task if each pixel represented one tissue and each tissue was represented by a single greyscale value. If this were the case then it would simply be a matter of counting the number of pixels with these individual values. However, an image is never an exact representation of the object under observation as the output of any system is corrupted by the system itself. In addition, each tissue is represented by a range of values (due to the method of imaging, image noise, inhomogeneity of the tissues and other factors) that may overlap with the ranges of other materials, making it difficult to quantify the amount of each tissue in the image. Therefore, it is important to know how the imaging system affects the input. However, in part due to the finite resolution of the X-ray machine, averaging takes place between the tissue types, resulting in mixed pixels. These mixed pixels are a combination of two or more tissue types and this mixing of tissues makes it difficult to obtain an accurate quantification of the image.

In Section 5.2, recent papers which model the distribution of pixel values, predominantly concerning medical images, are reviewed. In Subsection 5.3.2, three representations of the point spread function for the X-ray CT machine are fitted to the data and an isotropic bivariate normal density is chosen. This is used to

derive a new probability density function for the mixed pixels, termed the *mixed pixel density*: see Subsection 5.3.4. This distribution is combined with the distributions for the pure pixels to obtain a density function of pixel values in the image.

## 5.2 Review of recent literature on estimating the probability distribution of pixel values

Santago and Gage (1995) briefly review current work on the classification of pixels and the quantification of a given tissue. However, for the purposes of this work it is not necessary to use the classification approach as it is sufficient to estimate the overall proportions of each tissue without estimating the exact type of tissue/tissues present in each individual pixel.

A method of quantification is described by Gage et al. (1992), based on examining the histogram of pixel values in the image, and from this they develop a finite-mixture density model (see Everitt and Hand, 1981) for the distribution of the pixel values. The histogram of pixel values for a typically segmented lumbar image (such as Figure 4.3), is shown in Figure 5.1. This histogram will be discussed in more detail in the following section. Each tissue type has a probability density function with different parameter values and with a different probability. This approach of using a finite-mixture density will be adopted to estimate the tissue proportions in the X-ray lumbar images. Having estimated the parameters, Gage et al. (1992) use a classifier which has prior probabilities for the tissues, in order to sort the pixels into the different tissue types, although the proportions of the mixture density could have been used directly. The pixels which are determined to be mixed pixels are allocated among the appropriate tissues equally. Therefore, the total amount of each tissue is found by adding the pure pixels and the appropriate number of pixels from the mixed pixels.

Thaler et al. (1978) consider estimating the proportions of the three types of matter present in voxels (volume elements) in the brain. They assume that pure voxel values have independent Normal densities with different means,  $\mu_i$ , and variances,  $\sigma_i^2$ , and the greyscale values for mixed voxels, composed of all three types of tissue, (in proportions,  $p_i$ ) follow a Normal distribution with mean and variance given by  $N(\sum_{i=1}^3 p_i \mu_i, \sum_{i=1}^3 p_i \sigma_i^2)$ . They provide no derivation of the probability distribution of the mixed voxels to justify the assumption. They also assume that the variation in the proportions from voxel to voxel follows a Dirichlet

probability distribution.

To date, the approaches which use the finite mixture distributions either do not take into account the modelling of mixed pixels, (e.g. see Lei and Sewchand, 1992a), or they include them as being a different material, where the parameters of the probability density function of the mixed pixels are unconnected to the parameters of its pure pixels elements (e.g. see Luiting et al., 1995).

Luiting et al. (1995) also consider a finite mixture density to model the pixel values, and incorporate mixed pixels into their model in order to estimate the amount of fat and muscle present in pig carcasses. They assume that the observed greyscale frequency distribution is the sum of independent normal density functions for both fat and muscle tissues. They also assume that the values for mixed pixels follow a normal density function, but they do not express the parameters in terms of the means and variances of the pure tissues, but instead appear to treat the mixed pixel density function as a separate distribution with parameters independent of the pure tissue parameters. After maximising the likelihood to estimate these parameters, they differ from Thaler et al. (1978) by splitting the proportion of mixed pixels over fat and muscle according to the ratio of the proportions of pure fat to pure muscle. They also perform their analysis without the inclusion of mixed pixels, assuming that only pure pixels exist. They find that the parameter estimates became less biased and the estimated proportions were greatly improved through the inclusion of the mixing distribution.

Lei and Sewchand (1992a) approach the estimation of tissue types in an image in a slightly different manner. They also use a finite mixture of Normal densities and a classifier with prior probabilities for the tissue types but they do not assume to know the number of distinct regions present in the image. They use an information criterion to estimate this number initially but do not explicitly include mixed pixels in their model. On estimating the number of regions, they find the optimal number is in fact larger than the number of pure tissues expected to be present and they explain this result by the existence of mixed pixels.

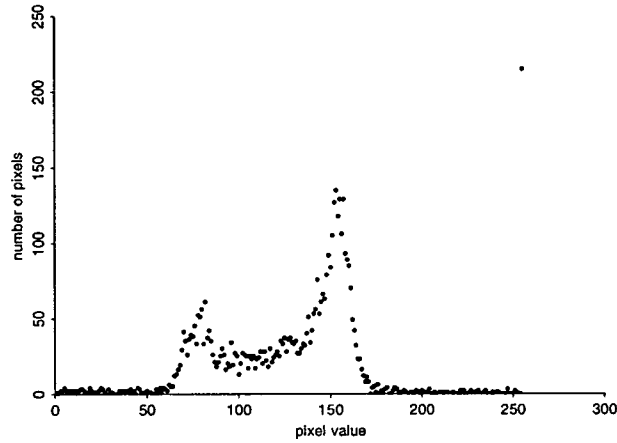


Figure 5.1: Shows a typical histogram of the pixel values in a segmented lumbar image, similar to Figure 4.3.

## 5.3 Approximating the probability density function of the pixel values

### 5.3.1 Review of literature on estimating the point spread function, (PSF)

Although each pixel in the image is represented by a single greyscale value, it is actually the average value of the amount of transmitted X-rays for a small region around that point in the image. As a result of this averaging, there is an area between two adjacent tissues whose greyscale values are different from those in either of the two tissues regions. The pixels in this area are defined as mixed pixels, and their greyscale values are usually between those of the two adjacent tissues, i.e. an ideal sharp edge in greyscale values does not exist on the boundary between the two tissues. Generally this averaging extends more than the distance between pixels, and Lei and Sewchand (1992a) state that the width of this averaging in X-ray CT imaging is about 1-2 pixels. Due to this averaging, many pixels in the CT images are actually responses to mixtures of two or more tissue types. Therefore, in trying to estimate the tissue proportions effectively it is necessary to take this mixing into account.

A typical histogram of a segmented lumbar image is shown in Figure 5.1. The greyscale values are displayed in the range  $[0,255]$  rather than in the original Hounsfield units: see Chapter 1. The peak at 0 (i.e. the background pixels) has been removed to allow a magnified view of the other greyscale values. The smaller peak around pixel value 75 corresponds to the fat pixels and the larger

peak around 150 corresponds to muscle pixels. Many, but not all, of the pixels lying between these peaks correspond to pixels which are a mixture of fat and muscle. The peak at greyscale value 255 corresponds to the bone pixels.

It can also be seen from Figure 5.1 that there are negligible amounts of mixed pixels between the muscle and the bone peaks and similarly between the fat peak and the background pixels. Also, since the bone pixels and the background pixels can be easily identified in an image using thresholding methods, our estimation of tissue amounts is restricted to fat and muscle only. It is assumed that all the mixed pixels are a combination of fat and muscle only and the derived probability density function for the greyscale values in an image (in Subsection 5.3.4) makes provision for pure fat and muscle pixels and a mixture of these tissues only.

In order to develop a distribution which accurately represents the pixel values shown in Figure 5.1 (excluding the background and bone pixels), and which takes into account the mixed pixels, it is first necessary to model the spatial response of a pixel for the X-ray machine. This means it is necessary to understand and model how the average value is obtained, i.e. is it a linear combination of all pixels in some local neighbourhood or is there a larger weighting to regions very close to the point? This spatial response is known as the point spread function (PSF).

The standard use of the point spread function in image analysis is in deconvolution to restore a blurred image. The main aim of restoration is to improve a given image in some sense by modelling the degradation and applying the inverse process in order to recover a better approximation to the original image. Early techniques for digital image restoration were derived mostly from the frequency domain (see Gonzalez and Woods, 1993).

A CT scanner blurs the images and also introduces noise (see Dore et al., 1997). A common model for studying such a system is to consider it as linear, spatially invariant and corrupted by additive output noise, such that

$$g(x, y) = (f * h)(x, y) + n(x, y), \quad (5.1)$$

where  $g(x, y)$  is the observed (degraded) image,  $f(x, y)$  is the original image,  $n(x, y)$  is random output noise, which may or may not be present,  $h(x, y)$  is the point spread function and  $*$  is a two-dimensional convolution. The assumption that the observed picture  $g(x, y)$  is a linear function of the ideal image  $f(x, y)$  as in (5.1) is approximately correct only over a small dynamic range of grey levels (see Rosenfeld and Kak, 1982). In most cases with various imaging modalities, it is assumed that the point spread function is position-invariant, i.e. the result depends only on a value of  $f(x, y)$  at a given point in the image and not on the



position of the point. If it is assumed that there is no noise present, then (5.1) reduces to

$$g(x, y) = (f * h)(x, y). \quad (5.2)$$

The PSF can be estimated in the frequency domain by taking Fourier transforms of both sides and using the convolution theorem to obtain

$$G(u, v) = F(u, v)H(u, v), \quad (5.3)$$

where  $G$ ,  $F$  and  $H$  are the Fourier transforms of  $g(x, y)$ ,  $f(x, y)$  and  $h(x, y)$  respectively:  $H(u, v)$  is also known as the *modulation transfer function* of the system that transforms the ideal image  $f(x, y)$  into  $g(x, y)$ . Rosenfeld and Kak (1982) demonstrate how in some cases the physical phenomenon underlying the degradation can be used to determine  $h(x, y)$ , for example in optical imaging systems or in photography. If the blurring is of an unknown nature or if the phenomenon underlying the degradation is too complex for analytical determination of  $h(x, y)$ , the only possible alternative is to estimate it from the degraded picture itself. Many experimental PSF identification techniques adopt one-dimensional linear system methods based on impulse methods, step functions or frequency methods. Dore et al. (1997) give details on these methods.

Glasbey et al. (1994) look at estimating the PSF of a desktop scanner using frequency domain methods. They examine digitised versions of binary images which show intermediate greyscale values because of blurring. To estimate the PSF they calculate the Fourier amplitudes of both the blurred and true versions of the binary images. Assuming there is no noise, as in (5.2), the ratios of these Fourier amplitudes are the Fourier amplitudes of the PSF. On examination of these ratios they find the PSF is isotropic and they approximate it by an integrated form of a bivariate Cauchy distribution.

The approach used to identify the PSF of the SAC-BioSS CT scanner is based on the knowledge of a true edge that is known to exist in the image.

A system modelled by (5.1) is completely characterised by the PSF and the noise, which accounts for the blur in the spatial domain. Output noise from the CT system is due to a variety of factors, resulting from the interaction of X-ray photons with the tissues, as well as noise at the detectors. It is generally assumed for many imaging modalities that output noise is uncorrelated to the original image  $f(x, y)$ , but Dore et al. (1997) state that this is not true of X-ray CT scanners. In general, the level of noise increases with the X-ray attenuation of

a tissue. They propose to identify the PSF of a CT scanner using a correlation-based method, through using the Wiener-Hopf equation, (see Dore et al., 1997). Although correlation-based methods have been used in other fields to estimate the PSF, they had not been used for medical imaging systems prior to this paper. In it they construct a ‘phantom’ consisting of a series of randomly located holes in order to estimate the PSF. They assume that the PSF is position invariant and that it is of finite size. A ‘phantom’ is an object made specifically for use either to check the resolution of the X-ray machine or as in Dore et al. (1997) to estimate the PSF. A phantom generally consists of objects of different sizes and the exact form of the input image is known before scanning. Therefore, this allows the PSF to be determined, given the output image after scanning.

Rathee et al. (1992a) define the PSF of a CT system as the image of an infinitesimally small point object. The main factor in image blurring for small objects in CT is the finite size of the X-ray beam profiles. In an earlier paper, (Rathee et al., 1992b), they assume that the PSF is spatially invariant for an idealised CT scanner. However, in their second paper they state that the PSF is in reality spatially variant due to the divergent and spatially variant fan beams of X-rays: see Subsection 2.3.1. Therefore, two point objects located at two different positions in the object plane are blurred differently. In general, the PSF of a CT system is spatially variant and is a function of the radial position of the point object.

Although the noise in CT images is spatially correlated and data dependent, recent formulations assume a noise-free image degradation model. For the purposes of estimating the PSF of the SAC-BioSS CT scanner, this assumption will also be made here, together with the PSF being spatially invariant. The edge method mentioned earlier is also used by assuming that a step edge is known to exist in reality in the image.

### **5.3.2 Estimation of the PSF of the SAC-BioSS CT scanner**

Here a simple approach is used in order to estimate the PSF. It is known that in reality the boundary from air to the cradle is a step edge, (see Figure 5.2(a)), i.e. so in the image if there was no PSF then all the pixels in this region should have a greyscale value corresponding to either air or the cradle only. A subsection of the cradle in Figure 5.2(a) is shown in Figure 5.2(b), which shows two boundaries between cradle and air pixels. Since the cradle is made from a homogeneous material it would be assumed that the greyscale values of each of the cradle

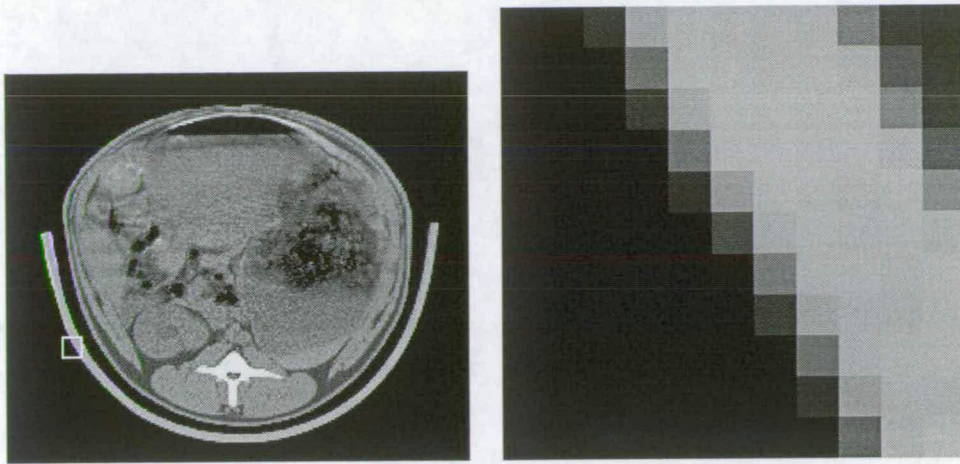


Figure 5.2: (a) An X-ray CT image before processing with the white square identifying the sub-image of the cradle in (b): (b) is used to estimate the point spread function.

pixels should be the same if there was no averaging taking place. It can be seen that this is obviously not the case. It can be seen in Figure 5.2(b) that the cradle pixels which lie on the two boundaries with air are slightly darker than those cradle pixels slightly further from the boundaries (i.e. the pixels in the middle of the cradle). It is assumed that the averaging taking place between the air and cradle pixels is the same at each of the two boundaries shown in Figure 5.2(b) and the averaging is also assumed to be the same over the rest of the image i.e. we are assuming a spatially invariant PSF, similar to Dore et al. (1997). Therefore, it is only necessary to consider one of these boundaries and here the larger of the two boundaries is used, i.e. the cradle/air boundary which lies on the left of Figure 5.2(b). It was assumed in Section 5.3 that all the mixed pixels in the carcass are a combination of fat and muscle only. From examination of Figure 5.2(a), the majority of the mixed pixels in the carcass region lie relatively close to the cradle pixels. Therefore, if this assumption of a spatially invariant PSF is slightly incorrect, then the estimated PSF found from this air/cradle region would still be quite accurate.

In order to estimate the PSF, the outer boundary of the cradle was assumed to be approximated by the arc of a circle with radius  $\gamma$  and centred at  $(\kappa_1, \kappa_2)$ . Since we are restricting ourselves to one of the boundaries, it is only necessary to model the  $N$  pixels near the edge of the cradle. For the boundary of interest, 44 pixels were chosen.

The greyscale values of these 44 pixels which lay close to the boundary with pixel

location  $(i, j)$ , were modelled by

$$f_{ij}^* = \alpha + \beta \psi \left( \frac{d_{ij}}{\tau} \right), \quad (5.4)$$

where

$$d_{ij} = \gamma - \sqrt{(\kappa_1 - i)^2 + (\kappa_2 - j)^2}$$

is the distance from pixel  $(i, j)$  to the outer edge of the cradle. The air and cradle regions are modelled as homogeneous materials with mean pixel values  $\alpha$  and  $(\alpha + \beta)$  respectively and  $\psi$  is a function with range  $[0, 1]$  that specifies the response of a pixel after scaling by  $\tau$ . It is a one-dimensional integral of the two-dimensional PSF. The expected greyscale value of pixel  $(i, j)$  is  $f_{ij}^*$ , with observed greyscale value  $f_{ij}$ . It should be mentioned here that in order to model the PSF the original Hounsfield unit values (approximately in the range  $-1024$  to  $+200$ ) in this section of cradle were used rather than the transformed values which lie in the range  $[0, 255]$ .

The six parameters in (5.4), i.e.  $(\alpha, \beta, \gamma, \kappa_1, \kappa_2, \tau)$  are estimated by minimising the root mean square error, (RMSE), between the data and the model, i.e.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i,j} (f_{ij} - f_{ij}^*)^2},$$

where  $N$  is the number of pixels which lay close to the boundary and the sum is over the values of  $i$  and  $j$  corresponding to the selected 44 pixels. Three choices of  $\psi$  were tested:

- a linear ramp where

$$\psi(x) = \begin{cases} 0 & \text{for } x < -1 \\ \frac{x+1}{2} & \text{for } -1 \leq x \leq 1 \\ 1 & \text{for } x > 1 \end{cases}$$

- a logistic function where  $\psi(x) = \frac{\exp x}{1 + \exp x}$
- a normal integral where  $\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-y^2/2) dy$

Table 5.1 shows the results of minimising the RMSE for the three forms of  $\psi$ , over the same 44 pixels. Figure 5.3 (a), (c) and (e) show the greyscale values of the 44 pixels near the outer edge of the cradle, plotted against the distance from the fit of the model in (5.4), for each of the three forms of  $\psi$ . The fitted greyscale values have been superimposed as a continuous line. Figures 5.3 (b), (d) and (f)

$\psi$	RMSE
ramp	34.6
logistic	29.7
normal integral	28.6

Table 5.1: *The minimised RMSE of (5.4) for the three  $\psi$  functions.*

Parameter	$\alpha$	$\beta$	$\gamma$	$\kappa_1$	$\kappa_2$	$\tau$
Estimate	-1025.76	1161.99	224.81	34.89	242.06	0.41

Table 5.2: *Estimated values for the six parameters which minimised the RMSE with  $\psi$  as a normal integral.*

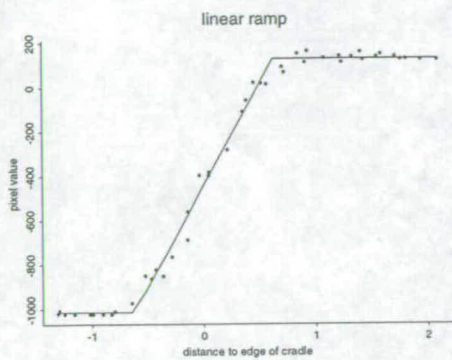
show the sub-image of the cradle from Figure 5.2(b) with the estimated cradle boundary superimposed.

It can be seen from Table 5.1 that the normal integral performs best, although there was little difference between this and the logistic function. However, the normal integral is the only form of  $\psi$  that corresponds to a simple two-dimensional PSF, i.e. a bivariate normal density: see Subsection 5.3.3.

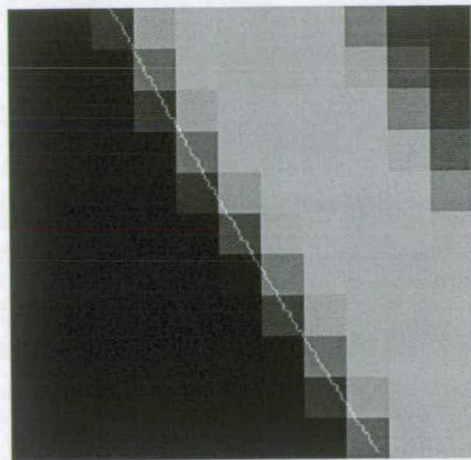
Table 5.2 shows the six parameter estimates obtained for the normal integral. The most important parameter here is  $\tau$ , which is the standard deviation of the isotropic bivariate normal density: see Subsection 5.3.3. The estimate of 0.41 pixel units is important because it is used later in order to derive the distribution of pixel values in an image. The estimate,  $\hat{\alpha}$ , of the mean greyscale value of the air pixels is -1025HU and accurately reflects the true greyscale value of air, which is -1024HU: see Section 1.2. The estimate,  $\hat{\alpha} + \hat{\beta}$ , of the mean greyscale values of the cradle pixels is 136HU. It can be seen from Figure 5.2 (a) that the cradle pixels are more dense than the muscle pixels and hence have larger Hounsfield numbers. On comparison with Figure 1.2 and the range of HU for muscle, this value of 136HU appears an accurate estimate of the mean value of the cradle pixels. Also, the estimate of  $\gamma$  corresponds to the estimated radii of the arc of the circle.

The estimated PSF is used in Subsection 5.3.4 in the derivation of the *mixed pixel distribution*. At present, as stated earlier, the method used here only considers two tissues in the model, although this may be extended to deal with more.

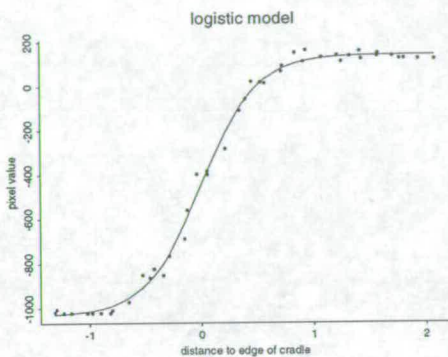




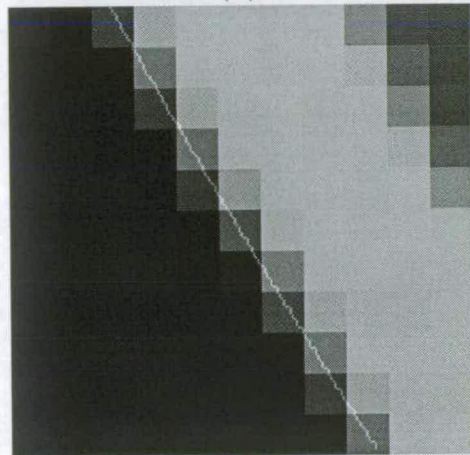
(a)



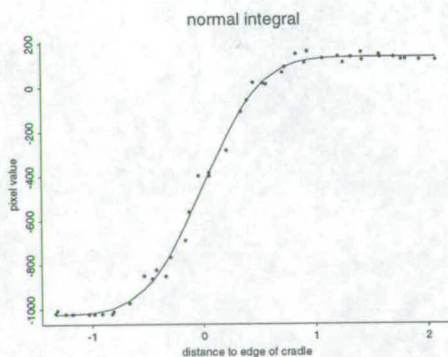
(b)



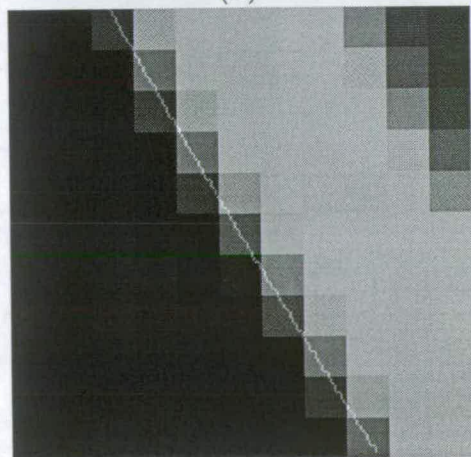
(c)



(d)



(e)



(f)

Figure 5.3: (a), (c) and (e) show the greyscale values for  $N$  pixels near the edge of the cradle plotted against the distance from the edge of the cradle, with the fitted model in (5.4), using the linear ramp, logistic and normal integral functions. (b), (d) and (f) show the sub-image from Figure 5.2(b), with the estimated edge of the cradle superimposed, again using the three forms of  $\psi$ .

### 5.3.3 Observed image for a circular normal PSF and a linear boundary

The normal integral which gave the best fit is also the only one of the three cases for  $\psi$  that corresponds to an analytically simple two-dimensional point spread function, which in this case is a bivariate isotropic normal density. In this section we will show that applying a bivariate normal PSF to an image produces the one-dimensional normal integral which minimised the RMSE in Subsection 5.3.2 (see Table 5.1).

On examination of segmented images similar to Figure 4.3, it is clear that the positioning of the layers of fat and muscle and hence the amount of each tissue varies between sheep. The boundaries between two tissues are reasonably smooth and continuous, and therefore are assumed to be locally linear. It can also be seen that the three layers of muscle/fat at the side of the carcass are approximately parallel to each other and to the outside of the body. These assumptions will be used in the simulation study in Chapter 6. It was decided due to the random position of the boundaries and the assumption of them being locally linear, that the boundaries separating the tissues could be approximated by random lines. The process of constructing random lines is given in Section 6.2.

Consider a locally linear tissue boundary,  $ax + by = c$  passing through an image, where a typical pixel in the image is denoted by  $(x, y)$ . If it is assumed that the pixels on the left and right of the line represent say fat and muscle, then using an indicator function we can denote fat as 1, and muscle as 0. It is assumed that the fat and muscle pixels have negative and positive distances respectively from the line. Therefore, the image may be represented as

$$r(x, y) = \begin{cases} 0 & \text{for } ax + by < c \\ 1 & \text{for } ax + by \geq c \end{cases} \quad (5.5)$$

An isotropic bivariate normal PSF, with standard deviation  $\tau$ , centred at  $(0,0)$  is given by

$$w(x, y) = \frac{1}{2\pi\tau^2} \exp \left\{ \frac{-(x^2 + y^2)}{2\tau^2} \right\}$$

Therefore, convolving this point spread function, centred on a pixel  $(x_0, y_0)$ , with (5.5) gives the weighted average greyscale value from the X-ray machine,  $\rho$ , on a pixel as



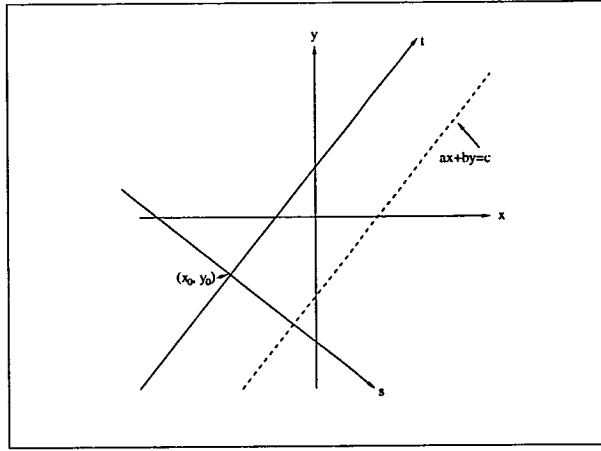


Figure 5.4: Geometrical representation of the new axes,  $s$  and  $t$ , after the change of variables.

$$\rho(x_0, y_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(x - x_0, y - y_0) r(x, y) dx dy, \quad (5.6)$$

where  $\rho$  may be interpreted as the proportion of fat in a pixel.

In order to shift the origin to the pixel  $(x_0, y_0)$  and to rotate the axes such that one axis is parallel to the boundary  $ax + by = c$ , a change of variables is introduced. This will simplify the integral in (5.6) where  $s = a(x - x_0) + b(y - y_0)$  and  $t = -b(x - x_0) + a(y - y_0)$ , (see Figure 5.4 for geometrical representation), and substituting into (5.6) gives

$$\rho = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w \left( \frac{as - bt}{a^2 + b^2}, \frac{bs + at}{a^2 + b^2} \right) r \left( \frac{as - bt}{a^2 + b^2} + x_0, \frac{bs + at}{a^2 + b^2} + y_0 \right) \frac{1}{J} ds dt \quad (5.7)$$

where the Jacobian,  $J$ , is given by

$$J = \left| \frac{\partial(s, t)}{\partial(x, y)} \right| = a^2 + b^2.$$

Using (5.5),

$$r \left( \frac{as - bt}{a^2 + b^2} + x_0, \frac{bs + at}{a^2 + b^2} + y_0 \right) = \begin{cases} 0 & \text{for } s < l \\ 1 & \text{for } s \geq l, \end{cases}$$

where  $l = c - ax_0 - by_0$ .

Therefore, (5.7) simplifies to

$$\rho = \int_l^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi\tau^2} \exp \left\{ -\frac{s^2 + t^2}{2\tau^2(a^2 + b^2)} \right\} \frac{1}{a^2 + b^2} ds dt.$$

After integration, we obtain

$$\rho = 1 - \Phi\left(\frac{l}{\tau\sqrt{a^2 + b^2}}\right) = \Phi\left(-\frac{l}{\tau\sqrt{a^2 + b^2}}\right) = \Phi\left(-\frac{D}{\tau}\right) \quad (5.8)$$

where

$$D = \frac{l}{\sqrt{a^2 + b^2}} = \frac{-ax_0 - by_0 + c}{\sqrt{a^2 + b^2}}$$

which is the perpendicular distance from the pixel, at position  $(x_0, y_0)$ , to the tissue boundary, with equation  $ax + by = c$ . This result is exactly the normal integral which minimised the RMSE in Subsection 5.3.2.

### 5.3.4 Using the estimated PSF to approximate the probability density function of pixel values

Using the estimated value 0.41 of  $\tau$  in the Normal PSF, see Table 5.2, it is reasonable to assume that any pixel which is more than a perpendicular distance of 1 (which is approximately  $2.5 \times \tau$ ), from the boundary between the two tissues is either pure fat or pure muscle. This is because if  $|D| > 1$  then  $\rho \approx 0$  or 1, since  $\Phi(-1/\tau) = 0.007$  and  $\Phi(1/\tau) = 0.993$ . Conversely, all pixels within a perpendicular distance of 1 of a boundary are assumed to be a mixed pixel (part fat and part muscle). In order to derive the probability density function and hence the cumulative distribution function of  $\rho$ , it is assumed that the perpendicular distance  $D$  is uniformly distributed, conditional on it taking a value between  $-1$  and  $+1$ . This assumption will be justified in Section 5.4.

The cumulative distribution function of  $D$ , denoted  $F(D)$  is given by

$$F(D) = \frac{1}{2}(D + 1) \quad (-1 \leq D \leq 1).$$

Therefore, the cumulative distribution function of  $\rho$ ,  $F(\rho)$ , is given by

$$\begin{aligned} F(\rho) &= 1 - Pr(D < -\tau\Phi^{-1}(\rho)) \\ &= \frac{1}{2}\{1 + \tau\Phi^{-1}(\rho)\} \quad (\Phi(-1/\tau) \leq \rho \leq \Phi(1/\tau)). \end{aligned} \quad (5.9)$$

The probability density function of  $\rho$ , denoted by  $f(\rho)$ , is found by differentiating (5.9) with respect to  $\rho$ . Since  $\rho$  in (5.8) is a differentiable decreasing function of  $D$ , its probability density function is given by

$$f(\rho) = \frac{dF(\rho)}{d\rho} \frac{dD}{dD} \quad \text{where} \quad \frac{dF(\rho)}{dD} = -\frac{1}{2}$$

and

$$\frac{d\rho}{dD} = -\frac{1}{\tau}\Phi' \left( -\frac{D}{\tau} \right)$$

where

$$\Phi' \left( -\frac{D}{\tau} \right) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{-D}{\tau} \right)^2 \right\} = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \Phi^{-1}(\rho)^2 \right\}.$$

Therefore, the probability density function of  $\rho$  is given by

$$f(\rho) = \sqrt{\frac{\pi\tau^2}{2}} \exp \left\{ \frac{[\Phi^{-1}(\rho)]^2}{2} \right\} \quad (\Phi(-1/\tau) \leq \rho \leq \Phi(1/\tau)). \quad (5.10)$$

Figures 5.5 (a) and (b) respectively show the probability density function and cumulative distribution functions of  $\rho$  under this assumption.

In order to consider the effect on  $f(\rho)$  and  $F(\rho)$  as  $\tau$  varies, it may be easier to re-express these distributions by defining  $D$  in terms of  $\tau$ . Above, the limits of the uniform distribution of  $D$  were chosen to be between  $-1$  and  $+1$  as this is approximately  $2.5 \times \tau$ . However, rather than substituting in these values of  $-1$  and  $+1$ , the distribution of  $D$  may be defined in terms of  $\tau$ . Therefore,  $F(D)$  is given by

$$F(D) = \frac{1}{2} + \frac{D}{5\tau} \quad \left( -\frac{5}{2}\tau \leq D \leq \frac{5}{2}\tau \right).$$

Therefore, the cumulative distribution function of  $\rho$ ,  $F(\rho)$  is given by

$$F(\rho) = \frac{1}{2} + \frac{\Phi^{-1}(\rho)}{5} \quad (\Phi(-5/2) \leq \rho \leq \Phi(5/2)). \quad (5.11)$$

Therefore, similar to above

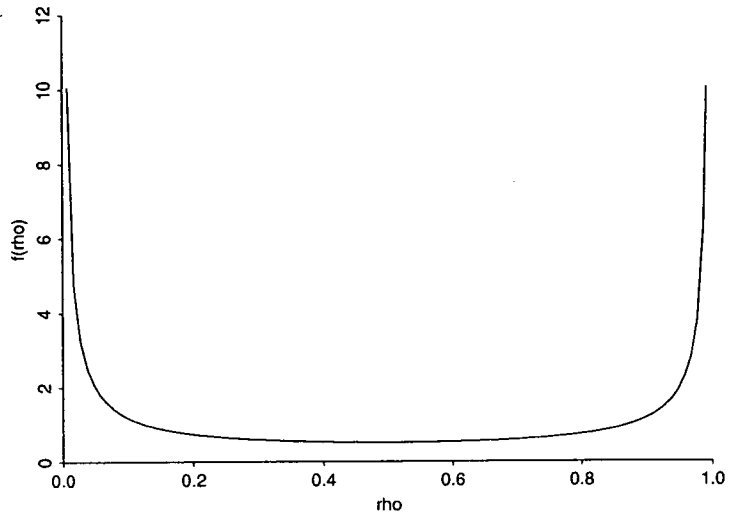
$$f(\rho) = \frac{dF(\rho)}{d\rho} \frac{dD}{dD} \quad \text{where} \quad \frac{dF(\rho)}{dD} = -\frac{1}{5\tau}.$$

Hence, the probability density function  $f(\rho)$  is given by

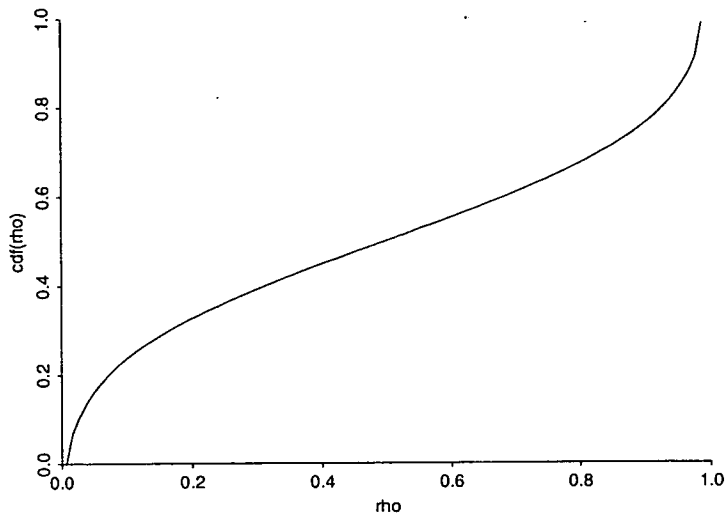
$$f(\rho) = \frac{\sqrt{2\pi}}{5} \exp \left\{ \frac{[\Phi^{-1}(\rho)]^2}{2} \right\} \quad (\Phi(-5/2) \leq \rho \leq \Phi(5/2)). \quad (5.12)$$

From (5.11) and (5.12) it can be seen that as  $\tau$  tends to zero or infinity then  $f(\rho)$  and  $F(\rho)$  remain the same for all values of  $\tau$ .

In reality it is known that fat and muscle do not have constant greyscale values. Lei and Sewchand (1992b) assume that in X-ray CT images, the pixel values for pure tissues are normally distributed within an image with different means and variances. Therefore, it is assumed that the greyscale values for pure fat pixels



(a)



(b)

Figure 5.5: (a) Probability density function (b) cumulative distribution function of  $\rho$ , the proportion of fat in a mixed pixel.

are normally distributed with mean  $\mu_f$  and variance  $\sigma_f^2$  and those for pure muscle pixels are normally distributed with mean  $\mu_m$  and variance  $\sigma_m^2$ . Assuming that these values are independently distributed, similar to Phillips and Smith (1994) and Glasbey (1998), then the greyscale value  $y$ , for a mixed pixel conditional on  $\rho$ , is also normally distributed, given by

$$y|\rho \sim N(\rho\mu_f + (1 - \rho)\mu_m, \rho\sigma_f^2 + (1 - \rho)\sigma_m^2). \quad (5.13)$$

This is similar to the distribution proposed by Thaler et al. (1978), who provided no proof of their result. Justification of this distribution is obtained by considering each pixel to be made up of  $n$  smaller ‘sub-pixels’ whose values are independently and identically distributed. So, if we know that for a pure fat pixel, the mean is  $\mu_f$  and the variance is  $\sigma_f^2$  then the sum of the means of the sub-pixels has to equal  $\mu_f$  and the sum of the variances of the sub-pixels has to equal  $\sigma_f^2$ . This leads to the result that the values for the  $i$ th fat sub-pixel, denoted by  $y_{fi}$ , has the following distribution

$$y_{fi} \sim N\left(\frac{\mu_f}{n}, \frac{\sigma_f^2}{n}\right). \quad (5.14)$$

Similarly for the sub-pixel values of a pure muscle pixel, denoted by  $y_{mi}$ ,

$$y_{mi} \sim N\left(\frac{\mu_m}{n}, \frac{\sigma_m^2}{n}\right). \quad (5.15)$$

Therefore, a mixed pixel, with proportion  $\rho$  of fat and  $1 - \rho$  of muscle can be regarded as  $n\rho$  fat sub-pixels and  $n(1 - \rho)$  muscle sub-pixels. Therefore, the mean and variance for the mixed pixel is the sum for the means and variances of the fat and muscle sub-pixels, hence providing the result in (5.13).

The probability density function for a mixed pixel,  $f(y)$ , (*mixed pixel density*) is found by convolving the conditional distribution (5.13) and the marginal distribution of  $\rho$  (5.10), i.e.  $f(y) = \int f(y|\rho)f(\rho)d\rho$ , which after some cancellation leads to

$$f(y) = \int_{\Phi(-\frac{1}{\tau})}^{\Phi(\frac{1}{\tau})} \frac{\tau \exp\left\{\frac{-1}{2} \left(\frac{[y - \rho\mu_f - (1 - \rho)\mu_m]^2}{[\rho\sigma_f^2 + (1 - \rho)\sigma_m^2]} - [\Phi^{-1}(\rho)]^2\right)\right\}}{2[\rho\sigma_f^2 + (1 - \rho)\sigma_m^2]} d\rho \quad (5.16)$$

This integral appears to have no analytic solution but can be computed by using numerical integration routines (Numerical Algorithms Group, 1993) for any

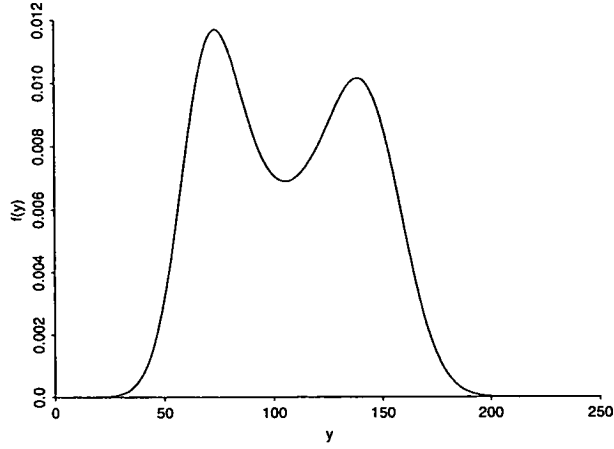


Figure 5.6: *The probability density function of mixed pixels,  $f(y)$ , for  $\mu_f = 65$ ,  $\sigma_f^2 = 150$ ,  $\mu_m = 150$ ,  $\sigma_m^2 = 250$  and  $\tau = 0.41$ .*

specified values of the parameters  $\mu_f$ ,  $\sigma_f^2$ ,  $\mu_m$  and  $\sigma_m^2$ . Figure 5.6 shows this distribution, where  $\mu_f = 65$ ,  $\sigma_f^2 = 150$ ,  $\mu_m = 150$  and  $\sigma_m^2 = 250$ . The choice of these values is justified in Chapter 6.

Finally, having derived the density (5.16) for mixed pixels in an image based on previous assumptions, it is possible to combine this with the distributions for the pure pixels, to obtain the probability density function of all pixel values in the image. Assuming that in any image the proportions of pure fat pixels and pure muscle pixels are  $\pi_f$ , and  $\pi_m$ , then it is assumed also that the remainder,  $(1 - \pi_f - \pi_m)$ , of the pixels are mixed pixels. Therefore, the probability density function of all pixel values in the image is given by

$$g(y) = \pi_f \frac{e^{-(y-\mu_f)^2/(2\sigma_f^2)}}{\sqrt{2\pi\sigma_f^2}} + \pi_m \frac{e^{-(y-\mu_m)^2/(2\sigma_m^2)}}{\sqrt{2\pi\sigma_m^2}} + (1 - \pi_f - \pi_m)f(y). \quad (5.17)$$

This distribution was fitted to the histogram of pixel values in Figure 5.1, by maximising the likelihood using a numerical optimisation routine, E04JAF (Numerical Algorithms Group, 1993). The maximisation was performed with  $\pi_f$ ,  $\pi_m$  and  $1 - \pi_f - \pi_m$  all guaranteed to be positive. The fitted probability density function is shown in Figure 5.7. The estimates of the two pure tissue proportion parameters were found to be  $\hat{\pi}_f = 0.05$ ,  $\hat{\pi}_m = 0.39$ , hence  $(1 - \hat{\pi}_f - \hat{\pi}_m) = 0.56$ . These results are fairly typical of the fitted proportions. Having obtained the maximum likelihood estimates for the six parameters, it was decided to adopt the approach similar to Gage et al. (1992) and Thaler et al. (1978) by estimating

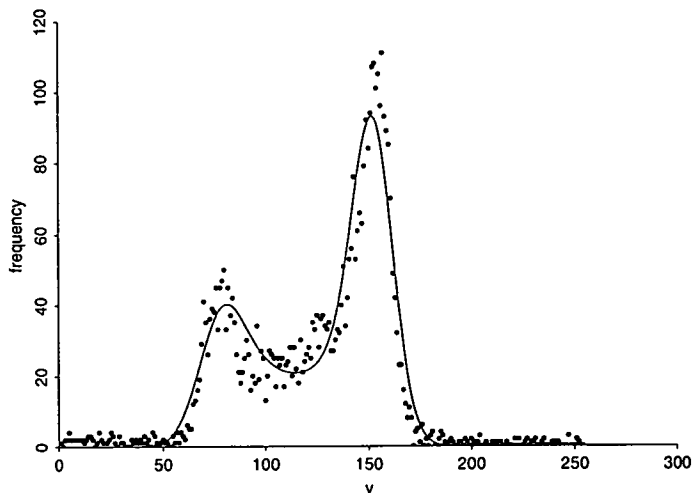


Figure 5.7: A histogram of greyscale values for one of the segmented lumbar sheep images. The superimposed probability density function of pixel values, given in (5.17), was fitted by maximum likelihood.

the total proportion of fat in the image as

$$\hat{\pi}_f + \frac{(1 - \hat{\pi}_f - \hat{\pi}_m)}{2}, \quad (5.18)$$

where half of the estimated proportion of mixed pixels are counted as fat and half as muscle. It was decided not to adopt the approach of Luiting et al. (1995), i.e. splitting the mixed pixels relative to the estimated proportions of pure fat and pure muscle. This is because, say for example the estimated proportion of fat was zero, i.e. no pure fat pixels. Using the method of Luiting et al. (1995) would result in there overall being no fat pixels at all and only muscle pixels, which would be incorrect. Given that it was shown above that the estimated proportion of pure fat can be very small, it was decided to chose to adopt the approach of Thaler et al. (1978).

## 5.4 The distribution of perpendicular distances of pixels from a random line

In Subsection 5.3.4 it is assumed that any pixel which lies within a perpendicular distance of 1 from a random line is a mixed pixel (based on the estimated value of  $\tau$  in Table 5.2). Then, to estimate the mixed pixel density, (5.16), and hence the probability density function of pixel values in (5.17), it is assumed that these signed perpendicular distances are uniformly distributed, subject to  $D \in [-1, +1]$ . This latter assumption of uniformity is now justified.



It is assumed that a square image, of size  $2n \times 2n$  pixels, consists of two tissues and they are separated by a straight line of the form  $y = \lambda(x - c)$ . For a given  $\lambda$ , the line passes through the image with a random perpendicular distance from the origin. The perpendicular signed distance of a pixel at position  $(i, j)$ , for  $-n \leq i, j \leq n$ , in the image to the line is given by

$$D_{i,j} = \frac{\lambda(i - c) - j}{\sqrt{1 + \lambda^2}} \quad (5.19)$$

and the line intersects the  $x$  axis at  $x = c$ . Without loss of generality, the origin can be chosen such that  $c \in [-1/2, +1/2]$ , by shifting the  $y$  axis by  $m$ , the nearest integer to  $c$ . It can be assumed that the positioning of each sheep in the X-ray machine has sufficient variability to allow the assumption that the boundary is positioned randomly with respect to the pixel lattice. Therefore, it can be stated that

$$c \sim U\left(-\frac{1}{2}, +\frac{1}{2}\right).$$

Using this and (5.19), it follows that

$$D_{i,j} \sim U\left(\frac{\lambda(i - \frac{1}{2}) - j}{\sqrt{1 + \lambda^2}}, \frac{\lambda(i + \frac{1}{2}) - j}{\sqrt{1 + \lambda^2}}\right). \quad (5.20)$$

If only the points which lie on the  $x$  axis are considered, (i.e.  $j = 0$ ), and if  $i$  is chosen at random from the set  $-(n + m), -(n + m - 1), \dots, n - m$ , for  $n \geq 1$ , then by restricting the  $D_{i,j} \in [-1, +1]$  it can be seen that

$$D_{\cdot,0} \sim U(-1, +1), \quad (5.21)$$

where  $D_{\cdot,0}$  denotes that is the distances for all possible values of  $i$ , when  $j = 0$ . If the slope,  $\lambda$ , is assumed to satisfy  $|\lambda| \geq 1$ , then the same result applies for other values of

$$j \in (-(n - 1), -(n - 2), \dots, (n - 1)). \quad (5.22)$$

If the slope does not satisfy this condition then the line does not intersect all the rows of the image and hence (5.22) does not hold as there will be no pixels in those rows which lie within a perpendicular distance of 1 from the line. However, if  $|\lambda| < 1$ , simply exchange  $x$  and  $y$  to make this condition hold. The result does not hold for  $j = \pm n$  because at these points the edges of the lattice cause edge effects and the point on the line required to measure the perpendicular

distance falls outside the lattice. Therefore, if we choose  $j$  at random from the set  $(-(n-1), -(n-2), \dots, (n-1))$  then

$$D_{.,.} \sim U(-1, +1), \quad (5.23)$$

giving the required result.

## 5.5 Summary and conclusions

In this chapter, the point spread function (PSF) of the SAC-BioSS CT machine was estimated using a known edge between the cradle and air pixels in the image. The estimated, spatially invariant PSF was an isotropic bivariate normal, and this was used to derive the *mixed pixel density*. This was combined with the distributions for the pure pixels in order to form an estimated probability density function of the greyscale values in the segmented lumbar images.

This distribution was fitted by maximum likelihood to histograms of greyscale values of lumbar images and appeared to provide a good fit to the data. However, since the model included provisions for only two tissue types, and no true proportions are available for comparison, a simulation study is carried out. This allows us to compare the estimated proportions by using this newly proposed probability density function, with those obtained using the method currently in use and also with the true proportions of the tissues. This simulation study is presented in detail in the following chapter.

# Chapter 6

## Estimation of tissue proportions

### 6.1 Introduction

In Chapter 5, the probability density function of pixel values for segmented sheep images is estimated. It is noted that the boundaries between two tissues are reasonably smooth and continuous: we assume that they are locally linear and the boundaries can be approximated by parallel random lines. In Subsection 5.3.1, it is assumed that all the mixed pixels are a mixture of fat and muscle, and the probability density function in (5.17) makes provision only for these two tissues. However, this could be extended to deal with more tissue types.

In Section 6.2, the construction of random lines in the plane is discussed. In Section 6.3, simulations are performed in which random parallel lines represent boundaries between areas of fat and muscle, so as to approximate the portion of the histogram of segmented images that corresponds to fat and muscle (i.e. greyscale values between approximately 40 and 200 in Figure 5.1). Several methods are used to estimate the amounts of fat and muscle in each simulated image. Such methods include the method currently in use at the SAC-BioSS CT unit and the fitting to the histogram of pixel values of the probability density function proposed in Subsection 5.3.4. These results are compared with the true tissue amounts in each simulation.

The accuracy of the assumptions made in order to derive the probability density function of pixel values is assessed by examining sub-images of a typical segmented lumbar image, and these assumptions are modified in Subsection 6.6.1. The simulations are repeated to assess the effect on the root mean square error of the estimation methods: see Subsection 6.6.2. Spatial information around each pixel is also considered in order to classify the pure pixels using certain thresholds on

the local variances. From this, the amount of fat present in the mixed pixels is estimated: see Section 6.7.

## 6.2 Random processes of geometrical objects

Random point patterns (or point processes) have an important role in stochastic geometry. They may be used in analysing random patterns of geometric objects through using suitable representation spaces, where a point from the representation system represents a particular geometric object in the original pattern. An example of this, described in Stoyan et al. (1995), is a random pattern of lines which can be viewed as equivalent to a point process, with the corresponding points lying on a cylinder in  $\mathbb{R}^3$ . A line process is a random collection of lines in the plane which is locally finite, where only finitely many lines hit each compact planar set.

### 6.2.1 Representation space for directed lines in the plane

Stoyan et al. (1995) define a directed line as a line together with a preferred direction along that line. They let  $F(2, 1)$  denote the family of all undirected lines and  $F^*(2, 1)$  denote the directed lines in the plane. There is an obvious 2:1 correspondence between the elements of  $F^*(2, 1)$  and  $F(2, 1)$  which is obtained by ignoring the direction of the line. The representation space for directed lines is described below: and the results transfer easily to the undirected case.

All directed lines in the plane can be put in a 1:1 correspondence with the set of points on a cylinder in  $\mathbb{R}^3$ . A convenient set of coordinates for the directed line,  $l$ , in the plane is based on its perpendicular distance from the origin,  $o$ , and the angle which it makes with the  $x$  axis. The signed perpendicular distance,  $p$ , of the line,  $l$ , from  $o$  is denoted positive if the origin lies to the left of the line and negative if it lies to the right of the line  $l$ . The angle between  $l$  and the  $x$  axis is denoted by  $\alpha$  and is measured in an anti-clockwise direction. See Figure 6.1 for an illustration of this coordinate system.

The cylinder in  $\mathbb{R}^3$  corresponding to  $F^*(2, 1)$  is denoted by  $C^*$  and defined by

$$C^* = \{(\cos \alpha, \sin \alpha, p) \in \mathbb{R}^3 : p \in \mathbb{R}, \alpha \in (0, 2\pi]\}.$$

Each member of  $F(2, 1)$ , the family of undirected lines in the plane, corresponds to a pair of directed lines, and hence a pair of points in  $C^*$ . These two points are

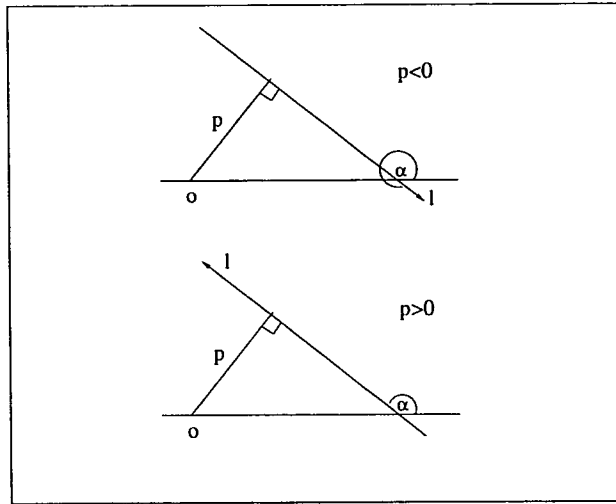


Figure 6.1: Taken from Stoyan et al. (1995, p245), this shows the construction of the representation space for directed lines in the plane. The upper diagram shows the case when  $p$  is negative and the lower shows  $p$  as positive.

reflections of each other in the origin. To represent an undirected line only one of these points is selected. The point which is selected is the one which is lying in the half cylinder which lies to one side of a fixed plane and which includes the cylinder axis. Therefore, the corresponding representation space for  $F(2, 1)$  is denoted by

$$C = \{(\cos \alpha, \sin \alpha, p) \in \mathbb{R}^3 : p \in \mathbb{R}, \alpha \in (0, \pi]\}.$$

It is worth noting that as the angle of the undirected line passes through 0, the point in  $C$  jumps from one edge of the half cylinder to the other, while the value of the  $p$ -coordinate is multiplied by  $-1$ .

Stoyan et al. (1995) further discuss the symmetries of  $C^*$ . The representation spaces described above depend on the particular choice of the origin,  $o$ , and the  $x$  axis for the plane, and therefore the stochastic geometry must account for these choices, since they correspond to important symmetries. To complete the discussion on these line processes, Stoyan et al. (1995) include a study of the transformations induced on the representation cylinder  $C^*$  by translations and rotations of the plane. The effect of choosing different origins and  $x$  axes for a square lying in the plane is discussed later.

Stoyan et al. (1995) formally describe a directed line process as equivalent to a random subset of the representation space,  $C^*$ . Therefore, to select a random line that passes through an image is the same as selecting a random point from the corresponding possible subset of  $C^*$ , where the random point has a uniform distribution over this subset. Point processes such as these can be considered as

Range of $\alpha$		Range of $p$ , given $\alpha$	
$0$	$\leq \alpha \leq \pi/2$	$\sqrt{2} \sin(-\pi/4 - \alpha)$	$\leq p \leq \sqrt{2} \sin(\pi/4 + \alpha)$
$\pi/2$	$\leq \alpha \leq \pi$	$\sqrt{2} \sin(\pi/4 - \alpha)$	$\leq p \leq \sqrt{2} \sin(\alpha - \pi/4)$
$\pi$	$\leq \alpha \leq 3\pi/2$	$\sqrt{2} \sin(3\pi/4 - \alpha)$	$\leq p \leq \sqrt{2} \sin(\alpha - 3\pi/4)$
$3\pi/2$	$\leq \alpha \leq 2\pi$	$\sqrt{2} \sin(-3\pi/4 - \alpha)$	$\leq p \leq \sqrt{2} \sin(\alpha + 3\pi/4)$

Table 6.1: For directed lines, the ranges of  $p$ , the signed perpendicular distances from  $o$ , which ensure that the lines pass through a square of side 2, with the origin positioned at the centre of the square.

particular cases of point processes in  $\mathbb{R}^2$ , because using the suggested parametrisation of Stoyan et al. (1995) i.e.,  $(p, \alpha)$  of  $C^*$ , the cylinder can be cut and embedded as a subset  $\mathbb{R} \times (0, 2\pi]$  of  $\mathbb{R}^2$ .

### Subset of $C^*$ corresponding to a square

As stated in Section 6.1, the objective is to perform simulations which approximately reproduce the histograms of segmented sheep images by simulating random parallel lines to represent the boundaries between fat and muscle. Using a square grid to approximate the sheep images, it is possible to find the corresponding subset of  $C^*$ , and hence the subset of  $\mathbb{R}^2$ , which corresponds to lines which pass through these square grids. These subsets depend on the choice of  $o$  and the  $x$  axis and we have chosen the origin to be at the centre of the image. Table 6.1 shows the range of  $p$ 's, for the given ranges of  $\alpha$ , which ensures that the line passes through the square of side 2. It can be seen from these results that the subset of  $\mathbb{R}^2$  is symmetric about  $p = 0$  when the origin is in the centre of the image.

## 6.3 Simulation study

The main objective of the simulation study is to simulate images whose histograms of pixel values approximate those of the segmented lumbar images, lying in the greyscale range of between approximately 40 and 200. Then, for these simulated histograms, the proportion of fat and muscle is estimated using various methods, which include the method currently in use at the SAC-BioSS CT unit (see Subsection 6.4.1) and the proposed probability density function of pixel values which is introduced in (5.17).

The greyscale values in the range 40 to 200, from the segmented lumbar images

are used to estimate the means and variances of pixel values corresponding to pure fat and pure muscle in each of the 24 training images, which are described in Chapter 1. The average value, over the 24 images, of each estimate is calculated and rounded, giving the estimates of the means and variances for pure fat and muscle respectively as

$$\mu_f = 65, \sigma_f^2 = 150, \mu_m = 150, \sigma_m^2 = 250. \quad (6.1)$$

It was decided to simulate images of size  $60 \times 60$ , which provides a close approximation to the total number of pixels in the histograms of the sheep carcass which lie in the greyscale range 40 to 200. The range of this total number of pixels in the 24 lumbar images is [3297, 5897].

A random line, representing a boundary between fat and muscle, is simulated using the representation described in Section 6.2. The origin is taken to be at the centre of the simulated image, and for a given  $\alpha$ ,  $p$  is selected, using the range defined in Table 6.1, to ensure the line passes through the square. Due to the symmetry of the chosen square, it is only necessary to consider  $\alpha$  in the range  $[0, \pi/4]$  and positive  $p$ . To select a pair  $(\alpha, p)$  at random from this specified region for a square of size  $60 \times 60$ , simulate  $\alpha \sim U(0, \pi/4)$  and  $p \sim U(0, 30\sqrt{2})$ . For the given  $\alpha$ , if  $p \in (0, 30\sqrt{2} \sin(\pi/4 + \alpha))$ , then this  $(\alpha, p)$  pair corresponds to a line which passes through the square. Otherwise, repeat the simulations of  $\alpha$  and  $p$  until a suitable  $(\alpha, p)$  pair is selected. All other possible images with one simulated line passing through the square, for other values of  $\alpha$  and  $p$  may be obtained by reflection and/or rotation of the image obtained with  $\alpha \in [0, \pi/4]$  and  $p \in [0, 30\sqrt{2}]$ . It is reasonable to use these restrictions on  $\alpha$ , and hence  $p$ , since it is only the overall histogram shape that is of interest in the simulation, rather than the individual pixel values or their spatial position within the image. This also reduces the computations required to calculate the overall true amount of each tissue in the simulated images.

For a simulation containing a single random line ( $\alpha \in [0, \pi/4]$  and  $p$  positive), and given the estimated PSF (see Subsection 5.3.2), pixels more than a perpendicular distance of one from the line and lying to the left and right are represented as pure muscle and fat respectively. This is chosen only to ensure that there is more muscle than fat in the simulated image. The remainder of the pixels are represented as mixed pixels. Depending on the position of this random line in the image, there can be at most  $2\sqrt{2}/60\%$  or 4.7% mixed pixels within the image. This is much less than the actual percentage of mixed pixels present in the histograms of the sheep images, which is approximately 20-30%. This value is estimated from



fitting (5.17) to a typical histogram of greyscale values from the segmented sheep images, such as Figure 5.7. In addition, the proportions of the pure tissues in these simulations do not accurately reflect those in the sheep images. Therefore, in order to increase the percentage of mixed pixels it is possible to either reduce the size of the simulated images, or to increase the number of lines in the  $60 \times 60$  grid. The first proposal is undesirable as the number of pixels in the increased grid would no longer accurately reflect the truth; therefore, the second method is used. A set of seven randomly positioned parallel lines is simulated which pass through the image to represent alternate layers of fat and muscle tissue.

Based on the earlier assumption that there is a negligible number of pixels affected by more than one boundary, and due to the estimated variance of the PSF, the parallel lines must be at least two units apart. Therefore, to ensure the seven lines pass through the square, with neither the first nor the last line lying too close to the bottom right or top left corners (hence reducing the proportion of mixed pixels), we take  $p_1 \sim U(20, 30)$ , where  $p_1$  is the perpendicular distance of the first line from the origin. The choice of this upper limit is based on the feasible range of  $p$ , given  $\alpha$  (see row one of Table 6.1). Each remaining line is then chosen to be within a perpendicular distance  $d_i$  from the previous line, with  $d_i \sim U(2, 9)$ , for  $i = 1, \dots, 6$ . Therefore, this gives

$$p_{i+1} = p_i - d_i \quad (i = 1, \dots, 6),$$

where  $d_i$  is the perpendicular distance between two consecutive lines at distances  $p_i$  and  $p_{i+1}$  from the origin. The eight regions bounded by these lines and the sides of the square represent alternately fat and muscle tissue, starting in the bottom right hand corner. Each pixel's greyscale value, depending on its position from a line and the region in which it is lying, is generated from either the pure fat, pure muscle or mixed pixel probability (5.16) density functions, using (6.1). Six examples of these simulations are shown in Figure 6.2 and the proportions of fat and muscle in these simulations are shown Table 6.2.

The corresponding histograms are constructed with the greyscale values rounded to the nearest integer: see Figure 6.3. It can be seen from these histograms that they accurately reflect the actual histograms, shown in Figure 5.1, from the sheep images as the greyscale values are in the appropriate range and the proportion of muscle is larger than that of fat, which is always the case for the lumbar images.

The true tissue areas in each simulated image are calculated fairly easily using standard formulae from coordinate geometry. Having restricted  $\alpha$  to the interval  $(0, \pi/4)$  and  $p_1$  to be positive, there are only three possible ways each of the seven

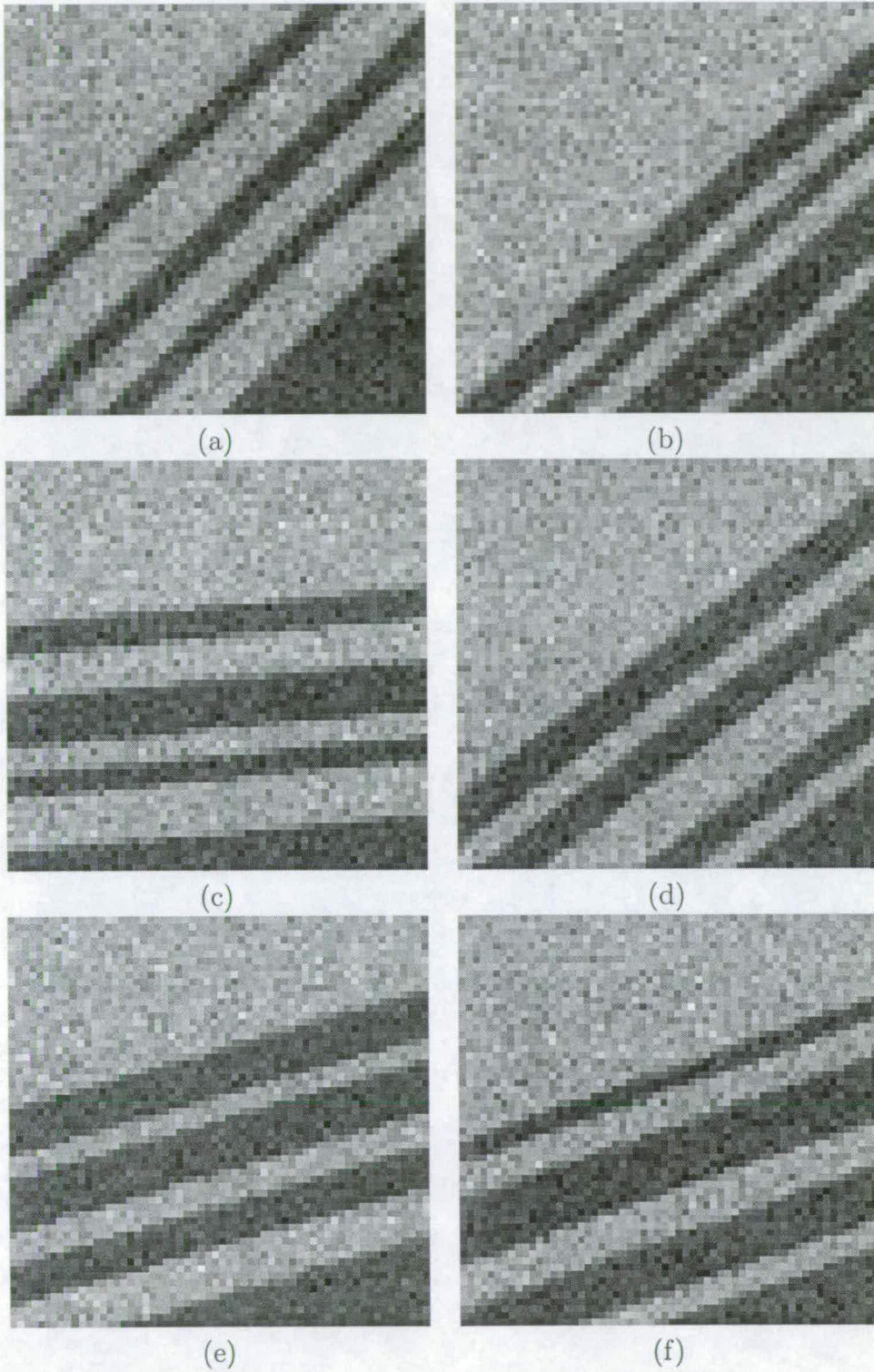


Figure 6.2: Six examples of the simulations from the proposed probability density function of pixel values in (5.17). As with the true sheep images, fat and muscle pixels are displayed as dark grey and light grey respectively.

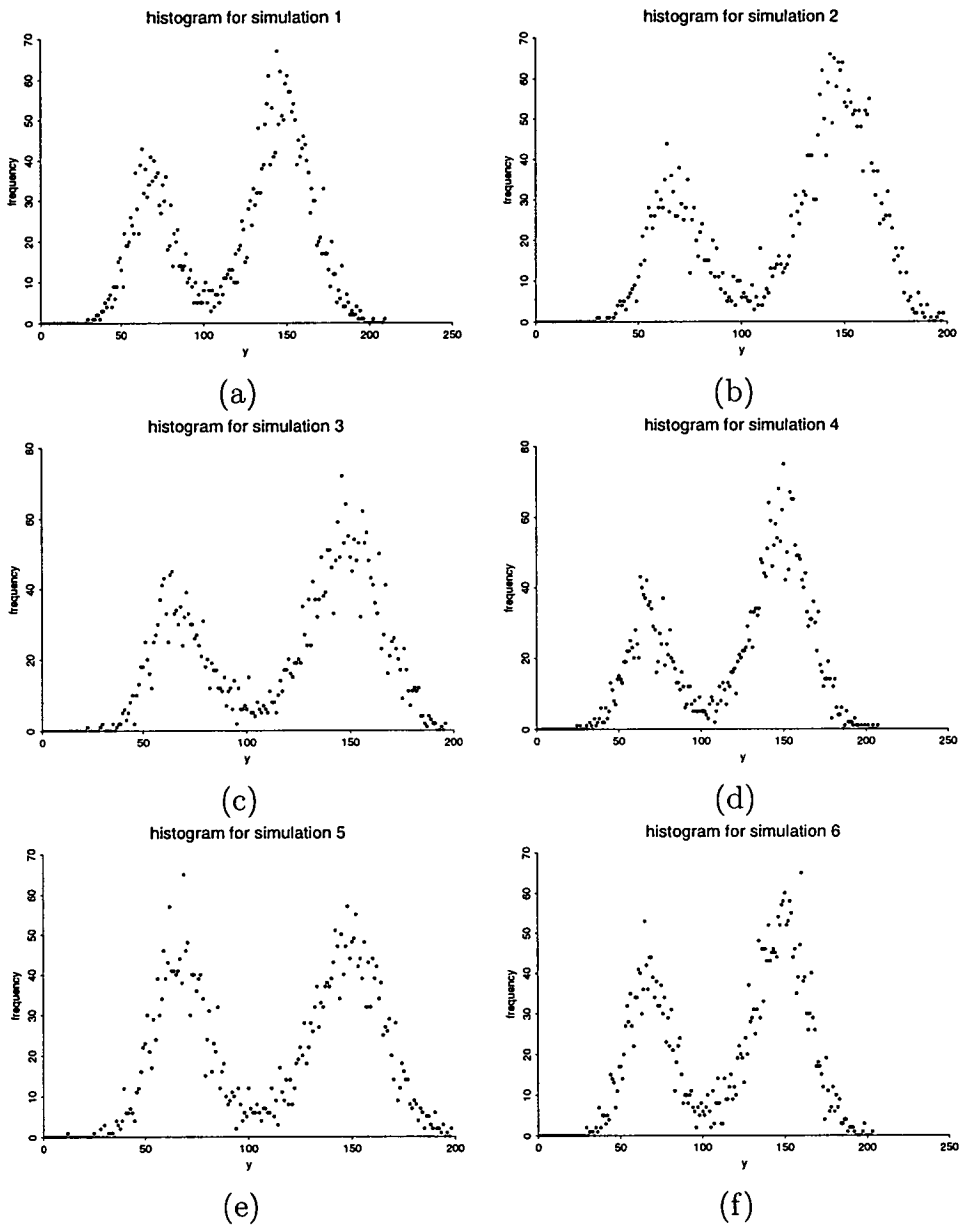


Figure 6.3: *Histograms of the greyscale values for the six images shown in Figure 6.2.*

Simulation number	Proportions of fat	muscle
1	0.22	0.53
2	0.20	0.58
3	0.23	0.53
4	0.21	0.57
5	0.32	0.45
6	0.28	0.50

Table 6.2: *Proportions of pure fat and pure muscle generated in the simulated images.*

lines may pass through the image:

- passing through the right side and the bottom of the image,
- passing through the right and left sides,
- passing through the left side and the top of the image.

## 6.4 Methods used for estimating tissue proportions

This section describes four methods used to estimate the amount of each tissue type. The results are compared using the root mean square error of the estimated amount of fat over all the simulations in Section 6.5. These methods are divided into two groups: two threshold methods which exclude provision for mixed pixels within the model and two methods which allow for mixed pixels. For the first three methods, two sets of results are obtained, depending on whether the four parameters,  $\mu_f$ ,  $\sigma_f^2$ ,  $\mu_m$  and  $\sigma_m^2$  are either assumed to be known or are required to be estimated.

### 6.4.1 Threshold methods

#### Method 1: Threshold at the average of the two tissue means

In this method it is assumed that the distribution of pixel values is a mixture of two normal distributions,  $N(\mu_f, \sigma_f^2)$  for fat with proportion  $\pi_f$  and  $N(\mu_m, \sigma_m^2)$  for muscle, with proportion  $\pi_m$ , such that  $\pi_f + \pi_m = 1$ . No provision is made for mixed pixels in this method. The method employed here is quite simply taking the average of the two means as a threshold for the image. If the means are

assumed known then the threshold is denoted by  $t_1$ , where

$$t_1 = \frac{\mu_f + \mu_m}{2}.$$

Using this threshold is the method that is currently in use at the SAC-BioSS CT unit. If the means are estimated then the threshold used to estimate the amount of fat is denoted by  $\hat{t}_1$ . Therefore, any pixel with greyscale value  $\leq t_1$  (or  $\hat{t}_1$ ) is considered to be fat, and any pixel with greyscale  $> t_1$  (or  $\hat{t}_1$ ) is considered to be muscle.

In the case of known means, the same threshold is applied to each histogram, without considering the individual counts of greyscale values. Obviously, this is not an ideal method of estimation as the means and variances of the pure tissues vary slightly between sheep, depending on for example, the obesity of the sheep, or how long it has been without food or water before scanning.

However, when the means and variances are assumed to be unknown they are estimated by maximum likelihood, where the individual counts in each histogram need to be considered. We denote the probability density function of pixel values by  $h(y; \theta)$ , (which is a mixture of two normal distributions), where  $\theta = (\mu_f, \sigma_f^2, \pi_f, \mu_m, \sigma_m^2)$ . Therefore, the likelihood,  $L$ , is given by

$$L(\theta; \mathbf{y}) = \prod_{\text{all pixels}} h(y; \theta).$$

Hence, for each histogram the likelihood is given by

$$L(\theta; \mathbf{y}) = \prod_j h(y_j; \theta)^{n_j},$$

where the product is taken over the range of simulated greyscale values,  $j$ , and  $n_j$  the frequency with which the greyscale values occur in an individual histogram. Taking logs gives a log likelihood  $l$ , such that

$$l(\theta; \mathbf{y}) = \sum_j n_j \log h(y_j; \theta). \quad (6.2)$$

This is maximised using numerical optimisation routine E04JAF (Numerical Algorithms Group, 1993). Although the estimated proportion of fat,  $\hat{\pi}_f$ , (found by maximum likelihood), could be used directly to estimate the total amount of fat (using  $60^2 \times \hat{\pi}_f$ ), the threshold value  $\hat{t}_1$  is used in preference to provide comparison with results obtained using  $t_1$ .

## Method 2: Threshold at the point the two normal probability density functions cross

Again assuming a mixture of two normal distributions as the probability density function of pixel values, another proposed method for estimating the amount of fat is to threshold the image at the greyscale value at which the two normal probability density functions cross. This is almost equivalent to quadratic discrimination with equal prior probabilities (see McLachlan, 1992). If the means and variances of the pixel values for fat and muscle tissues are assumed known then the points of intersection occur when

$$\frac{1}{\sigma_f} \exp \left\{ \frac{-(y - \mu_f)^2}{2\sigma_f^2} \right\} = \frac{1}{\sigma_m} \exp \left\{ \frac{-(y - \mu_m)^2}{2\sigma_m^2} \right\}.$$

After rearrangement and simplification, the points at which these two densities meet are given by

$$t_2 = \frac{(\sigma_f^2 \mu_m - \sigma_m^2 \mu_f) \pm \sigma_f \sigma_m \sqrt{(\mu_f - \mu_m)^2 + 2(\sigma_f^2 - \sigma_m^2) \log(\sigma_f / \sigma_m)}}{\sigma_f^2 - \sigma_m^2}.$$

Therefore the threshold,  $t_2$ , is chosen such that  $\mu_f < t_2 < \mu_m$ . In these simulations, the required value of  $t_2$  is found using the  $-$  from this  $\pm$ . If  $\sigma_f^2 = \sigma_m^2$  then  $t_2 = t_1$ . If the means and variances are assumed unknown, then the estimates  $\hat{\mu}_f$ ,  $\hat{\mu}_m$ ,  $\hat{\sigma}_f^2$ , and  $\hat{\sigma}_m^2$  obtained by maximum likelihood in method 1 are used to calculate the threshold value: this will be denoted  $\hat{t}_2$ . As before, an estimate of the total amount of fat is the number of pixels with greyscale values  $\leq t_2$  (or  $\hat{t}_2$ ).

### 6.4.2 Mixed pixel methods

#### Method 3: Fitting probability density function given by (5.17) by maximum likelihood

This method of estimation is the first we have considered to make provision for mixed pixels: the probability density function of pixel values in the image is assumed to be given by (5.17). The vector of parameters to be estimated is denoted by  $\theta_K$  or  $\theta_U$ , depending on whether the means and variances are known or unknown respectively. Therefore,  $\theta_K = (\pi_f, \pi_m)$ , and  $\theta_U = (\mu_f, \sigma_f^2, \pi_f, \mu_m, \sigma_m^2, \pi_m)$ , where the  $\pi_f$  and  $\pi_m$  are the proportions of the pure tissues. As in method 1,  $\theta$ , for the means and variance of the pure tissues either known or unknown, is estimated by maximum likelihood. By analogy with (6.2) the log likelihood is

given by

$$l(\mathbf{y}|\boldsymbol{\theta}) = \sum_j n_j \log g(y_j, \boldsymbol{\theta}),$$

where  $g(y_j, \boldsymbol{\theta})$  is the probability density function given in (5.17). The log likelihood, constrained by  $\pi_f \geq 0$ ,  $\pi_m \geq 0$  and  $0 \leq \hat{\pi}_f + \hat{\pi}_m \leq 1$ , is maximised using numerical optimisation routine E04JAF (Numerical Algorithms Group, 1993), thereby providing an estimate of  $\boldsymbol{\theta}$ . Using the estimates for the two proportion parameters,  $\hat{\pi}_f$  and  $\hat{\pi}_m$ , an estimate of the total amount of fat in the simulated image is obtained by adding the number of pure fat pixels to half the number of mixed pixels (see 5.18). Therefore, an estimate of the total amount of fat is given by

$$60^2 \times \left( \hat{\pi}_f + \frac{1 - \hat{\pi}_f - \hat{\pi}_m}{2} \right). \quad (6.3)$$

#### Method 4: Finding the first moment of the greyscale values of pixels in the image

It is known from (5.13), that

$$y|\rho \sim N(\rho\mu_f + (1 - \rho)\mu_m, \rho\sigma_f^2 + (1 - \rho)\sigma_m^2).$$

Therefore, the expectation,  $E(y)$ , of the greyscale values in an image is given by

$$\begin{aligned} E(y) &= E[E(y|\rho)] \\ &= \mu_f E(\rho) + \mu_m (1 - E(\rho)) \\ &= \mu_m + (\mu_f - \mu_m)E(\rho). \end{aligned} \quad (6.4)$$

Rearranging (6.4), an unbiased estimate of the proportion of fat in the image is given by

$$\hat{\rho} = \frac{\widehat{E(y)} - \mu_m}{\mu_m - \mu_f}, \quad (6.5)$$

where  $\widehat{E(y)}$  is estimated by  $\bar{y}$ , the average of all pixel values in the image. Hence, an unbiased estimate of the total amount of fat in the image is given by

$$3600 \times \left[ \frac{\mu_m - \widehat{E(y)}}{\mu_m - \mu_f} \right]. \quad (6.6)$$



## 6.5 Results and discussion

100 simulations are performed, by varying  $\alpha$ ,  $p_1$  and the  $d_i$  independently and for each simulation an estimate of the amount of fat is found using each of the four methods. As stated earlier, the distributions of greyscale values for fat and muscle pixels are  $N(\mu_f, \sigma_f^2)$  and  $N(\mu_m, \sigma_m^2)$ . Throughout all the simulations the means are fixed at  $\mu_f = 65$  and  $\mu_m = 150$ , (see Section 6.3), but  $\sigma_f^2$  and  $\sigma_m^2$  are varied for each set of 100 simulations. The true variances for the fat and muscle tissues are estimated from the 24 training images to be  $\sigma_f^2 = 150$  and  $\sigma_m^2 = 250$ . However, to check the sensitivity of the results to the particular choices of parameters, 100 simulations were repeated for eight other pairs of values of  $\sigma_f^2$  and  $\sigma_m^2$ , without varying  $\mu_f$  and  $\mu_m$ . The other pairs of variances were chosen to be centred around the estimated values.

The performance of the methods is compared on the basis of the root average mean square error of the estimated fat area, averaged over the 100 simulations, denoted by RMSE. Let  $\theta_k$  and  $T_k$  denote the true amount and estimated amounts of fat for simulation  $k$ , for  $k = 1, \dots, 100$ . In each simulation the amount of fat varies, so strictly speaking it is not the standard mean square error that is used for comparing the methods. Hence, for each of the four methods the average mean squared error (MSE), is denoted by

$$MSE = \frac{1}{100} \sum_{k=1}^{100} (T_k - \theta_k)^2,$$

which can be rearranged as

$$= (\bar{T} - \bar{\theta})^2 + \frac{1}{100} \sum_{k=1}^{100} (T_k - \bar{T} - (\theta_k - \bar{\theta}))^2 = \text{bias}^2 + \text{variance}.$$

These RMSEs of estimated fat areas from 100 simulations, for each estimation method, are shown in Table 6.3. As in Section 6.4, the results are divided according to whether the method is a threshold or mixed-pixel method, and whether the moments (the means and variances) of the pure tissues are assumed known or are estimated. To distinguish these results in the threshold methods, the threshold used will always be stated. In the case of Method 3, the results will be denoted MLK and MLU, depending on whether the moments of the pure tissues are known or unknown respectively.

It can be seen from Table 6.3 that the ML method out-perform both of the threshold methods, irrespective of whether the parameters are known or estimated. This

$\sigma_f^2$	$\sigma_m^2$	moments known				moments unknown			
		threshold		mixed		threshold		mixed	
		$t_1$	$t_2$	MLK	Moment	$\hat{t}_1$	$\hat{t}_2$	MLU	
75	125	6.0	25.2	4.7	6.5	9.5	51.4	5.0	
	250	13.5	57.0	5.4	8.1	8.5	74.6	8.1	
	500	51.4	79.7	7.3	10.5	27.0	95.0	18.6	
150	125	5.8	11.2	5.2	8.0	8.4	13.4	5.5	
	250	10.9	23.5	5.8	9.3	9.4	38.3	8.5	
	500	48.6	47.6	7.7	11.5	30.3	64.5	22.3	
300	125	21.1	39.0	6.6	10.3	19.6	33.9	8.7	
	250	13.3	9.4	7.1	11.3	18.8	18.5	9.9	
	500	30.9	26.4	8.9	13.1	22.1	46.1	24.9	

Table 6.3: *Root mean square errors of the estimated amount of fat, by applying the four methods to the data from 100 simulations for a range of values of  $\sigma_f^2$  and  $\sigma_m^2$ .*

is hardly surprising since the probability density function for this method was assumed to be the same as that from which the simulated values were drawn. The results also show that Method 3, MLK has estimated fat to within 5-9 pixels, which is approximately 0.2% of the total pixels.

Method 1, with threshold  $t_1$ , is currently being applied at the SAC-BioSS CT unit in order to estimate the overall amount of fat present in each segmented sheep image. Adoption of the proposed probability density function of pixel values to estimate the overall amount of fat ought to provide substantially improved performance for all pairs of variances, in particular for  $\sigma_f^2 = 150$  and  $\sigma_m^2 = 250$ , the estimated true variances of fat and muscle tissue. In reality the true means and variances vary from sheep to sheep, see Subsection 6.4.1, and are therefore unknown. Therefore, the performance of Method 1, with threshold  $t_1$ , is directly compared with Method 3 MLU, rather than Method 3 MLK, because if this mixed pixel method was adopted it would be assumed that the moments of the pure tissues are unknown.

On average using the RMSE scale, Method 1 with threshold  $t_1$  estimates fat to within approximately 28 pixels, which is approximately 0.8% of the total image. Although this seems a very small amount, this method is more sensitive to large changes in the variances than Method 3 MLU, in particular when  $\sigma_m^2 = 500$ . In comparison, Method 3 MLU estimates fat to within 14 pixels (0.4% of the total image).

Figure 6.4 shows the fitted probability density functions of the pixel values for

$\sigma_m^2$	$\sigma_f^2$		
	75	150	300
125	102.4	109.3	115.7
250	96.0	102.7	109.1
500	90.8	97.0	103.3

Table 6.4: Thresholds,  $t_2$ , for Method 2 for the pairs of variances  $(\sigma_f^2, \sigma_m^2)$ . These results can be compared with the threshold,  $t_1$ , used in Method 1 which is 107.5

Method 1, threshold  $t_1$  (solid line), and Method 3 MLU (dotted line), superimposed on the histograms shown in Figure 6.3. It can be seen that the fitted histogram for Method 3, MLU has a much better fit to the pixels lying between the two peaks, whereas the fit in this region is very poor with Method 1, threshold  $t_1$  (which makes no allowance for mixed pixels).

Figure 6.5 shows the true (solid line) with the estimated (dotted line) probability density functions of pixel values (using Method 3 MLU) superimposed on the six histograms shown in Figure 6.3. It can be seen from these that there is little difference between the estimated fit and the true probability density function.

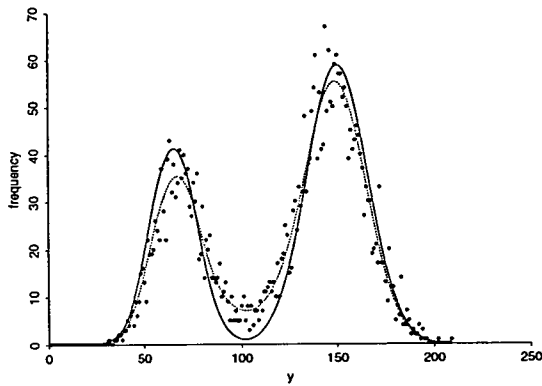
Method 2 ( $t_2$  and  $\hat{t}_2$ ) consistently performs worse than Method 1 ( $t_1$  and  $\hat{t}_1$  respectively), the other threshold method. This is due to the fact that the thresholds for Method 1 are unaffected by the choice of the variances for the two tissues (i.e. fat and muscle), unlike the thresholds chosen for Method 2. Table 6.4 shows the thresholds,  $t_2$ , for Method 2, for the different pairs of variances. It can be seen that when the variances are approximately equal, e.g.  $\sigma_f^2 = 300$  and  $\sigma_m^2 = 250$ , or  $\sigma_f^2 = 150$  and  $\sigma_m^2 = 125$ , then the RMSE of Method 2, with threshold  $t_2$ , is closer to that of Method 1, with threshold  $t_1$ , which uses a constant threshold of 107.5. Method 2, with threshold  $t_2$ , performs surprisingly well when  $\sigma_f^2 = 300$  and  $\sigma_m^2 = 250$  in comparison to the other methods.

The RMSE for the moment method is fairly consistent over the range of variances for the two tissues and this method is in general more accurate at estimating the amount of fat than either of the threshold methods.

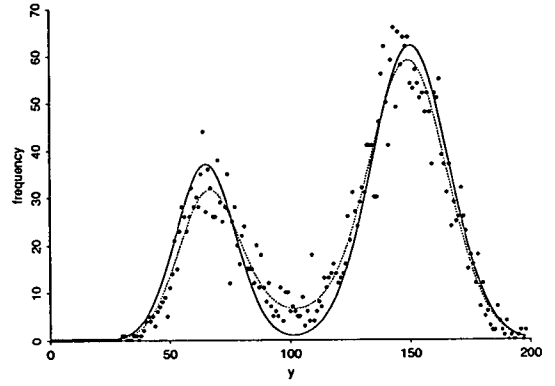
The MSE was split into its bias and variance components for each of the four methods of estimation, again for the means and variances of the pure tissues either known or estimated. Rather than showing the results for all nine sets of variances, three sets have been selected. These are

$$\sigma_f^2 = 75, \sigma_m^2 = 125; \quad \sigma_f^2 = 150, \sigma_m^2 = 250; \quad \sigma_f^2 = 300, \sigma_m^2 = 500.$$

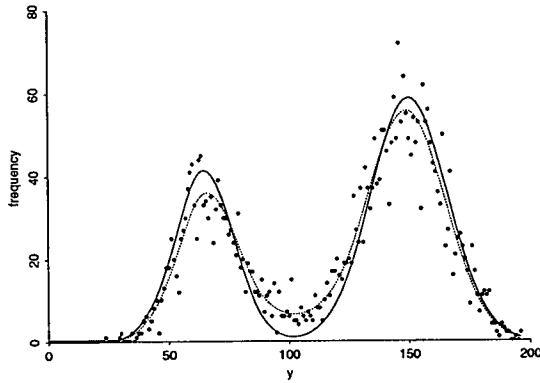
These results can be seen in Tables 6.5 (a), (b) and (c). The biases for the



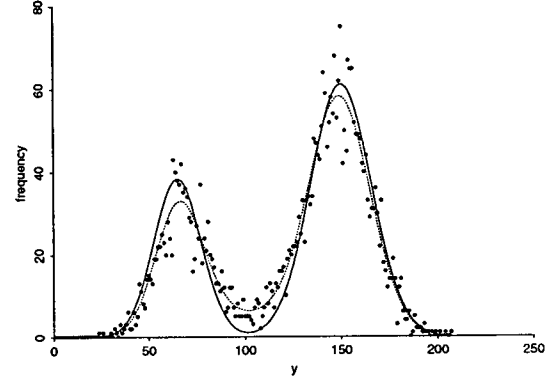
(a)



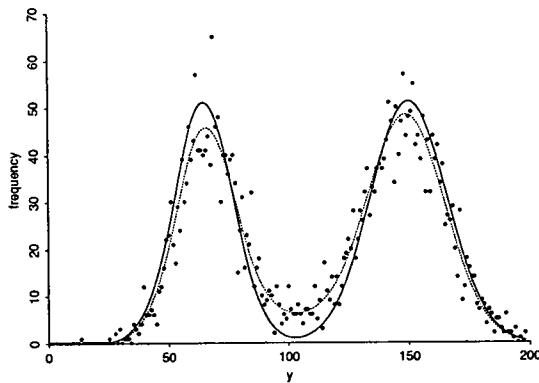
(b)



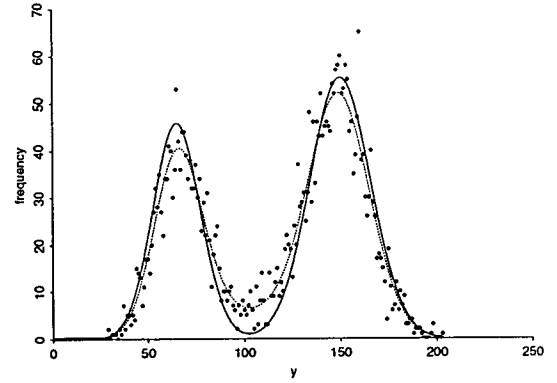
(c)



(d)

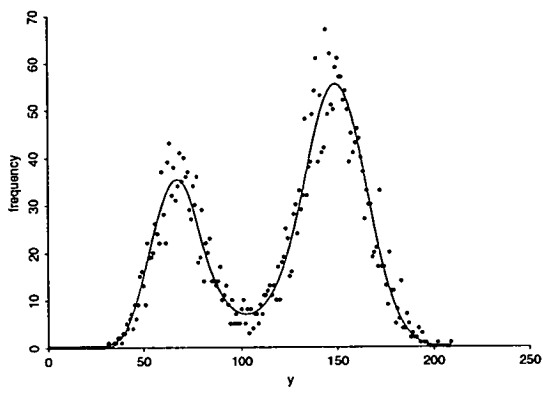


(e)

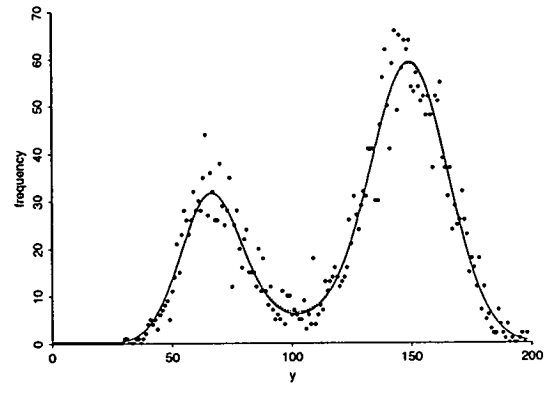


(f)

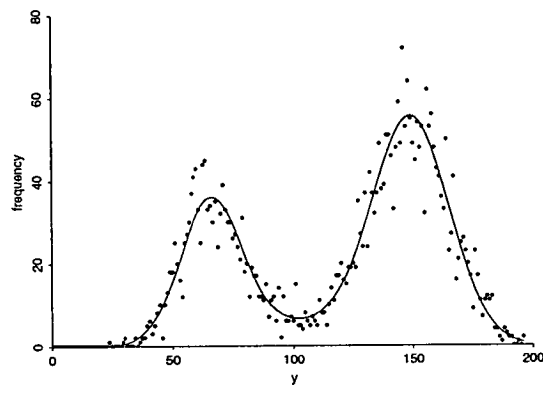
Figure 6.4: The six simulated histograms shown in Figure 6.3 with the estimated probability density functions of pixel values for Method 1 (—)(with only  $\pi_f$  estimated) and Method 3, MLU (- - -) superimposed.



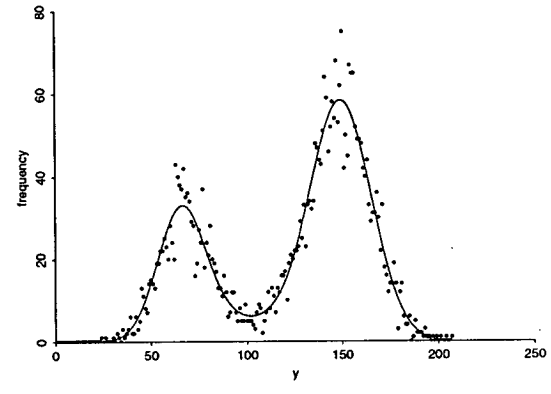
(a)



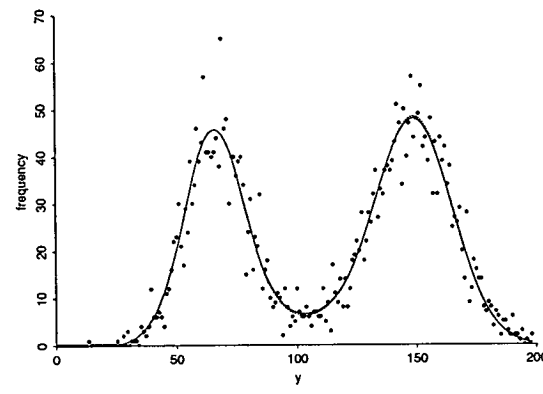
(b)



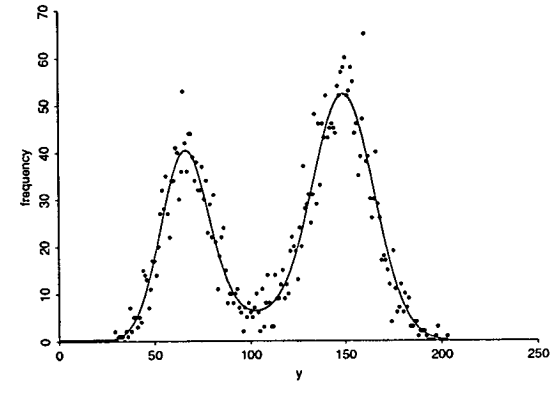
(c)



(d)



(e)



(f)

Figure 6.5: *The six simulated histograms shown in Figure 6.3, with the true (—) and estimated (---) probability density function of pixel values (using Method 3, MLU) superimposed.*

$$\sigma_f^2 = 75; \sigma_m^2 = 125$$

component	moments known				moments unknown		
	threshold		mixed		threshold	mixed	
	$t_1$	$t_2$	MLK	Moment	$\hat{t}_1$	$\hat{t}_2$	MLU
bias	2.7	-24.6	1.0	0.6	-6.6	48.9	1.0
variance	29.1	31.8	21.5	42.4	46.7	249.0	24.4
RMSE	6.0	25.2	4.7	6.5	9.5	51.4	5.0

$$\sigma_f^2 = 150; \sigma_m^2 = 250$$

component	moments known				moments unknown		
	threshold		mixed		threshold	mixed	
	$t_1$	$t_2$	MLK	Moment	$\hat{t}_1$	$\hat{t}_2$	MLU
bias	8.3	-22.6	1.2	0.9	-0.8	-35.9	1.2
variance	50.3	45.6	32.1	86.3	89.3	173.5	71.4
RMSE	10.9	23.5	5.8	9.3	9.4	38.3	8.5

$$\sigma_f^2 = 300; \sigma_m^2 = 500$$

component	moments known				moments unknown		
	threshold		mixed		threshold	mixed	
	$t_1$	$t_2$	MLK	Moment	$\hat{t}_1$	$\hat{t}_2$	MLU
bias	27.8	-22.8	1.4	1.3	10.3	-38.0	2.9
variance	183.7	176.7	77.1	171.1	384.1	685.4	613.9
RMSE	30.9	26.4	8.9	13.1	22.1	46.1	24.9

Table 6.5: *RMSE for each of the four methods, with the bias and variance components displayed for three pairs of variances.*

maximum likelihood methods all have the same sign, i.e. they tend to overestimate the fat in each image due to the same 100 lines being simulated over each set of pure tissue variances.

Since the maximum likelihood estimators, (MLK and MKU) are asymptotically unbiased and the moment estimator unbiased, their relative merit may be expressed as the ratio of their RMSE's. The efficiency of an estimator, say  $T_2$ , relative to another estimator of the parameter, say  $T_1$ , is defined as

$$\text{relative efficiency} = \frac{\text{RMSE}(T_1)}{\text{RMSE}(T_2)},$$

where  $T_2$  is less efficient than  $T_1$  if efficiency  $< 100\%$ .

In estimating the overall amount of fat using the moment estimator, it is assumed that the expected pixel values for the tissues are known and therefore, to provide direct comparison with the ML estimator, we shall estimate the efficiency of the moment estimator relative to the ML estimator under the same assumptions

$\sigma_m^2$	$\sigma_f^2$		
	75	150	300
125	72	65	64
250	67	62	63
500	60	67	68

Table 6.6: *The relative efficiencies of the moment estimator relative to the maximum likelihood estimator, MLK.*

i.e. Method 3, MLK. Table 6.6 shows the estimated efficiency of the moment estimator relative to the maximum likelihood estimator, MLK. It can be seen that the moment estimator is approximately 35% less efficient than the ML estimator. The higher efficiency of the maximum likelihood estimators is to be expected because they have asymptotic variance no greater than that of any other unbiased estimator. However, these results are included more for comparison with later results, see Table 6.13, after the moment method has been modified.

The estimates of the proportion of fat by both the maximum likelihood and moment methods are found by assuming that the probability density function of pixel values is given by (5.17). The assumptions that are made in order to estimate this probability density function are assessed by examining a typical segmented lumbar image in detail. This is presented in Section 6.6.

## 6.6 Modelling thin layers of tissue

The assumptions that are made in Chapter 5 in order to derive the probability density function of pixel values are now summarised:

- Given the estimated point spread function (isotropic bivariate normal with standard deviation  $\tau = 0.41$ ), a pixel lying within a perpendicular distance of one from the boundary between fat and muscle is defined to be a mixed pixel.
- The perpendicular distances,  $D$ , of the pixels to the tissue boundary within this region of mixed pixels are distributed as  $U(-1, +1)$ .
- Given that a pixel is classified as a mixed pixel, the mixing proportion,  $\rho$ , is affected by at most one boundary, which requires that the boundaries between the tissues are at least two units apart. Under this assumption, the proportion of fat in a mixed pixel is given by  $\rho = \Phi(-D/\tau)$ .



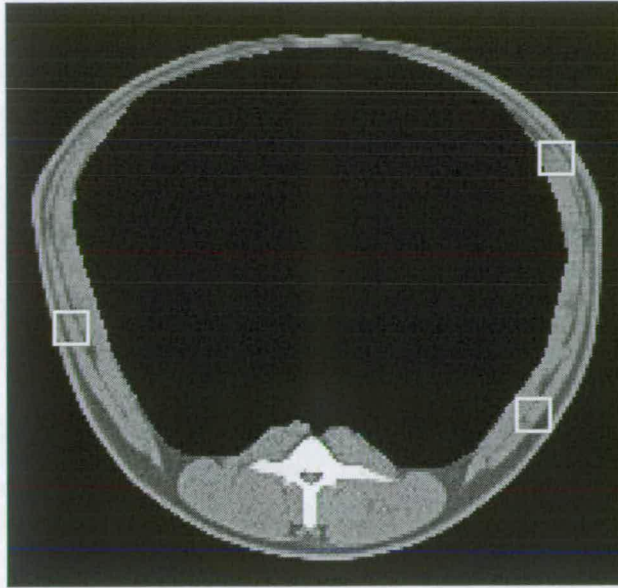


Figure 6.6: *Positions of three sub-images in the carcass region, which are magnified and shown in Figure 6.7.*

- The boundaries between tissues can be represented by random lines which are approximately parallel.

### 6.6.1 Assessing the accuracy of the original assumptions

The validity of the assumption relating to pixels being affected by only one boundary, that boundaries are more than a perpendicular distance of two apart, is investigated by examining three sub-images of a typical segmented lumbar image: see Figure 6.7. The positions of the three chosen sub-images are shown by the white boxes in the segmented lumbar image in Figure 6.6 and are magnified to show the distances between boundaries that exist in reality. These particular sub-images are chosen because they lie in the area where many layers of muscle exist around the sides of the sheep, see Figure 4.1. It can also be seen from the sub-images that the boundaries between tissues are approximately linear and parallel. Very thin layers of fat can exist between these regions of muscle, depending on the fatness of the individual sheep. It is known from the true position within the carcass that two layers of fat exist in each sub-image, some of which are just visible. It is not necessary to determine the exact distance between these boundaries, but rather show that some are closer than a distance two. This is achieved by simulating an image with known distances between the tissue boundaries, and comparing it by eye with the images in Figure 6.7.

An image containing alternate layers of fat and muscle is simulated, and as before,

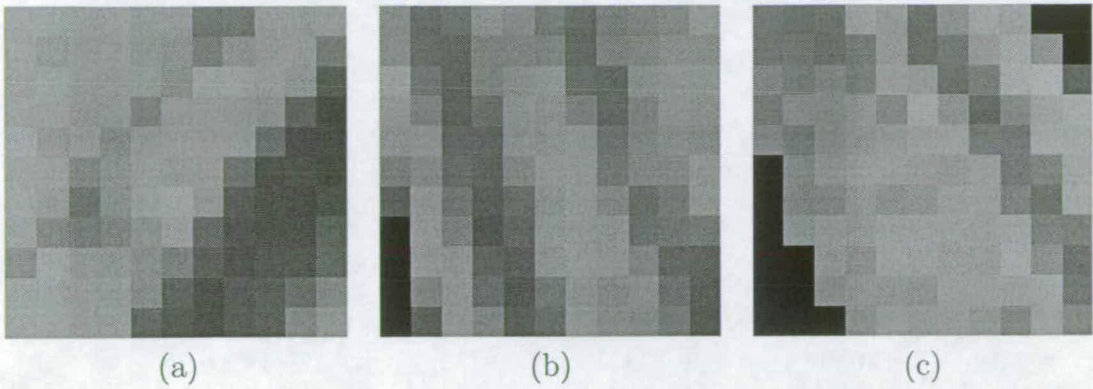


Figure 6.7: *Three sub-images taken from a segmented lumbar image. The positions of the three images within the carcass are shown by the white boxes in Figure 6.6.*

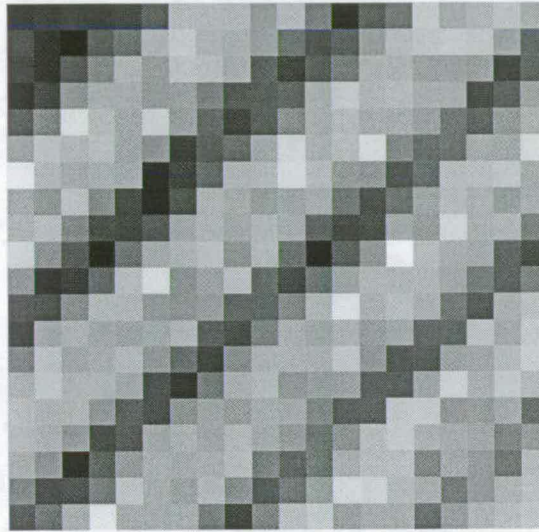


Figure 6.8: *A simulation of parallel lines with increasing distances between the layers of fat. This is for comparison with Figure 6.7.*

the layers are drawn parallel to each other. The layers of muscle are displayed with constant thickness of four units, whereas the distance between the thin layers of fat increases by 0.5 each time. These distances are given by (0.5, 1.0, 1.5, 2.0). This simulation is shown in Figure 6.8. On comparison with Figures 6.7 (a), (b) and (c) it can be clearly seen that the boundaries in the chosen sections can in reality lie closer than two pixels apart. Therefore, assuming that the standard deviation of the PSF is  $\tau = 0.41$ , then the original assumptions are inaccurate in some regions of the carcass. However, it is already assumed that most of the mixed pixels lie in the area from which these sub-images were chosen. Given that a mixed pixel can be affected by more than one boundary, the proportion of fat in some pixels will no longer be able to be estimated using (5.8), hence altering the probability density function of pixel values.



## 6.6.2 Modifying the simulations

In Subsection 6.6.1, it was shown that a mixed pixel can be affected by more than one boundary between tissues. Therefore, it has been decided to modify the previous simulations from Section 6.3. This is achieved by allowing for boundaries which lie closer together than a perpendicular distance of two (given the estimated standard deviation of the PSF). It is decided that it is realistic to assume that the layers of fat are in general less than a perpendicular distance two thick, whereas the layers of muscle are in general more than a thickness of two. This is particularly evident from Figures 6.7 (a), (b) and (c). Therefore, this ensures that the proportion of fat in a mixed pixel can be affected by at most two boundaries.

As before, the simulated images are of size  $60 \times 60$ , and the angle of the parallel lines is simulated from  $\alpha \sim U(0, \pi/4)$  with the first radius  $p_1 \sim U(20, 30)$ . In order to simulate images which contain a significant number of pixels affected by two boundaries, 9 parallel lines are used for this set of simulations and the distances between them,  $d_i$ , are selected alternately from  $\sim U(0, 2)$  and  $\sim U(3, 10)$  to produce four thin layers of fat. As before, the remaining radii  $p_i$  are found using  $p_{i+1} = p_i - d_i$ , for  $i = 1, \dots, 9$ . An example of this simulation is shown in Figure 6.9 (a) and the corresponding histogram of pixel values for the image is shown in Figure 6.9 (b). It can be seen from this image that in some cases, if  $d_i$  is very small, then the layer of fat is only just visible. This is similar to the sub-images in Figure 6.7 (a). The simulated  $d_i$  in Figure 6.9 (a) to produce the thin layers of fat, from the bottom right to top left corners are 1.7, 0.8, 1.8, 0.4. Again, 100 simulations are performed for each pair of tissue variances, but the amount of fat is estimated using only two methods: the moment method and the maximum likelihood method (parameters known,  $\theta_K$ ). In any subsequent results, these will just be referred to as the Moment and the ML methods. The RMSE's of the estimates are shown in Table 6.7. On comparison with Table 6.3 (column number 5), it is clear that the RMSE of the ML method has dramatically increased as a result of the modified assumptions, whereas the moment estimator results have remained constant and now out-perform the ML method. In addition, it can be seen that the ML method is now producing biased estimates, by over-estimating the amount of fat in each simulation, but the moment estimator remains approximately unbiased. This indicates that the moment estimator is more robust and not as sensitive to changes in assumptions. The reason behind this is now explained.

It was shown in Subsection 5.3.3 that under an isotropic bivariate normal PSF, a

mixed pixel close to only one boundary has a proportion of fat of  $\rho = \Phi(-D/\tau)$ . Given the assumption that  $D \sim U(-1, +1)$ , the probability density function of  $\rho$  was given and then used in turn to estimate the probability density function of pixel values in an image. However, under the revised assumptions, a mixed pixel can lie sufficiently close to two lines to affect the proportion of fat. If the boundaries between the tissues are considered and if we label the three regions such that fat is the middle tissue, then the proportion of fat in such a mixed pixel is given by

$$\rho_1 = \Phi\left(-\frac{D_1}{\tau}\right) - \Phi\left(-\frac{D_2}{\tau}\right),$$

for  $D_1$  and  $D_2$  simultaneously in the range  $[-1, +1]$ , where  $D_1$  and  $D_2$  are the perpendicular distances from the two lines, and line 1 lies to the right of the fat section. Obviously, if  $D_1$  or  $D_2$  are outside the range  $[-1, +1]$  then the proportion of fat is reduced to a single integral. Due to the spacing of the boundaries there are no pure fat pixels within the four strips of fat. The only pure fat pixels exist in the top left hand corner of the image. Given this definition of  $\rho_1$  the density function for  $\rho_1$ , denoted by  $f(\rho_1)$  will change, as will the overall probability density function of pixel values within the image. Therefore, the pixel values in the new simulations will be from a different distribution than (5.17), which is fitted to the histogram of pixel values to produce the ML results. Therefore, biased results would be expected. However, the moment estimator only considers the expectation of the pixel values to estimate the proportion of fat.

On further examination of Table 6.7, it can be seen that although the moment estimator has a lower RMSE than the ML method and is unbiased, the variance of the estimates of the amount of fat is considerably larger than for the ML method. It is particularly evident from Tables 6.3 and 6.7 that as the variances of the pure tissues (i.e.  $\sigma_f^2$  and  $\sigma_m^2$ ) increase, so does the variance of the estimates (and hence the RMSE) of the moment estimator. Hence, if the pure pixels could be identified in some way and we are required only to estimate the amount of fat present in the mixed pixels, then the variance of the moment estimator could be reduced, hence making it even more efficient.

The moment estimator method used to date has considered only the frequency of pixel values from the histogram and has not incorporated any spatial information on each pixel, i.e. it has not included any information on the positioning of each pixel in the image or information from the neighbouring pixels. It must be decided how much spatial information should be used in order to classify the pixels, i.e. should information on the rest of the image be used to classify each pixel or

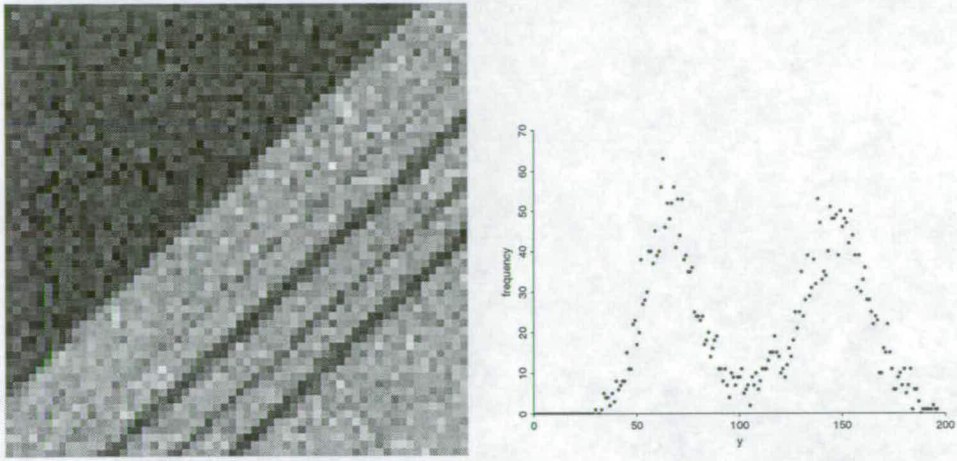


Figure 6.9: (a) A modified simulated image, with four thin layers of fat visible amongst the muscle. This is simulated using  $\sigma_f^2 = 150$ ,  $\sigma_m^2 = 250$  and as usual  $\mu_f = 65$  and  $\mu_m = 150$ , (b) The histogram of pixel values for the simulated image in (a).

just a small neighbourhood? In addition, it is desired that the amount of incorporated spatial information does not dramatically increase the overall computation time for estimating of the tissue proportions. We consider examining a small neighbourhood centred around each pixel to allow the pure pixels to be classified more easily and hence remove some of the variability that exists within the pure pixels, with the overall aim of reducing the variance of the moment estimator. The inclusion of spatial information on each pixel is investigated in Section 6.7.

## 6.7 Incorporating spatial information of pixels to estimate the tissue proportions

Having decided to use spatial information on each pixel to estimate the proportion of each pure tissue, we use this additional knowledge to estimate the overall amount of fat present in the image. This is achieved by calculating the variance of the pixel values within a  $3 \times 3$  square centred on each pixel. This is an example of a non-linear spatial filtering: see Subsection 4.6. The edge and corner pixels do not have eight nearest neighbours and there are several ways of dealing with these pixels: see Glasbey and Horgan (1995). Here, this variance filter has been modified to deal with incomplete neighbourhoods by ignoring the parts of the  $3 \times 3$  square which fall outside of the image. Hence, for the corner pixels the three nearest neighbours (and the pixel itself) are used and for the remaining edge pixels the five nearest neighbours (and the pixel itself) are used to calculate

$\sigma_f^2$	$\sigma_m^2$	RMSE		bias		variance	
		ML	moment	ML	moment	ML	moment
75	125	7.6	6.9	-3.9	0.7	43.3	46.5
	250	11.9	8.7	-8.5	0.9	68.7	75.7
	500	15.7	11.7	-11.7	1.2	108.7	135.0
150	125	9.8	8.0	-7.0	0.8	46.9	63.8
	250	12.9	9.7	-9.8	1.0	71.0	92.9
	500	15.0	12.4	-11.0	1.3	103.9	152.2
300	125	12.2	10.0	-9.0	1.0	67.8	99.7
	250	13.5	11.4	-10.0	1.2	84.1	128.4
	500	14.4	13.7	-9.7	1.5	114.3	186.7

Table 6.7: *The RMSE, bias and variance for all pairs of variances of pure tissues, for both the maximum likelihood method and the moment method. The results are for simulated images (similar to Figure 6.9) and with 9 lines and 4 thin layers of fat.*

the variance of that pixel. Figure 6.10 shows the effect of the variance filter applied to the simulated image in Figure 6.2 (a). Small variances are displayed as dark/black pixels, and large variances as light/white pixels. On comparison with Figure 6.2(a), it can be seen that the pure pixels of either tissue type correspond to the darker pixels in Figure 6.10 and the lighter pixels are those lying close to the boundaries, and hence are the mixed pixels.

By using this spatial information to identify the pure pixels, the variance of the moment estimator results should be reduced. The original simulations will be used to investigate this approach (i.e. seven lines all more than a perpendicular distance of two apart) and therefore the results of the unbiased maximum likelihood, ML ( $\theta_K$ ), can be used to re-calculate the estimated efficiency of the moment estimator.

Due to the original assumptions that the boundaries are all more than a distance of two apart and mixing extends only a distance of one in all directions, it seems justified to use a  $3 \times 3$  square to estimate the variance of each pixel. Therefore, if the variance of a pixel is large then it is to be considered a mixed pixel, as some of the neighbours will be pure pixels and some will be mixed. Similarly, a pixel which has a small variance can be assumed to be a pure tissue pixel since most of the neighbours have similar greyscale values and must therefore be the same tissue. A pixel with a small variance cannot be considered as a mixed pixel based on the model assumptions (all lines  $> 2$  apart) as a mixed pixel cannot be totally surrounded by other mixed pixels. An intermediate value of the variance may suggest pure pixels with mixed neighbours or a mixed pixel with mostly pure neighbours. Hence, a decision criterion, based on these local variances, needs to



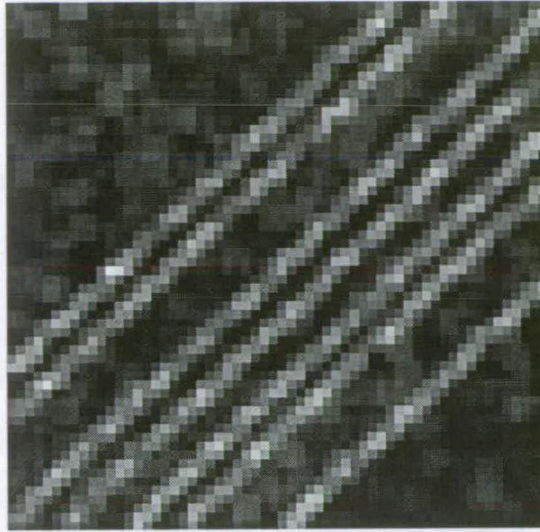
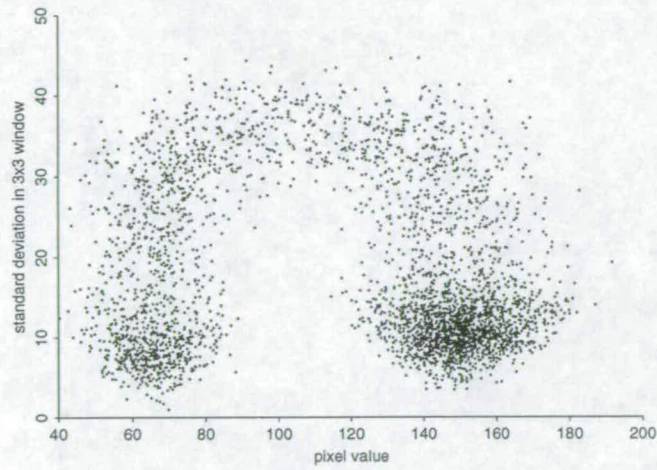


Figure 6.10: *The effect of the variance filter (using a  $3 \times 3$  square) applied to the simulated image in Figure 6.2 (a).*

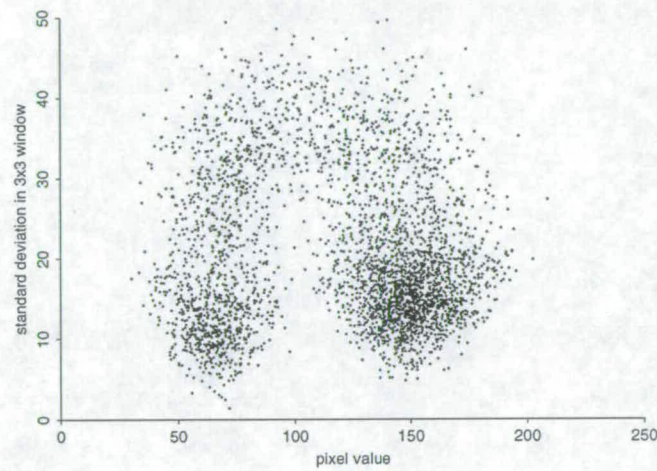
be established to distinguish between the types of pixels.

The new variance information for each pixel is used in combination with the original greyscale values in attempting to classify the pure and mixed pixels. Figure 6.11 shows the standard deviation of each pixel value plotted against the corresponding greyscale value, for three pairs of pure tissue variances, i.e.  $\sigma_f^2$  and  $\sigma_m^2$ . The clusters of pixels with small standard deviation at the bottom left and bottom right hand sides of the image correspond to the pure fat and pure muscle pixels respectively. The remainder of the pixels, (the arc of pixels connecting the two clusters) correspond to mixed pixels or pure pixels which have mixed neighbours and therefore have a larger variance. It can be seen from Figure 6.11 that the larger the values of  $\sigma_f^2$  and  $\sigma_m^2$ , the larger the overlap of pure and mixed pixels and the harder it becomes to distinguish the pure pixels from the mixed pixels. The next stage is to propose a decision criterion which can be used to classify all the pixels into these three groups. The choice of decision criterion is discussed in Subsection 6.7.1.

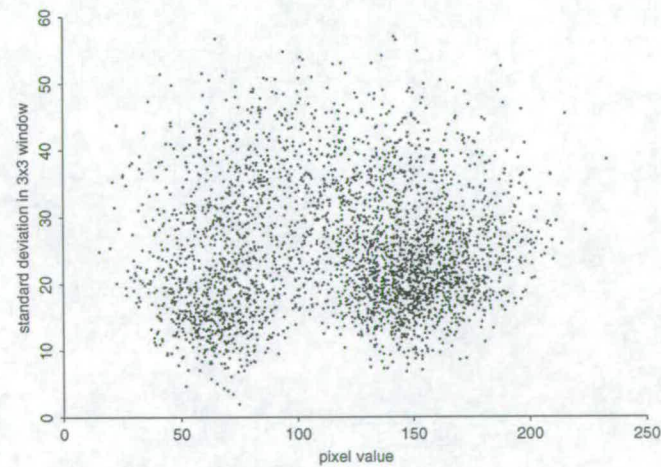
If, after classification, the estimated total number of pixels belonging to each of the three possible types is obtained and  $n_f$  pixels are classified as pure fat and  $n_m$  as pure muscle, hence  $n_{mix} = 3600 - n_f - n_m$  are classified as mixed pixels. Previously, in Subsection 6.4.2, the overall proportion of fat in an image, using the moment method, is given by (6.6). However, since the number of pure pixels  $n_f$  has already been estimated (using the decision criterion), then (6.6) may be used to obtain an estimate of the proportion of fat, in the pixels classified as



(a)



(b)



(c)

Figure 6.11: *Bivariate plots of the standard deviation in each pixel, given a  $3 \times 3$  window, against the original greyscale value of each pixel for (a)  $\sigma_f^2 = 75$ ,  $\sigma_m^2 = 125$ , (b)  $\sigma_f^2 = 150$ ,  $\sigma_m^2 = 250$  and (c)  $\sigma_f^2 = 300$ ,  $\sigma_m^2 = 500$ . All simulated images had seven lines which were all more than two pixels apart.*



mixed pixels and is given by

$$\pi_{mix} = \frac{\mu_m - \bar{y}_{mix}}{\mu_m - \mu_f}, \quad (6.7)$$

where  $\bar{y}_{mix} = \frac{1}{n_{mix}} \sum \sum f(i, j)$  is the mean of the greyscale values  $f(i, j)$  for the  $n_{mix}$  pixels which are classified as mixed pixels. Hence, an estimate for the total amount of fat in an image, using this method, is given by

$$n_f + n_{mix}\pi_{mix} .$$

This estimator will be referred to as the modified moment estimator.

### 6.7.1 Selecting thresholds for the variances of pure tissues

On examination of Figure 6.11(a), it seems clear that this plot can be split into three regions corresponding to pure fat, pure muscle and mixed pixels. We approach this by selecting thresholds on both axes. The threshold,  $g$ , on the pixel value axis is taken to be the average of the two pure tissue means, i.e.  $g = \frac{\mu_f + \mu_m}{2}$ . However, selecting the threshold on the variance axis is not so obvious. A simple, but subjective approach is to examine the histogram of the pixel variances in each image and select an appropriate threshold value by eye. However, this is a very approximate method of estimating the amount of fat in an image and would need to be standardised for all images.

Therefore, using the fact that the variances of the pure tissues,  $\sigma_f^2$  and  $\sigma_m^2$ , are assumed to be known in the moment method (see Subsection 6.4.2), it is known that the estimated variances,  $s^2$ , from a window containing  $n$  pure pixels has distribution

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi_{(n-1)}^2,$$

where  $\sigma^2$  is the known variance of a given tissue. Given that the variance of the majority of the pixels is found by considering a  $3 \times 3$  square and the mean of these values is estimated, it is reasonable to assume that most of the pure pixels in the image, with estimated variance  $s^2$ , have distribution

$$\frac{8s^2}{\sigma^2} \sim \chi_8^2.$$

This result is based on the assumption that the pixel values for both pure tissues are normally distributed and are independently and identically distributed.

Therefore, for a probability  $\beta$ , and a known true variance of the tissue,  $\sigma^2$ , an appropriate threshold on the variance axis, denoted  $v$ , is given by

$$v = \frac{\sigma^2}{8} \chi_8^2(\beta),$$

where  $\beta$  is the probability that  $\frac{8s^2}{\sigma^2} \geq \chi_8^2(\beta)$ . However, in these images there are two pure tissues, with different known variances and therefore it raises the question how to select a value for  $\sigma^2$ ; either to have two thresholds, one for each tissue type, or to have one threshold value, where  $\sigma^2$  is taken as a compromise between the two known variances  $\sigma_f^2$  and  $\sigma_m^2$ . In both methods below, the estimated variance of a pixel at location  $(i, j)$  in the image is denoted as  $s^2(i, j)$ .

### Two-threshold method

In this case a different variance threshold is used for fat and muscle, based on the two pure-tissue variances,  $\sigma_f^2$  and  $\sigma_m^2$ . Letting  $v_1$  and  $v_2$  be the variance thresholds for fat and muscle respectively, the decision rule is

- if  $s^2(i, j) < v_1$  and  $f(i, j) < g$  then  $(i, j)$  is classified as pure fat,
- if  $s^2(i, j) < v_2$  and  $f(i, j) \geq g$  then  $(i, j)$  is classified as pure muscle,
- otherwise,  $(i, j)$  is classified as mixed,

$$\text{where } v_1 = \frac{\sigma_f^2}{8} \chi_8^2(\beta) \text{ and } v_2 = \frac{\sigma_m^2}{8} \chi_8^2(\beta).$$

### One-threshold Method

This method uses an average of the two pure tissue variances,  $\sigma_f^2$  and  $\sigma_m^2$ , to obtain just one threshold. Therefore, for a common  $\beta$  and using a combined threshold, denoted  $v_3$ , the decision rule is

- if  $s^2(i, j) < v_3$  and  $f(i, j) < g$  then  $(i, j)$  is classified as pure fat,
- if  $s^2(i, j) < v_3$  and  $f(i, j) \geq g$  then  $(i, j)$  is classified as pure muscle,
- otherwise,  $(i, j)$  is classified as mixed,

$$\text{where } v_3 = \frac{\sigma_f^2 + \sigma_m^2}{16} \chi_8^2(\beta).$$

$\beta$ (%)	Two thresholds					One threshold			
	$v_1$	$v_2$	RMSE	bias	variance	$v_3$	RMSE	bias	variance
20.0	103	172	4.3	0.4	18.3	138	4.4	0.6	19.4
10.0	125	209	3.9	-0.1	14.8	167	4.2	0.3	17.6
5.0	145	242	3.8	-0.1	14.3	194	3.9	0.1	14.9
2.5	164	244	3.7	0.1	14.0	219	3.7	0.1	13.9
1.0	188	314	3.5	0.3	12.4	251	3.6	0.2	13.0
0.5	206	343	3.5	0.4	11.8	274	3.6	0.2	13.0
0.1	245	408	<b>3.2</b>	0.1	10.4	327	<b>3.4</b>	0.3	11.5

Table 6.8: *The RMSE, bias and variance for various probabilities,  $\beta$ , and corresponding thresholds when  $\sigma_f^2 = 75$  and  $\sigma_m^2 = 125$  for the pure tissue variances. For each threshold method, the optimal RMSE is displayed in bold.*

## 6.7.2 Results from using spatial information

Three pairs of variances of pure tissues are chosen to provide an indication of the trends in the RMSE. The three pairs and the reasons for choosing them are:

- $\sigma_f^2 = 75, \sigma_m^2 = 125$  are chosen to show the effect on the results when there is clear definition between both pure tissues and also mixed pixels.
- $\sigma_f^2 = 150, \sigma_m^2 = 250$  are chosen as these are the values which were estimated to be the variances of the pure tissues, estimated from the training set used in Chapter 1.
- $\sigma_f^2 = 300, \sigma_m^2 = 500$  are chosen to show the effect on the results when the variances of the pure pixel values are large. In this case there is not such clear distinction between the pure pixels and the mixed pixels.

The results of using both one and two thresholds are shown in Tables 6.8, 6.9 and 6.10. They display the threshold/s (rounded to the nearest integer) corresponding to various values of  $\beta$  and show the RMSE, bias and variance for the new estimates of the amount of fat present in an image using the modified moment estimator.

From Table 6.8 it can be seen that for both the one and two threshold approaches, and the chosen values of  $\beta$ , the optimal RMSE is achieved when  $\beta = 0.1$ . A small value of  $\beta$  should be expected from examining Figure 6.11(a), since there is very clear distinction between the pure fat and pure muscle pixels, and also the mixed pixels.

In Table 6.9, for the two threshold case, again small values of  $\beta$  produce a small RMSE. The optimal  $\beta$  in the one-threshold case is considerably higher than for the previous case. This is again due to fairly clear definition between the types

$\beta(\%)$	Two thresholds					One threshold			
	$v_1$	$v_2$	RMSE	bias	variance	$v_3$	RMSE	bias	variance
20.0	207	345	6.0	0.6	35.7	276	6.2	0.9	38.2
10.0	251	418	5.4	0.5	28.5	334	5.7	0.7	32.5
5.0	291	485	5.2	0.5	26.2	388	<b>5.1</b>	0.7	26.0
2.5	329	548	5.0	0.5	25.3	438	5.4	1.3	27.2
1.0	377	628	4.9	0.2	23.9	502	5.6	1.8	27.6
0.5	412	686	<b>4.8</b>	-0.1	23.1	549	5.5	1.9	26.6
0.1	490	816	5.1	-1.2	24.4	653	5.7	2.3	27.2

Table 6.9: *The RMSE, bias and variance for various probabilities,  $\beta$ , and corresponding thresholds when  $\sigma_f^2 = 150$  and  $\sigma_m^2 = 250$  for the pure tissue variances. For each threshold method, the optimal RMSE is displayed in bold.*

$\beta(\%)$	Two thresholds					One threshold			
	$v_1$	$v_2$	RMSE	bias	variance	$v_3$	RMSE	bias	variance
40.0	313	522	10.9	0.9	117.5	418	10.5	1.4	108.6
30.0	357	595	10.9	0.5	118.5	476	10.5	0.8	109.8
20.0	414	689	<b>9.9</b>	0.7	97.5	552	<b>10.4</b>	0.9	108.3
10.0	501	835	10.0	0.8	98.4	668	10.9	2.8	110.6
5.0	582	969	10.3	0.4	105.3	776	12.2	5.7	117.3

Table 6.10: *The RMSE, bias and variance for probabilities,  $\beta$ , and corresponding thresholds when  $\sigma_f^2 = 300$  and  $\sigma_m^2 = 500$  for the pure tissue variances. For each threshold method, the optimal RMSE is displayed in bold.*

of tissues: see Figure 6.11(b).

In Table 6.10, the values of the optimal  $\beta$  have greatly increased. This is due to the fact that the variances of the two pure tissues are large. On examination of Figure 6.11(c), there are no obvious clusters corresponding to pure fat and pure muscle. It has become more difficult to distinguish a mixed pixel from a pure pixel with a large variance, due to the large variability within the pure pixels alone. In summary of these three tables, as expected, the two-threshold method performs slightly better for all three pairs of pure tissue variances than the one-threshold method.

Table 6.11 shows the average proportion of pixels assigned to the pure tissue categories over 100 simulations, using the two threshold approach. The results are again for the three pairs of variances and are displayed only for the optimal  $\beta$  found earlier for these given variances. The true average proportions over 100 simulations for pure fat and pure muscle are 0.28 and 0.49 respectively.

Table 6.12 shows the optimal RMSE obtained when using spatial information of pixels to modify the moment estimator approach (with one and two thresholds)

$\beta(\%)$	Tissue variances		Average proportions	
	$\sigma_f^2$	$\sigma_m^2$	fat	muscle
0.1	75	125	0.24	0.48
0.5	150	250	0.25	0.49
20.0	300	500	0.20	0.38

Table 6.11: *The average proportions of pixels assigned to pure fat and pure muscle using the two threshold approach, for the optimal  $\beta$  using the three pairs of variances. The actual average proportions (over 100 simulations) are 0.28 and 0.49 for fat and muscle respectively.*

$\sigma_f^2$	$\sigma_m^2$	without spatial information		with spatial information			
		ML	Moment	Two threshold		One threshold	
				RMSE	$\beta(\%)$	RMSE	$\beta(\%)$
75	125	4.7	6.5	3.2	0.1	3.4	0.1
	250	5.4	8.1	4.3	0.5	4.6	0.5
	500	7.3	10.5	6.9	1.0	8.5	30.0
150	125	5.2	8.0	4.2	0.1	4.0	0.1
	250	5.8	9.3	4.8	0.5	5.1	5.0
	500	7.7	11.5	7.0	1.0	9.7	20.0
300	125	6.6	10.3	5.8	10.0	6.0	5.0
	250	7.1	11.3	7.0	10.0	7.3	10.0
	500	8.9	13.1	9.9	20.0	10.4	20.0

Table 6.12: *The optimal RMSE for each pair of pure tissue variances and using either 1 or 2 thresholds. The corresponding  $\beta$  for each method is also shown. The RMSE values for the maximum likelihood method and the original moment estimator values (before pure pixel classification, taken from Table 6.3) are included for comparison.*

for each of the nine original pairs of tissue variances. The corresponding optimal probability,  $\beta$ , is shown for each method. The RMSE from the moment estimator when spatial information has not been included and the maximum likelihood method (taken from Table 6.3) are included for ease of comparison. Overall, it can be seen that for each pair of tissue variances and for each threshold method, the RMSE for this classification approach (columns 5 and 7) has considerably reduced in comparison to the RMSE for the original moment method (column 4). In all cases (except  $\sigma_f^2 = 150$  and  $\sigma_m^2 = 125$ ) the two-threshold method performs better than the one-threshold method. It is also evident from Table 6.12 that there is no common optimal  $\beta$  for all pairs of variances, using either threshold method. If this has been the case, then a single value of  $\beta$  could be used for all simulations, irrespective of the variances of the pure tissues. However, it is interesting to notice from Table 6.12 that in the two-threshold case, the new

		Relative efficiency of modified moment estimator (%)	
$\sigma_f^2$	$\sigma_m^2$	2 thresholds	1 threshold
75	125	147	138
	250	126	117
	500	106	86
150	125	124	130
	250	121	114
	500	110	79
300	125	114	110
	250	101	97
	500	90	86

Table 6.13: *The relative efficiency of the modified moment estimator, (which uses spatial information), relative to the maximum likelihood estimator, Method 3 MLK (see Subsection 6.4.2).*

RMSE for this classification moment method is now less than the original RMSE for the ML method, except for  $\sigma_f^2 = 300$  and  $\sigma_m^2 = 500$ . Similarly, when the tissue variances are both fairly small, the same applies in the one threshold case.

Due to the extra information incorporated into the modified moment estimator, it becomes more efficient than the ML method, in the majority of 100 simulations for the pairs of variances for the pure tissue pixels. The efficiency of the estimator using spatial information and the moment estimator, relative to the ML method (from Section 6.5), is shown in Table 6.13.

The results in Tables 6.12 and 6.7, show that using the moment method in conjunction with spatial information from the pixels provides more efficient and robust estimates of the amount of fat, especially when the original model assumptions do not always hold and thin layers of tissue are modelled in the image.

## 6.8 Summary and conclusions

We have compared threshold-based methods with mixed-pixel methods (using the probability density function of pixel values derived in (5.17)) by means of simulating images which contain only two tissue types. It was found that the threshold-based methods (one of which is currently in use at the SAC-BioSS CT unit) produced biased results and were less accurate than the mixed-pixel methods for the estimation of tissue proportions. The two mixed-pixel methods

produced approximately unbiased estimates of the amount of fat present in the images.

However, the assumptions that were made to estimate the probability density function in (5.17) are not plausible in all sections of the segmented images, as thin layers of fat exist between the layers of muscle: see Subsection 6.6.1. Therefore, the simulations were modified to account for this and the two mixed-pixel methods, maximum likelihood and the moment method, were used to estimate the amount of fat. Given these modifications, it was shown that the maximum likelihood method produced biased results (on average underestimating the amount of fat), whereas the moment method remained unbiased. However, the variance of this moment estimator was very large in comparison to the maximum likelihood method. This variance was significantly reduced by considering a pixel's spatial context which examines the estimated variance of the greyscale values of its neighbours. Using two thresholds on the local variance, we were able to classify many of the pure pixels and therefore use the moment method from Subsection 6.4.2 to estimate the proportion of fat in the remaining pixels. It has been shown that by considering spatial information on each pixel, the modified moment estimator is more efficient than the maximum likelihood estimator for estimating tissue proportions.



# Chapter 7

## Conclusions and further work

The aim of this thesis was to develop and apply statistical methods for automating the segmentation of relevant tissues in X-ray CT images of sheep, and from this, provide estimation of the appropriate tissue proportions. In this final chapter we review the current methods along with the strategies we adopted to improve on their results. In addition we bring together our conclusions as we discuss the effectiveness of our approach. Finally, we suggest possible further work in Section 7.2.

Chapter 1 outlined the three-stage process currently in use at the SAC-BioSS CT unit in order to estimate the overall volume of each tissue present in the carcass of a sheep. The three stages of the process are summarised below.

**Stage 1:** Extract the carcass region of the sheep which contains the fat, muscle and bone, from the areas containing the cradle and the internal and external organs at three anatomical positions: ischium, lumbar and thorax. Prior to this research, a skilled operator performed this segmentation manually.

**Stage 2:** The proportions of air, fat, muscle and bone are estimated in the three segmented images (from Stage 1). To achieve this, three thresholds on the greyscale values of the pixels are selected.

**Stage 3:** These estimated tissue proportions, together with the liveweight of the animal are regressed on the dissected tissue weights to predict the overall amount of each tissue type within the whole carcass.

Although all three stages have been outlined, this thesis is concerned only with improving the first two. However, this is a continuing project at the SAC-BioSS CT unit, and further research may allow for Stage 3 to be reassessed and further developed.

## 7.1 Conclusions

In Chapter 2 we introduced and briefly outlined the history of X-ray CT. The scanning system of a typical CT machine was discussed and in particular we studied the SAC-BioSS CT scanner, together with the various scanning strategies that are in operation for sheep.

Low-level segmentation techniques were described in Chapter 3 and their weaknesses for this project were highlighted. To overcome these problems, deformable templates were proposed as an attractive approach to modelling anatomic structures, as they provide compact representations of the object shape. A comprehensive, but not exhaustive, review of both free-form models and parametric deformable templates was provided, although in particular focused on the latter of the two as this approach was adopted in Chapter 4.

In Chapter 4, deformable templates were used to automatically segment the lumbar images. Previously, this was performed by a skilled operator who manually traced the region of interest using a computer mouse. A training set of 24 manually segmented images was used to construct a stochastic model for the distribution of the boundaries. The images in the training set were from Suffolk sheep, all at approximately the same age.

The manually segmented boundaries were parametrised using Fourier coefficients and took into account both the rotation of the sheep in the cradle and the size of the animal. The latter was achieved by representing the radii of the inner boundary as a proportion of the distance to the outer boundary from the centroid of the image. This helps to correct the asymmetry of the animal due to gravity when it is not upright in the cradle. A reduced rank approximation to the matrix of Fourier coefficients, similar to principal component analysis, was used to reconstruct the manual boundaries, and from this we estimated the distribution of the parameters of the template. We demonstrated that the fitted boundaries approximated the manual boundaries very accurately.

To define the measure of fit of the template, to a new image not in the training set, in terms of local edge information, Prewitt's gradient filter was applied. This highlighted tissue boundaries in the images. The objective function, which combined this measure of fit with the fitted distribution for the Fourier coefficients, was optimised using the Nelder-Mead algorithm, which produced in most cases, very swift and accurate results. Overall, it was found that the template was more accurately and consistently located by smoothing the image with a Gaussian filter

prior to optimisation.

With one image in particular this approach provided a poorly fitting boundary, but this was thought to be due to the obesity of the sheep, where layers of muscle were 'hidden' by large areas of fat. The model was validated on an independent set of images of the same breed and age as the training set and was found to produce results with similar quality of fit.

The mirror image of each inner boundary was also included to double the size of the training set, motivated by the near symmetry of the images. It was found that the eigenvalues which corresponded to the matrix of sums of products of the coefficients of the sine coefficients about their means could be viewed as measuring the asymmetry in the images.

We have shown that the deformable template approach accurately fits the boundaries of interest in the lumbar images when incorporating information from both the inner and outer manual boundaries in the model. This method removes the subjectivity and tediousness of the manual approach and has since been adopted at the SAC-BioSS CT unit to replace the manual approach, providing the option to accept or reject the fitted boundary selected by the optimisation routine. If the deformable template method produces a poorly fitting boundary, similar to the image mentioned earlier, then the fit is rejected and segmentation is performed manually for that particular image. This considerably reduces the time required to analyse large quantities of images. Points for further work in this stage of the process are given in Section 7.2.

Having automated Stage 1 of the current procedure in Chapter 4, we developed statistical theory in Chapter 5 which has been shown to improve upon the threshold based method for estimating the proportions of fat and muscle. This was achieved by estimating the probability density function of the pixel values within the image. In Chapter 5, we have restricted our attention to only two tissue types, namely fat and muscle, by assuming the air and bone pixels could be easily identified in the histogram of greyscale values.

Due to the resolution of the X-ray CT machine, averaging takes place between tissue types and hence greyscale values of many pixels are in fact responses to mixtures of two or more tissue types. Therefore, it was necessary to model the spatial response of a pixel for an X-ray CT machine. We approached this by examining a known edge in the image, between the cradle and air, which was modelled as an isotropic, bivariate normal density. We assumed, similar to many authors, that the spatial response was spatially invariant.

From the estimated spatial response, we derived a new probability density function for the values of mixed pixels, (see Glasbey and Robinson, 1999) which to our knowledge has not previously been proposed. This was combined with the distributions for the pure pixels to estimate the probability density function of the greyscale values in the images.

In Chapter 6 we applied the proposed probability density function of pixel values to simulated data which contained only two tissue types. We compared various threshold methods, including that used at the SAC-BioSS CT unit with mixed-pixel methods, which included the new method from Chapter 5. The probability density function of pixel values was fitted to histograms by maximum likelihood, and it was shown that the mixed-pixel methods produced unbiased estimates of the amount of fat, and out-performed the threshold based methods which produced approximately biased results.

We also developed a moment estimator based on the distribution of mixed pixels. This method and the maximum likelihood method were re-evaluated having shown that some of the previous assumptions about the widths of tissue layers do not always hold, and results indicated that the moment method out-performed the maximum likelihood method when the image included thin layers.

By considering information on each of the pixels neighbours, the estimation of the amount of fat using this moment estimator could be improved. To take into account a pixel's spatial context, we calculated the variance of the greyscale values of its neighbours. Selected thresholds on these local variances were used to separate the pixels most likely to be pure fat or pure muscle. The moment estimator was then used to calculate the proportion of fat in the remaining pixels which were not considered to be pure tissue. This procedure produced more efficient estimates of fat than the maximum likelihood method under the original assumptions. Although we have considered only estimation of two tissue proportions, we have provided a strong platform for any future work on images containing more than two tissue types. Modelling the histogram of pixel values in an image containing two tissue types using the estimated probability density function in Chapter 5, has been shown to more accurately estimate the tissue proportions than the SAC-BioSS CT method. Therefore, this mixed pixel method could be extended to deal with more tissues, hence improving on the current Stage 2 results.

## Recommendations

Although further work has been suggested to extend the mixed pixel method to handle more than the two tissues, it is recommended that the SAC-BioSS CT unit combine their threshold method with the mixed pixel method, until such research has been carried out. This is proposed as follows:

Currently their method involves three threshold values. However, it is proposed to only use two, at  $-204$  HU and  $176$  HU, to identify the air and bone pixels respectively in the usual manner. This leaves fat and muscle only to be classified. At present the SAC-BioSS CT unit make no provision for mixed pixels in their method, and it is proposed at this stage to assume that mixed pixels exist, but only between fat and muscle. Therefore, the moment method described in Subsection 6.4.2 could be used to estimate the fat and muscle proportions in the remaining pixels. Whether the spatial information is incorporated, would depend on how difficult this would be to implement and the extra computation time required.

These recommendations to improve Stage 2 of the process are justified as the overall goal for meat sheep is to examine the ratio of fat to muscle and reduce the fat proportions, either by genetic breeding or by diet. Hence, the estimation of fat to muscle ratio needs to be more accurately estimated than the bone proportions.

## 7.2 Further work

The methodologies presented throughout this thesis have produced several promising results regarding the further development of the methods currently employed in this three-stage process. However, there are several areas which may benefit from further work. These are discussed in the order which they appear in the thesis.

- To date, only the lumbar images have been segmented by utilizing parametric deformable templates. A similar strategy could be employed to automate the identification of the relevant tissues in the two other anatomical positions; ischium and thorax.
- The selected matching criterion of the deformable template only comprises local edge information from the image. Authors such as Yuille et al. (1992) have also incorporated peaks and valleys of greyscale values in addition to this local edge data. The inclusion of additional knowledge from the image

into the matching criterion may improve the overall segmentation results.

- Prior to optimisation of the objective function, many authors such as Jain et al. (1996) have adopted a multi-resolution approach, from a coarse to a fine resolution in order to smooth the image. It is suggested that a Gaussian filter with a smaller variance could be used as a second stage in the optimisation procedure of the sheep images to locate a more accurately fitting boundary.
- Investigation of the spatial response of the CT machine could have been approached by alternative techniques, for example using Fourier domain methods. In addition, future work may include reassessing the assumption of this response being spatially invariant.
- It was assumed that the perpendicular distance,  $D$ , was uniformly distributed between  $(-1, +1)$ , given that  $\tau = 0.41$ , i.e.  $(-2.5\tau, +2.5\tau)$ . It would lead to an interesting investigation to see how the probability density function,  $f(\rho)$ , and hence the mixed pixel distribution, would change if  $D$  had a different distribution.
- To date, we have only considered estimating the proportions of two tissue types. The moment method, together with the incorporation of spatial information, could be extended in order to estimate the proportions of fat, muscle and bone in the segmented sheep images, at all three anatomical positions. This requires the distribution of bone greyscale values to be estimated.
- Upon establishing the distribution of bone pixel values, a simulation study could be performed to approximate the histogram of greyscale values in Figure 5.1. This would allow for direct comparison between the current threshold method in use at the CT unit and the method in Section 6.7, which combines the local spatial information of each pixel and the moment estimator.

# Bibliography

- Anton, H. and Rorres, C. (1991). *Elementary Linear Algebra: Applications Version*. Wiley, New York, 6th edition.
- Baldock, R. A. (1992). Trainable models for the interpretation of biomedical images. *Image and Vision Computing*, 10:444–450.
- Ballard, D. H. and Brown, C. M. (1982). *Computer Vision*. Prentice-Hall, London.
- Basilevsky, A. (1983). *Applied Matrix Algebra in the Statistical Sciences*. North-Holland, New York.
- Bracewell, R. and Riddle, A. (1967). Inversion of fan-beam scans in radio astronomy. *Journal of Astro Physics*, 150:427–434.
- Cootes, T., Cooper, D. H., Taylor, C., and Graham, J. (1992). Trainable method of parametric shape description. *Image and Vision Computing*, 10:289–294.
- Cootes, T., Hill, A., Taylor, C., and Haslam, J. (1994). Use of active shape models for locating structures in medical images. *Image and Vision Computing*, 12:355–366.
- Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models - their training and application. *Computer Vision and Image Understanding*, 61:38–59.
- Cormack, A. M. (1963). Representation of a function by its line integrals, with some radiological applications. *Journal of Applied Physics*, 34:2722–2727.
- Dore, S., Kearney, R. E., and DeGuise, J. A. (1997). Experimental correlation-based identification of X-ray CT point spread function. Part 1: method and experimental results. *Medical and Biological Engineering and Computing*, 35:2–8.
- Everitt, B. and Hand, D. (1981). *Finite Mixture Distributions*. Chapman and Hall, London.



- Fischler, M. A. and Elschlager, R. (1973). The representation and matching of pictorial structures. *IEEE Transactions on Computers*, C-22:67–92.
- Fisker, R. and Carstensen, J. M. (1998). On parameter estimation in deformable models. In *International Conference on Pattern Recognition*, pages 762–766.
- Gage, H. D., Santago, P., and Snyder, W. E. (1992). Quantification of brain tissue through incorporation of partial volume effects. *SPIE Medical Imaging V1-Image Processing*, 74:84–96.
- Glasbey, C. and Robinson, C. D. (1999). Estimation of tissue proportions in x-ray ct images using a new mixed pixel distribution. *Task Quarterly*, 3:409–418.
- Glasbey, C. A. (1998). Ultrasound image segmentation using stochastic templates. *Journal of Computing and Information Technology*, 6:107–116.
- Glasbey, C. A. and Horgan, G. (1995). *Image Analysis for the Biological Sciences*. Wiley, Chichester.
- Glasbey, C. A., Horgan, G., and Hitchcock, D. (1994). A note on the grey-scale response and sampling properties of a desktop scanner. *Pattern Recognition Letters*, 15:705–711.
- Glasbey, C. A. and Robinson, C. D. (2000). Inference from x-ray ct images of sheep. In *Bulletin of the International Statistical Institute*, volume Book 1, pages 469–472.
- Glasbey, C. A., Robinson, C. D., and Young, M. J. (1999). Segmentation of x-ray ct images using stochastic templates. In *Proceedings of 10th International Conference on Image Analysis*, pages 746–751.
- Glasbey, C. A. and Young, M. J. (2000). Maximum *a posteriori* estimation of image boundaries using a shortest path algorithm. Submitted.
- Gonzalez, R. C. and Woods, R. E. (1993). *Digital Image Processing*. Addison-Wesley, New York.
- Grenander, U. (1993). *General Pattern Theory: A Mathematical Study of Regular Structures*. Oxford University Press, Oxford.
- Grenander, U., Chow, Y., and Keenan, D. M. (1991). *Hands. A Pattern Theoretic Study of Biological Shapes*. Springer-Verlag, New York.
- Harville, D. A. (1997). *Matrix Algebra from a Statistician's Perspective*. Springer-Verlag, New York.

- Herman, G. T. (1980). *Image Reconstruction from Projections: the Fundamentals of Computerised Tomography*. Academic Press, London.
- Hill, A. and Taylor, C. J. (1992). Model-based image interpretation using genetic algorithms. *Image and Vision Computing*, 10:295–300.
- Hiriyannaiah, H. P. (1997). X-ray computed tomography for medical imaging. *IEEE Signal Processing Magazine*, 14:42–59.
- Hounsfield, G. N. (1972). A method of and apparatus for examination of a body by radiation such as x or gamma radiation. *The Patent Office, London*.
- Jain, A. K., Zhong, Y., and Dubuisson-Jolly, M. (1998). Deformable template models: a review. *Signal Processing*, 71:109–129.
- Jain, A. K., Zhong, Y., and Lakshmanan, S. (1996). Object matching using deformable templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:267–277.
- Joyce-Loebl (1985). *Image Analysis: principles and practice*. Joyce-Loebl, Gateshead.
- Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: active contour models. *International Journal of Computer Vision*, 1:321–331.
- Kuhl, D. E. and Edwards, R. Q. (1968). Rapid brain scanner with self-contained computer and CRT display for both rectilinear and transverse section viewing. *Journal of Nuclear Medicine*, 9:332.
- Lei, T. and Sewchand, W. (1992a). Statistical approach to X-ray CT imaging and its applications in image analysis. 2. A new stochastic model-based image segmentation technique for X-ray CT image. *IEEE Transactions on Medical Imaging*, 11:62–69.
- Lei, T. and Sewchand, W. (1992b). Statistical approach to X-ray CT imaging and its applications in image analysis. 1. Statistical analysis of X-ray CT imaging. *IEEE Transactions on Medical Imaging*, 11:53–61.
- Lipson, P., Yuille, A. L., O’Keeffe, D., Cavanaugh, J., Taaffe, J., and Rosenthal, D. (1990). Deformable templates for feature extraction from medical images. In *Proceedings of the First European Conference on Computer Vision, ECCV*, volume 81, pages 413–417.
- Luiting, P., Kolstad, K., Enting, H., and Vangen, O. (1995). Pig breed comparison

- for body composition at maintenance: analysis of computerised tomography data by mixture distributions. *Livestock Production Science*, 43:225–234.
- McInerney, T. and Terzopoulos, D. (1996). Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1:91–108.
- McLachlan, G. J. (1992). *Discriminant Analysis and Statistical Pattern Recognition*. Wiley, Chichester.
- Nelder, J. A. and Mead, R. (1965). A simplex method for function minimisation. *Computer Journal*, 7:308–313.
- Numerical Algorithms Group (1993). *Fortran Library Manual Mark 17*. Oxford.
- Phillips, D. B. and Smith, A. F. M. (1994). Bayesian faces via hierarchical template matching. *Journal of the American Statistical Association*, 89:1151–1163.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1996). *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, volume 1. Cambridge University Press, 2nd edition.
- Radon, J. (1917). Uber die Bestimmung von Funktionen durch ihre Integralwerte entlang gewisser Mannigfaltigkeiten. *Berichte Saechsische Akademie der Wissenschaften*, 69:262–279.
- Rathee, S., Koles, Z. J., and Overton, T. R. (1992a). Image restoration in computed tomography: estimation of the spatially variant point spread function. *IEEE Transactions on Medical Imaging*, 11:539–545.
- Rathee, S., Koles, Z. J., and Overton, T. R. (1992b). Image restoration in computed tomography: the spatially invariant point spread function. *IEEE Transactions on Medical Imaging*, 11:530–538.
- Robb, R. A. (1982). X-ray computed tomography: from basic principles to applications. *Annual Review of Biophysics and Bioengineering*, 11:177–201.
- Rosenfeld, A. and Kak, A. (1982). *Digital Picture Processing*, volume 1. Academic Press, London.
- Rueckert, D. and Burger, P. (1997). Geometrically deformable templates for shape based segmentation and tracking in cardiac MR images. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, number 1123 in Lecture Notes in Computer Science, pages 83–97. Springer, Berlin.
- Santago, P. and Gage, H. D. (1995). Statistical models of partial volume effect. *IEEE Transactions on Image Processing*, 4:1531–1540.

- Schalkoff, R. J. (1989). *Digital Image Processing and Computer Vision*. Wiley, New York.
- Shepp, L. and Kruskal, J. B. (1978). Computerized tomography: the new medical X-ray technology. *American Mathematical Monthly*, 85:420–439.
- Shepp, L. and Logan, B. F. (1974). The Fourier reconstruction of a head section. *IEEE Transactions of Nuclear Science*, NS21:21–43.
- Shrimpton, P. and Wall, B. F. (1995). The increasing importance of X-ray computed tomography as a source of medical exposure. *Radiation Protection Dosimetry*, 57:413–415.
- Simm, G. (1998). *Genetic improvement of cattle and sheep*. Farming Press, United Kingdom.
- Staib, L. H. and Duncan, J. S. (1992). Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:1061–1075.
- Stewart, G. W. (1973). *Introduction to Matrix Computations*. Academic Press, New York.
- Stoyan, D., Kendall, W. S., and Mecke, J. (1995). *Stochastic Geometry and its Applications*. Wiley, Chichester, 2nd edition.
- Sugakkai, N. (1987). *Encyclopedic Dictionary of Mathematics: by the Mathematical Society of Japan*, volume 2. MIT Press, Cambridge, MA.
- Thaler, H. T., Ferber, P., and Rottenberg, D. (1978). A statistical method for determining the proportions of gray matter, white matter and CSF using computed tomography. *Neuroradiology*, 16:133–135.
- Yuille, A. L. (1991). Deformable templates for face recognition. *Journal of Cognitive Neuroscience*, 3:59–70.
- Yuille, A. L., Hallinan, P. W., and Cohen, D. S. (1992). Feature extraction from faces using deformable templates. *Journal of Computer Vision*, 8:99–111.