Proceedings of the Second International Conference on 3-D Digital Imaging and Modeling, IEEE Computer Society, 1999, pp. 388-397.
4-8th October 1999, Ottawa, Canada. Sponsored and Organised by the National Research Council of Canada.

Building Symbolic Information for 3D Human Body Modeling from Range Data

L. Dekker I. Douros B. F. Buxton P. Treleaven
Department of Computer Science
University College London
London, UK, WC1E 6BT

Abstract

This work is concerned with the signal-to-symbol problem of building skinned, segmented, landmarked and labeled 3D models of the whole human body from range data. A fully automated model-based process is presented that takes raw range data, cleans and skins it, and then locates "interesting" features, to enrich the surface with symbolic information for specific applications. The method is validated via volumetrics in medicine and surface anthropometry.

1. Introduction

Range imaging systems designed specifically for the living human body [9] are becoming increasingly available, providing high-resolution surface data (approximately 5×5 mm or higher) with potential to build numerically and perceptually accurate digital body models. Such models have wide-ranging applications. Body size and shape, in all its variation, are of importance in ergonomic design and to improve comfort and fit in clothing [24]. In medicine, these data can be used in orthotics and prosthetics design [7], to determine correct doses of medication [13], and in monitoring posture and growth [3]. At the other extreme there is intense research activity whose goal is the construction of *dynamic* three-dimensional digital models that exhibit physiological and behavioral realism [12] [17].

These applications provide strong motivation for exploring methods to generate personalized digital body models on a large scale. A recurring theme is the need to augment the 3D digitized surface images with symbolic information, in a fully automated way. For example, segment labeling can be used to generate control parameters for cleaning, smoothing and surface reconstruction, examples of which are described here. Similarly, sets of homologous landmarks can be used as registration points for subsequent analysis, for example, via thin-plate splines [4], Procrustes alignment [15] or the training of point distribution models (PDMs) [8]. In the development of virtual animated humans, surface features can be used to infer joint locations for articulation [12] to make standards-compliant dynamic models.

This paper describes a model-based approach that has been developed to take human whole-body range data, and generate useful symbolic information, in particular, high-level body segments and landmark points used for anthropometry, such as hip points, the waist, center clavicle, bust points, and so on [21] (figure 1).

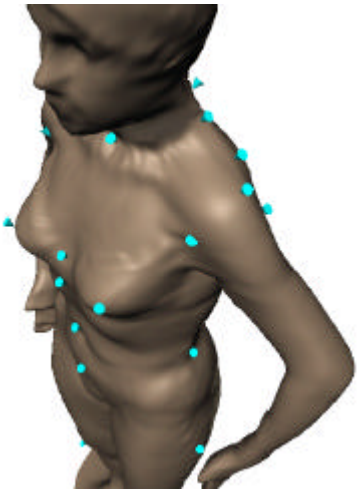


Figure 1. Digitized and skinned scan subject, with anthropometric landmarks.

Two general approaches are discussed: *global deformable template models* and *discriminant functions for modeling individual features*. The latter approach is taken here. The overall method is summarized below:

1. The raw range data set is cleaned, smoothed and skinned to provide a base, structured representation for feature detection. Two new representations are described: a compound polygon mesh model composed of deformed cylindrical segments, and a smooth closed B-spline representation.

2. The low-level and skinning processes are validated by comparison of the enclosed volumes with hydrodensitometric and plethysmographic methods [13]. Results are presented for women, men and children (down to eight years of age).
3. A number of operators are explored for their discriminatory power in detecting features of interest, and are composed into a library for constructing higher-level feature detection procedures.
4. The feature detection techniques are validated by developing a set of procedures that extract forty or so linear and curvilinear measurements on the body, according to the anthropometric standard ISO 8559 [21]. Training data is composed of scanned and manually measured female adult volunteers, of body mass index 14.6-27.2 (BMI = $weight/height^2$, *weight* in kg, *height* in m).

The processes developed show an accuracy of within a few per cent on sample data, are fast and require no manual intervention. These attributes are considered crucial to their application on a large scale.

2. Feature modeling approaches

2.1. Global deformable templates

Where some *a priori* information is known about the shape of the object, global deformable models have been used for some time for segmentation and surface reconstruction in three dimensions. The prior information 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 some sample data. Deformation constraints can be constructed from statistical information, for example, as point distribution models (PDMs) [5].

In deformable template models key landmark points are usually placed manually and additional points inferred from them, for example by spacing at regular intervals on cross sections of the body [24]. Similarly, construction of PDMs usually begins with the definition of sets of key landmark points by hand in training data [8], although attempts are being made to automate this process for 3D models, for example, using pairwise surface correspondence [5]. Taking quite a different approach, Geisen *et al* have developed automated techniques to detect manually placed fiducial markers from 3D body scans [14]. Whilst this has shown good results, it is limited in that it only provides the pre-placed landmarks, and any other features must still be detected via some other technique.

It is at first sight appealing to think of using anthropometric landmarks as the key points to initialize 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, 3D surface data. However, anthropometric landmarks are often based on skeletal features (such as particular vertebrae) or subtle visual features (such as the crease at the back of the knee) [21]. 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 model.

Although this looks like a potentially promising approach, it provides no clear mechanism for making use of all the 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.

2.2. Feature discriminant functions

An alternative approach is to construct functions to characterize and search for features individually. The potential of this in surface anthropometry has already been demonstrated elsewhere [25] [26]. Low-level attributes, including local shape descriptions, can be used as input attributes, in conjunction with contextual and other global information. Many learning techniques have been used in this way to construct models, for example neural networks [29], support vector machines [27] and genetic programming [16]. In Turk and Pentland's face recognition work [29], low-level operators are used to extract salient image attributes in 2D, and after reducing the dimensions of the data using principal components analysis, the new attributes are fed into a neural network, which attempts to model the object classification relationships. Similarly, Pontil *et al* use support vector machines to recognize 3D objects from a series of 2D grayscale images [27]; Draper [11] generates hierarchical functions from a set of machine vision primitives, for recognizing objects in 2D images; and Harris demonstrates the use of genetic programming in the synthesis of task-specific edge detectors [16].

This approach provides flexibility in that each feature detector can be modeled using a different representational form. The input attributes may be of any type (shape, non-shape, local, global), and various shape description methods can be used, for example, local differential geometry [23], pairwise geometric histograms [2], simple shape descriptors such as convexity and elliptic variance [20], and curve characterization using Bezier curves [6].

While this 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. To attempt to improve the tractability of the problem a progressive modeling approach is proposed. The first step in imposing structure on the initial body

data is to design a complete surface or solid model which will form the base for detection of body features. Primary landmarks, such as the branching points at the crotch and armpits, will be located, then these will be used to provide context for subsequent detectors. Part of this work will be to determine a set of effective operator primitives to construct higher-level procedures for feature detection, thereby reducing the combinatorial explosion encountered, for example, in attempts using genetic programming with low-level primitives [16]. In addition, this exploratory, function construction approach also gives scope for discovering interesting emergent properties of the human body.

3. Body surface representation

As the properties of the 3D image data and the skinned surface are of paramount importance in choosing the feature detectors and transformations to be used, we first describe the processes from image capture to surface reconstruction.

3.1. Image capture and preprocessing

The source of whole-body images used in our work is the 8-head Body Line Scanner developed by Hamamatsu Photonics [19] (figure 2). It uses near infrared position-sensitive detectors, arranged in eight sensor heads, each with an array of 32 laser diodes (LEDs). The scanning mechanism moves 5 mm vertically during the cycle time to pulse the 8×32 LEDs, which, over the full ~2000 mm working height of the system, gives a total of 102,400 sampled points. The full scan time is approximately 10 seconds. The active measurement range is 350 to 650 mm. Distances are measured in units of 1×10^{-1} mm from the sampled surface point to the sensor. Intensity values are also recorded and are used in the low-level processing.

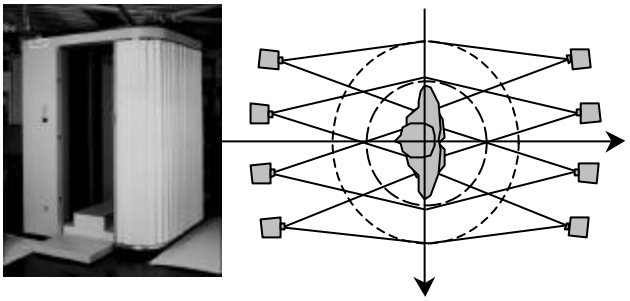


Figure 2. Hamamatsu whole-body scanner: whole system and cross-section.

The scanning heads are positioned to maximize coverage and provide as uniform a sampling as possible over the whole body (including, as far as possible, the insides of the arms and legs) when subjects stand in the center of the scanner space (figure 2). Various postures have been tested for coverage and accessibility of body surfaces. In the anthropometric posture used for the experiments described here, the subject stands with the feet between 10 cm and hip width apart, and arms held slightly away from the body (figure 3). The aim is to keep a few centimeters between the arm and torso to aid segmentation, but without distorting the shoulder and upper torso shape excessively. Subjects are scanned in a swimming hat and close-fitting underwear, or a swimsuit, to expose the body shape.

Figure 3. Adult and child subjects, showing reference scan posture.

Each unprocessed whole-body image is 300kB in size (uncompressed) and is stored initially in a tagged 256×400 matrix for preprocessing. Intensity thresholding, spatial thresholding, and reflection detection are carried out in several operations of $O(n)$ complexity, followed by outlier cleaning, of $O(n^2)$ complexity. The density of the cleaned data is approximately 7×5 mm (horizontal × vertical), with some sparser areas, for example on the insides of the arms and legs, and under the chin.

The data sets are in a common coordinate frame, where 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).

Figure 4. Body coordinate system.

3.2. Body representation criteria

For large-scale practical use the whole-body representation must be generated and rendered rapidly (within a minute or so) from the initial scan data, with no manual intervention, and should allow subjects to interact with their 3D image (rotate, zoom, etc.) in real-time on a moderately powerful computer. Since the initial data set is relatively compact there is no urgency for data compression, although this might be desirable in the future, for storage and for transmission over the Internet.

The representation must be expressive enough to convey the level of detail in which we are interested, and appropriate for the resolution and precision of the data. In the reconstructed model we would also like to represent body semantics with anatomically meaningful body segments, and to represent connectivity information for neighboring points on the surface. However, we do not want to over-specify the representation, or this might restrict the ability to explore certain feature detection approaches subsequently.

The human body is composed mainly of smooth surfaces, with points of discontinuity, for example at the cusps of 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, for example, on the face and fingers. The accuracy of the data acquired is approximately ± 1 mm [19]; the data density after cleaning is approximately 7×5 mm. This makes it appropriate to reconstruct individual features that are 3 cm or more in diameter, and to preserve features that deviate 3 mm or so in depth from the surface. 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.

It is immediately possible to dismiss simple constructive forms such as generalized cylinders, since they do not provide the representational flexibility to define the types of surface encountered on the body in a sufficiently compact way. Instead this discussion focuses on polygon mesh and parametric surface descriptions. Several good methods for general-purpose surface reconstruction from range data have been developed [28] [18] [1] but these have limited means of inferring the original surface behavior where data are sparse or missing. 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). In fact, without additional information it is not possible to infer the difference between these types of cusp. Consequently it is necessary to use a specific model of the human body to process the data correctly. For this work, it is justifiable to use *a priori* information in these low-level processes, as we are dealing only with the human body. In addition, we can, to a large extent, control the subject's stance during scanning, and capture the body scans in a reference posture, within a consistent coordinate frame.

A specific surface model for the human body has been developed by Li and Jones [24]. In their early work on whole-body surface imaging they devised a very compact torso representation containing useful symbolic information. The full surface was interpolated by cubic B-splines from key cross sections located at significant features (crotch, hips, waist, underbust, maximum bust, bideltoid and neck). However these features were located manually for each subject. In our work we would like to avoid such manual intervention. The following sections describe automated processes to generate whole-body surface representations, from the cleaned image data to a polygon mesh model, which is then used to construct a closed B-spline model.

3.3. Compound deformed cylinder model

The first surface reconstruction method uses compound deformed cylinders to represent each body segment (figure 5).

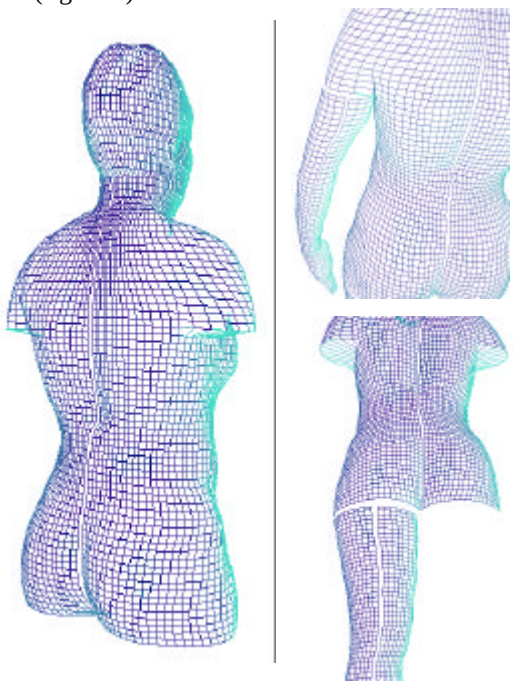


Figure 5. Segment meshes.

The method proceeds according to the following steps:

1. detect primary landmarks,
2. segment data into high-level body parts,
3. calculate medial axes for each segment,
4. convert each segment to a local cylindrical coordinate system,
5. bin data into *p* sectors,
6. remove redundant points in sectors,
7. fill gaps in the mesh,

8. join segment meshes,
9. mean smoothing.

With the exception of the primary landmark detection (which is described in more detail in section 4) all operations have *O(n)* complexity.

Primary landmarks (top of head, seventh cervical vertebra, left and right armpits, crotch, and the ends of the arms and legs) are located first, and used to label the data as belonging to the head, torso, left/right legs, left/right arms. After segmentation the data are converted to local cylindrical coordinates, and the medial axes are generated by concatenating the centroid points for each slice of data along the primary axis.

Each segment is represented by a $p \times q$ matrix of vertices, where p is the number of horizontal sample points on the segment, selected to give approximately uniform density across body segments; q is the vertical dimension of the matrix and varies according to the length of the segment (using a uniform, but selectable, sampling interval). The mesh segments are fitted to the segment data by first binning the points (figure 6) into fixed interval sectors (where the sector angle, θ, relates to p). Redundant points in each sector are removed by applying a median filter with respect to the radial distance from the segment centroid. Gaps in the matrix are filled by a simple radilinear interpolation method, which imposes an elliptical bias over interpolated areas.

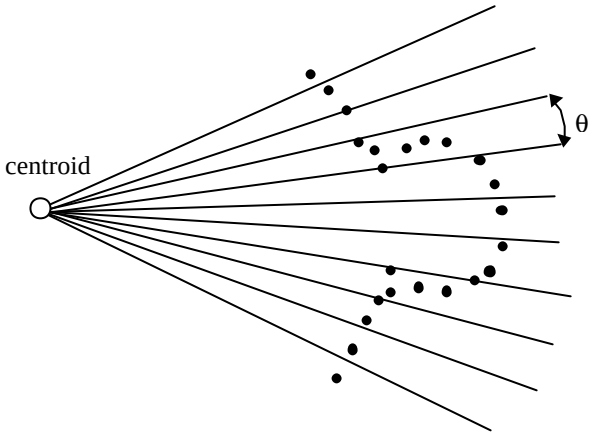


Figure 6. Data binning, showing edge of torso and arm at underarm level.

The connectivity and geometry of the vertices ensures that points in the same column of the $p \times q$ 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.

The resulting complete meshes are smoothed using a mean filter, with the window size variable for each segment. This can be selected depending on the level of detail it is necessary to preserve. For example, larger windows (diameter ~7 samples) are used on the arms and legs, while smaller windows (diameter ~3 samples) are used for the head and torso segments, where there is potentially more detail. Finally the border rows of adjacent segments are joined using a nearest neighbor method. Figure 7 shows two views on the same polygonal mesh reconstruction.

Figure 7. Segmented polygon mesh models.

3.4. B-spline surface fitting

For many applications and scales of display a polygonal mesh is an adequate representation. However, it is a faceted surface, and it would be desirable to produce a surface description that is smooth at any level of scale (except for areas where we want to allow cusps or folds, such as at the crotch and the armpits). B-splines are ideal for such a task, since they are parametric and easy to process [24]. Body surface information is represented by a set of control points, which reduces the storage demands and facilitates enforcement of continuity and smoothness along the boundaries of the segments.

In the polygon mesh structure described so far the torso is defined as a deformed cylinder from crotch to neck. For B-spline surface reconstruction this must be split into two torso segments (an upper segment from neck to armpits and a lower segment from armpits to crotch), which are truly branch-free structures with the topology of a deformed cylinder. This generates the appropriate “holes” for the arms to join the upper torso.

The next step is to resample each segment’s mesh. This is required in order to avoid high frequency artifacts, such as surface ripples, being introduced. Fitting a B-spline surface along each regular grid is then a standard interpolation problem with linear complexity, as it only requires the solution of a number of tridiagonal equation systems. If the B-splines are cubic, the degrees of freedom in the formulation of the interpolation equations can be exploited to enforce smoothness and continuity

constraints between the segments [10]. The result is a compound, multi-segment, smooth and closed surface, as shown in figure 8.

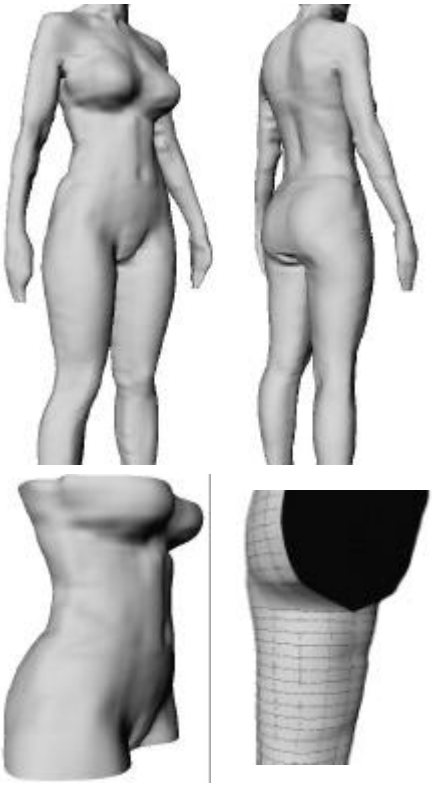


Figure 8. Spline surface models. Top: Whole body. Bottom: Closer views of torso and crotch branching (left side only).

3.5. Validation via volumetrics

From the surface representations it is possible to calculate whole-body measurements such as volume and surface area. These may be used in medical applications to determine body composition (for example, lean and fat percentages) which in turn is used in order to diagnose and monitor nutritional disorders, prescribe medication and design appropriate diets. Pressure chamber plethysmography and underwater weighing are currently used to determine body volume [13]. Although these methods are considered to give highly accurate results, they are quite intrusive on the subject, and cannot be used on the very young, old or unwell. It is therefore desirable to find an alternative technique.

The regular matrix structure of both the compound deformed cylindrical mesh and the B-spline surface representations makes it possible to calculate the volumes and surface areas for each slice separately, then integrate them to acquire an overall measurement [31]. In the polygonal mesh representation the total volume can be determined by summing the volumes relating to each quadrilateral face, if they are treated as "cheese wedges" about the slice centroid. 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. The wedge volume is then calculated as a scalene triangle extruded along the principal body segment axis.

Results in table 1 below are from a sample of approximately equal numbers of women, men and children (down to eight years of age). The first three columns show the measured volumes over a sample of 18 subjects, in liters. V_p is the volume measured using plethysmography. V_h is the volume measured using hydrodensitometry, and V_s is the volume measured using the implemented technique on 3D scanner data; e_1, e_2 and e_3 are the squared errors between the techniques, with the RMS error displayed at bottom right. Figures for scan-generated volumes are generally within less than a liter of the reference methods, with no significant trend in the error with respect to absolute body volume.

Table 1. Comparison of volume techniques.

Subject No.	V_P	V_h	V_S	$e_1 = (V_P-V_S)^2$	$e_2 = (V_h-V_S)^2$	$e_3 = (V_P-V_h)^2$
1	52.21	52.67	52.58	**0.13**	**0.01**	0.21
2	33.51	33.49	32.82	**0.48**	**0.44**	0.00
3	77.03	78.39	78.39	**1.85**	**0.00**	1.84
4	27.43	27.50	28.82	**1.93**	**1.73**	0.01
5	65.62	65.71	66.19	**0.32**	**0.23**	0.01
6	54.65	54.91	53.91	**0.54**	**1.00**	0.07
7	63.30	63.65	62.55	**0.56**	**1.20**	0.12
8	67.73	67.79	68.08	**0.12**	**0.08**	0.00
9	58.39	58.52	58.00	**0.15**	**0.28**	0.02
10	42.38	42.23	41.19	**1.41**	**1.08**	0.02
11	50.53	50.39	49.26	**1.61**	**1.29**	0.02
12	24.92	24.84	25.99	**1.15**	**1.32**	0.01
13	64.85	64.22	63.72	**1.26**	**0.25**	0.39
14	26.11	26.08	26.38	**0.08**	**0.09**	0.00
15	64.33	64.33	64.60	**0.07**	**0.07**	0.00
16	55.47	56.08	55.14	**0.11**	**0.88**	0.37
17	71.59	71.87	70.68	**0.82**	**1.41**	0.08
18	56.62	56.60	57.32	**0.48**	**0.51**	0.00
				0.73	***0.66***	*0.18*

As the techniques are under constant refinement, the mean errors tend to become smaller as development progresses. The B-spline representation makes it possible to take into account the curvature of the surface. If the surface integration is then done in a third-degree manner, we can expect a more accurate volume estimate since the fluctuation of the volume value across the slices should be described more accurately than with the polygon-mesh-and-cheese-wedge method. In the case of surface area, calculation via recursive subdivision is a well-known technique that may be used for spline surfaces.

Surface area is more difficult to validate, as there are no accurate and acceptable reference methods, although work is currently being carried out to devise a robust and repeatable technique for validation of the surface measurements [30]. Part of this includes determining an appropriate level of scale at which the surface is to be

measured. This is because the surface of any non-trivial object (anything that is not a sphere, cone, cube, block or some other simple geometrical entity) is actually a fractal-like quantity whose value depends strongly on scale.

4. Building feature detection functions

Once a high-level segmented surface description of the body is available it is possible to start progressively detecting and labeling features.

4.1. Starting assumptions: a generic model

As the subject will always be in approximately the reference posture and in a common coordinate frame, we make the assumption that these 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. Since we would like them to work well on women, men and children of medium BMI (~15-25), the operators should be tolerant to fairly large differences in body size and shape. For now a limitation is placed on non-standard body shapes, and it is assumed that the subject has two arms and legs, although exact symmetry is not assumed. The data used for exploring operators is therefore made up of a mixed group of subjects, from 8 to 58 years of age, and a range of body shapes and weights.

4.2. Exploration of feature detection operators

The aim of the exploratory phase is to find useful reusable operator primitives with which to construct higher-level detectors. Many low-level operators have been investigated to determine their discriminatory power. Recurring “idioms” are “collapsed” into reusable, parameterized operators, which are summarized at the end of this section. To illustrate this we first describe the development of detectors to locate the branching points at the armpits and crotch, which are used to segment and label points as belonging to the arms, legs and torso, according to the generic body model. The resulting segmented form is the basis of the surface reconstruction processes described earlier in section 3. In detecting these features many interesting problems are encountered.

The main characteristic of the branching points is that they have cusps, and therefore produce highly occluded regions at the very point that we want to detect. At the underarms partial cusps continue above the true armpit point, so the method must be able to detect the transition point. Once the correct level of the branching point is found, it is also necessary to detect and bisect the cusp in the horizontal plane, to label the points appropriately.

At first sight a simple frontal projection and gap detection method looks attractive, however, in practice this cannot be used on its own except in a very few cases. Such a method must be able to detect a true gap on the body within the finite sampling of data on the surface. The minimum detectable gap is significantly greater than the expected sampling density, and therefore at the armpit the gap will tend to be detected too low down the body (figure 9).

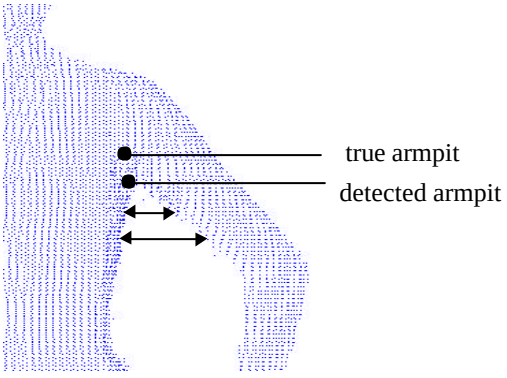


Figure 9. Frontal projection and gap detection.

Detecting the crotch point with such a method will usually fail because the legs often touch below the crotch unless the feet are placed very widely apart (which in turn distorts hip shape). Similarly, simple turning point methods fail in the same way.

Exploration of inter-point distance data reveals some interesting characteristics relating to cusps and transitions at branch points. Taking an arbitrary slice of data points, the mean Euclidean distance of each point to all other points can be computed. At cusp points, such as occur on the legs below crotch level and at the armpits, this value is low, and the value increases for points on the outside of the body, away from the cusp (figure 10). If the mean inter-point distance values for all data points in a given slice are plotted against the distance along the primary body axis, the transition points at the underarms and crotch appear as points of inflection. In practice this has not been used, however, because the points of inflection have been difficult to detect reproducibly in a fully automated way.

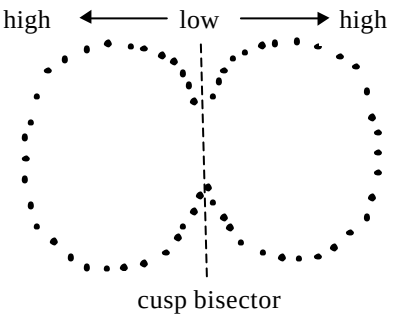


Figure 10. Inter-point means on a horizontal slice at the top of the legs; low values occur at cusp points.

A more fruitful approach has been to detect the reentrancy in the data that occurs at branching points, and where multiple body segments are present in a given horizontal slice. If the data are binned in sectors about the centroid, reentrant sectors can be detected. A threshold on reentrancy can then be used to detect the transition at the underarm level (figure 11).

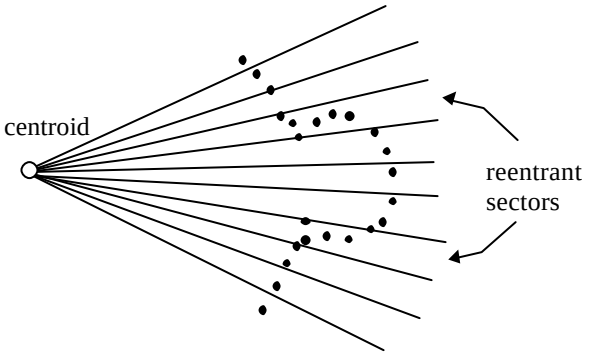


Figure 11. Reentrancy detection at the underarm level.

For detecting the crotch level it is possible to make use of changes in depth on the torso. If a constrained volume of data is sampled through the approximate center of the torso with respect to the secondary axis (body width), it is possible to detect the crotch level where the torso narrows in depth. The lowest point on the torso can be detected where the depth becomes zero. On subjects where the legs touch below this level, it is not possible to detect the zero-depth point, so in the stopping criterion it is necessary to use a threshold, which is dependent on an estimate of the subject's approximate body type. This is obtained from a function which takes as input the curve lengths of a sample set of convex hulls around the mid-torso, and the subject's detected stature, and outputs a body type indicator. A high indicator value gives a high stopping threshold and *vice versa*.

Once the level of the branch-point transition (with respect to the principal axis) has been located, it is necessary to find the cusp bisection in the horizontal plane (figure 10). Nurre [25] devised a computationally efficient method for cusp detection, which can be used to separate sets of data linearly. A similar method is used here. This is normally effective on the legs, where the geometry is relatively simple. However, on more complex areas, such as the underarm, a linear bisector is not always sufficient, and other bisector functions are being investigated to deal with this.

Other types of surface feature have been explored in a similar way, investigating the discriminatory power of different operators. It is beyond the scope of this paper to describe each in detail, but a set of selected reusable operators, used later, is summarized below:

- Point detection, transformations and attributes:

- Cartesian/cylindrical coordinates of indexed point
- centroids and medial axes
- body segment of a given point on the surface
- surface normals and derivatives at a given point
- detection of extrema in arbitrary coordinate systems
- detection of saddle, peak, pit points
- projections with respect to arbitrary axes
- point translation
- snap-to-surface, using radial, axis perpendicular and nearest point distance metrics

- Space curve generation and attributes:

- detection of ridge and valley lines
- open and closed contours, given two or more guide landmarks, in arbitrary plane
- multi-segment open curves in multiple planes
- straight line segments
- convex and surface lengths
- detection of minimum/maximum convex hulls
- simulating the effects of gravity on an elastic hull

- Partitioning and search constraints:

- surface and volume constraints via landmarks
- data partitioning by body segment
- data partitioning left/right, front/back with respect to centroid
- cusp detection
- reentrancy detection
- gap detection

4.3. Feature detection in surface anthropometry

Operators from the library can be combined to generate higher-level procedures. For example, the acromion (shoulder point) can be located as *the first point on the torso, traversing down from the nape on the ridge line of maximum z where there is a gradient less than g and then a gradient greater than h*. At present, generation of these high-level detectors is largely manual, and numeric search is used to optimize the parameters. However, fully automated synthesis is important, and other research suggests that this is a realistic goal [11] [16].

The approach is validated in surface anthropometry for garment design and manufacture, the aim being to extract 40 measurements to a few millimeters accuracy, using the ISO 8559 standard [21] and other *de facto* standard specifications. 45 women of medium build (height 156–188 cm, weight 46–74 kg and BMI 14.6–27.2) were scanned and manually measured. 25 of the data sets were used to build and train the feature detectors; the remaining 20 were used for validation. All automated measurements were extracted from a single scan of each subject, in the reference posture.

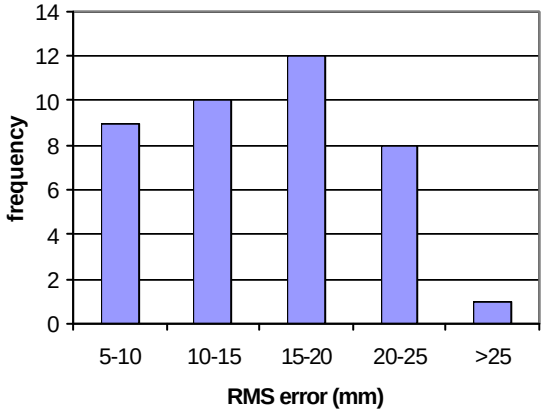
In cases where the search space is small, exhaustive search can be used for parameterization. However, dependencies exist between many of the landmarks and this gives rise to a combinatorial explosion when attempting to parameterize the operators that detect them. In these cases *morphic search* [22] is used, a form of hill-

climbing with a dynamic encoding strategy, designed to reduce the bias imposed by static search space encodings. It requires minimal parameter setting, which makes it ideal for this problem, where each evaluation cycle takes several minutes to execute. The objective function is a simple discrepancy measure. It takes as input the set of training data, applies the hypothesized landmarking functions and generates automated contour length measurements using the landmarks. The objective value is the RMS error between the manual reference measurements and the automated ones over the training set. Note that this value does not necessarily constitute *actual* error, since some measurements are difficult to take reproducibly by hand. Results on the validation set of 20 adult female subjects are shown in table 2 below. Measurements are grouped by body segment; column two shows the mean RMS error, in millimeters, for each group of measurements. This indicates, for example, that a slightly higher discrepancy appears on upper torso measurements than on the leg and arm measurements. Figure 12 shows the frequency of RMS errors over the 40 measurements.

Table 2. Automated anthropometry results.

Measurements, by body segement	RMS error (mm)
whole body	12.29
neck and shoulders	16.19
upper torso	18.14
lower torso	14.46
legs	11.93
arms	11.92

Figure 12. RMS error frequency.



The contribution of error from a given landmark cannot be isolated in a straightforward way, since most measurements depend on several landmarks, and each landmark may be used for several measurements. For example, branch-point detection determines the underarm point and crotch point location, and therefore measurements such as the crotch length from front to back waist, and underarm depth. However, these measurements also depend on the location of other features, such as the waist and the seventh cervical vertebra.

The results illustrate some of the problems that arise when comparing automated methods from scan data against manual methods, among them:

- sensitivity of landmark location and surface length to posture during scanning, particularly on areas of high deformation;
- use of a *single* reference posture to locate landmarks and extract measurements;
- lack of tactile feedback for skeletal features;
- limited access to occluded areas;
- lack of deformation of scan data, for example, with loose folds of clothing and soft flesh.

However, there appear to be many advantages to the scan-derived landmark detection processes, including:

- speed, and therefore high throughput;
- less intrusion on the subject;
- ability to implement an anthropometric standard in software, therefore potentially high reproducibility;
- ease of analytical detection of features such as maxima and minima.

5. Summary and future work

Fully automated techniques have been presented for preprocessing, skinning and progressively detecting and labeling landmark features from digitized surface data of the human body. The skinning techniques have been tested on a wide range of subjects, including women, men and children, and the resulting surface models have been used to determine body volumes. These show low error rates when compared with reference methods.

The feature detection methods have been tested by building procedures to detect key landmarks and use these to extract 40 linear and curvilinear measurements on women, according to the ISO 8559 specification [21].

The implemented techniques, from initial data, through cleaning, skinning and anthropometric processing run in less than a minute on a moderately powerful personal computer, which meets the stated speed criterion. Although these methods have generally proven to be accurate and reliable, it would be desirable to improve on this. Some improvements might be obtained by partitioning the data sets further and optimizing the landmark location for more specific subsets, rather than the relatively wide range used here.

Some of the assumptions used to reduce the complexity of processing do not hold in certain cases, for example, in subjects with scoliosis, where there may be a significant twist to the spine. In such cases the choice of axes is not fully valid, since the body “slices” may be rotated with respect to each other, and no affine transformation exists to convert the data to the appropriate representation (although this in itself is of interest in

diagnosis and monitoring). A more general approach might be to locate the secondary and tertiary axes separately for each slice, perpendicular to the primary axis but without the assumption that they lie in a plane parallel with the floor.

Now that a library of reusable operators has been developed and tested, the aim is provide more automation, by using machine learning to synthesize and optimize detectors from scratch. Another planned area of work is to investigate the combination of global deformable models with initial registration landmarks detected via feature discriminant functions.

Acknowledgements

This work has been funded by Hamamatsu Photonics and the Wellcome Trust (GR/053831). Thanks are due to Jonathan Wells at the Institute of Child Health and Nigel Fuller at the Dunn Nutrition Unit, Jeni Bougourd and Janet Ward at the London College of Fashion, and numerous others who volunteered as subjects and helped with this study.

References

[1] Amenta N, Bern M & Kamvysselis M, "A new Voronoi-based surface reconstruction algorithm", *SIGGRAPH '98*, pp 415-421.

[2] Ashbrook AP, Fisher RB, Robertson C, Werghi N, "Finding Surface Correspondence for Object Recognition and Registration Using Pairwise Geometric Histograms", *Proc. ECCV* 1998.

[3] 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.

[4] Bookstein FL, "Principal warps: thin-plate splines and the decomposition of deformations", *IEEE Trans. Patt. Anal. and Machine Intelligence*, Vol. 11, 1989, pp. 567-585.

[5] Brett AD, Taylor CJ, "A Method of Automated Landmark Generation for Automated 3D PDM Construction", *Proc. BMVC*, BMVA Press, 1998, pp. 914-923.

[6] Cinque L, Levialdi S, Malizia A, "Shape description using cubic polynomial Berzier curves", *Pattern Recognition Letters, Vol. 19*, Elsevier Science, 1998, pp. 821-828.

[7] Commean P, Smith K, Vannier M, "Automated Limb Prosthesis Design", *Visualization in Biomedical Computing, SPIE, Vol 2359*, 1994, pp. 493-503.

[8] Cootes TF, Cooper DH, Taylor CJ, Graham J, "Active Shape Models - their Training and Application", *Comp. Vis. and Image Understanding, Vol. 61 No.1*, 1995, pp. 38-59.

[9] Daanen HAM, van de Water GJ, "Whole body scanners", *Displays*, Vol. 19, Elsevier Science 1998, pp. 11-120.

[10] Douros I, *B-spline surface reconstruction of the human body from 3D scanner data*, MRes Thesis, UCL, 1998.

[11] Draper B, "Learning Object Recognition Strategies", *Symbolic Vision Learning*, Oxford University Press, 1996.

[12] 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.

[13] 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", *Clin. Sci.*, 82, 1992, pp. 687-693.

[14] Geisen G, Mason GP, Houston V, Whitestone J, McQuiston B, Beattie A, "Automatic Detection, Identification, and Registration of Anatomical Landmarks", *Proc. 39th Annual Meeting of the Hum. Fact. and Erg. Soc. (2)*, The Society, San Diego, 1995, pp. 750-753.

[15] Gower JC, "Generalised Procrustes Analysis", *Psychometrika, 40*, 1975, pp. 33-51.

[16] 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.

[17] Hodgins JK, Wooten WL. "Animating Human Athletes", *Robotics Research: the Eighth International Symposium*, Springer-Verlag, Berlin 1998, pp. 356-367.

[18] Hoppe H, *Surface Reconstruction from Unorganized Points*, PhD Thesis, University of Washington, 1994.

[19] Horiguchi C, "BL (Body Line) Scanner", *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXII, Part 5, Hakodate 1998.

[20] Iivarinen J, Peura M, Särelä J, Visa A, "Comparison of Combined Shape Descriptors for Irregular Objects", *Proc. BMVC*, BMVA Press, 1997, pp. 430-439.

[21] *ISO 8559: Garment Construction and Anthropometric Surveys – Body Dimensions*, International Organization for Standardisation, 1989.

[22] Kingdon J, Dekker L, "Morphic Search Strategies", *Proc. 1996 IEEE International Conference on Evolutionary Computation*, IEEE, NJ USA, 1996, pp. 837-841.

[23] Koenderink JJ, van Doorn AJ, "Surface Shape and Curvature Scales", *Image and Vision Computing*, Vol 10, No 8, Butterworth-Heinemann, 1992, pp. 557-565.

[24] 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.

[25] Nurre JH, "Locating Landmarks on Human Body Scan Data", *International Conference on Recent Advances in 3D Digital Imaging and Modeling*, IEEE, NJ USA, 1997, pp. 289-295.

[26] Pargas R, Staples N, Davis S, "Automated Measurement Extraction for Apparel from a 3D Body Scan", *Optics and Laser Technology, Vol. 28(2)*, Elsevier Science, 1997.

[27] Pontil M, Rogai S, Verri A, "Recognizing 3-D Objects with Linear Support Vector Machines", *Proc. ECCV*, 1998, pp. 469-483.

[28] Stoddart AJ, Hilton A, Illingworth J, "Slime: A new deformable surface", *Proc. BMVC*, BMVA Press, 1994, pp. 285-294.

[29] Turk M, Pentland A, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, 3(1), MIT 1991, pp. 71-86.

[30] Wells J, *private communication*, 1999.

[31] Yamanaka D, *Segmentation, Volume and Surface Area Calculations in a 3D Human Body Model*, MSc Thesis, UCL, 1998.