3D Human Body Modelling from Range Data

Laura Dekker

a thesis submitted for the degree of
Doctor of Philosophy in Computer Science
University of London

Department of Computer Science
University College London
2000

Abstract

This thesis describes the design, implementation and application of an integrated and fully automated system for interpreting whole-body range data.

The system is shown to be capable of generating complete surface models of human bodies, and robustly extracting anatomical features for anthropometry, with minimal intrusion on the subject. The ability to automate this process has enormous potential for personalised digital models in *medicine*, *ergonomics*, *design and manufacture* and for *populating virtual environments*. The techniques developed within this thesis now form the basis of a commercial product.

However, the technical difficulties are considerable. Human bodies are highly varied and many of the features of interest are extremely subtle. The underlying range data is typically noisy and is sparse at occluded areas. In addressing these problems this thesis makes five main research contributions.

Firstly, the thesis describes the design, implementation and testing of the whole integrated and automated system from scratch, starting at the image capture hardware. At each stage the trade-offs between performance criteria are discussed, and experiments are described to test the processes developed.

Secondly, a combined data-driven and model-based approach is described and implemented, for surface reconstruction from the raw data. This method addresses the *whole* body surface, including areas where body segments touch, and other occluded areas.

The third contribution is a library of operators, designed specifically for shape description and measurement of the human body. The library provides high-level relational attributes, an "electronic tape measure" to extract linear and curvilinear measurements, as well as low-level shape information, such as curvature.

Application of the library is demonstrated by building a large set of detectors to find anthropometric features, based on the ISO 8559 specification. Output is compared against traditional manual measurements and a detailed analysis is presented. The discrepancy between these sets of data is only a few per cent on most dimensions, and the system's reproducibility is shown to be similar to that of skilled manual measurers.

The final contribution is that the mesh models and anthropometric features, produced by the system, have been used as a starting point to facilitate other research, such as registration of multiple body images, draping clothing and advanced surface modelling techniques.

Acknowledgements

I would like to thank Dr Jason Kingdon, for too many reasons to list and Dr Zeluiz Ribeiro-Filho for inspiration to start. I would also like to thank Professor Philip Treleaven for making it possible for me to do this work and for supporting me throughout; Professor Bernard Buxton for his incredibly meticulous feedback; Dr Janet McDonnell, John Taylor, Dr Andrew McCulloch, Cecelia McCulloch, Jungwon Kim and Kanta Vekaria, for reading whole and parts of drafts, and providing invaluable advice and encouragement. Thanks also to my two examiners, Dr Peter Burger and Professor Peter Jones, for their detailed and very constructive comments.

I gratefully acknowledge the help and expertise of many other people: Jeni Bougourd, Pip Bull, David Carruthers, Elena Djendova, Ioannis Douros, Nigel Fuller, Surayya Khan, Asako Kohno, João Oliveira, Amani Siyam, Janet Ward, Dr Jonathan Wells and Diccon Yamanaka.

Special thanks are due to Hamamatsu Photonics—in particular, Chiyoharu Horiguchi—for providing the hardware that made this work possible, and for their very kind support. The volume assessment trials were funded by the Wellcome Trust and were carried out by Ioannis Douros, Nigel Fuller, Dr Jonathan Wells and others at the Dunn Nutrition Unit. I'd like to thank the many representatives from Bhs who provided their expertise in anthropometry for clothing design.

Finally, I'd like to thank colleagues at UCL, friends and family for their encouragement and support, and all the volunteers who donated their virtual bodies to research.

Table of Contents

List of Tables and Figures

Chapter 1

Introduction: the Digital Human

Chapter 1 presents the background and objectives for this thesis. It begins with a brief history of human body modelling, from traditional approaches to recent developments in 3D digitisation technology. This background reveals the motivation for the work presented in this thesis: the development of an automated system for feature detection from whole-body image data. The final section provides an overview of the thesis—the research goals, the thesis structure and its research contributions.

1.1 A Brief History of the 3D Digital Human

This thesis is about building 3D digital models of people from opto-electronic range data. It explores automated techniques to take raw range data of the whole body, then clean, segment and skin it, and detect "useful" symbolic features for applications, such as medicine, surface anthropometry for clothing design and entertainment.

1.1.1 Motivations for Digital Human Models

Recent developments in three-dimensional surface imaging techniques [Davd98] present enormous potential for building numerically and perceptually accurate 3D digital models of the human body. Such models have wide-ranging applications [JoRi97]. For example, the human body's size and shape, in all its variations, drives the ergonomic design of car interiors and the workplace [Seit98]. Such information can be used to improve comfort, fit and drape [Gray98] in made-to-measure and off-the-peg clothing [Wink97]. In medicine, detailed three-dimensional body data is used in prosthetics design [CoSV94], to determine correct doses of medication [LaST97], and in monitoring posture and growth [Bato92] [SDNV90]. At the other extreme there is intense research and standardisation activity [HAWG99] [Iann99] [Koen99] whose goal is the construction of *dynamic* three-dimensional digital models that incorporate sufficient

physiological and behavioural realism to be believable, as individually recognisable personalities [FGPA98] [HoWo98]. In design and retail, animated digital humans are being explored for virtual clothing design and catwalks [WeMT93] [ADBW96], fitting rooms [Gray98] [TrBS98] and as virtual shopping advisors who help to find and select items on the Internet [ChJo99]. Figure 1.1 summarises requirements of a digital body modelling system for some of these applications.

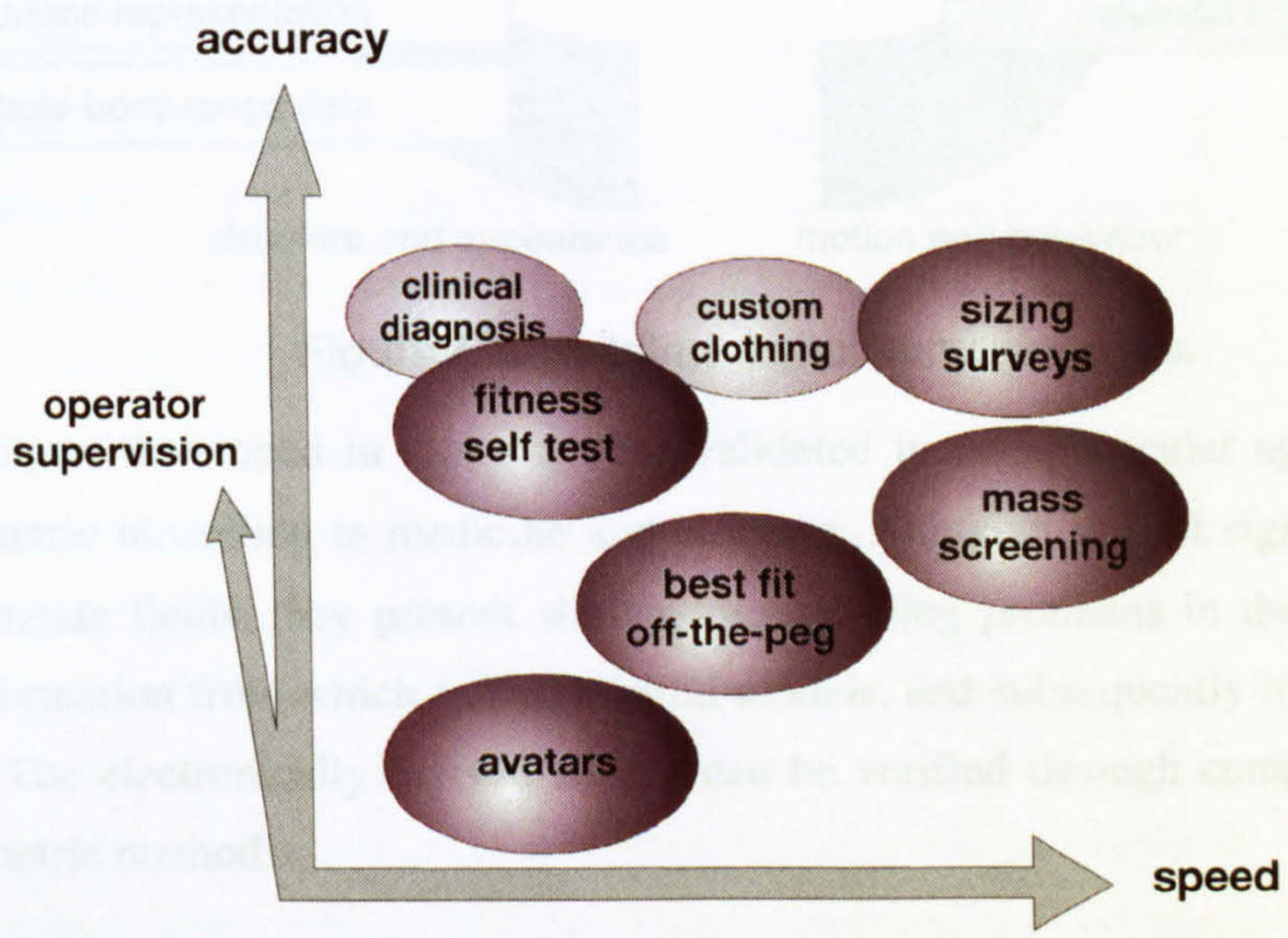


Figure 1.1: Requirements for some applications of 3D body scans.

However, the development of digital human models for these applications is typically meticulous and labour-intensive. In order to exploit the full potential of 3D images in creating personalised models on a large scale, the data must be processed *automatically*, from a raw, *unstructured* state to a form that contains *explicit knowledge* about the human body. Information of interest includes labelled body segments such as the legs, arms, head and torso; anatomically meaningful contours include closed curves through hip sections and open curves across the shoulders; point landmarks include specific vertebrae, and surface features near joints used in anthropometry. Extracting such information in an automated way is the focus of this thesis.

In the broader context of developments in synthesising and analysing virtual humans, this thesis is concerned with the stages from 3D image capture through to the detection of surface features for use in applications, e.g., those defined in the humanoid animation standard, H-Anim [HAWG99]. The transformation and construction of the many "layers" of structure, appearance, motion and behaviour are represented graphically in figure 1.2. This thesis addresses the unshaded layers, at bottom left, which integrate with, and facilitate, other higher-level stages of modelling. This work does *not* attempt to address issues such as extreme realism for visualisation, modelling dynamic properties, or the drape and synthesis of clothing on 3D models.

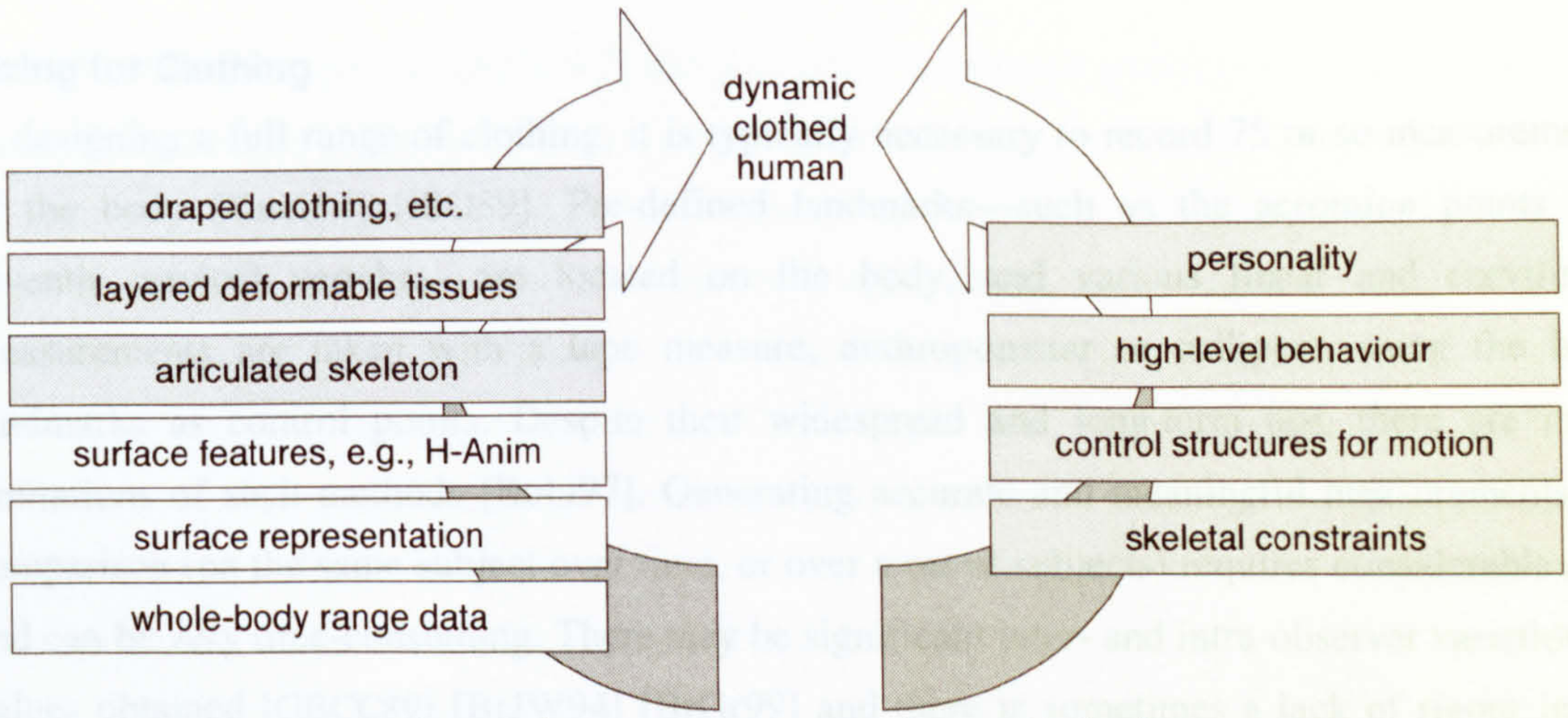


Figure 1.2: Building digital human models.

The techniques developed in this work are validated in two particular application areas—3D anthropometric modelling in medicine and clothing. Although at first sight these appear to be rather disparate fields, they present similar long-standing problems in the capture of accurate spatial information from which to build useful models, and subsequently in the interpretation of this data. The electronically derived results can be verified through comparison with existing anthropometric methods.

Medicine

Body size and shape have numerous implications in the context of human medicine. Surface anthropometrics, such as biceps girth, have an important role in diagnosis and monitoring, for example, in studying fat patterning and nutritional health. The spatial relationships of landmarks, and the description of curvature are important in large-scale screening and monitoring of conditions such as scoliosis [Bato92] [SDNV90]. Body volume is used to calculate density, which, in combination with other techniques, can be used to derive body composition (fat, fat-free mass, total body water, protein and body mineral) [Siri61] [FJLC92], which can be used to determine energy and fluid requirements. Body surface area [dBdB16] is used to assess the dose required in radiation therapy, drug dosage [LaST97], burns treatment [LuBr44] and to estimate metabolic rates. Although simple in concept, volume and surface area are extremely difficult to measure *in vivo* in the majority of human subjects. Body volume is generally measured by weighing the subject underwater and subtracting the measured lung volume using a gas dilution technique [FJLC92]. An alternative method, whole-body plethysmography [MGBM95], estimates body volume from air pressure changes within a capsule. Although these are considered to be the most accurate methods for measuring body fatness in fit and willing volunteers, they are sensitive to many operational factors, and quite intrusive on the subject. Such methods are therefore unsuitable for unwell patients and for the very young or old.

Sizing for Clothing

In designing a full range of clothing, it is typically necessary to record 75 or so measurements of the body [Kuni84] [ISO89]. Pre-defined landmarks—such as the acromion points and seventh cervical vertebra—are located on the body, and various linear and curvilinear measurements are taken with a tape measure, anthropometer or callipers, using the body landmarks as control points. Despite their widespread and long-term use, there are many limitations of such methods [Robi97]. Generating accurate and meaningful measurements for comparison (on the same subject over time, or over a set of subjects) requires considerable skill and can be very time-consuming. There may be significant inter- and intra-observer variation on values obtained [GBCC89] [BrJW94] [BrGr99] and there is sometimes a lack of rigour in the way that procedures are defined, leaving it up to the observer to draw their own interpretations. Despite the existence of measurement standards such as [ISO89] [Kuni84] [SAE99], different organisations often use their own procedures [Wink97]. Tailors can quantify three-dimensional information, such as curvature and spatial relationships of body features (so-called body "attitude" and "figuration"), by eye or by using adapted protractors, to improve fit and drape for made-to-measure clothing. A major drawback is the high level of skill and experience required in making these analyses. These problems make the collection of such data very costly, prohibitively time consuming, potentially error prone and inconsistent. All this is amplified when attempting large-scale surveys of the population and, not surprisingly, there is a scarcity of such extensive up-to-date information. Many designers and manufacturers still base their products and sizing categories on data collected in the 1950's [Kuni84]. It is well documented, though [HaVS96], that size and shape throughout the world have changed significantly since then, through changes in diet, work practices, exercise, and so on.

Consequently there is an enormous incentive to develop techniques for collecting and processing accurate, consistent and fully three-dimensional data of the human body, rapidly and with minimum intrusion on the subject. If suitable methods can be devised, this potentially opens up new fields of study, with a new level of accuracy and consistency not possible before.

1.1.2 Recent Developments in Human Body Data Capture

The past few decades have seen a great deal of research in the capture of accurate 3D data of the human body with which to build digital models [JoRi97]. The progress from one- to three-dimensional non-contact techniques is charted in table 1.1. This problem is similar to that of reverse engineering of manufactured items, but many characteristics of the *living person* make three-dimensional image capture particularly problematic [Riou98]. The human body is a flexible, deformable object, constantly in motion. With a living organism it is not easy to control the conditions under which image capture takes place, and it is not ethical to subject people to dangerous operations.

The Visible Human project [AcMa92] has provided the research community with an exceptional set of data of the total human body, most notably the two bodies transformed by cryosection and 2D colour image capture. Clearly, the destructive acquisition process precludes its use on living humans. Methods such as magnetic resonance imaging (MRI) and computed tomography (CT) have been used for several decades to provide detailed three-dimensional information, but they present several obstacles to widespread use. The image capture times are typically many tens of minutes; the data acquired is excessive for the purposes of surface modelling, and the deformations resulting from the prone scanning position make the data inappropriate for accurate surface anthropometrics and visualisation.

Attempts at capturing *surface* information include electronic spatial sensor pens, for example, mounted on a 3D pantograph [SaRS91], which can record accurate locations of selected points on the body surface. While these methods can be effective on static objects, the time-consuming individual sampling of points makes this an unrealistic method for capturing sufficient data about a living human. "Pin boxes" with arrays of displaceable rigid rods have been used for high density sampling of surface contours, and can be connected to devices to store the data directly in electronic form. These methods can capture three-dimensional information in an electronic form, facilitating analysis, processing and archiving. In addition, they reduce the potential for human error, but they are too slow for large-scale use, and constitute an unacceptable intrusion on the subject because of their use of full contact with the body.

Table 1.1: From one to three dimensions in human body data capture.

Computed tomography	+ -	Internal volumetric 3D; can differentiate between tissue types, disclosing internal skeletal information. Slow, expensive, intrusive.
Magnetic resonance imaging		Internal volumetric 3D; can differentiate between tissue types. Slow, expensive.
Non-contact opto-electronic surface imaging	+ -	Fast, relatively unintrusive, high resolution. Captures surface information only, some occlusion, expensive.
Contact 3D spatial sensors	+ -	Devices with fast response time can track points in motion. Low-density data, slow.
Contact "pin-box" methods.	+ -	Can provide medium-density output in electronic form. Difficult to capture whole-body information, intrusive.
Hydrodensitiometry and plethysmography	+ -	Accurate methods for determining body volume. No discrimination of shape.
Silhouette/2D imaging	+ -	Fast. Not truly 3D unless multiple images combined.
Callipers, tape measure, anthropometer	+ -	Low-cost, easily accessible tools, directly extract measurements. Slow, only 1D information.

Equivalent opto-electronic methods present numerous potential benefits. More rapid image capture is possible, with current sampling rates of approximately 10^5 or more points per second.

This translates to less than a second [SiPa98] to ~17 seconds [Cybe93] to capture a whole-body image, depending on the optical technique, the resolution and macro-level design of the imaging system. Because the methods do not require contact, they can be used on sensitive subjects, for example post-burns patients [WhGM97]. Some work has been carried out to reconstruct dynamic 3D models from orthogonal 2D data, for ergonomic [Seit98] and anthropometric applications [Telm96] [Gray98] and to build *H-Anim*-compliant avatars [HAWG99] for virtual environments [HBGS99]. These methods are attractive in that they use fast, low-cost imaging components (often simply a single video camera), but inevitably there is some loss of information because they only capture 2D projections of a 3D form.

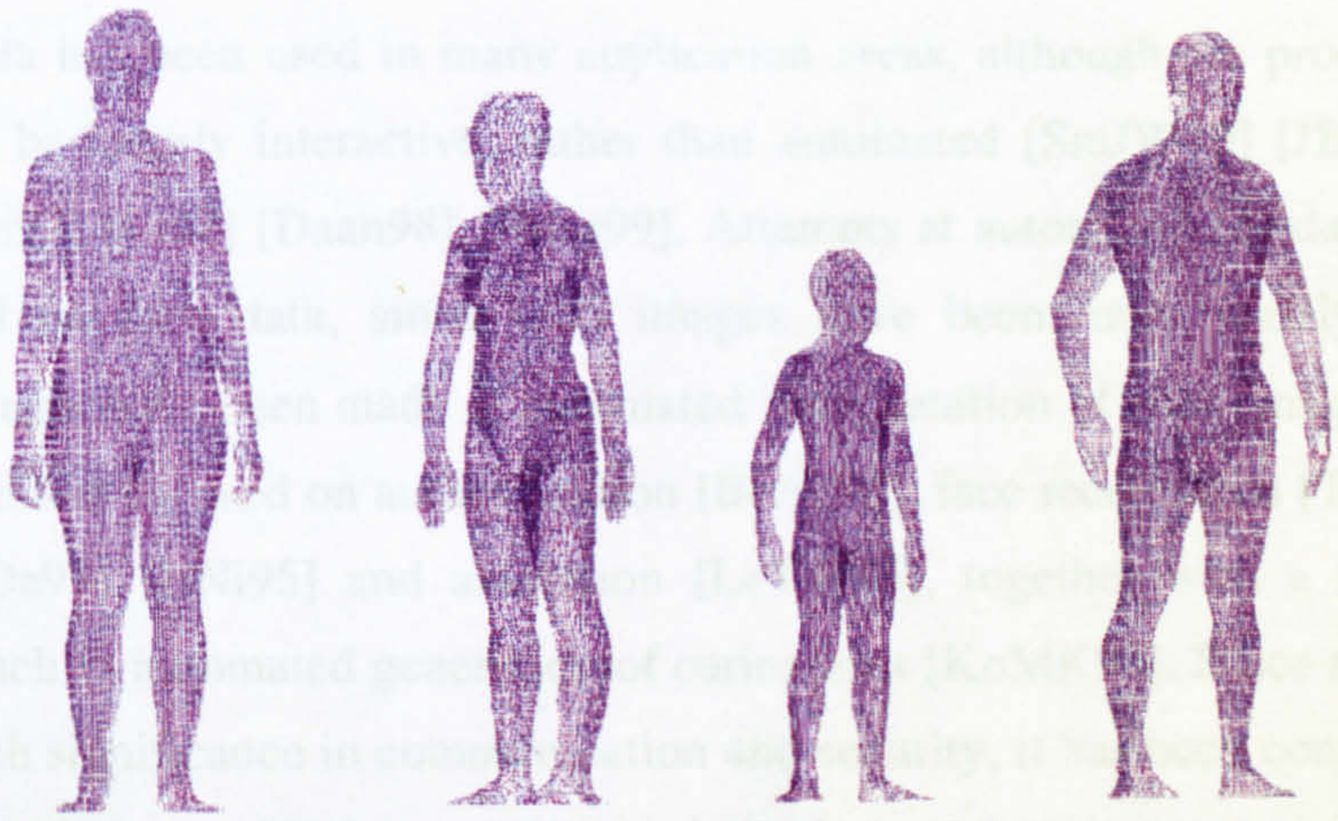


Figure 1.3: Typical whole-body scans.

The first full 3D systems were developed for parts of the human body, for example, the head [LTRC92] [WhGM97] and the torso [Bato92] [SDNV90]. Specific issues of *full body* 3D surface imaging are only more recently being addressed [JWHR89] [BrDR97] [DaBR97] [CoHu98] [Riou98] and are discussed in more detail in the next chapter. Such issues include occlusion because of the body's many degrees of freedom, the large imaging volume and fast capture times required. Detailed surface information can now be captured, of the order of 100,000-6,000,000 sample points over the whole surface of the body, with 1-7 mm resolution, at a high accuracy of ± 1-2 mm for each sampled surface data point. Figure 1.3 shows some typical 3D data sets as point clouds, captured using the Hamamatsu Photonics system [Hori95]. Some systems can also capture full colour texture data [Cybe93] [SiPa98] [StMi98] [InSp99], of particular interest for visualisation applications. Suitably processed image data can be stored in a compact digital form (a few KB [LiJo94] to a few MB [Riou98]).

1.1.3 Human Body Surface Image Processing

Having the whole body shape in a high-density, digital form has numerous important implications. One of the key advantages of archiving electronic images is that they can subsequently be revisited repeatedly and processed by increasingly sophisticated processing techniques and for multiple applications. Perhaps of most significance is that this rich source of

size and shape information can be queried in many different ways. For large-scale surveys this is particularly important. The images can be captured, and the choice of which information to extract can be made at a later date (though limited to some extent by the posture of the subject during image capture).

However, before the information of interest can be extracted from the data a great deal of processing is required. The whole-body range data tends to be noisy and sparse in areas where occlusion occurs [BrDR97] or where the body surfaces are almost parallel to the viewing axis. These characteristics are important in designing processing techniques specifically for this type of data.

Whole-body data has been used in many application areas, although the processing techniques have tended to be largely interactive, rather than automated [SmJW90] [JBHM94] [BrJW94] [Vann97] [Burn97] [Li97] [Daan98] [BrGr99]. Attempts at automated 3D data processing have mostly focused on face data, since such images have been more widely available. Many successful attempts have been made at automated interpretation of such images, both in 2D and 3D, with most effort focused on authentication [BeAc98], face recognition [TuPe91] [CCRA96] [BrPo93] [YaDa93] [JiNi95] and animation [LeTW95], together with a few more esoteric applications, such as automated generation of caricatures [KoMK97]. Since the human face and body are of such significance in communication and security, it has been considered worthwhile to hand craft techniques for interpreting such images and even to process single instances manually, given the still novel nature of 3D whole-body digitisation and its applications. Often manually generated feature extraction algorithms are combined with a statistical or learning technique [JiNi95] and more recently fully generic techniques have been used [WaCT98] [SaCi97]. Because of the similarities with the whole human body image understanding problem, many of the techniques used in analysis of face data, in 2D and 3D, will be discussed in subsequent chapters.

More recently efforts have been made to develop completely automated techniques for whole-body surface data. Li and Jones developed semi-interactive and automated techniques for building a fitted "skin" to their surface data [LiJo94] [LiJo97] with anthropometric applications in mind [JoLi98]. In postural diagnosis, in particular for scoliosis, surface scanning methods have been used [SDNV90], while [Bato92] presents a knowledge-based system for diagnosis from surface data. Geisen *et al* [GMHW95] and Lewark and Nurre [LeNu98] have developed techniques to detect manually placed optical fiducial landmarks from body images, using a Bayesian approach to identify each landmark. This method is being used to process large-scale data collected in the US and Europe [Robi98]. However, it requires rather intrusive pre-placement of markers against a background of standardised uniformly coloured clothing and may be prone to slight inconsistencies, as with traditional manual measurement practices. Therefore it is desirable to pursue methods that do not require pre-marking. Nurre [Nurr97] has

developed algorithmic methods to segment the body into primary components—the head, torso, arms and legs. Pargas *et al* [PaSD97] have developed a macro language that allows the user to define their own algorithmic methods for locating and measuring body features from 3D image data. This approach is demonstrated using moiré fringe data [DHWG97]. More recently Stuetzle and Certain presented work on a proposed low-cost image capture system for the lower body, where a surface model is fitted to patches of surface data, and anthropometrics are extracted automatically [CeSt99].

Despite this work of direct relevance, there is still a lack of demonstrably robust, automated methods for processing whole-body surface data. In particular, there is no method that can easily be adapted to locate new features of interest, without specific reprogramming effort. From a more abstract point of view, the problem is to transform the raw image, containing a vast amount of *implicit* information, into forms with knowledge held *explicitly*, as illustrated in figure 1.4. Framed in this way, the problem is part of a much larger research field of machine vision. In purely research terms, it is challenging to build a system that can deal with the general problem of feature detection on free-form surfaces, and to find what level of automation is possible.

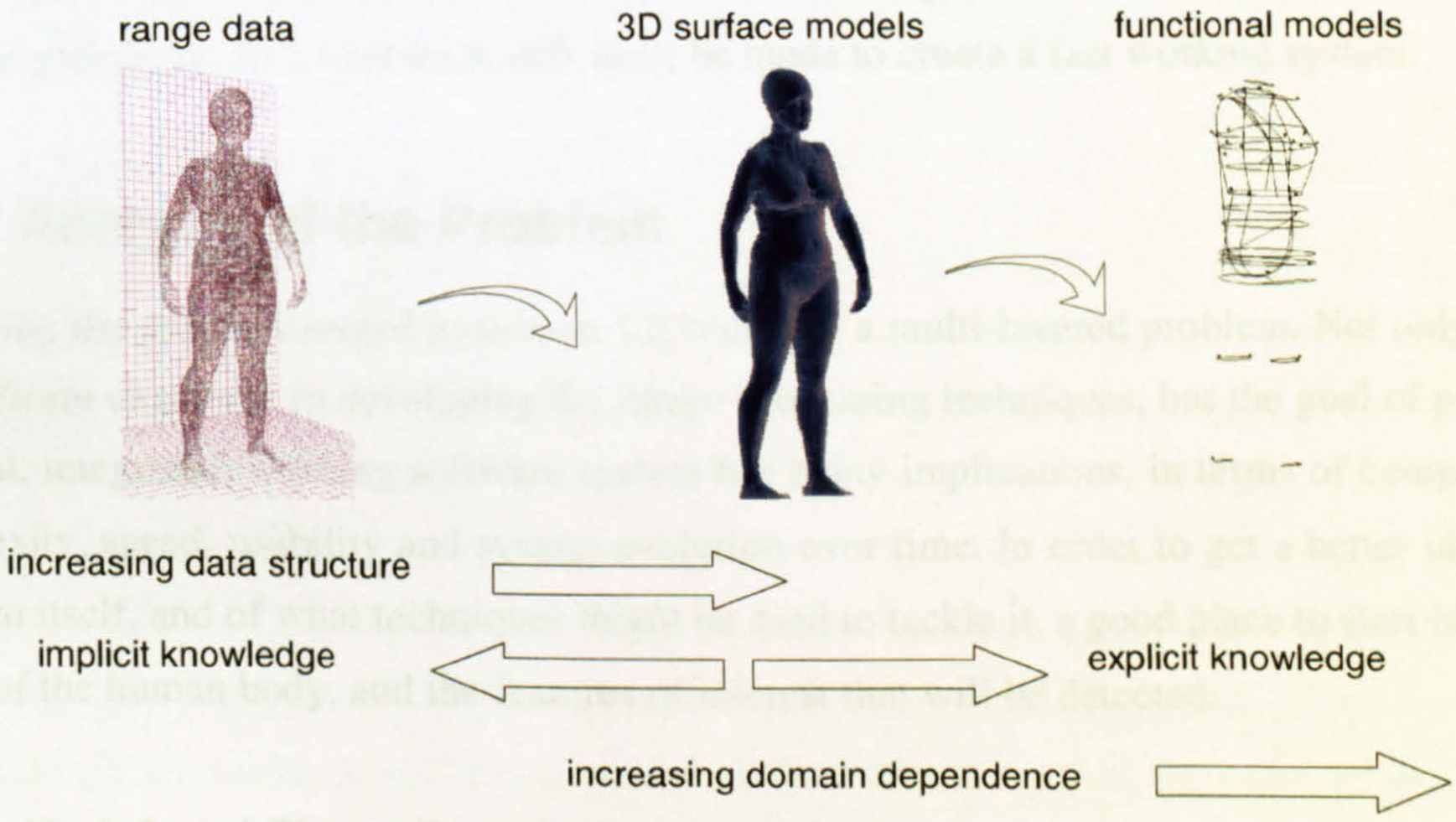


Figure 1.4: Building symbolic information from human-body range data.

If robust and accurate techniques can be devised to locate and label body features, then this provides a basic pre-processing tool for analytical techniques and future applications. Point landmarks can be used as markers for surface anthropometry [BDGW00]. Furthermore, software incorporating these techniques can incorporate a given standard to ensure consistency of measurement and analysis. At a lower level, symbolic information such as segment labelling can be used to determine parameters for cleaning, smoothing and surface reconstruction [DDBT99]. In terms of numerical analysis, sets of homologous landmarks can be used as registration points, for shape warping [Book89], for Procrustes alignment [Gowe75], or for training Point

Distribution Models [CTCG95]. In the development of virtual animated humans, surface features can be used to infer joint locations for an articulated skeleton, and ultimately for animation [HAWG99] [Carr98] [HBGS99].

Thus, there is clearly a need for further research, and it would appear to be necessary to look further afield at other machine vision techniques that might be adapted for this problem. Some machine vision issues are briefly discussed later in this chapter, in section 1.3.3, and potentially useful techniques are investigated in Chapter 3.

1.2 Research Goals

The goal of this work is to develop techniques to take 3D whole-body opto-electronic range images and transform them into models containing explicit symbolic information. A key part of this work is to engineer and build a fully automated system that takes as sensory input the raw range images, and which outputs surface models with features of interest labelled, from which relevant body metric information can be extracted. Two test problems will be used for validation: volumetry [FJLC92] [MGBM95] and surface anthropometry [Kuni84] [ISO89]. An issue of particular interest is the use of *a priori* knowledge, how constraints can be relaxed to increase generality, and what trade-offs must be made to create a fast working system.

1.3 Analysis of the Problem

Achieving the goals presented in section 1.2 is clearly a multi-layered problem. Not only is there a significant challenge in developing the image processing techniques, but the goal of producing a useful, integrated, working software system has many implications, in terms of computational complexity, speed, usability and system evolution over time. In order to get a better idea of the problem itself, and of what techniques might be used to tackle it, a good place to start is with the shape of the human body, and the features of interest that will be detected.

1.3.1 High-Level Shape Description

Although many characteristics of size and shape are shared across people, there is also enormous variation. The key to this problem is to characterise this variation and invariance, so that the features of interest can be detected within noisy and sometimes sparse range data. Several sources of information can help in a first analysis of this problem: the surface shape of living human volunteers; anthropometric specifications [Kuni84] [ISO89] (see Appendix A and B); and observations of, and discussions with, expert anthropometricians. For the moment, the focus is on the human body itself, independently of the context of any particular data capture medium, except to focus on surface geometry, ignoring tactile characteristics, colour, texture,

and so on, since the primary source of modelling data will be whole-body opto-electronic range images.

The body is a branched structure, composed of smooth surfaces separated by points of discontinuity, for example at cusps such as the crease of the underarm. Much of the body is approximately elliptical in cross section (parallel to the floor, in the transverse plane), becoming more complex at branching points and through areas such as the buttocks, the chin, and bust on women. Variability in the surface characteristics arises from differences in fat, bone structure, musculature and posture. Heavier deposits of fat and extreme muscle development give rise to new features where abrupt changes of curvature and folds occur. Features such as the knees, supra-clavicular notch or ankle bones may either be rather prominent, or almost indistinct in terms of surface shape.

A completely generalised method would take as input a 3D image of part or whole of the human body in an arbitrary posture and detect the appropriate features. To reduce the complexity of the problem it is necessary to lose some of this generality, and assume that the image is captured in a given reference posture. By relaxing certain restrictions and losing the corresponding assumptions, it may be possible, in the future, to increase the generality of the model, but for now it is necessary to reduce the problem to a tractable starting point.

1.3.2 Target Features

So far the discussion has dealt with the body at a relatively high level. It is also necessary to explore the particular characteristics of the types of target feature of interest (as described in Appendix A and B). To be reliably detectable, landmark features should be *reproducible* across the samples, within the working medium [WaCT98] [Book91]—the imaging or measurement technique. For example, in traditional anthropometrics skeletal features are appropriate as landmarks since they can be found reliably by tactile feedback [Kuni84]. In the case of opto-electronic range data, ideally each feature would correspond to some distinct and analytically detectable characteristic of surface geometry, such as a saddle point, or a sharp valley line that can be represented by transformations of the image data. Many researchers have contributed to work on description of surface and contour shape, in particular for 2D and 3D object recognition, and this topic is discussed in more detail in Chapter 3.

Bookstein adds the requirement that the choice of landmark features should encourage useful subsequent analysis [Book91]. In other words, such landmarks should have relevance for the particular analysis method or application. Recognition merely requires that a particular example be reliably discernable from similar, but different instances, without any particular requirement to explain why particular features are discriminatory. Conversely, Bookstein relates the choice of appropriate homologous landmarks to functional form and therefore their implications in the

context of developmental anatomy and evolutionary change [Book91]. The same is true for other methods that attempt to characterise variation amongst classes of objects [Gowe75] [CTCG95].

Unfortunately many of the standard landmarks for anthropometry [ISO89] relate to the skeleton (e.g., the seventh cervical vertebra, trochanter and acromion), and are not necessarily visible on the skin surface by direct spatial analysis. This is an important difference from many other morphometric problems, where the choice of features is left to the analyst, or to some automated "saliency" measure [WaCT98]. In this case the anthropometric requirements make the problem more challenging. Despite this, it is possible that other reliably detectable features may prove useful in locating the specific pre-defined features. In addition, what is apparent from this initial analysis is that most features of interest are context dependent. They are either *defined* by their context (e.g., the calf is the maximum girth parallel with the floor on the lower leg), or setting them in a relational context would be expected to *reduce the complexity* of the operations needed to locate them. This analysis leads to the following classification of target features:

- *Intrinsic surface features*, for example, "pits", such as the navel;
- *Secondary features* are based on intrinsic surface features, and may be found either directly or indirectly from them (e.g., the "upper hip" is defined as the girth, one quarter of the way from the waist to the full hip level);
- *Non-surface features* include many of the predefined landmarks of interest, where there are no reliable intrinsic surface shape characteristics. Most are skeletal features (e.g., the seventh cervical vertebra).

This makes it possible to order features with decreasing level of intrinsic surface characteristics, and a corresponding increase in the expected difficulty in their detection from surface geometry.

1.3.3 The Machine Vision Context

In first approaching the feature detection problem it is useful to look at other machine vision issues, in whatever medium or dimension, to fix this work in context and to appreciate why the problem is hard. Examples abound, from closely similar problems in detecting natural surface features in 2D and 3D, in particular on the human face [TuPe91] [Saly92] [KaCH93] [JiNi95] [OsFG97], to less obviously related problems such as the interpretation of hand-written musical scores [NgBo96]. These are all concerned with recognising discrete objects or features from a physically sensed signal, such as light or sound. Interpretation of the image typically takes a bottom-up approach, following Marr's ideas of the early visual processes [Marr82]. First low-level sub-features are extracted and classified, and then specific domain knowledge is used for higher-level construction and interpretation of features. Contextual information is often used to reduce the search space and to resolve ambiguities.

The first problem is to capture the whole-body range data. It is possible to control the conditions for this, to some extent, by specifying a reference posture for subjects. In addition, it can be assumed that only the subject of interest is contained in the image (albeit possibly with some noise and areas of occlusion). This prior knowledge can be used to make certain assumptions in the processing, such as the topology of the subject. Unlike many 3D image processing problems it is assumed that the 3D coordinate data is contained either explicitly in the input data, or is fairly straightforward to generate, i.e., that it is not necessary to generate it from other 2D data. It is expected that the data will contain some noise and be incomplete in places, in particular, if occlusion occurs [BrDR97]. Dealing with variable resolution and noise are major concerns in processing range data [Besl88] [HDDM92] [LeHD93] [StIW95] [GuMe97], and will be explored again in subsequent chapters. For example, with very high-resolution data and low noise, subtle features, such as the crease at the back of the knee, might be detected. However, if the sampling resolution is low, or the data is very noisy and requires smoothing, then such fine detail will not be observed. This problem of appropriate scale is encountered in many data interpretation problems, and must also be addressed in this work.

However, since the data space contains only the whole body in three dimensions, there is no concern with clutter, changes in illumination, or viewpoint variation in shape and visibility of features (except specific ones that tend always to be occluded, such as the underarm points). In addition, it is possible to know in advance what features are present, and in what quantity. Of course, this will not always be a valid assumption (for example, on subjects with limb loss, or where features are hidden by clothing).

A particular difficulty with human-body feature detection, mentioned previously, is that many of the target features are extremely subtle. Few of the target features are intrinsic *surface* features familiar in differential geometry, such as pits and saddle points. The surface cannot be broken down into simple geometric components—as is often the case for manufactured objects—and still retain its usefulness for feature detection. In addition, there is a great deal of variation in the local shape and location of features amongst subjects, even those with superficially similar size and shape, giving continuous variation within a feature. This variation and the invariances of each feature must be captured somehow within the operators that process the 3D image data.

The fact that the input data corresponds to a sampling of a single continuous surface poses certain difficulties. In a straightforward representation of the surface in Euclidean space, the features of interest are not discrete objects, and it is not immediately clear how a subspace of the data would be selected as input to model each feature. Ideally there would be a one-to-one correspondence between the input attributes for each data set. This would facilitate the use of learning techniques such as neural networks [RuHW86] and support vector machines [BoGV92] that expect uniform length input vectors, usually with some semantic correspondence between each *i*th element. However, there are examples where such techniques are used with non-

corresponding data [OsFG97] [TuPe91]. In the initial image data used in this work there is no particular relationship between the elements in each set. One way of dealing with this would be to extract a vector of features from the raw data, and use this as input. This also has the advantage of reducing the dimensionality. The problem then is to find an effective method to transform the input image into a useful representation (for example, as discrete features). One possibility is to extract key surface points and contours by virtue of their local shape characteristics [AsBr86] [Besl88] [KovD92]. In the case of the human body, it is clear that such local information alone would not suffice for detecting features, since features with similar local characteristics appear in many places over the surface of the body. Consequently, it will be necessary to make use of relational information for feature detection.

So far the problem has been discussed in the context of bottom-up signal processing methods. Clearly there are many other possible approaches, for example:

- a statistically-based deformable whole-body template, with pre-affixed landmarks, to which input data is matched;
- several filters or local templates, possibly one for each feature;
- a system that uses a form of "grammar" and data parsing to detect features in context;
- hierarchical context-based searching functions;
- a combination of techniques.

However, it is not appropriate to be too concerned yet with the particular technique for implementation. For example, a neural network could be used to build a parameterised surface patch template, or a discriminant function for classifying features; a genetic algorithm or other general-purpose search method could be used to optimise the matching process, or to synthesise a shape grammar, and so on. These issues will be discussed again in more detail in Chapter 3.

1.3.4 System Input for Training and Run Time

To train the system, various types of data may need to be collected, in addition to the whole-body range images:

- reference data ("ground truth") for the information of interest: locations of landmark features, the paths and lengths of contours [Kuni84] [ISO89], and volume measurements [FJLC92] [MGBM95];
- other information about the subject, such as weight, gender, "handedness" and other anthropometric and socio-economic information;
- knowledge, for example, from expert anthropometricians, about how they perform anatomical labelling and measurement tasks [Kuni84] [ISO89].

It might be legitimate to use any of these data sources in *developing* the image processing operators. However, since the system should be a closed loop at *run time*, it would be desirable to make use only of the range image of a subject and avoid, or at least minimise, the use of additional information.

At the start of this work no large-scale database of 3D whole-body data was in existence. Subsequently, within the CAESAR initiative [Robi98] [SAE99], body data is being collected, first in the USA, then in the Netherlands and Italy. Approximately 9000 subjects in total are being scanned in three postures, with pre-affixed landmarks manually placed at key features, such as the iliac crest. However, this data is currently accessible only to members of a closed consortium, so for this work it will be necessary to collect alternative sets of digitised body data, together with reference data. Collecting each sample, together with reference data, will be time consuming. Therefore the training set will be rather small, of the order of 50 data points, which will necessarily place limits on the type of technique that can be applied.

1.4 Assessment Criteria

In order to assess the methods developed in this work, it is important to list some more specific objectives, which will be revisited in Chapter 7. Some of the criteria listed below can be approximately predicted in advance, and can therefore be used to guide the choice of method, as discussed further in Chapter 3, when the overall approach to the problem is explored. Other criteria can only be assessed by testing a given implementation, and therefore provide a basis for designing the run-time tests.

- *Minimal intervention.*

It is desirable that image capture be as fast as possible and cause minimal intrusion on the subject. Consequently the aim will be to derive as much as possible from a single 3D image, and to avoid the use of any external markers on the body, or extraneous data. The execution of the processing techniques should also require minimal intervention; in other words, they should be as automated as possible. This may necessitate an explicit compromise on other performance criteria, such as accuracy. The advantages of rapid, non-contact image capture and automated analysis are such that some error, when compared with existing methods, would be considered acceptable for certain applications. However, if suitable fully automated and accurate image processing techniques can be devised, then this could remove human error, and potentially produce systems with higher accuracy and reproducibility than is obtained even from manual measurements by skilled anthropometricians.

- *Geometric accuracy.*

The whole-body representation and the anthropometric information extracted from it should be geometrically accurate. The accuracy will be determined by comparison with reference data

collected by alternative means. Precise requirements will be determined by consultation with practitioners in the field. Since there are potentially errors in all reference methods, it is expected that difficulties will arise in making comparisons. It is important that at each point the sources of error are isolated and, as far as possible, tested separately (e.g., noise in initial image data, artefacts of pre-processing, errors in several stages of feature generation and measurement, errors in training data, observation data, etc.).

- *Occam's Razor: "It is vain to do with more what can be done with fewer".*

All things being equal, the simplest technique should be used. This is of interest as a research strategy, but also to encourage computational efficiency. From image capture to visualisation and output of the information of interest, the processing must execute in a matter of a few (perhaps one or two) minutes on a moderately powerful personal computer. The simplest method will first be applied, in order to assess what kind of results can be achieved. More complex methods will only be used where they can be justified in terms of improved accuracy or robustness. This has particular significance for the use of prior knowledge: if the use of a constraint or assumption improves the time complexity significantly, and does not degrade performance, then it can be used. It is always possible to relax such constraints later.

- *Reproducibility.*

Results should be reproducible on a given individual, and be robust over a range of subjects and operating conditions. "Robustness" is used to cover the many issues relating to consistent performance over a specified range of conditions: tolerance to noise; variations in lighting conditions, body type (women, men and children of varying size, shape and age); the density of image data; and the location, orientation and posture of the subject during image capture.

- *A flexible, working software system.*

The central part of this work is the development of a working software system, which should be engineered in a modular way to enable each component to be replaced by alternative or improved techniques, and to allow its use with different image capture hardware. In addition, the system should be extensible, so that new components can be added to extend the functionality, or exploit it in a new way.

- *Broad applicability.*

The information generated by transformation of raw surface image data should be suitable for use in successive phases of human body modelling, for example, animation. Perhaps the most challenging goal is to develop a framework or methodology that is not simply hard-wired for a particular application, but can be used to detect a range of features. Thus it will not be sufficient to build a system that works only for a given instance, but is demonstrably applicable in a broader sense.

- *Emergent properties.*

Rather than simply simulating existing anthropometric procedures, it is desirable that the method should reveal interesting properties of the human body of its own accord.

1.5 Research Contribution

In addition to the immediate and proven benefits to the target applications areas of clothing design and manufacture, and medicine, the work makes five main research contributions:

1. The thesis describes the design rationale, development, testing and assessment of a fully automated system for interpreting and visualising whole-body range images. This is arguably the first fully integrated and automated system of its type, and the first reported in detail in the public domain.

2. A new, combined data-driven and model-based approach is described for low-level processing and surface reconstruction specifically for human-body range data. The techniques address the whole body surface and deal with occluded areas. Tests are carried out on a range of body shapes and sizes, and the processing techniques are demonstrated to be robust for women, men and children.

3. A library of computationally-efficient operators is developed and tested, designed specifically for human-body shape description and measurement on surface models from range data. The library includes a simulated tape measure to extract curvilinear measurements, and operators to extract curvature maps of the body surface and to describe relational information.

4. The library is used to build a set of a wide range of feature detectors for surface anthropometry, according to the ISO 8559 specification. The automated system can detect anatomical features, such as the seventh cervical vertebra, the navel and the knee joints. The system is tested on a set of women, by comparing expert manual measurements against the automated output from scans. The results are discussed in the context of specific uses of the automated system, and are shown to be comparable to similar systems that require considerable manual intervention in image capture and processing. The reproducibility is shown to be similar to expert manual measurers.

5. The system and its outputs have been used to facilitate research by others in digital human modelling, body composition and clothing technology. As such the system has become a basic tool for research problems, such as registration of multiple body images, draping clothing and advanced surface modelling techniques.

1.6 Thesis Organisation

This thesis is organised in eight chapters, with additional information provided in appendices, where it would otherwise break the flow of the main text. Survey material is presented in Chapters 2 and 3. Experimental work is presented in Chapters 2, 4, 5 and 6. Chapters 7 and 8 provide an assessment and conclusions. A brief overview of each chapter follows.

Chapter 2: Whole-Body Image Acquisition

Chapter 2 presents a survey and assessment of opto-electronic surface imaging techniques, in particular, systems for whole body imaging. After describing the system to be used in this study, some qualitative experiments explore the problems and conditions of capturing useful whole-body data of *living* humans, particularly in the context of automated feature detection. An analysis of these data sets provides the necessary starting point for choosing the overall approach to the problems tackled in this thesis.

Chapter 3: Techniques for Whole-Body Image Interpretation

Based on the analysis of the image data, Chapter 3 explores and assesses some computational techniques for shape description and shape matching. A progressive, procedural pattern recognition approach is chosen as a way of attempting to make the problem tractable and exploit prior information sources. To provide a flexible environment for experimentation and exploration, an interactive framework is designed, which interfaces with the image capture hardware, and provides facilities for 3D visualisation and analysis. The chapter concludes by outlining the issues for investigation. Each theme is tackled in the chapters that follow.

Chapter 4: From Points to Surfaces

The processes from raw data to surface reconstruction are explored. The chapter first discusses the choice of the generic whole-body representation, in the context of encouraging the discovery of the best feature detection techniques. A hierarchical compound mesh structure is proposed, segmented at a high-level. After visual assessment of the skinned images, they are validated by comparison of the enclosed body volume with hydrodensitometry and air plethysmography.

Chapter 5: From Surfaces to Features

Once the raw image has been converted to a surface representation it is possible to start analysing the surface information through various transformations in attribute space. Of particular interest are operations that can be used as discriminants for detecting interesting features. This chapter presents an exploration of some shape description techniques and an "electronic tape measure" to extract anthropometric information from the surface.

Chapter 6: Application and Evaluation

Feature detection operators for surface anthropometry are constructed and tested, based on the surface attributes explored in Chapter 5. A training set of scanned and manually measured

bodies is used to assess the accuracy of the computed values. A number of important issues arise in the generation and testing of computed measurements because of the many potential sources of error in obtaining validation data and in determining appropriate performance criteria.

Chapter 7: Assessment

Chapter 7 provides an assessment of the work carried out in this thesis, firstly in terms of the research approach, and an assessment of the system developed in this work, according to the criteria outlined in this chapter. Finally, the thesis is assessed in terms of its research contributions.

Chapter 8: Conclusions and Future Work

The thesis concludes, in Chapter 8, by summarising the main points of the thesis and discussing plans and opportunities for future work.

Appendices

Appendix A describes the anthropometric procedures, Appendix B provides a glossary of anthropometric terms and Appendix C presents results for the automated system.

1.7 Summary

This chapter has discussed some of the motivations for, and problems of, building accurate surface models of the human body. Techniques for capturing high-resolution surface images are outlined. While these potentially provide high quality raw data, they cannot be exploited to the full without appropriate processing techniques. In particular, automated feature detection and labelling is identified as a challenging research problem on which to focus, with numerous potential benefits for large-scale application and analysis of personalised body models. The chapters that follow present the design, implementation and testing of an automated system to perform these tasks.

Chapter 2

Whole-Body Image Acquisition

This chapter explores how 3D body images are captured. It begins by discussing the requirements of three-dimensional digitisation systems for capturing whole body surface images of living humans. This is followed by a review of opto-electronic techniques and a survey of some whole-body surface imaging systems. The near infrared scanning system used in this study is then described in more detail. Experiments are carried out to assess appropriate scanning postures, and to analyse the data captured.

2.1 Issues in Human Surface Digitisation

To capture and process the human body range data effectively, it is necessary first to explore some of the specific issues of human surface digitisation. This will provide a good basis from which to:

- select the appropriate image capture system;
- identify and avoid potential problems in image capture (subject's posture, etc.);
- enable a more informed analysis of the most appropriate techniques for processing the data, particularly for feature detection;
- design efficient operators to process the data.

An opto-electronic range imaging device—whether for teeth, the human body or a fold mountain—requires some source of electromagnetic radiation (sometimes simply natural lighting) which strikes the surface of the subject. Depending on the properties of the light source and the surface (including the angle of incidence, the surface's chemical make-up, its texture, internal structure and colour) the radiation is transmitted, absorbed, scattered and reflected to varying degrees. The reflected component is detected by one or more sensors, which convert the radiation to an electrical signal. The timing, coordinates or intensity of the signal are then used

to determine depth information. As well as the basic requirements of obtaining data in a clean, accurate, reliable and repeatable way, there are many specific problems when capturing 3D images of the living human body. These are recurrent themes that drive much of current research in range imaging [Davd98]. Before describing some techniques in more detail, this chapter first discusses some of these concerns, and how they affect the choice of technique and overall system design. Many of these issues are revisited toward the end of the chapter in designing low-level processing operations for the data captured.

2.1.1 The Dynamic Living Body

The human body is in constant change, even when "standing still". It is built from soft deformable tissue over a rigid, articulated skeleton, with motor reflexes continually at work to keep the body upright in dynamic equilibrium [JoRi97]. This introduces movement from sway, which, especially at the top of the head, can be considerable (~2 cm) [DaBR97] [CoHu98] in the ten or so seconds that many imaging systems take to capture a whole body image. Although methods, such as a contact pointer or grab rails have been suggested to reduce movement artefacts [DaBR97] [Hori98], these may be unacceptable to the subject, and can compromise the posture in other ways. Changes in respiration states typically change the girth of the torso by several centimetres, and so the breathing state during image capture ideally should be standardised. These changes take place over a second or so; however, it may also be relevant to take into account changes over the day through loss and gain of water content, intake of food, and so on, if the data captured are to be compared over time.

Some of these issues of movement are discussed again in section 2.2.6, where they are of importance in the high-level design of a whole-body image capture system.

2.1.2 Optical Properties of Skin, Hair and Fabric

Human skin and hair have particular properties that are relevant to its behaviour—reflectance, absorption, scattering and transmission—in the visible and near infrared regions [JoRi97]. The skin typically has a high diffuse reflectivity and low specular reflectance. It is particularly transparent at longer wavelengths, but it is desirable to use this range of the spectrum, rather than the blue region, for reasons of safety especially for the eyes [Besl89] [Hori98]. Other characteristics that may need to be considered are variations in subcutaneous fat, skin pigmentation and texture, which affect its albedo (diffuse surface reflection coefficient). Similarly, head and body hair, with its range of pigmentation and textures, will potentially cause problems, because of absorption and scattering [JoRi97].

These problems are further complicated when attempting to capture images of clothed or partially clothed subjects, where it is impossible to predict the range of fabric textures, chemical make-up, and colours that might be encountered. In most surface imaging systems, absorption

effects are most pronounced with certain black materials (e.g., leather and rubber). Surfaces with very high specular reflection, such as glasses and jewellery can also lead to undesirable artefacts with most optical imaging techniques [StMi98]. The sensors consequently need a good dynamic range [Riou98] to extract accurate depth readings over the range of surface properties expected to be encountered.

2.1.3 Body Geometry and Surface Sampling

The orientation of the human body's surface can cause problems of occlusion, where, as a result of the light source and sensor geometry, either component cannot "see" the relevant part of the surface. The most obvious of these problems can potentially occur at the inner surfaces of the arms and legs, especially given the many degrees of freedom of movement. Posture during scanning is clearly of significance, since this determines the orientation of body surfaces and potential occlusion effects.

Where the angle of incidence or that of the reflected ray is glancing with respect to the surface, the data may be rather sparse, as shown in figure 2.1. This occurs, for example, at the tops of the shoulders if the light source or the sensor's line of sight is parallel with the plane of the floor. More problematic still are undercuts from the bust and chin [BrDR97]. Problems of this type can be reduced by using paired detectors oriented slightly off-axis from the light source [Hori98] [StMi98] to give some redundancy. In addition to the sparseness of the data in such areas, the data are typically of lower accuracy.

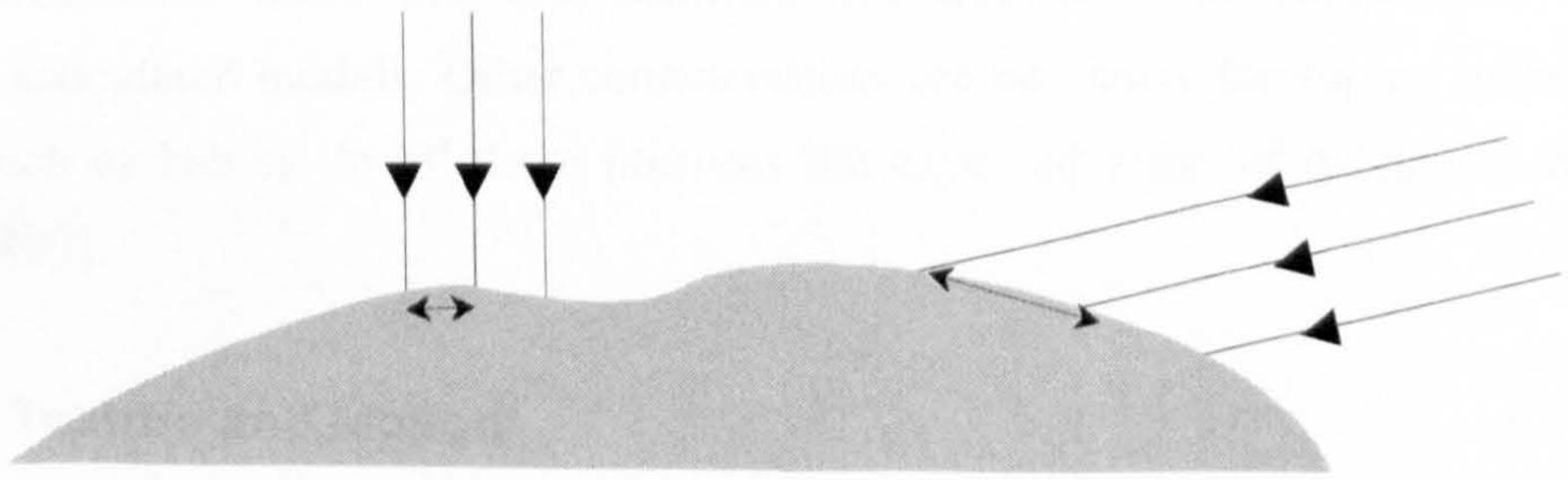


Figure 2.1: Angle of incidence/view and sampling density.

A known problem with any optical technique arises from mutual illumination at dihedrals, such as under the chin, at the armpit, or the buttocks. In such situations the perceived position of sample points may be several millimetres from the actual surface [Hori97]. The effect is dependent on the albedo of the surfaces, inversely proportional to the angle between the planes, but independent of the orientation of the angle. Since this occurs on *inside* surfaces, convex hulls will often not be affected. Nevertheless, it is expected that some problems might arise when attempting to reconstruct and measure the area around the armpits, inner legs, chin and neck, where the effect would occur on data points that lie on a convex hull.

Although it is generally desirable to obtain as high a resolution as possible over the whole surface of the human body, *appropriate* resolution is more important. For example, in

underwear design, high-resolution data on the head and hands is not needed. In addition, for scanning systems in which some or all of the data are gathered sequentially, problems of body movement mentioned previously may necessitate a compromise between the resolution of surface sampling and the length of the image capture time. In some situations it may be preferable to accept a lower resolution if this means that the data set *as a whole* has greater integrity. Sampling resolution should be sufficient to reconstruct curves and surfaces for visualisation and anthropometry with good fidelity to the accuracy required. Complex parts of the body and areas with a high rate of curvature, such as the face or hands, require higher sampling resolutions [StIW95] [JoRi97] [SiPa98]. Variations in resolution and noise ratio are of importance when designing processing operations such as data cleaning, smoothing, and curve and surface reconstruction. These are investigated in more detail in Chapters 4 and 5.

2.1.4 Working Range and Volume

For whole body imaging the active scan volume must be sufficient to cater for the target body sizes. Every imaging device has a given range within which it can be considered to deliver accurate and stable data [Riou98]. The limitations of a given system set-up may conflict with the desire to accommodate a wide range of body sizes. An ideal system would be capable of capturing whole images of the body in many different postures, not just standing, but for example, sitting, reaching, bending, and so on [TSHH98] [Seit98]. This requires the working volume of the sensor system to be quite broad, with a depth and width of 1.5 metres or so. This is essential for wheelchair users, and also desirable for ergonomic anthropometrics and in building realistic articulated models. Other considerations are necessary for supine scanning of some subjects, such as babies. In all these postures the expected areas of occlusion must be considered [BrDR97].

2.1.5 Colour, Texture and Motion

In addition to the requirements relating to the accuracy and integrity of the whole body data, the ability to map colour and texture onto the 3D form is useful for many applications. Examples include the assessment of burns or wound healing [WhGM97], and in enhancing the sense of presence in virtual environments [SiPa98] [InSp99]. Although this is not crucial for anthropometric applications, some form of intensity data is useful in the detection of manually placed markers. Geisen *et al* [GMHW95] and Lewark and Nurre [LeNu98] use flat matt white discs to mark key anatomical features, such as the trochanter and iliac crest, for subsequent detection by image processing.

Although none of the techniques described in the next section operates at the speeds required, one of the ultimate aims is to capture full 3D images of moving humans at high levels of

accuracy [SiPa98]. Another potential problem is in the amount of data acquired, especially if colour is involved.

2.1.6 System Design and Operation

Since a whole body imaging system potentially has such wide ranging applications outside of a controlled medical or industrial environment, it must be easy to operate and robust, capable of operating under a range of conditions. It may be subject to changes in ambient temperature, occasional vibration, and various sources of ambient light, from natural daylight and artificial lighting. The system should operate with minimal recalibration or other maintenance. These requirements are most critical for large-scale surveys, where the system may be required to deal with a throughput of a hundred or so subjects a day. These issues are also significant in building mobile systems, which have advantages for large-scale surveying [Tait98] [HQL95] [SiPa98] [InSp99].

Outer dimensions and weight of the whole system are also important, especially if it is to be used in a retail environment where space is usually expensive [Hori98]. Hence there is a requirement for lighter, more compact imaging components, and reduction of the weight contribution from any movement mechanism. Ideally the subject should be protected from any moving components. Unfortunately this can be difficult to engineer because the optical and geometrical properties of the system may then give rise to reflection artefacts.

The psychological factors of the imaging system are also crucial to its practical use, and relate to the criterion of minimum intrusion, stated in Chapter 1. Accurate technology is effectively useless if its presentation to the subject makes it unacceptable [Hori98] [Djab97]. This is especially true if it is to be used voluntarily. Since, for anthropometry, subjects will usually be processed in underwear, they must feel comfortable and private. Preparation should require minimal interference with the subject, if possible avoiding the need for wearing any special garment or the placement of body markers. This is not just important for long-term use of such systems, but also within the time span of this study.

2.2 Surface Imaging Techniques for the Human Body

Before describing the higher-level structure of specific systems, it is useful to describe the measurement principles employed. Many broadly-scoped surveys of these techniques exist [Besl89] [Isda98], but rather than reiterate these, this section focuses on techniques used to capture 3D data of the whole human body. Daanen and van de Water's survey [Davd98] describes many commercially available whole-body systems. The survey in this chapter is organised by imaging technique, with a short summary of systems, providing an update on the Daanen and van de Water paper. The techniques described here all use a structured lighting

approach, and can be broadly classified into active point and light-stripe sensors, moiré and stereophotogrammetric methods.

2.2.1 Single-Point Position-Sensitive Detectors

Conceptually the simplest of the active structured light techniques is the position-sensitive detector (PSD). The whole-body scanner developed by Hamamatsu Photonics [Hori98] uses this method. Each imaging head has an array of near infrared light emitting diodes (LEDs), combined with paired segment PSDs. The light emitted by each LED is focused onto the subject and the reflected light is focused onto the PSD, as illustrated in figure 2.2. The magnitudes of the currents drawn at each end of the PSD are determined by the location at which the light is incident on the PSD. By triangulation this can be used to determine the distance of the surface sample point from the sensor.

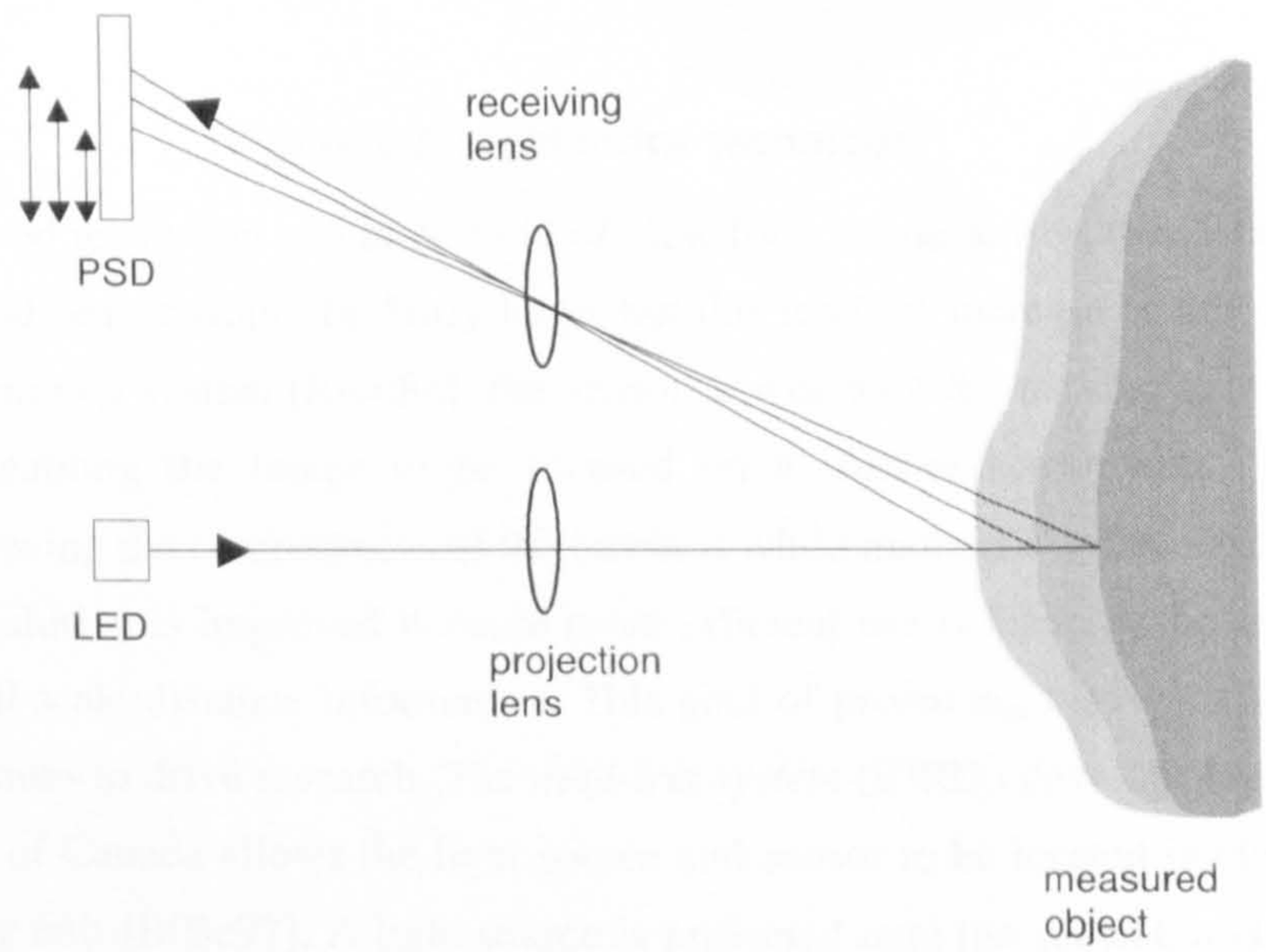


Figure 2.2: PSD triangulation.

Each sensor has two PSDs, oriented off-axis with respect to the emitting lens, and asymmetrically to each other. This set-up is designed to reduce occlusion effects. Although this method samples points one at a time, they are pulsed in rapid sequence at an interval of the order of 30 μsec (30×10^{-6} sec), and response times are such that speeds are obtained that are comparable to light stripe and charge-coupled device (CCD) methods.

2.2.2 Light Stripe Methods

Light stripe methods, employing laser and other light sources, are used in several commercial and research systems. A plane of light is projected onto the subject and where it hits the surface of the object, a planar contour is generated, as shown in figure 2.3. A 2D image of the scene is captured by an off-axis CCD camera. Since the equation for the light plane is known, and the

vector from each CCD pixel is known, it is possible to determine the 3D location of each sampled point at the unique intersection of the plane and vector. With this method the whole stripe image is captured simultaneously. In order to obtain the full surface at that field of view, the light source and sensor are swept across the surface of the subject. A limitation of this method is the image capture time of the cameras (typically about 50 Hz, although faster cameras do exist [StMi98]).

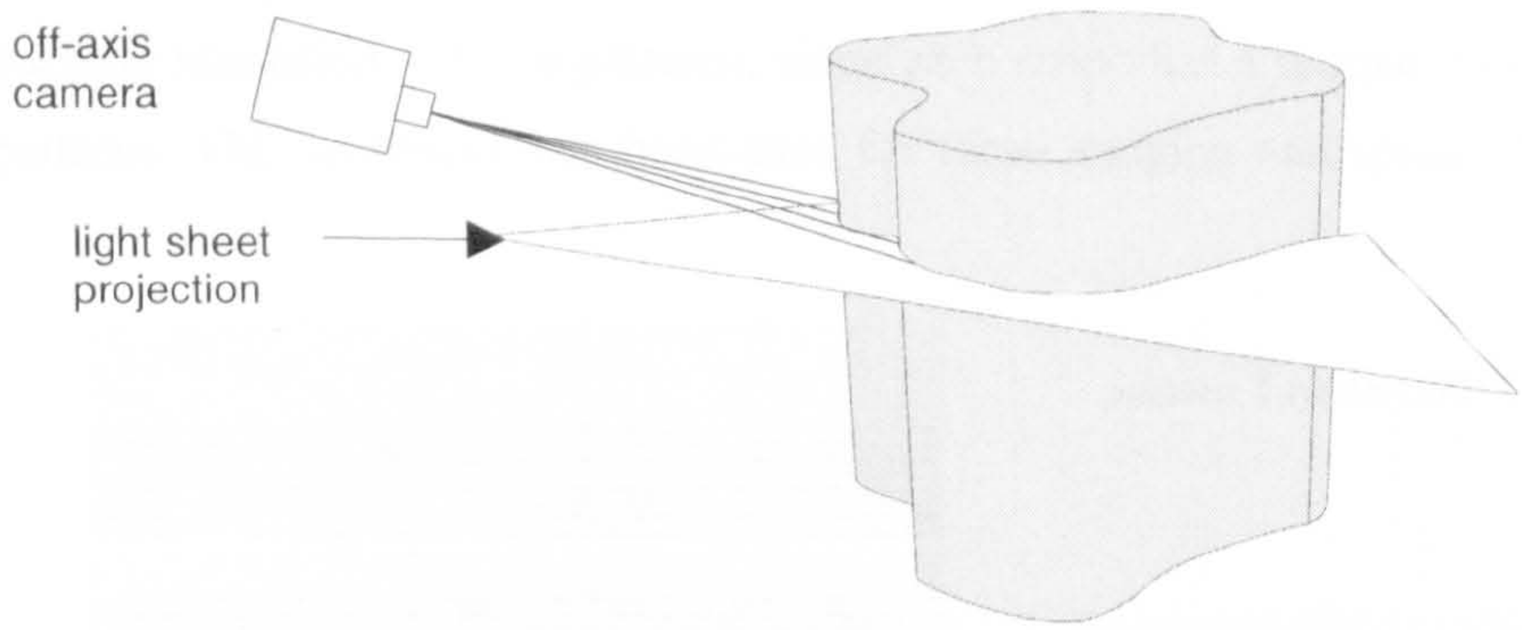


Figure 2.3: Light stripe technique.

In order to get a good resolution and large field of view for a single sensor, the distance between the light emitter and sensor would be fairly large, but this tends to increase occlusion effects. In a *synchronised scanning* system [Riou84], the sensor moves with the moving scan angle of the projected light, enabling the image to be focused on a smaller sensor area. This has the advantage of improving the compactness of the cameras while maintaining the depth of field and field of view. Resolution is improved because more efficient use is made of the sensor's width for resolving small-scale distance information. This goal of producing highly compact imaging components continues to drive research. The *dual-iris system* (BIRIS) developed at the National Research Council of Canada allows the light source and sensor to be located in close proximity within the imaging unit [BlBe97]. A light source is projected onto the subject, and the reflected light is split at a double-aperture mask in front of the lens. The image captured by CCD shows two intensity peaks for each sample surface point, one corresponding to each aperture. The separation of these peaks on the CCD array indicates the depth of the point in space.

Whole-body systems that employ variants of the light stripe method are the Loughborough Anthropometric Shadow Scanner (LASS) [JWHR89], Vitronics' VIRO-3D [StMi98], the TecMath 3D Body Scanner [Tecm97] and Cyberware's WB4 [Cybe93].

2.2.3 Multiple Light Stripe Projection

The previous methods sample point-by-point or line-by-line, and build a full surface by sequential scanning. Since speed is such an important issue, it is preferable to reduce the number of images required, and consequently speed up the process. By projecting multiple light stripes, images from each illuminated area can be captured simultaneously, for example by CCD

camera. Ambiguity between the lines must be then resolved so that each image pixel can be associated with a known light stripe source.

One approach to this has been to use a set of binary-coded light patterns, each pattern composed of n-1 light stripes. Within a given pattern each stripe from 1 to n-1 is assigned a binary value. If the bit value is 1 the light stripe is set to *on*, and if the value is 0 the stripe is set to *off*, as illustrated in figure 2.4. For example, for $n = 16$, the first pattern will have stripes 1 to 7 off and stripes 8 to 15 on. The patterns are projected one after the other, which allows any light stripe to be unambiguously identified in $\log_2 n$ patterns, since each stripe has a unique on/off sequence in the set of patterns. This technique has been used for torso imaging and spinal shape diagnosis [SDNV90].

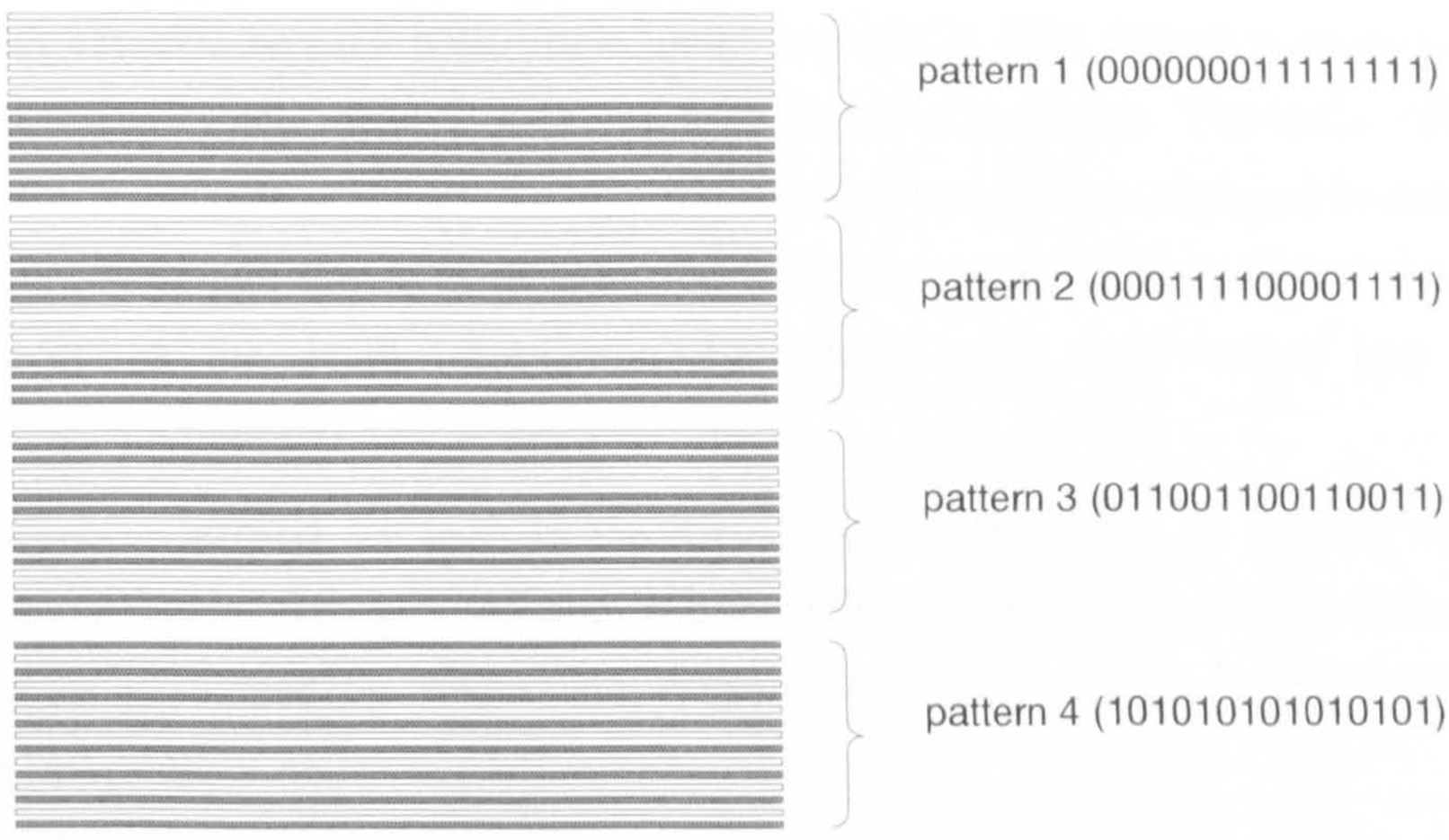


Figure 2.4: Set of binary-coded light patterns, each with n-1 stripes.

2.2.4 Moiré Fringe Techniques

Many methods exist to derive depth information from the intensity data in 2D images, captured, for example, by video. Shape from shading [Horn75] applies certain assumptions of surface smoothness, uniform albedo and Lambertian reflectance, and builds depth information based on changes in intensity of the radiated light. Photometric stereo [Wood78] allows some of these assumptions to be dropped, by using multiple images under different lighting conditions. Phase-shifted moiré [HaHs89] [DHWG97] can be seen as an extension to these methods, taking several images captured with known shifted intensity patterns.

A moiré method used by Demers *et al* [DHWG97] projects a sinusoidal stripe pattern of light intensity onto the subject, with a CCD camera capturing the 2D image. The intensity at each pixel is a function of the reflectivity of the surface and the phase of incident light intensity. Since the reflectivity of each area of the subject's surface cannot be controlled or known in advance, and the phase at any given point is dependent on the surface depth, it is necessary to

remove these from the equation. By capturing four images at $\pi/2$ intervals of the projected sinusoidal pattern, the function can be simplified, giving phase as a function of the four intensity levels at each pixel.

At the pattern periodicity required for accurate point location there will be several possible planes to consider, each with the same phase value. This ambiguity can be removed by also taking images using a lower periodicity pattern, to identify the particular area in which the first plane is located. The 3D coordinates are then determined by the intersection of each identified plane and pixel vector. It is also possible to resolve the ambiguities via prior knowledge of the surface, but for the variations expected in shape and posture of the human body, this is not considered practical.

Moiré techniques can also be used with single image capture. A light pattern is projected onto the subject, and the deformed image is captured with an off-axis camera. The initial reference pattern is overlaid (usually in software) and the interference patterns are analysed to determine depth information. Although this is faster than the four-image method described above, many ambiguities remain in the depth information, for which more sophisticated and slower processing methods are required.

Variants of the moiré method are used in the torso scanner developed by the UK National Engineering Laboratory, St. George's Hospital and the Ministry of Defence DCTA [Alle98], by the TriForm series developed by Wicks and Wilson [WiWi98], and the InSpeck 3D Body Builder system [InSp99].

2.2.5 Stereophotogrammetry

Stereophotogrammetric methods use multiple cameras at differing viewpoints to capture images of a subject, from which some attempt can be made to determine the depths of surface points [Long81]. From a single image it is possible to determine the x and y coordinates by tracing each pixel vector. There are an infinite number of depth (z) values to which each vector could correspond, but by finding the disparity between the pixel locations in two images corresponding to a given surface point, the z value can be determined by triangulation, as illustrated in figure 2.5. This potentially allows depth information to be extracted simply with the use of low-cost camera units and natural lighting.

However, the problem of locating corresponding points in multiple images is not a trivial one. Typically this method works by matching points on distinct features such as edges and line segments, which are notably scarce on the human body. One means of overcoming this problem is to project an "interesting" pattern, such as a grid, onto the surface of the object. One example is a random texture pattern projection [ScFo86]. In the C3D system developed by the Turing Institute [SiPa98] a flash texture unit illuminates the surface with a speckle pattern that enables the system to match corresponding pairs of points. To avoid loss of data from occlusion, the

camera pairs should be positioned close together, which then requires several imaging components to cover the whole body area, and a fairly large overall system design. Image capture time can be very fast (of the order of a few milliseconds), but the image processing time is typically quite long (approximately 45 minutes for C3D [SiPa98]).

A model-based stereophotogrammetry technique has been developed by Hilton *et al* [HBGS99], where orthogonal video images are mapped onto a simple parameterised 3D model. Similar techniques have been used to translate two-dimensional images [Seit98], such as those captured by Telmat's 2D system [Telm96] into a three-dimensional form for ergonomic modelling [TSHH98]. The major advantages of these approaches are a fast image capture time and use of low-cost components.

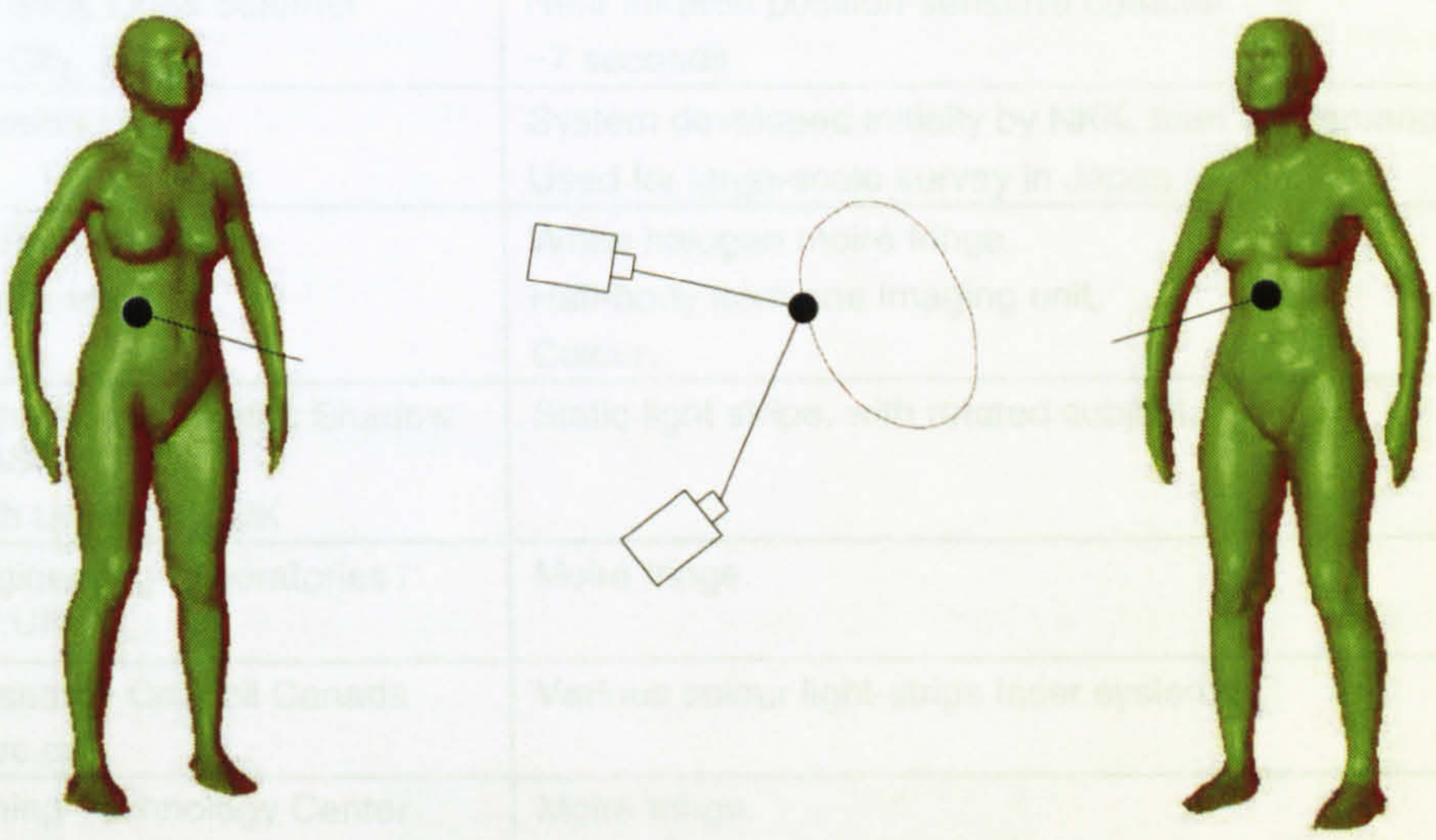


Figure 2.5: Depth from stereo.

2.2.6 High-Level System Design

In all the techniques described above, each imaging operation captures only part of the surface at one time—single points, lines or surface patches. To capture the whole body surface, several approaches can be taken. Table 2.1 summarises commercial and research systems for whole-body surface imaging. In the simplest systems the subject presents two or more (usually orthogonal) views to the system [Telm96] [HBGS99], and the two-dimensional images are fused to build a single 3D one. Because the subject moves, it is not possible to guarantee a close correspondence, and therefore image registration can be problematic. In early work at Loughborough the subject was rotated to capture the full 3D image [JWHR89]. More commonly, the imaging components scan across a given area to generate a partial surface image [Hori98] [Cybe93] [StMi98]. Several imaging components are used, each of which captures the surface visible from a given angle of view. Introducing more imaging components can reduce occlusion effects [SiPa98], provided that they are located and oriented appropriately with regard

to the expected postures of subjects [BrDR97]. Hand-held scanners and systems using terrain following have many interesting implications for capturing images of the human body in arbitrary postures, although current speeds are insufficient for living humans.

Table 2.1: Systems for Whole-Body Surface Imaging.

System and supplier	Key attributes
AvatarMe, UK www.avatarme.com	2D colour image capture. Model-based processing to generate polygon mesh avatar.
C3D Turing Institute www.turing.gla.ac.uk	Stereophotogrammetry with structured speckle pattern. Various sized systems, from face, up to (currently) half body. Colour.
Cyberware WB-4 Monterey, USA www.cyberware.com	Light stripe laser, 4 imaging heads. Colour.
Hamamatsu Body Lines Scanner Hamamatsu City, Japan	Near infrared position-sensitive detector. ~7 seconds.
Hamano Voxelan / NKK	System developed initially by NKK, then by Hamano. Used for large-scale survey in Japan.
InSpeck 3D Body Builder Montreal, Cananda	White halogen moiré fringe. Half-body from one imaging unit. Colour.
Lougborough Anthropometric Shadow Scanner (LASS) Lougborough University, UK	Static light stripe, with rotated subject.
National Engineering Laboratories / MoD DCTA, UK	Moiré fringe.
National Research Council Canada www.vit.iit.nrc.ca	Various colour light-stripe laser systems.
Textile/Clothing Technology Center www.tc2.com/RD/RDBody.htm	Moiré fringe.
TecMath 3D Body Scanner www.tecmath.de	Light stripe laser.
Telmat 3D SYMCAD	Structured light. Model-based processing to generate polygon mesh.
Tricorder www.tricorder3d.com	Stereophotogrammetry.
Vitronic VIRO-3D, Vitus www.vitronic.de	Light-stripe laser. Variable resolution, colour.
Wicks and Wilson TriForm www.wwl.co.uk	Moiré. Colour.

In *scanning* systems the direction of movement influences the effects of body movement on data quality. In most systems the scanning components move from head to foot, so reducing the effects of sway [DaBR97]. Some attempts have been made to model the expected motion and compensate for this in software [StMi98] [JWHR89]. It is preferable, however, to reduce these effects by reducing image capture times. In order to achieve this, much research has been devoted to improving the response time of the imaging components [StMi98] [Hori98]. In

systems with moving components, the drive mechanism must be correspondingly speeded up, and this poses additional problems of dealing with inertia, vibration and so on [Hori98]. In addition, the movement mechanisms are often bulky and heavy, there is a potential safety problem because of the moving parts, and correspondingly more parts to go wrong. Although systems composed of static imaging components [DHWG97] [SiPa98] avoid the above problems, a major disadvantage is the overall size of such systems, since the imaging components must be located at a distance from the subject in order to obtain the necessary field of view.

2.3 Capture, Conversion and Storage of Range Data

This section describes the whole-body imaging system to be used in this study. Since appropriate software is not available, it is also a necessary task of this thesis to develop the basic "operating system", to interface with the scanner and view the 3D images. This forms the basis of the automated image interpretation system, whose development and testing is described in the subsequent chapters of this thesis.

2.3.1 The Whole-Body Imaging System

Data for this study is captured using the 8-head Body Lines Scanner developed by Hamamatsu Photonics [Hori98], shown in figure 2.6. A system has been installed at University College London, enabling easy access to capture images of subjects.

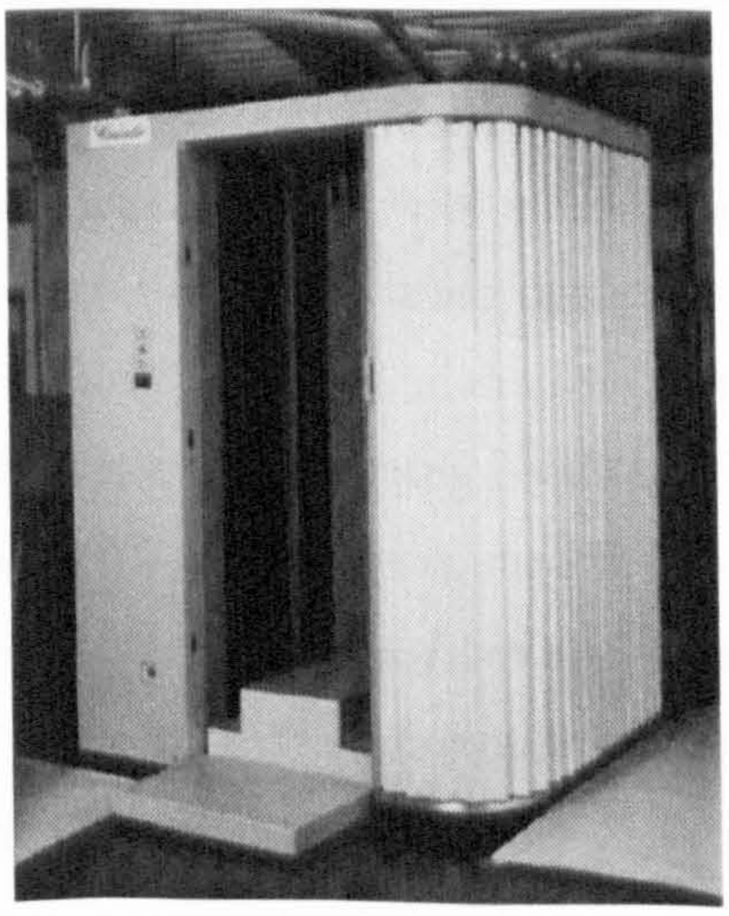


Figure 2.6: The Hamamatsu 8-Head Body Lines Scanner.

The scanner uses near infrared position-sensitive detectors, arranged in eight sensor heads, each with a 32-LED array light source. In order to obtain full 360° coverage on variable sized objects, there is an overlap of the fields of view of each head, giving some data redundancy.

The reflected component at the receiving lens is focused onto the position-sensitive detector. Two receiving lens/PSD pairs, angled symmetrically off-axis to the projected light, are used to

minimise occlusion effects. The pulse time is 30 μs and the timing between pulses is 30 μs, giving a total scan cycle for a single head of 1.92 ms (32×(30+30) μs). The scanning mechanism moves 5mm vertically during the cycle time to pulse the 8×32 LEDs. The resulting scan is therefore a staggered spiral. Over the ≈2000 mm working height of the system 400 "slices" (turns of the spiral) are obtained, with 102,400 sampled points in total. The full scan time is approximately 10 seconds, determined by the response time of the detectors and the acceleration and retardation period of the movement mechanism. The active measurement range is 350 to 650 mm, and the working scanning volume of the system used for this study is 460×2000×690 mm, making it suitable for scanning human subjects in a standing posture.

The eight sensor heads and the movement mechanism are contained between a light outer shell and an internal acrylic resin curtain wall, which protects the subject from moving parts. The system is certified with a CE mark as eye-safe, and safe for unborn children. The subject stands in the centre of the scanner, facing in the $-x$ direction, as shown in figure 2.7. The scanning heads are oriented to maximise coverage around the whole body and on the insides of the arms and legs.

2.3.2 Download and Storage

The data captured by the system is downloaded via a general-purpose interface board (GPIB), and consists of a block of 102,400 distance values and intensity values. Distances are measured in units of 1×10^{-1} mm from the sensor to the sampled surface point; intensity values range from 0 to 255 and correspond to the proportion of the emitted light reflected back to the sensor.

One of the first tasks is to devise a compact and lossless file format, so that scans can be stored and retrieved. In this format, each scan, together with information on the scanner configuration, totals 300 kbytes. The scanner configuration data are used in low-level processing operations. This allows the software to be used, without alteration, with several different types and configuration of scanner. The data from all scanning heads is downloaded together. Because of the common coordinate system of the imaging heads, no specific registration step is needed—the data set is simply the union of data from all the heads. There are some issues of redundancy where the fields of view of two heads overlap. This is discussed further in Chapter 4, when the low-level processing operators are designed.

2.3.3 Conversion to Common Coordinate System

Distance values provided by the hardware are first converted to a common Cartesian coordinate frame, using the camera geometry and point sequence information. The origin is a point at floor level in the centre of the scan volume (usually between the subject's feet), the primary axis is the body height (y), the secondary axis corresponds to body width left to right (z) and the tertiary axis corresponds to depth (x) (see figures 2.7a and 2.7b).

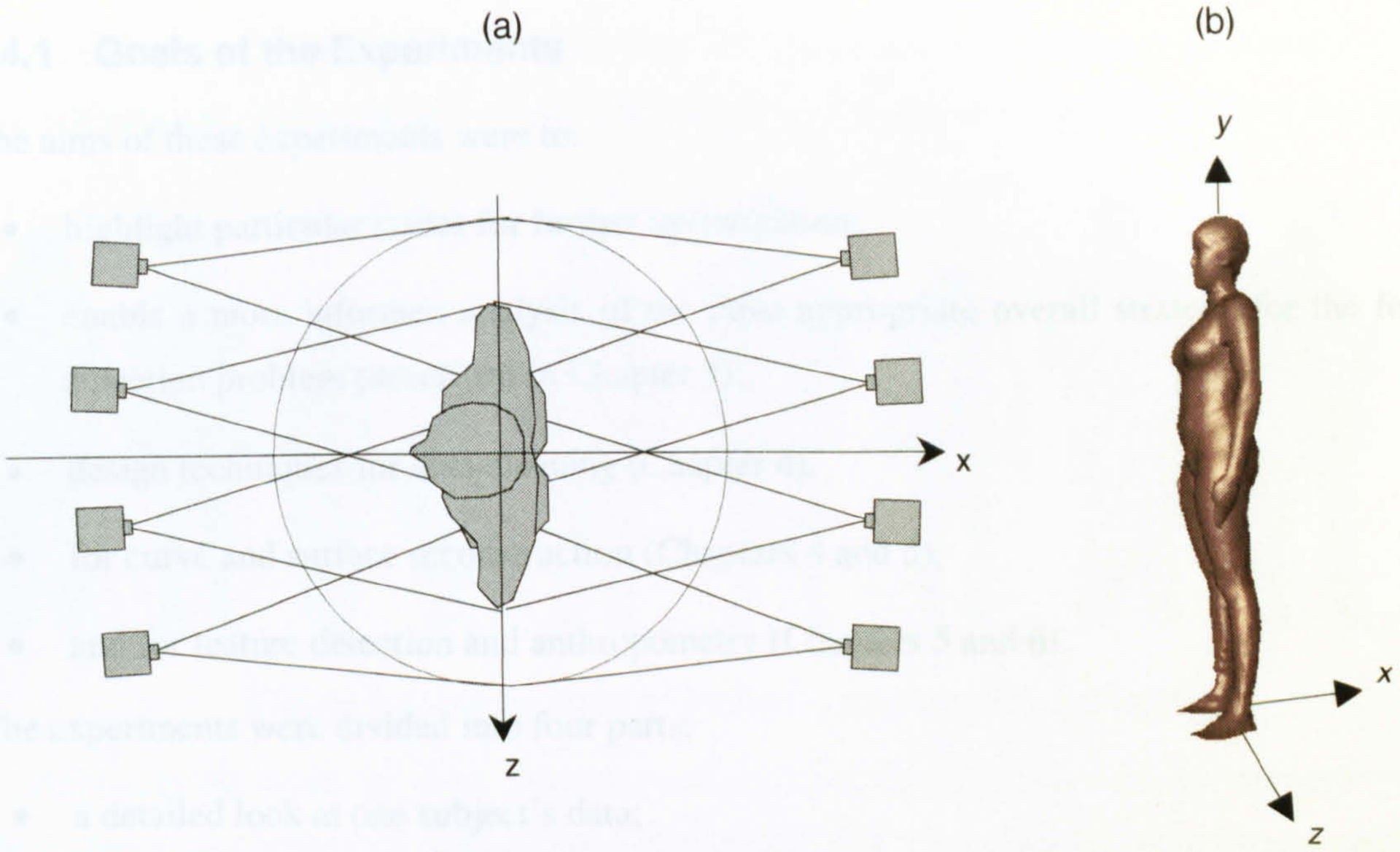


Figure 2.7: Hamamatsu scanner (a) cross section, showing field of view for each imaging head; (b) coordinate system

2.3.4 Initial Processing

After the data has been converted to Cartesian coordinates, two cleaning processes are carried out, each with $O(n)$ time complexity:

- threshold on the intensity values, as low intensity data is generally considered less reliable than higher intensity values (a threshold of 10, recommended by the scanner manufacturers, was used);
- removal of data points located outside the working scanning volume.

These steps provide a point set that is clean enough to analyse its characteristics, and to carry out experiments in scanning a range of subjects in various postures. Extensions were made to the simple scanner interface software to provide the following functions to visualise the data:

- "slices" of points through any of the three sets of planes perpendicular to the x, y and z axes;
- 2D projections of the data as greyscale images displaying intensity values.

2.4 Experiments on Data Capture and Analysis

This section describes some preliminary experimentation in whole-body image capture, to give a largely *qualitative* assessment of the behaviour of the data from a wide range of subjects and data capture conditions. These qualitative aspects are made quantitative later, in the relevant chapters as noted below.

2.4.1 Goals of the Experiments

The aims of these experiments were to:

- highlight particular issues for further investigation;
- enable a more informed analysis of the most appropriate overall strategy for the feature detection problem (described in Chapter 3);
- design techniques for data cleaning (Chapter 4),
- for curve and surface reconstruction (Chapters 4 and 5),
- and for feature detection and anthropometry (Chapters 5 and 6).

The experiments were divided into four parts:

- a detailed look at one subject's data;
- assessment of fabric properties, hair, skin, etc.;
- scan posture, and
- artificial landmarking.

The implications of the results of these experiments are discussed at the end of the section. Personal responses from subjects were also noted, in order to ensure comfort and privacy during the experiments, but these are not reported or analysed here. For an analysis of these issues, see Djabri [Djab97].

2.4.2 Subject Selection and Data Capture

For these experiments 41 subjects were scanned. Table 2.2 summarises the characteristics of the sample set. A number of postures were captured, and subjects wore varying amounts and types of clothing, from undressed, to wearing underwear, swimwear or outer garments, of varying looseness of fit, fabric texture, colour and material composition. Depending on the particular issues of interest, between one and ten postures were captured of each subject. All subjects were of normal health, with no limb loss. Some subjects had significant postural problems, such as scoliosis and lordosis, but no cases were severe.

Extreme cases in terms of size were one pregnant female subject at full term, three specialist outsize female models, and several heavily built men. At the other end of the scale, several young children were scanned, the youngest and smallest being a two-year-old. Some sample images are shown in figure 2.8 below. Each 3D image is shown in projection on the *xy* plane, with intensity values, stretched in the lower range with a sigmoidal function. The vertical bands of low intensity visible on most subjects result from the LEDs at the edges of the array, where the angle of incidence on the body surface tends to be oblique, and therefore gives lower

intensity readings. Some areas of missing data are apparent, most notably at the top of the head on some subjects, where the hair causes scattering and absorption in the near infrared region.

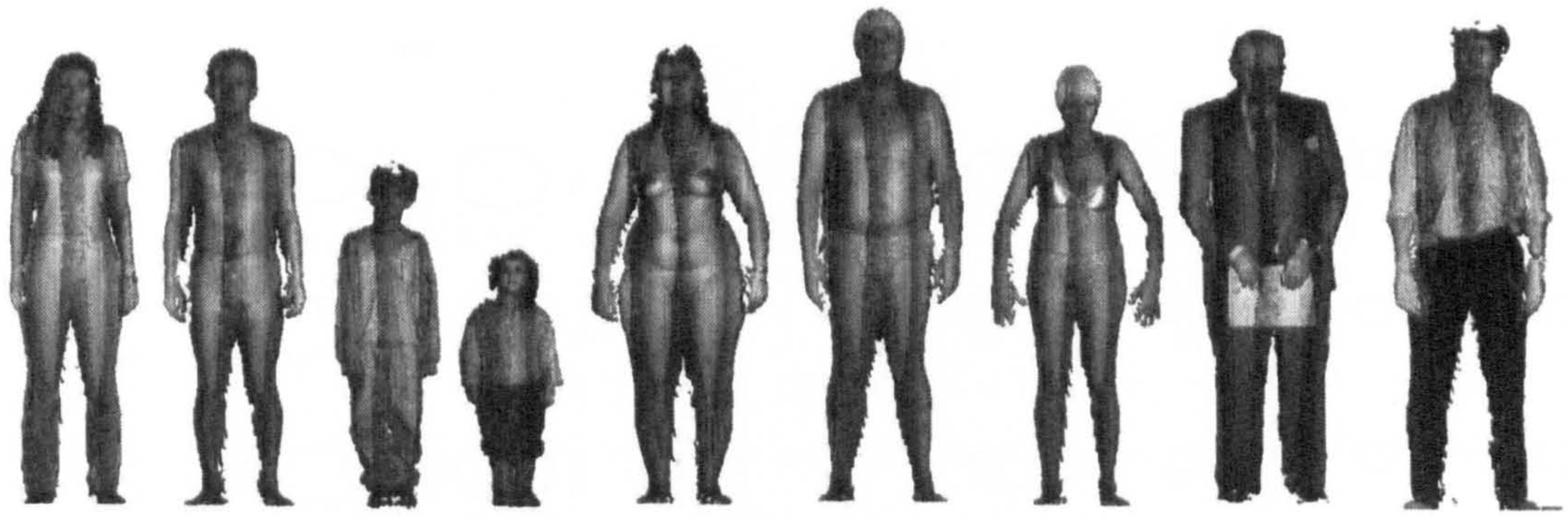


Figure 2.8: Projections of intensity maps on the *xy* plane.

In addition to the scans, subjects were also weighed and their heights were measured. Age, and approximate hair and skin colour were also noted. For five subjects who were scanned in underwear, the heights of the branch points for the arms and legs were recorded. Since it was expected that occlusion would be most significant and problematic in these areas, knowing the location of these features would enable closer analysis of the behaviour of the data.

Table 2.2: Subject characteristics for initial image analysis.

Gender	**female**				**male**		
Frequency	21				20		
Age (years)	**< 8**	**8-12**	**13-17**	**18-30**	**31-45**	**46-60**	**61+**
Frequency	3	1	0	28	5	3	1
Weight (kg)	**< 25**	**26-45**	**46-70**	**71-85**	**86-100**	**100+**	
Frequency	3	4	16	15	2	1	
Height (m)	**< 1**	**1-1.5**	**1.51-1.65**	**1.66-1.8**	**1.81-2.0**	**2.1+**	
Frequency	1	3	10	17	9	1	
Skin colour	**light**		**medium/olive**		**dark**		
Frequency	34		4		3		
Hair colour	**light**		**medium**		**dark/black**		
Frequency	9		17		15		
Clothing	**none**		**form fit / underwear**		**full clothing**		
Frequency	4		21		16		

2.4.2 Analysis of a Sample Scan

The first set of experiments was designed to take a detailed look at one subject's scan. The issues raised at the start of this chapter provide pointers for analysing the images, primarily how the different surface types on the human body affect the quality of the data sampled. A female subject of average height and weight was scanned in underwear, in several postures. The first of these is discussed here, where she adopted a relaxed stance, with the arms held slightly away from the body, and the feet approximately 20 cm apart. Figure 2.9 shows cross sections of data in the *xz* plane, partially processed as outlined in section 2.3.4. In figure 2.9a sections are shown

at 50 mm intervals across the whole body height on a female subject. Figure 2.9b shows sections between the lower head and upper torso, sampled at 5 mm intervals on a male subject.

Figure 2.9: Cross sections through raw data, (a) 50 mm intervals over whole body height; (b) 5 mm intervals between head and upper torso.

Features such as the nose and navel are visible, but small body segments, such as individual fingers, are not visible, given this resolution of data. In order to quantify the resolution over the body, data density at each cross section is computed for each body segment, as:

$$density = (surface\ length) / (number\ of\ points\ in\ sample).$$

The method required for computing surface length is described later, in Chapter 5. The maximum "gap" between adjacent data points within each horizontal section are also computed. The results for each body segment are shown in table 2.3. Sampling density on the front and back of the body is slightly higher than on the sides. This is to be expected, given the arrangement of the scanning heads, described earlier. This is more pronounced on smaller diameter parts of the body, such as the head and neck, where the LED rays have diverged to a greater extent (see figure 2.7a).

Table 2.3: Horizontal data density by body part (mm)

	head	torso	arms	legs
Mean distance	8.0	6.5	13.1	7.6
Maximum gap	42.1	51.6	49.9	56.3

It is evident from close inspection of the data, and from figure 2.9 and table 2.3, that occlusion effects are present on the inner surfaces of the legs and arms, where there can be gaps of up to five centimetres or so. This can be contrasted with the behaviour of data on the head, where the average density is lower, but the maximum gap is smaller and less frequent. In both cases the ratio of sample points to surface length is similar, but the spread of data is more uniform in the latter case. This is of significance for cleaning and surface reconstruction, and will be discussed again at the end of this chapter. For the time being this gives a useful indication of the distribution of data in each sample. In Chapter 4 a thorough quantitative analysis of the data will be provided.

It was noted earlier that the fields of view of the scanning heads overlap. This is illustrated for a horizontal cross section through the legs in figure 2.10. Data is shown for three scanning heads, where the data overlaps. Pairs of data points at a similar location, from different sensor heads, are consistent in most cases, but very slight discrepancies (~1-2 mm) do occur. One possible explanation is that the accuracy of the sensors is expected to be higher where the angle of incidence is small, with a decrease in accuracy as the angle of incidence increases.

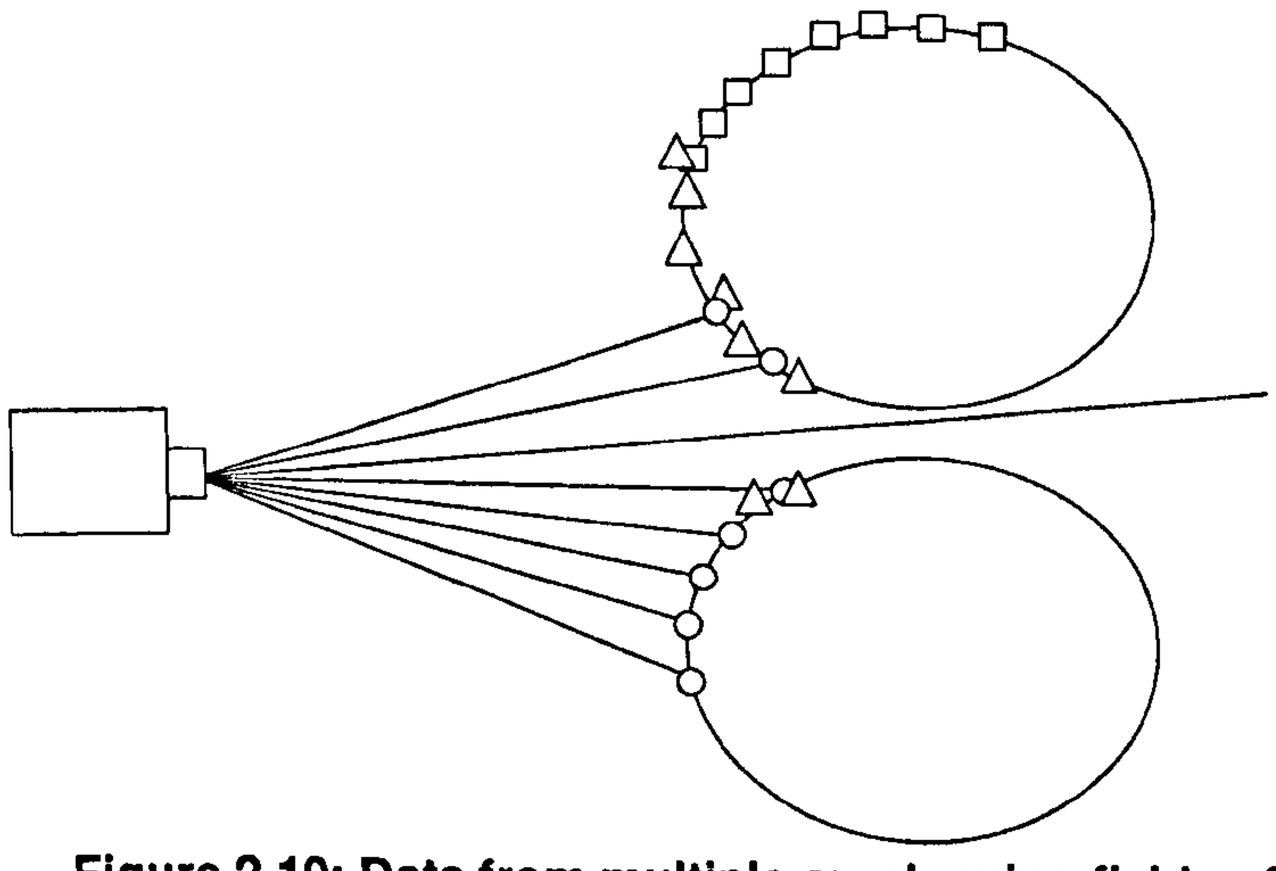


Figure 2.10: Data from multiple overlapping fields of view.

The potential importance of the dihedral problem (see section 2.1.3) is reduced in that the effect occurs in concavities, so that it does not affect closed curve convex hull measurements if the two surfaces are part of the same body segment (e.g., on the buttocks). Where the two surfaces correspond to *different* body parts (e.g., the inner surfaces of the left and right legs), any errors *could* be significant. Although no systematic tests were carried on for the dihedral problem, observations of cross sections through the sample subjects' data indicate no particular problems. This is therefore not considered further.

It was noted earlier that some noise is present in the data, which must be removed by filtering out the noisy data points or by applying a smoothing function to correct them. In some areas the data is sparse. In such areas this must be corrected by an appropriate technique for inferring

where the missing points should be. Correction for noise and sparseness are related, in that appropriate assumptions must be made about what the data *should* look like in order to be able first to *detect* the problem, and then to *correct* for it. This will be explored further in Chapter 4.

Other researchers have studied body movement in relation to data fidelity in whole-body surface imaging and noted the potential problems [DaBR97] [CoHu98]. One of the recommended solutions was the use of a pointer touching the crown of the head to reduce sway. For this work it is not considered appropriate, because of the minimum intervention criterion noted at the end of Chapter 1. Since the scanner used in this work is significantly faster than the one used in these previous experiments, it is assumed that movement will be significantly less (although not a *linear* reduction with respect to scan time), and so no specific tests were carried out to assess movement artefacts.

2.4.3 Scan Posture

Since posture has a strong effect on the shape of the body, and on occlusion effects, experiments were carried out to determine the most appropriate, reproducible posture to capture good image data, and to convey meaningful anthropometric information. For example, to locate the armpits, the arms should be held well away from the body to allow the scanner "access" to these areas. In women in particular, this can cause significant changes to the shape of the upper body.

Ideally, all the information of interest would be extracted from a single scan. However, a full anthropometric procedure, e.g., ISO 8559, carried out using traditional techniques, requires the subject to adopt various different postures to capture the full set of information. Consequently, as noted by several researchers [BrDR97] [BrGr99] [BDGW00], a single scan cannot reproduce this exactly. One way to tackle this problem is to take several scans and extract a subset of information from each, as is done by the CAESAR project [Robi98]. In contrast, for this work the aim is to simplify the procedures, and attempt to extract as much information as possible from a single scan. It is worth noting that overly prescriptive instructions about the required image capture posture appear to be counter-productive, since few subjects are able to absorb very detailed instructions, and translate them into the correct stance.

A compromise posture is therefore necessary, whereby the subject holds the arms just far enough away from the body to provide a visible gap between the arms and the torso, to aid in segmentation of the data. Qualitative tests showed that if the arms are rotated back slightly, the underarm area is opened up, without changing the shoulder or upper torso shape significantly. This posture is illustrated in figure 2.11, with reconstructions of scans of an adult and child subject.

For larger subjects this arm position occasionally results in truncation of the arms in the image, because of the particular scan volume of the system used here (see the dotted circle in figure 2.7a). Possible methods for dealing with this would be to "zip" multiple scans, based on

common landmarks. However, as in general the scan volume appeared to be adequate for the work presented here, this is not considered further.

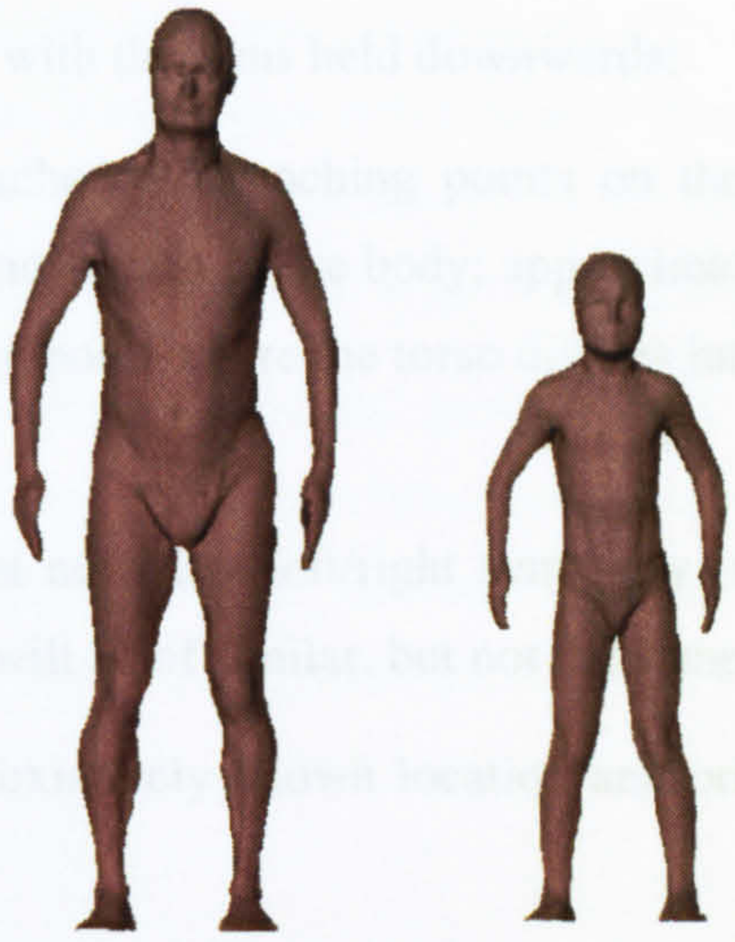


Figure 2.11: Adult and child subjects, showing reference scan posture

Similar problems are encountered when defining the appropriate position for the feet and legs. If the feet are held too closely together, segmentation of the data into the left and right legs might be problematic, and detection of the leg branch point would be extremely difficult. However, if the legs are held greater than hip width apart, the shape of the hips changes significantly. Again, a compromise position is needed, where the subject stands with the feet between 10 cm and hip width apart. It is not possible to eliminate the problem of the legs touching, when the legs are held hip width apart, even on some subjects of light build. Therefore an effective technique for branch-point detection and leg segmentation must be devised, which does not rely simply on gap detection [Nurr97].

The head is held with the chin approximately horizontal, as described in the ISO 8559 specification. This has the advantage of minimising loss of data in the concave chin area, although natural variation amongst individuals means that this effect cannot be avoided entirely.

Since key dimensions can change considerably during breathing, it is important to capture the scans at a consistent respiration phase. This, of course, is true with any other anthropometric procedure, and it is appropriate to standardise on the traditional specifications [ISO89] [Kuni84], where the subjects are at a relaxed mid-tide lung volume.

The hand in its relaxed position usually creates a concavity, where the fingers are slightly flexed. If the hand is opened out flat, the concavity is removed, thereby simplifying the reconstruction of the inner surface of the hand, and reducing the tendency for the fingers to approach too closely to the torso. Since information on individual fingers is not required, the fingers and thumb can be closed laterally, which again aids in simplifying the requirements for surface reconstruction.

If subjects are captured in this approximate reference posture, it is valid to make the following sort of assumptions in the subsequent processing:

- the subject is standing upright, with the arms held downwards;
- the subject has two arms attached at branching points on the side of the body, a short distance down from the uppermost point of the body; approximately about halfway down the body, there is another branching point where the torso divides into two, below which the legs reach down to the floor;
- there will be approximate, but not exact left/right symmetry (shoulders will be at similar, but not identical levels, arms will be of similar, but not the same length, etc.);
- the subject stands in an approximately known location and orientation with respect to the imaging system;
- there will be some variation in the orientation of the principal axis of the body, through postural differences and conditions such as scoliosis and lordosis [SDNV90] [Bato92];
- the hip, waist, shoulder alignment may not be even: twisting, shifting, etc. may occur; girths may not be parallel to floor (e.g., iliac crests may be at different heights, etc.);
- the arms may be held at differing distances from body, and may or may not be fully in the image, etc.

Similarly, other variations from the mean size, shape and proportion must be considered, and either incorporated into the model, or explicitly defined as being out of its scope.

The three axes chosen for the coordinate system correspond closely with the principal axes of the body when the subject stands in the reference posture described above. However, in some (not uncommon) cases these assumptions do not hold fully. For example, subjects with scoliosis may have a lateral twist to the spine and/or a rotation about the main axis, therefore a slice taken horizontally at hip level may be oriented as expected, but slices taken above this may be progressively rotated. In such cases there is no simple linear transformation that can map the image back to the reference stance. Variation in posture also has implications for specific feature detection. For example, it is not valid to assume that the bust is always the most prominent area on the torso, because a slight change in posture may make the abdomen the most prominent area. For example, in cases of even mild lordosis, there is a curvature of the spine in the sagittal plane, which tends to make the abdomen shift forward with respect to the hip and upper torso.

One possible solution to such variations is to use principal components analysis (PCA) individually on each subjects' image data to extract three axes. However, this would not fully account for the curvature found in subjects with scoliosis and lordosis. Another approach would be to use a non-PCA-based method to determine the (non-linear) principal axis, and to

determine the secondary axes at closely spaced intervals orthogonal to this axis. This removes the constraint of a simple linear coordinate system, and would allow for lateral curvature or rotation. However, unless the arms are previously segmented from the data, their many degrees of freedom might distort the results, lending spurious rigour to such an approach. It is assumed, then, that a fixed linear coordinate system will be used for this work, to maintain simplicity, but in future work it might be interesting to investigate shape change arising from postural variation.

Anthropometry on obese subjects brings additional problems, where, typically, folds of flesh might need to be moved to gain access to the appropriate areas for measurement. This will cause obvious problems with a static surface image of the body. Therefore no attempt will be made to deal with this in this work, although it is certainly important to be able to cater for all body shapes in the future.

2.4.4 Skin, Hair and Fabric

No noticeable differences in the quality of the data were found with dark or light skin. It might be expected that darker skin would show less noise at dihedrals, because of the lower albedo [Hori97], however, this was not tested quantitatively. Dark hair absorbs infrared strongly, and certain textures scatter the data. There appears to be a threshold of hair density and coverage area at which this occurs. Body hair, for example, does not appear to be a problem in terms of capturing data points of sufficient intensity, but some of the measured points may correspond to hairs, rather than the underlying skin. A swimming hat can be worn to avoid the problem of absorption and scattering on the head. This also has the advantage of flattening the hair so that in most cases an accurate height can be measured, and it enables hair to be held away from the neck and shoulders where it would interfere with body data in those areas. Loose clothing is not appropriate for this work, since this simply produces images of the draped fabric, with little information on the body shape. The best type of clothing would seem to be underwear or swimwear, as long as it is close fitting, but not so tight as to distort the body shape.

Most black materials absorb strongly in the near infrared region, in particular, some woven fabrics, leather, rubber and PVC. Retroreflective material also gives rise to data points with very low intensity, presumably because most of the light is reflected back to the source and little is captured by the sensors. All underwear and swimwear fabrics tested here appeared to give adequate image quality.

Objects with very high specular reflection, such as jewellery and glasses, give rise to noisy points of high intensity offset slightly from the surface. Points of this kind that are isolated are not expected to be problematic for cleaning and smoothing operations, but larger aggregations of such points would present difficulties. Consequently, if possible, subjects should be scanned without these items. Reflections can also arise because of the structure and geometry of the

scanner itself. Most have low intensity and are filtered out by the intensity thresholding, but some artefacts have high intensity.

2.4.5 Reference Data for Feature Detection

Since the central theme of this work is to detect body features, reference data on these features is needed in order to develop the techniques. These experiments are designed to test the feasibility of detecting manually placed markers within the intensity data of the image, and to use this to develop the processing techniques to extract the features of interest. However, the final system at *run time* aims *not* to make use of such artificial landmarks. Pre-placed protruding markers were used by Brook-Wavell *et al*, who detected the markers in the scanned image by virtue of their shape [BrJW94]. Li mapped texture onto the surface to enable flat marks to be detected manually in the resulting image [Li97]. Geisen *et al* [GMHW95] and Lewark and Nurre [LeNu98] detect flat disks pre-affixed to the body, using standard image masking techniques to detect the markers automatically by their colour or intensity characteristics. Subjects wear close-fitting, grey lycra clothing, to provide a controlled background against which to detect the markers. Some locations, such as the tops of the shoulders, require protruding markers, because of the difficulty of detecting flat markers in these areas of lower sampling density.

In this work the scan data is of lower density, and therefore it would be expected that larger markers would be required, since sufficient data points must be reliably identified as belonging to the marker. If the size of markers used by Geisen *et al* [GMHW95] are scaled up to give a similar number of sample points in the scanner used here, the markers would have to be 2 cm or more in diameter. In addition, since the aim is to have minimal interference with the subject, it would not be acceptable to have them wear special clothing. This means that the marker must be detected within higher levels of variation in the background intensity values, potentially adding further difficulties.

Artificial markers of various materials, size, colour and shape were tested, using visibility by eye within the scan image as an initial indicator of the likely success of an automated image processing technique. Materials tested included paper, rubber, plastic tape, silvered paper and retroreflective material. Semi-matt black plastic tape was found to be the most visible material, against a range of patterned fabrics and skin colours, giving an area of intensity below the lower threshold. Crosses approximately 4 cm in diameter were found to be easily visible, allowing a landmark to be located manually on the scan to within an accuracy of 5 mm, even in areas of sparse data, such as the tops of shoulders. When many such landmarks were placed on the body, interference between the markers occurred, making it difficult to distinguish each one. In addition, the placement and removal of the markers was considered to be intrusive to the subjects. Although these tests were only very cursory, enough problems were encountered to

suggest that this would not be an acceptable approach to the collection of reference data on landmarks. An alternative approach might be simply to collect the anthropometric information, such as surface contour lengths, without directly collecting reference data on the landmarks themselves.

Another possible method would be to use 3D position sensor equipment [SaRS91] [Polh98] to determine the spatial locations of selected landmarks. This may be of limited use because of the slow speed of information capture, as movement from sway between collection of each landmark might render the overall integrity of the data invalid. An alternative would be simply to measure only the heights of landmarks with an anthropometer. This might yield more valid data, since the height values are more stable with respect to sway than are the width or depth readings. The third method would be to mark features manually on the reconstructed images of the training set. There are clear drawbacks to this, since there is no tactile feedback with the electronic image, and the surface visualisation would need to be good to enable such interactive placement of landmarks.

2.5 Summary

This chapter has discussed some of the problems and techniques for surface digitisation of the living human body, and a survey of whole-body systems was presented. A near infrared whole-body scanning system has been used to collect sample images of subjects under various conditions, to explore the data quality, how it changes under these conditions, and some of the implications for developing automatic processing techniques. Appropriate postures for scanning have been investigated, and a reference stance has been devised, which provides good data coverage, and allows the subject to assume a fairly relaxed stance. Although the scan data is of fairly high density, some noise is present, and areas of sparseness occur, which must be dealt with later. However, the images appear to be of sufficient quality for this work. With this understanding of the 3D image data, it is now possible to explore techniques that might be used in transforming it to detect surface features.

Chapter 3

Techniques for Whole-Body Image Interpretation

This chapter discusses methods for representing and interpreting the image data described in the previous chapter. It starts by critically reviewing methods for representing the body surface data as a whole, and then discusses techniques for describing and matching specific features on those surfaces. A progressive pattern recognition approach is introduced to address the tractability of the problem and to exploit knowledge acquired during image interpretation. The chapter concludes by outlining the issues for further investigation.

3.1 Whole-Body Representation

At this stage each whole-body image is represented as a set of points, which are noisy, with some gaps and redundancy, but little explicit structure. This data set is limited because the close proximity of some of the body segments (e.g., the upper arms and the torso) makes it difficult to extract anthropometric information without interference between the data from these segments. Before attempting to extract more information, it is necessary to:

- segment the data in an anatomically meaningful way;
- restore data in sparse areas;
- impose some structure so that the body is represented as a whole object, with appropriate connectivity information.

This section discusses representations for the whole human body, specifically in the context of feature detection and the extraction of body metrics.

3.1.1 Body Representation Criteria

The primary function of the whole-body representation is to provide a starting point for feature detection, and therefore it should represent the whole body, including *all* the surfaces of interest. It was noted in the previous chapter that some surfaces are usually occluded, and therefore gaps in the data occur. These areas must be restored with as high a level of accuracy as possible. Ideally the representation should be a fully closed surface.

The representation should be expressive enough so that it does not prejudice what techniques can be used later for feature detection. In other words, it should provide as much generic information as possible, without loss of information and without being over-specified. For example, it should incorporate neighbourhood relationships of surface points, to facilitate extraction of sets of points according to some neighbourhood metric. This will allow extraction of local surface information, such as derivatives. However, it is desirable to keep the data compact, and therefore this precludes simply building a superset of information. Instead, the generic representation should be the basis from which other forms can be generated as required. If the data for each subject can be limited to approximately 1 MB, then this has the advantage of being suitable for current transmission rates over the Internet.

Where two surfaces are in close proximity (e.g., inner tops of legs, underarms), it is necessary to segment the data. It may be useful, for example, to represent body semantics with anatomically meaningful segments. Since this is part of a larger set of research and development activities, it is important to align with standards, such as *H-Anim* [HAWG99] and remain open to a range of applications. Either the representation should conform to the appropriate standards, or allow an easy transformation path to them.

The representation must convey as high a level of detail as possible. The human body can be considered at many levels of detail, and it is necessary to determine what is appropriate for this work. In practice, the level of detail is bounded by the resolution, accuracy and precision of the initial image data[1]. The accuracy is approximately ± 1 mm [Hori98]. As discovered in the previous chapter, the data density after initial cleaning is approximately 7×5 mm. At this resolution and accuracy, the human body is composed mainly of smooth surfaces, with some cusps (for example, at the underarms and buttocks). Much of the body is approximately elliptical in cross section (for example, the mid torso, head, legs and arms) but with varying rates of curvature. Some fine-level detail is present on the face and fingers. Given the expected characteristics of the body surface, and the precision of the imaging technique, it should be possible to preserve features that deviate 3 mm or so in depth from the surface. In order to

---

[1] An exactly similar issue arises when measuring the length of surface contours. Convex hulls do not suffer in the same way from this ambiguity of detail.

discriminate a distinct component against noise, it is assumed that at least 3×3 sampled surface points are required, and therefore it is appropriate to reconstruct individual body segments that are 2 cm or more in diameter. Thus, for example, it is not appropriate to reconstruct each finger individually, but to treat the hand as a single entity. However, it should be possible to preserve detail such as the edges of clothing against the body.

A secondary function of the whole-body representation is for visualisation. For large-scale practical use the whole-body representation must be generated and rendered rapidly, and allow interaction (rotate, zoom, etc.) in real-time on a moderately powerful computer. A good surface representation would have the benefit of allowing manual interaction with the image, to place landmarks for building training sets for feature detection (but since the aim is for full automation, manual intervention should not be required at *run time*). It should be noted, though, that visualisation itself is not considered to be a high priority, so this work will not be concerned with texture mapping or striving for full realism.

Although this work is based on data from a particular family of scanning systems, it is worthwhile to consider data from other sources when designing the body representation, or choosing an existing one. This will increase the usefulness of the work done here, by allowing the same techniques to be applied to data imported from other image capture systems. In order to give this flexibility, the processing techniques should take into account the varying density, distribution and sampling structure used by different systems, and also any known characteristics of noise, measurement precision and accuracy.

Since surface reconstruction is a large research field in its own right, it is not desirable to get too immersed in this topic. Rather the aim is to find a simple and sufficient representation, and a data transformation method that is easy to implement and computationally efficient.

3.1.2 Representational Forms for Humans

Jones and Kilpatrick proposed a simple, extendable representation of range data, whereby points are segmented and labelled according to the relevant body part [JoKi97]. The structure is hierarchical, allowing further subdivision down to any appropriate level of detail. However, it is based on point data, with no low-level structure implied.

For structured surface representations, it is immediately possible to dismiss simple constructive forms such as fully generalised cylinders, superquadrics and geons. Such representations have been used successfully for *recognition* of complex objects [BoFi97], and for low-complexity representation of dressed human avatars [JGMS98], however, they do not provide the representational flexibility to define the types of surface encountered on the body at a high level of detail in a sufficiently compact way. Similarly, volumetric representations such as octrees and voxels, used, for example for CT and MRI data, are not appropriate, since they do not give relevant surface connectivity information, unless this is explicitly added.

Certain and Stuetzle have used paired stacks of ellipses to model the lower body, and report good results for anthropometry on simulated scans [CeSt99]. However, this representation would be inappropriate for the whole body, particularly the upper torso and shoulders.

Other researchers working with whole-body range data have typically used generalised triangulated forms, as shown in figure 3.1. Interactive editing is usually required to reconstruct complex and occluded areas. The resulting meshes use adaptive polygon sizes to model large smooth areas and areas of detail in an efficient way. No specific symbolic information is contained in these triangulated representations.

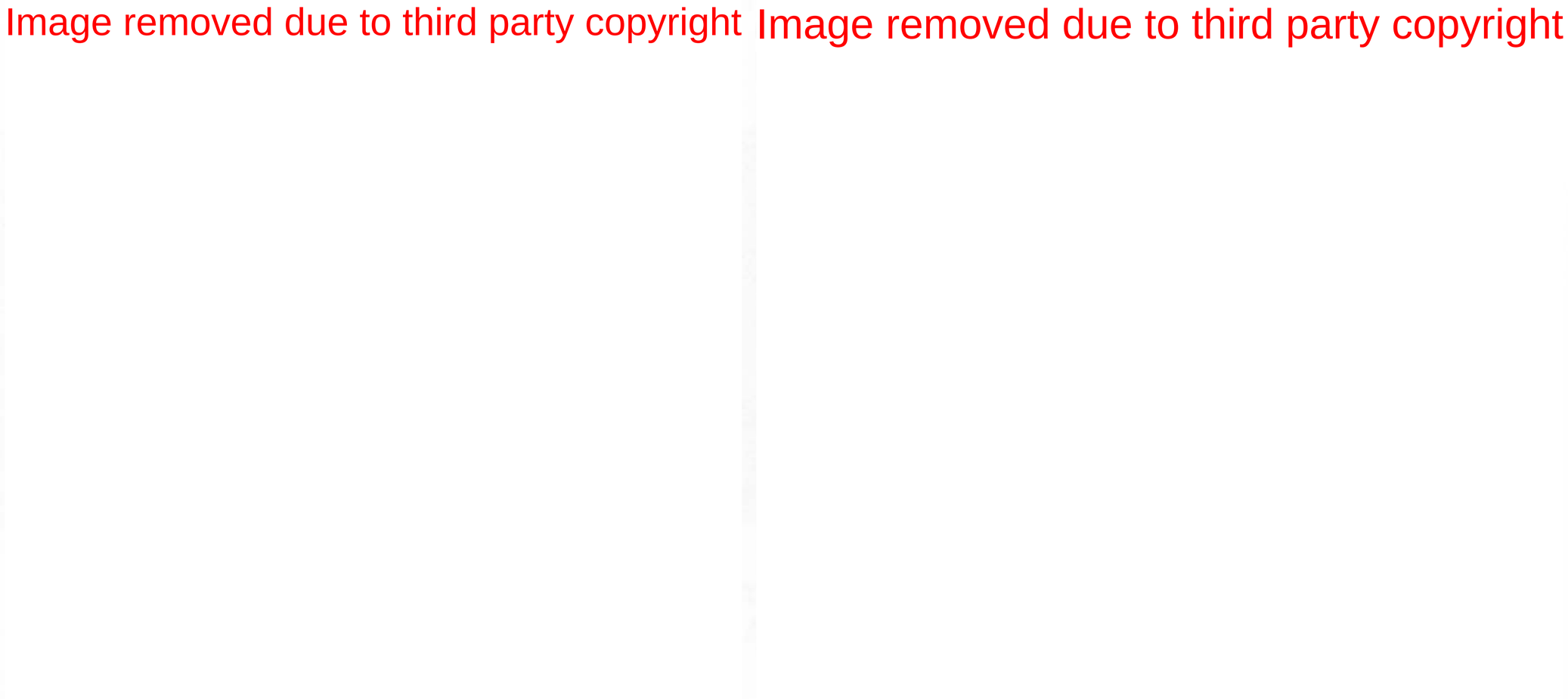


Figure 3.1: Polygonal mesh output of the Cyberware WB4 scanner.
Reproduced courtesy Steve Addleman, Cyberware.

A spline model for the human body has been developed by Li and Jones [LiJo94]. In their early work on whole-body surface imaging they devised a torso representation containing symbolic information of anthropometric significance. The torso is reduced to a matrix of eight cross sections corresponding to key anthropometric positions (neck, bideltoid, maximum bust, underbust, waist, hips and crotch), each using 16 non-uniform rational B-splines (NURBS) control points to define one half curve of the body. The full NURBS surface is interpolated from the key cross sections. This provides a high level of detail of the torso in an extremely compact form. In this form it assumes lateral symmetry, but can be extended with 32 control points to define the full cross section, without the symmetry constraint. This representation only describes the torso, so it is necessary to find a method to describe the whole body, including the limbs and head. Many other algebraic surface representations [StHI94] could be used for efficient and high detail representation of the human form, but little work has been published on this. More recently work has been done on B-spline representations, specifically for automated reconstruction of human whole-body range data [West97] [DoDB99].

Articulated 3D models have been in practical use for some time, beginning with SAMMIE, developed at the University of Nottingham [CaPB90], followed by JACK [BaHG93], RAMSIS [Bubb98] and numerous similar systems [Yead97] [AmBu99]. These are aimed at modelling properties such as reach and centres of mass for performance and ergonomic applications, rather than attempting to represent the human form at high geometric detail. Therefore the outer appearance is typically a set of simplified polyhedral segments. For example, the model developed by Yeadon has eleven segments and is used to model gymnasts in motion [Yead97]. The geometry was derived from 3D scan data captured by LASS [JWHR89]. Amos's model for body positioning, with 17 segments and 16 ball-and-socket joints [AmBu99], is shown in figure 3.2.

Figure 3.2: An articulated model for ergonomic analysis.
Reproduced courtesy Gideon Amos, University College London.

In order to standardise on these models and avatars for virtual environments, *H-Anim* defines skeletal segments and joint structures [HAWG99], as illustrated in figure 3.3. Although the current specification has minimal reference to outer appearance, the next version, will also include a "seamless avatar" format—a regular mesh representation [Smit99]. It is not yet clear if the resolution of this format will be sufficient to deal with the level of detail required for accurate anthropometry. The jointed representation ensures a known data structure when sending command streams for a virtual human's behaviour. Since this is often performed in a distributed manner across heterogeneous systems, the emphasis is on choice of representation appropriate to the application, bandwidth and computational resources available. In addition, *H-Anim* defines surface features based on the CAESAR landmark set [SAE99] with additional features identified during this work [Dekk98].

Complementary to this, a great deal of research is devoted to developing synthetic actors with hyper-realistic appearance. On top of the articulated skeleton, layers of deformable musculature and other body tissue are modelled [SPCM97], for example, using springs attached between the rigid skeleton and deformable skin surface [NeTh98] and metaballs [FGPA98]. Using appropriate biomechanically-based control structures, the aim is to produce movement that is both believable and anatomically correct, and is demonstrated for actions such as running, vaulting and cycling [HoWo98].

Image removed due to third party copyright

Figure 3.3: H-Anim standard for joints and surface sites.
Reproduced courtesy Bernie Roehl, *H-Anim* group.

Skin representations used with these models are typically spline patches and irregular triangle meshes, whose properties can be tuned to produce the appropriate mechanical behaviour; some attempts have also been made to map texture for micro-structural realism of the skin [WuMT95]. Virtual synthesis of clothing is an area of intense research, with the goal of realistic real-time modelling of clothing, conveying the great range of properties of different fabrics, their drape, movement and interaction with the surface of the body [MTYT91] [HiMW92] [Tsop93] [NgGr96] [ADBW96] [JGMS98].

This discussion has explored the types of data structures in use for modelling humans. The transformation and construction of the many "layers" of structure, appearance, motion and behaviour are represented graphically in figure 1.2 in Chapter 1. For this work, the important point to note is the need for a body surface. Since no standard body surface representation is available, the precise details for this are still open.

3.1.3 Human Surface Reconstruction Techniques

In the previous section representational *forms* for humans were discussed; this section deals with the *processes* for generating those forms. Again, this is a vast area of research, so this is not intended as a full survey, but to ascertain the types of approaches used. Different surface reconstruction methods can apply various levels of assumption, and therefore bias, on the reconstruction, thereby determining the specificity or generality of the technique. The discussion begins with generalised methods for surface reconstruction from unorganised points sampled from a surface. Minimal prior assumptions are made about the object to be reconstructed. To make sense of this data, the following sub-problems must be addressed:

- determine the topology of the object;
- infer the surface behaviour in areas of missing data, and interpolate across those areas;
- associate appropriate points to define surface connectivities, either by interpolating or approximating between existing points;
- possibly apply some smoothing operations;
- it might also be desirable to reduce the complexity of the resulting object, by culling detail in areas of low curvature, while maintaining detail in other parts with high-frequency features.

Many good methods for general-purpose surface reconstruction have been developed, based on polyhedral and algebraic surface patches [ORou81] [Bois84] [LoCl87] [Besl88] [HDDM92] [StHI94] [FaPi95] [CuLe96] [AmBK98] [BBCS99]. Although these do not explicitly make assumptions about the *overall* object being reconstructed, by necessity they must at least make local assumptions and apply heuristics to drive the point association or surface fitting process.

For example, points in close proximity *usually* lie on the same surface segment, and nearest neighbours *usually* should be connected. In the whole-body range data this will not be the case where two distinct segments touch, but in order to detect these cases, some semi-local or global knowledge must be applied. The notion of coherence can be used, by preferring to generate surfaces that preserve an observed consistent characteristic, such as a more or less constant curvature. Framed in this way, the problem is effectively one of optimal decision making. A surface is wrapped, deformed, grown or sculpted, driven by one or more objective functions that must be minimised.

Where the reliability of any given data point is in question, the decision about whether or not to include it in describing the surface may be moderated by a heuristic, such as minimal surface area, combined with a "reliability" measure of the points included in the alternative configurations. Uncertainty in the data can arise in active imaging systems, for example, where the angle of incidence of the light is large. Such points are typically considered to be less accurate than those sampled by a sensor operating at a small angle of incidence. In this case the cost function makes trade-offs between the two contributing criteria, which may be in opposition. Besl's region growing method was based on extracting local curvature information and growing spline surface patches by connecting points of similar characteristics [Besl88]. O'Rourke started with the convex hull around the point set, shrink-wrapped the object iteratively by including interior points, while minimising the overall surface area—in other words, maintaining a low bending energy [ORou81].

In Stoddart *et al*'s surface reconstruction method based on the Slime representation [StHI94], the initial surface shape is seeded by a voxel representation, where voxels are placed in cells occupied by a data point. The voxels are then decomposed into a triangulated control mesh that is then deformed with respect to the data. The voxel seeding avoids prior assumptions about topology, but has the drawback of requiring that the data set have no large unsampled areas.

Free-form deformable models [Burr81] operate by deforming an initial shape to fit the sampled points. In its generic form, no object-specific information is used to constrain the deformations. The data imposes external forces on the template, with salient features in the data acting as attractors. As an opposing force, the template is constrained by internal forces to preserve properties such as smoothness and continuity. This class of technique has been used widely and successfully for segmentation and reconstruction in 2D [KaWT87] [JaZL96], 3D [LeTW95] [RBFM97] [DeMo98] and 4D [McTe94]. If the starting template is a closed form, then the resulting deformed shape will also be closed. This has the advantage of guaranteeing a complete surface when data is sparse or missing, in contrast to surface "growing" techniques, which often leave holes where data is sparse.

However, as with any other data fitting method, all these general-purpose surface reconstruction techniques have limited means of inferring the original surface behaviour where large areas of

data are missing. In Chapter 2 some of the key characteristics of the specific whole-body range data were identified, in particular, that the data contains noise, is incomplete in places, and has areas of redundancy arising from the overlapping fields of view of the sensors. In Chapter 1 some characteristics of the human body shape were discussed. To a large extent, it is possible to control the subject's stance during scanning, and to capture the scans in a reference posture, within a consistent coordinate frame. This constrains the general surface reconstruction problem, as the branched topology is known. In addition, some characteristics of curvature were also noted, which should be of help in reconstructing areas of missing data, making decisions about noise, and inferring the appropriate connectivity of points.

For example, the human body has several cusps, some of which indicate separate body segments (such as at the top of the legs), whilst others are only partial (such as on the buttocks), as shown in figure 3.4. There is no qualitative difference between the behaviour of the data in the two slices, despite the fact that the first slice represents two separate segments (the left and right legs) and the second represents a single body segment (the torso). In fact, without additional, higher-level information it is not possible to infer the difference between these types of cusp. By using the expected structure of the data it may be possible to devise an algorithm that is more robust than those dealing with general point sets, and computationally less expensive. In fact, the use of a specific human body model is not only convenient and justifiable, but it may be *necessary* in order to process the data correctly.

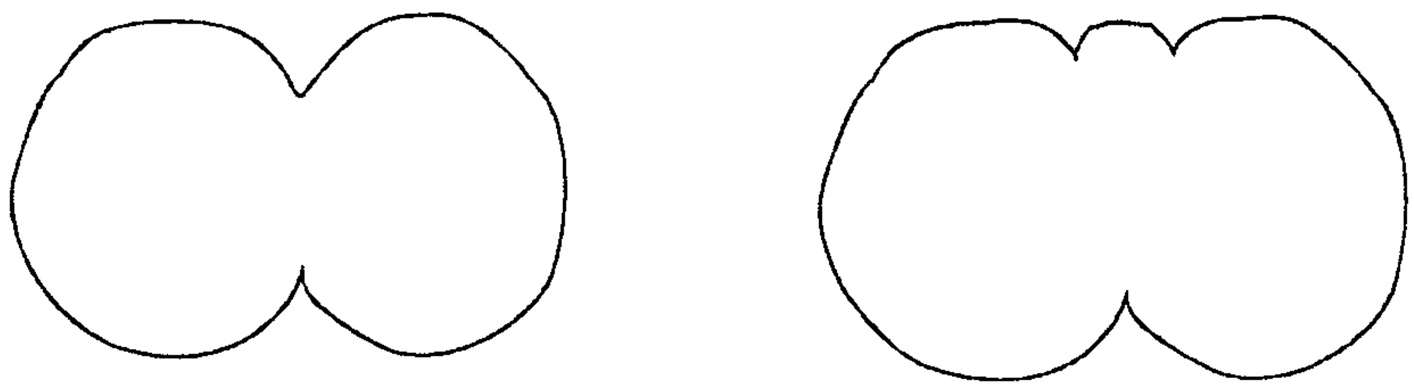


Figure 3.4: Cross sections through the upper legs and lower torso.

In cases such as this, where some *a priori* information is known, *object-specific* deformable models can be constructed. The prior knowledge is used to define the initial shape of a polygon mesh or parametric surface template, and to constrain or bias its deformations with respect to the sample data. These constraints can be programmed by hand, and tailored to the specific class of objects, if the shape variation is well defined, for example for faces [YuCH92]. Deformation constraints can also be constructed from statistical information, for example, via principal components analysis [ScPe95] [CTCG95], to characterise the observed variation within a training set of sample objects. This constrains the deformations to forms that fall within the eigenmode space. Since the specific deformable approach also has implications for feature detection, it will be discussed again in more depth later.

The deformable template approach has limitations, in that the shapes being modelled must be well defined, and suitable for representing in a parameterised form. Fortunately, this appears to be the case for the human body surface reconstruction problem. Li and Jones [LiJo94] use a deformable template method to reconstruct the torso from range data, based on first manually locating key sections, then spline fitting to generate the smooth surface. This method is interactive, not automated, and is designed only for the torso, but it does indicate a way forward for surface reconstruction in this work.

It appears then that no fully satisfactory published method exists that can deal with surface reconstruction for the whole human body from the type of range data available for this study, because of the need for segmentation and reconstruction of missing data. Therefore a suitable method must be introduced that can incorporate the appropriate prior knowledge. A parameterised deformable template approach to surface reconstruction will be explored further in Chapter 4, which is concerned with transforming the initial image data to a complete surface representation of the human body. Of particular interest is the use of anthropomorphic knowledge at various levels, and how this affects the outcome of surface reconstruction.

3.2 Feature Detection Strategies

The next chapter will describe in detail techniques to transform the initial range data to a complete surface representation of the body. For now it is assumed that such a representation is available. This section discusses techniques for extracting features from that free-form surface, despite variation in the overall size and shape of the subjects, and variation in the local shape of features themselves.

The available sources of modelling data and prior knowledge have a strong influence on the choice of approach, in particular, whether it is largely *data driven* or *model based*. The former is appropriate when attempting to build general-purpose systems, with few *a priori* assumptions, or where prior knowledge is simply not available. One difficulty with a data-driven technique is that it normally requires a substantial amount of good quality training examples, which may not be available in this case. Of course, no data-driven method is totally without some underlying model, however simple, so it is not possible to avoid bias entirely. In fact it is essential for any system that attempts to generalise about unseen data. There is a balance between the gains made by capitalising on *a priori* knowledge, and loss through the limitations required to make those assumptions valid. Since it is *not* the aim of this work to produce a general-purpose feature detection method, it is valid to make certain restrictions and optimise the method specifically for the human body, while sacrificing generality. This is in contrast to a great deal of work in machine vision, where the aim is specifically to minimise the *a priori* information built into the system [Besl88]. Nevertheless, it would be desirable to minimise these restrictions, since it

would counteract many of the potential advantages of rapid image capture, and reduce the applicability of the system.

Given the subtlety of the features, as noted in Chapter 1, it is likely that this is a rather demanding task. There is little preceding work in feature detection on the whole body, but the literature does contain a good deal of examples from 2D and 3D feature detection in faces and other free-form objects. These can be categorised into two main strategies, which are discussed here:

- global *deformable surface* methods, and
- *feature-based* methods, which can be further subdivided into:
  - — templates making use of local shape and overall contextual information;
  - — functions that describe or search for features, based on local and global shape, and possibly non-structural attributes.

While these approaches have characteristics in common, and indeed they have been used in combination, it is useful to discuss them in this way to highlight the varying use is made of local shape, global structure and other information.

3.2.1 Deformable Surface Methods

The previous section discussed deformable template methods for surface reconstruction, and so it is natural to see if they could also be extended to use for feature detection. The general idea of the approach would be to:

- define a starting template, with features of interest appropriately pre-placed;
- generate deformations of the template;
- use characteristics of the scan data to affect the deformations;
- constrain the deformations in ways appropriate to the class of object,
- until a termination criterion, e.g., convergence, is met.

Much of the work on deformable templates is based on principal components analysis to derive a model that constrains the deformations in "plausible" ways. Each example from an object class is discretised into a set of descriptive parameters, for example, corresponding landmark points. From the landmark points a mean shape and covariance matrix for the landmarks can be derived, from which the principal components are determined by taking the largest eigenvectors. Effectively, the technique finds the principal axes of the data set by successively minimising the squared distances from the data point to each axis. This generates a new set of variables which, in linear combinations, can characterise a high percentage of the shape variation found in objects of similar type. In the approach devised by Cootes *et al* [CTCG95], this is called a point

distribution model (PDM). When used for object recognition, the model deforms against a new sample of data, and the candidate with the lowest deformation energy, and within the constraints of the training set for that object class, is deemed to be the best match [ScPe95].

This approach models linear relationships successfully, but is not always applicable, especially when bending or local rotation within the object is encountered. In such cases a linear analysis will yield inappropriate shapes that are not representative of the training set. Non-linear methods have been devised, using polynomial regression and multi-layer perceptrons [SCTD97]. These methods still employ linear PCA as a starting point, but then model the relationships between the components in a non-linear way, with better success on object classes for which the purely linear methods fail.

One of the long-standing difficulties of this approach is in placing the initial landmarks with which to build the PDM. The number and location of the landmarks should be chosen so as to model corresponding points on the body, and give sufficient detail to the deformations. In training the model, often key landmarks are placed manually and additional landmark points inferred from them, for example by spacing at regular intervals over the curve or surface [CTCG95]. At first sight it is appealing to think of using anthropometric landmarks as the key points to initialise a PDM, but this has not been done. The landmarks used to train PDMs must be reliably locatable on the training data in the processing medium, in this case surface images. However, as noted in Chapter 1, many of the anthropometric landmarks are based on skeletal features (such as particular vertebrae) or subtle visual features (such as the crease at the back of the knee) [ISO89]. Such landmarks are distinct and fairly reliable when tactile feedback is possible, but they are not "visible" in a straightforward way purely from surface information. If a PDM were to be used it would therefore be necessary to construct the model from reproducible surface control points, and then to attach anthropometric landmarks to the trained model.

Attempts have been made to automate the landmark selection and placement process, as an end in its own right, and for training statistical shape models. Hill and Taylor [HiTa94] devised a method to generate landmarks on objects in 2D, starting from partially marked objects. Pairs of training samples are matched, using various methods of pairwise non-rigid correspondence, iteratively matching the results of those to generate a mean shape. A full landmark set placed on the mean shape can then be propagated backwards to the branches of the binary merging tree to any of the training set. Since these landmark placements tend to be only approximate, optimisation techniques are then used to refine the landmark locations. In that case the objective function relates to the compactness of the point distribution model. A different, application-specific function could be used in its place, for example, based on optimising with respect to anthropometric values. This automated landmark generation method has been adapted in 3D [HiBT97], using a variant of Besl and McKay's iterative closest point algorithm [BeMc92].

Sclaroff and Pentland present a method for point correspondence on unmarked objects [ScPe95], using a finite element method to define the parameter space. Corresponding points can then be found by finding the strongest matches under the movements within each eigenmode. Although good results are reported, it was subsequently not considered to be robust enough, for example, for generating PDMs [HiBT97]. Other approaches have used a measure of "saliency" in scale space [WaCT98] to select features with a high likelihood of detection in subsequent data. This has the advantage of being applicable to higher-dimensional data sets, where "manual" placement of features is not possible.

Magnenat and Delingette proposed a different template method for the reconstruction of range data, based on an iterated, coarse-to-fine matching process [MoDe97]. Prior information is built into the initial template, which then undergoes iterated registration processes, from rigid, to similarity, affine, "hybrid" deformation, and finally free-form deformation. In essence, the degrees of freedom of the deformation are gradually increased. Davatzikos devised a non-statistical technique for mapping tomographic brain images with an elastic surface of variable deformability across regions [Dava97]. Selected features are pre-placed on the initial model, and after deformation, the landmarks are located on the target image to a few millimetres accuracy. These approaches are attractive, in that they avoid the use of corresponding landmarks required to train PDMs.

These methods have mostly been designed to deal with the *general* case of shape analysis. In the case of feature detection on the human body, there is a great deal of prior information available, such as expert anthropometric knowledge and non-structural information such as weight, body mass index, handedness, gender or age, which might affect feature characteristics. Others have devised more specific methods that make use of more prior information. For example, Lee *et al* [LeTW95] used a combined feature detection and deformable template method for animating faces starting with range images. Key features were detected in the 3D data using a Laplacian operator to find edges around the eyes, lips and chin. A generic face mesh was then fitted to the range data, using the known positions of the features to control the mesh adaptation. This makes use of more prior knowledge than, for example, Magnenat and Delingette's method, and provides a potentially more feasible approach for the human body feature detection problem.

Although these look like potentially promising approaches, they still do not escape the problem of accurately placing the anthropometric landmarks on the template model. This could be done by manually locating the features from the surface image, but this is prone to error, and this error would be exacerbated once the model had been deformed. Although this might be sufficiently accurate for the surface reconstruction problem (as discussed in the previous section), or for object recognition, it may not be sufficient for anthropometric feature location. An important point in favour of a deformable template approach is that it is likely to give good worst case behaviour. It will always find the landmarks, and in roughly the correct place, as the

relational information of the features is implicit in the model. In addition, this approach has the advantage of combining surface reconstruction with feature detection in one step.

At this point, these deformable template methods appear to generate a number of problems. This is true, in particular, for the techniques that are based on landmarked training sets, since this is precisely the desired *result* for this human body modelling problem. The methods are not dismissed for future work, but since they do not appear to be fully satisfactory at this point, it is necessary to look at other strategies for feature detection.

3.2.2 Feature-Based Methods

An alternative approach is to characterise and search for features individually. This is interesting in its own right, but could also be used in the future as a precursor to other methods, such as a deformable template. Several representational forms might be used, for example, filters or local templates or a shape syntax. Perhaps the simplest case is a filter or template, designed to match a particular feature. For example, this might take the form of a parameterised spline patch, with statistically derived constraints, characteristic of that feature. If there is sufficient local shape information, and the feature is unique, then it may be possible to match the feature without reference to contextual information. In the case of human body features there is non-rigid variation across a given feature class, so either the template, or the matching procedure must accommodate observed variability [Marq98].

Once a suitable template has been created, the problem is then how to carry out matching in a meaningful and efficient way. This is discussed in the context of face recognition in 3D by Beumier and Acheroy, who devised a computationally efficient way of matching facial contours against data captured by a structured light imaging system [BeAc98]. The generalised Hough transform has been used to match many types of feature from templates, for example, to locate eyes in images [Nixo85]. Genetic algorithms (GAs) have also been used for local template matching, to extract features from greyscale images [RoLe94]. The encoded representation is a set of indexed points that minimally describe the shape being matched (thus giving an extremely compact encoding); the fitness function indicates how well the candidate region matches the template. Lanitis *et al* [LHCT95] used GAs for image search with a statistical shape model, to locate faces in cluttered 2D images. The PDM's encoded pose and shape parameters are manipulated by the GA, to optimise the template's placement and shape within the image.

Figure 3.5 illustrates how a template might be applied, by traversing the body surface until a suitable match is found. For most features an unambiguous match will only be found by making use of contextual information, such as relational characteristics of other features, to help resolve ambiguities or reduce the search space. This has been used in face recognition, where shape and context are well understood, and is similarly applicable here.

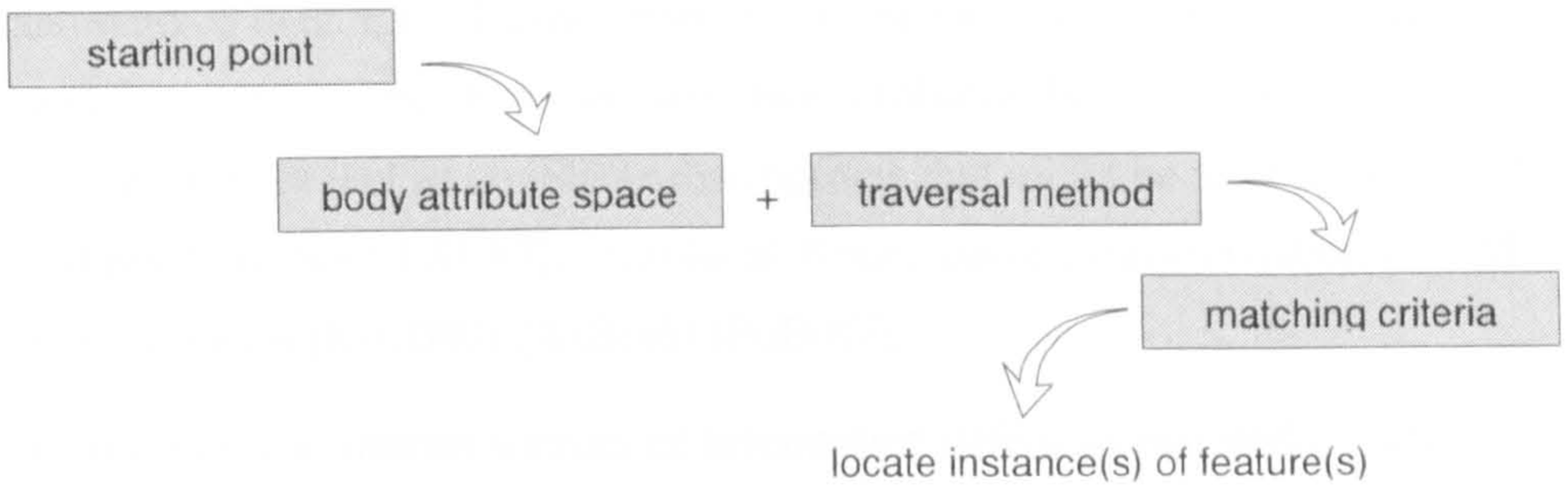


Figure 3.5: Template-based feature detection

A simple, structural syntactic approach could be extended to define a full pattern grammar for each feature, if the essential shape components can be broken down into a suitably parsable form. Hsu and Hwang classify shapes by decomposing the shape boundary into contour segments, which are assigned simple attributes such as segment height and convexity [HsHw97]. Relational characteristics, such as segment angles, are also extracted. Each shape is then transformed into a symbolic representation, as a list of predicates. Asada and Brady's curvature primal sketch [AsBr86] generates a set of distinct curvature event components from a contour, which can then be described in syntactic form as a semantic net showing relational characteristics.

Rather than building a template and a separate matching procedure, a feature detection function can be defined that describes how to search for the feature. This has many similarities with the template-based approach, but instead the discriminant attributes are embedded in the matching process itself, rather than a template. Figure 3.6 illustrates how this approach might be applied.

The potential of this approach in surface anthropometry from range images has already been demonstrated elsewhere [Nurr97] [PaSD97], but little specific detail is given of the technique in these cases. However, numerous examples can be cited in characterisation of faces for recognition and authentication, both in 2D greyscale images and full 3D [KoMK97] [BrPo93].

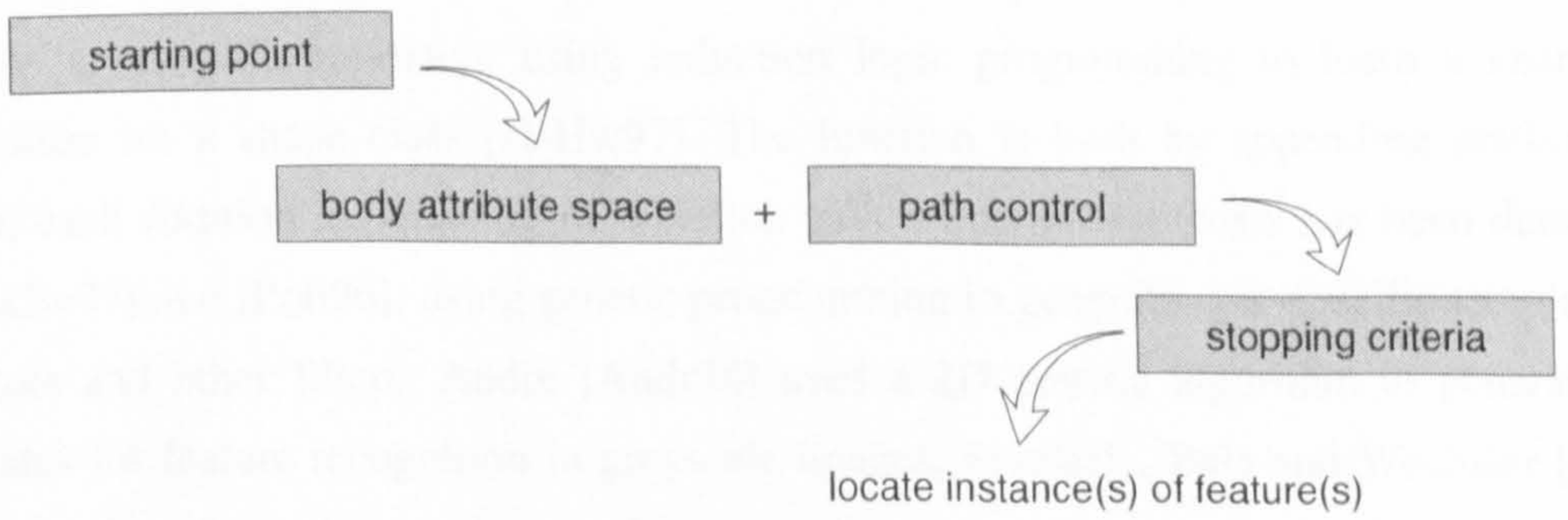


Figure 3.6: Function-based feature detection.

The procedural approach is particularly attractive because anthropometric specifications are presented as procedures for locating features, and therefore lend themselves to translation into procedural form for a machine. In addition, a significant advantage of this approach over deformable template methods is that it can be implemented without the need to build a model

from hand-marked data. Each feature detector can be modelled with a different representational form: input variables can be shape and structure attributes, but may also include attributes such as age, gender, weight and so on. Shape descriptions that might be used include scalar attributes such as elliptic variance [IPSV97], semi-local Bezier curve characterisation [CiLM98], or local curvature description [KovD92] [AsBr86] [PoBr87].

While the use of these various sources of information offers an extremely expressive hypothesis space for feature detection, it creates the complementary problem of generating and selecting the appropriate attributes and representational form for each feature detector. Fortunately, contextual information can be used to constrain the search space. This is used heavily in Brunelli and Poggio's work on face recognition in 2D [BrPo93], where feature-based functions based on edge maps, and local greyscale templates were compared. Kondo *et al* also used specific context for feature detection in 3D to generate facial caricatures [KoMK97]. In both cases, progressive use was made of newly discovered features in localising the search, from an initial global space, to increasingly smaller and well-defined contexts.

Much of the work discussed in this section is directed towards the analysis of facial images in 2D and 3D. Since this is a problem domain of particular significance, it is perhaps justifiable that a great deal of effort has been spent on hand crafting techniques for it. The analysis of whole-body data is of similar significance. Nevertheless, it is clearly desirable to employ techniques that can be tuned to the problem in an automated way, to reduce the work required, and to allow better refinement than might be possible with a purely manual approach. Many learning techniques have been used to construct models to recognise features and objects, for example, neural networks [TuPe91], support vector machines [OsFG97], inductive logic programming [HsHw97] and evolutionary computation [Andr94] [Poli96] [BaWe96] [HaBu97]. Of particular interest are learning techniques applied to the synthesis of procedures or syntactic pattern descriptions from primitives. Draper [Drap96] used Markov decision models to generate hierarchical functions from a set of machine vision operator primitives. Hsu and Hwang describe a syntactic approach, using induction logic programming to learn a characteristic description for a shape class [HsHw97]. The function is built by appending predicates to a clause, each addition maximising information gain. Function synthesis has been demonstrated by [HaBu97] and [Poli96], using genetic programming to generate task-specific templates, edge detectors and other filters. Andre [Andr94] used a 2D genetic algorithm to generate logical templates for feature recognition in greyscale images. Similarly, Bala and Wechsler [BaWe96] evolved morphological elements and the sequencing instructions for applying those operators to extract discriminant features. These features were then fed in to an inductive learning system to generate classifiers, which, in turn, were also optimised via evolutionary computation.

Numerous other examples of evolutionary computation in machine vision are given in [BoAl97]. These serve to demonstrate the potential for automated synthesis of new feature

detectors. The immediate problem, though, is to gain enough understanding to generate the initial shape, structure and process primitives that could feed into such an automated method.

3.3 Shape Description Techniques

The previous section discussed overall strategies for feature detection. This section looks at specific methods for describing shape characteristics that could be employed for feature detection. Shape characterisation is a major topic in machine vision, being central to object recognition and feature detection. Loncaric [Lonc98] provides a recent survey of shape analysis techniques, which covers many descriptive forms. Early attempts at surface shape characterisation from range data focused mainly on detection of particular features such as planes, edges and corners. In contrast, this discussion is focused on techniques that could be applied to the human body feature detection problem. This discussion makes the distinction between techniques that give a high-level description of shape (e.g., body mass index), and *information preserving* methods that permit full reconstruction of contours or surfaces from the description (e.g., a parameterised spline surface model).

3.3.1 Differential Geometry in Shape Description

Differential geometry has been used as a basis for many general curve and surface shape description techniques [DrNa81] [AsBr86] [PoBr87] [Besl88] [Fish89] [KovD92] [Mokh95] [LeTW95] [DoJa97]. For example, in Lee *et al*'s work on face animation from range data, local changes in curvature corresponding to eyes and lips, etc., were detected by zero crossings in the Laplacian field [LeTW95]. This then allowed a deformable mesh model of the head surface to be matched against those features.

At any given point on a surface the shape can be characterised by the principal curvatures, corresponding to extrema of the normal curvature function, in other words, the minimum and maximum curvature, usually denoted κ_1 and κ_2 respectively. Gaussian and mean curvature can be derived from the principal curvatures by taking their product and arithmetic average. These quantities provide position- and orientation-invariant descriptions of surface features. At umbilic points all normal curvature values are equal (peaks and pits); saddle points are where the principal curvatures are of opposite signs; ridges and valleys are where one value goes to zero.

If a Gaussian smoothing filter is applied to a contour or surface, its shape can be studied at multiple levels in "scale space", by increasing the width of the filter [Witk83]. Changes as a function of scale may then be tracked as significant features in the scale space. Examples of such features include contours that indicate where the zero crossings of curvatures exist and disappear. This has been used as a function of contour length and scale, for recognition and

retrieval of 2D shapes [Mokh95]. Based on Marr's primal sketch to extract low-level features from greyscale images [Marr82], Asada and Brady devised the curvature primal sketch [AsBr86], which describes a contour in terms of key "events" of significant changes in curvature at multiple scales. The contour is decomposed into discrete shape primitives, which may be represented by single events, such as a point of inflection, or multiple events, such as a dent in a flat surface. The primal sketch treatment was subsequently extended to the *surface* by Ponce and Brady [PoBr87]. This approach is of particular interest for the human feature detection problem. For example, consider the neck and shoulder shape on which the neck base and acromion features are located, as illustrated in figure 3.7. The challenge would be to find the appropriate contours to study, a suitable range of scales under which to analyse the contours, and to determine the discriminant curvature events that allow the features of interest to be detected.

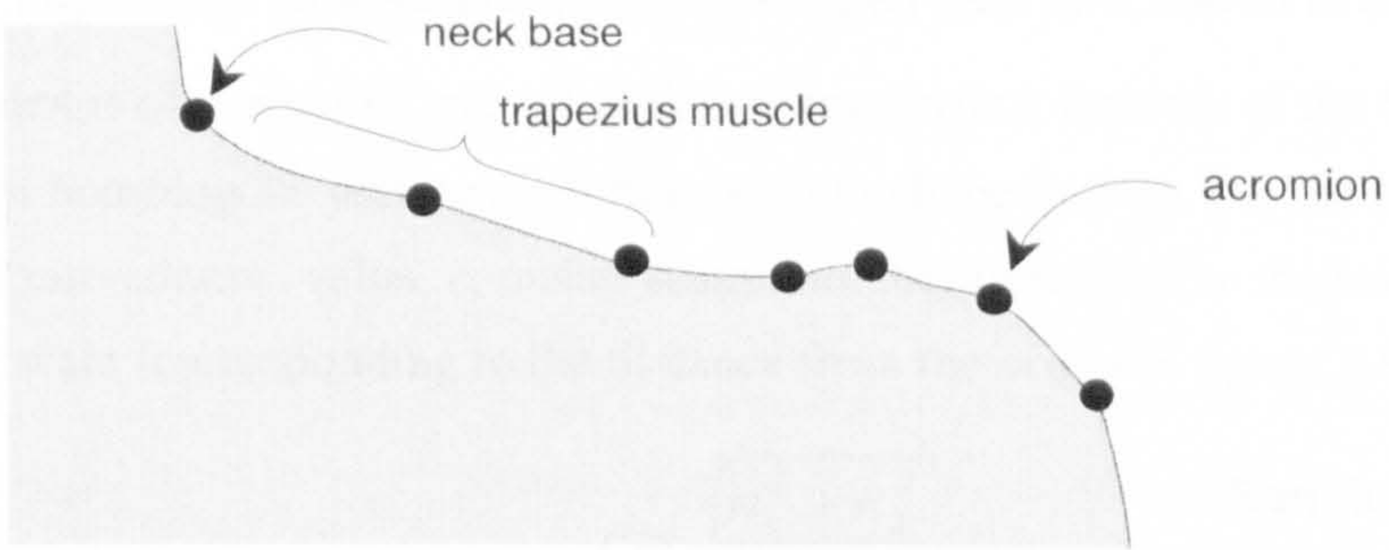


Figure 3.7: Curvature events on the shoulder contour.

In order to describe the shape of a surface both Gaussian and mean curvatures, or both principal curvatures, are needed. Koenderink and van Doorn [KovD92] noted that if the κ_1, κ_2 parameter space is considered as a polar coordinate system then the angular value corresponds to unique shape, and the radial value is a scaling factor. This is illustrated in figure 3.8.

Thus, most of the useful information can be represented by a single value, the shape index, s, which is defined as:

$$s = \frac{2}{\pi}\arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right), \text{ with } \kappa_1 \geq \kappa_2 \qquad \textbf{Eqn. 3.1}$$

which ranges between ±1. Between –1 and –0.5 are the fully convex shapes (peak to ridge); between –0.5 and +0.5 are the family of saddle shapes, with the symmetric saddle at 0; between +0.5 and +1 are the fully concave shapes (valley to pit). Complementary shapes have equal and opposite s values.

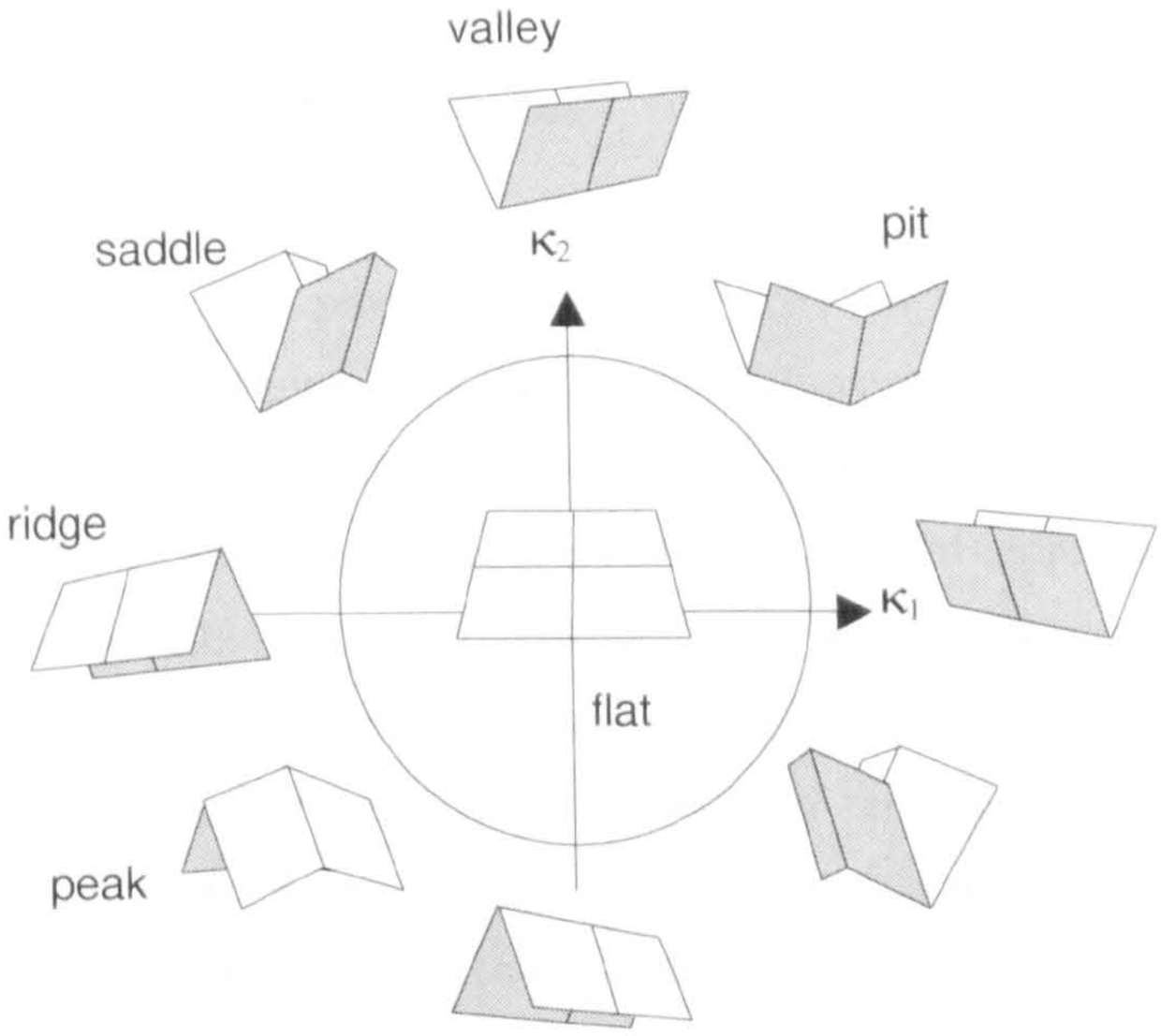


Figure 3.8: Characteristic surface types as a function of κ_1 and κ_2.

The shape index is of particular interest in describing surface features of the human body, where it is likely that homologous features will have similar shape indices, but the scale of the features may vary. A "curvedness" value, c, makes distinction between similar shapes of differing rate of curvature, or scale (corresponding to the distance from the origin in figure 3.8):

$$c = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}} \qquad \textbf{Eqn. 3.2}$$

A modified version of Koenderink and Van Doorn's shape index [KovD92] is used by Dorai and Jain [DoJa97] to produce frequency histograms of surface shape—known as "shape spectra"—for object recognition. This work is similar to Horn's extended Gaussian image surface, which uses normal histograms to characterise convex shapes [Horn84].

The aim of these curvature-based shape description techniques is to describe *viewpoint invariant* properties of shape, independent of position and orientation. However, in the case of the human body the position and orientation of a feature are of particular importance. If shape is defined with respect to a *specific* coordinate system, it should be possible to exploit the differences in position and orientation of points *relative to each other*, to give a richer vocabulary of characterisation.

Where the data points are discrete samples of a surface, rather than an analytic description, a difficulty arises in determining the derivatives, particularly in the presence of noise [Besl88] [FlJa89] [LeHD93] [StIW95] [TrFi95] [GuMe97] [KaHä98] [TaMe99]. Typically this has been done by local curve and surface fitting, and generating the derivative and curvature information from the analytic form. This is rather computationally intensive, and it would be desirable to find some more efficient method, without sacrificing too much in the way of accuracy or robustness. These are important implementation issues, explored further in Chapters 4 and 5.

3.3.2 Semi-Local and Global Attributes

At a larger scale, whole contours and surfaces can be described in fully information preserving ways. Moment invariants were proposed by Hu [Hu61] for pattern recognition. Low-order moments can be used to determine global attributes such as centre of mass and principal axes; with higher-order moments, the whole contour can be reconstructed fully. Transformations to forms such Fourier descriptions [ZaRo72] often provide a useful representation, and have been used by Li [Li97] for describing cross-sections of the human body. Separation of the description into components of increasing frequency level makes it possible to analyse curves at an appropriately chosen level. For example, truncation to remove high-frequency components serves to remove noise and local distortions, while preserving the characteristics of the underlying shape. This has been used, for example, in face recognition, to describe key high-level invariants [JiNi95]. Wavelets allow a similar multi-resolution treatment, and have been used to analyse spinal shape in children with scoliosis [AyMa96].

The parametric form of representations such as NURBS, thin-plate splines and Bezier curves makes them convenient for shape description and comparison. Jones *et al*'s shape matrix for the human torso [JoBW95] describes the surface by NURBS control points at anatomically useful cross sections, and so allows for comparison of corresponding parts of the body in a meaningful way. By decomposing objects to surface patches and curve segments in such parametric forms, semi-local features can be characterised. This approach has been used for many applications, including face recognition, by [SaCi97] [BeAc98] [CiLM98].

It is also possible to describe the shape of contours based on some simple scalar attributes, as illustrated in figure 3.9. Simple attributes such as these have often been used in industrial machine vision systems, for example, for the recognition of flat engineering parts. Of interest here is the fact that by integrating such attributes over the whole 3D shape, simple scalars such as surface area, volume and surface area to volume ratio may be obtained and used to characterise objects partially [LiPe82].

Histogram-based techniques are useful in characterising the distribution of simple attributes across a larger entity. The chain-code histogram [IPSV97] characterises the frequency of each chain type on a contour represented in Freeman's chain code [Free60]. Closely related to this is the pairwise geometric histogram, which characterises shape in 2D based on a frequency histogram of angles and distances between pairs of line segments. This is reported to show good properties for object recognition of rigid shapes in the presence of noise, occlusion and clutter. This can be extended to 3D by analysis of polyhedral faces, and has been used for surface matching to register multiple range images [AFRW98].

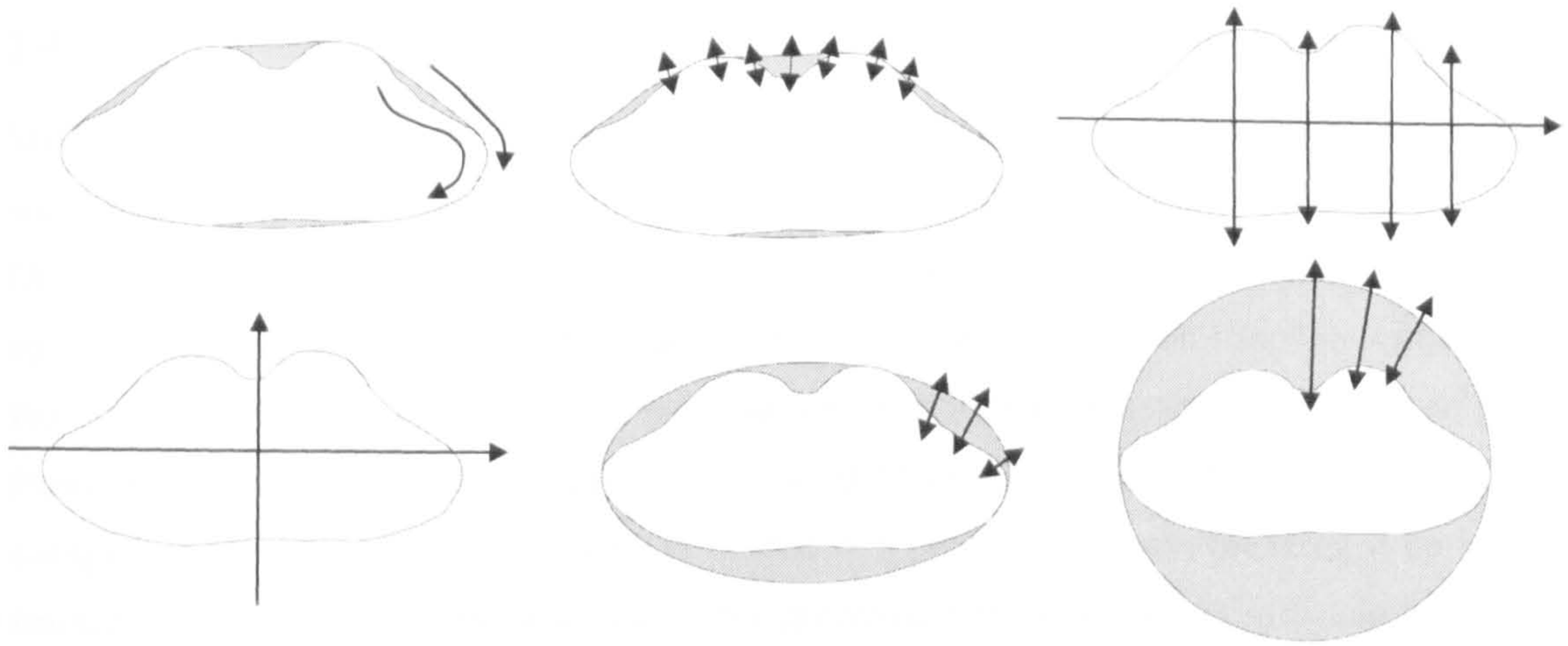


Figure 3.9: Global shape attributes

(a) convexity = l(c)/l(s), where l(c) = length of convex hull, l(s) = shape boundary
(b) convex deficiency = (a(c)–a(s))/a(c), a(s) = area of shape, a(c) = area of convex hull
(c) symmetric width difference = mean and variance of width perpendicular to principal axis
(d) aspect ratio = of lengths of principal axes
(e) elliptic variance = variance of shape boundary with respect to an elliptical boundary
(f) circular variance = variance of shape boundary with respect to a circular boundary

Although high-level scalar characteristics are attractive for their simplicity and compactness, such attributes inevitably allow ambiguity, since many contours of different shape will possess, for example, the same eccentricity measure. While none of these simple descriptors is likely to describe a feature or object uniquely, they may be useful in combination. Iivarinen et al [IPSV97] explored operators for recognition of irregular objects and found that a combination of five simple descriptors (convexity, principal axes, compactness, circular and elliptic variance) showed good performance with low computational complexity as compared with more computationally expensive operators such as chain-code histograms and pairwise geometric histograms. Eggleston used the aspect ratio of the principal components and the convexity measure [Eggl92]. Similarly, Mokhtarian *et al* use eccentricity and circularity, combined with a curvature scale space method, to focus the search in an image database [MoAK96]. In detecting point features on the human body, it is possible that these operators could be combined to locate features in an unambiguous way.

Clearly there is a rich set of shape description techniques that could be employed in designing feature detection methods. The main challenges are to find appropriate techniques that are effective, reliable and efficient. As will be seen in subsequent chapters, differential geometry is used as the primary means of shape description in the work presented in this thesis.

3.4 A Strategy for Progressive Human-Body Range Interpretation

Techniques have been discussed for representing the human body, and the preceding section has briefly reviewed methods for detecting specific features. This section describes and provides a rationale for the approach taken in this work. The proposed research is outlined, together with key issues for investigation in the chapters that follow.

3.4.1 The Overall Approach

Structure will first be imposed on the image data by segmenting it and reconstructing the skin surface. This will be done using a template method designed for human-body range images. Once the initial data has been converted to a surface form, it is then possible to start detecting specific features on that surface. For each feature a pattern recognition function will be devised that searches for using local shape information. Previously amassed knowledge of labelled features and critical surface points will be used to constrain the search. In this sense, the complex image interpretation problem is treated in a progressive way, breaking it up into more tractable sub-parts. In terms of a system for processing the body scans, this can be translated into the following steps:

1. clean the data;
2. segment into high-level body components (head, torso, left leg, etc.);
3. reconstruct data in sparse areas;
4. generate the skin surface;
5. locate first set of features;
6. locate next set of features, using the context of existing known features;
7. repeat 6 until all features are found.

From a more abstract view, each point in the chain of processing corresponds to a particular model of the human body. The segmented and reconstructed surface corresponds to a generic body model, from which other, more specific, models are derived. Each model corresponds to a different "view" imposed by the particular application. This is illustrated in figure 1.4 on page 19. The model gradually becomes more specific, makes more assumptions, has a higher derived knowledge content, and provides more support for the particular application area.

3.4.2 Rationale

This progressive modelling approach is proposed primarily as a means of making the problem more tractable by breaking it down into sub-problems. Once the body has been segmented, the branched central axes can be constructed, the surface can be skinned, and gross information about the body, such as height, body volume and surface area can be determined. After this, a search for specific features can made, in progressively refined volumes, as symbolic information is gradually amassed. This symbolic information can also be used to adapt the cleaning and interpolation operators for different parts of the body, based on prior knowledge of susceptibility to noise, expected shape, and so on. This reflects the processes used by a human performing the same tasks, and also the typical structure of anthropometric specifications

[ISO89]. A similar approach has been used in several cases for detection of facial features for recognition [BrPo93] [KoMK97], and is proposed by Nurre [Nurr97] for the human body.

The assessment criteria set down in Chapter 1 present some opposing influences. For example, while generality, minimum intervention, broad applicability and robustness will tend to increase the complexity of the system, the demand for speed and low computational complexity will make it necessary to make trade-off decisions at many stages. The framework presented here should make these choices easy to implement, and to change in the future for new applications with different performance criteria.

The use of differential geometry as the main shape description mechanism is justified on the basis of a great deal of prior work on similar free-form shape description problems [AsBr86] [PoBr87] [LeTW95]. In Chapter 5 various transformations will be explored for their usefulness in locating key feature types, with the purpose of building a library of feature detection primitives. This exploratory, function construction approach potentially gives scope for discovering interesting emergent properties of the human body. The function building approach is convenient in that it lends itself to automated refinement by parameter optimisation, and the automated synthesis of new feature detection procedures [Drap96] [Poli96] [HaBu97], however this will be left as a topic for future work.

3.5 An Integrated System for Human Body Modelling

In order to explore and test the approach outlined in the previous section, an integrated system is needed, which will facilitate experimentation and demonstrate fully automated processing. Since no appropriate software is available, a system must be built as part of this work. Thoughtful design is needed at this stage, in terms of data structures and modular components, to ensure full flexibility and extensibility for a range of potential future applications.

In particular, the system design is such that modules can be replaced if more effective and efficient techniques are devised in the future, and if the computational platform has sufficient resources. The modular design also allows the same internal components to be used for a range of applications. Figure 3.10 illustrates this layered architecture, showing the components that will be developed in the chapters that follow.

Although it is useful for computational efficiency to use as much knowledge as possible about the image capture process, it is desirable for an open system that these hardware-dependent parts are localised as much as possible at a low level of processing. This is addressed by providing a *hardware interface* layer, which contains most of the processing that is specific to the image capture process. This makes it easier to replace the input data by another hardware system. In this work the automated system will be linked to three different versions of a family of whole-body scanners, with varying imaging components, numbers of data points, data density and scan

volume. This is managed by a configuration file, which contains appropriate parameters for each hardware type.

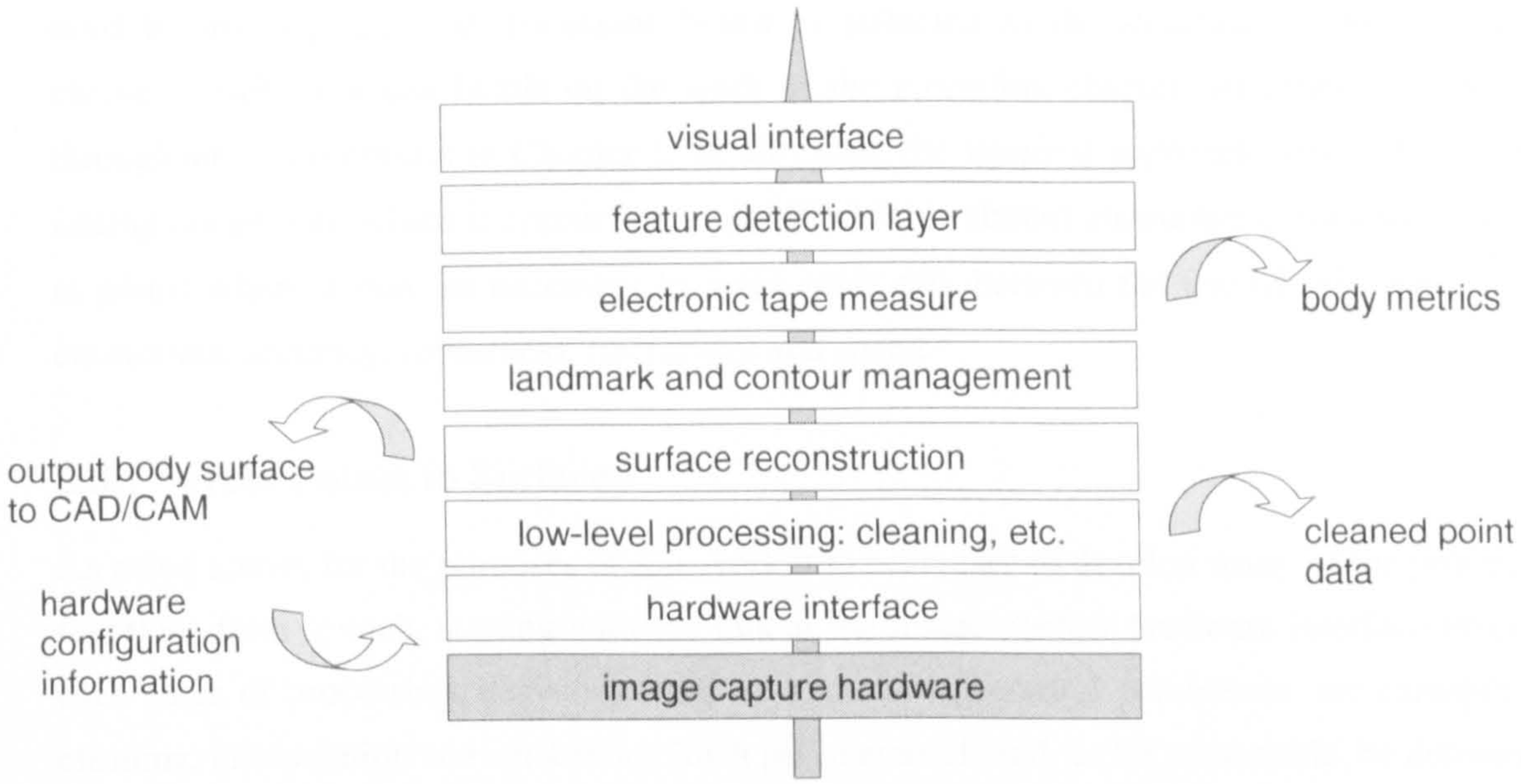


Figure 3.10: Layered architecture of the body modelling system.

All work is carried out on a Dell Optiplex GX1 with a 400MHz Pentium II processor and 128MB RAM. Implementation is in C++, using the Open Inventor 3D graphics libraries. This provides a good range of object-oriented graphics primitives at a convenient level of abstraction, thereby allowing this work to concentrate on the problem in hand, rather than expending effort on low-level graphics functionality. In addition, the library is portable to Silicon Graphics machines, which may be of benefit in the future.

In order to encourage other users to experiment and provide feedback during development, the user interface should be straightforward and intuitive. Where possible, operations should be fully automated, but it should also be possible to override any automated decisions (e.g., to relocate inaccurately placed landmarks), and to extract size and shape information interactively by manually placing landmarks and using these as anchors for an "electronic tape measure". Interaction with, and navigation of, the 3D image should be as direct as possible, so that it can be rotated through any angle, zoomed, etc. As a development environment, it must also provide diagnostic tools and full access to parameters for data cleaning and pruning.

To be useful to others, the system is designed to provide output in a number of formats. This has the benefit of encouraging other researchers to build on the results, add extensions and develop more sophisticated techniques for human-body surface modelling [West97] [Dour98] [Carr98] [Yama98] [DoDB99] [OlBu99].

3.6 Issues for Investigation

In exploring techniques for feature extraction, and building the automated system, many issues must be investigated. The discussion below is reflected in the structure of the subsequent chapters, each of which builds on the work of the preceding chapter. Attention will be paid throughout to the criteria in Chapter 1. In all cases, the simplest approach will be taken, only adding complexity where it appears to be justified. This should encourage conscious decisions at points where it may be necessary to make trade-offs between the use of prior knowledge, constraints, accuracy, robustness, tractability and speed.

3.6.1 From Points to Surfaces

As noted above, for the purposes of this work it is necessary to develop many of the processing functions from scratch, starting with the link at the image capture hardware interface level. At each stage of processing, decisions must be made about control parameters, for example, for cleaning, interpolation and smoothing. Such parameters should, as far as possible, be determined and justified by measurable characteristics intrinsic to the data, such as the sampling density. It will also be useful to note sensitivity to each parameter.

The underlying representation used for each whole-body image directly affects the operation of the modelling technique, and should be chosen so as to encourage the discovery of the best feature detection operators. It is expected that the restoration of missing data may be problematic.

Measuring the performance of cleaning and surface reconstruction algorithms on live subjects is difficult, since reference data cannot easily be obtained. Initial evaluation will thus be purely visual, which should give a good first impression of the performance of the techniques. Volumetric comparison will be used to give quantitative results, using underwater weighing and air plethysmography techniques as reference standards [FJLC92] [MGBM95]. Performance must be compared across different body types, to determine if there are any significant biases or trends. For example, the test set should include male and female adults and children, of a wide range of body mass indices.

Part of the requirement of minimal intervention is to devise a system that can operate fully automatically. For each scan it would be useful to generate a confidence measure to indicate if the data is likely to generate accurate results. For example, if the arms are held too close to the sides, then the upper torso measurements are likely to be inaccurate. It may be possible to devise an automated method to determine this, for example by a kind of "syntax checking". This is particularly important for self-scanning and for very large-scale processing, where there would be no reliable human observer to notice "obvious" errors.

Many of the issues raised above relate not only to these early stages of the processing, but will be of relevance throughout this work.

3.6.2 From Surfaces to Features

In Chapter 5 techniques for shape description will be investigated, and a library of operator primitives appropriate for the human feature detection problem will be built. The purpose of the library is to provide computationally efficient components with which to build detectors for specific features and specific applications, as demonstrated later in Chapter 6. Where appropriate, strict requirements will be relaxed to reduce complexity. For example, it might be quite valid to assume that in most cases a simulated tape measure is constrained to describing planar curves (e.g., the arc between the acromion points), unless explicitly stated otherwise (e.g., the multi-plane, multi-segment contour from nape to acromion, to elbow, to wrist).

To provide sufficient means of expression, the library should contain low-level operators, to extract quantities such as the coordinates of a given point on the surface, surface curvature values, open and closed contours through a given set of points, and attributes of those contours. In other words, the library should be expressive enough to describe the features fully, but should also allow a compact expression. This is particularly important, if automated synthesis is to be attempted in the future, to reduce the combinatorial explosion encountered, for example, when using genetic programming with low-level primitives [HaBu97]. As yet there is no clearly defined primitive set for this. Other researchers using such a constructive approach have started with library functions or a well-judged set of primitives chosen by an expert in the application domain [Koza92] [Drap96] [Poli96] [HaBu97].

At this stage it is not necessary to enforce strict language design constraints on the library. In the future, if automated synthesis is to be done, then it will be necessary to consider these issues in a more rigorous way, by showing sufficiency, closure, and so on. For now, the most important issues are to provide an expressive, but compact representation.

3.6.3 Application and Evaluation

In Chapter 6 the operators developed in Chapter 5 will be combined and applied to a real anthropometric problem. Anthropometric data, such as curvilinear distances between landmarks [ISO89] will be taken from a number of subjects, to validate the location of features. One of the most important challenges is to devise a systematic method for building feature detection functions, based on the anthropometric specifications, local shape attributes and contextual information provided by labelled features.

In measuring performance of the automated system some very practical problems must be addressed. Automatically generated anthropometrics will be compared against manually collected measurements. This raises many problems. Perceived “errors” will, in many cases,

simply be a result of the different media in which the data are collected. For example, while a real tape measure deforms soft tissue slightly, the electronic method presented here works on a still image, which does not deform. The compromise on choice of scan posture may be significant on certain measurements that are particularly sensitive to movement. These issues should be noted and dealt with appropriately. In addition, it is important to obtain information on the reproducibility of the automated system itself, as distinct from inter-technique variation.

The techniques will be tested on a specific group of body types—women of medium build—rather than attempting to deal with all body sizes and shapes. This choice is typical of any pattern recognition system, where a decision must be made between generality and specificity. If a notional "body type" variable could be isolated, performance could be plotted against this variable, as sketched in figure 3.11. The "no free lunch" theorems of search and learning [WoMa95] show that neither can be considered "better" in a general sense (the areas below the curves are the same). It is therefore the particular applications' requirements that should dictate the choice of performance profile of the system. Note that this brings another requirement for the system to detect the body type and select the appropriate model to apply, adapt accordingly, or recognise the limits of its performance and degrade gracefully.

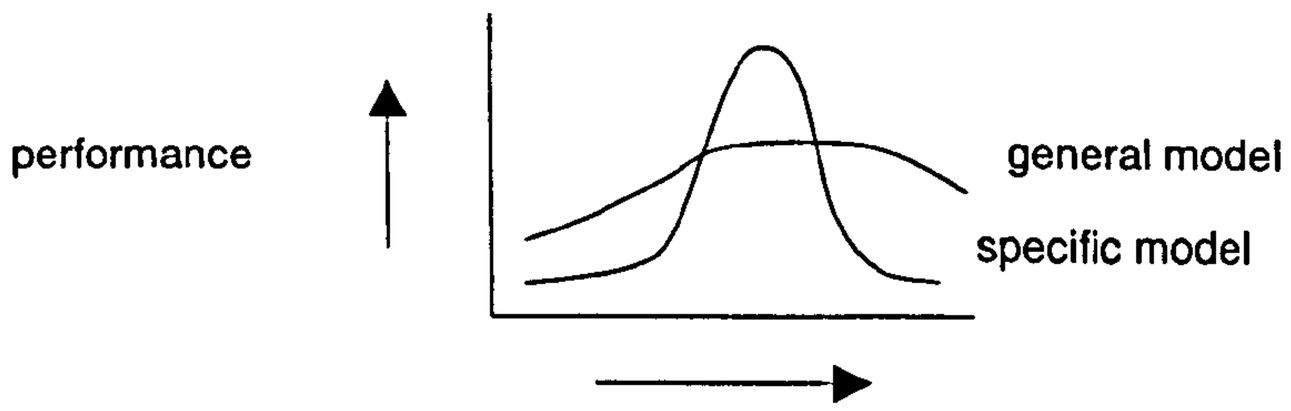


Figure 3.11: Choice of generality versus specificity

Various criteria have been proposed to judge the performance of anthropometric techniques [GBCC89] [BrGr99]. It is thus necessary to know what is *relevant* to the applications addressed in this work, and others illustrated in figure 1.1 of Chapter 1. In selecting the right off-the-peg clothing sizes to fit, the "grading" between sizes is typically about 5 cm, and therefore millimetre accuracy is not required, but centimetre accuracy is. These "acceptability" measures are discussed in detail at the end of Chapter 6.

Certain other performance characteristics will be difficult to quantify, for example, intrusion on the subject, and ease of processing. It will be assumed that any constraints on the subjects' clothing, or restrictions on posture, and time for processing will count against the performance.

3.7 Summary

Whole-body representations have been discussed for feature detection and for other body modelling activities such as animation. While the latter is outside the scope of this thesis, it is important to place this work in context, by considering the requirements of other activities. A polygonal surface mesh has been chosen, as a simple representation that will provide the required functionality. A deformable template approach will be used to generate a suitable mesh from the raw scan data.

Strategies for feature detection were then discussed. A progressive, function-based approach was chosen, which uses search functions to find specific features, thereby interpreting the body surface in increasing detail. Local curvature will be used to control the search functions, together with the amassed contextual information. An overall system architecture has been proposed, and finally, some issues have been identified for investigation at each stage in the chapters that follow.

Chapter 4

From Points to Surfaces

This chapter describes the transformations from the raw sampled data to a structured surface. Starting with surface reconstruction techniques that use minimal assumptions, further a priori knowledge is applied to improve the results. After visual assessment of the skinned images, they are validated by comparison of the enclosed body volume with hydrodensitometry and air displacement plethysmography.

4.1 Data Modelling for Surface Reconstruction

In Chapter 2 some qualitative experiments were carried out to assess the behaviour of the range data over the various surfaces of the body. At this point it is possible to define some basic characteristics of the data:

- The n data points are sampled from the surface, S, giving a point set, $P = \{p_1,.....,p_n\} \subset \Re^3$, which is assumed to contain noise.
- P is represented as a matrix of points $p_{[row][column]}$, where *row* and *column* correspond to the sampling sequence from top to bottom, and from the right side in an anti-clockwise direction. When referring to a single row of points, the first index will usually be dropped from the notation.
- Each point, p_i, has an intensity value, I_i.

Because of the overlapping of the imaging units' field of view, in the horizontal plane, the horizontal sequence does not imply a true ordering of points in space. At this point in the processing, very little has been done in the way of cleaning, except intensity and spatial thresholding, described in section 2.3.4. It is clear that considerable work is now needed to reconstruct a complete and smooth surface for visualisation and, more importantly, feature detection. To build this surface, several related sub-problems must be addressed in this chapter:

- *detection, cleaning* and *smoothing* of systematic errors and noise;
- *segmentation* and labelling according to body part;
- *restoration* of the surface in sparsely sampled areas;
- building a smooth *surface* structure on the data points.

For these processes to be effective, they must adapt to the particular properties of the data, which varies over the body, and with different image capture systems. This section investigates how these processes can be controlled by the known and measurable characteristics of the scan data, such as:

- error characteristics;
- point intensities;
- sampling density;
- expected size, shape and location of features.

4.1.1 Range Measurement and Errors

Coordinate values are measured in tenths of a millimetre, corresponding to the length that a bit represents in the original distance values (see section 2.3.2). The same units are used in all internal calculations. The mean error of each depth measurement within the working range of the system is taken to be ±1 mm [Hori98]. In modelling for surface reconstruction it is desirable to make use of as much prior knowledge as possible of these expected errors, in terms of magnitude, frequency, location and distribution. This section first discusses some known sources of error, and then how to make use of observed data characteristics and prior knowledge to detect and remove systematic errors and noise.

The investigations in Chapter 2 indicate that the error is likely to be higher in some isolated areas, in particular, on surfaces with very high specular reflectance and at dihedrals, such as on the inside of the arms and legs and around the chin. This prior information might usefully be taken into account when designing operators for cleaning, smoothing and reconstruction, for example, by applying a different smoothing technique in areas where higher levels of error are expected.

When the light beam hits the body surface its diameter is a few mm, so that where the angle of incidence is very large, the depth range, δ, over the beam width becomes significant (figure 4.1, case b). In these cases the expected depth precision would be less. In addition, further from the LED source the beam diverges slightly, so that the point size increases, with a corresponding decrease in precision. This, in part, determines the valid working range of the sensors. If it is

- *detection, cleaning* and *smoothing* of systematic errors and noise;
- *segmentation* and labelling according to body part;
- *restoration* of the surface in sparsely sampled areas;
- building a smooth *surface* structure on the data points.

For these processes to be effective, they must adapt to the particular properties of the data, which varies over the body, and with different image capture systems. This section investigates how these processes can be controlled by the known and measurable characteristics of the scan data, such as:

- error characteristics;
- point intensities;
- sampling density;
- expected size, shape and location of features.

4.1.1 Range Measurement and Errors

Coordinate values are measured in tenths of a millimetre, corresponding to the length that a bit represents in the original distance values (see section 2.3.2). The same units are used in all internal calculations. The mean error of each depth measurement within the working range of the system is taken to be ±1 mm [Hori98]. In modelling for surface reconstruction it is desirable to make use of as much prior knowledge as possible of these expected errors, in terms of magnitude, frequency, location and distribution. This section first discusses some known sources of error, and then how to make use of observed data characteristics and prior knowledge to detect and remove systematic errors and noise.

The investigations in Chapter 2 indicate that the error is likely to be higher in some isolated areas, in particular, on surfaces with very high specular reflectance and at dihedrals, such as on the inside of the arms and legs and around the chin. This prior information might usefully be taken into account when designing operators for cleaning, smoothing and reconstruction, for example, by applying a different smoothing technique in areas where higher levels of error are expected.

When the light beam hits the body surface its diameter is a few mm, so that where the angle of incidence is very large, the depth range, δ, over the beam width becomes significant (figure 4.1, case b). In these cases the expected depth precision would be less. In addition, further from the LED source the beam diverges slightly, so that the point size increases, with a corresponding decrease in precision. This, in part, determines the valid working range of the sensors. If it is

assumed that the surfaces lie within this working range, the data may be assumed to conform to the stated error. Outside of this range the error is undefined.

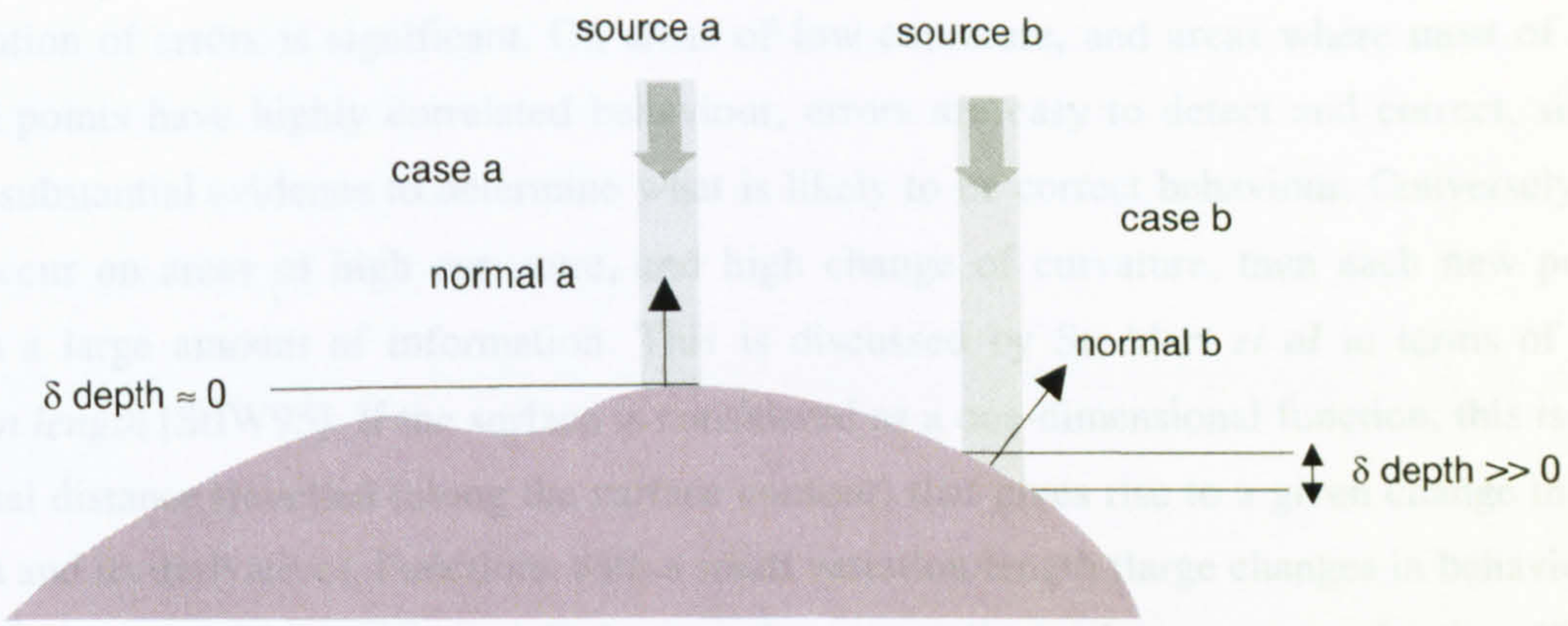


Figure 4.1: Angle of incidence and depth uncertainty.

Noise can be considered separately from the known sources of systematic error. In the absence of any other knowledge, the errors due to noise are considered to be normally distributed, with zero mean.

The main problem when processing the data automatically is to discriminate between errors and subtle features. A way of formulating this might be to define a maximum local deviation that a sampled point may take with respect to the characteristics of neighbouring samples, without being considered to be noise. In information theoretic terms, this would be a measure of the information content of those points. Points with too high an information content would be treated as noise. This similarity value can be used as a threshold in logical smoothing operators to remove undesirable points, or in probabilistic terms, to determine values for a weighted sum filter.

The first question then is what metric should be used. Candidates include changes in curvature and Euclidean distance between points. For example, if the points in a horizontal cross section are considered, a simple choice might be to impose a low curvature bias on the reconstructed surface. In fact, it is likely that multiple metrics are needed, and that these will vary over different parts of the body. The next question is what size of neighbourhood should be used to determine what surface characteristics are considered "normal". If the neighbourhood is too large, then the effect will be to smooth out subtle surface features; if it is too small, then results will be more prone to contamination by noise [LeHD93].

If systematic errors occur, there may be two or more points with similar, but anomalous characteristics with respect to the overall behaviour of the sampled points. This situation is harder to detect, since there is no obvious local indication in the data that an error has occurred. In this case prior information, or the use of other, more global data characteristics, are required

to resolve the ambiguity. Such information could include the expected surface characteristics and an application-dependent decision on what type of feature it is appropriate to preserve. This is discussed further, in section 4.1.4, in the context of what types of feature can be modelled.

The location of errors is significant. On areas of low curvature, and areas where most of the sampled points have highly correlated behaviour, errors are easy to detect and correct, since there is substantial evidence to determine what is likely to be correct behaviour. Conversely, if errors occur on areas of high curvature, and high change of curvature, then each new point contains a large amount of information. This is discussed by Stoddart *et al* in terms of the *variation length* [StIW95]. If the surface is considered as a one-dimensional function, this is the horizontal distance travelled (along the surface contour) that gives rise to a given change in the function and its derivatives. Functions with a small variation length (large changes in behaviour) require higher sampling to be modelled correctly, especially in the presence of noise. If the sampling density is fixed, then regions with small variation length are modelled with lower confidence.

4.1.2 Intensity

As well as putting a threshold at the lower intensity limit, I_{lower}, as discussed in Chapter 2, intensity can be used to indicate the confidence of the depth value ($confidence \propto I$). In particular, when the angle of incidence is large, the intensity is expected to be low, and the precision is also expected to be low. Similarly, multiple reflections will often, but not always, produce low intensity values.

In logical processing operations, points of higher intensity should therefore be selected in preference to those of lower intensity. In weighted sum operators, a legitimate choice would be to assign a weight proportional to the intensity value. Processing in conjunction with other objective functions, intensity could be used to *moderate* the effect of an operator. For example, if outliers are to be removed based on spatial isolation, there is a danger of cleaning out a point that is most especially needed, precisely because it *is* the only point in a large neighbourhood. Since this problem is most likely to occur in areas where the intensity values tend to be low, it is possible to prevent spatial "outliers" being removed unless they are also below an upper intensity threshold. Unfortunately, it is not easy to choose a particular value of this upper intensity threshold, I_{upper}, or to determine it in an automated way.

4.1.3 Sampling Density

It was noted in Chapter 2 that sampling is sparse on surfaces whose tangent lies close to the direction of the incident light beam. This affects regions such as the tops of the shoulders, the insides of the arms and legs, and the underside of the chin and bust. In addition, because of the

divergence of the light beams from the LEDs on each sensor (as shown in figure 2.7a on page 43), sampling on the sides of the head and neck, and on small children is also slightly lower.

These effects will now be investigated in a quantitative way, in order to devise appropriate parameters for processing. Density along a particular surface contour is estimated here by first reconstructing the contour (this is discussed in detail later, in Chapter 5). The parameter u is defined as the distance along the contour, and the *separation*, λ, between connected points on the contour can be computed as $\lambda_i = |p_{i+1}-p_i|$.

Results are shown in table 4.1 for cross-sectional contours sampled at key sections on an adult female subject:

- head (at nose level),
- neck (narrowest part),
- shoulder (~2 cm below acromion),
- upper torso (~2 cm below axilla),
- lower torso (mid hip),
- right arm (at elbow),
- right upper leg (mid thigh),
- right lower leg (mid calf).

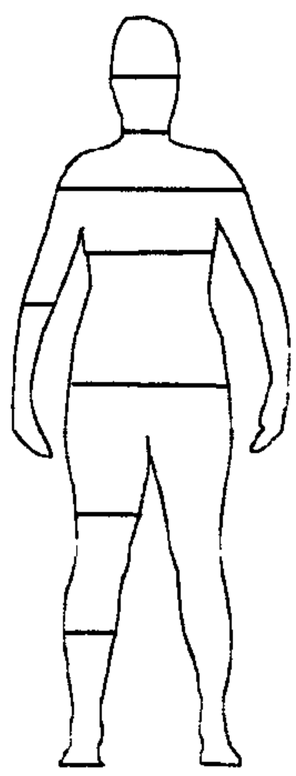


Table 4.1: Separation between points (λ) on sampled contours (10^{-1} mm)

	total pts	contour length	max λ	median λ	mean μ_λ	stdev σ_λ
head	70	5688	318	86	81	57
neck	49	3600	599	82	73	201
shoulders	174	9835	157	75	57	27
upper torso	133	9798	561	76	74	58
lower torso	154	9931	171	72	64	30
right arm	19	2357	541	79	124	113
right upper leg	66	5068	577	81	77	75
right lower leg	50	3550	427	80	71	68
			overall	**76**	**70**	**60**

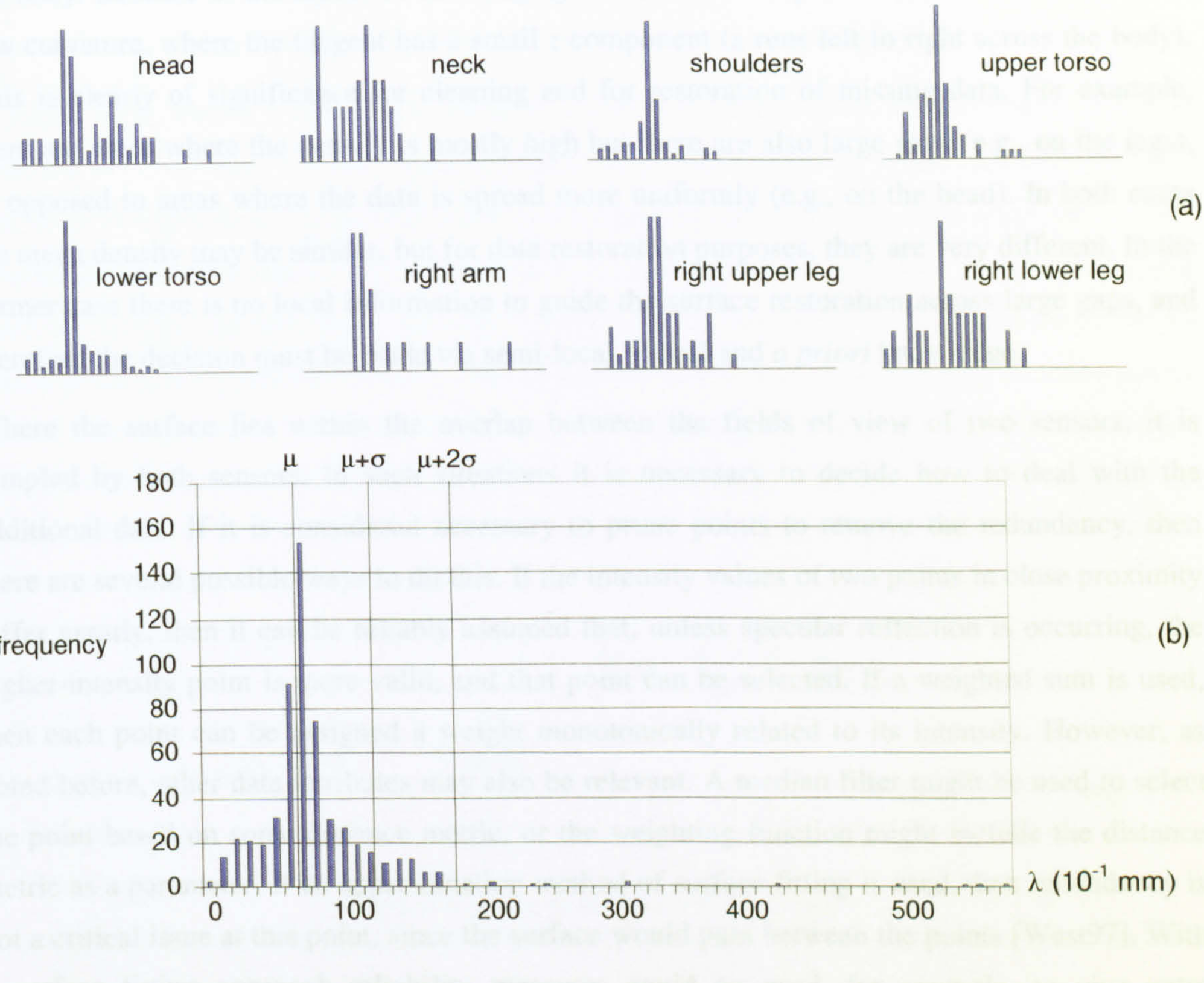


Figure 4.2: Frequency histogram for λ (a) over each sample, (b) combined.

If the samples from each body segment are considered separately, it is clear that the distributions of λ are very skewed and a normal distribution is not an appropriate model. However, when all body segment samples are considered together, by inspection a normal distribution (albeit with a long positive tail) appears to be a workable assumption. Therefore the overall mean, μ_λ, and standard deviation, σ_λ, will be used in various places where the density of the data is important in controlling operations, and can be used to obtain the probability, p(sample(d)), on sampling a point within a given distance d in u, where u is the surface contour length:

if $d = \mu_\lambda$ p(sample(d)) = 0.5

if $d = \mu_\lambda + \sigma_\lambda$ p(sample(d)) = 0.842

if $d = \mu_\lambda + 2\sigma_\lambda$ p(sample(d)) = 0.977

Note that in this work a single value for d is used over the whole body surface. A more sophisticated system would make use of local information on the different observed sampling densities over the body.

If λ is considered as a function of *u* it is possible to observe the distribution of densities around the body. Because of the layout of the imaging units, low-density areas often occur on areas of low curvature, where the tangent has a small *z* component (*z* runs left to right across the body). This is clearly of significance for cleaning and for restoration of missing data. For example, there are areas where the density is mostly high but there are also large gaps (e.g., on the legs), as opposed to areas where the data is spread more uniformly (e.g., on the head). In both cases the mean density may be similar, but for data restoration purposes, they are very different. In the former case there is no local information to guide the surface restoration across large gaps, and therefore the decision must be made via semi-local, global and *a priori* knowledge.

Where the surface lies within the overlap between the fields of view of two sensors, it is sampled by both sensors. In such situations it is necessary to decide how to deal with the additional data. If it is considered necessary to prune points to remove the redundancy, then there are several possible ways to do this. If the intensity values of two points in close proximity differ greatly, then it can be reliably assumed that, unless specular reflection is occurring, the higher-intensity point is more valid, and that point can be selected. If a weighted sum is used, then each point can be assigned a weight monotonically related to its intensity. However, as noted before, other data attributes may also be relevant. A median filter might be used to select the point based on some distance metric, or the weighting function might include the distance metric as a parameter. If an approximation method of surface fitting is used, then redundancy is not a critical issue at this point, since the surface would pass between the points [West97]. With a surface fitting approach reliability measures could be used, for example, to give extra weighting to higher intensity points.

4.1.4 Modellable Features

Prior knowledge of human shape, and expected size, shape and distribution of features can be used to determine the limits of what can be usefully modelled, in terms of:

- individual entities, for example, the possibility of modelling separate fingers;
- local features *on* the surface, for example, the sharpness of cusps, detail of buttons on clothing, etc.

The limits of what can be modelled are a function of the measurement precision, data density, expected shape characteristics and the particular requirements of the application.

Thus the size of the individual entity that can be reconstructed, in terms of its *height* and *width*, is determined largely by the density of the data. It is also affected by the measurement precision, but since this is considerably smaller than the sampling density, it does not significantly affect this decision. According to the Nyquist sampling theorem [Vase96], the smallest feature that can be recovered is twice the sampling frequency. In other words, to be 98% confident of

reconstructing a feature, it must be at least $\mu_\lambda+2\sigma_\lambda$ in each dimension. For the subjects used in this work, this would be approximately 2 cm. Thus, for example, it is not appropriate to reconstruct each finger individually, but to treat the hand as a single entity.

In terms of a feature's *depth*, the smallest value is determined by the precision, which is assumed to be ±1 mm [Hori98]. Thus it should be possible to preserve features that deviate 3 mm or so in depth from the surface, such as the edges of clothing against the body, but not any smaller features (see figure 4.3b). The *largest* local feature that should be preserved is based more on the *expected* attributes, and what is necessary for the applications. If a series of sample points gives a fairly smooth surface, and the next sample is significantly far from those points, should it be assumed to be a valid feature, or outlier (see figure 4.3c and 4.3d)? For this work it is not necessary to model very fine features, or deep into cusps, and therefore it is acceptable to smooth them.

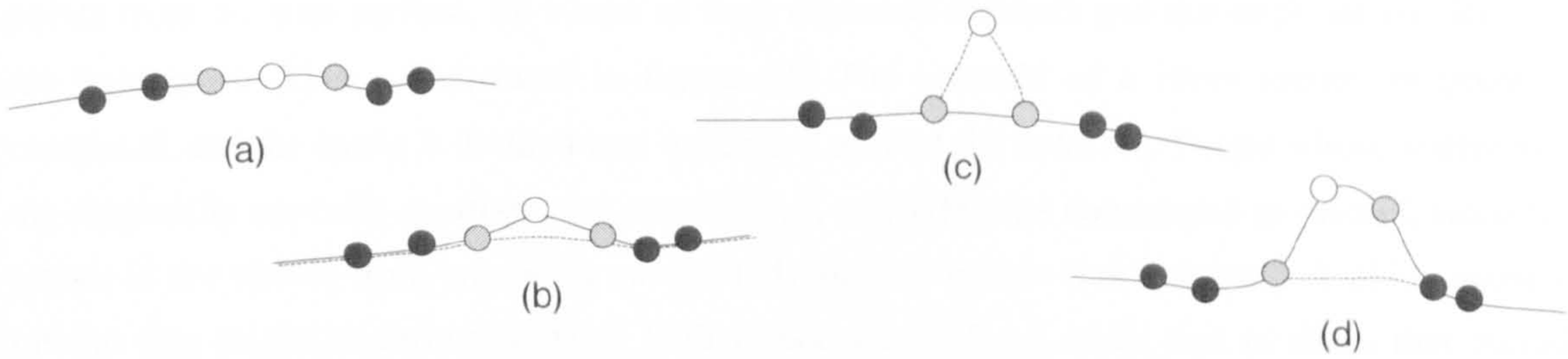


Figure 4.3: Modelling small features.
(a) featureless surface, (b) possibly a feature, (c) probably an outlier, (d) probably a feature.

It is expected that data reconstruction may be difficult on sections of very small circumference, such as the wrists, where the rate of curvature is higher, and the overall data density is smaller. Again, this relates to the variation length, discussed by Stoddart *et al* [StIW95]. It might be possible to use information on curvature to control the properties of a smoothing operator. For example, an anisotropic operator could be devised that biases the smoothing to lower values in directions of high curvature. This would preserve detail in directions of small variation length, but smooth areas where the variation length is large.

4.1.5 Surface Normal Prediction

By knowing the approximate stance of subjects, as discussed in Chapter 2, and the geometry of the sensors, it is possible to use some of this information to place bounds on the estimation of the surface normals. For this discussion the data is considered slice by slice.

For a point to be viewable by the sensor, the dot product of the surface normal and the LED vector must be greater than zero, as illustrated in figure 4.4. At first sight this might appear to be a useful aid for segmentation. Unfortunately the visibility constraint does not provide enough information to determine, for example, if two points, adjacent in sequence, belong to the left and

right leg respectively, or to the same leg, since the two sampled points may legitimately correspond to surfaces whose normals are anything less than π apart.

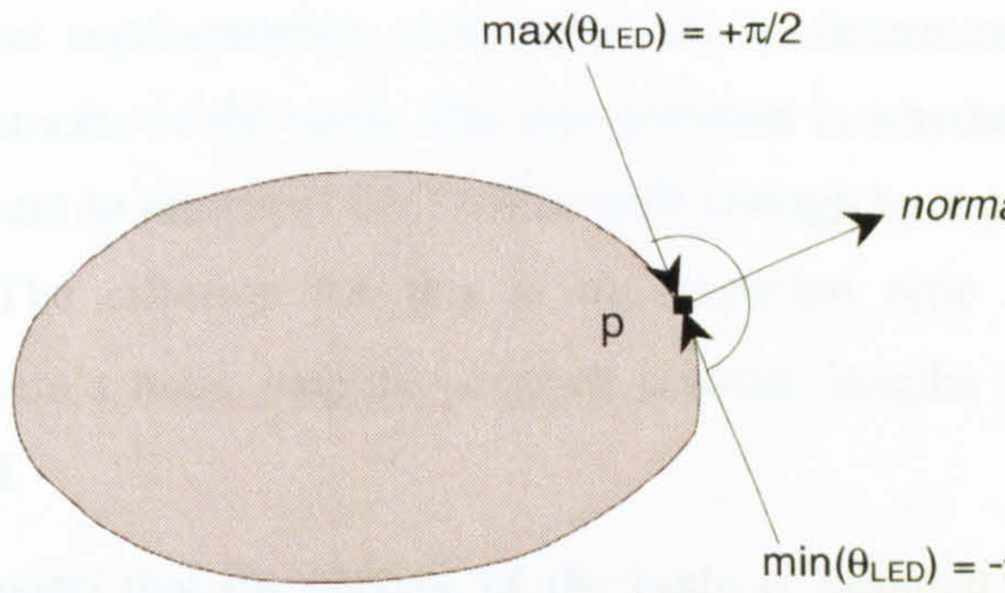


Figure 4.4: Limits of LED vector and surface normal.

Assumptions about body geometry can be used to discount (with a high probability) certain points from the true surface, by virtue of their expected normals and the angle of incidence of the light beam. This is illustrated in figure 4.5. The centroid of a cross section of points is computed, and the space is divided into quadrants around the centroid. Points whose source is in the diagonally opposite quadrant, for example, p_1, can safely be considered as invalid, since that source is not visible from any point on the body surface within that quadrant. A likely cause of invalid data might be reflection from highly specular surfaces. Note that in slices that include the arms, these assumptions are less valid, because the arms' many degrees of freedom mean that the body centre cannot be well approximated via the bounding box of all the points. After generating an hypothesised surface, decisions about point validity can be based on the normal of the proposed surface and the angle of the light source.

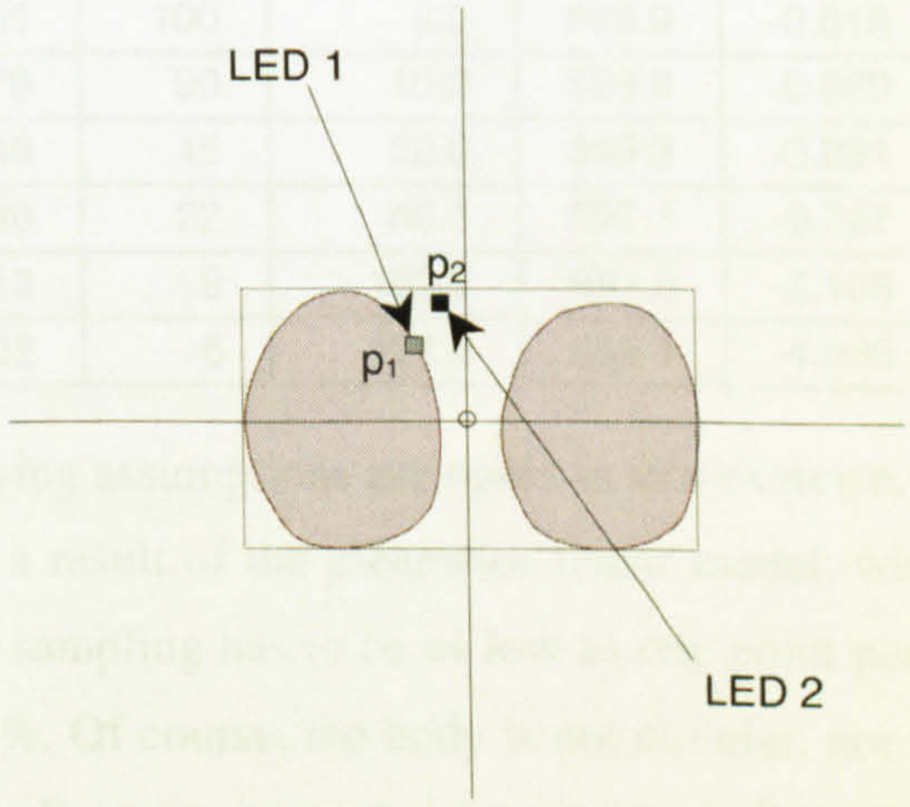


Figure 4.5: LED angle, expected surface normals, and invalid points.

4.2 Piecewise Linear Representations

In the preceding discussion, the surface representation had not yet been determined. The main requirements are for a complete and faithful representation that can be generated in a

computationally efficient way, with low memory requirements. In many ways, the simplest choice would be a polygonal mesh, which also has the benefit of being fast to render on a machine of modest computational power. Before proceeding to develop techniques to generate such a piecewise linear representation, it is necessary to determine if it is suitable for the anthropometric requirements of the work. The first question is whether the density of the initial sampled data is sufficient to represent the surface well enough by a polygonal mesh with nodes of uniform density. The criterion for this is the expected type and magnitude of errors introduced in the system's main outputs—convex contour lengths and volumes—by using a piecewise linear model.

If it is assumed (for now) that the surface of the body is sampled with uniform density and perfect depth measurement, then it is possible to determine what kind of errors are incurred by the use of a polygonal mesh. Table 4.2 below shows the errors on estimates of piecewise linear circumference (PLC) and area (PLA) on an example cross section of the body. The cross section is assumed here to be circular, which, though of course not true in reality, is a useful assumption for this purpose. The data separation values in the leftmost column correspond to the Euclidean distance between samples, in other words, the chord length between points along the surface contour. The circumference is chosen to be 900 mm, with an area of 644.58 cm^2.

Table 4.2: Errors for piecewise linear circumferences (PLC) and areas (PLA)

chord length (mm)	angle (radians)	no. points	arc length (mm)	PLC (mm)	PLC %error	PLA (mm^2)	PLA %error
6	0.042	150	6.0	899.9	-0.007	64439	-0.029
7	0.049	129	7.0	899.9	-0.010	64432	-0.040
8	0.056	112	8.0	899.9	-0.013	64424	-0.052
9	0.063	100	9.0	899.9	-0.016	64415	-0.066
10	0.070	90	10.0	899.8	-0.020	64405	-0.081
20	0.140	45	20.0	899.3	-0.081	64248	-0.325
40	0.280	22	40.1	897.1	-0.327	63618	-1.303
100	0.713	9	102.2	881.0	-2.106	59131	-8.263
150	1.102	6	157.9	855.1	-4.985	52178	-19.051

Although several simplifying assumptions are made in this exercise, it serves to show how small an error is introduced as a result of the piecewise linear model, with perfect data on a circular cross section. In fact, the sampling has to be as low as one point per 50 mm for the error on the contour length to reach 1%. Of course, the body is not circular, nor fully convex, but if no noise is assumed, the piecewise linear estimate for the convex hull will always be an underestimate, whether the shape is convex or not. An opposing effect comes from noise, which will tend to cause overestimates on convex hulls. Isolated *positive* errors on depth measurements will result in an overestimate on the convex hull, but isolated *negative* errors will not affect it at all; only *systematic* negative errors will affect the convex hull. On area and volume estimates, concavities and convexities tend to cancel each other out in a piecewise linear model.

With the observed sampling densities, the discussion above suggests that a piecewise linear model is a valid representation for body contour measurement. Therefore the goal of the next sections will be to reconstruct the body surface as a polyhedral mesh. If the mesh has a uniform connectivity then this simplifies future operations. The *c*-connected, oriented, piecewise linear surface is represented by, $R = \{r_1,.....,r_m\} \subset \Re^3$, such that for all $i = 1...m$, and for all $j = 1...c$, where $|r_i - r_j| \leq d$, where d is the maximum separation permitted between connected points.

The surface need not be fully closed. For example, the under surfaces of the feet may be ignored, and there is no particular requirement that the extremities, such as the fingertips, or the very top of the head, are closed, as long as a minimum distance condition holds. This relaxation of the closed surface condition should not affect volume or surface anthropometrics, although in the future it may be necessary to restore the closure condition, for example, in determining surface area.

4.3 Generic to Humanoid Surface Reconstruction

A theme that runs through this work is the use, or not, of prior knowledge. This is particularly interesting to observe in surface reconstruction. A general surface reconstruction method that makes minimal assumptions [HDDM92] [CuLe96] [AmBK98] [BBCS99] can be applied to the semi-cleaned data. As discussed in Chapter 3, such a method would be expected to give good, but not adequate, results, with particular problems at the occluded branch points of the body. By adding more assumptions, for example, segmenting the data, it would be expected that increasingly better results might be obtained. In addition, the use of assumptions and constraints makes it possible to simplify processes and reduce the computational complexity. This progression will be explored through the following steps:

- generalised surface reconstruction methods;
- better cleaning and simple triangulation;
- segmentation and simple triangulation;
- closed compound deformed cylinder meshes.

more assumptions,
more constraints

At each stage the results will be assessed visually, and finally a volumetric analysis will be carried out to validate the enclosed volumes against reference methods.

The drawback of adding assumptions to the processing is that it puts limitations on the applicability, for example, that the subject must not be wearing loose clothing, that they must be standing in a certain posture, etc. By developing more sophisticated processing operators in the future, some of these conditions might be relaxed, thereby extending the applicability. For example, a broader range of postures and body shapes might be processed successfully.

4.4 Generalised Reconstruction

Experiments are first presented using Hoppe's surface reconstruction technique [HDDM92], to determine how well such a generalised technique can deal with the images. Data sets were partially cleaned by intensity thresholding, as described in Chapter 2, then normalised. In Hoppe's technique tangent planes at each data point are estimated as the least squares best planar fit over k nearest neighbours; the sign of the normal is chosen to give consistent orientation to neighbouring planes. This NP-complete problem [GaJo79] is approximated by the heuristic of propagating the normal alignment first over areas of low curvature, and then to higher curvature, which tends to give successful results. The triangular mesh surface is then generated by a version of the marching cubes algorithm [LoCl87]. Outlier noise is handled using assumptions of a given minimum separation between connected points, so that a point greater than d distant from other points is not considered part of the surface. The method assumes smoothness (i.e., the surface does not contain spikes or sharp cusps) and connectivity where possible (i.e., it does not contain holes smaller than d in diameter), within this constraint. The distance parameter can be varied to deal appropriately with data of various sampling densities and noise characteristics.

The results of Hoppe's method, using scan input after intensity thresholding, are illustrated in figure 4.6a. The strips of points above the body and in front of the left foot correspond to points sampled on the physical structure of the scanner itself. Since the geometry of this is known, the points can be removed by spatial thresholding. High-intensity points appear in the raw data above the head, but since these are relatively few in number, the top of head can be correctly detected if a threshold is imposed on the number of data points in the cross section. By starting at the first, topmost data points captured, the first slice of head data can be detected as the one in which the total number of points, after initial cleaning, is above the threshold. All points below this are assumed to belong to the body. The results of these processes are shown in figure 4.6b.

On most of the reconstructed images an artefact appears, extending from the inner surface of the left leg downwards, as illustrated in figure 4.6a and 4.6b. At first sight it is difficult to see how such a structure could be removed without losing valid data, since the artefact actually joins the leg. On closer inspection, it is apparent that the structure corresponds to data points above the lower intensity threshold, which recur at a similar location at the front of the leg. By tracing them back to their source, they were found to be the result of LEDs diagonally opposite, and therefore can be assumed to be reflections, as discussed in section 4.1.5, and illustrated in figure 4.5. It appears that they are caused by light passing between the legs and reflecting off the inside of the scanner itself. Such points can be detected and removed by a simple test whose time complexity is $O(n)$. This method is valid as long as the subject is standing with both feet at approximately the same location in x. Results of this reflection cleaning step are shown in figure 4.6c.

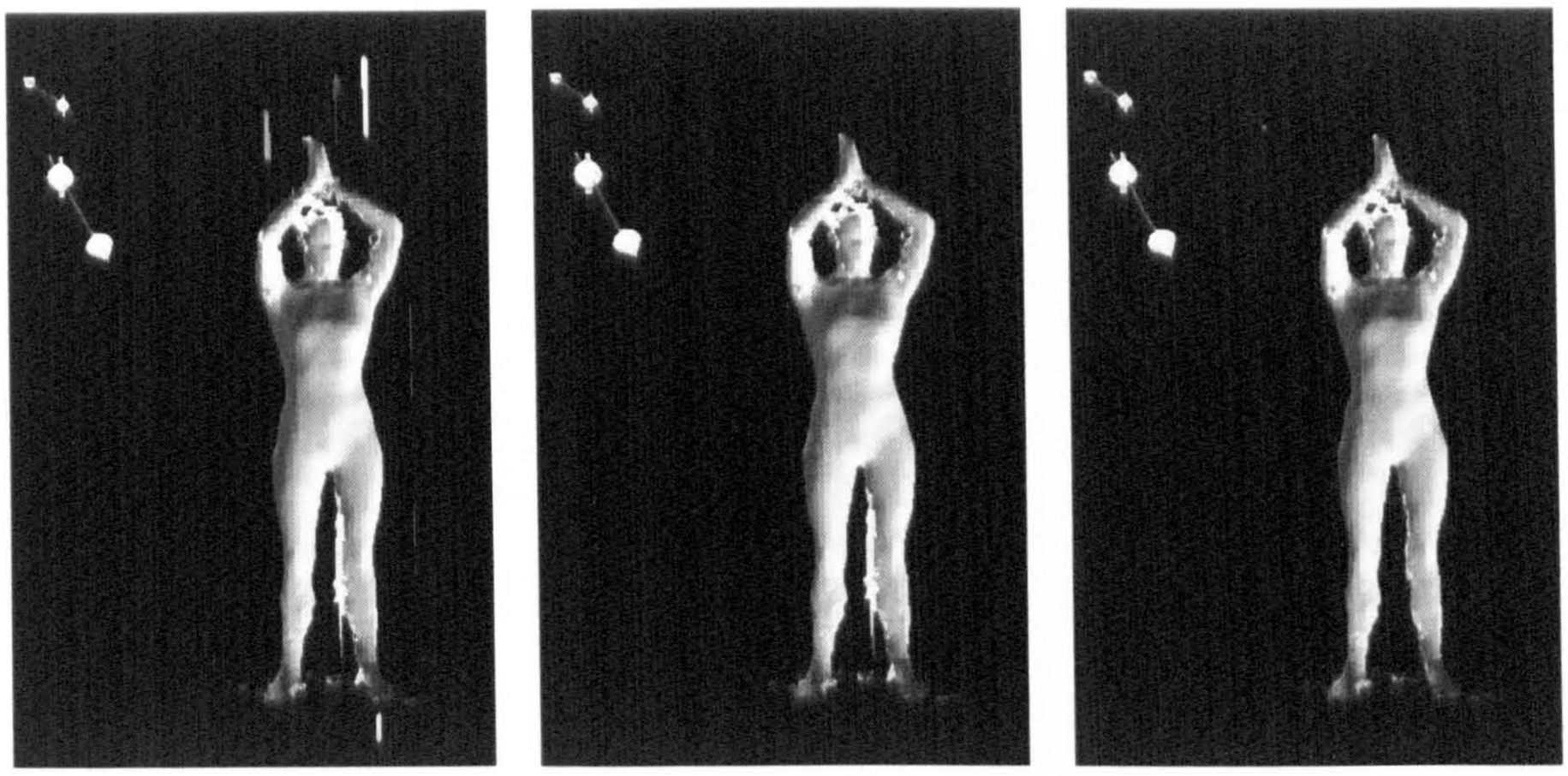


Figure 4.6: Hoppe's reconstruction, (a) with intensity thresholding, (b) with spatial thresholding, (c) with reflection detection.

Hoppe's technique does not make any assumptions about surface connectivity, or about the number or location of surface boundaries. It therefore quite justifiably leaves gaps where data is sparse, and connects body segments where points are in close proximity. Conversely, it has the advantage that it can be used on any posture, as shown in figure 4.6, since it does not have prior assumptions. Although a deformable template method might avoid the problem of holes, via a closure constraint, it is not expected that other generalised surface reconstruction techniques would perform significantly better in terms of dealing with the problem of nearby points originating from different body parts.

As discussed in section 4.1, there is a great deal of prior and measurable information that can be applied by developing an improved surface reconstruction technique specifically for human-body range data. In the spirit of the minimal complexity criterion, it is useful to start with a simple scheme. The known structure in the data can be exploited by reconstructing data from each sensor head separately. This is illustrated in figure 4.7. The line of points in each sampled slice, s, are treated as being essentially parallel with adjacent slices, and the known sampling sequence for each head is used. Thus, adjacent points, $p_{[s][i]}$ and $p_{[s][i+1]}$ are connected if they meet the distance criterion, $|p_{[s][i]}-p_{[s][i+1]}| \leq d$, where d is the distance parameter corresponding to $\mu_\lambda+2\sigma_\lambda$, discussed in section 4.1.3. The third vertex of each triangulated face is connected at $p_{[s+1][i]}$ if $|p_{[s+1][i]}-p_{[s][i]}| \leq d$,. This automatically avoids the problem of large spatial outliers, in a similar way to Hoppe's technique, but it means that a few holes will occur. The number of such holes depends on the probability of obtaining a suitably connected triple of points, which in turn relates to the sampling density and the chosen d value. The orientation of the surface normal for each polygon is much more straightforward to determine, since the direction of the light source is known, the local visibility constraint, $\theta_{LED}-\pi/2 < \theta_{norm} < \theta_{LED}+\pi/2$, discussed earlier, holds. This avoids Hoppe's NP-complete surface orientation problem, and the overall time complexity is $O(n)$.

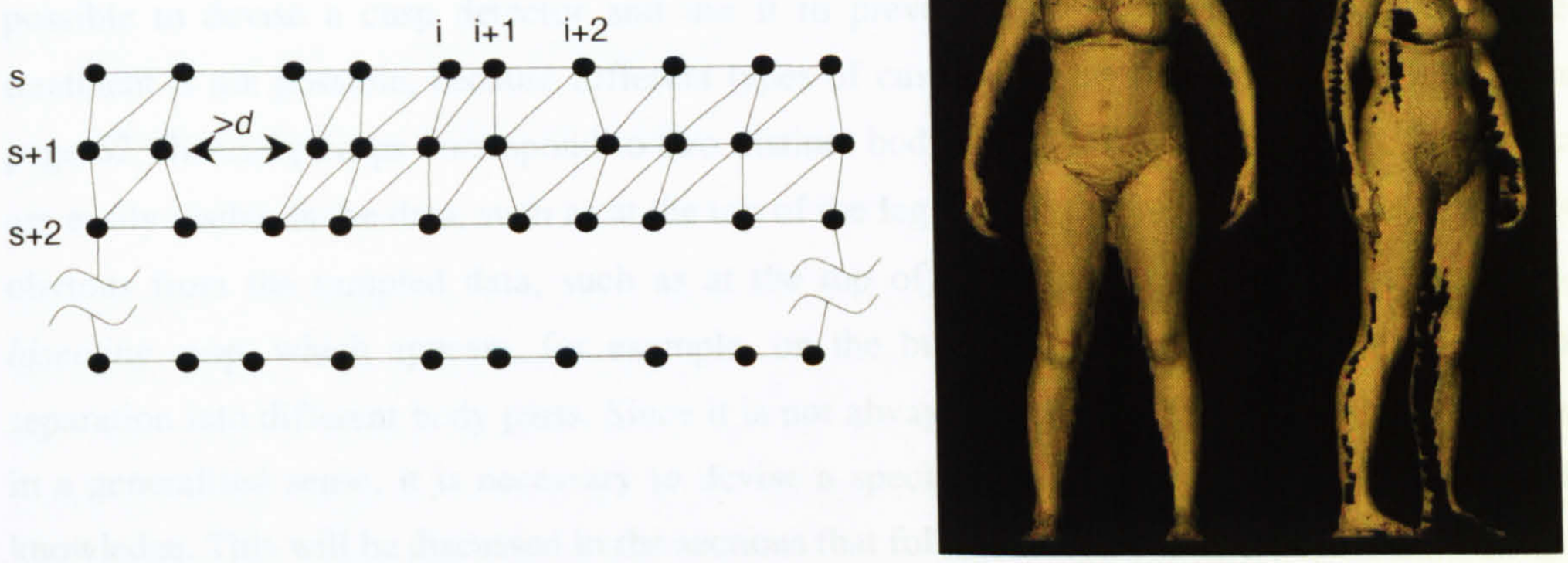


Figure 4.7: Surface reconstruction by polygon growing.

This starting point makes it possible to determine what is essential in the surface reconstruction process. At this point the problem of redundant data is side-stepped, since the data from each sensor head is reconstructed separately, and each generates a fairly smooth surface patch, which overlaps at the edges with the patches from the adjacent sensor heads. Although this is not a problem visually, it clearly must be addressed if any useful body metrics are to be extracted.

This simple reconstruction process exploits the information implicit in the data capture sequence, so that points are connected in an anti-clockwise manner, unless they fail the distance condition. If data from all sensors is treated as a union set, then the assumption of a sequence cannot be fully exploited, since the whole data set makes up a staggered, overlapping spiral. The simple triangulation method on the union set tends to produce small ripples at these areas of overlap. If such a data interpolation method (as opposed to surface fitting by approximation) is to be pursued, then the data redundancy problem must be addressed. The pruning technique shown in figure 4.8 uses a mean filter on a radial neighbourhood, which ensures data reduction to any given density, depending on the filter radius r. Data reduction takes $O(n^2)$ time, where n is the number of points in the slice. Alternative methods would be to select the point with highest intensity, or weight the mean filter according to intensity. However, it will be shown later that this cleaning step is not needed, as it can be carried out as part of another operation.

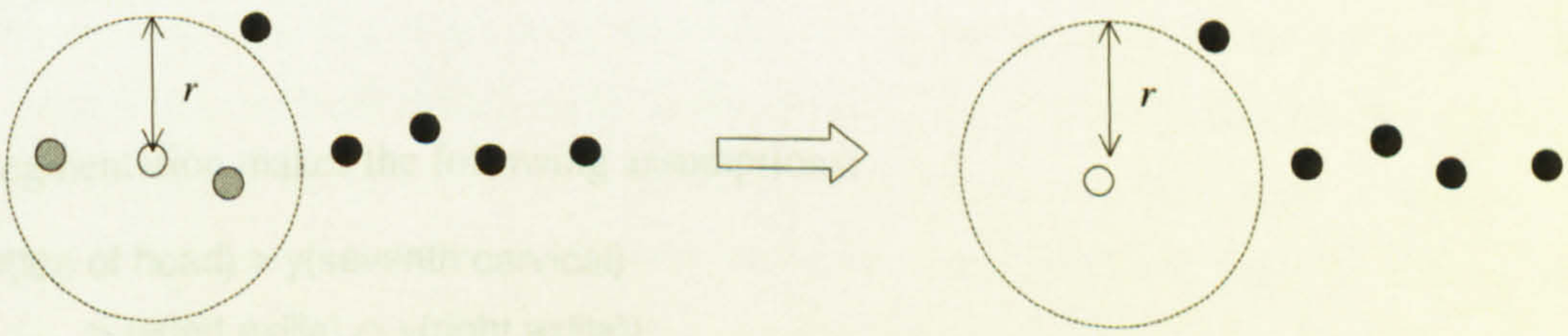


Figure 4.8: Data reduction by a mean filter in a *d* radial neighbourhood.

This simple reconstruction method is clearly limited because of the proximity of body segments. If the connectivity distance, d, is increased, fewer gaps occur in the surface, but inappropriate connection between body segments increases. Since these always occur at cusps, it might be possible to devise a cusp detector and use it to prevent this. Unfortunately such a simple treatment is not possible, because different types of cusp occur, as illustrated in figure 3.4 on page 62. *Bisecting* cusps correspond to two distinct body segments. Some features of this type are easily visible in the data, such as at the top of the legs; some are very subtle, and not always obvious from the sampled data, such as at the top of the arms. The second type is a *non-bisecting* cusp, which appears, for example, on the buttocks, and does not correspond to a separation into different body parts. Since it is not always possible to distinguish these features in a generalised sense, it is necessary to devise a specific method that incorporates this prior knowledge. This will be discussed in the sections that follow.

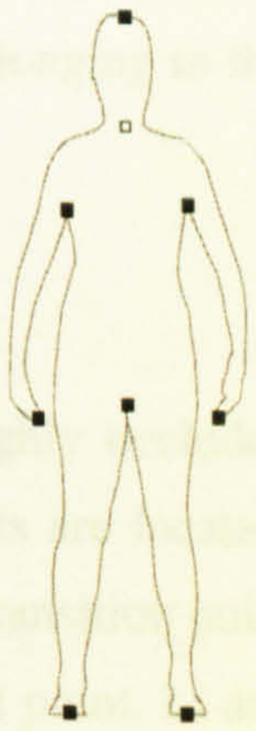
4.5 Segmentation

Some of the problems of *generalised* surface reconstruction techniques may be avoided if the process incorporates specific knowledge of the human body. This will be incorporated into segmentation procedures that detect the key branching points and then assign an appropriate label to each sampled point, indicating to which part of the body it belongs.

4.5.1 Primary Landmarks and Segment Choice

The key body segments have been chosen here to break the body into simple components that can be skinned more easily. The primary landmarks mark the boundary between key body segments:

- top of head,
- seventh cervical vertebra (the top of the torso proper),
- branch points of the left and right arms (which may be at different levels),
- torso/leg branch point,
- ends of the hands (which may be at different levels),
- bottoms of the feet.



This segmentation makes the following assumptions:

y(top of head) > y(seventh cervical)
> (y(left axilla) ∩ y(right axilla))
> (y(end of left hand) ∩ y(end of right hand) ∩ y(leg branch point))
> (y(bottom of left leg) ∩ y(bottom of right leg))

y(left underarm point) not necessarily = y(right axilla)

z(left underarm point) > z(leg branch point) > z(right axilla)

y(end of left hand) not necessarily = y(end of right hand)

y(top of left leg) = y(top of right leg) = y(leg branch point)

y(bottom of left leg) = y(bottom of right leg) = 0

z(bottom of left leg) > z(leg branch point) > z(bottom of right leg)

At this point it is necessary to make a decision about how the boundary between the segments should be defined. For greatest anatomical meaning, the partitions should correspond to the way that the tissue is associated with each part of the skeleton. To do this correctly requires considerable investigation beyond the scope of this work, so a more simple approach is taken here, by making a horizontal partition at each branch point. This is useful in optimising the surface reconstruction and shape analysis operations, as will be discussed in further depth in section 4.6.2 and in Chapter 5.

The top of the head is trivial to locate if it is assumed that the arms are held downwards from the shoulders, and nothing extends above the head. The top of head landmark is placed at the nearest sample point to the centroid of the top slice of data.

The seventh cervical vertebra marks the separation between the torso and the neck, which would normally be located by feel. Nurre located this as the smallest convex hull on the body [Nurr97], although he noted that this was not a robust method on very muscular subjects. In this work it is found by detecting features in the convex hull space. If horizontal convex hulls are generated at close intervals down the body, the nape corresponds to a discontinuity in the first derivative of the convex hull length with respect to the height of the slice. This will be discussed in more detail in Chapter 5, in section 5.4.7. Points above this can then be assigned as belonging to the head. The remaining landmarks relate to the main branching points on the body.

4.5.2 Branch-Point Detection

The main characteristic of the branching points is that they lie at cusps, in highly occluded regions. There is therefore very little data at the point at which the branch points are located. Partial cusps continue above the axilla, so the method must be able to detect the transition point between the non-bisecting cusp above the axilla and the bisecting cusp below that point. Li and Jones used a step detection method to segment the arms [LiJo97]. The results appear to be effective, although no quantitative validation is reported. At first sight a simple frontal projection and gap detection method looks attractive. Such a method must be able to detect a true gap on the body, as opposed to the gaps between the finite sampling of data on the continuous surface. The minimum detectable gap must therefore be significantly greater than the expected sampling density. Unless the arms are held far away from the body (which distorts the

"natural" shape of the shoulders) the gap will tend to be detected too low down. This problem is worsened, of course, on data with a lower sampling density.

In addition, detecting the branch point of the legs with such a method will usually fail because the legs often touch unless the subject is of very slight build. This can be avoided if the feet are placed widely apart, but this distorts the hip shape. Simple turning point methods fail for the same reasons. West [West97] used a spline fitting method for curve reconstruction and detected the branch point where the splines from the left and right legs start to overlap, working upwards from a point of guaranteed separation. Again, this is only suitable where the subject's legs do not touch. Nurre's method [Nurr97] was similar, in that the branch points were designated as the point at which the convex hulls of the separate body segments merged. More recently Certain and Stuetzle have used a method that detects the branch point of the legs, based on their representation of the lower body as a paired stack of ellipses fitted to the range data [CeSt99]. If the distance between the centres of the ellipses of the right and left side is considered as a function of height, the branch point appears at a discontinuity in the function.

Exploration of inter-point distance statistics reveals some interesting characteristics relating to cusps and transitions at branch points. Within a slice of data points, the mean Euclidean distance of each point to all other points can be computed. At cusp points this value is low, and the value increases for points on the outer side of the body, away from the cusp, as shown in figure 4.9. If the mean inter-point distance values for all data points in a given slice are described as a function of the distance along the body's vertical axis, the transitions at the branch points appear as points of inflection. Tests have been carried out to detect this automatically, but it has not been possible to do it reproducibly, so this method has not been used for transition detection. However, it is used for detecting the cusp bisector, as described later in this section.

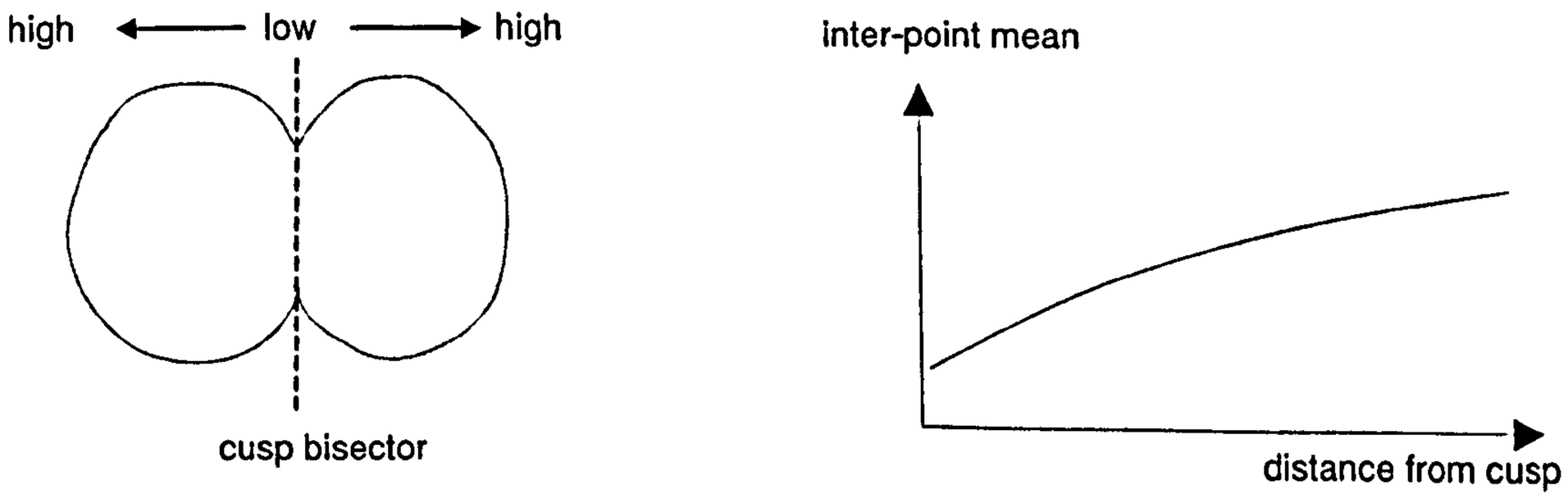


Figure 4.9: Inter-point means on a horizontal slice through the legs.
Low values occur at cusp points.

Re-entrancy has also been investigated by transforming the data to a cylindrical coordinate system about the centroid of each slice of data. Re-entrancy can be detected where two or more points fall close to the same radial line, but with significantly differing distances from the centroid. This is implemented by binning the data into sectors of fixed size $\delta\theta$, and detecting the

re-entrant sectors (figure 4.10). A threshold on the level of re-entrancy is used to detect the branching point transition at the underarm level. This method allows each arm branch point to be detected independently, without the assumption of symmetry.

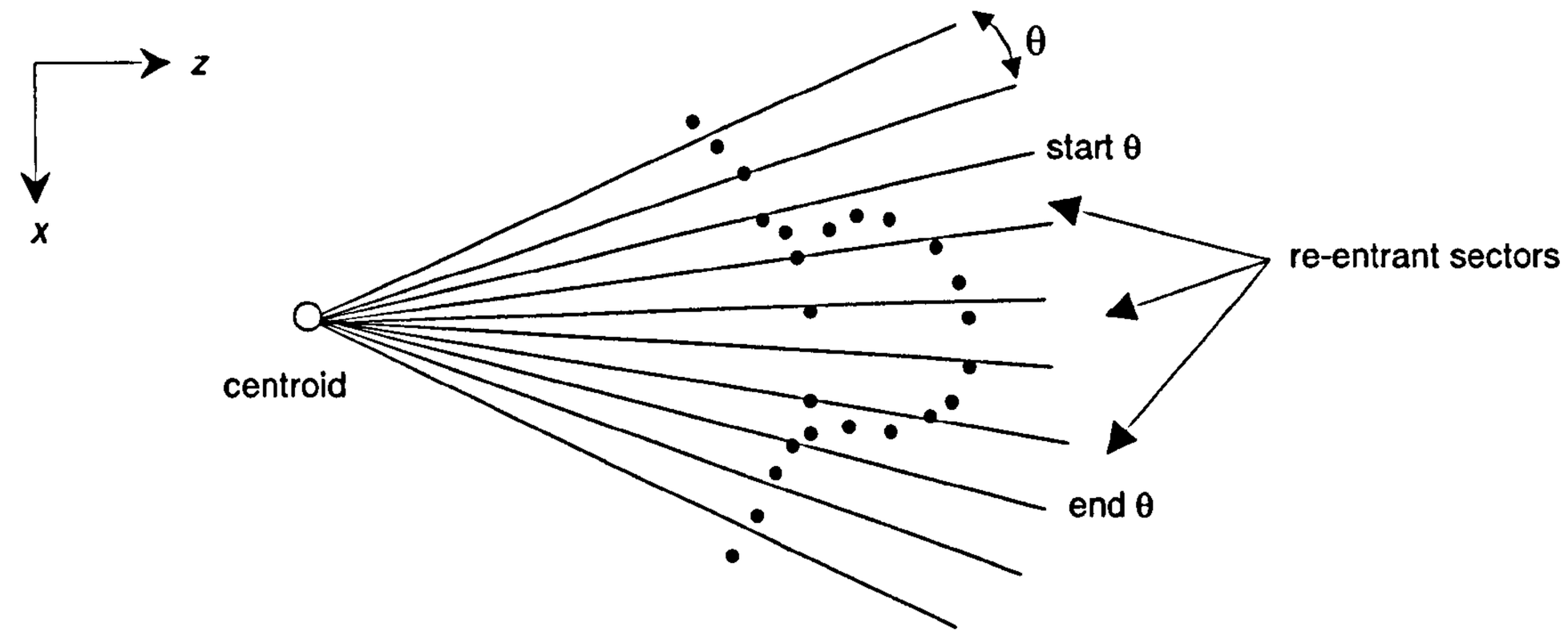


Figure 4.10: Re-entrancy detection at the axilla level.

Once the correct level (y) of the branch points has been detected the precise location in the xz plane must be found. This is considered to be the centre (with respect to x, the depth) of the cusp at that level, as illustrated in figure 4.10. The centre axilla is then designated as the inner point at approximately $(\theta_{end}-\theta_{start})/2$.

It is not possible to use this method directly to detect the branch point on the legs because the same type of radial re-entrancy does not occur. A more straightforward approach is to make use of the changing depth of the torso through the centre line of the body (central with respect to body width, along the z axis), and to detect the bottom-most point on the torso. In order to find this point, a constrained volume of data is sampled through the approximate centre of the torso along the z axis. It is assumed that the subject may be standing off-centre when scanned, and the arms might be held in any position, and so the true centre of the torso may be significantly different from the centroid of the data in each slice. A better approximation is found by finding the centroid of the head, and using the same z value, cz, as the approximate centre line of the torso. When the data through $cz \pm (\mu_\lambda+2\sigma_\lambda)$ is projected onto the xy plane, the changes in depth can be extracted from the contour. This width, $\mu_\lambda+2\sigma_\lambda$, is sufficient to sample enough points to reconstruct the full contour, and to allow for the fact that the approximated central line is not necessarily the true centre of the torso.

Consider the depth of this contour, from front to back, starting at the top of the torso, and moving downwards along the body's main axis. On subjects where the legs do not touch each other (figure 4.11 (b), (c) and (e)), the lowest point on the torso can be detected where the depth is zero. On subjects where the legs touch below this level (figure 4.11 (a) and (d)), the zero-depth point occurs below the target point. The zero-depth stopping criterion must therefore be modified to deal with this wider range of body shape. Since the larger values correspond to subjects of heavier build (whether because of fat or muscle bulk), body mass index, BMI =

weight/height^2, was investigated as an indicator. The correct level of the branch point, and the width of the contour at that point, were extracted manually from scans of subjects with a range of BMI values. A linear relationship, *stopping depth* = $k \times$ *BMI*, was tested, and a piecewise linear regression fit was found on three body mass ranges—light, medium and heavy—as shown in figure 4.12.

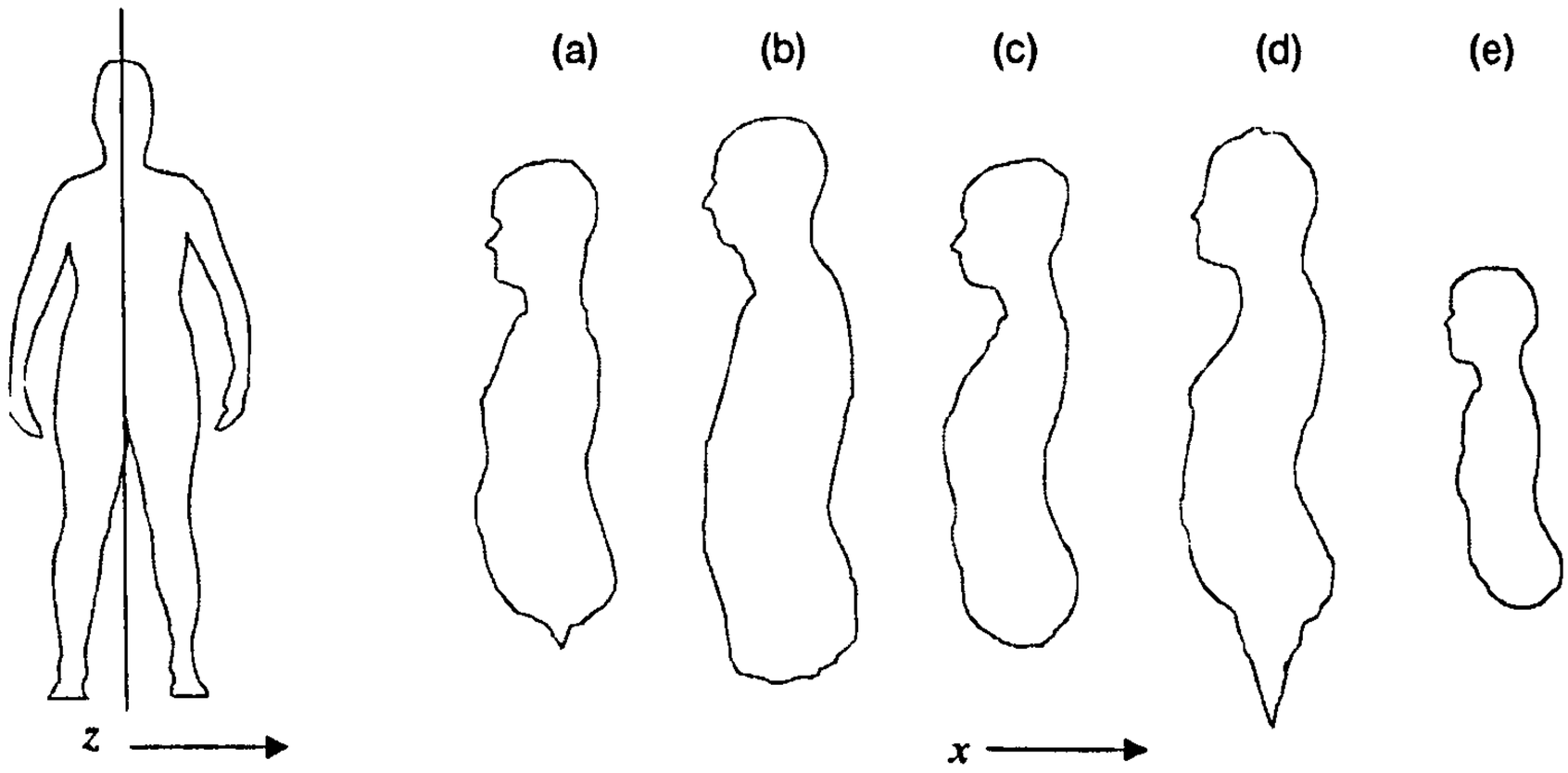


Figure 4.11: Contours through the body's centre *z* plane.

To calculate BMI, the subject's stature can be determined in an automated way from the top of head landmark, but the weight must be supplied (which contravenes the minimum intervention criterion), since it cannot be detected from the image. An alternative estimate of body type was investigated by sampling contour lengths on the torso. The curve lengths of a sample set of convex hulls around the mid-torso, together with the height value, can be used to approximate volume: *volume* ≈ *sample contour length* × *height*. This can then be used as an approximation of weight in the BMI calculations.

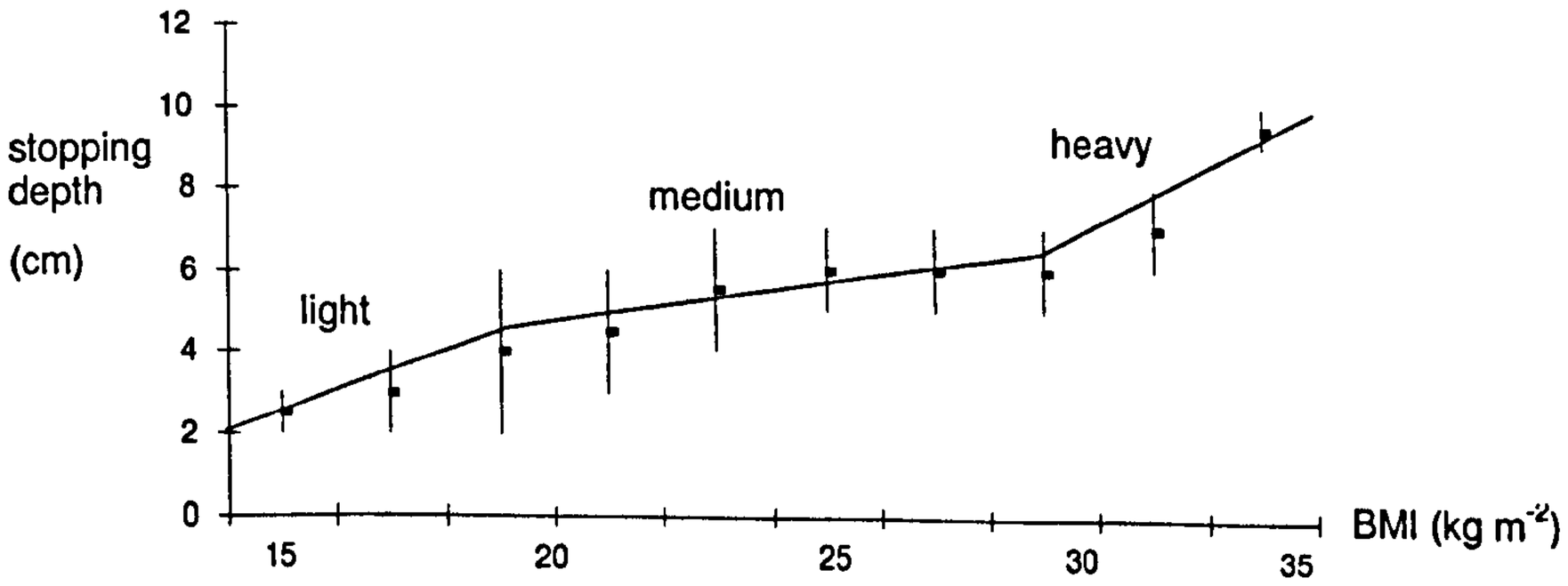


Figure 4.12: Piecewise linear fit of stopping parameter versus BMI.

The precise landmark position is defined as the centre of the cusp at the detected level. This is found by computing the mean of the n smallest inter-point distance values, discussed earlier in

this section. As the distance between each point and every other point must be calculated, this takes $O(n^2)$ time, where n is the number of points in the relevant slice.

4.5.3 Data Partitioning

Once the branch points have been found, data must be partitioned so that each point can be labelled. In order to detect the gap between the arm and the torso, the data are projected onto the yz plane. If a gap greater than a given threshold is detected in the data, it is considered to be a gap between body segments. It is desirable that the threshold be as small as possible, in order to detect the smallest gap possible, but it must also be large enough that it indicates, with a high probability, a *true* gap on the surface, and not a gap in the sampling. With the data projected onto yz, any point may correspond to the front or back surface of the body (i.e., there are two chances of sampling a point). The probability of not sampling, p(¬sample), within a distance d in z is therefore:

$$p(\neg sample(d, z)) \approx 1 - p(\neg sample(d, u)) * p(\neg sample(d, u))$$

where: $p(sample(d, z)) = p(\exists\ p_l \text{ such that } z_{min} < z(p_l) < z_{min}+d)$

and similarly: $p(sample(d, u)) = p(\exists\ p_i \text{ such that } u_{min} < u(p_i) < u_{min}+d)$

with: $p(sample(d, u)) = 0.842, \text{ if } d = \mu_\lambda + \sigma_\lambda$

thus: $p(\neg sample(d, z)) \approx 1-(1-0.842)*(1-0.842) = 0.975$

In practice, however, the probability is actually slightly higher than this because μ_λ and σ_λ are calculated around the *whole* contour of the body. Since it is known that the sampling density is high in areas where the surface normal has a high x component, and *vice versa*, the density of a *projection* on z would be proportionately higher.

Starting from a point sufficiently far below the ends of the arms, each slice is processed. From the centre of the body outwards, space occupancy should conform to the following regular expressions, where 1 = data present, 0 = data not present, + = repeat one or more times:

`1+0+` = above underarm

`1+0+1+0+` = arm

`1+0+` = below arm

It is necessary to ensure that the gap detection starts slightly away from the centre of the body, so that the gap between the left and right leg does not confuse matters. For example, if the hands were held above or below the leg branch-point, the pattern from the centre would be 0+1+0+1+0+. Points on the outside of the detected gap are assigned to the arms, and points on the inside are assigned to the torso (unless they have already been assigned to the legs). From the highest level at which a gap is detected, segmentation proceeds by interpolating between the innermost arm point at the top of the detectable gap, and the arm branch point as detected

previously. Points on the outside of this z value are assigned to the arm, and on the inside to the torso.

Segmentation of the legs is more straightforward, if it is assumed that the z value of the branch point can be used to define a plane separating the legs. Points on either side of the plane below the branch point can then be assigned to each leg as appropriate.

4.5.4 Sensitivity to Noise

These methods occasionally produce erroneous results because of sensitivity to noise. Since the methods described make heavy use of local information, and little use of global or semi-local information, this is not surprising. An immediately useful contribution to solving this problem is to apply an outlier cleaning process before segmentation. A simple method makes use of the Euclidean distance between points to detect spatial isolation. A point, p_i, is considered to be an outlier if there is no point p_j such that $|p_i - p_j| \leq d$, for all $j = 1...n$, where $d = \mu_\lambda + 2\sigma_\lambda$. A drawback of this method was mentioned earlier, in that isolated points in sparsely sampled areas might be incorrectly identified as outliers. If an upper intensity threshold is incorporated, this helps to reduce the over-cleaning. For example, a point is considered to be an outlier if there is no point p_j such that $|p_i - p_j| \leq d \cap I_i < I_{upper}$ for all $j = 1...n$. For simplicity, I_{upper} is defined here as $2 \times I_{lower}$. This is an $O(n^2)$ operation, where n is the number of points in a slice.

4.5.5 Syntax Checking as a Confidence Measure

The processes described above make several assumptions about the body posture, and so on. Therefore, to have a good level of confidence in the automated output, it is necessary to determine if the subject and their posture actually conform to the assumptions. For this a simple "body syntax checking" has been devised. After segmentation, the spatial relationships of the key landmarks are fed into a function that tests for the conditions described in section 4.5.1. In addition, some approximate tests are made on the maximum and minimum distances between the 7^{th} cervical vertebra, the crotch point and the axillae. The range values have been determined statistically from the sample subjects in this work, with a margin of error to avoid false negatives.

Although the tests described here are extremely simple, it is interesting to consider how this might be made more sophisticated, for example, via statistical shape models [CTCG95].

4.5.6 Validation of the Segmentation Processes

Segmentation, using the processes described above, was carried out on 26 male and female subjects, with a range of ages (8-48) and BMIs (14.3-34.2). The results were assessed visually, as shown in figure 4.13.

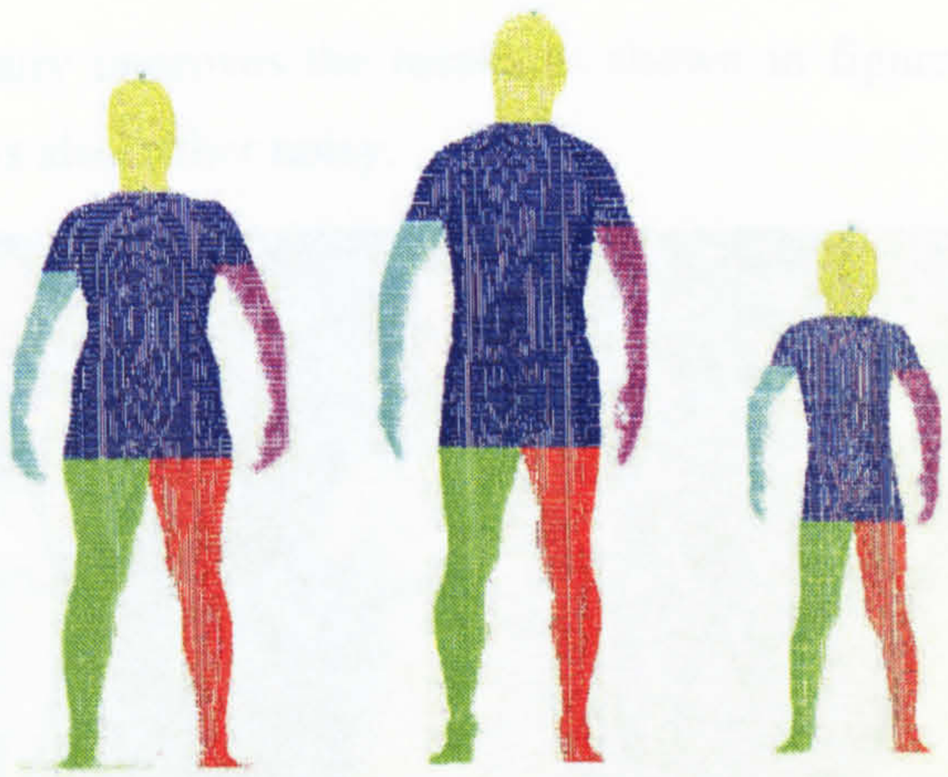


Figure 4.13: Segmented scans.

In addition, a random sample of ten subjects was tested quantitatively. Branch points were detected manually on the 3D surface and by the automated method, and their heights, y, were compared, as illustrated in figure 4.14. The branch point of the legs shows very good agreement, but the branch points of the arms show a rather high error, and are usually 2 cm or more too low. Detection of the arm branch points appears to be sensitive to the arm position, with better results if the arms are held further from the body. The leg branch point detection appears to be fairly insensitive to posture. These tests give an initial indication of the effectiveness of these procedures. Further validation will be carried out later in this chapter, by volumetric methods, and in Chapter 6, by surface anthropometry.

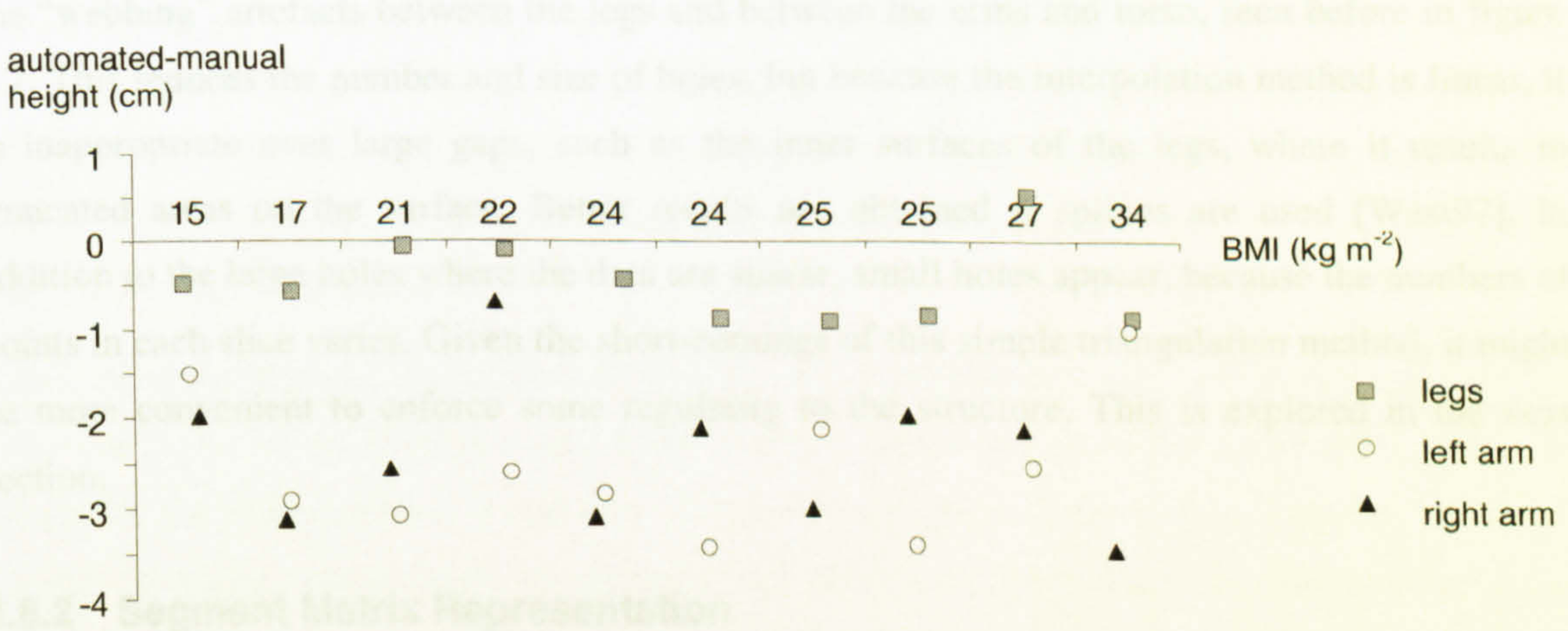


Figure 4.14: Heights of branch points (automatic–manual), versus BMI.

4.6 Surfaces on Segmented Data

Now that the images have been segmented, it is possible to explore how the results of this can be used to improve surface reconstruction.

4.6.1 Simple Triangulation

The same triangulation method, as described previously, can be used to reconstruct the surface, this time only connecting points belonging to the same segment, except at the specified segment

boundaries. This greatly improves the result, as shown in figure 4.15. Good detail is visible, although the surface is also rather noisy.

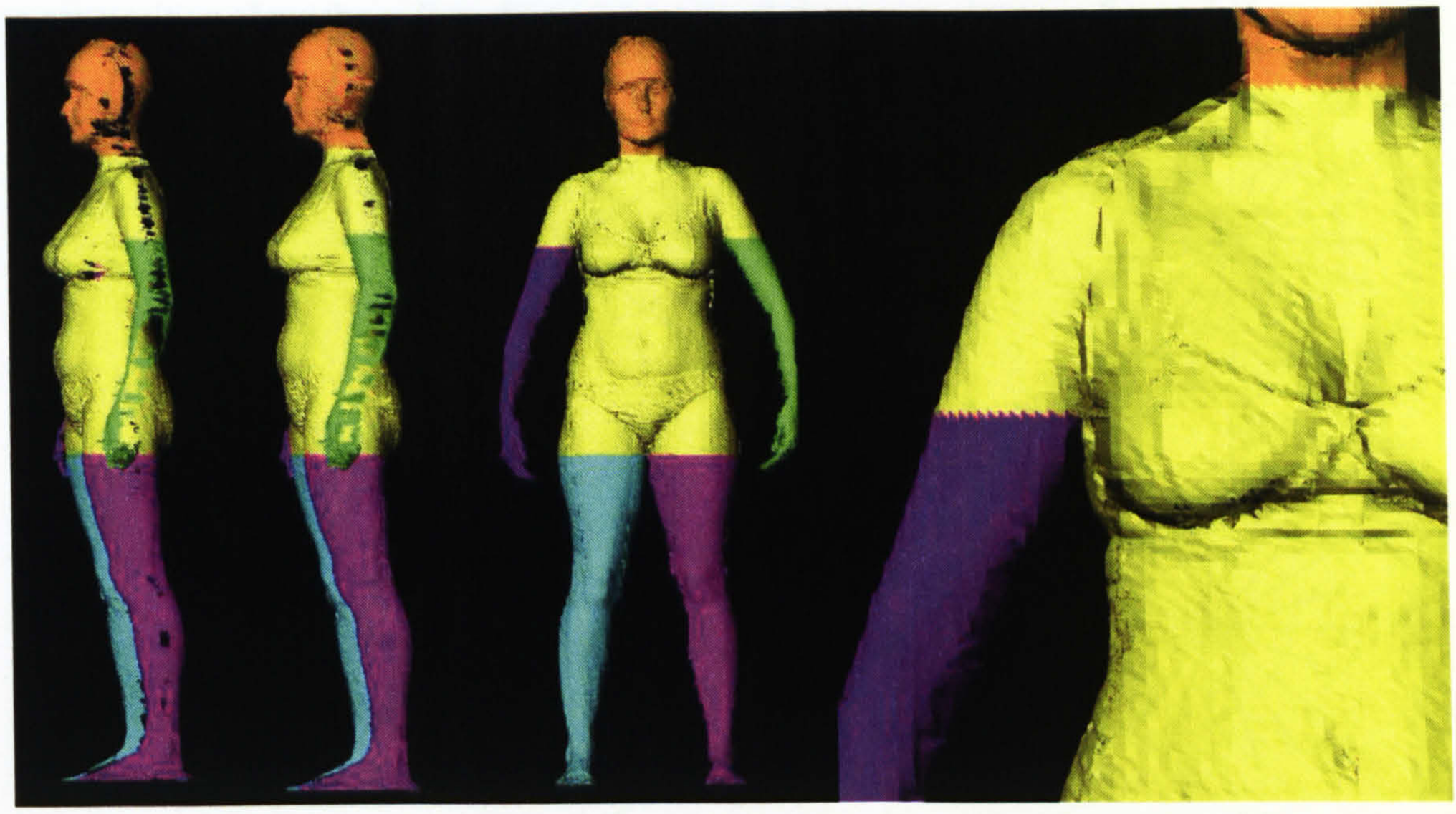


Figure 4.15: Triangulation on segmented data.
d increases in the images left to right.

Because the data have been segmented, *d* can be increased, to interpolate across gaps, without the "webbing" artefacts between the legs and between the arms and torso, seen before in figure 4.7. This reduces the number and size of holes, but because the interpolation method is linear, it is inappropriate over large gaps, such as the inner surfaces of the legs, where it results in truncated areas on the surface. Better results are obtained if splines are used [West97]. In addition to the large holes where the data are sparse, small holes appear, because the numbers of points in each slice varies. Given the short-comings of this simple triangulation method, it might be more convenient to enforce some regularity to the structure. This is explored in the next section.

4.6.2 Segment Matrix Representation

If a uniform number of points per slice is enforced, each segment can be represented as a deformed cylinder, which flattens out as a two-dimensional matrix of points. The data can be converted to local cylindrical coordinates if the centroids of each segmented slice are found. If it is assumed that no radial re-entrancy remains, the angle about the centroid imposes useful structure on the data. Each segment is represented by a $p \times q$ matrix of vertices, where p is the number of horizontal sample points on the segment, and q is the uniform vertical sampling. These parameters are selected to give an approximately uniform density, similar to the average density in the initial sample points, and a smooth change in density between segments.

At this point it is necessary to ensure that no undesirable constraints are placed on what can be represented. In effect, the representation is a radial expansion matrix, where each point can slide to any position along the radials. Some examples of shapes are shown in figure 4.16.

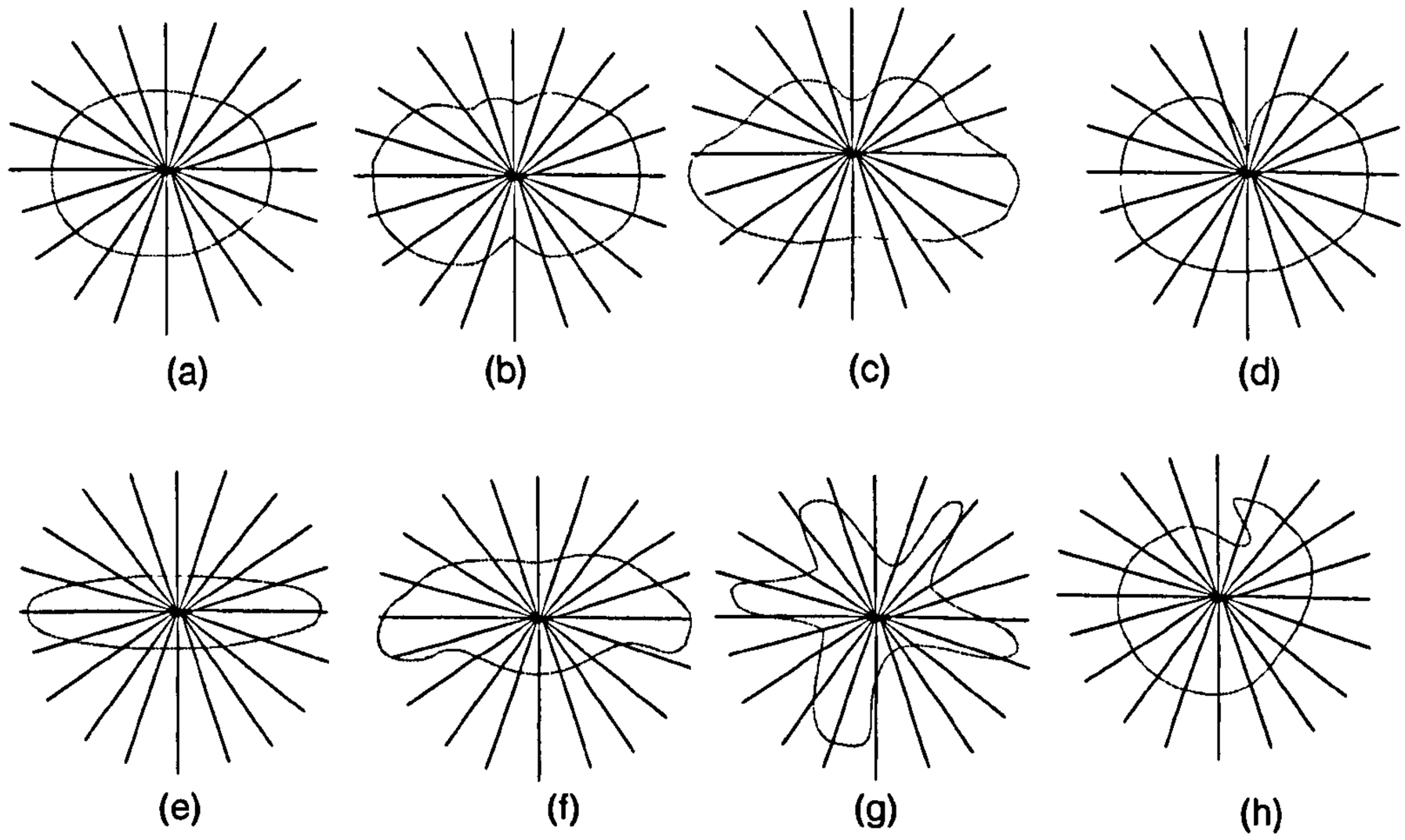


Figure 4.16: Representational constraints.

This representation is well suited to shapes that are approximately circular in cross section, where the tangent on the curve is close to perpendicular to the radial line (a, b and c). In such cases, the distribution of vertices is approximately uniform. Wherever the tangent at a given point is almost parallel to the radial line, the shape is less well represented, as the data will be sparse in these areas (e, f and g). This representation allows cusps to be retained (d), but does not allow radial re-entrancy (h).

Since the arms, legs and central part of the torso are approximately elliptical in cross section (a), there are no problems in these areas. Even on the complex shapes of the upper and lower torso (b and c), the representation is suitable. The two areas where re-entrancy sometimes occurs are on the chin, and on the lower shoulders where there may be a concavity between the upper humerus and the outer end of the clavicle (f). Since these effects appear to be very small, and are expected to have a minimal effect on the anthropometric information extracted in this work, they are not addressed further here.

In order to obtain the regular data structure, highly dense areas must be pruned and sparse areas must have missing data restored. Ideally the density would be proportional to the curvature (i.e., not necessarily uniform), but the requirement for simplicity of processing means that a compromise must be made, by erring on the side of higher than necessary density in some places. In practice, matrices with n columns on the arms, $2n$ on the legs and $4n$ on the head and torso have been used. To give a density similar to the initial sampling, n was set equal to 45 on all subjects for these experiments. The vertical dimension, q, of the matrix varies according to

the length of the segment, by default, the original 5 mm vertical resolution of the data. The resulting compound matrix representation is shown in figure 4.17.

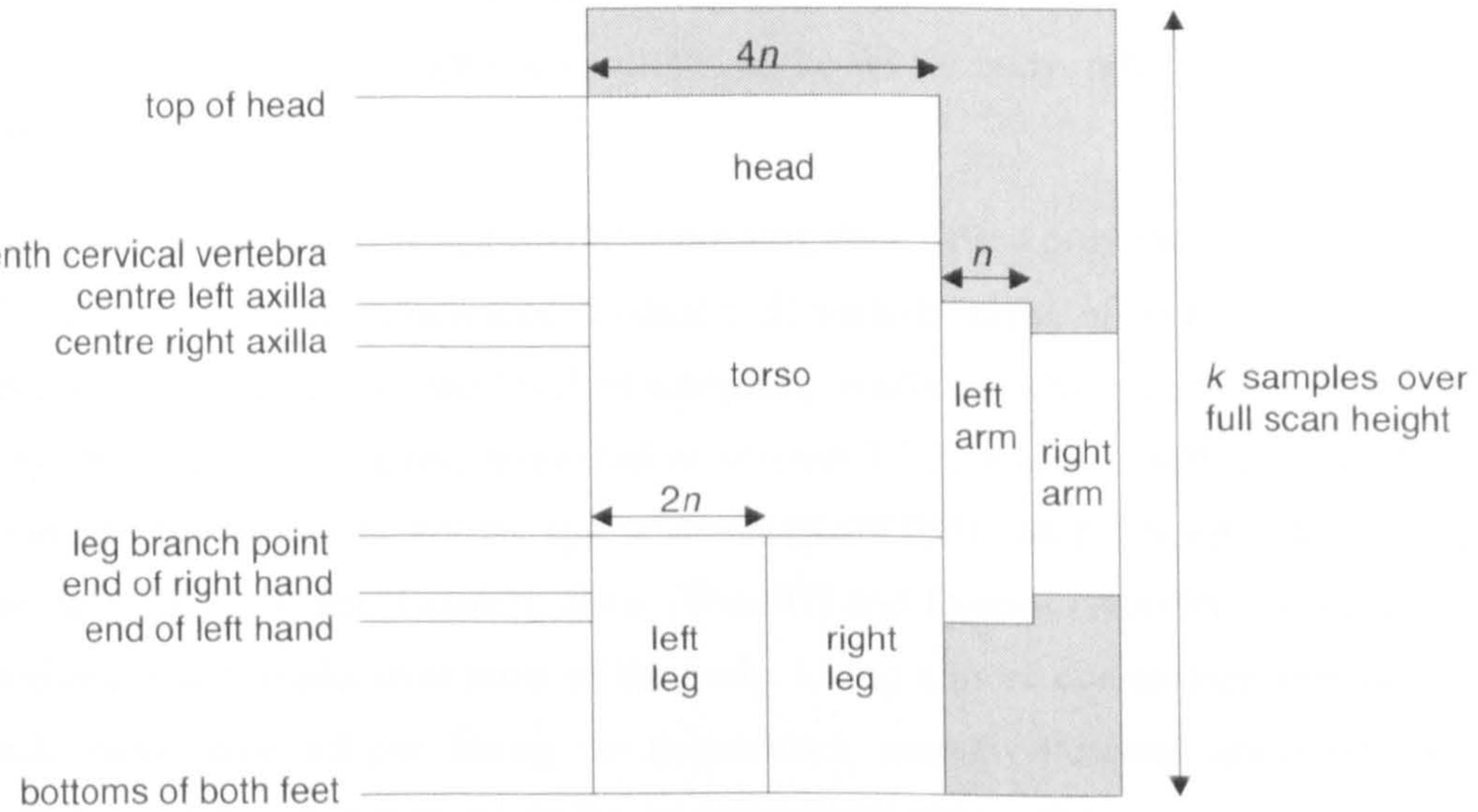


Figure 4.17: Hierarchical segment matrix.

The meshes are fitted to the segment data by first binning the points (figure 4.18) into fixed interval sectors, where the sector angle $\theta = 2\pi/p$. The radial distance of point $p_{[m][i]}$ from the segment centroid is denoted by $r(p_{[m][i]})$. Redundant points within each sector are removed by applying a median filter with respect to r, which also removes spatial outliers. In the next pass, a logical smoothing operator removes points where $|p_{[m][i]} - p_{[m][i\pm1]}| > d \wedge |p_{[m][i]} - p_{[m\pm1][i]}| > d$. All operations have $O(n)$ time complexity, where n is the number of points in the original sampling of the slice.

The connectivity and geometry of the vertices ensures that points in the same column of the point matrix have a similar radial value, and that vertices in the same row have the same y value. These constraints help to reduce the complexity of several subsequent operations, and automatically provide the horizontal contours.

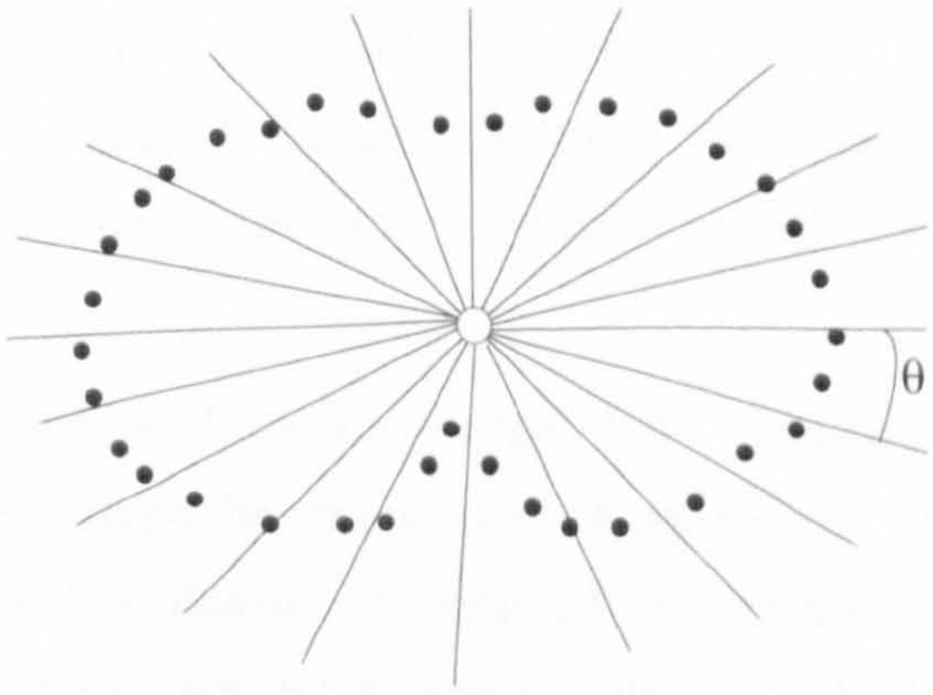


Figure 4.18: Radial binning, with bin angle θ.

4.6.3 Restoration of Missing Data

At this point missing elements in the data structure are tagged. Previous attempts at linear interpolation across gaps did not produce acceptable results. It is necessary to find a low-complexity method that gives acceptable results over the whole body, where different curvatures and data densities occur.

A commonly used technique is to approximate missing data with a polynomial curve or surface, usually of order two or three, which tends to deal well with the types of surfaces encountered on the human body, and the associated level of sampling available. This can be implemented, for example, by deformable templates, discussed in section 3.1.3, where a stiffness coefficient, or similar, controls the behaviour across sparse areas [KaWT87]. More specifically, on human-body range data, Li and Jones [LiJo97], West [West97] and Douros [Dour98] used cubic spline fitting, obtaining good results over most of the body. Using a more constrained method, Certain and Stuetzle have used ellipse fitting to reconstruct sparsely-sampled areas of the body [CeSt99]. Other valid techniques might be to interpolate across the gap imposing a smooth, perhaps linear, change of curvature or radial length.

By necessity, all these techniques use some sort of heuristic to infer the most likely behaviour over the missing data. The biases they impose are well understood, and given that they operate in the absence of prior knowledge, they produce good results. However, the results of spline fitting highlight some of the problems of using them on large gaps. On most of the arms and legs, a similar curvature applies across the whole cross section, so the existing data can effectively control the spline to reconstruct in a feasible manner. However, at the top of the legs, if they touch, the missing data is actually planar, and so a spline interpolating technique will fail to reconstruct accurately. On the insides of the hands the surfaces are also approximately planar or slightly concave, and therefore a linear interpolator might give acceptable results, but splines tend to produce convex surfaces, which are not representative of the true contours. These variations make it difficult to find any single method that would produce acceptable results; on some areas a technique such as cubic spline interpolation can be used, and on others prior knowledge appears to be required. In addition, these techniques tend to be computationally expensive, so it is interesting to investigate whether or not a simpler method can produce acceptable results.

As discussed earlier, in section 4.1.3, knowledge of the geometry of the imaging components can be used to advantage. Given the arrangement of the sensors, as illustrated in figure 2.7a, it would be expected that low density is obtained on surfaces whose normal has a small component in x, and surfaces that are almost parallel with the incident light beam, in other words, where the normal has a small y component. Areas that are occluded because of mutually shadowing body segments, such as the branch points of the arms and legs, also tend to have a small x component, so these do not have to be treated as special cases. For *large* gaps to occur,

the x and y conditions must hold over that area, and therefore this will tend to be in areas of low curvature. This suggests that the interpolator should impose a low curvature bias. Interpolating the radial lengths across the empty elements of the segment matrices imposes a circular bias, tending to over-expand the surface (4.19a), especially on the inner surfaces of the hands. Point-to-point linear interpolation truncates the surface (4.19b). A combined point-to-point linear and linear radial interpolation (4.19c) takes the mean of the radial and point-to-point interpolators, and produces more acceptable looking results in a computationally efficient ($O(n)$) way.

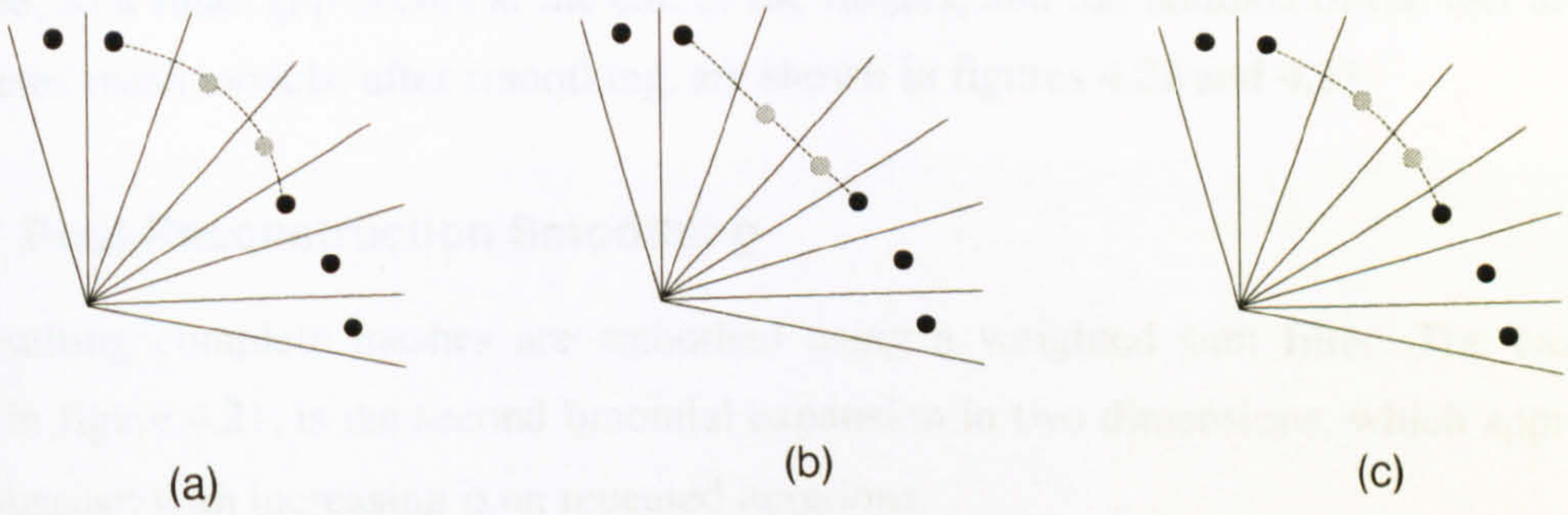


Figure 4.19: Linear radial, point-to-point linear and combined interpolation.

4.6.4 Closing and Rendering the Model

The surface is now composed of a set of meshes, which mostly describe each segment well, but do not represent connectivity across segments. For visualisation, the meshes can simply be rendered as they are, by generating a vertical seam to join the zeroth and pth points of each row, as shown in figure 4.20.

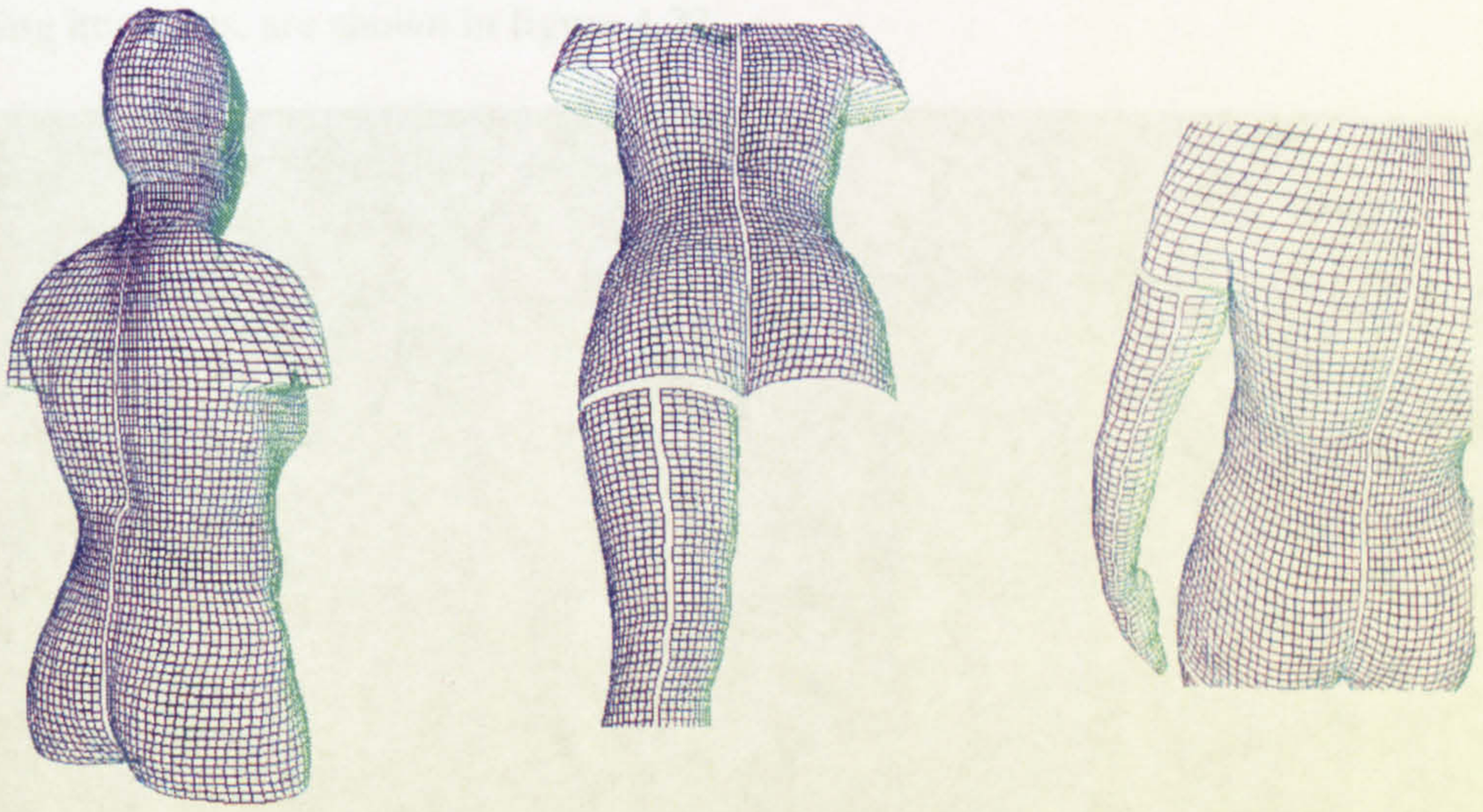


Figure 4.20: Segment mesh joining.

At the boundaries of each segment, visual or actual closure should occur. At the top of the head, closure is enforced by simply duplicating the top-of-head landmark p times to generate the first row of data. Since the head and torso both have the same p value, they join smoothly. However,

where the arms and legs join the torso this is not the case, because an unknown number of points on the torso should join with each arm and each leg. For visualisation purposes it is possible to obtain an acceptable result by simply overlapping the meshes, but it would be desirable to represent the connectivity information explicitly for further processing operations. The tops of arms join the torso at known rows (where the underarm points are located), and the legs join at the same, bottom-most, row of the torso. An array of nearest neighbours is generated for each direction across the join, which stores the relevant connectivities. Complete closure is not enforced, so a small gap occurs at the end of the fingers, and the bottoms of the feet are absent. Completed mesh models, after smoothing, are shown in figures 4.22 and 4.23.

4.6.5 Post-Reconstruction Smoothing

The resulting complete meshes are smoothed using a weighted sum filter. The basic filter, shown in figure 4.21, is the second binomial expansion in two dimensions, which approximates to a Gaussian, with increasing σ on repeated iterations.

0	1	0
1	2	1
0	1	0

1	
1 2 1	$\sigma = 0.58$
1 4 6 4 1	$\sigma = 1.03$
1 6 15 20 15 6 1	$\sigma = 1.23$
...	...

Figure 4.21: Basic 2D filter, and the binomial expansion for iterations.

It is possible to vary the window width for each segment, or equivalently, to keep the same window width and iterate a variable number of times. The results of this smoothing operator, at increasing iterations, are shown in figure 4.22.

Figure 4.22: 0, 1, 2 and 7 iterations of the smoothing operator.

In practice, the number of iterations has been chosen empirically, based on visual assessment of the results, and is ordered *head* < *torso* < *arm* < *leg*. Figure 4.23 shows output of the uniform smoothing operator at (a) 2 and (b) 7 iterations, and (c) output of the "adaptive" smoothing operator, which shows better smoothing on the insides of the legs than the 2-iteration operator, and better level of detail on the torso than the 7-iteration operator. There is, of course, more than one criterion for smoothing. Visual acceptability and appropriate scale for subsequent shape analysis are both of relevance here. For example, the smoothing scale is important in detecting underlying curvature against local changes in curvature. This will be discussed further in Chapter 5, when differential geometric properties of the surface are explored.

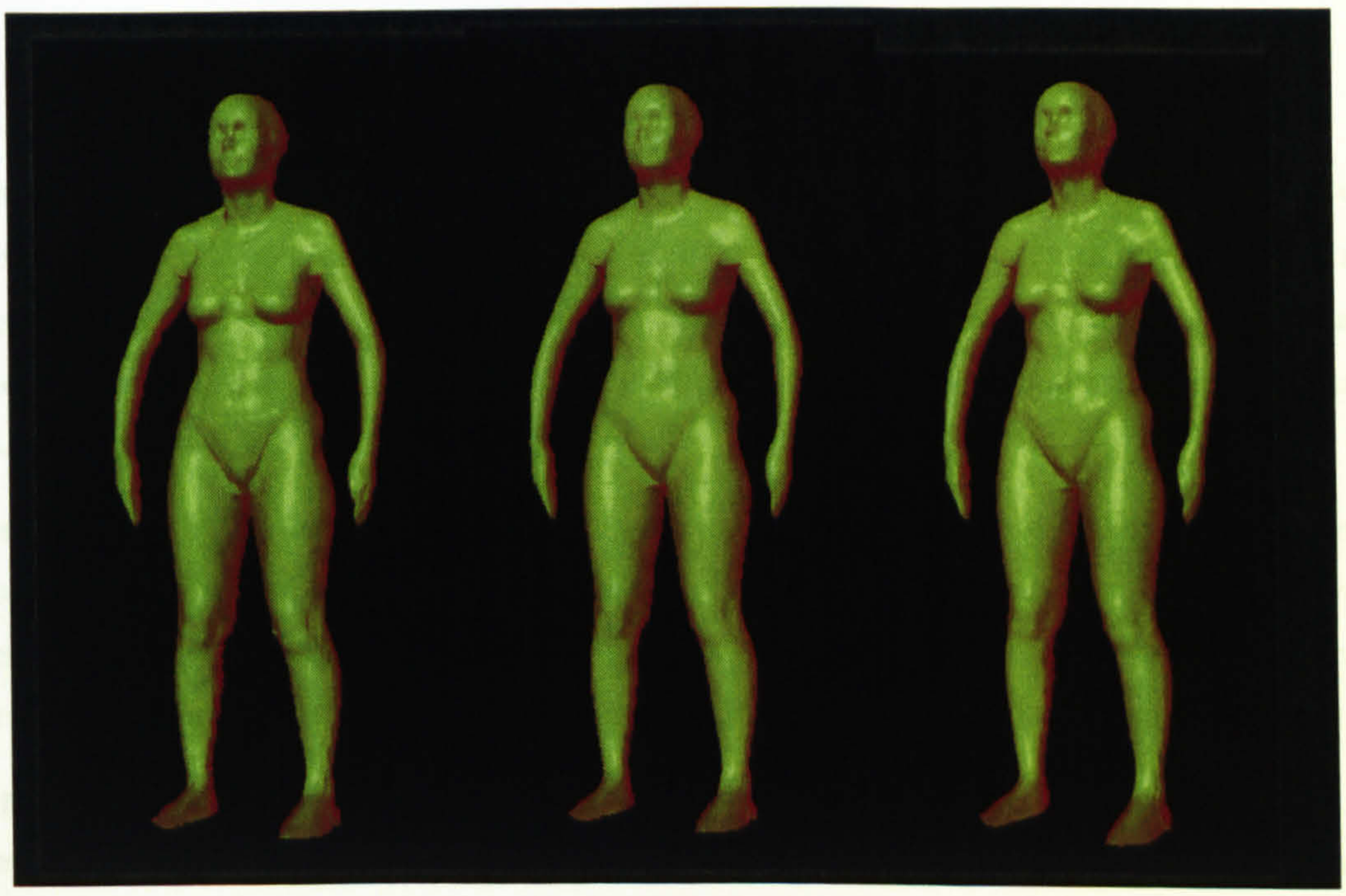


Figure 4.23: (a) uniform smoothing 2 iterations, (b) uniform smoothing 7 iterations, (c) "adaptive" smoothing.

At this point the body is segmented in a relatively simple way, but in the following chapters as more landmarks are found, more specific operators could be applied, based on detailed sub-division of the body. For example, the hands could be treated differently than the arms.

4.6.6 Summary of Surface Reconstruction Processes

Various experiments have been described to reconstruct the body surface from raw data. Some of these methods were discarded in favour of techniques that produced better results, so it is worthwhile now to summarise the processes that have finally been implemented:

1. remove very low intensity noise, by intensity thresholding;
2. spatial thresholding: remove data that correspond to known structures of the scanner itself;
3. use knowledge of expected surface normals to detect and remove multiple reflections;

4. spatial outlier cleaning;
5. detection of key landmark points;
6. segmentation: label points according to each body part;
7. radial binning: impose a regular radial structure on the data for each segment;
8. median redundancy filtering: reduce the data to a regular matrix of points;
9. median smoothing: remove noisy points, using the matrix connectivity information;
10. mesh interpolation: fill missing elements in the matrices;
11. iteratively apply a simple, isotropic, Gaussian smoothing filter.

4.7 Volumetric Validation

The visual inspection carried out during these stages of processing serves to detect "obvious" errors, but does not give quantitative results on the overall cleaning, segmentation and skinning processes. It would be possible to carry out tests on phantoms or run simulations to validate the methods. However, because of the many problems of the living human, it is necessary to validate based on live subjects if possible.

In-sample tests have already been carried out, in section 4.5.6, to develop the segmentation techniques and it would also be useful to validate the surface reconstruction. Since body volume can be determined by other methods [FJLC92] [MGBM95], which are considered to be accurate and reliable, this might be a useful method to contribute towards a quantitative assessment.

The processing operations being validated at this stage are effectively an implementation of the generic model of the human body, so it is reasonable to expect that equally good results will be obtained for women, men and children, of various builds. It will be of interest, therefore, to determine if any particular bias is present.

4.7.1 Data Collection

In these experiments data from 23 subjects have been collected. Each subject was scanned, and their volumes were determined by both underwater weighing [FJLC92] and air displacement plethysmography [MGBM95]. These are considered to be the most accurate methods available for body volume determination, and usually give results to about 1% error.

Underwater weighing measures the body tissue volume (without lung volume) but air displacement plethysmography measures body volume and residual lung volume (at expiration), so corrections have been made to derive comparable total body volume (including lungs) and total lung volume (including residual volume) for each technique.

Subjects were scanned in the "reference posture" as described in Chapter 2, at a mid-tide lung volume. To estimate total body volume including full lung volume (to enable comparison with

the reference techniques) half the independently measured tidal lung volume was added to the results.

Subjects were processed by each technique in as close a time frame as possible to ensure that minimal change in volume had occurred. In all cases the techniques were carried out within the same day, or at the same time of day on two consecutive days. Each process was carried out twice on each subject. Similar numbers of children, adult women and men are taken from the original data set to represent a wide range of builds, although the extremes are not represented.

Table 4.3: Characteristics of sample.

age/gender	7 x children, aged 8-16 7 x adult female, aged 24 – 42 9 x adult male, aged 22 - 48
weight range	24 - 78 kg
height range	1.28 – 1.84 m
BMI range	14 - 25

4.7.2 Volume Estimation from Scans

The regular matrix structure of the compound deformed cylindrical mesh representation makes it possible to calculate the volumes for each slice separately, then integrate them to acquire an overall measurement [Yama98]. In Yamanaka's method the total volume is determined by summing the volumes relating to each quadrilateral face, if they are treated as "cheese wedges" about the slice centroid, as illustrated in figure 4.24. Each pair of approximately parallel sides on the quadrilateral outer face of the wedge is averaged so that it can be treated as a rectangle in a plane orthogonal to the overall slice plane. The wedge volume is then calculated as the area of the resulting scalene triangle, *a'b'centroid'*, extruded along the *y* axis.

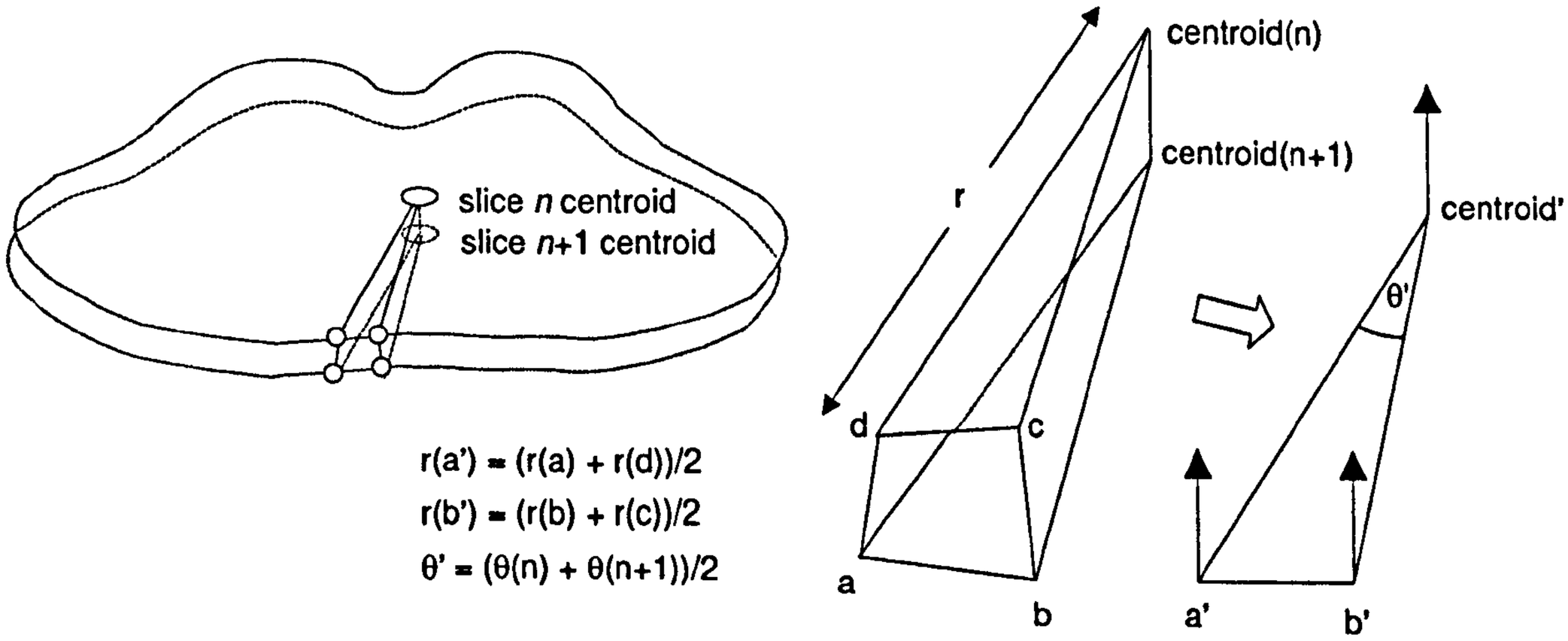


Figure 4.24: Integrated wedge method for volume estimation.

4.7.3 Results

Results are presented in table 4.4 and figure 4.25. Each method was carried out twice on each subject. The differences between each of the two runs allow a tentative reproducibility measure, averaged over all subjects, to be computed, as indicated in table 4.5.

Table 4.4: Body volumes (including lungs).
air displacement = ad, scanner = sc and underwater weighing = uw.

subject	mean volume (l)	lung volume (l)	ad-uw (l)	sc-ad (l)	sc-uw (l)
1	25.25	1.50	0.077	0.867	0.943
2	26.19	1.90	0.023	0.508	0.531
3	27.92	1.40	-0.074	0.494	0.420
4	33.27	1.90	0.029	-0.755	-0.726
5	41.93	2.70	0.149	-0.823	-0.674
6	50.06	2.90	0.135	-0.754	-0.619
7	52.49	3.30	-0.453	1.165	0.711
8	54.49	3.70	-0.266	-0.193	-0.459
8	55.56	3.10	-0.609	0.183	-0.426
10	56.85	3.70	0.022	1.567	1.589
11	58.30	3.20	-0.132	-0.103	-0.236
12	58.94	3.10	-0.723	-0.518	-1.241
13	59.15	3.80	-0.950	2.462	1.513
14	59.23	3.40	0.385	-2.702	-2.317
15	63.16	4.15	-0.345	0.089	-0.256
16	64.26	3.20	0.000	1.022	1.022
17	64.42	4.70	-0.086	1.643	1.557
18	65.69	3.90	0.621	0.011	0.632
19	65.84	3.90	-0.650	-0.166	-0.815
20	67.68	3.65	-0.063	0.880	0.817
21	67.86	4.00	-0.464	0.261	-0.203
22	71.38	4.20	-0.284	1.771	1.487
23	77.94	4.20	-1.358	1.485	0.127
		RMS	**0.233**	**1.329**	**0.997**

Table 4.5: RMS intra-technique difference.

	intra-technique difference (litres)
scanner	0.427
air displacement	0.191
underwater weighing	0.312

The body volume tests show a good agreement with current volumetric methods over a range of body types. Figure 4.25 shows the differences in the techniques, (a) as a function of mean volume by all techniques and (b) as a function of BMI. No particular bias appears to be present. However, the accuracy is not yet considered to be high enough to use the values for estimating quantities such as percentage body fat in a meaningful way [WDFE99].

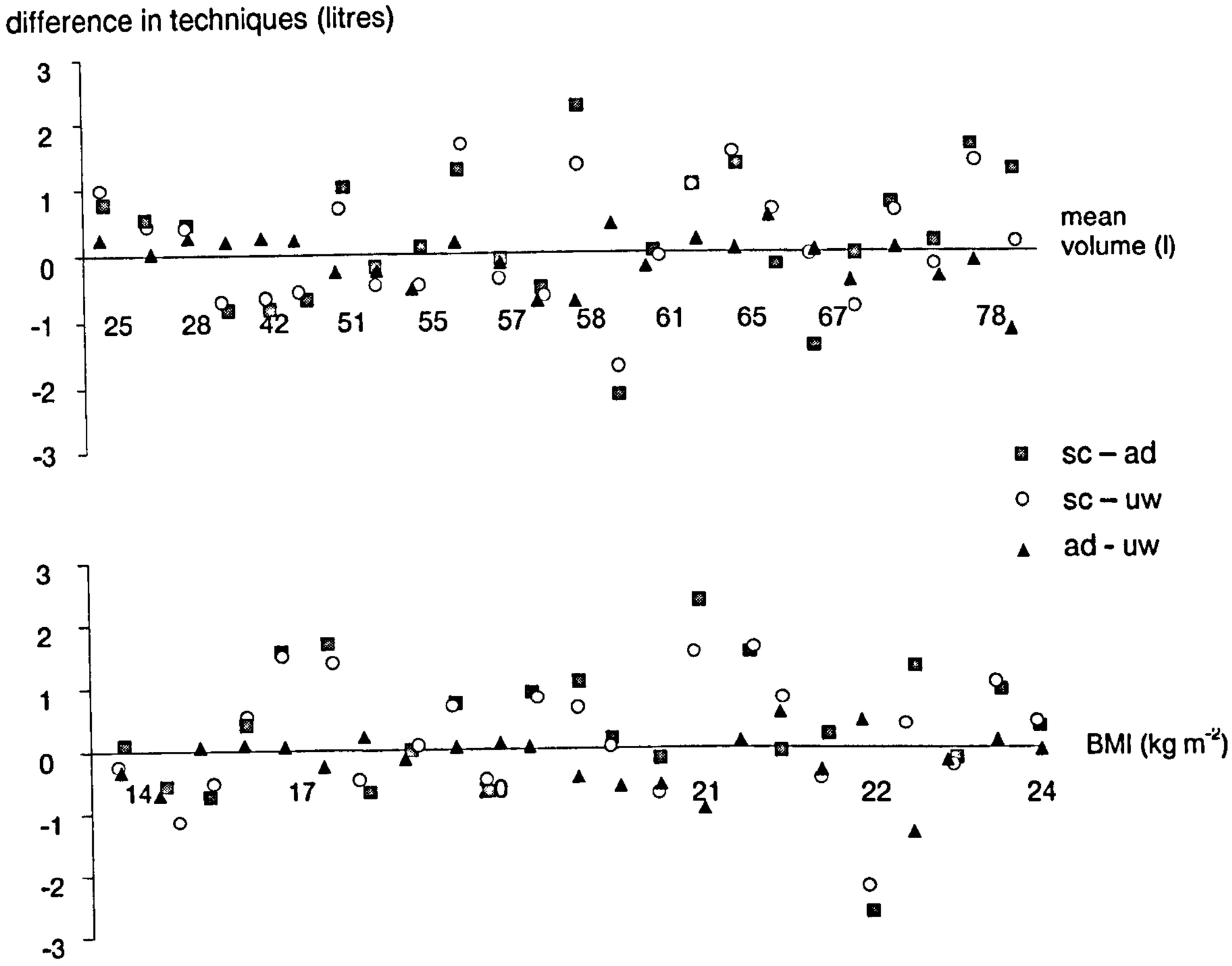


Figure 4.25: Inter-technique differences.
(a) versus mean volume by all methods, (b) versus BMI.

Data on reproducibility of the scanning technique are currently limited, but the intra-technique differences in table 4.5 show similar values for all the techniques. It is interesting to note that these *intra*-technique differences are similar to the *inter*-technique value of (ad-uw), whereas the inter-technique values of scanning versus the other two techniques are significantly higher. This suggests that the differences between the scan-derived measurements and the other two techniques might be the result of systematic differences that arise on a subject-by-subject basis.

A possible source of error is the variation in methods for determining lung volume. In particular, large differences in lung volume (up to two litres) could occur when subjects attempt to find and hold their mid-tide volume for each scan. Subsequent tests have been carried out by Wells *et al* to measure the lung volume held during scanning [WDFE99]. The results show an improved precision of the scan-derived measurements, but there is still a discrepancy between the techniques, which must be explained somehow. This was also noted by Smith *et al*, who carried out similar experiments [SmJW90].

There are several other possible sources of error, each of which might cause a loss or gain in terms of the volume measurements. Earlier, it was acknowledged that the representational form imposes a slight bias, by "filling in" small concavities, however, this is likely to cause only a few millilitres gain. The sparseness of data on the inner surfaces of the arms and legs might

result in a significant loss, if the surface restoration algorithms discussed in section 4.6.3 are not performing correctly. In addition, it is possible that the method used to calculate volume might impose a bias, which is positive in some body shapes and negative in others. If the body is treated in a simplified way, it is possible to estimate the effects of depth measurement error on the volume calculations. For the purposes of this discussion, the body is considered as a set of three cylinders: the head and torso together are represented as a cylinder of diameter 150mm and height 1000mm; the legs are represented by two cylinders of diameter 75mm and height 750mm. The total volume of this is 97.19 litres. If a systematic error of +1mm in depth measurements is present over the whole body, the volume increases to 98.85 litres, giving an error of +1.66 litres. Although it is assumed that errors in depth measurement are *not* correlated in this way, this simple example does serve to illustrate how high an error can arise in the volume measurement from a small error in the original sampling or post-processing of the surface.

4.8 Discussion

Visual assessment has been useful during development of the techniques, to indicate obvious errors in cleaning, segmentation and skinning. The segmentation processes have also been assessed by comparing the heights of the automatically detected branch points against those located manually on the scanned image. Although the leg segmentation method appears to give reproducible and accurate results, the arm segmentation method appears to be rather sensitive to arm position, giving better results if the arms are held further from the body. Several improvements might be made, perhaps using a step detection technique similar to that of Li and Jones [LiJo97]. The body volume tests showed a good agreement with current volumetric methods over a range of body types. However this method is limited as a validation technique because of the many potential sources of error in capturing the original image, and the many processing operations required to determine body volume from the scan.

The radial expansion matrix, described in section 4.6.2, provides a useful way of representing the body over most areas. A few limitations were noted, which means that certain areas are less well represented, such as at the branch points of the arms and legs, under the chin, and at the shoulder joints, where the body surface sometimes folds over slightly and radial re-entrancy occurs. While these short-comings are acknowledged, the representation has many advantages, notably the low time complexity to generate the surfaces, and the useful structure it imposes on the data to simplify subsequent processing.

The interpolation method appears to produce visually acceptable results over most of the body, although the "stepped" appearance on the inner surfaces of the arms and legs indicates the need to make more use of vertical coherence between the data slices. Application of a large-window

Gaussian filter makes a significant improvement here, but some ripples are still discernable on the resulting images.

However, as with the previous processing steps, the architecture of the system allows each processing operator to be replaced at any time by one employing a different technique as long as it can take the same structure of input and produce the same type of output. The choices taken here are simply those based on the criteria stated in Chapter 1, and the current state of understanding of the data and processing techniques. It is always possible to make use of another method in the future, as techniques are improved, and greater computational resources are available. For example, more use could be made of subsequent landmark labelling to control data interpolation operators retrospectively.

Validation of these processing techniques *in vivo* is a difficult task. Three methods have been used here:

- visual assessment of images;
- numerical comparison of automatically generated branch points versus manually located ones;
- volumetric comparisons of the fully skinned images.

Further quantitative validation will be carried out in Chapter 6, where automated surface anthropometrics will be developed. It would be useful to carry out more tests on the reproducibility of all these processes; this is identified as an immediately important task for future work.

4.9 Summary

This chapter has followed the transformation of the raw data to a polygon mesh surface representation. Starting from a simple triangulation method, with few built-in assumptions, more assumptions and constraints have been gradually added, where this can be justified in terms of the assessment criteria set down in Chapter 1. The resulting processes are partially data-driven and, where necessary, also include some *a priori* knowledge of the imaging system, the body geometry and the subject's stance.

Segmentation algorithms have been developed to detect the branch points of the body and partition the image into a set of high-level body segments. This facilitates the surface reconstruction process, and allows a simple two-dimensional matrix representation method to be used. Finally, quantitative validation has been carried out, by comparing the resulting enclosed volumes with existing volumetric techniques. Good correlation is found over a wide range of body types, although some discrepancies are still present, and these must be addressed if the technique is to be used in precision applications, such as the estimation of body composition.

Chapter 5

From Surfaces to Features

Now that the body is represented as a complete surface it is possible to start analysing it in terms of size and shape. Of particular interest are operations that can be used as discriminants for detecting features. This chapter presents an exploration of some surface transformation operations to extract low-level features and their attributes. These operators are collected together at the end of the chapter into a library for constructing feature detectors in Chapter 6.

5.1 Exploration of Shape Description Techniques

This chapter explores methods to extract features and their attributes from the surface. In this discussion "features" refers to entities on the surface: points, contours, patches, and so on; "attributes" refers to the characteristics of those features.

The goal of this chapter is to find useful, reusable *primitives* for feature detection. At the end of the chapter, in section 5.6, the operators are brought together into a library. In the next chapter this library will be used to build detectors to locate application-specific features defined in the anthropometric standard ISO 8559 [ISO89], and summarised in Appendix A. Some of these operators have already been used in the previous chapters to explore the data. For example, in the last chapter non-convex, planar, closed curves were used to determine the data density with respect to contour length.

The investigations can be broadly separated into three levels at which they describe the body:

- high-level structure and relational information (for example, central axes);
- local attributes from differential geometry (for example, critical points of curvature);
- surface lengths and convex lengths of open and closed curves (for example, to simulate a tape measure).

5.1.1 Representation and Basic Attributes

Each data set is now represented as a compound two-dimensional matrix of points, with connectivity within the segment implicit in the matrix, and connectivity between segments explicitly defined. From this representation it is possible to:

- extract the cylindrical and Cartesian coordinates of each vertex;
- find the nearest surface point to an arbitrary point in space, using various distance metrics;
- determine to which segment each vertex belongs;
- extract connectivity and proximity information;
- extract horizontal and vertical "slices" of data from the matrix.

5.1.2 Data for Experimentation

Data for experimentation in this chapter is made up of a mixed group of subjects, from 8 to 48 years of age, and a range of body shapes and weights, as described in Chapter 4. All subjects were scanned in the reference posture, to ensure that the investigations focused on the body in the appropriate stance.

At this point the experimental framework allows scans to be processed automatically from download (or, equivalently, read in from a stored file) to generate the surface structure. Landmarks can then be placed manually on the reconstructed surface and labelled, so that images can be "annotated". This makes it possible to assess the anatomical and anthropometric significance of the shape analysis processes explored in this chapter.

5.2 High-Level Structure

In the previous chapter the body was segmented into its main components. This section discusses other methods for explicitly describing structure at a high level, to provide context and reduce the search space for feature detection.

5.2.1 Central Axes

With the data segmented, the central axis of each segment can be generated. This is done by concatenating the centroids for each horizontal slice of data along the vertical axis, as shown in figure 5.1. One simple way to find the centroid is at the centre of the bounding box for each slice. For simplicity, the edges of the bounding boxes are constrained to align with the x and z axes.

An obvious enhancement would be to use the first-order moment [Nurr97] instead of the bounding box centroid. This would be more resilient to noise, and would be a true reflection of the actual data, taking into account slight rotations in posture.

This method is acceptable except where the branching points at the underarms occur at different levels (as is often the case). In such situations the axis makes a sudden shift, and then returns to the centre after passing the second branch point. A simple solution is to interpolate the axis between the two branch-point levels, although a true medial axis method, such as a 3D "prairie fire" algorithm, might be more acceptable overall [DuHa73].

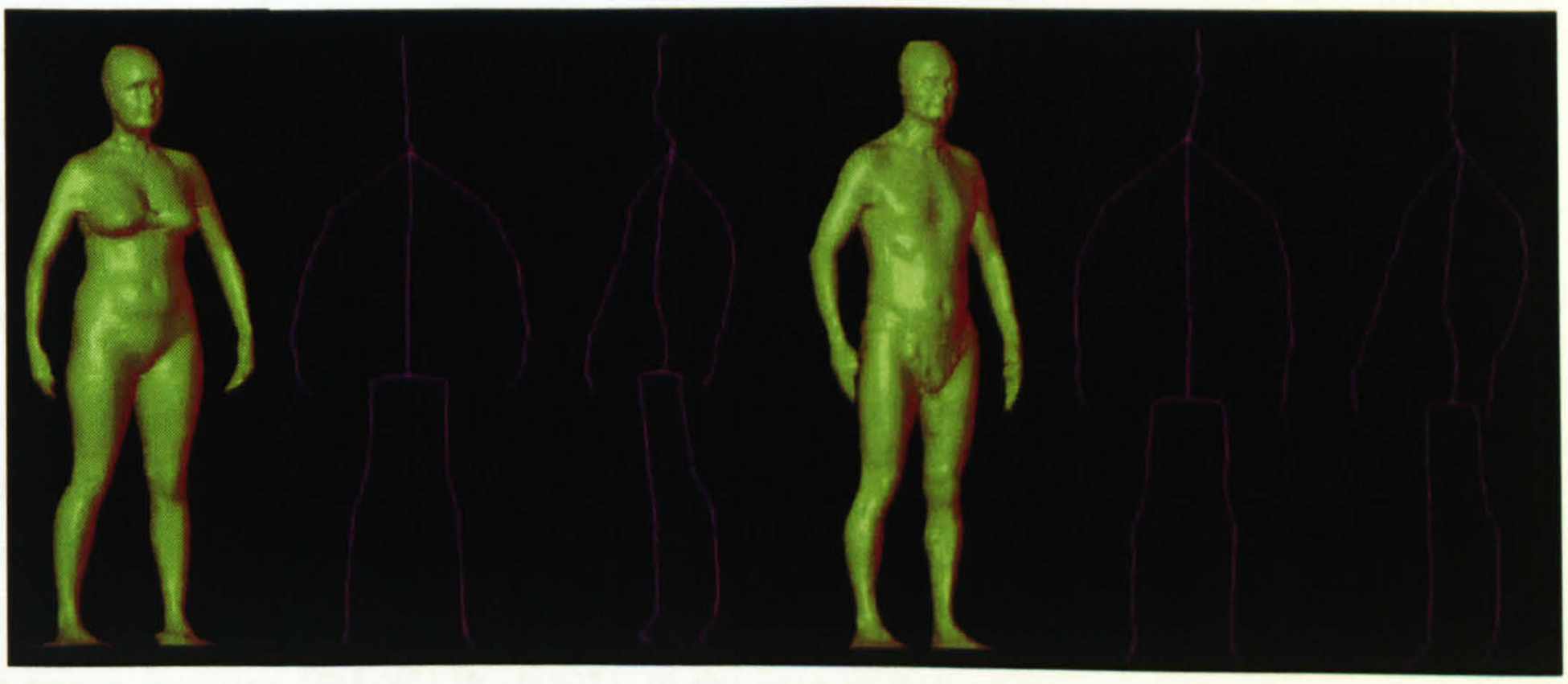


Figure 5.1 Central axes by centroid interpolation.

5.2.2 Space Partitioning and Search Constraints

Segmentation makes a significant contribution in constraining the search space for a particular feature. With the axes found, relational constraints can also be imposed. For example, it is then possible to constrain a search to the inner or outer side of a leg. If (cx, cy, cz) are the coordinates of the centroid:

$z > cz \Rightarrow$ left side

$z < cz \Rightarrow$ right side

$x < cx \Rightarrow$ front

$x > cx \Rightarrow$ back

It is also useful to be able to constrain the search to an area "close" to the centre of the body. For example, the navel is expected to be located close to (but not necessarily exactly *at*) the centre in z, in other words, where $z \approx cz$.

As new landmarks are found, they can be used to delineate sections of the body volume to further constrain the search. For example, the acromion will be located below the seventh cervical vertebra and above the axilla ($y(acromion) < y(7^{th}\ cervical\ vertebra) \cap y(acromion) > y(axilla)$).

5.3 Differential Geometric Properties

This section explores how differential properties can be extracted from the surface in a computationally efficient way

5.3.1 Derivatives and Curvature on Sampled Surfaces

In estimating derivative information on sampled surfaces, it is necessary to determine whether this can be determined directly from the polygon mesh data (for example, by a weighted window operator), or if it requires a function to be fitted to the data. Many researchers have used local least squares curve fitting techniques, on the basis that a better estimate of derivatives can be obtained that way, particularly on noisy data. Typically, a polynomial curve or surface is fitted in a local neighbourhood around the point (the so-called "facet model") and the derivatives of the point are calculated from the expression of the curve [Beau78] [Besl88] [Fish89] [LeHD93] [StIW95].

However, these methods are computationally expensive, and it would be desirable to use a simpler method if it does not compromise the ability to generate useful characterisations of shape. Others have discussed methods that do not make use of any kind of surface fitting, but these start with the assumption that the surface normals have been found for the points adjacent to the one in question [FlJa89] [KaHä98]. In the case of derivative and curvature measures it is arguable that the *relative* values are of more importance. If this is the case, then it might be acceptable to use a more computationally efficient method, as devised by Guy and Medioni, who determine sign and direction of principal curvature on sparse, noisy data by a voting method, based on various preferred surface characteristics [GuMe97]. However, in the case of the data for the human body, a surface, with connectivity and orientation information, has already been derived. In addition, since a significant amount of smoothing has already been done, noise should not be a critical issue.

Because the body is represented as a two-dimensional matrix, it is possible to take advantage of this regular structure to simplify operations significantly, not just for curvature estimation, but also for other derivatives. Discrete estimations are often used in two-dimensional greyscale image processing, and should be equally applicable here. For example, the Laplacian can be estimated by a convolution operator with a 3×3 mask, as shown in figure 5.2. In fact, the operator can be defined in many ways. Two main decisions must be made: (a) the extent (width) of the neighbourhood, and (b) what weighting to give each point. If the neighbourhood is too small the estimate is sensitive to noise; if it is too large, discontinuities and other high-frequency changes in curvature can affect the true estimate. Much of the noise should already have been removed in the earlier processing operations, so the window need not be very wide. Consequently the convolution masks used here are of length three, corresponding to approximately 1×1 cm on most of the body. An odd-numbered mask has the advantage of

generating the value associated with the centre element, while an even-numbered mask generates the value relating to a point midway between elements. Weightings on the elements are dependent on the purpose of the operator, and will be discussed separately in each section.

0	1	0
1	-4	1
0	1	0

Figure 5.2: A 3×3 convolution mask for discrete estimate of the Laplacian.

This work will only look at the first and second derivatives, which should provide a useful set of transformations on the surface for feature detection, without resort to higher derivatives, which have the disadvantage of being increasingly sensitive to noise.

5.3.2 Viewpoint-Invariant Curvature

The viewpoint-invariant curvature measures discussed in Chapter 3 are investigated here. The theoretical underpinning of curvature analysis must now be translated into an efficient method for computing the values. Curvature in the plane through the normal is determined by finding the radius of the circle fit to the contour (figure 5.3a). At any one point, the point and its two immediately adjacent neighbours on the planar contour (p_{i-1}, p_i, p_{i+1}) (figure 5.3b) can be considered as lying on this circle, which is estimated in its discrete form as a regular polygon. The angle θ at the centre of the circle and two points on the contour (p_{i-1} and p_i, or p_i and p_{i+1}) is therefore approximated by π-ϕ, where ϕ is the inner angle between the triple of points (figure 5.3c).

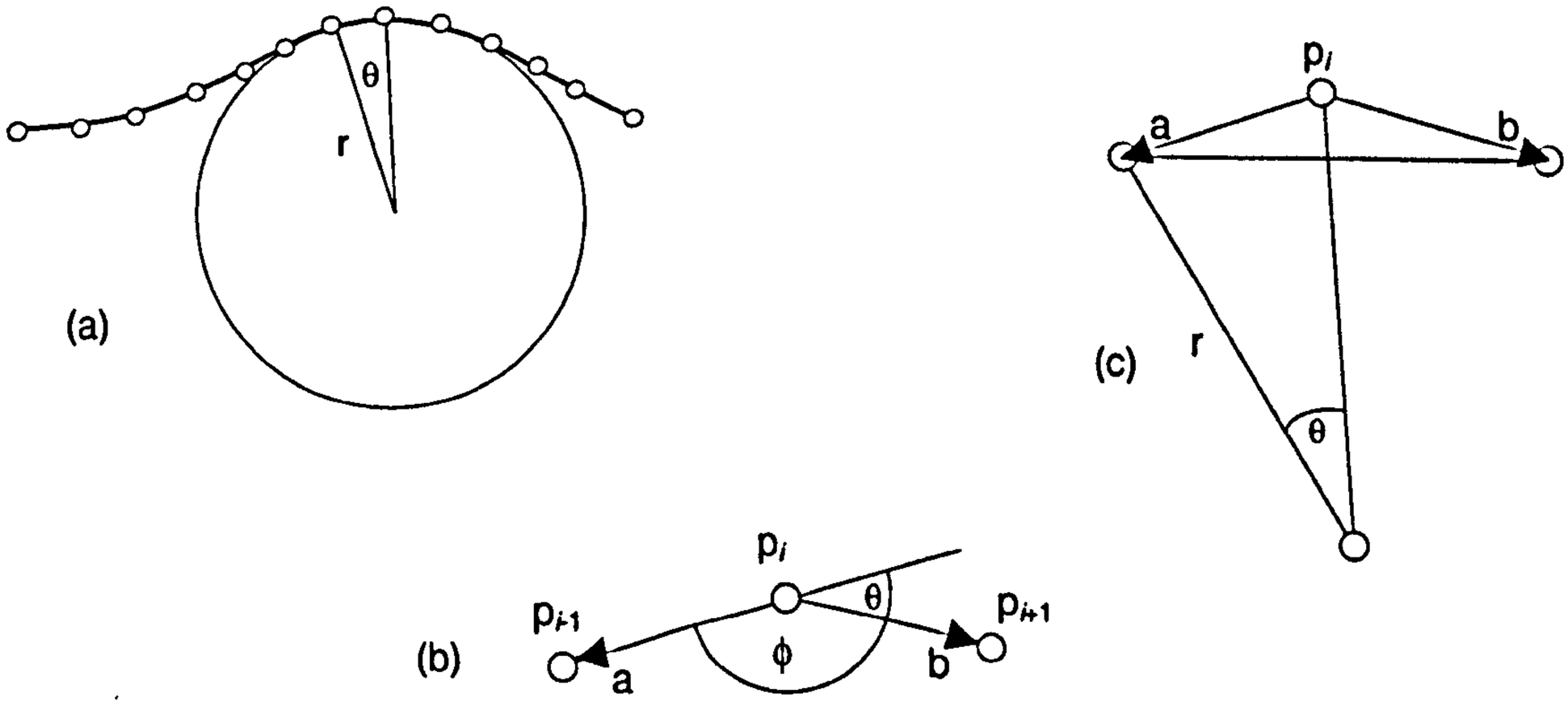


Figure 5.3: Curvature from discrete data.

This inner angle is determined from the dot product of the vectors between the sampled points:

$$\phi = \arccos\left(\frac{a \cdot b}{|a||b|}\right), \text{ where } \begin{matrix} a = p_{i-1}\text{-}p_i \\ b = p_{i+1}\text{-}p_i \end{matrix} \qquad \textbf{Eqn. 5.1}$$

The curvature is estimated from:

$$\theta = \pi - \phi \qquad \textbf{Eqn. 5.2}$$

$$r \sin\theta \approx (|b - a|)/2 \qquad \textbf{Eqn. 5.3}$$

$$1/r \approx \frac{2\sin\theta}{|b-a|} \qquad \textbf{Eqn. 5.4}$$

The principal curvature is the maximum on the curvature magnitude function about the normal. Since the representation is already quantised, it is possible to simplify identification of the maximum by computing the curvatures only at $\pi/4$ intervals, as shown in figure 5.4:

$$|\kappa_1| = max(|\kappa_\varphi|), \text{ with } \varphi = 0 \text{ to } 2\pi \qquad \textbf{Eqn. 5.5}$$

and κ_2 is found orthogonally to κ_1. If necessary, the direction of maximum curvature could be found by interpolation, as, for example, in the commonly used Canny edge detector [Cann86]. To calculate each planar curvature value convolution masks of length three are used, as discussed above:

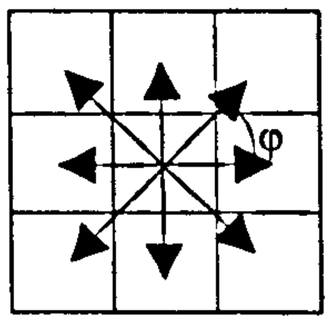


Figure 5.4: Length-3 convolution masks for curvature.

The sign of curvature can be assigned by determining whether p_i falls on the outside or inside of the chord p_{i-2} to p_{i+2} with respect to the centroid (figure 5.3c). The distance of the centroid to the closest point on the chord, and the distance of the centroid to p_i are calculated. If:

centroid-chord > *centroid-*p_i, the sign is positive,

centroid-chord = *centroid-*p_i, the curvature is zero,

centroid-chord < *centroid-*p_i, the sign is negative.

The only undefined case is where the two vectors are parallel, which cannot occur, since the representation precludes such radial re-entrancy.

A drawback of the fixed discrete window is that the vertical density of the vertices is fixed, but the horizontal density varies, so the width of the fixed sample window in terms of the surface contour length also varies. If the window is too large (if the horizontal density is low) this will tend to smooth out areas of mixed curvature. This will also affect the curvature values slightly if

the vectors *a* and *b* in figures 5.3b and 5.3c are of significantly different lengths, and a correction for this could be made by appropriate weighting. Aside from this effect, the method for determining the magnitude of curvature is invariant to the surface orientation. This method thus deals with positive, negative and zero curvatures uniformly, and presents no additional constraints over and above those imposed by the body representation itself.

Each vertex in the mesh can be assigned its signed curvature values, so that the body can be mapped according to its type:

- flat ($|\kappa_1| \approx 0, |\kappa_2| \approx 0$)
- parabolic points: ridge ($|\kappa_1| < 0, |\kappa_2| \approx 0$), valley ($|\kappa_1| > 0, |\kappa_2| \approx 0$)
- elliptic points: pit ($|\kappa_1| > 0, |\kappa_2| > 0$), peak ($|\kappa_1| < 0, |\kappa_2| < 0$)
- hyperbolic points: saddle ($|\kappa_1| < 0, |\kappa_2| > 0$ or $|\kappa_1| > 0, |\kappa_2| < 0$)

Results of this analysis are presented in figure 5.5 for a male and a female torso. Peaks are coloured green, pits are red, and saddle areas are yellow.

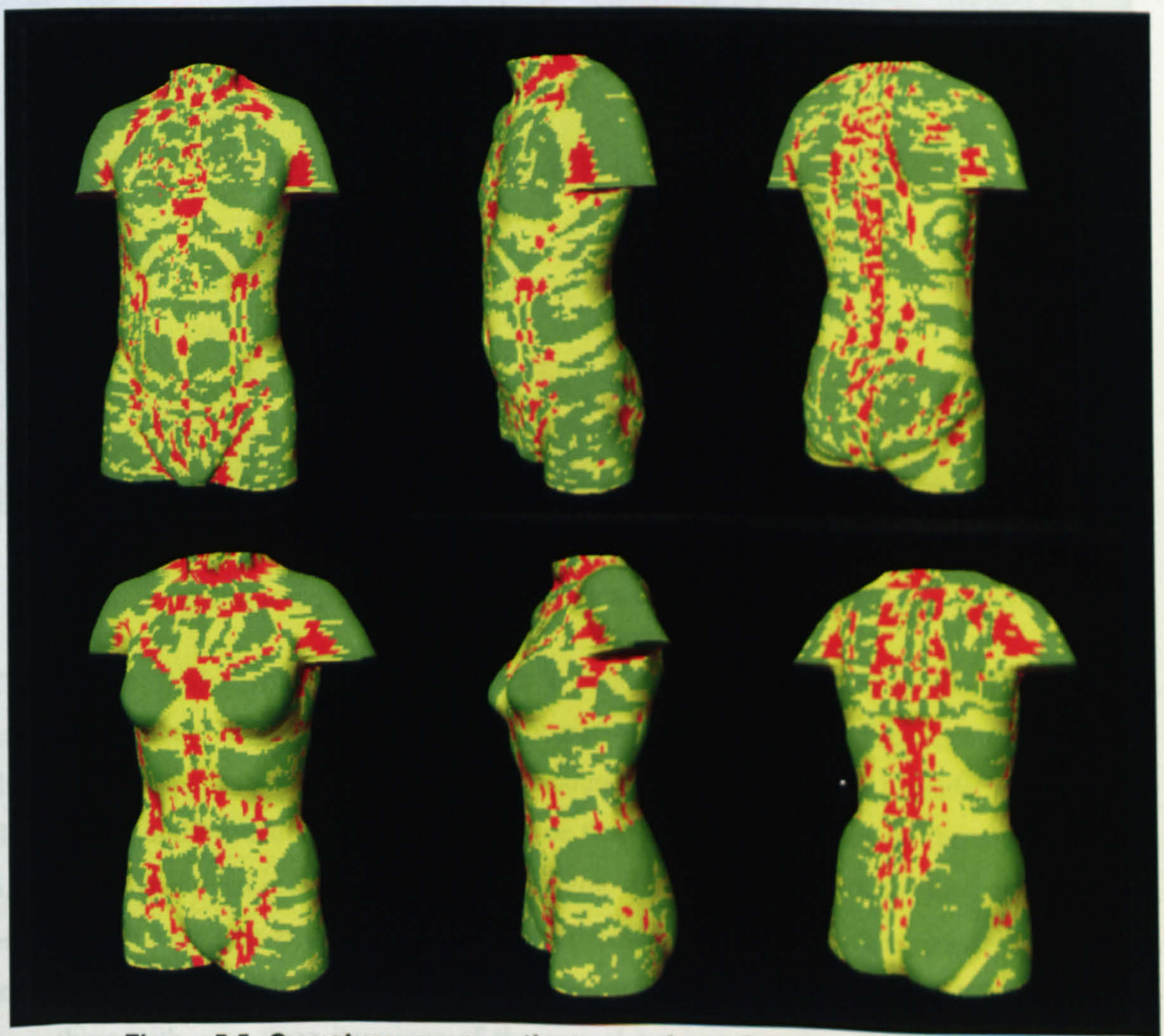


Figure 5.5: Curvature maps on the torso of a male and female subject.
Red = pit, yellow = saddle, green = peak.

Note that "flat" is designated here where $|\kappa| \approx 0$. There will be little that is truly planar on the human body, so a threshold, κ_{flat} is needed to discriminate flat regions. Figure 5.6 shows the resulting curvature maps when κ_{flat} is increased, and more areas are labelled as valley (pink), ridge (light green) and fully flat (white).

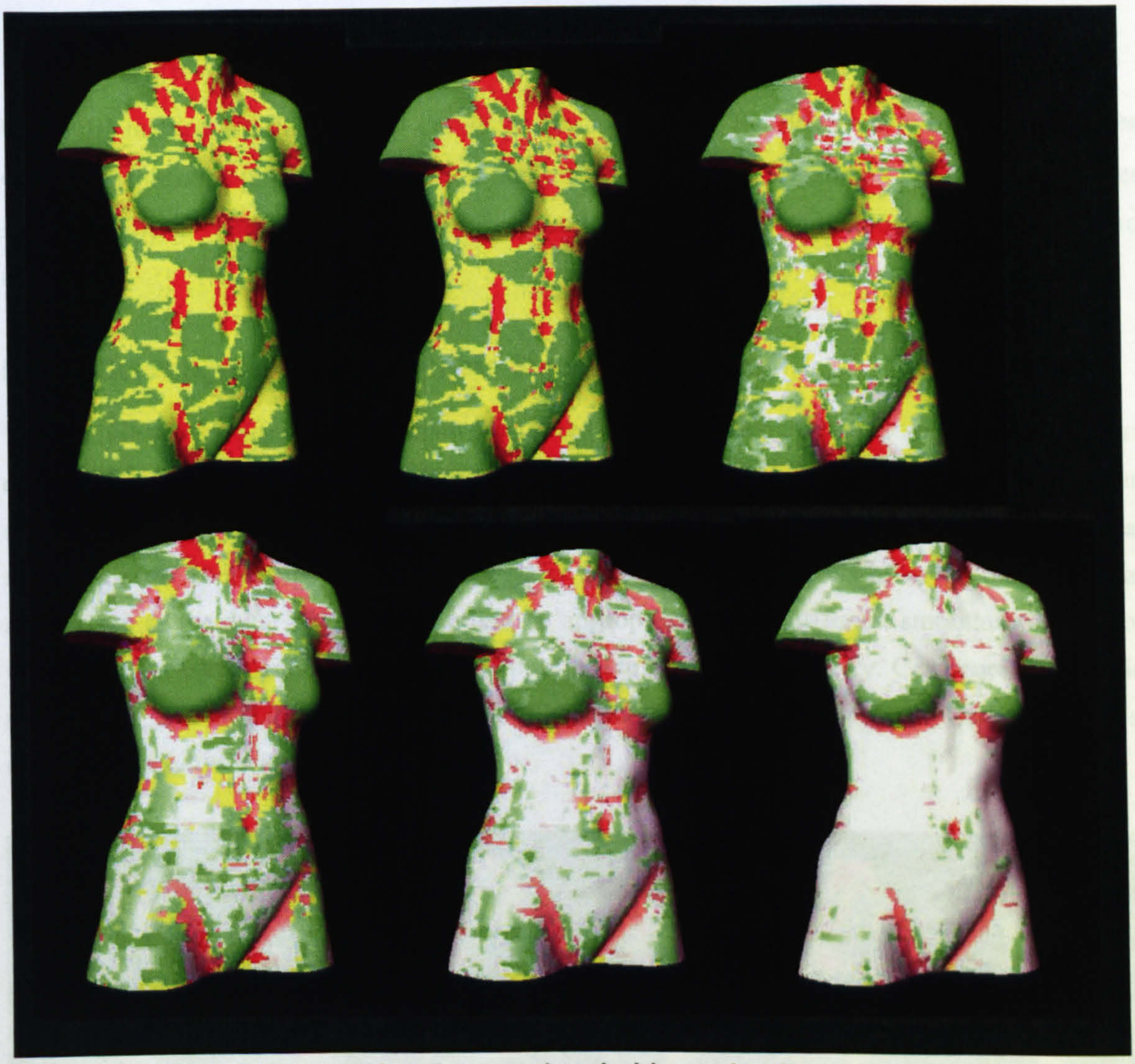


Figure 5.6: Increasing flatness threshold, κ_{flat}, for the curvature map.
κ_{flat} = 0.0, 0.1, 0.2, 0.4, 0.6, 0.8.
Red = pit, yellow = saddle, green = peak, pink = valley, light green = ridge, white = flat.

There is no "objective" value for κ_{flat}, but it relates to the precision of the original depth measurement, any errors introduced by the processes leading up to surface reconstruction, and the particular application of the data. A lower bound can be estimated by considering a perfectly flat surface, and estimating the curvature value that would be observed on this surface (figure 5.7). This would then indicate the lowest κ_{flat} value that should be used, to eliminate the effects arising from sampling errors and subsequent processing. The flat surface is sampled at 5 mm intervals, with uncorrelated errors on depth measurement, ε_1, of approximately ±1 mm. After smoothing with, say, three iterations of the smoothing operator (σ = 1.2), the error on depth measurements, ε_2, is approximately 0.12 mm.

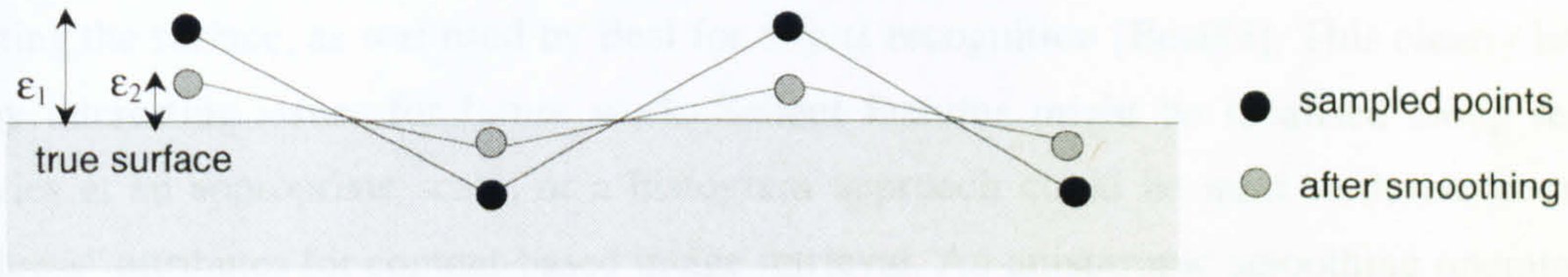


Figure 5.7: Sampling errors, before and after smoothing, on a flat surface.

Using the method for determining curvature described here, the observed normalised κ values would be approximately 0.2, which gives the lower value for κ_{flat}, as shown in figure 5.6c. Other suitable flatness values could be found in a similar way, taking application-specific needs into account.

5.3.3 Scale Space and Appropriate Smoothing Levels

The analyses in the previous sections were carried out on the body with the smoothing operators described in Chapter 4. The operators selectively smooth areas of the body with Gaussian approximations of varying width, which were selected for visual acceptability. It was noted that the choice of smoothing is not absolute, and should be chosen according to the particular application of the data. Therefore, it might be appropriate to use different smoothing levels for volumetric measurement, analysis of curvature or surface anthropometry. Curvature changes in scale space are shown in figure 5.8, starting from the unsmoothed surface (in other words, smoothing window width 1).

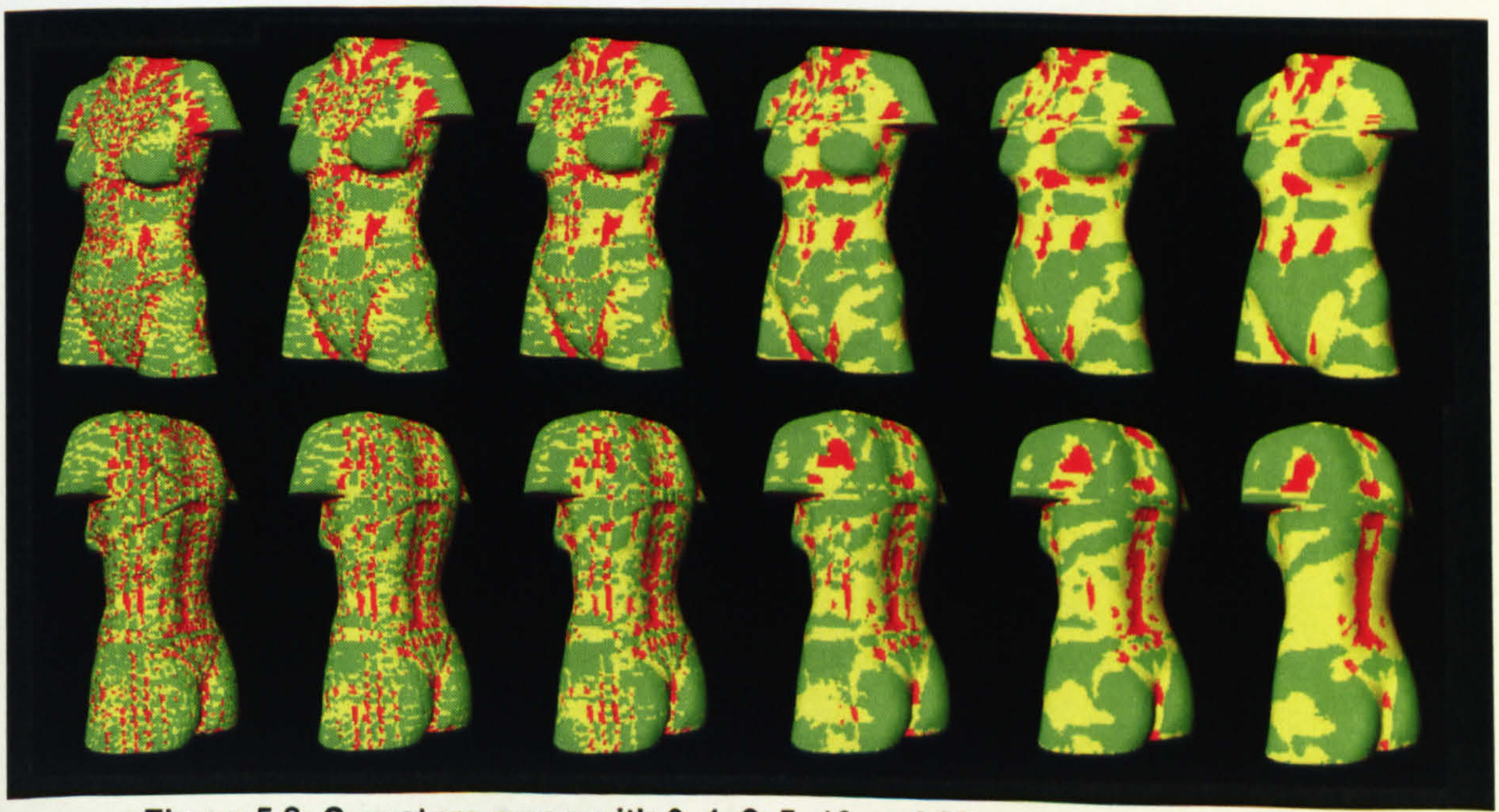


Figure 5.8: Curvature maps with 0, 1, 2, 5, 10 and 20 smoothing iterations.

Higher smoothing values lead to larger areas of similar shape characteristic, which is useful for segmenting the surface, as was used by Besl for object recognition [Besl88]. This clearly brings up many interesting issues for future work. Salient features might be localised along region boundaries at an appropriate scale, or a histogram approach could be used to derive compact "curvedness" attributes for content-based image retrieval. An anisotropic smoothing operator, to preserve fine detail, could be developed based on the curvature values. Such an operator might work by first smoothing the surface as before, and extracting the curvature values at the chosen level of scale. The surface would then be "unsmoothed", and a *directional* smoothing operator would be applied, with smoothing level inversely proportional to the curvature. This would provide a global criterion for selecting the smoothing level at any given point, thus dealing with the issue of appropriate smoothing levels, which was raised in Chapter 4.

5.3.4 Derivative Maps in Fixed Coordinate Systems

Analysis of derivative information in fixed coordinate systems provides a richer vocabulary for describing shape. Since subjects are captured in a common coordinate frame, it is possible to use derivatives in fixed coordinate systems in a common way across the sets of data. Cartesian and cylindrical systems are discussed here. The directional derivatives are of particular interest, as they facilitate description of shape in each coordinate separately, as opposed to the Laplacian, which does not distinguish the directional components.

In designing the partial derivative operators it is necessary to decide whether or not points orthogonal to the direction of curvature should influence the result. The effect of using orthogonal points is to smooth the result. Since the surface has already been smoothed, the convolution masks shown below in figure 5.9 are used:

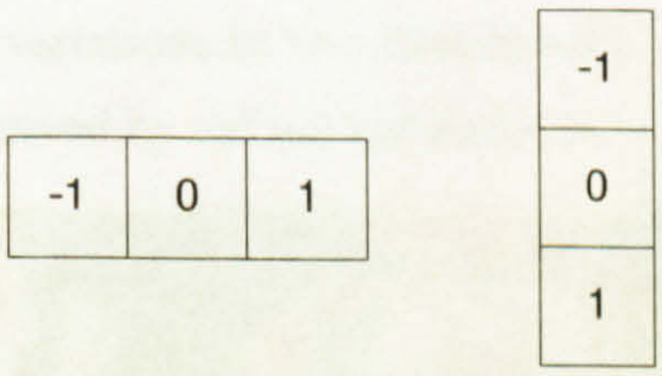


Figure 5.9: Horizontal and vertical convolution masks for the partial derivatives.

Each image is then mapped with:

- turning points in x, y, z and radius (zero crossings of first derivative);
- discontinuities in first derivative of x, y, z and r.

Figure 5.10 shows an example of the turning point maps at increasing smoothing levels.

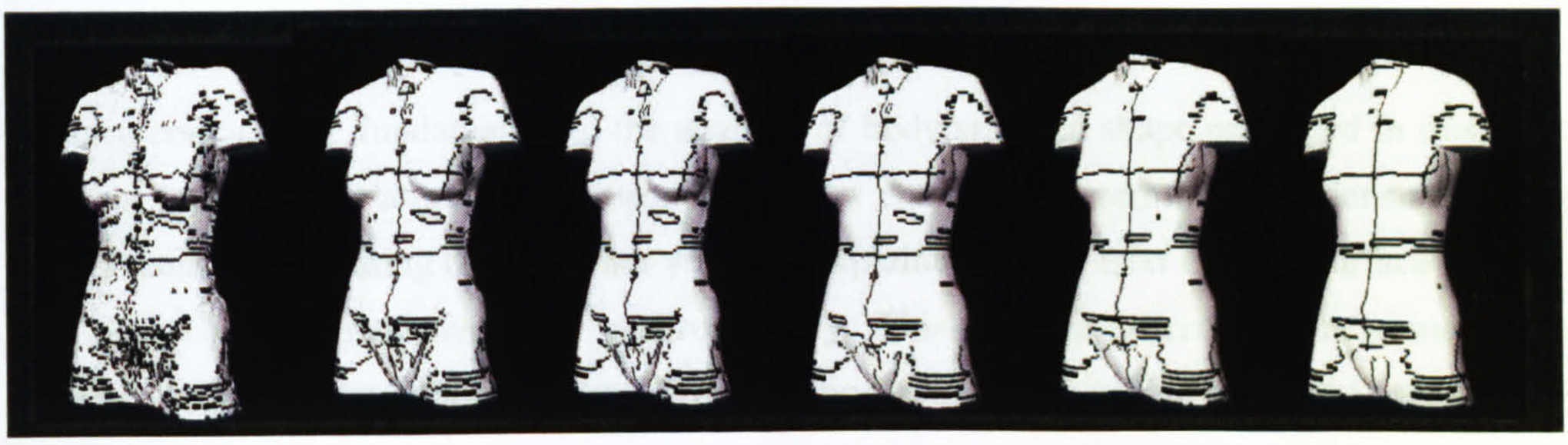


Figure 5.10: Turning point map in *y* and *r*, at smoothing levels 0, 1, 2, 5, 10 & 20.

5.3.5 Anatomical and Anthropometric Significance

The differential geometric properties explored in the previous sections provide useful mechanisms for making explicit certain shape information on the human body. Of particular interest is the potential to use this information in locating features of anthropometric relevance.

Several characteristic features are immediately apparent. The curvature maps clearly show the valleys between the tops of the thighs and the torso, concavities of the navel and large saddle regions around the waist area. The discussion on "curvedness" raised the issue of characterising areas by the magnitude of curvature. With an appropriately chosen threshold value, the strongly curved and the relatively flat areas of the body are visible. The turning point maps show features such as the contour along the top of the trapezius muscle (on the shoulder). The nipples occur at a crossing between two orthogonal local maxima contours.

One important question is which features are consistent across individuals, and how do these shape patterns vary? Some of this will be tested in the next chapter. A more thorough analysis is left for future work. Because the body is represented as a regular two-dimensional matrix structure, one interesting approach would be to register multiple body images and perform statistical analyses on the variations in two dimensions. An example of input data could be the flattened curvature map derived by cylindrical projection, shown in figure 5.11.

Figure 5.11: Flattened curvature map by cylindrical projection.

5.4 Surface and Convex Contours

Surface contours are fundamental to the analysis of body size and shape presented in this work. Contour lengths have already been used in the last chapter to determine data density, but the mechanism for generating them has not yet been explained. In the next chapter surface contours will also be used extensively for anthropometry. This section describes a new method to generate contours from a minimal set of "control points", and determine the convex and surface lengths of the contours.

5.4.1 An "Electronic Tape Measure"

This task can be considered as simulating the behaviour of a tape measure. The "electronic tape measure" tool should be interactive and flexible, enabling a user to place the tape measure over the surface of the body in much the same way as in real life. More importantly, the system must be able to perform the same tasks itself, and so the simulated tape measure must be controllable within the software by a set of high-level instructions.

The main two requirements are to extract the convex and the non-convex lengths of specified contours. The interface requirements are also important if the electronic tape measure is to be useful. It is necessary to determine what essential characteristics must be embodied to obtain the appropriate behaviour, and what type of input from the user is appropriate? At first sight this seems a relatively simple task, but on further investigation several interesting problems arise.

The first decisions relate to how the user will interact with the simulated tape measure. Is it desirable to provide an almost exact replication of the real thing? In practice this is not necessary: for example, the tape need not fall under gravity. However, it is useful to provide some constraining behaviour that is not available in reality, such as ensuring that the tape is parallel with the floor if necessary. For now no attempts will be made to model the deformable behaviour of the real human body, for example, when a taut tape is held against the skin.

In order to design the way that the electronic tape measure should work, consider how the tape is used in the real case, by holding the two ends and guiding it into place. It is then possible to define the different contours by the minimal number of control points that are required to perform each type of task, as in table 5.1.

Table 5.1: Minimal control point sets for an electronic tape measure.

Contour type	Control points	Example
closed curve, parallel to floor	1	waist, hips, bust, underbust
closed curve, in arbitrary plane	2 or 3	torso circumference via shoulder and crotch
open curve	2 or 3	shoulder to nipple, 7^{th} cervical vertebra to waist
multi-segment curve	3 or more	7^{th} cervical vertebra to shoulder to elbow to wrist
straight line segment	2	hip width, abdominal seat diameter, armscye width

By defining the contour types in this way, a simple user interface suggests itself, whereby the user (or internal instruction code) places landmarks on the body surface, and specifies what type of contour should be generated, using a selected set of one or more landmarks as control points. The same processes can be used by the system when carrying out these tasks in an automated way.

The next problem is to model the behaviour of the simulated tape measure between these control points. As well as geometric accuracy, an important issue is computational efficiency, since the system must be able to generate a hundred or more specified contours within a minute or so. Because the real tape has a non-infinitesimal width and lies flat on the skin surface, without crumpling or twisting, a reasonable approximation can be obtained by assuming that the contour lies in a plane. In fact many anthropometric contours are specifically defined as lying in a plane [Kuni84] [ISO89]. The multi-segment case can be considered as a concatenation of several such planar curve segments, but the overall contour may occupy $\Re^3$. This assumption of planarity should simplify much of the processing and reduce the computational cost.

There is now enough information to make a first attempt at designing a general strategy for generating contours. Having dismissed the need for curve *fitting* techniques earlier, in section 4.2, this work focuses on methods for piecewise linear contour *growing*.

5.4.2 Seed-and-Propagate for Closed Curves

The simplest case is a closed contour. If a single control point is supplied, it is assumed that the curve lies in a plane parallel with the floor. Three control points describe an arbitrary plane. The case of under-defined planes, with two control points, is discussed later.

In order to "grow" the curve, it is necessary to constrain it to vertices that are likely to be on the contour. In the case of the contour in the horizontal plane this is trivial, since the body representation is already structured in this way. Any contour of this type can be extracted directly as the row of data at a given y value. For the general case (not in the horizontal plane) a plane is passed through the control points, and vertices close to the plane are sampled. The definition of "close" is determined by the choice of tape measure width, and relates directly to the real case. In practice, this should be between one and two times the vertical density of the data, in order to ensure the sufficient points are sampled to generate a complete closed contour. Only those vertices on the same body segments as the control points are sampled. This allows, for example, the right thigh girth to be generated without vertices from the left side interfering, even if the points are close. The sampled points are then projected, so that they lie exactly on the plane.

Where the control points describe an arbitrary plane, it is also necessary to determine the sequence of the points that reflects the original connectivity on the surface. Starting with one of the control points as a seed, the nearest neighbouring vertex is computed, and the contour grows

to that point. This point is then removed from the search. This process continues until no more vertices are left, and the contour closes to the seed point. Closure is ensured, as each step is guaranteed to remove a point from the initial sampling list, and closure is obtained when no more points are left. This takes $O(n^2)$ time, where n is the number of sampled points in the plane.

The surface length can then be computed by summing the piecewise linear distances between points. It is necessary to smooth the resulting contour first, using a one-dimensional equivalent of the surface smoothing operator.

Some characteristics of the electronic tape measure are implicit in the description of these processes. It is infinitesimally thin (in its smallest dimension); it can be assigned any particular width (say 10 mm), corresponding to the sampling width of the plane; it can, of course, be of any length.

The method just described works well in almost all cases. However, in unusual cases this nearest neighbour contour growing method gives rise to "meandering" if several sampled points lie close to each other. This can largely be avoided by cleaning out the unused vertices in the neighbourhood behind the propagation path, so that the contour cannot double back on itself as it grows. This happens where the sampling plane lies close to the surface tangent at any point. For example, when the three control points are approximately collinear, a slight change in the position of one of the control points has a drastic effect on the orientation of the control-point plane, giving rise to chaotic behaviour. (Increasing the effective thickness of the sampling plane exacerbates this effect.) To avoid this, a test is introduced so that, if the initial sampling plane lies too close to the tangent, the control points can be adjusted by a small amount to ensure that the sampling plane is closer to the surface normals at the control points.

5.4.3 Closed Convex Hull Generation

Once the closed surface contour has been generated, its convex hull can be found. There are standard algorithms for this [Sedg90]. In this work, a simple, computationally undemanding method has been implemented.

For contours in an arbitrary plane, the coordinate system is first transformed to a reference plane, xz. The bounding box is then computed. Choosing any of the points on the bounding box, shown in figure 5.12, it is placed as the starting point in the list of convex hull points. For example, starting at the point with the lowest x value, min_x, the angle is calculated between that point, the next point in the list, and the line $x = min\_x$. This is repeated for all points. The point with the smallest angle can be guaranteed to lie on the convex hull, as shown in figure 5.13, so it is copied into the convex hull list, and designated as the new current point.

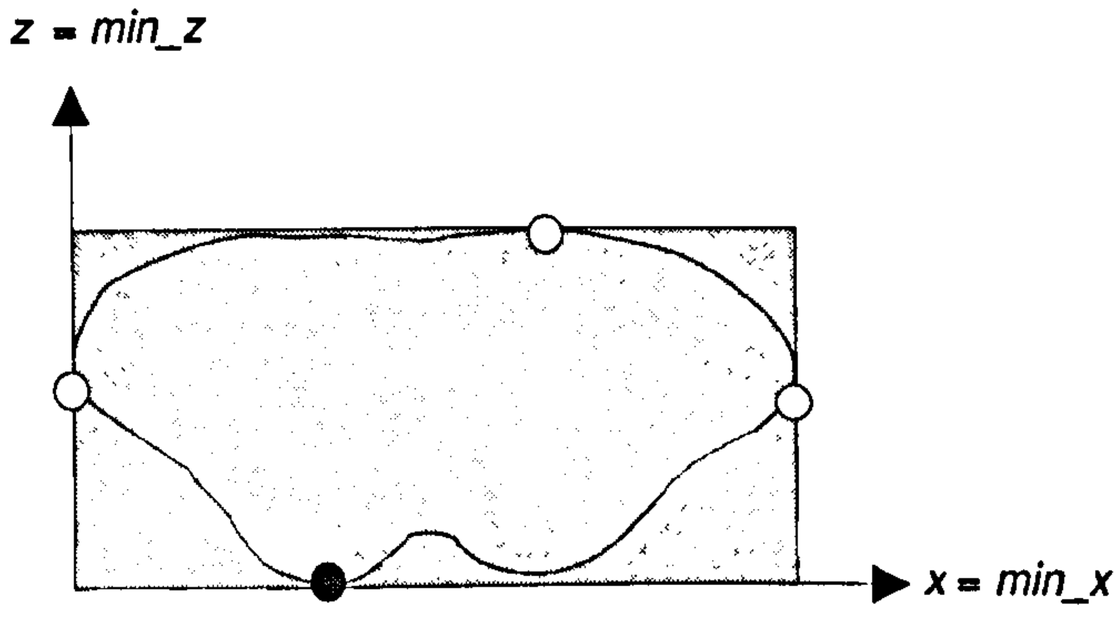


Figure 5.12: Initial bounding box for the convex hull.

All points on the non-convex contour between the last point and the current point, inclusive, are removed from the candidate list (the previous non-convex ordering of points makes this possible). Processing continues in the same way from the new current point. Closure is obtained when no more points are left in the candidate list.

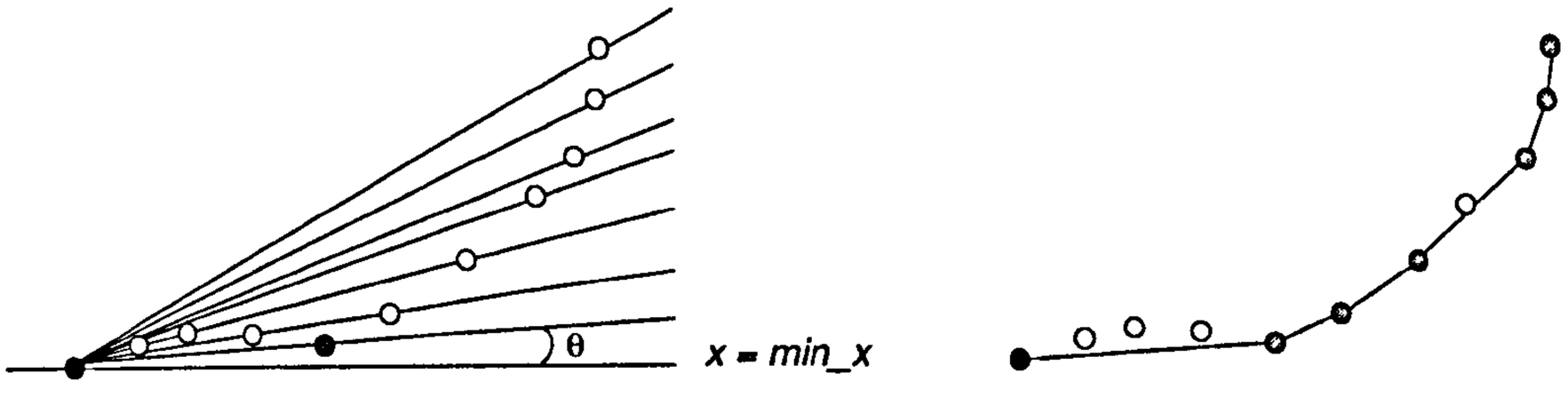


Figure 5.13: Current-to-candidate angle for the convex hull between two bounding-box points.

The convex length is computed by summing the piecewise linear distances between the sequenced convex hull points. The complexity is at worst $O(n^2)$, but usually substantially less than this, reducing as the convexity decreases.

The implications of a piecewise linear approximation to the surface contour were discussed in Chapter 4, and it was determined that, at this sampling density, a negligible underestimate was introduced as a result. The potential problem of a counter-effect (possible overestimate) from noise was also discussed, but it is argued that the logical and mean smoothing operations carried out in Chapter 4 will minimise this effect.

5.4.4 Open Curves

If an open curve is required, several conditions must be considered, as illustrated in figure 5.14, which shows a planar section through the body. If three control points are specified, they describe the open curve unambiguously. Consider the two cases in figure 5.14 (a) and (b). For each contour two endpoints are specified. The third point (the midpoint) identifies the plane through which the contour goes, and which segment of the whole closed contour is required.

The resulting non-convex contour runs from one endpoint, via the midpoint, and terminates at the second endpoint. There are four main cases to consider, where the endpoint chord has the following behaviour:

- it intersects the surface only at the two endpoints: (a) and (b);
- it is tangential at the two endpoints (c);
- it intersects the surface at the two endpoints and 2n other points: (d);
- it intersects the surface at the one endpoint, is tangential at the other, and intersects at 2n+1 other points: (e) and (f).

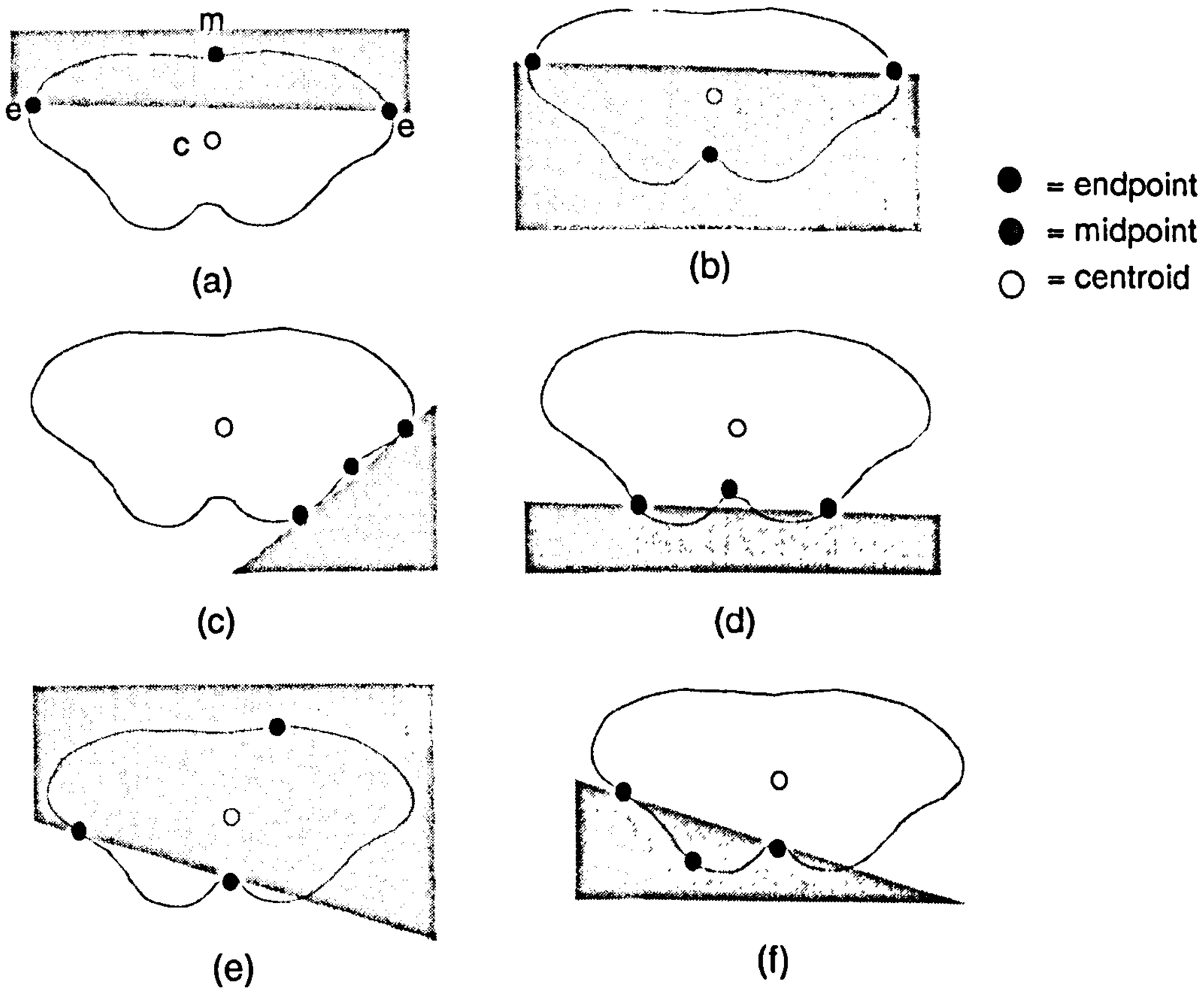


Figure 5.14: Generating open contours from three control points.
Shaded area shows the data selected by the endpoint chord partition.

It must be shown that these cases can be dealt with satisfactorily. At first sight it is appealing to generate the open contour by simply partitioning the data points on either side of the endpoint chord, and then generating the non-convex and convex hulls from those data. On further inspection, this only works in a few cases, for example (a) to (c), but would fail for (d) to (f). In (d) part of the non-convex hull would be truncated, but the convex hull method would be successful; in (e) part of both the non-convex and the convex hulls would be truncated; in (f) an extra segment would be added to the contour.

The open curve generation algorithm must take into account the solidity of the object on which it operates, without "cutting through" the surface, except at the endpoints. It must therefore generate the *full* closed contour first, and then select the appropriate segment from it. The convex hull can then be generated from the open curve in a similar way to the closed curve, as shown in figure 5.15. It is important to note that, in most cases, the convex hull of the open curve is *not* part of the convex hull of the parent closed curve.

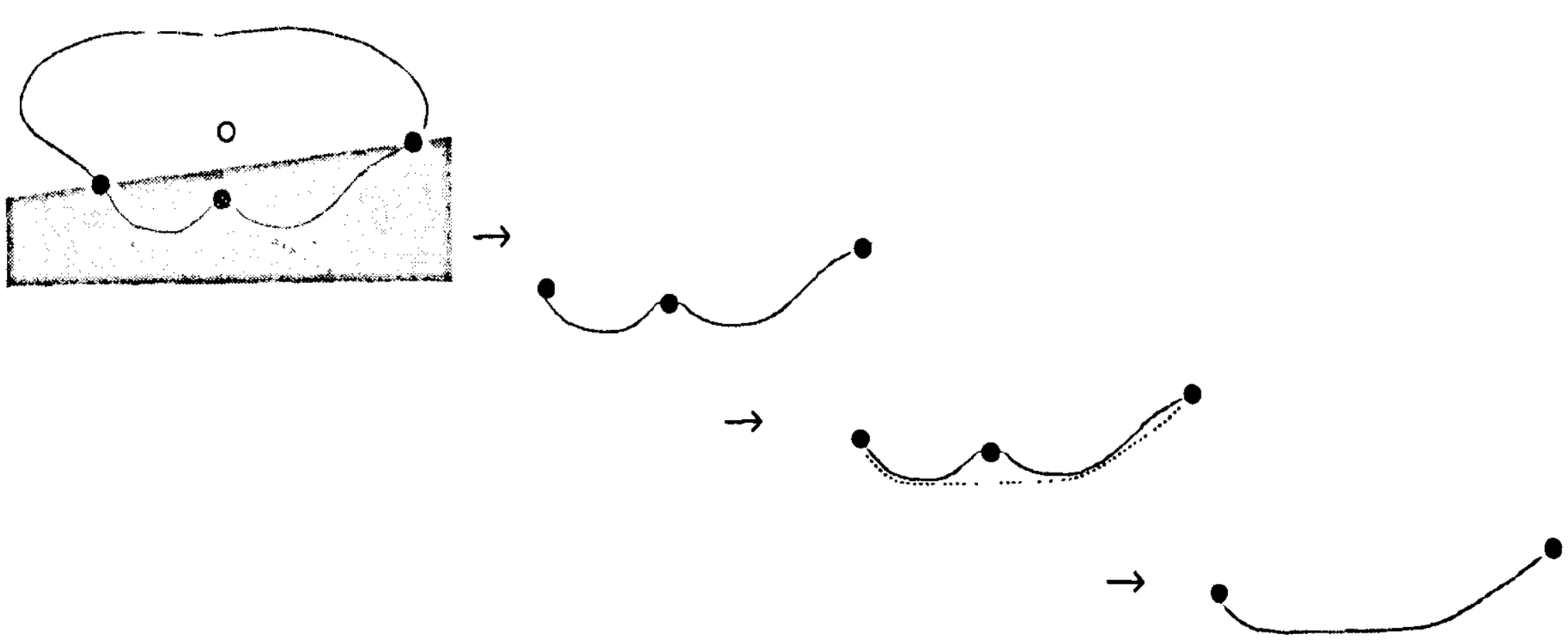


Figure 5.15: Generating a non-convex curve and its convex hull.

5.4.5 Inference of Under-Defined Curves

In many cases, especially for open curves, it is more convenient for the user to provide only two endpoints, and for the plane of the curve to be inferred from the behaviour of a real tape measure held between the corresponding points on the subject's body. Since the tape would be taut, it would be expected to follow a shortest path, with the additional constraint that the path is approximately planar. Because of the planarity constraint it is not possible to use the usual shortest path algorithms [Sedg90]. It is necessary therefore to find a computationally efficient way of simulating this.

If it is assumed that the y value of the third point can be chosen as the mean of those of the two endpoints, the problem then is to determine where on the horizontal plane the third point should be located. It would be possible to place the third point at each location on the horizontal slice, and then choose the shortest path. This would be computationally expensive, and it is possible to reduce the search space by selecting "likely" paths.

One candidate for the optimal position is to compute the point halfway along the chord between the two endpoints and then locate the nearest surface point to the interpolant in Euclidean space. Alternative distance metrics are to take a radial line out from the centroid, through the chord bisector, and to place the third point where the radial line intersects the surface, or to determine it by projecting the chord bisector in x onto the nearest surface. In practice, the appropriate path

lies somewhere between these, and the shortest path can be found by sampling between them. Figure 5.16 shows the two endpoints and the three contours generated by inferring the third point by Euclidean, projection and radial interpolation. In this case the shortest path lies on the middle contour. 5.16b shows a cross section through the body at the level of the third point.

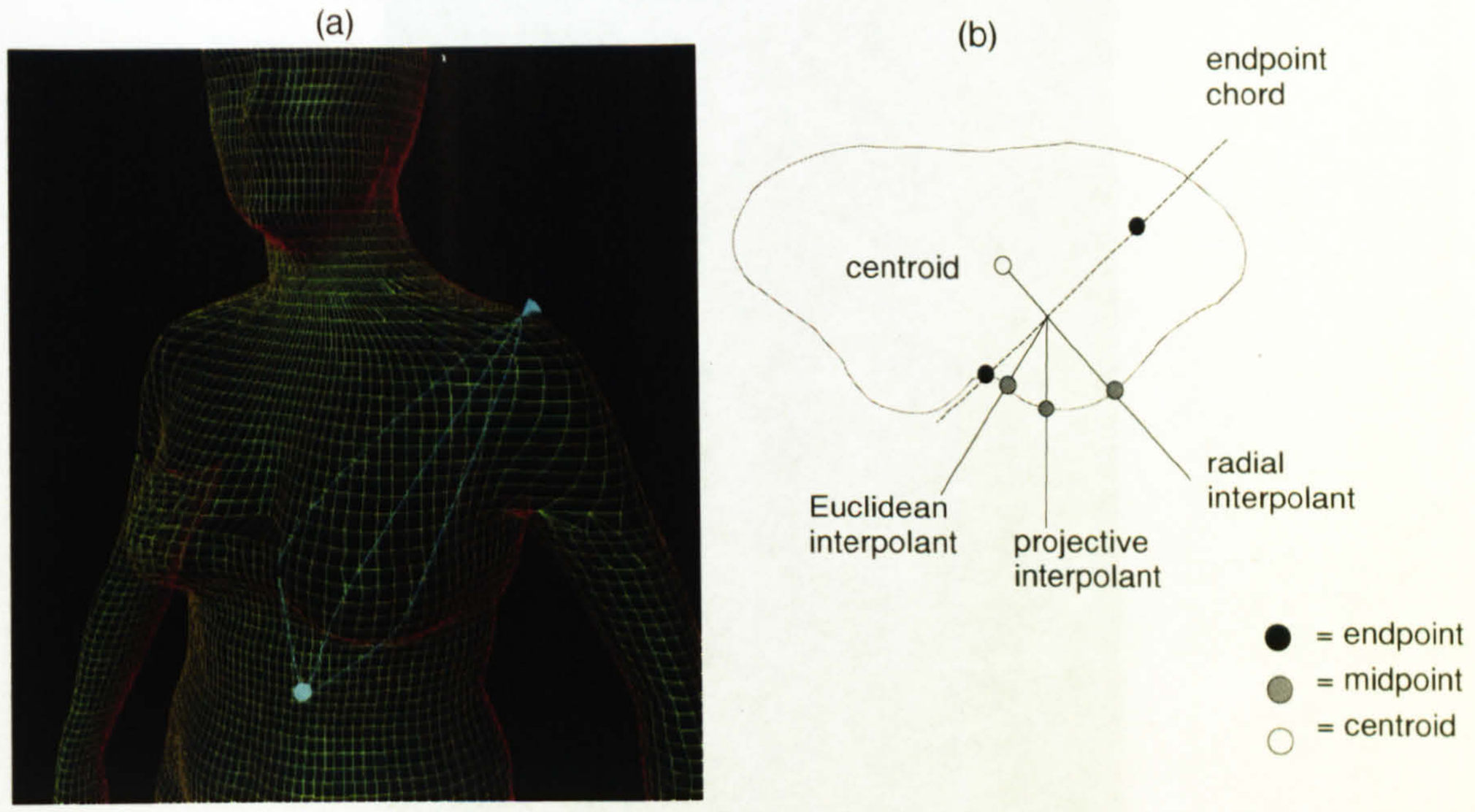


Figure 5.16: Shortest planar path by mid-point interpolation.

This two-point open curve algorithm is used as a basis for the multi-segment curve, whereby three or more points can be used to define a compound curve, which may lie out of a single plane. Each pair of points in sequence is used to generate a planar two-point open curve, as described above, and the multi-segment curve is composed by concatenating the segments.

5.4.6 Summary of the Electronic Tape Measure

The preceding discussion has explained the rationale behind the development of the electronic tape measure. Examples of each measurement type are shown in figure 5.17 below. It is worthwhile now to summarise the algorithms:

1. Use one or more landmarks to define the plane and body part(s) in which the contour lies.
2. Use the plane defined by the control points to select candidate points close to the plane, within a given "tape measure width". Ensure that only those points on the same body parts as the control points are used.
3. Generate the non-convex closed curve from these points.
4. (a) For closed curves, generate the closed convex hull.

   (b) For open curves, extract the appropriate segment of the whole non-convex closed curve, then generate the open convex curve.

(c) For multi-segment curves, generate each open segment, then concatenate them to make a single curve.

5. Compute the convex and non-convex lengths by summing piecewise linear values.

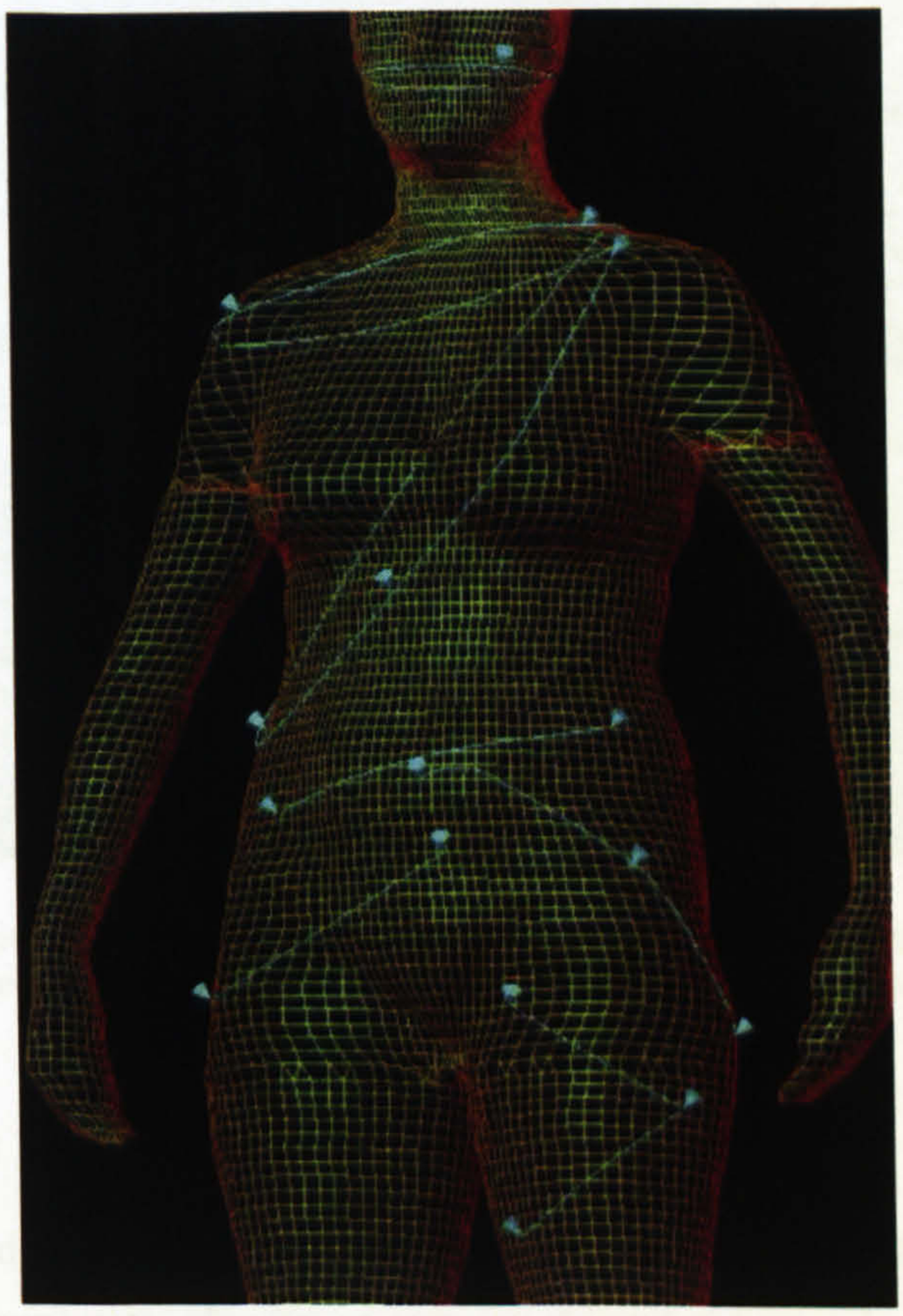


Figure 5.17: Electronic tape measure contour types.
From top to bottom: 1-point closed, 2-point closed, 3-point closed, 2-point open, 3-point open, straight line, multi-segment.

5.4.7 Generalised Cylinder Attribute Space

The body is represented as a set of discrete generalised cylinders, with arbitrary closed contours defined at regular intervals and parallel with each other. It is useful to consider attributes of the closed contours as a function of length along the axis (which is a space curve). This reveals interesting shape attributes. In particular, many of the anthropometric definitions are described in this way, such as convex hull lengths measured parallel with the floor. Only the convex hull function is explored here, but other attributes, such as surface length, convex deficiency (*convex area – non-convex area*), etc., could be used.

Derivatives on the convex hull lengths are computed in the usual way, and provide another set of useful attributes for shape analysis, as shown below in figure 5.18. In particular, turning points, such as maximum calf girth, are of interest for anthropometry. Discontinuities in the first derivative of the convex hull function relate to features such as the neck base and turning points on the shoulders.

Figure 5.18: Turning points and discontinuities in 1st derivatives of convex hulls.
(a) low-resolution mesh, (b) surface contours, (c) convex hulls.

5.5 Validation

In order to test the validity of this contour reconstruction method, tests have been carried out on human subjects and inanimate objects of varying geometric complexity. These experiments describe validation of the *convex* contour lengths. Validation of the *surface* contour lengths is problematic, as it is difficult to measure these quantities manually, therefore this is not addressed here. Numerous tests were also carried out to assess the results visually while the contour generation methods were under development. This section only shows results in detail for the final set of validation tests.

5.5.1 Convex Hull Validation Procedure

Comparison of manual and electronically derived measurements has been discussed by [BrJW94] [Daan98] and [BrGr99], where the electronic measurements are taken from markers placed manually on the scan image. A similar procedure is used here:

- control points are placed manually on the reconstructed scans;
- open and closed contours are generated from those control points;
- the computed lengths of the *convex* contours are compared against manual measurements taken directly on the object/subject.

Three types of objects are used for the tests:

- reference cylinder: green plastic, girth approximately 1000 mm,
- mannequin: papier maché covered, torso only,
- female human subject, in underwear.

Inanimate objects are used to test the algorithms to eliminate additional errors from movement and breathing. However, it is noted here that the surface properties of the inanimate objects are not representative of a human, and therefore the results need to be considered in this light.

5.5.2 Discrepancy Measures

Discrepancy measures between the manual measurements and the scan-derived measurements are described below, where:

$x_{t1}[i]$ = measurement of dimension x, on sample i, using technique $t1$,

m = manual measurement,

s = scan-derived measurement.

Mean absolute difference (MAD) has been used by Gordon *et al* in the context of inter- and intra-observer variation in manual anthropometric surveys [GBCC89]. Bradtmiller and Gross also present results for scan-derived measurements and acceptability criteria in terms of MAD [BrGr99]. To enable comparison against these values, MAD will also be quoted here:

$$MAD(x,t1,t2) = \frac{1}{n}\sum_{i=1}^{n} | x_{t1}[i] - x_{t2}[i] | \qquad \text{Eqn. 5.6}$$

The *signed* mean is used to detect systematic differences between the scan-derived and manual values:

$$\mu(x,t1,t2) = \frac{1}{n}\sum_{i=1}^{n} (x_{t1}[i] - x_{t2}[i]) \qquad \text{Eqn. 5.7}$$

Root mean squared (RMS) error emphasises the larger errors:

$$RMS(x,t1,t2) = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (| x_{t1}[i] - x_{t1}[i] |)^2} \qquad \text{Eqn. 5.8}$$

Standard deviation, σ, on the signed difference values:

$$\sigma^2(x,t1,t2) = \sum_{i=1}^{n} \frac{1}{n} (\mu(x,t1,t2) - (| x_{t1}[i] - x_{t2}[i] |))^2 \qquad \text{Eqn. 5.9}$$

To determine the *percentage* error with respect to the absolute contour length, a mean percentage difference value (M%D) is also computed:

$$M\%D(x,t1,t2)=\frac{100}{n}\left(\sum_{i=1}^{n}\frac{2|x_{t1}[i]-x_{t2}[i]|}{(x_{t1}[i]+x_{t2}[i])}\right) \qquad \text{Eqn. 5.10}$$

5.5.3 Results

Results are presented in table 5.2, showing RMS error, mean absolute difference (MAD), standard deviation on absolute difference, bias and mean percentage difference (M%D) for each type of contour on the three sample objects.

Table 5.2: Comparison of manual and electronic tape measurements.

Contour type	no. samples	MAD (mm)	μ (mm)	RMS (mm)	σ (mm)	M%D
closed curve parallel to floor	14	4.3	0.3	5.4	3.3	0.65
closed curve in arbitrary plane 2 control points	4*	11.0*	-0.5*	11.4*	3.6*	1.77*
closed curve in arbitrary plane 3 control points	5	8.8	7	11.9	6.9	2.20
open curve in arbitrary plane 2 control points	9	4.7	-1.3	6.0	4.1	1.00
open curve in arbitrary plane 3 control points	5	6.0	0.8	6.8	4.1	0.97
straight line between 2 control points	5	5.8	-1.2	5.6	3.3	0.95

*One sample with >100% error removed (816 mm manual measurements, versus 1941 mm electronic).

5.5.4 Discussion

As discussed by other researchers, the comparison between live measurements on a human subject and an electronic image is fraught with many difficulties, and this is discussed more fully in the next chapter [BrJW94] [Daan98] [BrGr99]. First of all, it is interesting to note that the MAD values are similar to those of inter-observer errors on manual measurements, as discussed by Gordon *et al* [GBCC89], and discussed again later in Chapter 6. The closest correlation between the manual and scan-derived measurements is on closed contours in the horizontal plane. If the data for the reference cylinder and phantom are considered separately, the results, as expected, are significantly better, with RMS error at 3.8 mm and M%D at 0.36%.

Errors in contour measurements may come from several sources. Some of the limitations of the surface representation were discussed in the last chapter. Because the body representation is discrete, there is a small error introduced on landmark placement. Landmarks "jump" to the nearest mesh vertex, therefore the mesh determines the precision of the landmark placement. As discussed previously, it would be expected that uncorrelated noise will tend to cause a slight overestimate in the convex length, although it is argued that the previous surface smoothing should reduce such effects.

The strongest bias appears on closed curves with three control points, on open curves with two control points, and on straight line measurements. Because of the small sample size, it is not immediately apparent if these are true biases, and if so, why they occur. Certain constraints built

into the electronic tape measure may introduce some discrepancy, such as the assumption of curve planarity. In the straight line case, it is possible that a large part of the error comes from the manual measurements.

Plane inference for under-defined contours with two control points appeared to give good results in almost all cases. However, occasionally very large errors are generated, as in one of the test cases, which has been removed from the results table. It is worthwhile noting, therefore, that the errors, when they happen, tend to be very large, and it might be possible to devise an automated way of detecting such conditions.

The number of samples in the final validation tests is relatively small, and therefore no attempt has been made to determine the statistical significance of the results in terms of comparing the accuracy and reliability of the different types of measurement.

5.6 Feature Detection Library

The investigations in this chapter have revealed a rich source of information for characterising shape on the human body. In order to use these techniques for detecting pre-defined features, it is necessary to compose them into a library of operators that can be used to build specific procedures. There are four main issues here, which provide various influences on the choice of such operators:

- compactness of expression: the library should contain high-level operators that perform useful generic jobs (such as generating and extracting attributes from complex contours);
- flexibility of expression: there should be low-level operators that perform simple tasks (such as point translation) and give access to simple attributes (such as the surface normal at a given point);
- sufficiency and closure: the library should be able to express all appropriate feature detection procedures;
- efficiency: no redundant operators should be present in the library.

The library of operators are listed below in table 5.3 below. In the next chapter these operators will be used to build specific feature detection procedures. The design of the library is open, to allow other operators to be added if necessary (and existing ones to be improved).

Table 5.3: Operator library.

Basic attributes
Cartesian/cylindrical coordinates of an indexed point
Centroid of a horizontal slice for a given body segment
Body segment of a given point on the surface
Point translation by vector(*x,y,z*)
Nearest point, using nearest radial, projective and Euclidean distance metrics
Derivatives, etc.
Surface normal at a given point
Surface derivatives at a given point
Principal curvatures
Shape index
Turning points and discontinuities in first derivative in Cartesian and cylindrical coordinates
Turning points and discontinuities in first derivative of convex hull lengths
Space curve generation and attributes
Open & closed curves, in horizontal plane
Open & closed curves, arbitrary plane
Convex & surface lengths
Multi-segment open curve
Straight line segments
Search constraints
Constrain search relative to landmarks
Constrain search by body segment
Constrain search left/right, front/back with respect to centroid

5.7 Summary

Methods have been explored to transform the surface of the body to extract local shape attributes and measure contours. A library of operators has been compiled that provides access to low-level information, such as individual vertices on the polygon mesh, the central axes of the body, and higher-level features related to distinct curvature types. In addition, an "electronic tape measure" has been developed—a system of landmark control points and surface contour generation functions, which allows body metrics to be extracted, effectively performing the tasks of traditional anthropometric tools. In the next chapter these operators will be put to use, to build and test specific feature detectors for anthropometry.

Chapter 6

Application and Evaluation

The operators explored in the previous chapter provide a basis with which to develop feature detectors for anthropometric mapping of the body. Training and validation sets of scanned and manually measured bodies are used to verify the accuracy of the computed values. A number of important issues arise in generating and testing computed measurements because of the many potential sources of error, in obtaining test data and the validation measurements, as well as in defining appropriate acceptability criteria.

6.1 Experimental Design

In the previous chapter a library of shape analysis operators was developed, and some of the implications for anthropometry were briefly discussed. In this chapter these operators will be used to build detectors for a specific anthropometric problem. The overall plan is to:

- collect training data;
- build feature detection procedures, using the training data to evaluate and refine the operators;
- validate the output with an independent set of data.

6.1.1 Target Anthropometric Problem

The anthropometric standard ISO 8559 [ISO89] is used for validation, because of its international relevance and the clarity of description. Several other dimensions have been added because of their widespread use in garment design. The full set is presented in *Appendix A: Anthropometric Specifications*.

The target group for these tests is women of medium build. As well as being most economically compelling, this group contains the UK industry standard size 12 from which other sizes are

"graded" up or down to produce garment patterns. This work therefore does not deal with all body types, but aims to demonstrate the automated techniques on a more limited subset, with the longer-term goal of wider application.

6.1.2 Validating Output

The problem of validating the system's output was raised in Chapter 2. Some experiments were carried out to place optical markers on subjects before scanning, in order to detect them in the image. This method was used by [BrJW94] [Li97] [Daan98] [BrGr99], who located the pre-placed markers manually on the processed images. Geisen *et al* [GMHW95] and Lewark and Nurre [LeNu98] devised an automated technique to detect the markers. However, this method is inappropriate for the imaging technique used here, because of the resolution of the data being captured. In addition, this approach is inconsistent with the goal of minimum intervention at run time, outlined in Chapter 1.

An alternative might be to use an electronic 3D position sensor [Polh98] to mark the location of features on the live subject and collect 3D position measurements. Tests were carried out to assess the accuracy and reproducibility of this method, by marking the ISO 8559 points on two subjects [ISO89]. The vertical measurements of the landmarks were found to be fairly accurate, but the locations in the other axes could not be reproducibly detected, because of sway during the elapsed time between taking readings for each point. Given these observations, an anthropometer or measurement stand would be adequate to collect the same data.

Another option is to place landmarks manually on each image for the training set, without marking subjects prior to scanning. However, this is potentially as time consuming as the traditional methods. There is also the drawback that the surface images lack texture mapping, or the tactile feedback used to detect some skeletal features, and it is not possible to adjust the posture of the image to facilitate location of landmarks. However, this method has advantages over using a 3D position sensor, because the image is static when the landmarks are located. The method could therefore provide fairly accurate landmark locations.

In order to validate the *final* outputs of the system, it is necessary to collect data on the linear and curvilinear measurements, using traditional measurement tools (tape measure, callipers, etc.). This also provides reference information indirectly on the landmark locations: if the electronically derived measurements match the manual measurements, then the landmarks can be assumed to be correctly located, and *vice versa*. It is possible that good agreement might be found in the measurements, using incorrectly located landmarks, but, given the number of measurements taken and the dependencies between them, this is very unlikely. On balance, it appears that this is an appropriate method for collecting reference data, and this is the one that is used here.

6.1.3 Implications of the Reference Data

The use of manual measurements of surface contours as reference data has certain implications in terms of training and validation. Firstly, it is generally accepted that error amongst even skilled measurers can be significant [GBCC89] [BrJW94] [BrGr99], and therefore the automated system will be trained and validated against data that is expected to contain some error. Each dimension has differing factors that determine the expected accuracy of manual and scan-derived measurements. Some dimensions may be relatively insensitive to these differences, but others may be very sensitive. It would be expected, for example, that the detection of maxima or minima would be highly accurate on the smoothed scan data, but might be prone to subjective error when measured manually (e.g., abdominal seat diameter).

Until realistic, real-time models of tissue and fabric deformation can be devised, electronic 3D images suffer from the drawback of having no tactile feedback. Manual measurements may differ from electronically derived ones because of the use of slight pressure in manual measurement, which pulls in loose fabric, and has differing effects on soft and firm body tissue [BrJW94] [Daan98] [BrGr99]. This will tend to produce smaller values for manual data versus electronic measurements. Because of this slight ambiguity, it is important to perform tests to attempt to isolate these discrepancies. As well as using numeric performance measures, the output will be assessed visually, to detect obvious errors in the electronic system.

Other problems arise if only a single image is used to extract measurements. *In vivo* measurement using such a standard as ISO 8559 requires the subject to adopt various postures appropriate for measuring different dimensions [BrGr99]. The choice of posture is very relevant here, since one of the stated objectives is to extract as much information as possible from a single scan.

Clearly any "error" in the automated measurements is not necessarily *actual* error, but a discrepancy, as acknowledged by Gordon *et al* when they compared inter-observer measurements [GBCC89]. However, for simplicity, the word "error" will be used here, to indicate the discrepancy between the two methods.

6.1.4 Performance Measures

Discrepancy measures between manual and scan-derived measurements were introduced in section 5.5.2, when the electronic tape measure was validated. The same set will be used here. Recall that:

$x_{t1}[i]$ = measurement of dimension x, on sample i, using technique $t1$,

m = manual measurement,

s = scan-derived measurement.

$$\text{range}(x) = \max(x_m) - \min(x_m) \qquad \text{Eqn. 6.1}$$

$$\%\text{variation}(x) = 100 \times \text{range}(x) / \bar{x}_m \qquad \text{Eqn. 6.2}$$

$$\mu(x,t1,t2) = \frac{1}{n}\sum_{i=1}^{n}\left(x_{t1}[i] - x_{t2}[i]\right) \qquad \text{Eqn. 6.3}$$

$$MAD(x,t1,t2) = \frac{1}{n}\sum_{i=1}^{n}\left| x_{t1}[i] - x_{t2}[i]\right| \qquad \text{Eqn. 6.4}$$

$$RMS(x,t1,t2) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\left| x_{t1}[i] - x_{t1}[i]\right|\right)^2} \qquad \text{Eqn. 6.5}$$

$$worst(x,t1,t2) = \max\left(\left| x_{t1}[i] - x_{t2}[i]\right|\right) \qquad \text{Eqn. 6.6}$$

$$\sigma^2(x,t1,t2) = \sum_{i=1}^{n}\frac{1}{n}\left(\mu(x,t1,t2) - \left(\left| x_{t1}[i] - x_{t2}[i]\right|\right)\right)^2 \qquad \text{Eqn. 6.7}$$

$$M\%D(x,t1,t2) = \frac{100}{n}\left(\sum_{i=1}^{n}\frac{2\left| x_{t1}[i] - x_{t2}[i]\right|}{\left(x_{t1}[i] + x_{t2}[i]\right)}\right) \qquad \text{Eqn. 6.8}$$

Each of the discrepancy measures conveys useful information that will be used during development of the feature detectors. For example, root mean squared (RMS) error tends to emphasise larger discrepancies, and therefore is used as a bias against the worst case errors, in the objective function for optimisation.

6.1.5 Collecting Training and Validation Data

Subjects were scanned wearing their underwear, and a swimming hat to hold the hair away from the neck and to reveal the true stature. In some cases this was not possible (because of hair styling), and the resulting stature measurements show some error. Breathing state was standardised to a relaxed mid-tide lung volume. Six scans were taken of each subject, although only those scans in the reference posture (as described in Chapter 2) were processed for the results presented here.

Manual measurements, as listed in Appendix A, were recorded, as well as the person's gender, weight and "handedness". The latter determines which side certain dimensions, such as the biceps diameter, are measured. Other information, such as lifestyle and response to the scanning process, was also recorded, but is not discussed further here. Hand measurement and scanning were carried out within one hour of each other, so that valid comparisons could be made between the sets of data.

Subjects for training and validation varied quite significantly, as shown in table 6.1. Recruitment for the training set was not limited in any way, so there is a wide variation amongst

that set, although the BMI values all fall within a medium healthy range. In terms of ages and garment sizing the validation set was more limited, but the body mass index values also include some very low values.

Each subject was hand measured by two skilled people, taking a total elapsed time of an hour per subject to scan, hand measure and record other information. In addition to the recruitment and scheduling of volunteers, this is time consuming. Consequently, the training and validation sets are small in comparison with the size of the problem space.

Different teams were used to collect the training data and the validation data, and therefore it is expected that some bias may be present in the measurement procedures between the two sets. Different subsets of dimensions were measured for the two sets. For example, the training set includes both left and right side measurements, where relevant, whereas for the validation set measurements were only taken on the dominant side.

Table 6.1: Characteristics of the training and validation sets.

	no. samples	age (years)	weight (kg)	height (cm)	BMI (kg/m^2)	UK clothing size
training set	18	18-49	46.0-76.5	151.3-188.2	18.4-29.4	8-16
validation set	18	18-25	40.6-67.5	153.9-175.0	15.3-25.0	10-14

6.2 Implementation of Detectors

This section describes a systematic method for building feature detectors, making use of shape and contextual information in the processed body images. First the overall approach is outlined, and then some specific detectors are described, based on various shape attributes.

As the reference posture is used, it is valid to make the assumption that the feature detection operators need not be invariant to large rotations and translations. This allows certain conditions to be relaxed and therefore simplifies the processing, since the search space is drastically reduced. However, the operators must be tolerant to small differences in position and posture.

6.2.1 Preparation of the Training Set

The first stage is to process each image in the training set in the following way:

- generate the cleaned, smoothed, segmented surface, as described in Chapter 4;
- generate the central axes and shape maps with critical points of curvature, etc., as described in Chapter 5;
- place the main target landmarks on each of the training set surfaces, as in figure 6.1.

Landmarks pre-placed on processed training images:

top of head
nape
supra-clavicular notch (centre front clavicle)
neck base points (left and right)
shoulder points (left and right)
axillae
bust level (centre front)
nipples (left and right)
underbust (centre front)
waist (centre front)
hips (centre front)
crotch point
knees (centre front left and right)
outer ankle bone (left and right)

Figure 6.1: Reconstructed and pre-landmarked subject.

Note that only the main landmarks need be placed manually—the remainder, listed in Appendix A and Appendix B, are automatically located using specific rules, e.g., the top hip is found halfway between the waist and full hip. It is assumed that some error is present in the landmark placement because of the lack of tactile feedback, but the placements should be accurate enough to start to devise the detector functions.

After pre-processing, information to build detectors is thus represented implicitly by:

- context defined by body segmentation, the previously detected landmarks found during segmentation, and the central axes;
- local shape attributes defined by the shape maps.

6.2.2 Translating the Anthropometric Specifications

Contextual and local shape descriptions are held implicitly and explicitly in the anthropometric specification for each dimension. The detector building problem therefore corresponds to finding a match between the anthropometric description and the information available via the shape maps. This information grows as more features are detected. This is done via a systematic manual process, shown schematically in figure 6.2.

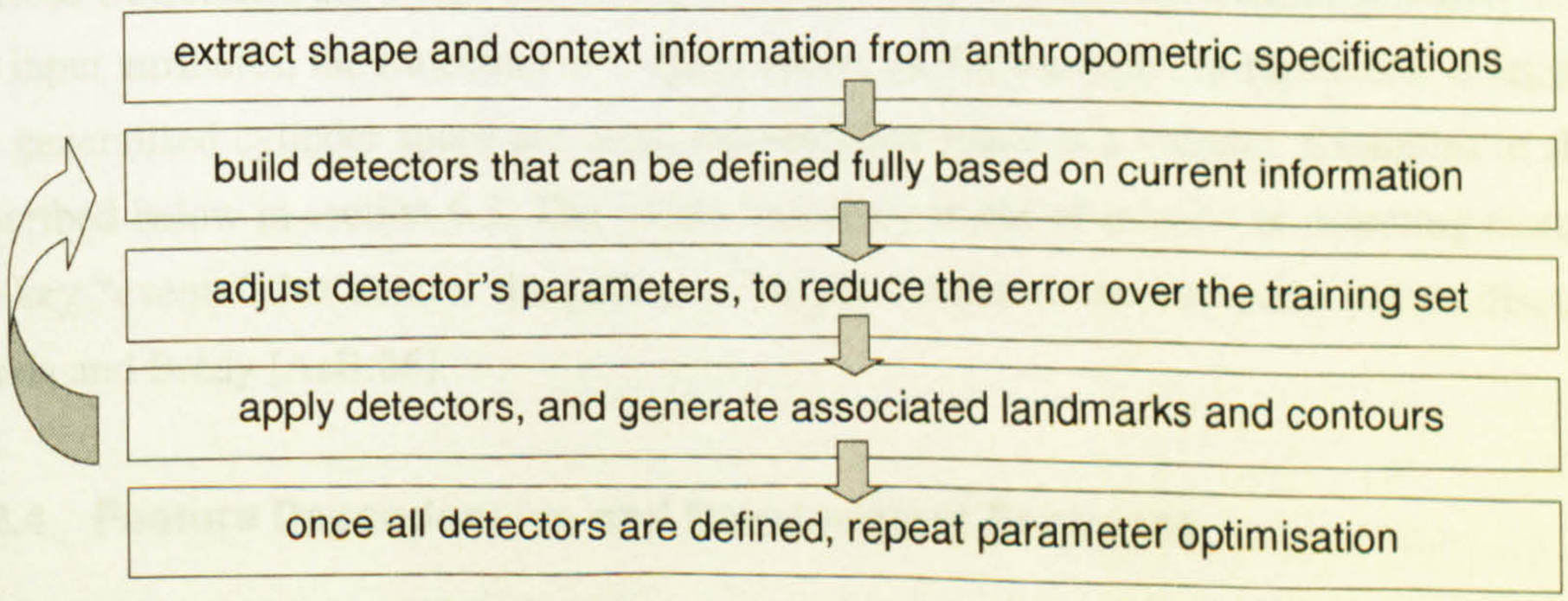


Figure 6.2: Iterative detector construction process.

6.2.3 Search Trajectories

All the detectors take a similar overall form. This is illustrated, in figure 6.3, by the detector to find the left acromion point:

- high-level context defines the starting point and constrains the search volume, e.g.:

  between nape and left underarm point

- intermediate context defines the search path control, e.g.:

  follow contour of shoulder ridge by concatenating turning points in z coordinate
  locate levelling off on top of shoulder, where gradient < g_1

- local shape description = defines the stopping criteria, e.g.:

  stop when gradient > g_2
  place acromion landmark

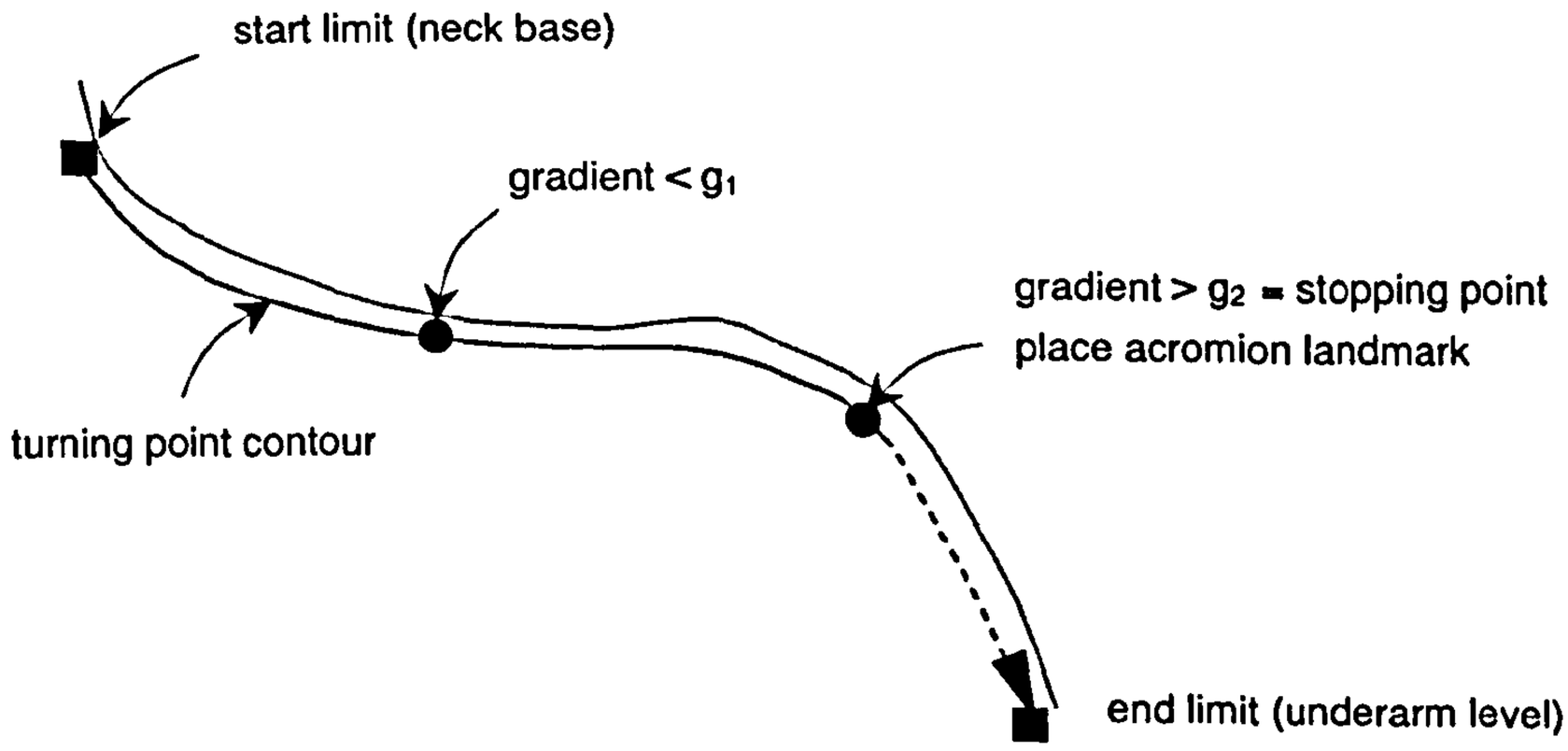


Figure 6.3: Shoulder region, showing search trajectory for the acromion.

In practice, some experimentation has been needed to derive the appropriate form for many of the search functions. It would be preferable to avoid this in the future, by learning the function.

Various trajectories are swept out during feature search. If local differential geometry is used as the input attributes, the trajectory is a space curve, as, for example, in figure 6.3. If attributes of the generalised cylinder space are used, the searched space is a volume. Examples of these are described below in section 6.3. The whole trajectory is not of interest in detecting features, but the key "events" that control the path are. This is similar to the curvature events discussed by Asada and Brady [AsBr86].

6.2.4 Feature Dependencies and Development Sequence

Clearly the sequence in which the detectors are developed is significant. The sequence has been determined here by analysis of the dependency relationships between features, which can be

represented as a directed graph, where $a \rightarrow b$ represents the relationship: b is dependent on a. This is illustrated in figure 6.4. Note that it is not a tree. Features with dependencies must therefore be developed before any of their dependent features. For example, the top of head is located first, which makes it possible to determine the level at which the body starts; the seventh cervical vertebra is then detected, so that the top of the torso can be located; then either the axillae or the crotch landmark can be found. These allow the limits of the torso and the limbs to be defined. It is then possible, for example, to detect features on the legs, since they do not depend on any other landmarks.

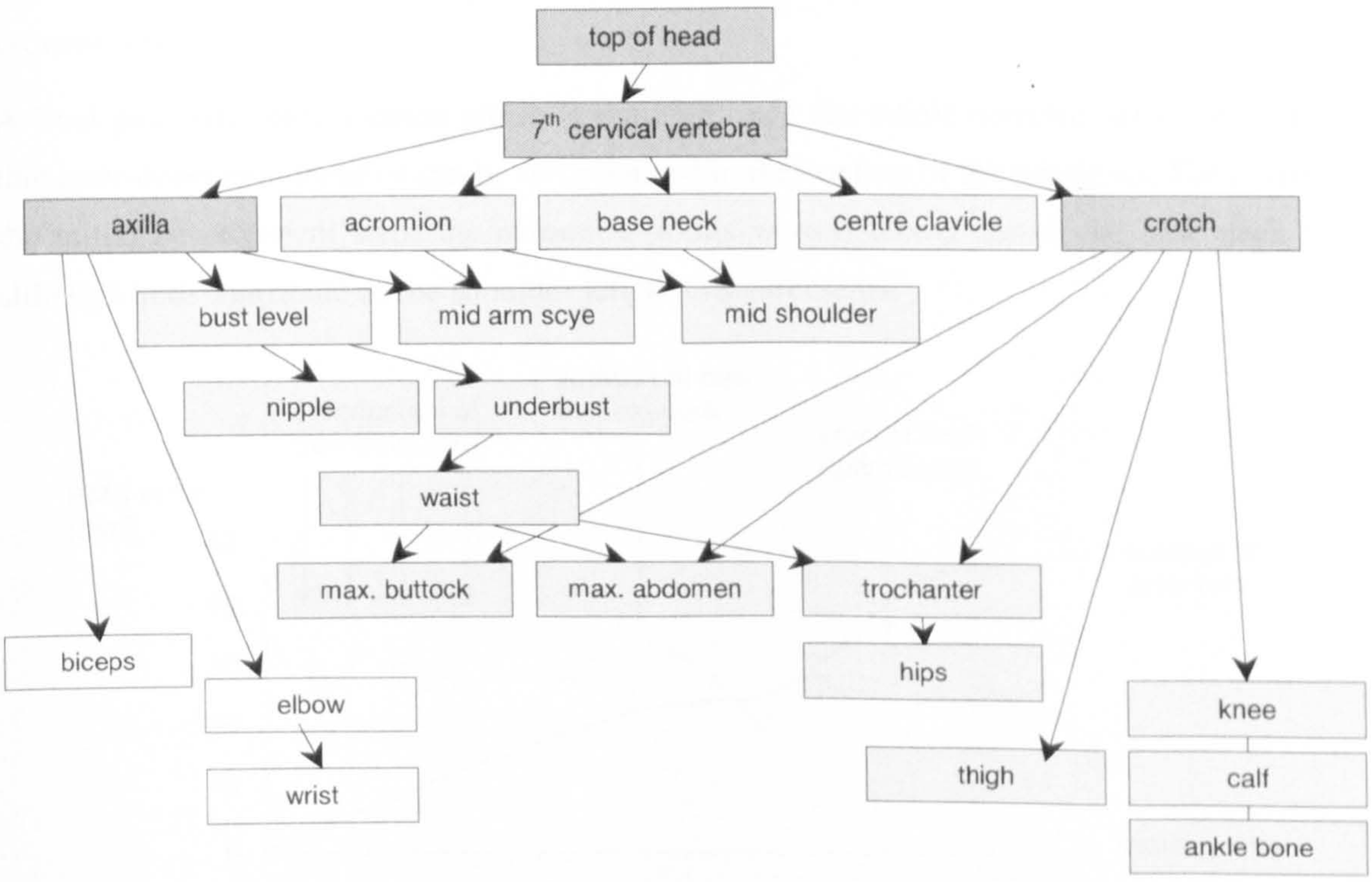


Figure 6.4: Dependency graph for feature detection.
Features found in segmentation are heavily shaded, unimplemented features are unshaded.

6.2.5 Parameter Setting

The acromion example illustrates a surface traversal function with two parameters, g_1 and g_2, which must be optimised. Because the detectors are built sequentially, each set of parameters can be roughly optimised by exhaustive search within a small range of values (testing, on average, eight values for each parameter). The RMS error value on the training set is obtained by applying the relevant tape measurement functions to the landmarks. For example, once the seventh cervical, base neck and nipple landmarks are found, the cervical-to-breast-point contour can be generated using the multi-segment tool to generate a contour between the three landmarks. Error on the acromion landmark corresponds to the error on the shoulder width and the two (left and right) shoulder lengths.

Figure 6.5 shows the reduction in the RMS error over training "time". Each timeslot represents one evaluation of the detectors, rather than elapsed time as measured in the usual sense. Discontinuities occur whenever a new detector is introduced, and some of these points are indicated by arrows. The broken line shows the increase in the number of detectors in the system over time. Typically the error increases, then decreases as the parameters are optimised. In practice, the training data was collected in two batches (of eight and ten subjects respectively). A sharp discontinuity appears where the new subjects were added to the training set. At this point the new data is effectively out-of-sample, and so this increase in error is to be expected. This is followed by a reduction in error as the parameters were reoptimised sequentially.

A final parameter optimisation phase is required once the whole detector set is developed, so that inter-dependent features can be re-optimised in the context of the whole set. For example, in the initial development steps the acromion points were detected before the base neck points, although both contribute to the shoulder length measurements.

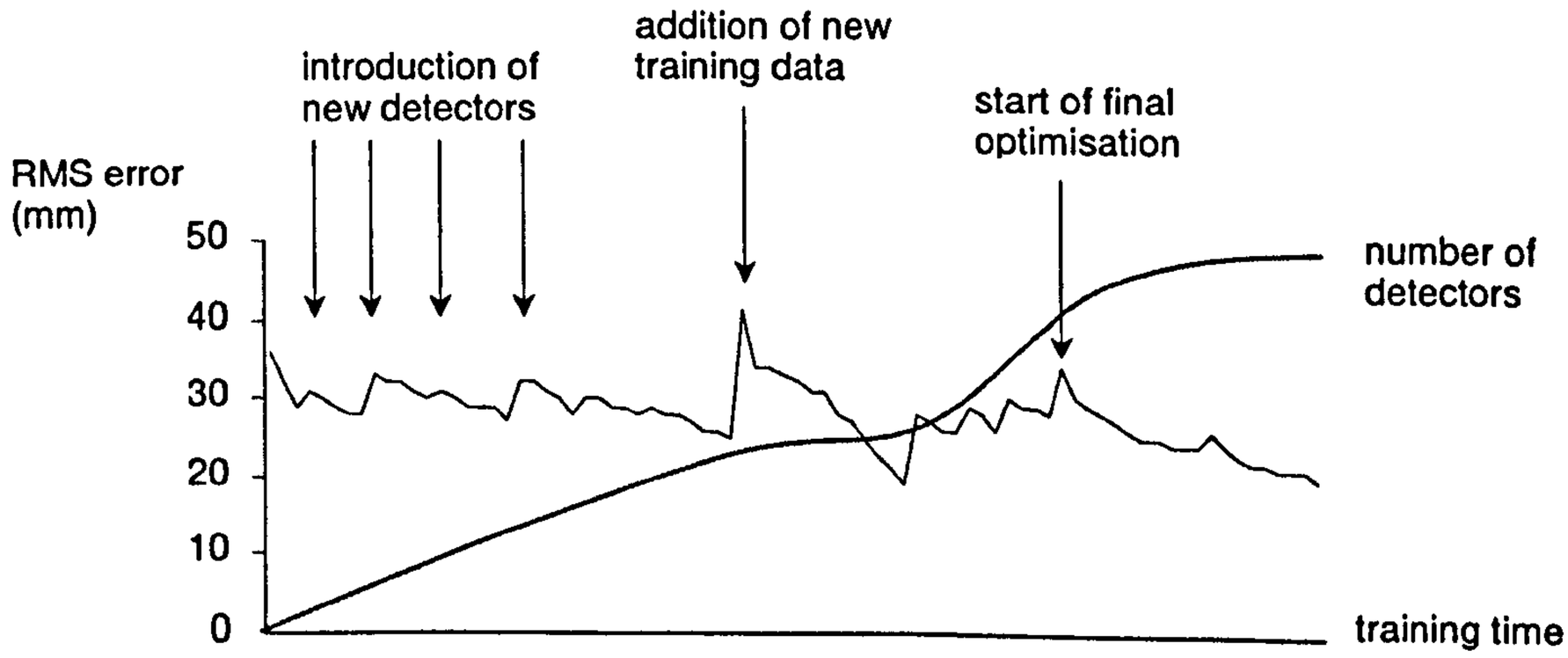


Figure 6.5: RMS error profile against training time.

However, it is important to note that the parameters are optimised here *sequentially*, rather than by *global* optimisation over the whole set. This approach has certain limitations because of the dependencies between landmarks, as illustrated in figure 6.4. A single landmark can be used to generate several dimensions, and most dimensions require more than one landmark. The search space thus expands non-linearly as new landmarks and contours are added. It might be assumed, therefore, that the resulting landmark configurations actually correspond to a *local* optimum in the solution space. Ideally all the inter-dependent feature detectors would be optimised simultaneously to deal with the non-linear relationships, and this is clearly an immediate task for future work. Because of the large size of the search space it will be essential to reduce its complexity by a linear separation into sub-problems, based on the dependency graph shown in figure 6.4. For example, if should be possible to optimise the shoulder measurements independently of the hip and leg measurements.

In this work, although several discrepancy measures have been discussed, only RMS error has been used in the objective function. It is possible that a multi-objective approach would produce better results. In particular, the use of multiple evaluation functions would allow more control over application-specific requirements. For example, worst-case errors might be considered more of a liability in an unsupervised system, than in a system where an operator is present to detect anomalies, and perhaps choose to rescan the subject.

6.2.6 Statistical Defaults

Certain features have proven extremely difficult to detect by a single search criterion over all of the training examples. The knees are one such feature, because of the subtlety and variation across subjects. In most cases, a particular gradient g in x can be used to detect the lower part of the knee, where the surface curves outwards. However, this is not reliable in all cases, and a default value, based on the statistical mean of the ratio between the crotch and the floor is needed where the search-based detector fails. This has been computed here from the manual data on the training set.

In practice, this is the only detector for which a default rule has been used. In the future it might be beneficial to use such statistical defaults for other features if this reduces the worst case error when a detector function fails. However, it is noted here that it will not be straightforward to detect the "failure" condition in all cases.

6.3 Example Detectors

Rather than describing each of the feature detector functions that generate this output, this section describes a few examples using psuedo-code to illustrate the typical structures and input attributes used.

6.3.1 Features in Generalised Cylinder Space

Many features can be detected by considering each body segment in terms of certain attributes as a function of y, distance along the vertical axis, as discussed in Chapter 5:

$c = f(y)$ = convex hull length at y, in a slice in the horizontal plane

In these detectors y is discrete valued. Several features can be obtained by detecting critical points in the convex hull function, as illustrated below.

Calf girth

The calf is detected as a turning point in the convex hull function between the knee and ankle:

between y(knee) and y(ankle)
find $dc/dy = 0$, $d^2c/dy^2 < 0$
calf = closed contour in xz plane

Seventh cervical vertebra

The seventh cervical vertebra is detected at a discontinuity in the first derivative of the convex hull function, i.e., an abrupt change where the main part of the neck joins the shoulders:

between y(head– m) and y(head-n)
find $d^2c/dy^2 > g$
seventh cervical vertebra = centre back point

The parameters *m* and *n* define the *y* range in which the landmark is expected to lie. It might be expected that this detection function will not work well on people with extremely well developed neck and shoulder muscles, but this has not been tested here.

Waist

The waist can be detected where the gradient of the convex hull length is increasing rapidly, corresponding to where the torso increases significantly at the top of the hips:

between underbust and crotch
find $d^2c/dy^2 > g$
waist = closed contour in xz plane

Again, there are obvious cases where such a detector might fail, for example, on very overweight subjects.

6.3.2 Features from Curvature and Derivative Maps

The curvature and derivative maps discussed in Chapter 5 can be used in a similar way to construct detectors.

Acromion Points

The acromion detector, which makes use of these attributes, has already been described above. In this description, *u* is the length along the contour that is generated when turning points in *z* are concatenated along the top ridge of the shoulder:

between y(seventh cervical vertebra) and y(left axilla)
generate contour on shoulder ridge by concatenating turning points in z
find $dz/du < g_1$
find $dz/du > g_2$
place acromion point

Bust Girth

The bust level can be found at the contour connecting the local turning points in *x* along the vertical axis of the torso:

between y(axillae) and y(crotch)
generate contour on front torso by concatenating turning points in x
find first $dx/du = 0$, $d^2x/du^2 > 0$
fullest bust = closed contour in xz plane

Note that this does not necessarily occur at a global maximum on the contour described above, because postural differences can result in the global maximum occurring on the abdomen. The nipples are then found as the turning points in the radial length parameter on the horizontal closed contour at that level.

Navel

In practice, the surface curvature attributes have not been used directly in many cases, but have been of more interest in providing clues about surface shape for designing the overall functional form for detectors. One place where surface curvature has been used directly is to detect the navel, which appears as a cluster of high positive curvature vertices on the centre front of the torso. The sum of curvature values in a 7×7 window are computed at each vertex in the vertical band close to the centre front of the torso, and the point at the centre of the window with the highest positive value is labelled as the navel.

6.4 Experimental Results

After development of all the detectors for the automated system, each set of scans was processed, using a Dell OptiPlex GX1 with a 400MHz processor and 128 MB RAM. Each scan took 45 seconds to process, from the raw image to surface reconstruction, feature detection, extraction of the measurement information and rendering.

For the training set one scan for each of the 18 subjects in the reference posture was used. For the validation set twelve of the subjects had two scans in the reference posture, and the remainder had one scan only in the reference posture. Although the automated system extracts about 50 measurements, manual reference data was not available for all these values, and therefore results are only presented for those dimensions with corresponding reference data.

A summary of the output is presented in tables C.1 to C.9 in Appendix C. C.1 to C.3 show results for measurements on the training set. C.1 summarises the manual measurements; C.2 summarises the corresponding automated measurements; C.3 presents comparisons between the manual and automated measurements. Tables C.4 to C.9 present results for the validation set. C.4 to C.6 summarise manual and automated measurements; C.7 and C.8 present comparisons between automated and manual measurements for each of the two scans; C.9 summarises results of the reproducibility experiments, comparing the automated output of the two scans.

6.5 Discussion

In this section the results are discussed, firstly in terms of inter-technique variation and reproducibility. Sources of discrepancy between the scan-derived and reference data are then discussed. This is followed by a comparison of various acceptability criteria, and how the different performance profiles relate to particular applications' requirements.

6.5.1 Visual Assessment of the Output

As the automated system processes the raw scans, it renders the results visually in 3D, showing the reconstructed skin surface, the detected and labelled landmarks, and the electronic tape measure. This makes it possible to assess the subject's posture, and to observe where the landmarks and the electronic tape measure have been located on the body. Figure 6.6 shows some of these outputs.

Although subjects were briefed about the reference posture for scanning, there were very noticeable variations, and in many cases the head was held downwards, possibly where they followed the motion of the imaging units. No obvious anomalies were observed in the electronic tape measure itself. However, some of the waist girth measurements appeared to be located slightly incorrectly by the automated system. In one case, where the subject had a significant twist to the spine, the nipple landmarks were placed incorrectly. Such variations in posture appear to make significant differences to the results. It would be possible to carry out repeat scans in order to obtain the "best" results, but this has not been done here, and so these differences are preserved in the results. Since sensitivity to posture is an extremely important issue, with many implications, it will be discussed in more detail in section 6.5.4.

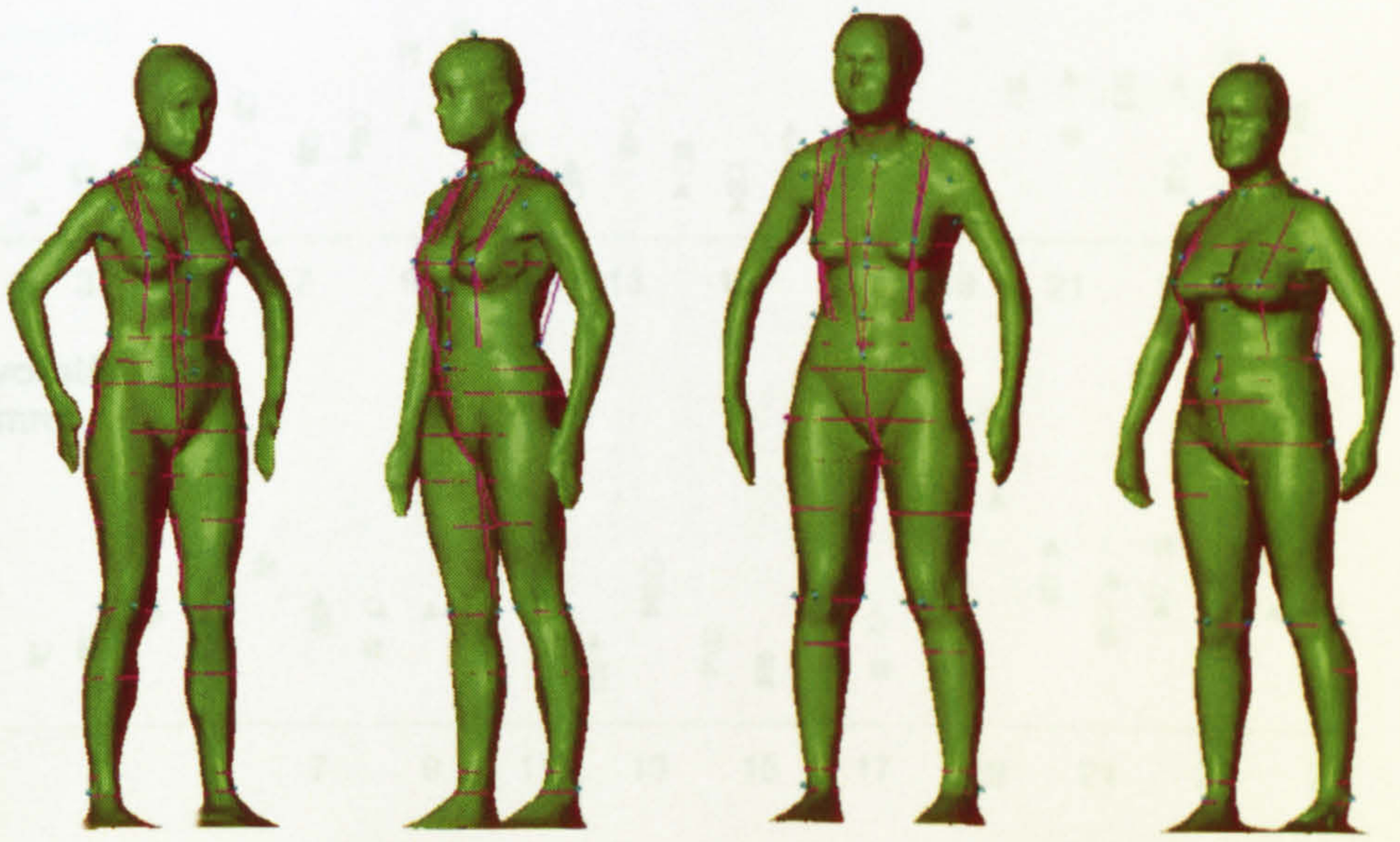


Figure 6.6: Sample output, showing landmarks and electronic tape measure.

6.5.2 Performance on Training and Validation Data

Measurements extracted from the scans have been compared against the equivalent manual measurements, and the results for *signed mean*, *RMS error*, *worst-case error* and *standard deviation* are presented in figure 6.7. Results are shown on all dimensions for which comparable reference data was available in both the training and validation sets. Two sets of results are presented for the validation set—one for each scan that was processed.

Dimensions

1. stature
2. nape height
3. crotch height
4. centre front waist height
5. right knee height
6. neck girth
7. bust girth
8. underbust girth
9. waist girth
10. upper hip girth
11. top hip girth
12. hip girth
13. right thigh girth
14. right knee girth
15. right calf girth
16. shoulder width
17. right shoulder length
18. bust width
19. back width
20. cervical to breast point
21. right mid shoulder to front waist
22. right mid shoulder to back waist
23. scye depth
24. total crotch length
25. body rise

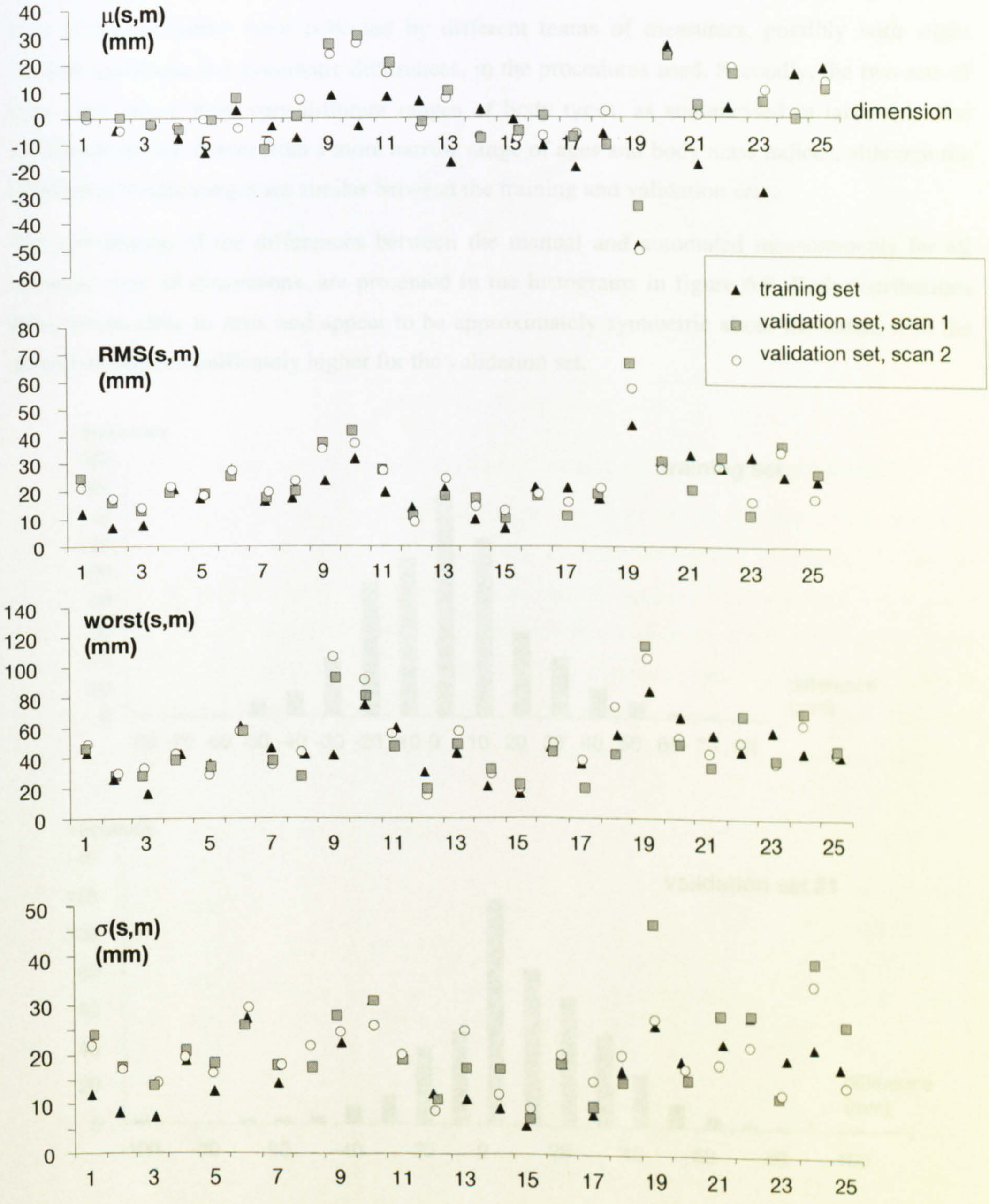


Figure 6.7: Comparison of automated versus manual measurements.

Values on RMS, worst-case error and standard deviation of error appear to be strongly correlated. Three dimensions stand out as having especially high values on these, corresponding to peaks on all the plots: the waist, upper hip and back width. A noticeable feature is the similarity between the performance of the two validation sets, and the difference between these and the training set. It would normally be expected that the results for training and validation data would be significantly different, since the validation data constitutes previously unseen samples. However, there are also other factors that can account for differences. Firstly, the two sets of measurements were collected by different teams of measurers, possibly with slight random variations and systematic differences, in the procedures used. Secondly, the two sets of data were drawn from very different ranges of body types, as summarised in table 6.1. The validation set was drawn from a more narrow range of ages and body mass indices, although the height and weight ranges are similar between the training and validation sets.

The distribution of the differences between the manual and automated measurements for all subjects, over all dimensions, are presented in the histograms in figure 6.8. Both distributions have means close to zero, and appear to be approximately symmetric about the mean, with the spread of values significantly higher for the validation set.

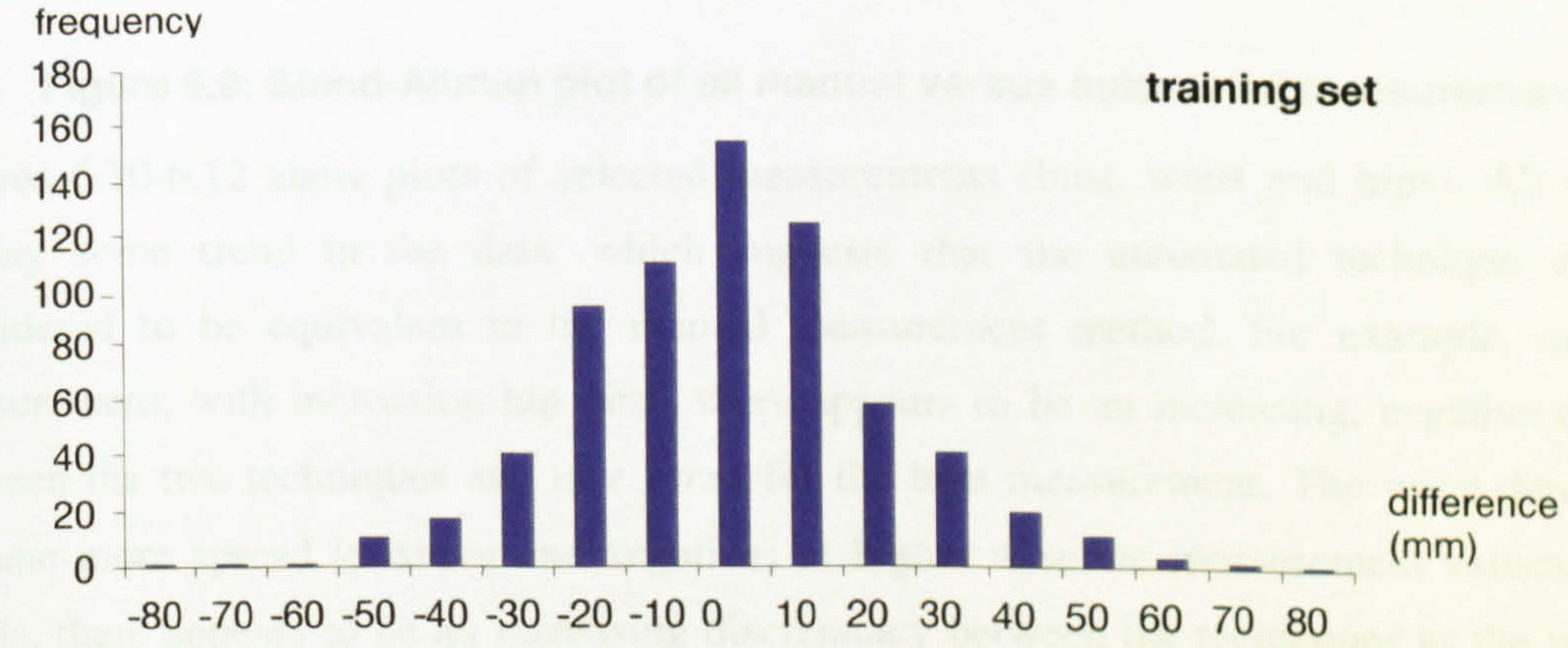


Figure 6.8: Distribution of differences between scan and manual measurements.

A Bland-Altman plot of all the data is shown in figure 6.9, comparing manual measurements and automated measurements from the validation set. If the two measurement techniques are to be considered equivalent the plot should show a random scattering of points close to the $y=\mu$ line [BlAl86]. The absolute measurement values appear to cluster in several groups along the x axis, and therefore it is not appropriate to draw too much from this combined plot, but to separate the data by attribute and by subject, as discussed below.

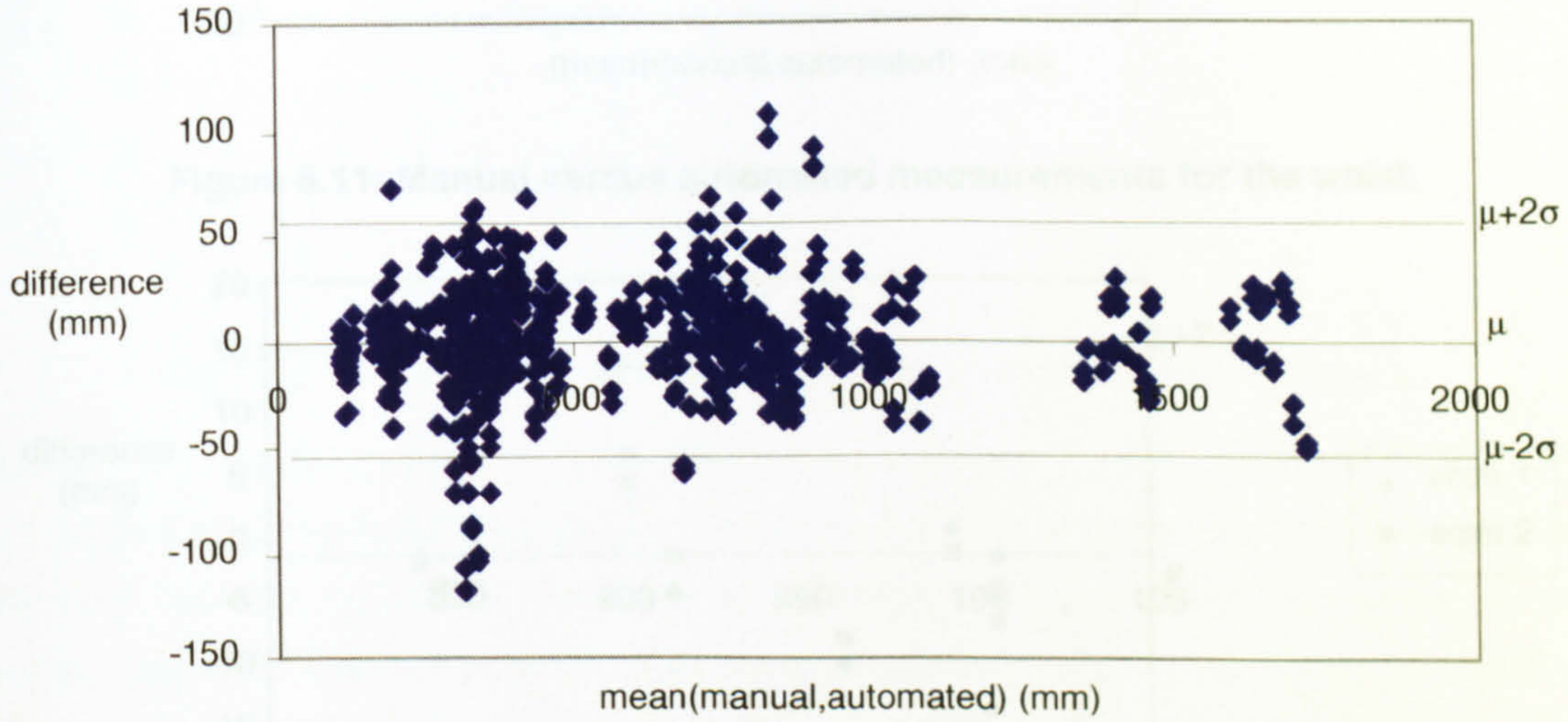


Figure 6.9: Bland-Altman plot of all manual versus automated measurements.

Figures 6.10-6.12 show plots of selected measurements (bust, waist and hips). All appear to display some trend in the data, which suggests that the automated technique cannot be considered to be equivalent to the manual measurement method. For example, on the hip measurement, with increasing hip girth, there appears to be an increasing, negative difference between the two techniques and *vice versa* for the bust measurement. The waist data tends to become more spread (positive and negative) at higher absolute measurement values. In other words, there appears to be an increasing discrepancy between the techniques as the waist girth becomes larger.

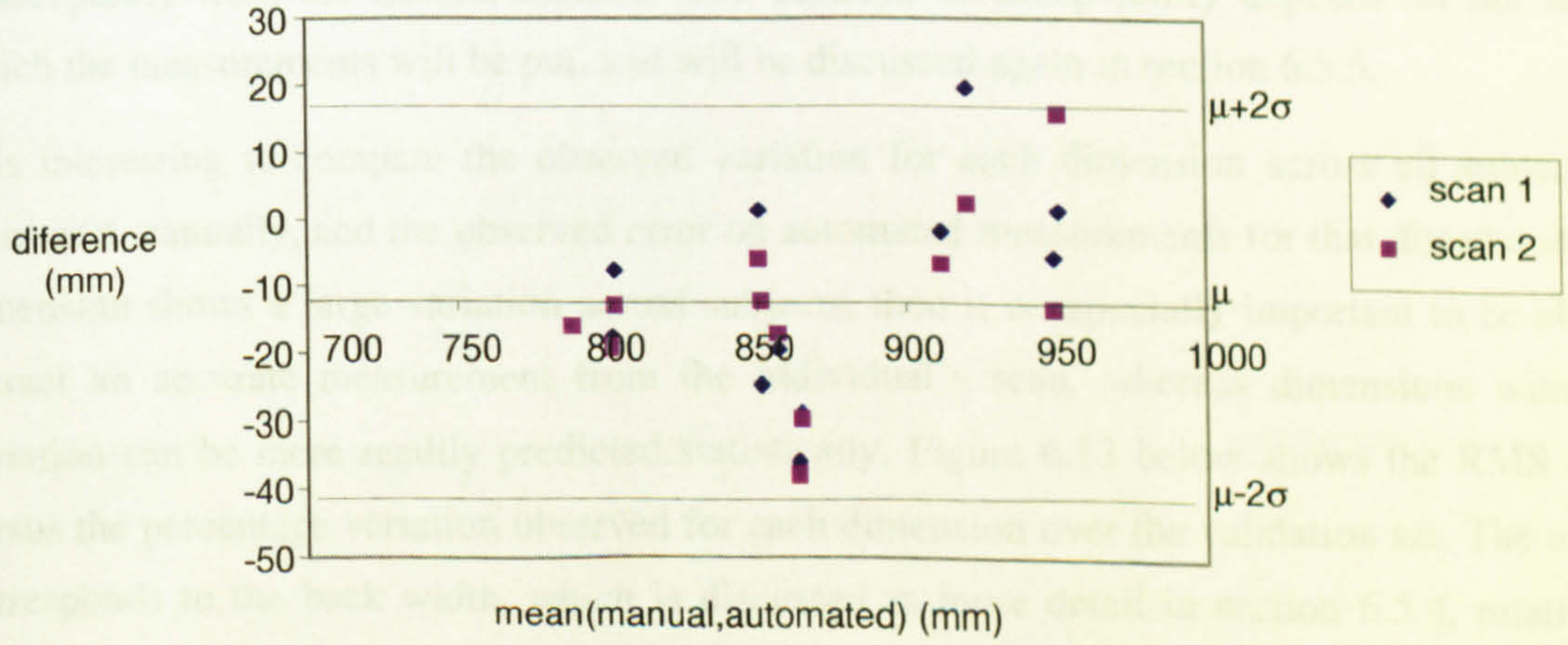


Figure 6.10: Manual versus automated measurements for the bust.

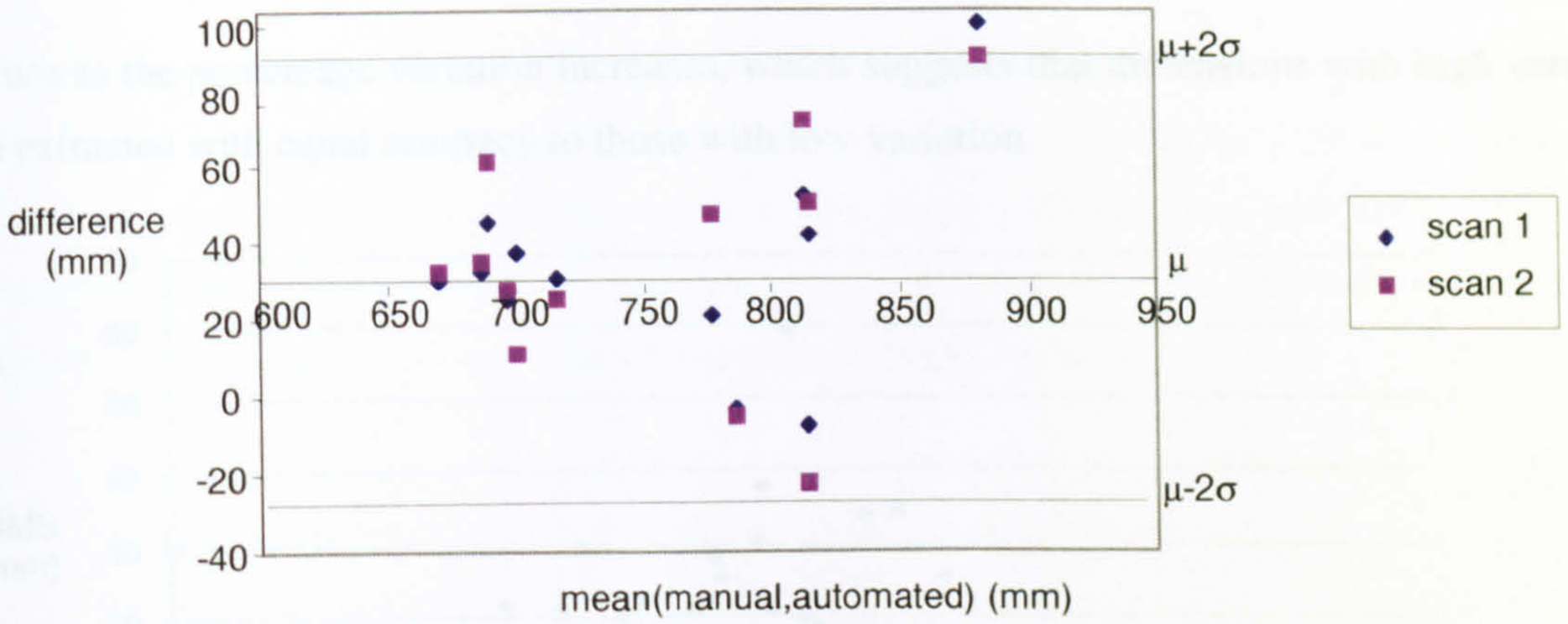


Figure 6.11: Manual versus automated measurements for the waist.

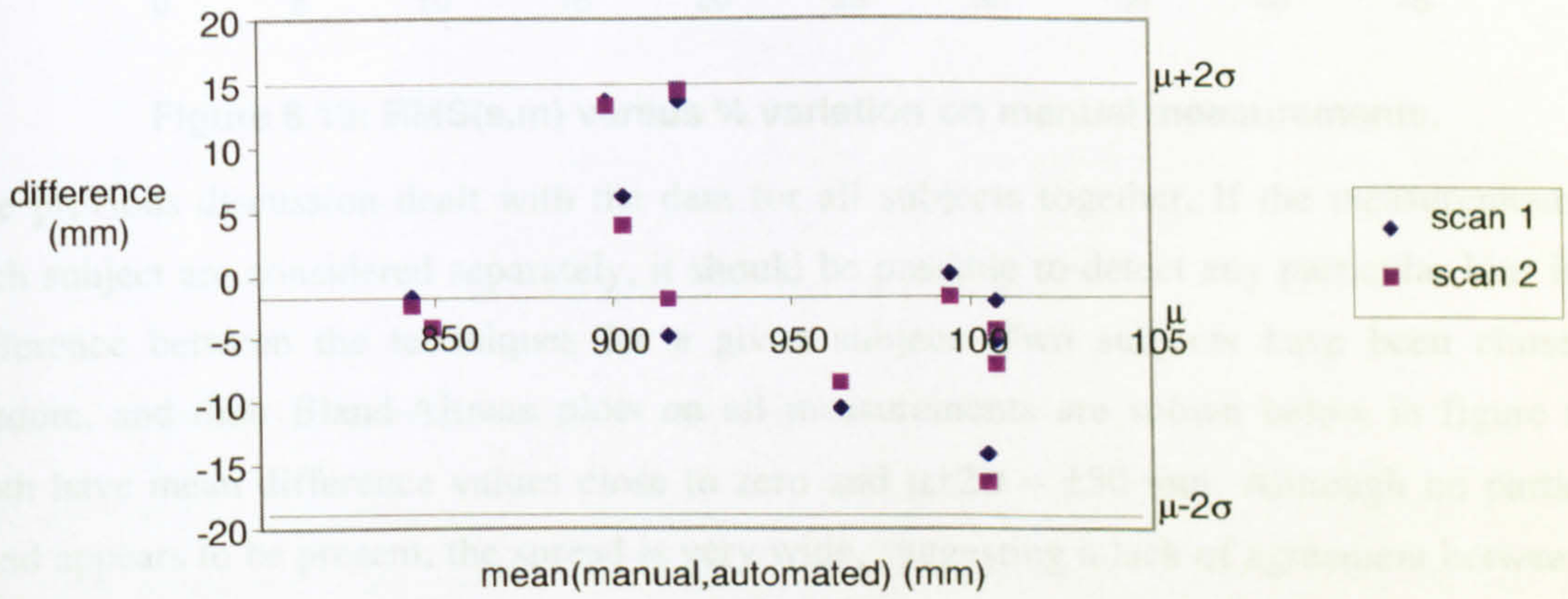


Figure 6.12: Manual versus automated measurements for the hips.

An adjustment can be made to the automated measurements by adding the mean difference value to improve the agreement between the two techniques. While leaving the spread of values the same, this shifts the data to give a mean difference of zero. If such a correction is made, the spread of values on the hip and bust measurements may be considered sufficiently small that the values obtained by the automated system are acceptably close to the manual measurements. The waist measurements show a wider spread, and therefore might be considered to differ unacceptably from the manual method. The decision of acceptability depends on the use to which the measurements will be put, and will be discussed again in section 6.5.5.

It is interesting to compare the observed variation for each dimension across all subjects as measured manually, and the observed error on automated measurements for that dimension. If a dimension shows a large variation across subjects, then it is especially important to be able to extract an accurate measurement from the individual's scan, whereas dimensions with low variation can be more readily predicted statistically. Figure 6.13 below shows the RMS value versus the percentage variation observed for each dimension over the validation set. The outlier corresponds to the back width, which is discussed in more detail in section 6.5.4, relating to scan posture and measurement sensitivity. There appears to be no significant increase in RMS

values as the percentage variation increases, which suggests that dimensions with high variation are extracted with equal accuracy to those with low variation.

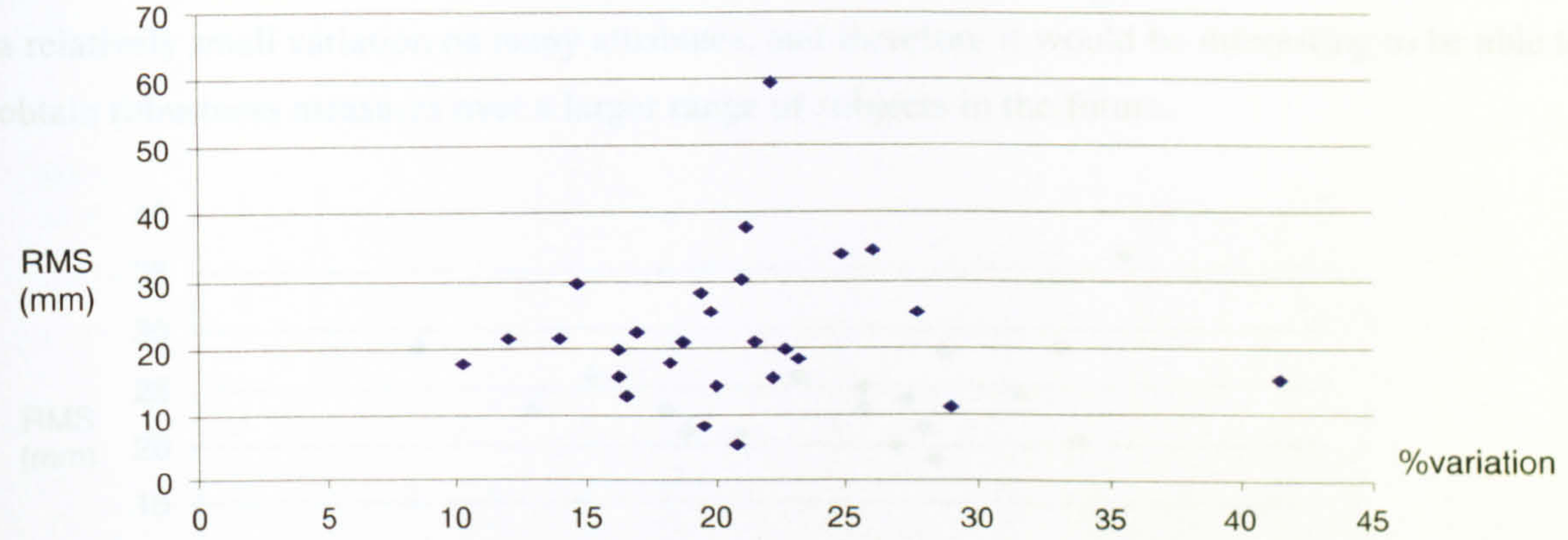


Figure 6.13: RMS(s,m) versus % variation on manual measurements.

The previous discussion dealt with the data for all subjects together. If the measurements for each subject are considered separately, it should be possible to detect any particular bias in the difference between the techniques for a given subject. Two subjects have been chosen at random, and their Bland-Altman plots on all measurements are shown below in figure 6.14. Both have mean difference values close to zero and $\mu \pm 2\sigma \approx \pm 50$ mm. Although no particular trend appears to be present, the spread is very wide, suggesting a lack of agreement between the techniques. When considering individual dimensions, however, as in figures 6.10-6.12, it appears that a correction factor could be used to improve the agreement between the techniques, and thereby reduce the spread observed in figure 6.14.

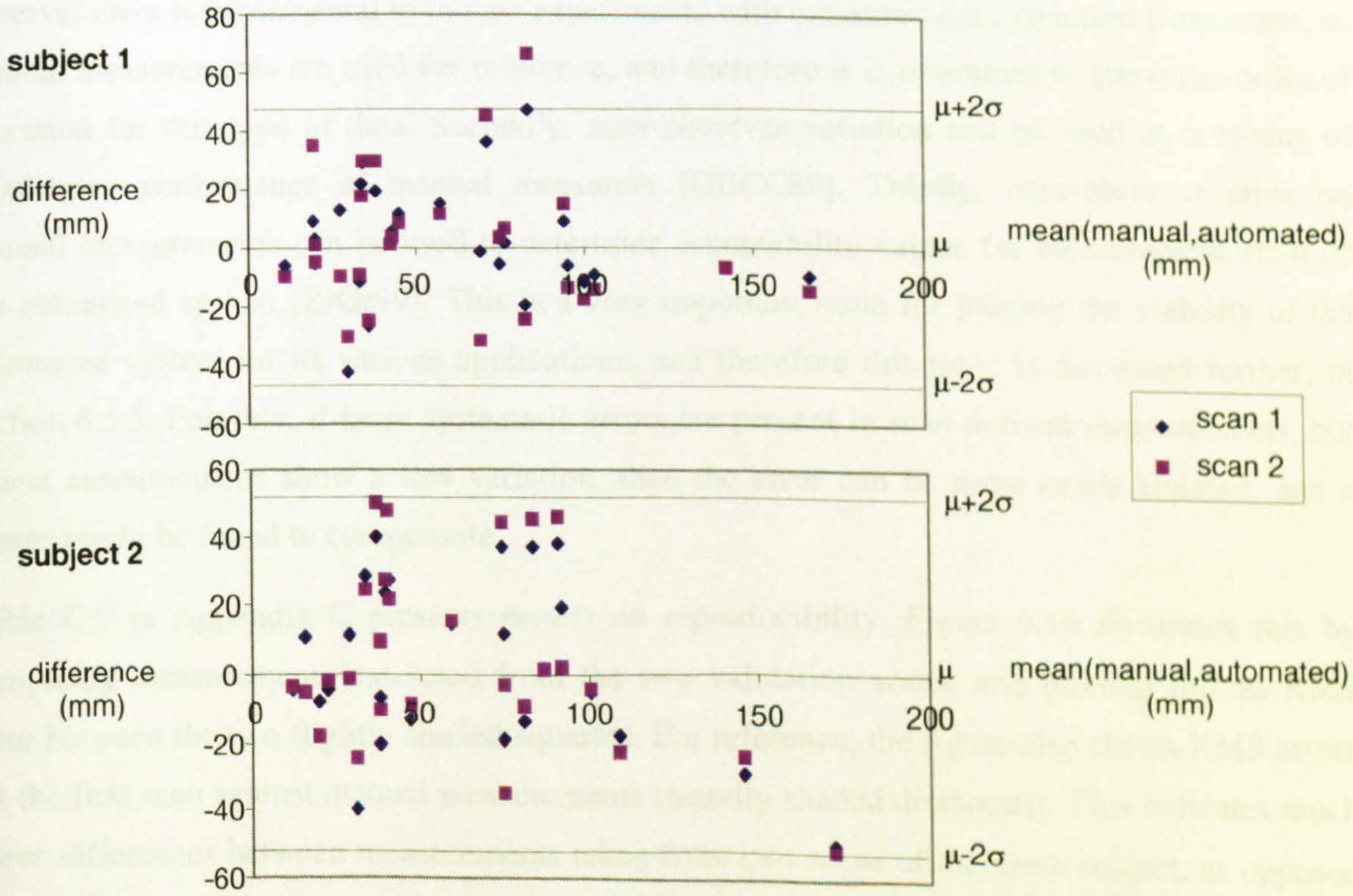


Figure 6.14: Manual versus automated measurements on all measurements, by subject.

Another important attribute of the automated system is its robustness over a varied range of subjects. Figure 6.15 shows the RMS on all measurements for a given subject versus their BMI. No trend is observable in the data. However, it was noted before that the validation set has only a relatively small variation on many attributes, and therefore it would be interesting to be able to obtain robustness measures over a larger range of subjects in the future.

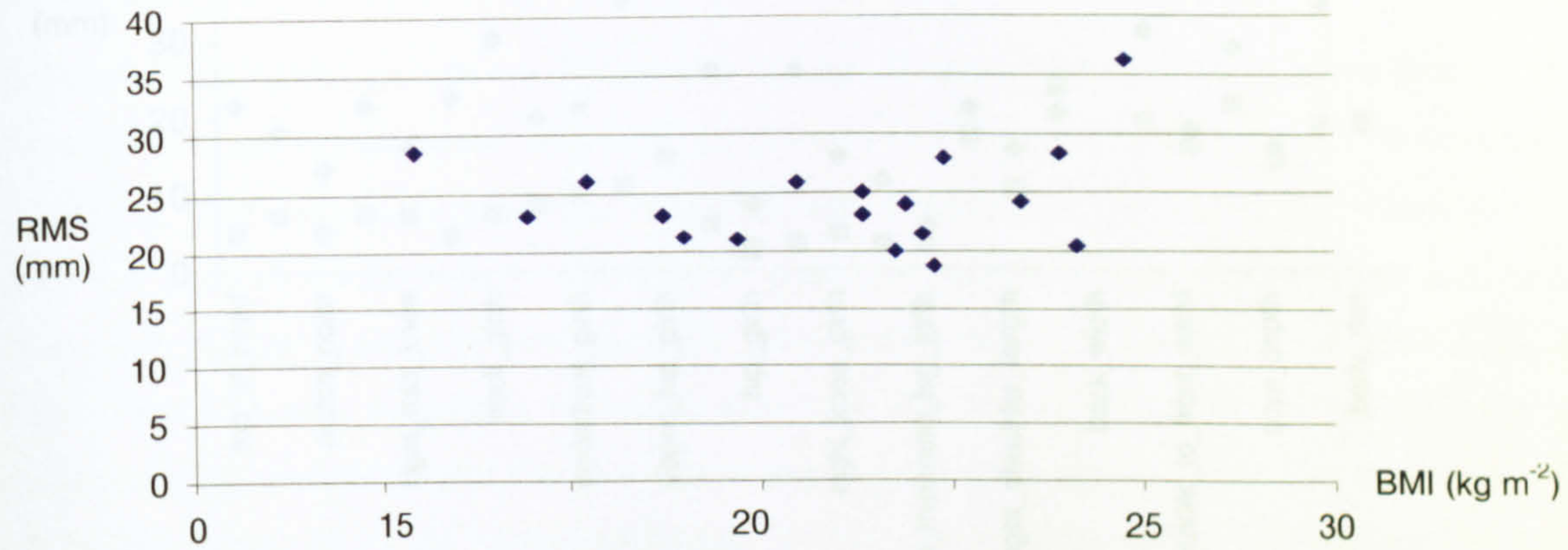


Figure 6.15: RMS(s,m) versus BMI.

6.5.3 Intra- and Inter-Technique Variation

This section first discusses the reproducibility of repeat measurements within a given technique, and then discusses some results for comparison of traditional measurements against those extracted from 3D scans.

Measurement of *intra*-technique variation is important for several reasons. Firstly, inter-observer error is fundamental to *in vivo* experiments with measurement extraction from scans, as manual measurements are used for reference, and therefore it is important to know the order of precision for this type of data. Secondly, inter-observer variation can be used as a means of monitoring performance of manual measurers [GBCC89]. Thirdly, inter-observer error on manual measurements can be used to determine acceptability values for measurement error of the automated system [BrGr99]. This is a very important issue for judging the viability of the automated system for its various applications, and therefore this topic is discussed further, in section 6.5.5. Fourthly, if large systematic errors are present in scan-derived measurements, but repeat measurements show a low variation, then the error can be more easily isolated, and a means might be found to compensate.

Table C.9 in Appendix C presents results on reproducibility. Figure 6.16 illustrates this by comparing measurements extracted from the two validation scans, and plotting this as RMS error between the two (lightly shaded squares). For reference, the figure also shows RMS errors for the first scan against manual measurements (heavily shaded diamonds). This indicates much lower differences between measurements taken from two scans of the same subject, as opposed to the differences between scan and manual measurements, which is an encouraging result.

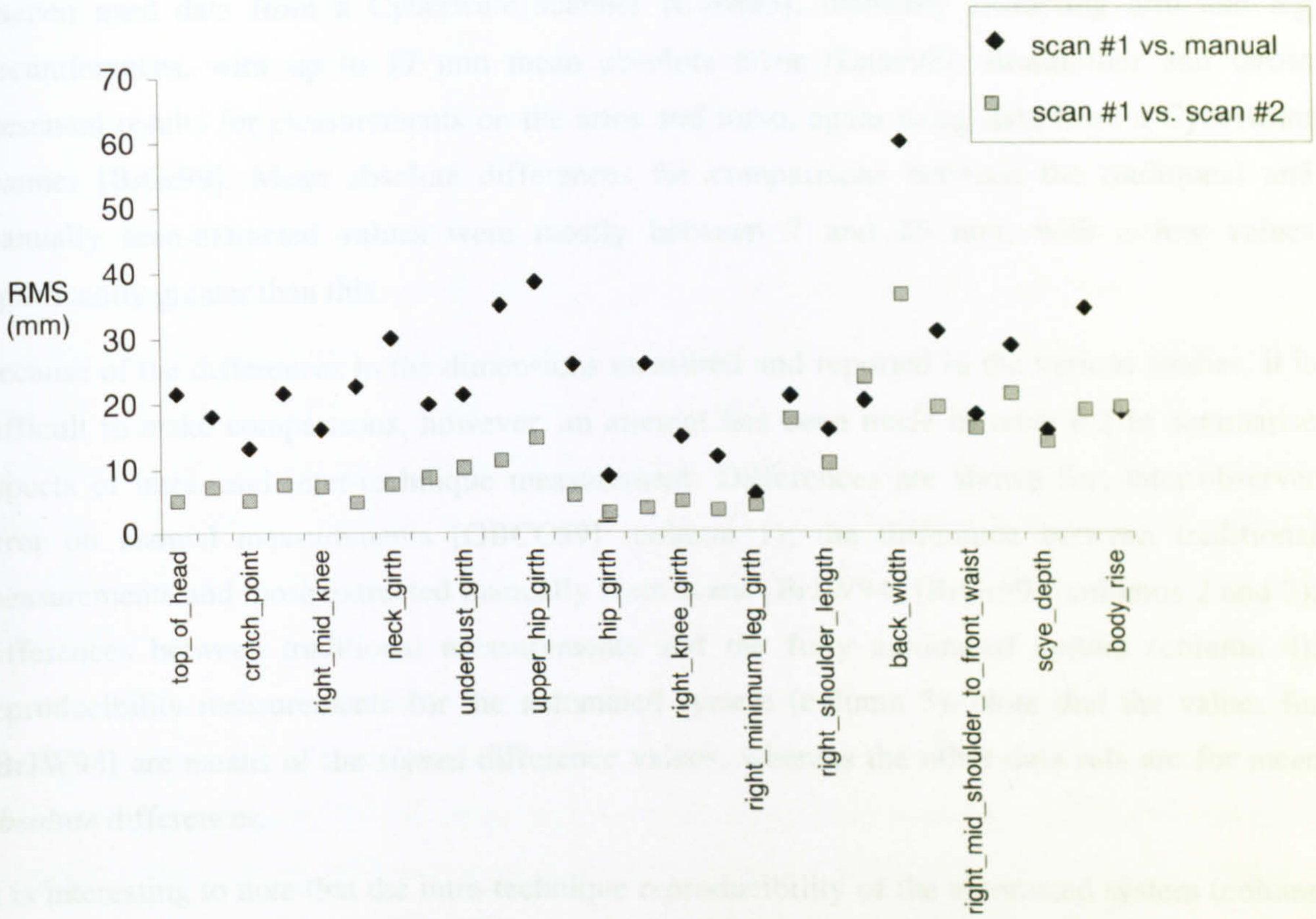


Figure 6.16: Reproducibility - RMS(scan #1, scan #2) and RMS(scan #1, manual).

Gordon *et al* discuss inter-observer error amongst skilled measurers using traditional anthropometric methods on a wide range of dimensions. Mean absolute differences are less than 10 mm on most dimensions, but a few have up to 24 mm error [GBCC89]. Brooke-Wavell *et al* describe experiments to record inter- and intra-observer variability, using traditional methods, and also by extracting measurements manually from scans with pre-placed marks at key points [BrJW94]. In all cases the differences are between approximately 2-13 mm over several measurements on the neck, shoulders and torso. Not surprisingly, the intra-observer difference is generally lower than inter-observer difference. Typically the *inter*-observer difference is greater on traditional measurements than on scans, and *intra*-observer difference is less on traditional measurements than on scans. Bradtmiller and Gross also report inter-observer differences on traditional measurements, with similar values to those found by [GBCC89] and [BrJW94].

Previous comparisons between anthropometric measurements from whole-body imaging systems and traditional measurements have been made by several groups of researchers, although few results have been published [BrJW94] [Daan98] [BrGr99]. In those studies subjects were scanned with flat or protruding landmarks pre-placed before scanning, and the measurements were extracted manually from the 3D scans. The results for the system described in this thesis are significant, as they appear to be the first to be published, to date, for a *fully automated* system [DKWB98] [DeTr98] [DDBT99].

Brooke-Wavell *et al* presented results using LASS [JWHR89] for measurements on the neck and torso, with mean differences mostly 11 mm or less, but some up to 28 mm [BrJW94].

Daanen used data from a Cyberware scanner [Cybe93], manually extracting arm and leg circumferences, with up to 17 mm mean absolute error [Daan98]. Bradtmiller and Gross presented results for measurements on the arms and torso, again using data from a Cyberware scanner [BrGr99]. Mean absolute differences for comparisons between the traditional and manually scan-extracted values were mostly between 7 and 25 mm, with a few values significantly greater than this.

Because of the differences in the dimensions measured and reported in the various studies, it is difficult to make comparisons, however, an attempt has been made in table 6.2 to summarise aspects of intra- and inter-technique measurement. Differences are shown for: inter-observer error on manual measurements [GBCC89] (column 1); the difference between traditional measurements and those extracted manually from scans [BrJW94] [BrGr99] (columns 2 and 3); differences between traditional measurements and the fully automated system (column 4); reproducibility measurements for the automated system (column 5). Note that the values for [BrJW94] are means of the *signed* difference values, whereas the other data sets are for mean *absolute* differences.

It is interesting to note that the intra-technique reproducibility of the automated system (column 5) is better than the skilled manual measurers (column 1) on these dimensions. In the inter-technique comparison (columns 2 to 4), where the measurements extracted from scans are compared against traditional measurements, the values are higher than the intra-technique values.

Table 6.2: Differences (mm) for measurements on women.

	MAD(m1,m2)[2] [GBCC89]	μ(s,m) [BrJW94]	MAD(s,m) [BrGr99]	MAD(s,m) automated	MAD(s1,s2) automated
bust girth	12	-17	16.5	16	6
crotch height	6	-	16.1	10	4
neck girth	5	11	7.1	26	6
hip girth	8	-6	7.8	7	2
waist circumference	10	11	16.4[3]	25	9
mean absolute	8	11	13	17	5

6.5.4 Sources of Discrepancy

Many potential sources of discrepancy between automated and traditional measurements have been mentioned earlier in this chapter. In this section these issues will be discussed in more detail.

---

[2] Each subject was marked up only once before repeat measurements by different observers.

[3] The waist is measured in two ways by [BrGr99] from a scan; the mean of the two values is given here.

The early processing of the scan data, described in Chapter 4, raised several potential sources of error. One of the main problems is in occluded areas. The areas most strongly affected are the underarms, crotch, and on the insides of the legs and arms. However, these areas are important in many anthropometric processes, and so it is important that the segmentation and reconstruction processes can deal with them adequately. The tests carried out in Chapter 4 indicated that the segmentation of the legs worked well, but that the underarms were often detected a centimetre or two too low on the body. In addition, the inner surfaces of the arms and legs still appeared slightly "rippled", even after smoothing. This might be expected to cause problems for the procedures that detect minima and maxima on the surface. The successful detection of dimensions such as the maximum calf girth and minimum ankle girth, indicates that this does not appear to be the case, however, no tests have been carried out here on arm dimensions.

Tests were carried out in Chapter 5 to isolate the differences between the electronic tape measure and its real counterpart, without errors from the *automated detection* of features, as explored in this chapter. The electronic tape measure was placed manually on the scan images to extract various types of surface measurement, and compared against the equivalent measured *in vivo*. In most cases a very small error (~1%) was reported, which provides a baseline error level that would be expected.

Discrepancies in measurements arising from contact and non-contact differences have been discussed before [BrJW94] [Daan98] [BrGr99]. The use of pressure from a taut tape measure tends to pull in loose fabric, which may cause some of the difference in dimensions such as the trunk circumference and total crotch length (although these measurements also have other potential sources of error, such as those arising from occlusion). Compression of tissue has also been raised by the researchers mentioned above, so it is of interest to determine if it contributes to errors here. Given the observed magnitudes of discrepancy between the automated and manual measurements presented here, and the many other sources of error, it is assumed that the compression issue does not contribute significantly. It is also possible that some errors would also be introduced because of the assumption of planarity in the electronic system, whereas some contours, such as the neck base, do not lie exactly in a plane, although this has not been tested here.

The largest source of errors in the automated system is perhaps the lack of tactile feedback in locating skeletal features such as the iliac crest, acromion points, seventh cervical vertebra and trochanters. This would be expected to contribute to error on most measurements, and have knock-on effects for dimensions that require the use of several landmarks that are "difficult" to locate. For example, the total crotch length is dependent on the waist location, but the error arising from an incorrectly located waist is doubled, because the measurement is taken from front to back. It also potentially suffers from errors because it requires access to the occluded

area around the crotch, and may be overestimated because of loose fabric. In addition, the images used here do not provide textural feedback, which was used for the manual detection of natural landmarks by [Li97] and for the automated detection of pre-placed fiducial markers by [GMHW95] [LeNu98].

It was noted earlier, in Chapter 2, that the posture used for this work does not match the range of postures adopted by a live subject when being hand measured. It is likely that this will cause some discrepancy between the two techniques, which may cause systematic differences. This was also noted by Bradtmiller and Gross [BrGr99]. Some measurements, such as the back width and bust width, show very high RMS values, and so the sources of these discrepancies must be investigated. When the arms are held away from the body, as in the reference posture, the subject may choose to hold them in a wide range of positions, and back width is strongly sensitive to the arm position. The standard posture for measuring this dimension *in vivo* is with the arms relaxed at the sides, which is much more reproducible. Bradtmiller and Gross also took traditional measurements of the back width with the arms held away from the body, to give a more appropriate correspondence between the manual and scan-derived data [BrGr99].

The arm position also affects the detection of acromion and axilla landmarks. If the arms are held too far away from the body, and the natural shape of the shoulders becomes less well defined, then the detector tends to find the feature at the top of the arm, rather than on the shoulder itself. If the arms are held too closely to the body, then the axilla landmarks tend to be detected too low down the body.

Postural disorders were briefly mentioned in Chapter 2. Such cases are not uncommon in the sample data, and result in significant twisting laterally, dorsally, and about the vertical axis. Differences in shape of the torso can also arise simply from changes in posture in a healthy person. This is illustrated in figure 6.17, where top views are shown for a subject whose body twists slightly about the vertical axis. In this subject the nipple landmarks were not detected correctly. The twist to the torso could account for the failure of the detection method, which uses turning points in the radial distance function around the centroid.

Differences in posture are of particular relevance in detecting contours that are defined in the horizontal plane. For example, the waist can be defined as being parallel with the floor. However, if the left and right iliac crests are not level, a large overestimate in the true waist girth is obtained. With human observers using a tape measure, it is possible that an unconscious compensation is made for small postural deviations such as this, whereas the software measures *exactly* parallel with the floor if it is programmed to do so. A significant improvement might be obtained by allowing the automated system some degree of freedom in the sampling plane for dimensions such as the waist, even if the specification defines it to be exactly horizontal. The neck girth is found by locating the minimum girth in the neck region, allowing two degrees of freedom: translation along the neck axis, and a pitching rotation about a pivot on that axis. This

works well if the subject holds their head with approximate lateral symmetry, but if there is a significant shift left or right, then the girth will not lie on a minimum. If the degrees of freedom are increased, to allow a rolling rotation, then it has a better chance of success, however, this does increase the computational requirements.

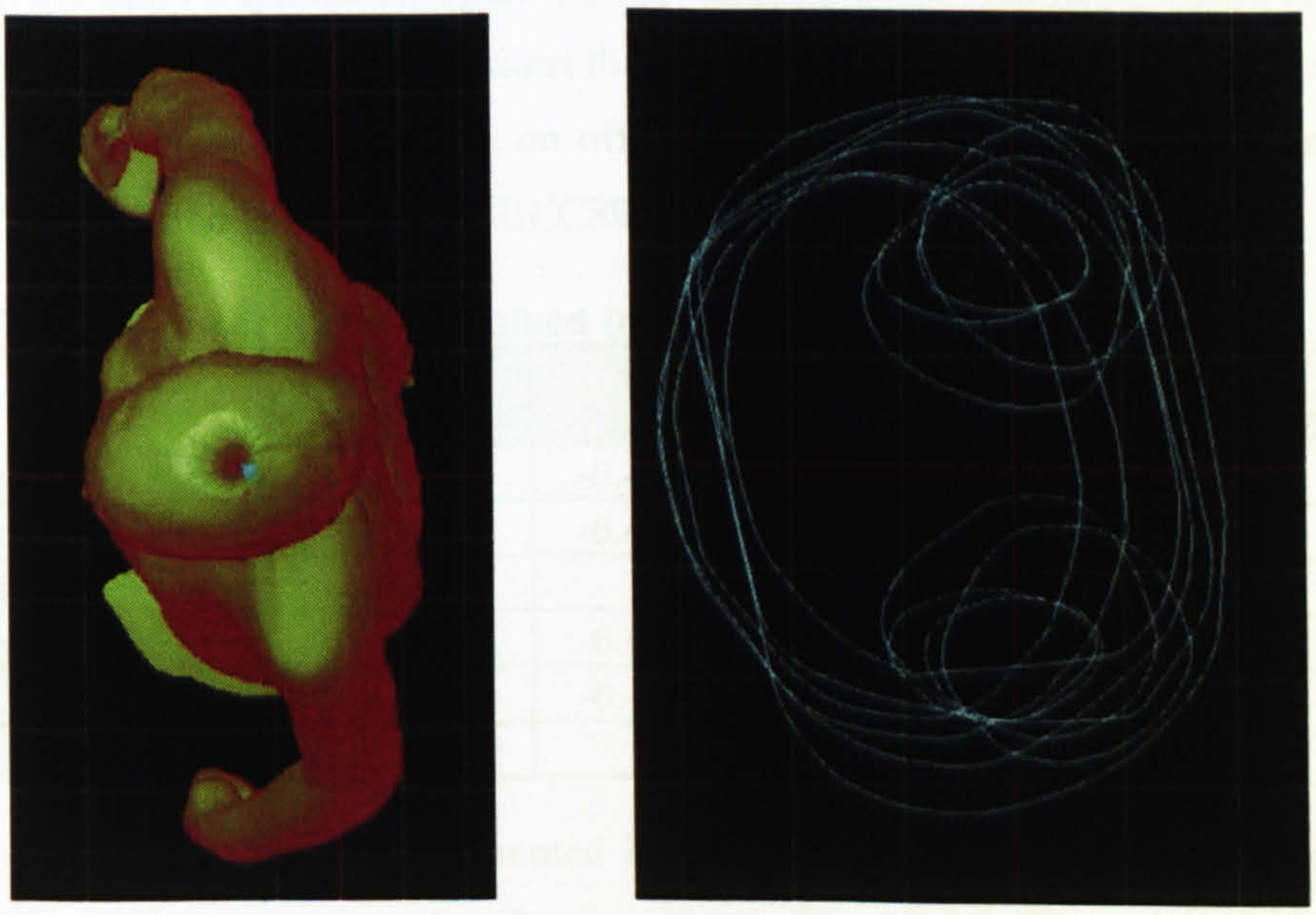


Figure 6.17: View from above, showing twist about the vertical axis.
(a) Surface model, (b) horizontal convex hulls on girths at intervals over the body.

Inter-observer and inter-technique differences have been discussed in section 6.5.3. On many dimensions a slight variation in the procedure does not result in a large difference on the measurement. For example, on many subjects there is a range of a few centimetres on the full hip area where a similar girth measurement is found, and therefore this dimension is not strongly sensitive to the exact vertical location at which it is measured. The upper hip, on the other hand, is located at a part of the body where the girth changes significantly, depending on the vertical location at which it is measured. Not surprisingly, this dimension has a very high RMS error in the automated system.

6.5.5 Acceptability Measures

The discussion so far has not addressed the acceptability of the results for any particular application. Of particular interest is how the results on the fully automated system relate to the requirements of the applications outlined at the very start of this thesis, in figure 1.1 on page 13.

Specific numeric limits on acceptability have been proposed by Bradtmilller and Gross [BrGr99], who discuss three sets of criteria:

- values based on inter-observer error on manual measurements from the 1988 Anthropometric Survey of US Army Personnel (ANSUR) [GBCC88],
- expert tailors' criteria,

- grading between sizes on main dimensions for off-the-peg clothing.

In addition, ISO 8559 states acceptable error levels as 1% or 5 mm, whichever is lower [ISO89].

A sample set of values on women's measurements for these different criteria are shown in table 6.3. Note that the tailor's quoted acceptable error values are not symmetric about zero, since an underestimate has more serious implications than an overestimate. The grading values show half the size difference between dimensions on off-the-peg clothing. The fourth column shows the acceptability levels listed in ANSUR [GBCC89].

Table 6.3: Acceptability values (mm) for measurements on women.

	ISO 8559 [ISO89]	tailor's [BrGr99]	grading [BrGr99]	ANSUR [GBCC89]
bust girth	5	-6.4 to +12.7	25.4	15
crotch height	5	-6.4 to +12.7	50.8	10
neck girth	3.5	±6.4	12.7	6
hip girth	5	-6.4 to +12.7	25.4	12
waist girth	4	-6.4 to +12.7	25.4	12
mean	**5**	**9**	**28**	**11**

Figure 6.18 summarises the results presented in tables 6.2 and 6.3. Each horizontal line shows the mean of the difference values for the five listed dimensions. The vertical lines show the quoted acceptability levels for errors.

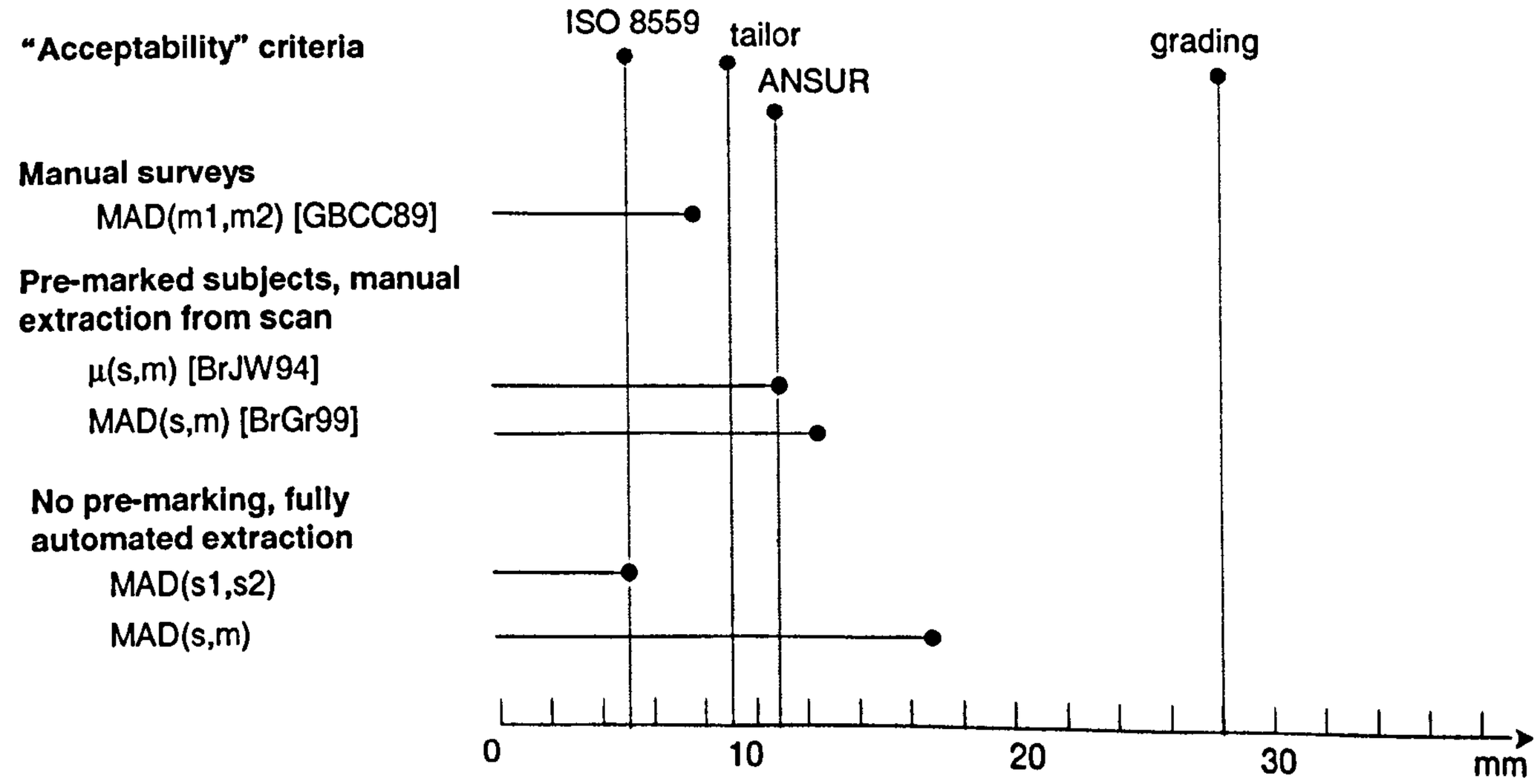


Figure 6.18: Summary of mean differences and acceptability levels.
s = scan, m = manual

The implication of the ISO 8559, tailor's and grading criteria is that these values are the *highest* acceptable error levels. However, since it is the *discrepancy* between manual and automated techniques that is being assessed, the confidence in these measures cannot be higher than the precision of the reference data itself. The ISO 8559 highest acceptable error values are, in fact,

much smaller than the *mean* inter-observer differences recorded for skilled measurers on several manual measurements by Gordon *et al* [GBCC89]. This suggests that the ISO error level is not achieved by skilled manual measurers, and therefore it is not even possible to *measure* errors to the levels prescribed by the ISO standard using *in vivo* measurements as reference data. However, the reproducibility of the automated system does appear to be close to this range, which is very encouraging.

Only a limited set of common results is available for the other studies, so this comparison only gives an approximate indication of the order of errors obtained by the different methods. A more thorough comparison would be interesting if more data becomes available in the future. In particular, it would be interesting to compare against another fully automated system, although at present it has not been possible to gain access to such data.

Intuitively it would seem that the strictest criteria would be needed for made-to-measure clothing, and this would need to be considerably better than the grading differences between sizes for off-the-peg clothing. However, if a person's whole set of individual measurements is used for a made-to-measure garment, each dimension contributes to the customised size of the garment. Therefore a good made-to-measure fit might be obtained even if the errors are similar to the grading differences.

In terms of prioritising the accuracy requirements, it would generally be expected that those dimensions that exhibit larger variation across the population would be more important to extract accurately by the automated method, while dimensions that have lower variation, such as the shoulder length, might be estimated from mean population values. Figure 6.13 shows the RMS error as a function of percentage variation on a given dimension, and no particular difference is found between the RMS on dimensions with low or high variation. However, the priority of accuracy clearly also varies across applications, each of which may have different critical dimensions.

One of the important attributes of the applications, illustrated in figure 1.1 on page 13, is the presence or not of system supervision. Worst-case error levels for each dimension are especially important if the system is to be used without any human intervention to check the validity of results. Paradoxically, large errors might not be problematic, as they would be easily noticed by an operator, or a means might be devised to detect them automatically for an unsupervised system. However, smaller errors would be more difficult to detect. Similar issues arise in manual methods, where transcription and arithmetic errors tend to result in large and obvious error values, but slight differences in procedure result in small errors.

6.5.6 Conclusions

This section has discussed many of the sources of discrepancy between manual measurements and the automated system, and the implications on acceptability of the results.

The sets of data used in this work are fairly small (18 subjects for training and 18 for validation), so the results must be considered in this light. The adult female subjects are taken from a range of low to medium body mass index, and a wide range of heights, and the results on various metrics appear to be consistent across this range, with no apparent degradation in performance at the extremes of the range.

The use of traditional measurement methods for collecting reference data raises several problems. Firstly, the acknowledged inter-observer error for even skilled measurers means that comparisons are being made against data of limited precision. Fortunately, such inter-observer error is well documented, and so it is possible to make an estimate of the precision of the reference data. 75% of discrepancies between the traditional measurements and the output of the automated system are less than 25 mm. This value has been chosen here as being significant in terms of garment grading, and therefore places the results in the context of target applications. Reduction of these worst-case errors is required if the system is to be used unsupervised for applications requiring high accuracy.

Work has been carried out by others to extract measurements from 2D and 3D scans, however, little has been published to date [BrJW94] [Daan98] [BrGr99]. In those studies artificial markers were placed on subjects prior to scanning and measurements were extracted manually from the scans. When compared against traditional measurements, the errors are typically higher than inter-observer error by human measurers, and similar to the automated system. This is especially significant, because the system described in this thesis is fully automated, whereas the procedures used in related studies required considerable human intervention.

When automated measurements from multiple scans of the same subject are compared, the results are similar to the inter-observer error on traditional measurements. This has implications in attempting to reduce the error on the automated system, which is one of the most important areas for immediate future work. Bland-Altman plots of the differences between manual and automated measurements for each dimension indicate that some improvement could be made by adding a correction factor to each measurement obtained by the automated system. The correction factor might be as simple as the mean difference between the automated and manual measurements for each dimension. This would not reduce the variance of the differences, but could reduce the RMS error obtained.

Many potential sources of discrepancy between traditional anthropometry and non-contact methods have been discussed. Errors may arise because of occlusion, deformation of loose fabric, or soft body tissue. However, the most important difference appears to be the sensitivity of the observed measurements to body posture [BrJW94] [BrDR97] [DKWB98] [BrGr99]. These issues arise not only in extraction of measurements from 3D scan data, but also for manual measurement. Posture is of particular significance for anthropometry using opto-electronic systems, as the subject is captured in a given stance for any one image, and the

requirements of anthropometry must be weighed up against the quality of the surface image in that stance. In this work subjects were assumed to be standing in an approximate reference position, but within that constraint they were permitted a considerable amount of freedom, which is important if the system is to be used unsupervised.

In order to deal more effectively with variations in posture, significant improvements might be made by allowing the detector functions a greater degree of freedom. For example, when detecting the waist girth, better results might be obtained if it is assumed that the waist might lie in any plane. Relaxing such constraints on search will increase the computational time required, but given the speed of the current system (45 seconds to carry out all the processing), it would be justified if such key features can be located more accurately and robustly.

6.6 Summary

Using the library of operators created in Chapter 5, a set of automated feature detection functions for surface anthropometry has been developed and tested against reference data collected manually on women. The results generally show good performance, with consistent behaviour over a range of subjects' size and shape, and consistent behaviour over dimensions that exhibit large and small variability. Although the reproducibility of the automated system is very good, it was noted that the worst-case errors are high for some dimensions, and certain dimensions are sensitive to posture.

Chapter 7

Assessment

This chapter provides a critical assessment of the work presented in this thesis, firstly in terms of the overall approach to modelling, feature detection, its implementation and testing procedures. This is followed by an assessment of the automated system itself, according to the criteria set down at the end of Chapter 1 and, finally, the research contribution made by this work is presented.

7.1 Research Goals

The goals of this work were outlined at the start of the thesis, in section 1.2 of Chapter 1, and are worth restating here. The aims were to develop techniques for processing opto-electronic range images of the human body, and to integrate these techniques into an automated system, which could be applied to problems in anthropometry. The work that has been presented in this thesis goes some way towards achieving these goals. Throughout this work the aim has been to test and analyse potential sources of error, before moving on to the next stage of development. Therefore each stage has been assessed as it appears in sequence in the thesis. However, it is worthwhile to stand back and discuss the work critically as a whole, now that the system has been developed and tested.

7.2 Assessment of the Approach

This section discusses the overall approach to the problem, and critically assesses each stage of the work.

7.2.1 Surface Modelling

Chapter 2 presented some qualitative experiments to analyse the raw scan data, and the data was analysed quantitatively in Chapter 4. This made it possible to make use of attributes of the data, such as the observed sampling density, to design partially data-driven operators for cleaning and surface modelling. By addition of some prior knowledge of the image capture system, the subject's stance, and a generic model of the human body shape, these low-level processing and surface modelling operations were made more effective and more computationally efficient.

After some experimentation with simple, generic surface growing techniques, a radial expansion matrix representation was introduced. This is a constrained form of generalised cylinder, where each body segment is represented by a two-dimensional array of points, wrapped about the principal axis of that segment. This representation appears to be suitable for the human body, although some limitations were noted. For example, slight artefacts are introduced where the surface "fills in" certain subtle concavities, which appear on some subjects between the upper humerus and outer clavicle, between the trapezius muscle and clavicle, under the chin, and at the axillae where the arms join the torso. This representation has many advantages in subsequent processing, because of its regular structure. For example, the quadrilateral grid makes it possible to use traditional two-dimensional image processing operators, for example, for smoothing and edge detection.

In order to transform the raw range data into this representation, it must first be segmented. Operators to detect the main branch points of the body, at the crotch and underarms, were devised and tested on a range of subjects, including men, women and children. The leg branch point detector initially required the use of an external input, the subject's weight, as a parameter, but a surrogate for this value was introduced, by sampling convex hull lengths on the torso. This made it possible for the whole process to operate without intervention, while adapting to the varying body size and shape of the range of subjects. The branch points for the arms were more problematic, as they are very sensitive to the position of the arms. If the subject's arms are held too close to the body, then the branch points are detected too low down the torso.

The mesh structure is then fitted onto the segmented data. In order to impose the regular structure, data must be pruned from regions that are densely sampled, and empty elements of the mesh must be filled in. A simple, computationally efficient interpolation technique was devised that makes use of knowledge about the image capture process and the predicted areas of data loss. The technique works well on small areas of missing data, but generates a "lathed" effect over large gaps in the data, for example on the inner surfaces of the legs. This arises because limited use is made of vertical coherence. Consequently, a more sophisticated interpolation technique might be appropriate over such areas, with a potential increase in computational complexity.

The smoothing operators appear to produce acceptable results at relatively low (between two and five) iterations, although no systematic tests were carried out to validate this process independently of the other surface generation processes. Some loss of useful detail, such as the edges of clothing, was observed because the operators used are isotropic. While this does not appear to affect the subsequent feature detection processes, it may be of importance for visualisation. An anisotropic version was proposed, whereby the isotropically smoothed surfaces are processed to extract curvature values; the surface is "desmoothed", and the curvature values are then used to control the iterations of the smoothing operator. In an ideal system an appropriate smoothing level would be obtained in a data-driven manner for each part of the surface. In this system the levels for each body segment were selected empirically. If the system is to be adapted for independence from the source of the range data, it will be necessary to make the smoothing fully data driven, to deal with differences in density, error levels, and so on.

The discrete polygon mesh representation has been justified for many reasons, but it does have some shortcomings. A more suitable form might be a spline surface, which allows direct calculation of derivatives. Another benefit of spline-fitting is that considerable compression—of the order of 1000 times—is possible. Previous work on spline representations of the human body has been done by Li and Jones [LiJo94] [Li97], and concurrent work has been carried out by West and Douros [West97] [Dour98]. West's spline fitting process provides a more conceptually straightforward solution to some of the cleaning and redundancy problems described in Chapter 4, but the iterative nature is computationally expensive, and therefore inappropriate for this work. Douros has subsequently built on this work, developing a fully closed, whole-body form, composed of spline patches with continuity and smoothness constraints imposed across the segment boundaries [DoDB99].

7.2.2 Shape Analysis and Feature Detection

Chapter 5 investigated techniques to extract size and shape characteristics of the polygonal mesh body surface, from high-level relational attributes, to local attributes such as curvature. Firstly some aspects of high-level structure were explored. By generating the central axes of each body segment, relational attributes such as front, back, left and right can be applied, with respect to the axes. This is of particular relevance in constraining the search for features and improving the computational efficiency. As features such as point landmarks and contours on the surface of the body are discovered, they are used to further constrain the search for other features. In this way the body surface is gradually mapped with an increasingly rich set of information.

In addition to the use of higher-level structure, local differential geometry was used as the main means of controlling the search for surface features. Curvature-based approaches have been used successfully in the past for feature detection in other domains [AsBr86] [PoBr87]

[LeTW95]. The segmented matrix representation of the body was shown to be convenient for extracting the differential attributes, as the two-dimensional format makes it possible to use versions of traditional image processing operators, for example, for gradient estimation. In addition, a computationally efficient means of calculating curvature was devised, which does not rely on analytical function fitting, or the use of surface normals, as required by most other approaches. The method devised here for curvature estimation is not a general-purpose one, however, the assumptions that it makes are shown to be valid within the modelling system used here for the human body.

Contours are fundamental to surface anthropometry. Chapter 5 presented a set of techniques for generating surface contours, based on a minimal set of control points, to simulate the behaviour of a tape measure. Although the main motivation was for extracting measurements within the fully *automated* system, attention was also paid to the interactive use of the "electronic tape measure", to make it a flexible tool for experimentation. In order to achieve computational efficiency, certain assumptions were made, such as planarity of the contours. Tests were carried out to compare the electronic tape measure with measurements taken using traditional methods, and a good correspondence was found, with only 1% difference in the large majority of cases. If only two control points are used to define a contour, the plane is under-defined. Occasionally the plane is inferred incorrectly, resulting in large errors. This problem needs to be addressed before that particular feature of the electronic tape measure can be used reliably.

These generic operators to extract local and global size and shape characteristics were then brought together into a feature detector library, which was designed to be compact yet expressive. In Chapter 6 a systematic method was described to use this library to develop detector functions for specific anthropometric features. Although the feature detection processes make certain assumptions about the subject's posture, there is no assumption of left/right symmetry. This degree of freedom is important in extracting meaningful information on *individuals*, rather than using an idealised view of the human body.

The functional form of each detector was designed by hand, but the parameters within the detectors were optimised automatically, by sequential exhaustive search. Because of the non-linear relationship between features, the performance of the resulting detector functions is almost certainly affected by the order in which each parameter was optimised. Therefore an obvious improvement would be to optimise inter-dependent features together, using, for example, a stochastic search technique in an attempt to deal with the extremely large search space [KiDe95] [DeKi96]. However, the longer-term goal is to use the feature detection library as the input primitives for a full learning system, for example, using genetic programming. The results for the interactively built functions described here, and some prior work in genetic programming and machine vision suggests that this might be an achievable goal [Poli96] [HaBu97].

Although the function-based approach has been the focus of this work, some other potentially interesting approaches have emerged. By reducing the three-dimensional body surface to a two-dimensional matrix representation, there appear to be many possibilities for exploring shape, which have perhaps been under-exploited in this work. One example was presented in which a template-based technique was used to detect a feature (the navel), using a local matrix of curvature values. However, other possibilities include the use of statistical techniques for *global* shape characterisation, based on the whole-body or body-segment matrices of curvature.

7.2.3 Performance Testing

In testing the automated system, the first issue is how to collect reference data against which to compare the output. As the system is to be used on living people, who breathe and move and adopt a range of postures, it must also be tested on living subjects.

In Chapter 4 the processes from raw data to the generation of whole-body surfaces were validated by comparing the enclosed volumes with existing state-of-the-art volumetric methods *in vivo*. Adult and child subjects were scanned and then measured by underwater weighing and air plethysmography [FJLC92] [MGBM95]. Results showed close correspondence between the three techniques, with a slightly higher discrepancy between the scan-derived measurements and the other two methods. Although the differences were small—less than two litres in all but two cases—the precision of the scan-derived method is not sufficient to be used in determining body fat, which is one of its potential applications. The differences might easily be attributed to the variable lung volume held by subjects during scanning, a point noted previously by [SmJW90]. Pilot tests have been carried out to assess the effects of lung volume on the scan-derived volumes [WDFE99], indicating better precision, but still some discrepancies, which are not yet resolved.

Although this method of assessment provided a general indication of the validity of the processing techniques, it is clearly a rather insensitive method. More convincing validation data might be derived using realistic humanoid phantoms for volume estimation, and with more extensive investigation of other measurements on live subjects, such as surface contour lengths, and so on. Chapter 6 presents a more detailed set of experiments on surface anthropometry.

Several approaches were discussed for collecting reference data for surface anthropometry, each of which appeared to have serious difficulties. The method used here was to compare the surface length measurements of the system against measurements taken by traditional methods. As indicated, this approach has some problems, notably that manual measurements are known to be prone to intra- and inter-observer error. This puts limitations on the precision at which errors on the automated system can be measured. Extensive work has been carried out by other researchers to assess such variations in traditional anthropometric procedures, so a good deal of data is available, which indicates a consensus on the expected inter-observer differences for

trained measurers [GBCC89] [BrJW94] [Daan98] [BrGr99]. These statistics provide a useful benchmark against which to test the performance of the automated system.

In addition, some limited information is available from previous studies where subjects were scanned with pre-placed markers at key points, and measurements were extracted by manually locating these markers on the processed images [BrJW94] [Daan98] [BrGr99].

Results obtained by the automated system developed in this thesis were presented for two sets of subjects. The first set of data corresponds to the hand measured and scanned subjects, whose manual measurements were used to train the automated system. The second set of results is for a validation set of subjects, who were also hand measured, and in each case two scans in the reference posture were taken. Results obtained from each scan were compared against the manual data. Results from the two scans were also compared against each other to measure the reproducibility of the automated system. Acceptability criteria were discussed to enable the significance of the results to be analysed, in terms of the potential applications of the automated system.

Differences between the manual and scan-derived data were analysed using several metrics. The absolute difference between the values obtained by the two techniques showed close agreement, and in most cases were within the range of grading steps between clothing sizes. However, some higher-magnitude discrepancies were also observed, which need to be addressed. In terms of reproducibility, the system appears to perform extremely well, with results similar to those recorded for inter-observer differences on purely manual measurements [GBCC89] [BrJW94] [Daan98] [BrGr99].

In addition, performance was analysed in terms of robustness over the range of body mass indices (BMI) in the sample set, and robustness over the range of different dimensions. In each case the performance appeared to be consistent across the range of BMI, with no observable degradation in performance. However, the experiments were carried out on a fairly small number of subjects (18 women of low to medium BMI in each set). The consistent results suggest that there should be good confidence of the system's performance on other similar subjects, and possibly on subjects just outside of the range tested. However, the results do not imply any expectation of the system's behaviour on children and men, or on women of very high BMI or very unusual body shape.

The comparison of the automated system with *in vivo* measurement has confirmed some problems that have been noted before [BrJW94] [BrDR97] [BrGr99]. An extremely important issue is posture, and reconciling the difficulties of capturing good quality 3D data with anthropometric requirements. In this work an attempt has been made to extract as many measurements as possible from a single scan, with the not unexpected result that some dimensions, such as the back width, show consistently high errors.

7.3 System Assessment

At the end of Chapter 1 some assessment criteria were defined for the automated system. Each of these will now be discussed to evaluate how well the system has met these objectives.

7.3.1 Minimal Intervention

In order for the system to be widely adopted by potential users, it is necessary that it is straightforward to process subjects and to extract the information of interest. The system developed here requires subjects to adopt an approximate reference posture, but within that posture a range of variation is permitted. No pre-marking of the subject and no special clothing are required, although better results are obtained if close-fitting clothes are worn, and a swimming hat is used, to hold the hair away from the neck and to flatten the hair to reveal the true stature. Once the subject as been scanned, the image is downloaded by a single mouse click, and another mouse click sets the automated software running, to clean the data, reconstruct the skin surface and extract the anthropometric information.

Thus the system would appear to meet the minimum intervention criterion very well. However, some limitations of the system, arising from sensitivity to posture, were noted. It would be interesting to explore what improvements could be made to allow the whole hardware and software system to be used by untrained subjects to capture their own 3D scan and extract their own body metrics unaided.

7.3.2 Geometric Accuracy

Volume and surface anthropometry measurements show results to within a few per cent of the reference data in most cases. For a machine vision system to work this well on an object as complicated and variable as the human body, with relatively little control of the image capture conditions, is an encouraging result. However, the applications tackled here are extremely demanding, and many potential users of such data would still consider these results to be unsatisfactory.

For estimation of body fat content the volume measurements must have less than 1% error [WDFE99], and therefore further improvements would need to be made for the system to be used successfully in that application. It is possible that loss of data from occlusion is a limitation, and better results might be obtained from raw data with better coverage, especially on the inner surfaces of the legs. However, it is also possible that the error might be reduced by improving the surface reconstruction processes, using the existing range data.

The results for surface anthropometry appear to be more immediately useful for the target applications. Most of the dimensions measured show errors below the grading steps between garment sizes, although the worst-case errors on some dimensions are unacceptably high. At

many stages, conscious decisions have been made to trade off accuracy for tractability, to enable the system to run with sufficient speed so as to be practical. Given the high error on a few critical dimensions such as the waist, it would seem necessary to put more computational resources into these key detector functions. For example, for the waist, it should be possible to obtain significantly better results by allowing the search function more degrees of freedom, to search in a greater range of planes, and possibly to use *non-planar* contours.

7.3.3 Occam's Razor

In order for the system to be practical, speed is an important design consideration. To achieve this, computational efficiency is crucial. As noted above, conscious trade-offs have been made in many of the processes to achieve this. The overall approach in designing each processing component is to use the simplest technique that produces acceptable results. Often this has required making simplifying assumptions, but the aim has been to make these decisions only where the performance, in terms of accuracy, robustness and so on, would not be compromised.

The resulting system is thus quite fast, taking only 60 seconds or so to scan the subject, access the raw data, clean it, reconstruct the surface and extract the anthropometric information, using only a moderately powerful computer (Dell Optiplex, 400 MHz Pentium II, 128 MB RAM). This is fast enough to allow for immediate assessment of a subject's scan, and to carry out repeat scans if necessary. It is also fast enough to process large-scale surveys. In fact, given that there are some key areas of processing, such as surface interpolation, that require improvement, it would be appropriate to experiment with alternative processing techniques that might require increased processing time, but would provide benefits in terms of geometric accuracy and robustness.

7.3.4 Reproducibility

Given the number of issues raised about comparisons of traditional and scan-derived measurements, it is especially important to investigate the system's reproducibility outside of such inter-technique variation. All the subjects in the validation set were scanned twice, and the measurements extracted from different scans show extremely good agreement.

This is interesting for several reasons. It was noted that some dimensions are strongly sensitive to posture, but some of these dimensions have very small differences across scans. This suggests that it might be possible to improve some of the measurement extraction processes by taking account of systematic errors. Perhaps most noteworthy, however, is that the reproducibility results for the automated system are similar to the inter-observer error recorded in studies on skilled measurers [BrJW94] [BrGr99].

The results presented here used scans taken within a few minutes of each other. It would be interesting to look at data collected over a longer period of time, to investigate postural variation and sensitivity of the automated system under a wider range of conditions.

7.3.5 A Flexible, Working Software System

The methods developed in this work now form the basis of a commercial system. As noted above, the system is relatively simple and fast to use. As well as the fully automated processing, it provides interactive tools for tuning the operational parameters, diagnostics and manual extraction of anthropometric information, with a direct manipulation interface.

The system is designed in a component-based way, to enable each processing operation to be replaced as necessary by improved techniques, whether for reasons of computational efficiency, accuracy or robustness. Throughout this work the components have evolved, and will continue to do so in the future. The aim has been to isolate the scanner-dependent aspects in the low-level components, to increase the potential for use of the system with other image capture systems. The software has been tested on three different configurations from the same family of scanners, and several different instantiations of each of these scanner types. Results appear to be consistent across the different systems, despite variations in data resolution. It is expected that further adaptation of the system will be required for it to function effectively with other image-capture hardware, not just at the low-level, but also in some of the higher-level functions.

7.3.6 Broad Applicability

This criterion relates to the range of conditions under which the system can operate successfully. The early processing stages from data cleaning to segmentation and surface reconstruction were tested over a range of healthy subjects, including women, men and children. The segmentation and volumetric results appear to show consistent error levels across all these subjects.

The feature detection and surface anthropometry functions have been designed specifically for women. Therefore they have only been tested on this target group, and no claim is made on this part of the system's performance outside of this range. The system has been trained and tested using data collected from subjects with BMIs ranging from low, to medium for women. Results indicate consistent behaviour over this set of subjects, with no observable degradation in performance at either extreme. No heavily overweight subjects were tested, so it would be interesting to test the system on a wider range of subjects in order to find its limits.

7.3.7 Emergent Properties

This assessment criterion has been added in the acknowledgement that the success of the work does not come entirely from the planned and expected outcomes, but also from unexpected

results and the potential it presents for future work. One aspect of the work that was noted as being under exploited was the use of the shape maps, and the transformation of the body into a regular two-dimensional structure. Visualisation and quantification of this information has numerous implications, not just for the applications addressed in this work, but for sculptural and design-related fields. For example, the curvature maps reveal interesting information for understanding subtle aspects of human body shape.

7.4 Research Contribution

The previous sections have discussed the work in terms of the research approach, and the performance of the resulting system that has been built. This section now discusses the thesis in terms of its five main research contributions.

Firstly, the thesis describes an automated system for processing whole-body range images for anthropometry. This is arguably the first fully integrated and automated system of its type, and the first reported in detail in the public domain. The development of the system is described from scratch, starting at the image-capture hardware and an assessment of the data available. Each of the processing operations is then described, to take the raw image, clean it, reconstruct the surface and extract the anthropometric information of interest. At each stage the rationale behind the design is discussed, and experiments are described to test the validity of the processes.

Secondly, a new, combined data-driven and model-based approach is described for surface reconstruction. This fully automated method is designed specifically for human-body range data in a "natural" reference posture. It addresses the *whole* body surface, including areas where body segments touch, and other occluded areas. Experiments have been carried out by volumetric measurement on a range of body shapes and sizes, and the processing techniques are demonstrated to be robust for women, men and children. Some limitations are noted, such as the slight rippling artefacts on the inner surfaces of the legs where large areas of missing data must be reconstructed, and this is flagged as an area requiring improvement.

Thirdly, a library of computationally efficient operators is developed, specifically for shape description and measurement of human-body surface models. The aim is to provide the necessary set of primitives with which to construct detectors for specific anthropometric features, either by manual construction or, in the future, by machine learning methods. The library provides simple high-level relational attributes, an "electronic tape measure" to extract linear and curvilinear measurements, as well as access to low-level shape information, such as curvature values at individual vertices on the polygon mesh surface.

Fourthly, application of the library is demonstrated by building a set of detectors to find a range of anthropometric features on the body, based on the ISO 8559 specification. These detectors

are integrated into an automated system for processing human body scans. Traditional manual measurements are used as reference data to measure the performance of the automated system. A detailed analysis of the results is presented, with comparisons to similar studies. In particular, inter- and intra-technique discrepancies are discussed, and the results are illustrated graphically in figure 6.13. The results are then assessed in the context of appropriate acceptability criteria and the needs of specific applications of the automated system.

Finally, although this thesis has focused on some specific problems in anthropometry, a side-effect of the work has been to facilitate research by others in digital human modelling. As such the system has become a basic tool for research problems, such as registration of multiple body images, draping clothing and advanced surface modelling techniques [DoDB99].

7.5 Summary

This chapter has provided an assessment of the work presented in this thesis. Firstly, the initial research goals, stated in Chapter 1 were revisited, in order to place the assessment in context. The second section presented an assessment of the work in terms of the overall research approach, and how the particular sub-problems have been tackled. This was followed by a brief discussion of each of the assessment criteria defined in Chapter 1, which focused mainly on the performance of the automated system developed in this work. The final section looked at the work in terms of its contribution to the research community.

Chapter 8

Conclusions and Future Work

This final chapter begins by drawing some final conclusions from the work described in this thesis, in the broader context of 3D human body modelling. The chapter concludes by outlining plans for specific future work.

8.1 Conclusions

At the start of this work the aim was to develop automated techniques to interpret surface information and detect key features on the human body. In practice, a much larger part of the work than expected has been spent on developing low-level processing techniques, before this higher-level work could be done, as no appropriate methods were available. In particular, no technique existed to take the raw range data and reconstruct the whole body surface, taking into account occluded areas, in an automated way.

The processing method developed in this work appears to be effective on the women, men and children on which it was tested, although it was not tested on subjects of extreme physique. Many improvements could be made, especially to the surface restoration across large occluded areas. Also, some decisions about parameterisation were left partially unsolved. For example, although most of the low-level processing operators were data-driven, the smoothing level was chosen empirically; however, a method has been proposed about how to address this, in section 5.3.3. Although subjects are scanned in a relatively natural and relaxed posture, several assumptions are made in these stages of processing, which constrain the system's use. It would be interesting to explore ways in which some of these conditions could be relaxed, to broaden the applicability, for example, to subjects in a wider range of postures, or even wearing clothing. Several other applications were discussed, in Chapter 1 and again in Chapter 6, when the integrated system was evaluated. One of the most interesting problems to address now is how to make the system fully reliable so that people can scan themselves and gain access to their own

body information. This is of particular importance, given the rapid expansion of Internet applications, games, and other virtual environments, which require personalised models, often with very high geometric accuracy or visual realism, at low computational costs [OlBu99] [HBGS99] [HAWG99] [Smit99]. Of course, this would almost certainly require more sophisticated and computationally intensive techniques. Given that the system developed here is relatively fast, it would seem to be quite acceptable to trade off some of this speed for improved accuracy and robustness over a wider range of conditions.

The overall approach to feature detection on the surface models appears to be a valid one, but many other possible approaches were discussed, which would be interesting to investigate. A pressing requirement is to integrate the approach with a learning mechanism. This would enable the detection of new features, and application to other body types not addressed in this work. Nevertheless, the library of feature detection operators provides a useful starting point for function-synthesis approaches, such as those explored by [Koza92] [Andr94] [Poli96] [HaBu97] [Drap96].

Although the results for the automated feature detection system appear to be fairly good, one argument that might legitimately be raised is, why extract one-dimensional metrics at all, given the richness of the three-dimensional information available? In addition, why use the skeletal features that are difficult to detect in this medium? Whole-body surface scanning is a relatively new technology, and for the clothing industry in particular, a vast amount of capital and knowledge has been invested in CAD/CAM systems that make use of traditional data. If 3D technology is to be adopted, it must be done in stages, so this work presents a contribution to making that possible. A great deal of other, truly three-dimensional information could be extracted, using the framework presented here. Examples include the spatial relationships of landmarks and characteristics of contours on the surface of the body. Li has already carried out work to investigate this higher-dimensional information, using shape vectors to characterise cross sections [Li97]. A great deal of work has also been carried out in shape analysis in other domains, such as medicine and taxonomy, and this should be equally applicable here [AsBr86] [PoBr87] [Book91] [CMLR91] [CTCG95] [AyMa96].

Given the many new whole-body surface imaging systems appearing, as outlined in Chapter 2, the ability to process data from a range of sources in a standardised way is extremely important. Of particular interest are the very low-cost systems being designed, such as the ones described by Hilton *et al* [HBGS99] and Certain and Stuetzle [CeSt99], and emerging techniques, such as Fisher's "magic wand" [FARW99]. Each system produces data with various levels of resolution, precision, noise, occluded areas, colour information, and so on. In order to process these effectively the system developed here would need some adaptation, both in the low-level processing, and at a higher-level. Those aspects that are dependent on the image-capture system would ideally be isolated in a low-level component, which would generate output in a standard

representation. The higher-level components would then take as input this standard format. The prime candidate for this would be the emerging "seamless avatar" standard being prepared by the *H-Anim* working group [Smit99]. This has the advantage of also tying in with animation standards.

Over the recent years there has been a marked change in the expectations of what can be done with 3D data of the human body. Potential users of the data are more aware of the limitations of the image capture systems and of the processing techniques. There is also a better understanding of the potential of the three-dimensional aspects, and how these could be exploited in a practical way [Li97] [JoRi97] [Gray98] [Tait98] [JoLi98] [RoDP99], rather than treating such systems simply as automated tape measures. Perhaps the most untapped potential comes from the ability to capture, archive and process large-scale longitudinal data sets, rapidly and with minimal intrusion. Together with recent advances in data mining techniques, this presents enormous potential in extracting patterns and testing hypotheses for epidemiological studies.

8.2 Future Work

The conclusions in the previous section highlighted many areas for future work. Some of these ideas are already being realised in new collaborative work. One such activity is the *3D Centre for Electronic Commerce*, funded jointly by the UK government and industry. This initiative has four main strands:

- a large-scale 3D survey of the UK population, carried out at several geographical locations;
- integrated systems to take body metrics from 3D scans and use them to make customised clothing;
- virtual clothes shopping applications, enabling people to try on clothes virtually, over the Internet and in-store;
- 3D infrastructure for these applications: a standard body representation to enable processing software to become independent of the image-capture system, automated image interpretation techniques, and secure Internet mechanisms for browsing and storing body data.

This on-going work has four main implications for future work arising out of this thesis:

- availability of training data for machine learning;
- investigation of three-dimensional aspects of anthropometry;
- relaxing operating conditions for the techniques and improving reliability;
- standardisation and links to other image-capture systems.

Most significantly for the work presented in this thesis is the eventual access to large sets of data, from a varied range of body types, including men and children. Lack of data has been a major limiting factor in this work, and the availability of good training sets will make it possible to investigate a wider range of techniques, in particular, the automated function synthesis approaches discussed earlier. Deformable template methods were also discussed as potentially useful methods [CTCG95] [MoDe97]. These could be applied to the body data in three dimensions, or using the two-dimensional matrix representation of the body. The second approach has many attractive features because of its relative simplicity.

One of the areas that was only explored very briefly was the extraction and quantification of two- and three-dimensional attributes of the human body. This is not just of interest because of the richness of the data available from 3D scans, but also because there is clearly a demand for this type of information in the potential application areas. In particular, it is proposed to investigate the use of curvature information and the spatial relationships of features on the body, in immediate future work.

Given the many and increasing numbers of potential applications of 3D body data, an important part of the future work is to relax some of the constraints and improve the reliability of the processing techniques developed in this thesis. One topic that was explored a little, in the work on segmentation and surface reconstruction, is automatically deriving confidence measures in the output of the system. It would be interesting to explore this further, to infer the body posture during scanning, and apply the appropriate surface reconstruction and feature detection techniques. The availability of more training data will be extremely useful in all of this work.

Another key activity is standardisation. Of particular importance is the work on developing techniques that are independent of the image-capture hardware, and integration with standards such as *H-Anim* [HAWG99]. The techniques developed in this thesis will be adapted to make use of such standard representations, which will then open up links to data from the many new 3D image-capture systems appearing. There would then be a complete and automated link from the scanned images to articulated and animated forms.

8.3 The Broader Context

At the start of this thesis some of the many applications of digital human body models were discussed. This work has provided an ideal opportunity to discuss the potential of these techniques with an enormously varied range of industrial technologists, designers, researchers, medical practitioners, artists and social scientists. It is apparent that 3D surface capture and processing are not just of relevance for their technical merits, but because they provide a whole new medium for communicating knowledge about the human form. This has been quickly recognised by artists and architects, who have been keen to integrate such data into two- and three-dimensional creative forms. In terms of personal and social perceptions, it is clear that

accurate 3D images of the human body have important implications for perception of real body shape, which, in turn, is of importance in such diverse areas as visual communication, diet, eating disorders, body composition, Internet shopping and postural diagnosis and monitoring. The ability to compare size and shape over time and perform predictive simulations would be of great benefit, not just because of the potential for numeric accuracy, but because of the ability to *visualise* the potential outcomes. The framework presented in this thesis goes some way towards automating the processing of surface images of the human body. This is just one step in this exciting, multi-disciplinary field.

References

[AcMa92] Ackerman MJ, Masys DR, "The Visible Human Project of the National Library of Medicine: The Need for Representation Standards for Volumetric Anatomy", *Proceedings of the Electronic Imaging of the Human Body Workshop*, CSERIAC, Dayton Ohio, 1992, pp. 77-87.

[ADBW96] Aono M, Denti P, Breen DE, Wozny MJ, "Fitting a Woven Cloth Model to a Curved Surface: Dart Insertion", *IEEE Computer Graphics and Applications*, 1996, pp. 60-69.

[AFRW98] Ashbrook AP, Fisher RB, Robertson C, Werghi N, "Finding Surface Correspondence for Object Recognition and Registration Using Pairwise Geometric Histograms", *Proceedings of the European Conference on Computer Vision*, 1998.

[Alle98] Allen R, *personal communication*, 1998.

[AmBK98] Amenta N, Bern M, Kamvysselis M, "A new Voronoi-based surface reconstruction algorithm", *SIGGRAPH '98*, ACM, 1998, pp. 415-421.

[AmBu99] Amos G, Buxton BF, "Positioning Human Body Models Using Lagrange Multipliers", *International Conference on Digital Human Modeling for Design and Engineering*, Society of Automotive Engineers, USA, 1999, 1991-01-1917.

[Andr94] Andre D, "Automatically Defined Features: The Simultaneous Evolution of 2-Dimensional Feature Detectors and an Algorithm for Using Them", *Advances in Genetic Programming*, MIT Press, 1994, pp. 477-494.

[AsBr86] Asada H, Brady M, "The Curvature Primal Sketch", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 8, 1986, pp. 2-14.

[AyMa96] Aykroyd RG, Mardia KV. "An MCMC Approach to Wavelet Warping", *Proceedings Image Fusion and Shape Variability Techniques*, Leeds University Press, 1996, pp. 129-140.

[BHG93] Badler NI, Hollick M, Granieri J, "Real-Time Control of a Virtual Human Using Minimal Sensors", *Presence*, Vol 2 No 1, 1993, pp. 82-86.

[Bato92] Batouche M, "A Knowledge Based System for Diagnosing Spinal Deformations: Moiré Pattern analysis and Interpretation", *Proceedings of the 11th International Conference on Pattern Recognition: Computer Vision and Applications*, IEEE Press, 1992, pp. 591-594.

[BaWe96] Bala J, Wechsler H, "Shape Analysis using Hybrid Learning", *Pattern Recognition*, Vol 29 No 8, Elsevier Science, 1996, pp, 1323-1333.

[BBCS99] Bernardini F, Bajaj CL, Chen J, Schikore DR, "Automatic Reconstruction of 3D CAD Models from Digital Scans", *International Journal on Computational Geometry and Applications*, Vol 9 Nos. 4-5, 1999, pp. 327.

[BDGW00] Bougourd JP, Dekker L, Grant-Ross P, Ward JP, "A Comparison of Women's Sizing by 3D Electronic Scanning and Traditional Anthropometry", *Journal of the Textile Institute, Part 2, Economics, Management and Marketing*, 2000, submitted.

[BeAc98] Beumier C, Acheroy M, "Automatic Face Authentication from 3D Surface", *Proceedings of the British Machine Vision Conference*, BMVA Press, 1998, pp. 449-458.

[Beau78] Beaudet PR, "Rotationally Invariant Image Operators", *Proceedings of the 4th International Conference on Pattern Recognition*, 1978, pp. 579-583.

[BeMc92] Besl PJ, McKay ND, "A Method for Registration of 3D Shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 14, 1992, pp. 239-256.

[Besl88] Besl PJ, *Surfaces in Range Image Understanding*, Springer-Verlag, New York, 1988.

[Besl89] Besl PJ, "Active Optical Range Imaging Sensors", in Sanz JLC (ed.) *Advances in Machine Vision*, Springer-Verlag, New York, 1989, pp. 1-63.

[BlA186] Bland JM, Altman DG, "Statistical methods for assessing agreement between two methods of clinical measurement", *Lancet*, Vol 1, 1986, pp. 307-10.

[BlBe97] Blais F, Beraldin J-A, "Calibration of an Anamorphic Laser Based 3D Range Sensor", *SPIE Proceedings, Videometrics V*, Vol 3174, SPIE, 1997, pp. 113-122.

[BoA197] Bounsaythip C, Alander J, "Genetic Algorithms in Image Processing – A Review", *Third Nordic Workshop on GAs and their Applications*, 1997, pp. 173-192.

[BoFi97] Borges DL, Fisher RB, "Class-Based Recognition of 3D Objects Represented by Volumetric Primitives", *Image and Vision Computing*, Vol 15, Elsevier Science, 1997, pp. 655-664.

[BoGV92] Boser B, Guyon I, Vapnik V, "A Training Algorithm for Optimal Margin Classifiers", *Fifth Annual Workshop on Computational Learning Theory*, ACM Press, 1992.

[Bois84] Boissonat J-D, "Geometric Structures for Three-Dimensional Shape Representation", *ACM Transactions on Graphics*, Vol 3 No 4, ACM, 1984, pp. 266-286.

[Book89] Bookstein FL, "Principal warps: thin-plate splines and the decomposition of deformations", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 11, 1989, pp. 567-585.

[Book91] Bookstein FL, *Morphometric Tools for Landmark Data*, Cambridge University Press, 1991.

[BrDR97] Brunsman MA, Daanen HAM, Robinette KM, "Optimal Postures and Positioning for Human Body Scanning", *Proceedings of the International Conference on Recent Advances in 3D Digital Imaging and Modeling*, Ottawa, 1997, pp. 266-274.

[BrGr99] Bradtmiller B, Gross ME, "3D Whole Body Scans: Measurement Extraction Software Validation", *International Conference on Digital Human Modeling for Design and Engineering*, Society of Automotive Engineers, USA, 1999, 1991-01-1892.

[BrJW94] Brooke-Wavell K, Jones PRM, West GM, "Reliability and Repeatability of 3-D Body Scanner (LASS) Measurements Compared to Anthropometry", *Annals of Human Biology*, Vol 21 No 6, Taylor & Francis Ltd, 1994, pp. 571-577.

[BrPo93] Brunelli R, Poggio T, "Face Recognition: Features versus Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 15 No 10, 1993, pp. 1042-1052.

[Bubb98] Bubb H, "From Measurement to Design and Evaluation: the Idea of the CAD-Tool RAMSIS", *Proceedings of the Workshop on 3D Anthropometry and Industrial Products Design*, Laboratoire d'Anthropologie Appliqué, Université Paris V, 1998, pp. 9/1-9/8.

[Burn97] Burnsides DB, "Software for Visualization, Analysis and Manipulation of Laser Scan Images", *Proceedings of SPIE 3023*, 1997, pp. 159-168.

[Burr81] Burr DJ, "Elastic Matching of Line Drawings", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 3 No 6, 1981, pp. 708-713.

[Cann86] Canny JF, "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 8, 1986, pp. 679-698.

[CaPB90] Case K, Porter JM, Bonney MC, "SAMMIE: A Man and Workplace Modelling System", *Computer-Aided Ergonomics*, Karowowski W *et al* (eds.), Taylor & Francis, London, 1990.

[Carr98] Carruthers D, *Conversion and Articulation of 3-Dimensional Body Scans*, MSc Thesis, University College London, 1998.

[CCRA96] Costen N, Craw I, Robertson G, Akamatsu S, "Automatic Face Recognition: What Representation?", *Proceedings of the European Conference on Computer Vision*, 1996, pp. 504-513.

[CeSt99] Certain A, Stuetzle W, "Automatic Body Measurement for Mass Customization of Garments", *Proceedings of the Second International Conference on 3-D Imaging and Modeling*, National Research Council of Canada, IEEE, 1999, pp. 405-412.

[ChDy86] Chin RT, Dyer CR, "Model-Based Recognition in Robot Vision", *Computing Surveys*, Vol 18 No 1, ACM, 1986, pp. 67-108.

[ChJo99] Cheyne T, Johnson G, "Shopping with Rei", *What + If: Research Journal of the NCR Knowledge Lab*, Autumn 1999, No 2, pp. 175-197.

[CiLM98] Cinque L, Levialdi S, Malizia A, "Shape description using cubic polynomial Bezier curves", *Pattern Recognition Letters*, Vol 19, Elsevier Science, 1998, pp. 821-828.

[CMLR91] Coombes AM, Moss JP, Linney AD, Richards R, James DR, "A mathematical method for the comparison of three-dimensional changes in the facial surface", *European Journal of Orthodontics*, Vol 13, 1991, pp. 95-110.

[CoHu98] Corner B, Hu A, "Effect of Sway on Image Fidelity in Whole Body Digitizing", *Proceedings of SPIE 3313*, 1998, pp. 90-99.

[CoSV94] Commean P, Smith K, Vannier M, "Automated Limb Prosthesis Design", *Visualization in Biomedical Computing, Proceedings of SPIE 2359*, 1994, pp. 493-503.

[CTCG95] Cootes TF, Taylor CJ, Cooper DH, Graham J, "Active Shape Models - their Training and Application", *Computer Vision and Image Understanding*, Vol 61 No 1, Academic Press Inc., 1995, pp. 38-59.

[CuLe96] Curless B, Levoy M, "A Volumetric Method for Building Complex Models from Range Images", *SIGGRAPH '96*, ACM, 1996, pp. 303-312.

[Cybe93] Cyberware, *3D Development*, Newsletter No 1, 1993.

[Daan98] Daanen HAM, "Circumference Estimation Using 3D-Whole Body Scanners and Shadow Scanners", *Proceedings of the Workshop on 3D Anthropometry and Industrial Products Design*, Laboratoire d'Anthropologie Appliqué, Université Paris V, 1998, pp. 5/1-5/6.

[DaBR97] Daanen HAM, Brunsman MA, Robinette KM, "Reducing Movement Artifacts in Whole Body Scanning", *Proceedings of the International Conference on Recent Advances in 3D Digital Imaging and Modeling*, IEEE Press, 1997, pp. 262-265.

[Dava97] Davatzikos C, "Spatial Transformation and Registration of Brain Images Using Elastically Deformable Models", *Computer Vision and Image Understanding*, Vol 66 No 2, Academic Press Inc., 1997, pp. 207-222.

[Davd98] Daanen HAM, Van de Water GJ, "Whole body scanners", *Displays*, Vol 19, Elsevier Science 1998, pp. 111-120.

[dBdB16] du Bois D, du Bois EF. "A formula to estimate the approximate surface area if height and weight be known", *Archives of Internal Medicine*, Vol 17, 1916, pp. 863-871.

[DDBT99] Dekker L, Douros I, Buxton BF, Treleaven P, "Building Symbolic Information for 3D Human Body Modelling from Range Data", *Proceedings of the Second International Conference on 3-D Digital Imaging and Modeling*, IEEE Computer Society, 1999, pp. 388-397.

[DeKi96] Dekker L, Kingdon J, "The Shape of Space: Topology and Trajectory", *UCL Technical Report*, 1996.

[Dekk98] Dekker L, *communication with H-Anim group*, 1998.

[DeMo98] Delingette H, Montagnat J, "General Deformable Model Approach for Model-Based Reconstruction", *IEEE International Workshop on Model-Based 3D Image Analysis*, IEEE, 1998, pp. 55-64.

[DeTr98] Dekker L, Treleaven P, "Learning Systems for Automated Anthropometric Landmarking", *Proceedings of the Workshop on 3D Anthropometry and Industrial Products Design*, Laboratoire d'Anthropologie Appliqué, Université Paris V, 1998, pp. 11/1-11/6.

[DHWG97] Demers MH, Hurley JD, Wulpern RC, Grindon JR, "3D Surface Capture for Body Measurement using Projected Sinusoidal Patterns", *Proceedings of SPIE 3023*, SPIE, 1997, pp. 13-25.

[Djab97] Djabri F, *Understanding User Perceptions of the Body Lines Scanner*, Ergonomics and HCI Unit, University College London, 1997.

[DKWB98] Dekker L, Khan S, West E, Buxton B, Treleaven P, "Models for Understanding the 3D Human Body Form", *IEEE International Workshop on Model-Based 3D Image Analysis*, IEEE, 1998, pp. 65-74.

[DoDB99] Douros I, Dekker L, Buxton B, "An Improved Algorithm for Reconstruction of the Surface of the Human Body from 3D Scanner Data Using Local B-Spline Patches", *mPeople Workshop Proceedings*, IEEE, 1999, pp. 29-36.

[DoJa97] Dorai C, Jain AK, "COSMOS – A Representation Scheme for 3D Free-Form Objects", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 19 No 10, IEEE Press, 1997, pp. 1115-1130.

[Dour98] Douros I, *B-Spline Surface Reconstruction of the Human Body from 3D Scanner Data*, MRes CVIPGS Thesis, University College London, 1998.

[Drap96] Draper B, "Object Recognition as a Markov Decision Process", *International Conference on Pattern Recognition*, Vol IV, 1996, pp. 95-99.

[DrNa81] Dreschler L, Nagel HH, "Volumetric Model and 3D Trajectory of a Moving Car Derived from Monocular TV-Frame Sequences of a Street Scene", *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, 1981, pp. 692-697.

[DuHa73] Duda RO, Hart PE, *Pattern Classification and Scene Analysis*, John Wiley, 1973.

[Eggl92] Eggleston P, "Content Based Feature Indexing and Retrieval for Image Databases", *Proceedings of SPIE 1819*, 1992.

[FaPi95] Fang T-P, Piegl LA, "Delaunay Triangulation in Three Dimensions", *IEEE Computer Graphics and Applications*, 1995, pp. 62-69.

[FARW99] Fisher RB, Ashbrook AP, Robertson C, Werghi N, "A Low-Cost Range Finder Using a Visually Located, Structured Light Source", *Proceedings of the Second International Conference on 3-D Digital Imaging and Modeling*, IEEE Computer Society, 1999, pp. 24-25.

[FGPA98] Fua P, Grün A, Plänkers R, D'Apuzzo N, Thalmann D, "Human Body Modeling and Motion Analysis from Video Sequences", *International Symposium on Real-Time Imaging and Dynamic Analysis*, Hakodate, Japan, 1998.

[Fish89] Fisher RB, *From Surfaces to Objects - Computer Vision and Three Dimensional Scene Analysis*, John Wiley & Sons, 1989.

[FJLC92] Fuller NJ, Jebb SA, Laskey MA, Coward WA, Elia M. "Four-component model for the assessment of body composition in humans: comparison with alternative methods, and evaluation of the density and hydration of fat-free mass", *Clinical Sciences*, Vol 82, 1992, pp. 687-693.

[FlJa89] Flynn PJ, Jain AK, "On Reliable Curvature Estimation", *IEEE Conference on Pattern Recognition*, IEEE, 1989, pp. 110-116.

[Free60] Freeman H, "Techniques for the Digital Computer Analysis of Chain-Encoded Arbitrary Plane Curves", *Proceedings of the National Electronics Conference*, Vol 17, 1960, pp. 421-432.

[Free90] Freeman H (ed.), *Machine Vision for Three-Dimensional Scenes*, Academic Press Inc., 1990.

[GaJo79] Garey MR, Johnson DS, *Computers and Intractability*, Freeman, 1979.

[GBCC89] Gordon CC, Bradtmiller B, Churchill T, Clauser CE, McConville JT, Tebbetts I, Walker RA, *1988 Anthropometric Survey of U.S. Army Personnel: Methods and Summary Statistics*, Anthropology Research Project Inc., Yellow Springs Ohio, 1989.

[GMHW95] Geisen G, Mason GP, Houston V, Whitestone J, McQuiston B, Beattie A, "Automatic Detection, Identification, and Registration of Anatomical Landmarks", *Proceedings of the 39th Annual Meeting of the Human Factors and Ergonomics Society*, Vol 2, The Society, San Diego, 1995, pp. 750-753.

[Gowe75] Gower JC, "Generalised Procrustes Analysis", *Psychometrika*, Vol 40, 1975, pp. 33-51.

[Gray98] Gray S, "Virtual Fashion, the New Look in Digital Design", *IEEE Spectrum*, 1998, pp. 19-25.

[GuMe97] Guy G, Medioni G, "Inference of Surfaces, 3D Curves and Junctions from Sparse, Noisy 3D Data", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 19 No 11, 1997, pp. 1265-1277.

[HaBu97] Harris C, Buxton B, "Low-Level Edge Detection Using Genetic Programming: Performance, Specificity and Application to Real-World Signals", *Evolutionary Computation in Vision*, Elsevier, 1997.

[HaHs89] Halioua M, Hsin C-L, "Optical Three-Dimensional Sensing by Phase Measuring Profilometry", *Optics and Lasers in Engineering*, Vol 11 No 3, 1989, pp. 185-215.

[HaVS96] Hauspie RC, Vercauteren M, Susanne C, "Secular Changes in Growth", *Horm Res*, Vol 45 suppl 2, 1996, pp. 8-17.

[HAWG99] Humanoid Animation Working Group, *Specification for a Standard Humanoid, Version 1.1*, http://ece.uwaterloo.ca/~h-anim/spec1.1/, 1999.

[HBGS99] Hilton A, Beresford D, Gentils T, Smith R, Sun W, "Virtual People: Capturing human models to populate virtual worlds", *IEEE International Conference on Computer Animation*, 1999, pp. 174-185.

[HDDM92] Hoppe H, DeRose T, Duchamp T, McDonald J, Stuetzle W, "Surface Reconstruction from Unorganized Points", *SIGGRAPH '92*, pp. 71-78.

[HiBT97] Hill A, Brett D, Taylor CJ, "Automatic Landmark Identification Using a New Method of Non-Rigid Correspondence", *15th International Conference on Information Processing in Medical Imaging*, 1997, pp. 483-488.

[HiMW92] Hinds B, McCartney J, Woods G, "Pattern Development for 3D Surfaces", *Computer-Aided Design*, Vol 23 No 8, 1991, pp. 583-592.

[HiTa94] Hill A, Taylor CJ, "Automatic Landmark Generation for Point Distribution Models", *Proceedings of the British Machine Vision Conference*, BMVA Press, 1994, pp. 429-438.

[Hori95] Horiguchi C, "Sensors that Detect Shape", *Journal of Advanced Automation Technology*, Vol 7 No 3, 1995, pp. 210-216.

[Hori97] Horiguchi C, *personal communication*, 1997.

[Hori98] Horiguchi C, "BL (Body Line) Scanner", *International Archives of Photogrammetry and Remote Sensing*, Vol XXXII, Part 5, Hakodate 1998.

[Horn75] Horn BKP, "Obtaining Shape from Shading Information", *The Psychology of Computer Vision*, McGraw-Hill, 1975, pp. 115-155.

[Horn84] Horn BKP, "Extended Gaussian Images", *Proceedings of IEEE*, Vol 72 No 12, 1984, pp. 1656-1678.

[HoWo98] Hodgins JK, Wooten WL. "Animating Human Athletes", *Robotics Research: the Eighth International Symposium*, Springer-Verlag, Berlin 1998, pp. 356-367.

[HQL95] Research Institute of Human Engineering for Quality Life, *Report About a Project to Construct a Database for Measurements of the Human Body* (in Japanese), 1995.

[HsHw97] Hsu J-C, Hwanag S-Y, "A Machine Learning Approach for Acquiring Descriptive Classification Rules of Shape Contours", *Pattern Recognition*, Vol 30 No 2, 1997, pp. 245-252.

[Hu61] Hu MK, "Pattern Recognition by Moment Invariants", *IRE Transactions on Information Theory*, Vol 49, 1961, pp. 1428.

[Iann99] Ianni JD, "A Specification for Human Action Representation", *International Conference on Digital Human Modeling for Design and Engineering*, Society of Automotive Engineers, USA, 1999.

[InSp99] InSpeck inc., *3D Body Builder*, 1999.

[IPSV97] Iivarinen J, Peura M, Särelä J, Visa A, "Comparison of Combined Shape Descriptors for Irregular Objects", *Proceedings of the British Machine Vision Conference*, BMVA Press, 1997, pp. 430-439.

[Isda98] Isdale J, "3D Scanning Systems", *VR News*, Vol 7 Issue 7, 1998, pp. 25-29.

[ISO89] *ISO 8559: Garment Construction and Anthropometric Surveys – Body Dimensions*, International Organization for Standardisation, 1989.

[JaZL96] Jain AK, Zhong Y, Lakshmanan S, "Object Matching Using Deformable Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 18 No 3, 1996, pp. 267-277.

[JBHM94] Jones PRM, Baker AJ, Hardy CH, Mowat AP, "Measurement of Body Surface Area in Children with Liver Disease by a Novel Three-Dimensional Body Scanning Device", *European Journal of Applied Physiology and Occupational Physiology*, Vol 68, Springer-Verlag, 1994, pp. 514-518.

[JGMS98] Jojic N, Gu J, Mak I, Shen HC, Huang TS, "Computer Modeling, Analysis and Synthesis of Dressed Humans", *Proceedings Computer Vision and Pattern Recognition*, IEEE 1998, pp. 528-534.

[JiNi95] Jia X, Nixon MS, "Extending the Feature Vector for Automatic Face Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 17 No 12, 1995, pp. 1167-1176.

[JoBW95] Jones PRM, Li P, Brooke-Wavell K, West GM, "Format for Human Body Modelling from 3D Body Scanning", *International Journal of Clothing Science and Technology*, Vol 7 No 1, MCB University Press, 1995, pp. 7-16.

[JoKi97] Jones PRM, Kilpatrick W, "Standards Relevant to 3D Surface Anthropometry", *3D Surface Anthropometry: Review of Technologies*, NATO AGARD, 1997, pp. 164-176.

[JoLi98] Jones PRM, Li P, "Three-Dimensional Anthropometry: Applications in Clothing Technology", *Proceedings of the Workshop on 3D Anthropometry and Industrial Products Design*, Laboratoire d'Anthropologie Appliqué, Université Paris V, 1998, pp. 22/1-22/6.

[JoRi97] Jones PRM, Rioux M, "Three-dimensional Surface Anthropometry: Applications to the Human Body", *Optics and Lasers in Engineering*, Vol 28, Elsevier Science, 1997, pp. 89-117.

[JWHR89] Jones PRM, West G, Harris D, Read J, "The Loughborough Anthropometric Shadow Scanner (LASS)", *Endeavour Vol. 13 No. 4 New Series*, 1989, pp. 164-168.

[KaCH93] Kang C-Y, Chen Y-S, Hsu W-H, "Mapping a Lifelike 2.5D Human Face via an Automatic Approach", *Proceedings of Computer Vision and Pattern Recognition*, 1993, pp. 611-612.

[KaHä98] Karbacher S, Häusler G, "A new approach for modeling and smoothing of scattered 3D data", *Proceedings of SPIE 3313*, 1998, pp. 168-177.

[KaWT87] Kass M, Witkin A, Terzopoulos D, "Snakes: Active contour models", *International Journal of Computer Vision*, Vol 1 No 4, 1987, pp. 321-331.

[KiDe95] Kingdon J, Dekker L, "The Shape of Space", *Proceedings of the First IEE/IEEE International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications (GALESIA '95)*, IEE, London, 1995, pp. 543-548.

[Koen99] Koenen R, "MPEG-4, Multimedia for our time", *IEEE Spectrum*, Vol 36 No 2, 1999.

[KoMK97] Kondo T, Murakami K, Koshimizu H, "From Coarse to Fine Correspondence of 3D Facial Images and its Application to 3D Facial Caricaturing", *Proceedings of the International Conference on Recent Advances in 3D Digital Imaging and Modeling*, Ottawa, 1997, pp. 283-288.

[KovD92] Koenderink JJ, van Doorn AJ, "Surface Shape and Curvature Scales", *Image and Vision Computing*, Vol 10 No 8, Butterworth-Heinemann, 1992, pp. 557-565.

[Koza92] Koza JR, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, Cambridge, MA, 1992.

[Kuni84] Kunick P, *Modern Sizing and Pattern Making for Women's and Children's Garments*, Philip Kunick Publications, 1984.

[LaST97] Lack JA, Stuart-Taylor NE, "Calculation of drug dosage and body surface area of children", *British Journal of Anaesthesia*, Vol 78, 1997, pp. 601-605.

[LeHD93] Lee C-K, Haralick RM, Deguchi K, "Estimation of Curvature from Sampled Noisy Data", *Proceedings Computer Vision and Pattern Recognition*, IEEE Press, 1993, pp. 536-541.

[LeNu98] Lewark EA, Nurre JH, "Automated fudicial [sic] labeling on human body data", *Proceedings of SPIE 3313*, 1998, pp. 82-89.

[LeTW95] Lee Y, Terzopoulos D, Waters K, "Realistic Modeling for Facial Animation", *SIGGRAPH '95*, ACM, 1995, pp. 55-62.

[LHCT95] Lanitis A, Hill A, Cootes TF, Taylor CJ, "Locating Facial Features Using Genetic Algorithms", *Proceedings of the International Conference on Digital Signal Processing*, 1995, pp. 520-525.

[Li97] Li P, *3D Surface Modelling of the Human Body*, PhD Thesis, Loughborough University, 1997.

[LiJo94] Li P, Jones PRM, "Anthropometry-Based Surface Modelling of the Human Torso", *Computers in Engineering*, American Society of Mechanical Engineers, Minneapolis, 1994, pp. 469-474.

[LiJo97] Li P, Jones PRM, "Automatic Editing and Curve-fitting of 3-D Surface Scan Data of the Human Body", *International Conference on Recent Advances in 3D Digital Imaging and Modeling*, IEEE, 1997, pp. 296-301.

[LiPe82] Lin C, Perry MJ, "Shape Description using Surface Triangularization", *Proceedings of Workshop on Computer Vision: Representation and Control*, IEEE, 1982, pp. 38-43.

[LoCl87] Lorenson W, Cline H, "Marching Cubes: a High Resolution 3D Surface Construction Algorithm", *Computer Graphics*, Vol 21 No 4, 1987, pp. 163-169.

[Lonc98] Loncaric S, "A Survey of Shape Analysis Techniques", *Pattern Recognition*, Vol 31 No 8, Elsevier Science, 1998, pp. 983-1001.

[Long81] Longuet-Higgins HC, "A Computer Algorithm for Reconstructing a Scene from Two Projections", *Nature*, Vol 293 No 5828, 1981, pp. 133-135.

[LTRC92] Linney AD, Tan AC, Richards R, Coombes AM, Gardener J, Lees WR, "The acquisition, visualisation and applications of three-dimensional data on the human body", *Proceedings of the Electronic Imaging of the Human Body Workshop*, Dayton, Ohio, 1992, pp. 38-48.

[LuBr44] Lund CC, Browder NC, "The Estimation of Areas of Burns", *Journal of Surgery, Gynaecology and Obstetrics*, Vol 79, 1944, pp. 352-358.

[Marq98] Marques JS, "A Fuzzy Algorthm for Curve and Surface Alignment", *Pattern Recognition Letters*, Vol 19, Elsevier Science, 1998, pp. 797-803.

[Marr82] Marr D, *Vision*, WH Freeman and Co., New York, 1982.

[McTe94] McInerney T, Terzopoulos D, "A Dynamic Finite Element Surface Model for Segmentation and Tracking in Multidimensional Medical Images with Application to Cardiac 4D Image Analysis", *Journal of Computerized Medical Imaging and Graphics*, 1994.

[MGBM95] McCrory MA, Gomez TD, Bernauer EM, Molé PA, "Evaluation of a New Air Displacement Plethysmograph for Measuring Human Body Composition", *Medicine and Science in Sports and Exercise* Vol 27 No 12, 1995, pp. 1686-1691.

[MoAK96] Mokhtarian F, Abbasi S, Kittler J, "Robust and Efficient Shape Indexing through Curvature Scale Space", *Proceedings of the British Machine Vision Conference*, BMVA Press, 1996, pp. 53-62.

[MoDe97] Montagnat J, Delingette H, "A Hybrid Framework for Surface Registration and Deformable Models", *Proceedings of Computer Vision and Pattern Recognition*, 1997, pp. 1041-1046.

[Mokh95] Mokhtarian F, "Silhouette-Based Isolated Object Recognition Through Curvature Scale Space", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 17 No 5, 1995, pp. 539-544.

[MTYT91] Magnenat-Thalmann N, Yang Y, Thalmann D, "The Problematics of Cloth Modeling and Animation", *Proceedings of the 2nd Conference on CAD and Computer Graphics*, International Academic Publishers, 1991.

[NeTh98] Nedel LP, Thalmann D, "Real-Time Muscle Deformations using Mass-Spring Systems", *Computer Graphics International*, 1998.

[NgBo96] Ng KC, Boyle RD, "Recognition and Reconstruction of Primitives in Music Scores", *Image and Vision Computing*, Vol 14, Elsevier Science, 1996, pp. 39-46.

[NgGr96] Ng HN, Grimsdale RL, "Computer Graphics Techniques for Modeling Cloth", *IEEE Computer Graphics and Applications*, Vol 16 No 5, 1996, pp. 24-41.

[Nixo85] Nixon M, "Eye Spacing Measurement for Facial Recognition", *Proceedings of SPIE 575, Applications of Digital Signal Processing*, 1985, pp. 279-285.

[Nurr97] Nurre JH, "Locating Landmarks on Human Body Scan Data", *International Conference on Recent Advances in 3D Digital Imaging and Modeling*, IEEE, 1997, pp. 289-295.

[OlBu99] Oliveira J, Buxton B, "Creating Lightweight Virtual Humans for Virtual Environments", *Eurographics '99*, Alberti MA, Gallo G, Jelinek I (eds.), Eurographics Association, 1999.

[ORou81] O'Rourke J, "Polyhedra of Minimal Area as 3D Object Models", *Proceedings of the International Joint Conference on Artificial Intelligence*, 1981, pp. 664-666.

[OsFG97] Osuna E, Freund R, Girosi F, "Training Support Vector Machines: an Application to Face Detection", *Proceedings of Computer Vision and Pattern Recognition*, IEEE, 1997, pp. 130-136.

[PaSD97] Pargas R, Staples N, Davis S, "Automated Measurement Extraction for Apparel from a 3D Body Scan", *Optics and Laser Technology*, Vol 28 No 2, Elsevier Science, 1997.

[PoBr87] Ponce J, Brady M, "Toward a Surface Primal Sketch", *Three-Dimensional Machine Vision*, Kanade (ed.), Kluwer Academic Publishers, 1987, pp. 195-240.

[Polh98] Polhemus Inc., *Isotrak II Tracker and 3D Digitizer*, 1998.

[Poli96] Poli R, "Genetic Programming for Image Analysis", *Proceedings of the First International Conference on Genetic Programming*, Koza *et al* (eds.), MIT Press, 1996, pp. 363-368.

[RBFM97] Ruckert D, Burger P, Forbat S, Mohiaddin RH, Yang GZ, "Automatic Tracking of the Aorta in Cardiovascular MR Images Using Deformable Models", *IEEE Transactions on Medical Imaging*, Vol 16 No 5, 1997, pp. 581-590.

[Riou84] Rioux M, "Laser Range-Finder Based on Synchronised Scanners", *Applied Optics*, Vol 23 No 21, 1984.

[Riou98] Rioux M, "Fundamental Limits of Optical 3D Digitizing Systems for the Measurement of the Human Body", *Proceedings of the Workshop on 3D Anthropometry and Industrial Products Design*, Université Paris V, Laboratoire d'Anthropologie Appliquée, Paris, 1998, pp. 3/1-3/6.

[Robi97] Robinette KM, "Chapter 1", *3D Surface Anthropometry: Review of Technologies*, Robinette *et al* (eds.), NATO AGARD Advisory Report 329, 1997, pp 1-18.

[Robi98] Robinette KM, "The Civilian American and European Surface Anthropometry Resource (CAESAR) Project", *Proceedings of the Workshop on 3D Anthropometry and Industrial*

Products Design, Laboratoire d'Anthropologie Appliqué, Université Paris V, 1998, pp. 18/1-18/3.

[RoDP99] Robinette KM, Daanen H, Paquet E, "The CAESAR Project: A 3D Surface Anthropometry Survey", *Proceedings of the Second International Conference on 3-D Digital Imaging and Modeling*, IEEE Computer Society, 1999, pp. 380-386.

[RoLe94] Roth G, Levine MD, "Geometric Primitive Extraction Using a Genetic Algorithm", *IEEE Transactions on Pattern Recognition and Machine Intelligence*, Vol 16 No 9, 1994, pp. 901-905.

[RuHW86] Rumelhart DE, Hinton GE, Williams RJ, "Learning Internal Representations by Error Propagation", *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, Vol 1, MIT Press, 1986, pp. 318-362.

[RVRJ97] Robinette KM, Vannier MW, Rioux M, Jones PRM, *3D Surface Anthropometry: Review of Technologies*, NATO AGARD Advisory Report 329, 1997.

[SaCi97] Sato J, Cipolla R, "Identifying Human Face Profiles with Semi-Local Integral Invariants", *Proceedings of Image Analysis and Processing*, Vol I, Springer LNCS 1310, 1997, pp. 414-421.

[SAE99] Society of Automotive Engineers, "Digitally Defining the Human Body – CAESAR 3-D Anthropometric Database", http://www.sae.org/technicalcommittees/caesar.htm, 1999.

[SaIy92] Samal A, Iyengar PA, "Automatic Recognition and Analyis of Human Faces and Facial Expressions: a Survey", *Pattern Recognition*, Vol 25 No 1, 1992, pp. 65-77.

[Sanz89] Sanz JLC (ed.), *Advances in Machine Vision*, Springer-Verlag, New York, 1989.

[SaRS91] Sachs E, Roberts A, Stoops D, "3-Draw: A Tool for Designing 3D Shapes", *IEEE Computer Graphics and Applications*, Vol 11 No 6, 1991, pp. 18-26.

[ScFo86] Schewe H, Forstner W, "The Program PALM for Automatic Line and Surface Measurement Using Image Matching Techniques", *Proceedings of the Symposium of the International Society for Photogrammetry and Remote Sensing*, Vol 26 Part 3/2, From Analog to Digital, pp. 608-622.

[ScPe95] Sclaroff S, Pentland AP, "Modal Matching for Correspondence and Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 17 No 6, 1995, pp. 545-561.

[SCTD97] Sozou PD, Cootes TF, Taylor CJ, Di Mauro EC, Lanitis A, "Non-Linear Point Distribution Modelling Unsing a Multi-Layer Perceptron", *Image and Vision Computing*, Vol 15, Elsevier Science, 1997, pp. 457-463.

[SDNV90] Smets C, De Groof M, Nuyts J, Vandermeulen D, Suetens P, Marchal G, Oostelinck A, "Interpretation of 3D Medical Scenes", *Machine Vision for Three-Dimensional Scenes*, Freeman H (ed.), Academic Press Inc., 1990, pp. 163-193.

[Sedg90] Sedgewick R, *Algorithms in C*, Addison-Wesley, 1990.

[Seit98] Seitz T, "The Optical Measurement System PCMAN", *Proceedings of the Workshop on 3D Anthropometry and Industrial Products Design*, Laboratoire d'Anthropologie Appliqué, Université Paris V, 1998, pp. 7/1-7/4.

[SiPa98] Siebert JP, Patterson JW, "Captivating Models", *Proceedings of the IEE Colloquium on Computer Vision for Virtual Human Modelling*, IEE, London, 1998, pp. 7/1-7/3.

[Siri61] Siri WE. "Body composition from fluid spaces and density: analysis of methods", *Techniques for measuring body composition*, Brozek J, Henschel A (eds.), National Academy of Sciences NRC, 1961, pp. 223-244.

[Smit99] Smith R, "Proposal for the Specification of Seamless Human Models in H-Anim", *Technical Report VSSP-TR-3/99*, Surrey University, 1999.

[SmJW90] Smith NSH, Jones PRM, West GM, "A New Tool in the Study of Human Body Composition", *Annals of Human Biology*, Vol 17, pp. 340.

[SPCM97] Scheepers F, Parent RE, Carlson WE, May SF, "Anatomy-Based Modeling of the Human Musculature", *SIGGRAPH '97*, pp. 163-172.

[StHI94] Stoddart AJ, Hilton A, Illingworth J, "Slime: A new deformable surface", *Proceedings of the British Machine Vision Conference*, BMVA Press, 1994, pp. 285-294.

[StIW95] Stoddart AJ, Illingworth J, Windeatt T, "Optimal Parameter Selection for Derivative Estimation from Range Images", *Image and Vision Computing*, Vol 13, 1995, pp. 629-635.

[StMi98] Stein N, Minge B, "VIRO 3D: Fast Three-Dimensional Full-Body Scanning for Humans and Other Living Objects", *Proceedings of SPIE 3313*, 1998, pp. 60-64.

[Tait98] Tait N, "Is Mass Customization Possible?", *Apparel*, The Textile Institute, 1998, pp. 22-24.

[TaMe99] Tang C-K, Medioni G, "Robust Estimation of Curvature from Noisy Data for Shape Description", *Proceedings of the Internaional Conference on Computer Vision*, IEEE, 1999, pp. 426-433.

[Tecm97] Tecmath, *High Speed 3D-Body Scanner, Technical Specification*, Tecmath, 1997.

[Telm96] Telmat, *Automated Body Measuring and Garment Size Allocating System*, Telmat Ed. 7, 1996.

[TrBS98] Treleaven PC, Buxton B, Slater M, *Centre for 3D Electronic Commerce*, Foresight LINK Award Application, University College London, 1998.

[TrFi95] Trucco E, Fisher RB, "Experiments in Curvature-Based Segmentation of Range Data", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 17 No 2, 1995, pp. 177-182.

[TSHH98] Trieb R, Seidl A, Hansen G, Hamfeld H, Stöhr M, Ovtcharova J, "RAMSIS/SCAN – A New Approach to 3D Body Scanning for Autoamted Anthropometric Measuring and Individual Product Design", *Proceedings of the Workshop on 3D Anthropometry and Industrial Products Design*, Laboratoire d'Anthropologie Appliqué, Université Paris V, 1998, pp. 10/1-10/6.

[Tsop93] Tsopelas N, *Modelling Thin-Walled Objects in Computer Graphics and Animation*, PhD Thesis, University of London, 1993.

[TuPe91] Turk M, Pentland A, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, Vol 3 No 1, MIT 1991, pp. 71-86.

[Vann97] Vannier MW, "Visualization, Modeling and Analysis", *3-D Surface Anthropometry: Review of Technologies*, NATO Advisory Group for Aerospace Research & Development, NATO Neuilly-sur-Seine, 1997, pp. 75-113.

[Vase96] Vaseghi SV, *Advanced Signal Processing and Digital Noise Reduction*, Wiley Teubner, 1996.

[WaCT98] Walker KN, Cootes TF, Taylor CJ, "Locating Salient Object Features", *Proceedings of the British Machine Vision Conference*, BMVA Press, 1998, pp. 557-566.

[WDFE99] Wells JCK, Douros I, Fuller NJ, Elia M, Dekker L, "Assessment of Body Volume Using 3-Dimensional Photonic Scanning", in press, 1999.

[WeMT93] Werner HM, Magnenat Thalmann N, Thalmann D, "User Interface for Fashion Design" *IFIP Transactions on Graphics Design and Visualization*, North-Holland, Amsterdam, 1993, pp. 197-204.

[West97] West E, *B-Spline Surface Skinning for Body Scanner Data*, MRes CVIPGS Thesis, University College London, 1997.

[WhGM97] Whitestone J, Geisen GR & McQuiston BK, "Three-dimensional Anthropometric Techniques Applied to the Fabrication of Burn Masks and the Quantification of Wound Healing", *Proceedings of SPIE 3023, Three-Dimensional Image Capture*, 1997.

[Wink97] Winks J, *Clothing Sizes: International Standardization*, The Textile Institute, Manchester, 1997.

[Witk83] Witkin A, "Scale-space filtering", *Proceedings of the 7th Joint Conference on Artificial Intelligence*, 1983, pp. 1019-1021.

[WiWi98] Wicks and Wilson Ltd., *TriForm 3D Scanners*, 1998.

[WoMa95] Wolpert D, Macready W, "No Free Lunch Theorems for Search", *Santa Fe Institute Technical Report* SF1-TR-95-02-010, 1995.

[Wood78] Woodham RJ, "Photometric Stereo: A Reflectance Map Technique for Determining Surface Orientation from a Single View", *Proceedings of SPIE 255, 22nd Annual Symposium*, 1978, pp. 136-143.

[WuMT95] Wu Y, Magnenat-Thalmann N, Thalmann D, "A Dynamic Wrinkle Model in Facial Animation and Skin Aging", *Journal of Visualization and Computer Animation*, Vol 6, 1995, pp. 195-205.

[YaDa93] Yacoob Y, Davis L, "Labeling of Human Face Components from Range Data", *Proceedings of Computer Vision and Pattern Recognition*, 1993, pp. 592-593.

[Yama98] Yamanaka D, *Segmentation, Volume and Surface Area Calculations in a 3D Human Body Model*, MSc Thesis, University College London, 1998.

[Yead97] Yeadon MR, "The biomechanics of the human in flight", *The American Journal of Sports Medicine*, Vol 25 No 4, 1997, pp. 575-580.

[YuCH92] Yuille AL, Cohen DS, Hallinan P, "Feature Extraction from Faces Using Deformable Templates", *International Journal of Computer Vision*, Vol 8, 1992, pp. 99-112.

[ZaRo72] Zahn CT, Roskies RZ, "Fourier Descriptors for Plane Closed Curves", *IEEE Transactions on Computers*, C-21Vol 3, 1972, pp. 269-281.

Appendix A: Anthropometric Specifications

The table below summarises the anthropometric extraction functions implemented in this work. The definitions are largely based on ISO 8559 [ISO89], with additional dimensions from *de facto* industry standards in the UK, and Kunick [Kuni84].

Dimension	ISO8559	Definition
Neck girth	2.1.2	Girth of neck below Adam's apple at level of 7th cervical vertebra.
Shoulder length	2.1.4	Distance from side neck base to outer edge of acromion.
Shoulder width	2.1.5	Distance between outer acromion landmarks across shoulders at back, following body surface.
Back width	2.1.6	Horizontal distance across back, measured midway between top and bottom of armscye.
Bust girth	2.1.8	Maximum horizontal girth, normal breathing, across nipples and shoulder blades.
Bust width	2.1.9	Horizontal distance between the nipples.
Underbust girth	2.1.10	Horizontal girth directly below breasts.
Waist girth	2.1.11	Girth of "natural" waist between iliac crests and lower ribs. Normal breathing, relaxed.
Hip girth	2.1.12	Horizontal girth at level of trochanters.
Thigh girth	2.1.18	Horizontal girth at highest level on thigh.
Mid-thigh girth	2.1.19	Horizontal girth midway between hip level and knee.
Knee girth	2.1.20	Horizontal girth around the knee, between top of tibia and bottom of femur.
Calf girth	2.1.22	Maximum horizontal girth on the calf.
Minimum leg girth	2.1.23	Minimum horizontal girth above ankle bones.
Ankle girth	2.1.24	Horizontal girth at the level of the centre of outer ankle bone.
Stature	2.2.1	Height of top of head from floor. Head held with chin approximately parallel with floor.
Trunk length	2.2.2	Vertical distance between 7th cervical vertebra and crotch level.
Waist height	2.2.3	Vertical distance between waist level and floor.
Hip height	2.2.4	Vertical distance between hip level and floor.
Body rise	2.2.5	Vertical distance between waist and crotch level.
Knee height	2.2.6	Vertical distance between knee level and floor.
Scye depth	2.2.9	Vertical distance between 7th cervical vertebra and the average level of the left and right axilla landmarks.
Back waist length	2.2.10	Curvilinear distance between 7th cervical vertebra and centre back waist level.
Cervical height	2.2.12	Straight line distance of 7th cervical vertebra to floor. (Adaptation of ISO8559 2.2.12).
Cervical to breast point	2.2.13	Curvilinear distance from 7th cervical vertebra to side base neck to nipple.
Front waist length	2.2.16	Side neck base, over nipple, then vertically straight down to front waist.
Waist to hips	2.2.17	Curvilinear distance at side of body between waist level and hip level.
Trunk circumference	2.2.18	Centre right shoulder down via front breast, crotch point, then back to centre shoulder.
Total crotch length	2.2.19	Curvilinear distance from centre front waist, via crotch, to centre back waist.
Thigh length	2.2.26	Vertical distance between crotch point and knee level.
Inside leg length	2.2.27	Vertical distance between crotch point and floor.
Shoulder slope	2.3.1	Angle between side neck base and outer edge of acromion, measured on the top of shoulder.

Upper hip circumference	N/A	Horizontal girth at level one quarter way down from waist to full hip.
Top hip circumference	N/A	Horizontal girth at level half way down from waist to full hip level.
Abdominal seat diameter	N/A	Horizontal distance between greatest abdominal protruberance and maximum buttock protruberance.
Centre front clavicle to waist	N/A	Curvilinear distance of centre front clavicle to centre front waist.
Bust level	N/A	Height of fullest bust level.
Underbust level	N/A	Height of underbust level.
Centre shoulder to bust	N/A	Curvilinear distance of side neck base to nipple.
Centre shoulder to underbust	N/A	Curvilinear distance of centre shoulder to nipple then directly down to underbust level.
Centre shoulder to front waist	N/A	Curvilinear distance of centre shoulder to nipple then directly down to waist level.
Centre shoulder to back waist	N/A	Curvilinear distance of centre shoulder over shoulder blade then directly down to waist level.

Appendix B: Landmark Glossary

7th cervical vertebra (also cervicale or nape)

The most prominent vertebra at base of neck at centre back (with head held upright).

Acromion (shoulder point or upper armscye marker)

Outer edge of shoulder blade at the turning point of the shoulder. This is situated close to the prominent outer end of the clavicle, which lies on the inner side, toward the neck.

Adam's apple

Cartilage projection at centre front of neck.

Ankle bones

Inner and outer bony projections at lower end of tibia (inner) and fibula (outer).

Axilla (also underarm, armpit or lower armscye)

Lowest level at branch point of the arm and torso, at centre under the arm.

Centre front clavicle

Central point between inner ends of left and right clavicle (collar) bones at front.

Centre shoulder

Point on the top ridge of the trapezius muscle (on the top of shoulder), halfway between side neck base and acromion landmark.

Cervicale

See 7th cervical vertebra.

Femur

Upper bone of leg.

Fibula

Smaller, outer bone of lower leg.

Humerus

Upper bone of arm.

Iliac crest

Upper front bone of pelvis.

Lower armscye

See axilla.

Mid-armscye markers

Landmarks at front and back, halfway vertically between upper armscye and axilla markers.

Nape

See 7th cervical vertebra.

Nipple (breast point or bust point)

Centre of nipple.

Side neck base

Points on the left and right side at the junction between the neck and the shoulder, on the top ridge of the trapezius muscle (on the top of shoulder).

Supra-clavicular notch (also suprasternale)

Dip at front base neck, where left and right clavicle bones join sternum.

Suprasternale

See supra-clavicular notch.

Tibia

Larger, inner bone of lower leg.

Trapezius muscle

Large muscle that follows the top edge of the shoulder, to the back of the neck and down between the shoulder blades.

Trochanter

Outer edge of top of femur (thigh bone) close to skin surface at hip level.

Underarm

See axilla.

Upper armscye

Landmark vertically above axilla, on top of shoulder, at centre.

Appendix C: Automated Anthropometry Results

Table C.1: Training set, manual measurements, 18 samples.

	μ (mm)	σ (mm)	range (mm)	% variability
stature	1695	83	319	17
nape height	1439	74	285	17
crotch height	779	50	200	22
centre front waist height	1056	50	230	19
left knee height	469	32	115	21
right knee height	470	33	130	23
neck girth	324	16	50	14
bust girth	897	46	190	20
underbust girth	757	41	140	17
waist girth	712	37	127	16
upper hip girth	820	43	167	18
top hip girth	890	35	125	13
hip girth	965	43	180	17
left thigh girth	576	31	110	17
right thigh girth	578	31	110	17
left knee girth	361	22	97	24
right knee girth	361	22	95	23
left calf girth	364	17	68	17
right calf girth	363	17	71	18
left ankle girth	247	12	45	17
right ankle girth	248	12	45	17
shoulder width	386	20	75	18
left shoulder length	137	12	40	26
right shoulder length	138	11	35	23
bust width	201	14	45	20
back width	350	26	105	26
back waist length	409	34	160	32
clavicle to waist length	361	26	115	27
cervical to breast point	344	23	80	21
left mid shoulder to bust	260	17	65	22
right mid shoulder to bust	258	17	60	21
left mid shoulder to underbust	341	26	105	26
right mid shoulder to underbust	340	26	100	25
left mid shoulder to front waist	451	29	120	23
right mid shoulder to front waist	448	26	92	18
left mid shoulder to back waist	416	32	140	28
right mid shoulder to back waist	414	32	145	29
scye depth	208	19	71	30
total crotch length	696	50	175	23
trunk circumference	1551	67	270	16
abdominal seat diameter	250	21	70	25
body rise	271	24	85	27

Table C.2: Training set, automated measurements, 18 samples.

	μ (mm)	σ (mm)	range (mm)	% variability
stature	1695	80	295	17
nape height	1434	72	270	19
crotch height	777	51	210	27
centre front waist height	1058	64	245	23
left knee height	458	25	105	23
right knee height	456	26	95	21
neck girth	327	28	116	35
bust girth	892	43	162	18
underbust girth	749	38	124	17
waist girth	704	38	149	21
upper hip girth	802	40	158	20
top hip girth	890	36	126	14
hip girth	970	45	183	19
left thigh girth	559	27	88	16
right thigh girth	559	27	92	16
left knee girth	365	19	65	18
right knee girth	363	20	70	19
left calf girth	363	20	70	19
right calf girth	361	18	61	17
left ankle girth	247	15	66	27
right ankle girth	248	16	71	28
shoulder width	375	19	80	21
left shoulder length	121	9	36	30
right shoulder length	118	9	29	25
bust width	195	17	68	35
back width	315	27	123	39
back waist length	388	21	76	19
clavicle to waist length	348	20	75	21
cervical to breast point	370	24	82	22
left mid shoulder to bust	250	23	88	35
right mid shoulder to bust	245	20	87	36
left mid shoulder to underbust	325	30	97	30
right mid shoulder to underbust	330	33	109	33
left mid shoulder to front waist	416	21	86	21
right mid shoulder to front waist	421	32	128	30
left mid shoulder to back waist	414	21	71	17
right mid shoulder to back waist	410	17	64	16
scye depth	180	19	82	45
total crotch length	726	41	174	24
trunk circumference	1582	64	254	16
abdominal seat diameter	255	17	55	21
body rise	297	17	69	23

Table C.2: Training set, automated measurements, 18 samples.

	μ (mm)	σ (mm)	range (mm)	% variability
stature	1695	80	295	17
nape height	1434	72	270	19
crotch height	777	51	210	27
centre front waist height	1058	64	245	23
left knee height	458	25	105	23
right knee height	456	26	95	21
neck girth	327	28	116	35
bust girth	892	43	162	18
underbust girth	749	38	124	17
waist girth	704	38	149	21
upper hip girth	802	40	158	20
top hip girth	890	36	126	14
hip girth	970	45	183	19
left thigh girth	559	27	88	16
right thigh girth	559	27	92	16
left knee girth	365	19	65	18
right knee girth	363	20	70	19
left calf girth	363	20	70	19
right calf girth	361	18	61	17
left ankle girth	247	15	66	27
right ankle girth	248	16	71	28
shoulder width	375	19	80	21
left shoulder length	121	9	36	30
right shoulder length	118	9	29	25
bust width	195	17	68	35
back width	315	27	123	39
back waist length	388	21	76	19
clavicle to waist length	348	20	75	21
cervical to breast point	370	24	82	22
left mid shoulder to bust	250	23	88	35
right mid shoulder to bust	245	20	87	36
left mid shoulder to underbust	325	30	97	30
right mid shoulder to underbust	330	33	109	33
left mid shoulder to front waist	416	21	86	21
right mid shoulder to front waist	421	32	128	30
left mid shoulder to back waist	414	21	71	17
right mid shoulder to back waist	410	17	64	16
scye depth	180	19	82	45
total crotch length	726	41	174	24
trunk circumference	1582	64	254	16
abdominal seat diameter	255	17	55	21
body rise	297	17	69	23

Table C.3: Training set: automated versus manual measurements, 18 samples.

	μ (mm)	MAD (mm)	RMS (mm)	worst (mm)	σ (mm)	M%D
stature	0	7	12	43	12	0.41
nape height	-5	5	6	24	8	0.34
crotch height	-2	6	7	16	7	0.77
centre front waist height	-8	18	21	42	20	1.70
left knee height	-11	12	16	38	13	2.56
right knee height	-14	14	18	37	12	2.98
neck girth	3	22	27	60	27	6.76
bust girth	-5	14	17	44	17	1.56
underbust girth	-8	15	18	42	17	1.98
waist girth	7	18	22	42	22	2.53
upper hip girth	-3	22	31	75	32	2.68
top hip girth	8	14	20	57	19	1.57
hip girth	5	10	13	29	12	1.04
left thigh girth	-16	17	21	44	14	2.95
right thigh girth	-19	19	22	42	10	3.29
left knee girth	4	7	10	28	9	1.94
right knee girth	2	8	9	18	9	2.22
left calf girth	-1	6	7	15	7	1.65
right calf girth	-1	5	6	15	6	1.38
left ankle girth	0	12	15	46	16	4.86
right ankle girth	1	13	16	36	16	5.24
shoulder width	-11	16	20	46	18	4.15
left shoulder length	-16	17	19	34	10	12.41
right shoulder length	-20	20	22	34	9	14.49
bust width	-6	12	17	51	16	5.97
back width	-35	41	45	86	28	11.71
back waist length	-12	15	18	38	14	3.67
clavicle to waist length	-2	12	15	27	15	3.32
cervical to breast point	26	27	31	68	17	7.85
left mid shoulder to bust	-10	15	17	26	14	5.77
right mid shoulder to bust	-13	17	20	39	15	6.59
left mid shoulder to underbust	-16	18	22	53	16	5.28
right mid shoulder to underbust	-11	21	24	39	22	6.18
left mid shoulder to front waist	-27	27	33	57	19	5.99
right mid shoulder to front waist	-18	29	33	58	28	6.47
left mid shoulder to back waist	7	15	18	40	17	3.61
right mid shoulder to back waist	4	24	28	44	29	5.80
scye depth	-28	29	33	58	19	13.94
total crotch length	14	21	25	44	21	3.02
trunk circumference	31	31	35	63	18	2.00
abdominal seat diameter	5	11	13	28	13	4.40
body rise	17	21	24	44	18	7.75
mean	**-3**	**17**	**21**	**46**	**16**	**4.70**

Table C.4: Validation set, manual measurements, 18 samples.

	μ (mm)	σ (mm)	range (mm)	% variability
stature	1642	51	211	12
nape height	1401	41	152	10
crotch height	754	36	136	17
centre front waist	1028	38	155	14
right mid knee	447	22	79	16
hip level	818	40	149	17
neck girth	315	13	50	15
bust girth	886	48	156	16
underbust girth	734	45	175	21
waist girth	689	57	210	26
upper hip girth	771	67	180	21
top hip girth	863	66	187	20
hip girth	956	64	205	20
right thigh girth	547	52	169	28
right knee girth	361	21	80	20
right calf girth	346	27	113	29
right minimum leg girth	210	13	48	21
shoulder width	373	21	75	19
right shoulder length	125	8	30	22
bust width	198	14	50	23
back width	356	27	90	22
cervical to breast point	338	24	80	21
right mid shoulder to front waist	404	29	83	18
right mid shoulder to back waist	397	27	88	19
scye depth	156	22	82	41
total crotch length	711	49	195	25
body rise	274	24	71	23

Table C.5: Validation set, automated measurements from scan #1, 18 samples.

	μ (mm)	σ (mm)	range (mm)	% variability
stature	1657	34	105	6
nape height	1408	37	125	9
crotch height	767	23	73	9
centre front waist	1038	33	95	9
right mid knee	452	12	45	10
hip level	829	21	70	8
neck girth	324	21	70	22
bust girth	865	58	166	19
underbust girth	723	46	145	20
waist girth	699	67	265	38
upper hip girth	786	71	236	30
top hip girth	867	70	213	25
hip girth	942	61	165	17
right thigh girth	552	44	127	23
right knee girth	349	25	90	26
right calf girth	335	30	108	32
right minimum leg girth	207	14	49	24
shoulder width	362	23	85	23
right shoulder length	114	10	39	34
bust width	198	17	64	32
back width	298	27	104	35
cervical to breast point	365	21	68	19
right mid shoulder to front waist	403	26	85	21
right mid shoulder to back waist	417	18	61	15
scye depth	186	13	38	20
total crotch length	705	61	203	29
body rise	284	26	79	28

Table C.6: Validation set, automated measurements from scan #2, 12 samples.

	μ (mm)	σ (mm)	range (mm)	% variability
stature	1656	31	97	6
nape height	1411	39	125	9
crotch height	769	21	73	9
centre front waist	1038	30	85	8
right mid knee	451	15	50	11
hip level	830	23	75	9
neck girth	322	24	76	24
bust girth	864	58	177	20
underbust girth	722	47	148	21
waist girth	700	65	252	36
upper hip girth	787	73	225	29
top hip girth	869	70	204	24
hip girth	941	60	163	17
right thigh girth	552	42	120	22
right knee girth	349	28	100	29
right calf girth	335	31	108	32
right minimum leg girth	208	16	53	26
shoulder width	370	26	88	24
right shoulder length	117	7	21	18
bust width	184	12	34	19
back width	306	29	117	38
cervical to breast point	364	28	82	22
right mid shoulder to front waist	411	21	79	19
right mid shoulder to back waist	412	28	94	23
scye depth	180	17	60	34
total crotch length	700	57	186	27
body rise	281	23	77	27

Table C.7: Validation set, automated measurements from scan #1 versus manual, 18 samples.

	μ (mm)	MAD (mm)	RMS (mm)	worst (mm)	σ (mm)	M%D
stature	2	17	21	49	22	1.04
nape height	-4	15	18	30	18	1.07
crotch height	-2	10	13	33	13	1.33
centre front waist	-2	19	21	43	22	1.85
right mid knee	-1	12	16	29	16	2.68
hip level	-3	18	22	38	23	2.20
neck girth	-2	26	29	61	30	8.29
bust girth	-10	16	20	36	17	1.81
underbust girth	5	17	21	40	21	2.32
waist girth	24	25	35	107	25	3.63
upper hip girth	28	32	38	91	26	4.15
top hip girth	17	20	25	52	20	2.32
hip girth	-3	7	8	14	8	0.73
right thigh girth	7	20	25	59	25	3.66
right knee girth	-8	11	14	29	12	3.05
right calf girth	-7	9	11	20	9	2.60
right minimum leg girth	-3	5	6	11	5	2.38
shoulder width	-7	16	21	50	20	4.29
right shoulder length	-8	12	15	35	14	9.60
bust width	3	11	20	73	20	5.56
back width	-53	53	59	107	28	14.89
cervical to breast point	26	27	30	52	16	7.99
right mid shoulder to front waist	0	13	18	42	18	3.22
right mid shoulder to back waist	19	23	28	50	22	5.79
scye depth	11	12	15	35	10	7.96
total crotch length	1	26	34	62	35	3.66
body rise	13	15	18	42	13	5.47
mean	**2**	**18**	**22**	**48**	**19**	**4.22**

Table C.8: Validation set, automated measurements from scan #2 versus manual, 12 samples.

	μ (mm)	MAD (mm)	RMS (mm)	worst (mm)	σ (mm)	M%D
stature	2	20	24	51	25	1.22
nape height	0	14	16	24	17	1.01
crotch height	-2	11	13	28	14	1.45
centre front waist	-4	17	21	39	21	1.60
right mid knee	0	13	17	34	18	2.87
hip level	-6	15	20	37	20	1.74
neck girth	8	23	28	61	28	7.43
bust girth	-13	16	18	38	14	1.80
underbust girth	1	15	17	27	17	2.02
waist girth	27	28	38	97	28	4.09
upper hip girth	30	37	43	83	31	4.83
top hip girth	20	23	27	46	19	2.70
hip girth	-3	8	10	18	10	0.84
right thigh girth	10	15	19	49	17	2.81
right knee girth	-7	13	18	32	17	3.73
right calf girth	-5	7	9	19	7	2.05
right minimum leg girth	-2	6	7	14	7	2.96
shoulder width	1	15	18	44	19	3.95
right shoulder length	-7	8	11	19	8	6.82
bust width	-10	12	18	41	15	6.38
back width	-50	61	67	118	47	17.0
cervical to breast point	26	26	29	50	14	7.63
right mid shoulder to front waist	4	17	20	34	21	4.18
right mid shoulder to back waist	18	27	33	69	28	6.90
scye depth	6	7	11	35	10	4.61
total crotch length	-1	30	36	69	38	4.30
body rise	10	20	26	46	25	7.23
mean	**2**	**19**	**23**	**45**	**20**	**4.23**

Table C.9: Validation set, automated measurements from scan #1 versus scan #2, 12 samples.

	μ (mm)	MAD (mm)	RMS (mm)	worst (mm)	σ (mm)	M%D
stature	2	4	5	10	5	0.26
nape height	-2	6	7	15	7	0.42
crotch height	-2	4	5	8	4	0.54
centre front waist	0	6	7	15	8	0.57
right mid knee	1	4	7	20	7	0.93
hip level	-1	3	5	10	5	0.41
neck girth	2	6	7	14	7	2.00
bust girth	1	6	8	17	9	0.72
underbust girth	1	7	10	18	10	1.01
waist girth	-1	9	11	20	11	1.34
upper hip girth	-2	11	14	26	15	1.47
top hip girth	-2	4	5	13	5	0.52
hip girth	1	2	3	6	3	0.22
right thigh girth	0	3	3	6	4	0.49
right knee girth	0	3	5	12	5	0.95
right calf girth	-1	2	3	7	3	0.67
right minimum leg girth	-1	2	4	12	4	1.02
shoulder width	-7	14	17	34	16	3.80
right shoulder length	-2	8	10	20	10	6.84
bust width	14	15	23	69	19	7.94
back width	-8	21	36	100	36	5.97
cervical to breast point	1	14	19	53	20	4.19
right mid shoulder to front waist	-8	13	16	28	14	3.10
right mid shoulder to back waist	5	17	21	43	21	4.25
scye depth	6	11	13	26	13	6.90
total crotch length	5	16	18	35	18	2.26
body rise	3	16	19	37	19	5.79
mean	**0**	**9**	**11**	**25**	**11**	**2.39**