From geometric to semantic human body models

M. Mortara[a]*, G. Patané, M. Spagnuolo

[a]Istituto di Matematica Applicata e Tecnologie Informatiche,
Consiglio Nazionale delle Ricerche
Genova - Italy

The paper introduces a framework for the automatic extraction and annotation of anthropometric features from human body models. The framework is based on the construction of a structural model of the body, built upon a multi-scale segmentation into main bodies (e.g., torso) and limb features (e.g., fingers, legs, arms). The decomposition is independent of the body posture, it is stable to noise, and naturally follows the shape and extent of the limb features of the body. The structural description of the human body is turn into a semantic description by using a set of rules and measures related to the features and by reasoning of their configuration. Results are shown both for scanned body models and virtual humans, and applications are discussed in relation to several tasks of the animation process.

1. Introduction

The automatic recognition of features in free form shapes is a challenging issue, especially when the semantics underlying the feature definition is related to an intrinsically not formalized context. This is the case of features of the human body: neck, legs, thigh, elbow, and many other terms which identify relevant body parts, refer to portions of the body shape which cannot be precisely coded or identified by a mathematical formulation. Also, some of the body features are composition of other features: a leg is defined by the shin, the calf, the thigh, its articulation depends on the knee and by the ankle and hip which connects it to the body. At the same time, in the last decade we assisted to a growing interest in computer-aided methods to study and analyse the shape of the human body in digital contexts. Due to the advances of scanning technology, it has been possible to carry out one of the largest anthropometric survey within the CAESAR project [13] which has made available a set of data of over 10.000 individuals in digital form. Traditional anthropometric practices largely rely on the knowledge of the expert performing the manual measuring of sizes and shapes, using tapes and calipers, and on the use of different postures to get a precise evaluation of the various anthropometric parameters. Body size measuring tools are generally limited to one-dimensional information while the new $3D$ body scanning technology provides capabilities such as segmental volumes and surface areas [13], and may support a more reliable and surely less expensive way to measure shapes. The potential impact of $3D$ surface anthropometry is therefore very high in fields related to ergonomics, cloth and prosthetic design, obesity studies, and many more [5].

Human body models, being either scanned or modelled, are of high interest for the animation industry as well. A modelling system based on human features would greatly improve several steps of the animation pipeline. A human body is usually animated by associating a so-called control skeleton to the $3D$ shape, which is a connected set of segments corresponding to limbs and joints, that is, the points where the connected limbs may move. Unfortunately, both the skeleton extraction and the establishment of the correspondence between the geometry and the skeleton can be strongly time-consuming. Some commercial software packages include tools for skeleton-based animation, like Maya and 3D Studio MAX. Nonetheless, the creation of a control skeleton may require several hours of work, and the user must possess a fair degree of proficiency with a package to obtain even a rudimentary motion. A decomposition of the human body into relevant features would therefore contribute to speed up these applications.

In this context, we present the results of a new

---

*E-mail: {mortara,patane,spagnuolo}@ge.imati.cnr.it

framework for automatically annotating a human body model with information related to the body features. The annotation and reasoning about the features is supported by the segmentation of the human body models into geometric features, by the creation of a skeleton which encodes the feature attachment relations, and by a measuring scheme which allows to attach quantitative descriptors to each part. The segmentation approach is based on the multi-scale method called *Plumber*, developed for segmenting a surface into generalized cones and cylinders [7,8]. *Plumber* defines the basic decomposition of the body model into tubular-like parts and main body, usually corresponding to the torso in the context of human body models. The surface patch corresponding to the torso is further segmented in order to extract symmetry regions and areas of influence of the various attachments of joints to the torso. Based on this geometric segmentation, a semantic model is built as an annotated shape-graph where each node corresponds to a relevant feature represented by its centreline skeleton and a set of cross-sections. Reasoning can be performed on the shape-graph to deduce further measures and identify compound of features, as well as to classify body models using standard anthropometric rules.

The main characteristics of the method proposed are the ability to produce a semantically consistent shape-graph of human body models independently of their posture, and the automatic association of skeletal lines to body limbs together with cross sections and size parameters. Due to the properties of the *Plumber* method, the segmentation is stable with respect to noise in the model.

The paper is organized as follows: first, previous work on the characterization of the human body is reviewed in Section 2; the segmentation approach underlying the presented framework is briefly described in Section 3, and the reader can find full details in [7,8]; in Section 4 the graph used to code the human body and its use for extracting and computing relevant anthropometric measures are described; in Section 5 the use of the framework is discussed in relation to the input data characteristics and applications. Finally, conclusions and future work are drawn.

2. Previous work

The paradigm of *shape segmentation* has been largely studied in the literature, both for generic and specific application contexts as well as for discrete and continuous shape representation schemes. In the specific context of human body, the segmentation has been often addressed in parallel with the automatic location of landmarks. The first attempts were devised for working on the point clouds resulting from the scanning sessions. The work presented in [12] approaches the problem of recognizing relevant body parts as an aid to the optimisation of the body measures themselves. $3D$ body scanners acquire data along horizontal slices and the quality of the resulting measurement obviously depends on physical limitations of the scanning device and the body posture. If the body is scanned using a natural standing position, indeed, the arms will generally touch the torso and in this area it is impossible to distinguish points on the arm from points on the torso. To solve this problem, the method proposed in [12] is aimed at the detection of sharp variations of the contour shape, which are used to trim the arm data and to reconstruct the missing torso data. The method does not have a general validity, as it depends on the specific posture, and the segmentation provides a quite poor description of the body features.

The method described in [9] and refined in [10] adopts an approach based on the alignment of a stick figure, representing the abstract skeletal structure of a body in a standard pose, to the raw data. The stick figure is composed by six linear segments, that are aligned to the scan data under user control. This method provides the segmentation and also the computation of the feature centrelines, but it suffers of the same dependence on the body posture as the method previously described. The segmentation indeed uses information of the horizontal scanning contours, and it produces a space-based and not a shape-based decomposition.

The analysis of horizontal slices has been more recently used in [16] within a framework which segments raw data according to their membership to a body part, reconstructs the shape of the body parts, and attaches them together in order to build the full body reconstruction. Again, the segmentation is functional to the reconstruction and the boundary of the

body parts is horizontal, as it is computed relying on the horizontal scanning slices.

The advantages of defining the segmentation directly on the body surface instead of on the space occupied by the body has been introduced in [17], where the body is segmented using concepts of Morse theory. A topological graph which codes the evolution of the level sets of a real-valued mapping functions is used to segment the body shape. Terminal and branching nodes of the graph correspond to critical values of the mapping function, which is chosen as the integral geodesic distance. The graph connectivity is used to segment the shape into parts which represent the relevant features of the body. As the authors point out, one of the main advantages of the method is to provide a posture independent segmentation of the shape.

3. Multi-scale geometric shape segmentation

The proposed decomposition of the human body and its semantic annotation rely upon a general shape segmentation method called *Plumber*, developed by the authors and fully described in [7,8]. For the sake of clarity, we summarize here the main aspects of the segmentation, while details can be found in the cited references.

The *Plumber* approach to shape decomposition is aimed at the extraction of tubular features of a 3D surface represented by a triangle mesh. The *Plumber* algorithm segments a surface into connected components that are either body parts or elongated features, that is, handle-like and protrusion-like features, together with their concave counterparts, i.e. narrow tunnels and wells. The recognition is based on the classification of vertices according to geometric and morphological descriptors evaluated on neighbourhoods of increasing size. The set of neighbourhoods associated to each vertex is defined by a set of spheres, centred at the vertex, and whose radii represent the scale at which the shape is analysed. The number of connected components of the intersection curve between each sphere and the surface gives a first qualitative characterization of the shape in a $3D$ neighbourhood of each vertex. Then, the evolution of the length ratio of these components with respect to the radius of the spheres can be used to refine the classification and detect specific features, such as sharp protrusions

or wells, mounts or dips, blends or branching parts. For example, for a thin limb, the intersection will be simply connected for a small radius and it will rapidly split into two components as the radii increase. For a point on the tip of a limb, the intersection will remain connected, but the ratio of its length to the radius of the sphere will be decreasing. See Figure 1 for an example of the process.

Plumber specializes this approach to the detection and extraction of tubular features. At the first step, seed vertices are located and clustered to form candidate seed regions which are then used to compute the first reliable tube section, called the *medial loop*. This loop is ensured to be around each candidate tube and works as a generator of the feature. Then, the medial loop is moved in both directions on the shape, by using spheres placed not on the surface but at the barycentre of the medial loop iteratively and until the tube is completely swept. The stop criteria of the iterative procedure are discussed in the paragraph *Shape Segmentation* at page 5. The tube detection works in a multi-scale setting, starting with the extraction of small tubes first. Assuming that the shape is represented by a mesh triangle $\mathcal{M}$ and that we are using a set of levels of detail $\{r\}$, the steps of the extraction procedure are presented in the following paragraphs.

Vertex classification.

For each vertex $v \in \mathcal{M}$ and scale r, we consider the surface region containing v and delimited by the intersection between $\mathcal{M}$ and the sphere $S(v,r)$ of center v and radius r; let γ be the boundary of this region and let us discard all other regions of intersection between the sphere and the mesh that might occur but do not contain v (see Figure 1(a)). If γ has only one connected component (see Figure 1(b)), then the surface around v is equivalent to a disc and its curvature at scale r is approximated by the non-negative ratio $G_r(v) := l_\gamma/r$ [4], where l_γ is the length of γ. Furthermore, v is classified as *planar* if $G_r(v) \approx \alpha$, *sharp* if $G_r(v) < \alpha$, and *blend* if $G_r(v) > \alpha$, where α is a given threshold.

Let us now suppose that γ has two connected components, and in this case the vertices are labelled as *limb*. The vertex v at scale r is classified as *cylindrical* when the ratio between the maximal and minimal length of γ_1 and γ_2 does not exceed a given threshold ϵ, that is $l_{\gamma_1} \approx \epsilon l_{\gamma_2}$; otherwise, it is labelled as *con-*
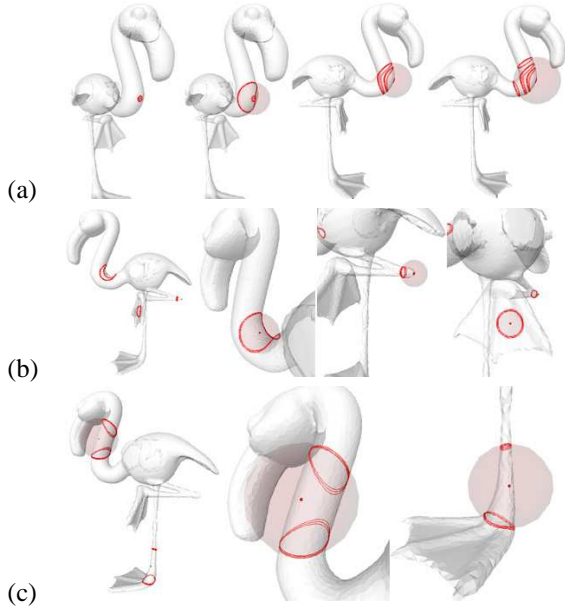
Figure 1. (a) Evolution of the intersection curves between the input surface and a set of spheres with the same centre and increasing radii, (b) classification of blend, sharp and planar vertices, (c) tubular features classified as cylinders and cones.

Figure 2. (a) Selection of a level of detail r, (b) classification of vertices, (c) identification of a seed limb region, (d) medial loop, (e) iterations, (f) extraction and abstraction of the tubular feature as a skeletal line and a set of contours.

ical (see Figure 1(c)). If γ has 3 or more connected components, v is a *branching* and we do not consider other geometric descriptors.

The set of radii is automatically set by uniformly sampling the interval between the minimum edge length and the diagonal of the bounding box of $\mathcal{M}$. These parameters, as well as those ones used for the classification of the vertices (i.e., $\alpha := 2\pi$, $\epsilon := 2$), can be selected by the user if an a-priori information on the input shape is available or if he/she is searching for some specific configurations (e.g., vertices whose sharpest angle is less than a given value). The choice of α and ϵ can obviously take into account a specific application context, as detailed in Section 4.

Shape segmentation.

The vertex classification is used for defining a shape segmentation into connected components which are either *tubular features* (i.e., regions which
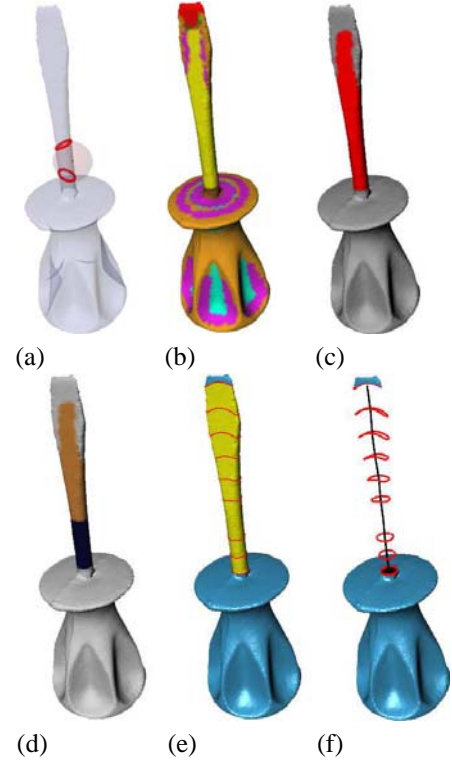
can be described as generalized cones or cylinders) or *body parts* (i.e., regions which connect tubular features). To this end, we proceed in the following steps: we select a level of detail r and we identify *seed limb-regions* as the maximal edge-connected regions of limb-vertices with respect to a depth-first search (see Figure 2(a-c)). Then, we compute the *medial loop* of each seed limb-region which represents the generator of the feature, and it is used for its expansion until a stop criteria is satisfied. In Figure 2(d), the medial loop is the boundary of the dark region, while the growing phase is shown in 2(e). Then, we iterate the process on $\mathcal{M}$ by considering the next level of detail.
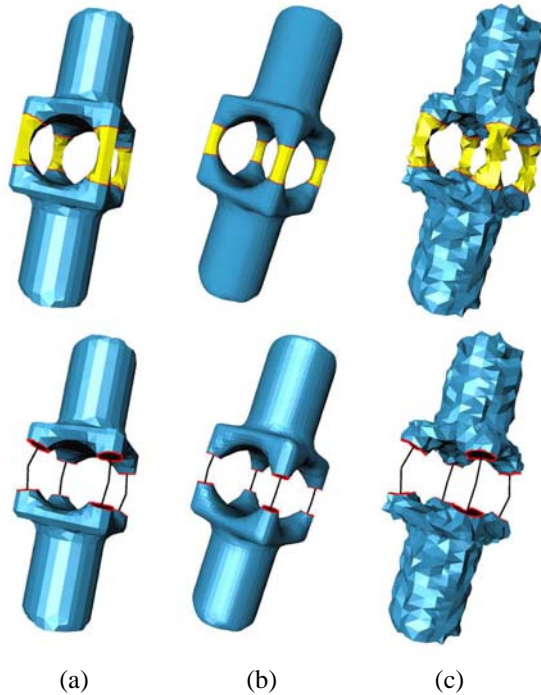
Figure 3. Shape decomposition when the geometry of the input surface is (a) coarse, (b) smooth, (c) and affected by noise.
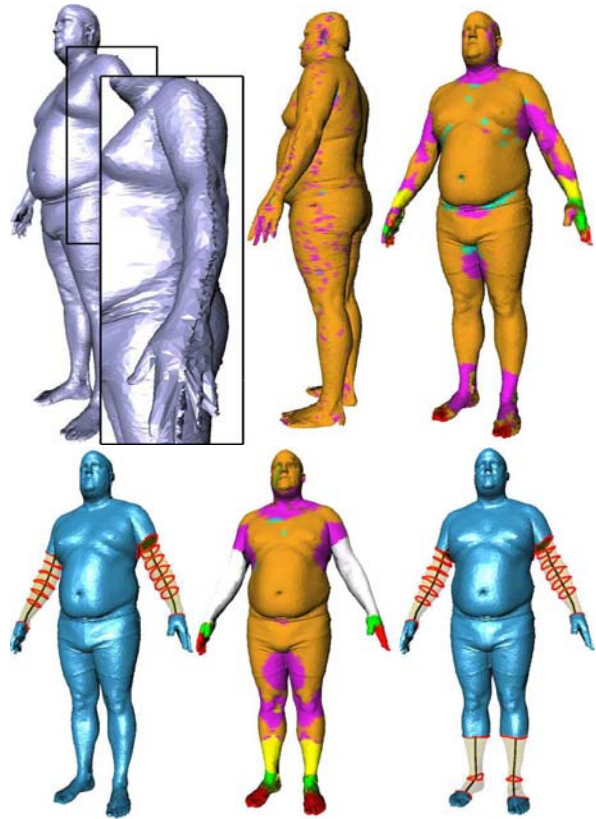

Figure 4. Main step and robustness to noise of the *Plumber* segmentation on a scan model consisting of 13790 vertices; timings are given in Table 1.

The radius, or scale, of the sphere influences two steps of the tube recognition process: once for the morphological analysis, to locate the *limb* vertices and candidate tube regions, and once for the tube growing phase. The stop condition of the tube sweeping phase is decided either by a threshold on the variation of the intersection length, by the ending of the tubular feature itself, or by the splitting of the tube at a branching site. If the tubular feature ends, the tube is called *cap* and it will have only one boundary, as it is shaped as a generalized cone. The extraction of tubes adopts a fine-to coarse strategy, marking triangles as visited while the tube grows so that they are not taken into account at the following steps. At the end of the whole process, tubes are labelled with respect to the scale at which they were found. The connected components of the shape which are not classified as tubular features define the body parts of the input surface. In Section 4, an extension and refinement of the *Plumber* segmentation for body parts will be presented.

Shape segmentation properties.

The described segmentation method is robust to noise and independent of the vertex sampling and connectivity regularity (see Figure 3). In fact, the computation of the intersection curves among $\mathcal{M}$ and the selected set of spheres uses the connectivity structure only for the computation of γ, while the classification of a vertex p as belonging to a tube at scale r (i.e., $\|p - c\|_2 \leq r$, with c centre of the current sphere) relies only on the set of vertices. The curva-

Table 1

Timings (in seconds) of the *Plumber* segmentation shown in Figure 4.

Task	$R = 1$	$R = 5$	$R = 8$
Tailor	11 sec.	51 sec.	61 sec.
Medial Loop	— —	21 sec.	50 sec.
Tube Construction	— —	3 sec.	3 sec.

ture evaluation on the real scan model shown in Figure 4 is performed at three radii; at the smallest radius, the segmentation presents many tiny regions, mainly composed by one vertex only, and due to the noise in the data. At larger scales, the influence of noise on the curvature computation and tube extraction becomes negligible; in fact, the intersection between the sphere and the mesh is computed exactly and it is not affected by the underlying mesh quality. Timings are reported in Table 1.

Performance of the algorithm.

The time complexity required by the tube extraction depends on the following stages: curvature evaluation, medial loop computation, and tube growing. In the worst case (i.e., when almost all the vertices fall inside the sphere), the curvature analysis at each vertex takes $O(n)$-time, with n number of vertices; therefore, the complexity for the whole model is $O(n^2)$. Even if the average case is less complex and depends on the radius size, this is the slowest phase of the workflow. The medial loop computation uses Dijkstra's algorithm and takes $O(m^2 \log m)$-time for each seed limb region, where m is its number of vertices. The tube growing phase, which actually constructs the tubular structures, is linear in the number of triangles belonging to tubes. Once the final segmentation is built, the shape-graph is evaluated in linear time with respect to the number of patches.

4. The Semantic Body Model

In this section, we describe how to extract the semantic content, which is implicit in the digital model, from the geometry, the structure, and the knowledge pertaining to the domain. Since the input model represents a human body, either virtual or scanned, its rel-

evant tubular features will identify arms, legs, neck, and fingers. The torso and its symmetries are also important data in the anthropometry domain. In general, *Plumber* will not find directly these features, but some of their parts. For instance, a hand will be segmented into five small tubes, possibly with their associated caps, all of them being attached to the same body part. Reasoning on the relative sizes of the features and on the attachment relations among them, makes possible to recognize and automatically measure semantically relevant parts of the human body. To this end, let us explain how the segmentation obtained using *Plumber* is coded as a shape-graph, and then we will show how different descriptors can be associated to the shape-graph in order to produce a semantic body model.

Centrelines of tubular features

Each tubular feature $\mathcal{T}$, extracted at scale r, either conical or cylindrical is abstracted by a skeletal line defined by joining the barycentres b_i of the intersection curves γ_i between $\mathcal{T}$ and the set of spheres used to sweep the tube. Since the shape and position of the tubular feature could be arbitrarily complex, the intersection curves are good descriptors for the cross sections of the tube along the centreline. At the same time, positioning the centrelines at the barycenters of the intersection curves allows us to follow the extent of the feature at the resolution that the application requires. It is important to notice that, from the point of view of the measures that will be introduced later on, all tubular parts are represented with a number of intersections that depends only on the scale and size of the feature, thus ensuring consistency of the classification among different parts of the body.

Refinement of the body parts

While conical and cylindrical tubular features have a specific shape, body parts can be arbitrary shaped. It might be interesting to further segment each of them in order to identify symmetries in their shape and define the influence areas of the attached features for supporting the localization of joints (see Section 5).

Let us consider a body primitive $\mathcal{B}$ with $k \geq 3$ boundary components γ_i, $i = 1, \ldots, k$; in order to determine the area of influence of each boundary, a reasonable approach is to cluster the vertices of $\mathcal{B}$ which
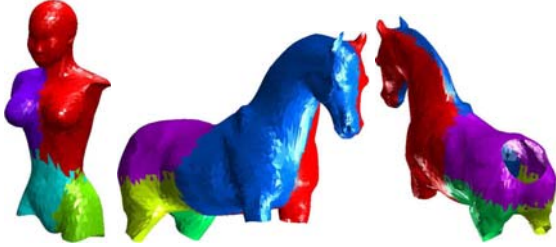
Figure 5. Voronoi-like regions of the body primitives shown in Figure 9 and 15 respectively.

Figure 6. Surface segmentation of the bi-torus into a body primitive and two tubular features and its shape graph.

are closer to the same boundary component with respect to the geodesic distance. Instead of using this approach, which is time consuming and sensible to the connectivity regularity of $\mathcal{B}$, we use a parameterisation of the body part on a planar domain Ω isomorphic to $\mathcal{B}$ and with respect to one of its boundary components, using the approach presented in [11]. If $\varphi : \mathcal{B} \to \Omega$ is such parameterisation, a vertex $p \in \mathcal{B}$ can be associated, i.e. clustered with respect to, the boundary γ_s such that:

$$\|p^\star - pr_{\beta_s}(p^\star)\|_2 = \min_{j=1,\ldots,k}\{\|p^\star - pr_{\beta_j}(p^\star)\|_2\}$$

where pr_{β_j} is the orthogonal projection onto the convex curve $\beta_j := \varphi(\gamma_j)$, and $p^\star := \varphi(p)$. Therefore, the use of the geodesic metric on $\mathcal{B}$ has been replaced by the evaluation of the Euclidean distance on the parameterisation domain. At this stage, regions with the same area reflect a symmetry of the patch (see Figure 5).

Coding the features in an adjacency graph

The surface decomposition and the skeletal lines of tubular features are coded in a connectivity graph which represents the spatial arrangement of the tubular features onto bodies. The *shape-graph* nodes are the extracted primitive shapes, while the arcs code the adjacency relations among them (see Figure 6). In general, each arc between two adjacent nodes falls into one of these cases: cylinder-body, cylinder-cylinder, and conical-cylinder. The cylinder-body or cylinder-cylinder adjacency is called H-junction (i.e., *handle-junction*) if both boundaries of the cylinder lay
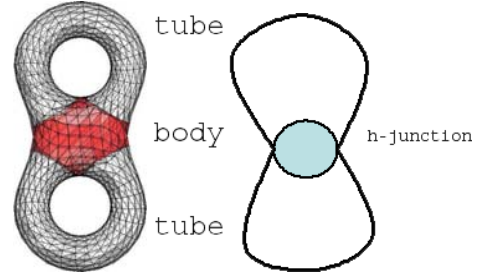
on the same body, or cylinder; in this case, the arc induces a loop and the cylinder locates a handle on the input model. In the case of human bodies, this might happen if, for example, the hand touches the leg or the torso. Finally, if only one boundary of the cylinder belongs to the cylinder-body, the adjacency is called a T-junction.

Semantic descriptors and reasoning

While the adjacency relations in the graph define the structure of the human body with respect to the decomposition, the geometry of the features are further characterized by the following descriptors. Each *cylindrical* node is uniquely labelled in the graph and it is stored with:

- the scale r at which the tube has been found;

- the set of its approximated cross sections (i.e., sphere-tube intersections) and the average radius;

- the set of its centreline points (i.e., barycentres of the sphere-tube intersections);

- the orientation of each segment of the centreline (*turning*);

- the centreline axis length and the approximated volume.

Each *conical* node is stored with the same attributes as tubes, except the radius of the average cross section which is replaced by the radius of its basis section; in
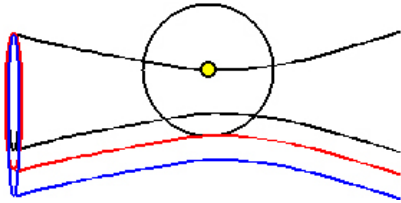
Figure 7. The sphere of radius r centred in the yellow vertex has two intersection curves if the maximum diameter of the tube section is less than $r/2$ (black line), and only one if it is greater than $r/2$ (blue line). The red line represents the limit value of $r/2$.

Figure 9. (a) Two seed regions for each leg are found at the same scale: the knee and the ankle. (b) Result of the tube growing on the seed regions in (a). (c-d) Result obtained by joining the adjacent tubes in a single tube.
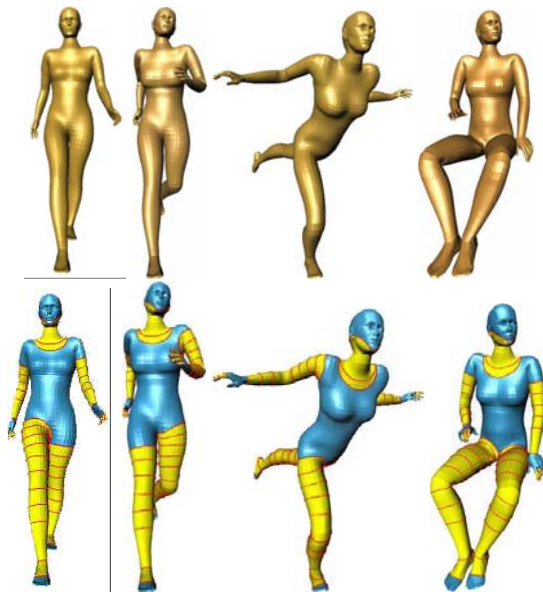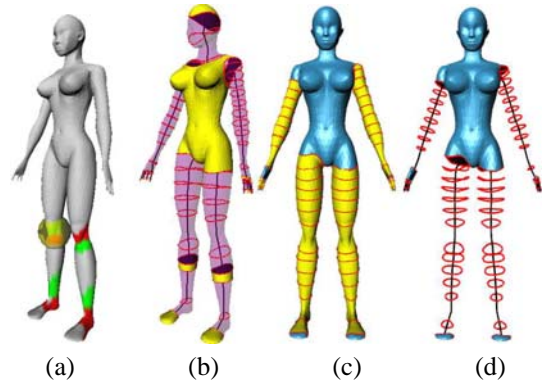

Figure 8. First row: input models of virtual humans in different postures. Second row: cylindrical features are depicted in yellow, conic and body features are shown in blue. The model consists of 5775 vertices and the algorithm takes 18 seconds for the curvature analysis and 2 seconds for the other phases.

this case, also the maximum of the Gaussian curvature in the region is stored. Each node of type *body* is stored with the number of its boundary components, its approximate volume, and its refined segmentation.

Based on the structural and geometric information, it is possible to reason on the semantic aspects of the body model. First of all, there is a strong relation between the scale r at which the feature is found and the minimum section size of the tube it represents. A sphere of radius r will label as *limb* all the vertices lying on a tubular part whose section has a maximum diameter less than $r/2$ (see Figure 7). Therefore, running *Plumber* from smaller to larger scales, we expect to recognize at first fingers, then arms, legs, and eventually the neck. Note that what is important for the identification of a tube is the minimum section size to start the growing: wrist are found first as candidate tubes, and then the tube growing constructs all the arms; the same for legs that are identified starting from the ankles, and so on. For this reason, the neck is likely to be the last tube found; its section is usually larger than that of the ankle (see Figure 8).

The arms and legs might be found as the composition of two tubes: this case happens rarely, and usually for designed virtual human where the shape of the body part can emphasize joints for artistic reasons,
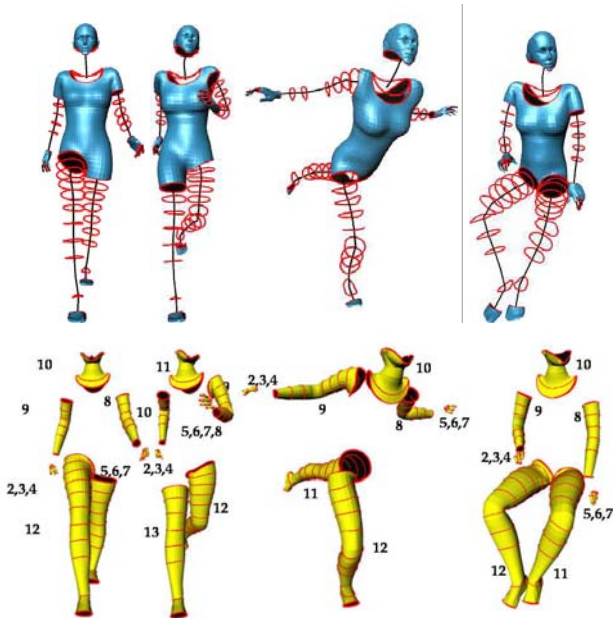
Figure 10. First row: semantic body model of the virtual humans shown in Figure 8. Second row: identifiers of the features, whose shape parameters are reported in Table 2.
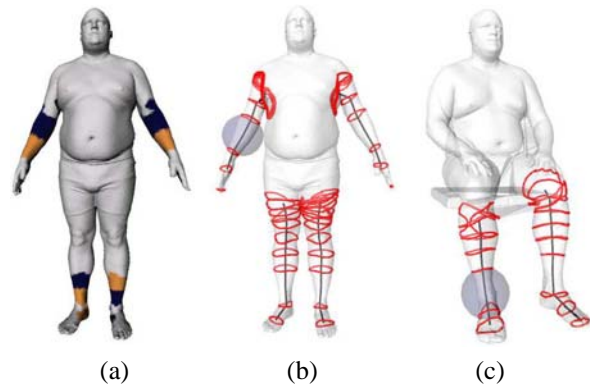

Figure 11. (a) Seed regions identified by *Plumber* on a fat man model. Note the low joint between the legs. (b) Identified tubular features. Tubes representing the legs stop in correspondence of the joint and the length of arms and legs becomes similar. (c) The same man in a sitting posture, with arms lying on the body. Arms will not be recognized as tubes, and legs stop when the seat is intersected.

like in the example given in Figure 9(a-b).

The knee sections may be so small to be recognized as candidate tubes at the same scale as the wrist and the ankle, respectively. This produces two tubes for each limb to grow, and where they meet, each one stops. This situation is handled by checking if the last computed tube section lies completely onto another tube; the two tubes are kept separately but a link is included in the shape-graph to underline that they form a single semantic tubular feature, while tube properties, such as length and section size, will be computed as they were two single tubular features (see Figure 9(c-d)).

The scale attribute together with the tube length, the information about the section size, and the shape graph make it possible to classify each tube as arm, leg, finger, or neck; this step is currently under development but it has already been validated by results shown in Figure 10 and Table 2 at page 12. All the

three parameters are needed because of the many possible types of result. It may happen that the neck is too short and wide to be recognized, and the same may happen for fingers. It is quite frequent, indeed, to identify only some of the fingers, usually missing the thumb which is the thickest and shortest. If at least some of the fingers are recognized, we are able to identify the arms through the shape graph, and consequently all the other limbs.

We cannot be completely confident on the tube length either: for fat humans, legs intersect before the hip (see Figure 11(a-b)) and the corresponding tubes are shorter, comparable with the arm length. This may also be caused by the posture, as it happens for the sitting man in Figure 11(c), where only the foreleg can be recognized. As shown in Figure 12, fatness represents the major problem, since *Plumber* does not classify as tubes limbs that are too short with respect to their thickness.

The axis inclination gives us a precious information about the body posture. In the context of virtual humans, we can exploit the fact that limbs are rigid except at the joints; therefore, the tube axis will be
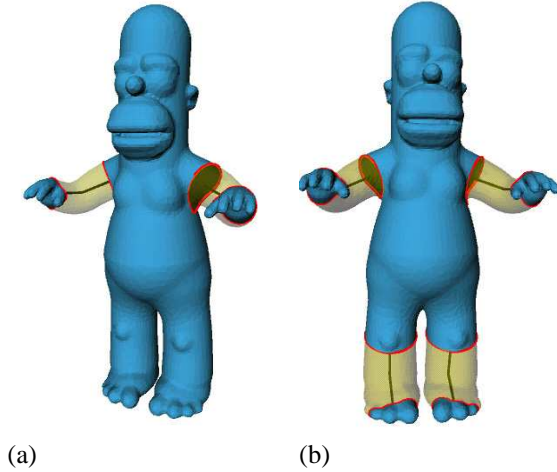
Figure 12. Tubes identified at (a) the first and (b) second scale.

surface models must be normalized[2]. For the triple (a, b, p), let $u := (a - p)$, $v := (b - p)$, and α be the turning value at p computed as previously described. When (a, b, p) lay on a straight line, u and v form at p an angle of π corresponding to a null turning; when this angle is 2π, i.e. $a \equiv b$, a maximum turning occurs in p (see Table 2).

Comparing the values in Table 3 obtained for virtual and real body models, some further conclusions can be drawn on the efficacy of the descriptors. For virtual humans, the turning value quite nicely discriminates between different postures and for real body models we may see that the different sizes of the body is reflected by the changes in the cross section size.

Finally, approximating each tube with truncated cones of circular bases, each having the same length of the corresponding tube cross section, enables to calculate an approximation of the feature volume as the sum of the volumes of each building part. Then, the ratio volume/length gives an hint on the human limb fatness, in analogy with the *Body Mass Index* (weight/length, see Figure 13), and the ratio volume/length discriminates between two individuals of same limb thickness but different height (see Figure 14 and Table 3).

5. Discussion

Our approach subdivides the model into limbs and body. The scales (i.e., the radius of the spheres used for the characterization and tube growing phases) can be automatically tuned on human anatomic measures such as approximated section of fingers, wrist, forearm, arm, ankle, calf, thigh, neck. Then, the geometric attributes of the recovered tube sections and axis can give information about the characteristics of the human model; for instance, the approximated volume of limbs with respect to their length may give the amount of fatness/thinness. The neck will be recognized as a tube only if it is thin and long, and the same applies to fingers. Also, the approximated volume of

nearly straight, except in a few points, which identify the torsion in the articulation sites. Note that a tube may have at most three articulations: for instance, a leg may comprehend the ankle, knee, and hip joints. Again, the shape graph is used to discriminate each joint, giving an "outward" ordering to each tube, from its attachment to the body towards the tip of the protrusion. We compute the turning α at each node p of the tube axis of $\mathcal{T}$ as $\alpha := \cos^{-1}(< u, v > /(\|u\|\|v\|))$ where u and v are the vectors of $\mathcal{T}$ which share p. For each triple (a, b, c) of consecutive points along $\mathcal{T}$, the cosine of the angle formed by ab and bc discriminates between an acute and an obtuse angle (a turning greater or smaller than $\pi/2$) but do not distinguish a "right" from a "left" turning with respect to a fixed coordinate system. Since the cosine function is bounded, each turning value belongs to the range $[-1, 1]$ and it can be directly used for comparing virtual humans in different postures. On the contrary, tube length, section, and volume depend on each model measure unit; then, before running the morphological analysis the

---

[2]Given two triangle meshes $\mathcal{M}_1$ and $\mathcal{M}_2$, we normalize them by applying a uniform scaling on their vertices in such a way that the new surfaces belong to the unit cube while maintaining their relative proportions, that is, we normalize the vertices with respect to the constant $C := \max\{c_1, c_2\}$ with $c_k := \max_{p_i \in \mathcal{M}_k}\{p_i^{(j)}, \ j \ = \ 1, 2, 3\}$, $k \ = \ 1, 2$, and $p_i^{(j)}$ j^{th}-component of p_i.
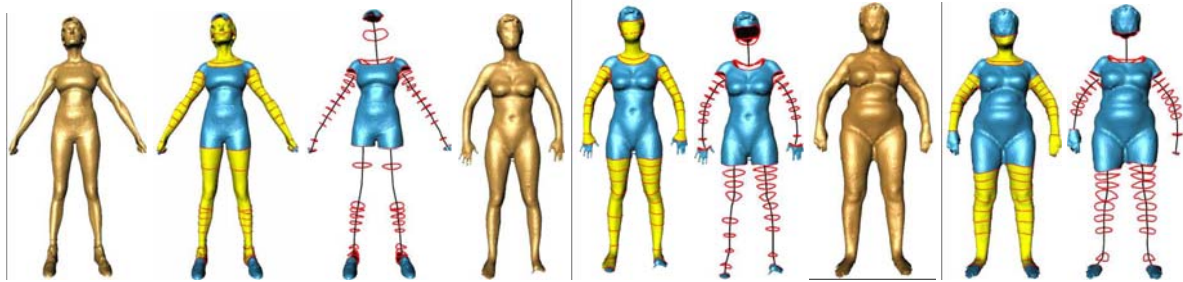
Figure 13. Analysis of the thinness and fatness based on the ratio between the volume, size, and length of each body feature.

the remaining body component can be checked in this direction.

In the proposed approach, we made no assumptions on the method used to produce the human body model, which can be either an acquisition or a modeling process. Differences in the results depend however on the type of input model.

In the case of a scanning process, the model is likely to consist of a huge amount of points. The morphological characterization which represents the first step of the tube recognition on such dense surfaces is indeed more precise. On the other hand, the mesh produced by the triangulator associated to the scanner device usually produces meshes with holes (e.g., due to occlusions which might occur in correspondence of armpits), or tends to join patches of surfaces which are separated but very close in space (e.g., the base of fingers or of the thighs). It is also true that usually human body scans have been acquired in a standard posture, which sees the human in a standing position, with legs and arms straight and completely stretched, and closed fists. In this case, it is not possible to get fingers as tubular features. We used models captured from real humans to prove the robustness of our algorithm to manage huge, noisy, or corrupted models.

In the case of the generation of the virtual human by the modeling act of a digital artist, the mesh quality is very different: it obviously consists of much less points and the surface is smoother. Generally, fine details are not provided and the computed morphological characterization is quite coarse; nonetheless, the tube extraction is facilitated and much faster. Furthermore, tube boundaries do not suffer of a poor

mesh resolution since intersection curves between the mesh and the sphere used for the tube growing are inserted as constraints to refine the mesh just in correspondence of tube extreme boundaries. The virtual human produced by modeling usually assumes different postures, which are set by the digital artist "by hand" using skeleton-driven animation packages (see *www.maya.com*); moreover, details like fingers are usually well formed. These models are interesting test cases to prove the ability of our approach to capture human limbs in arbitrary postures.

Another application of the proposed framework is the automatic detection of landmarks on the body model. In this case, the problem consists of deriving anthropomorphic features from a database of human models and identifying meaningful landmarks [9,13,14] for applications to database indexing and matching, and for animation purposes. Many of these features, such as concavities at eyes and navel, tips of fingers, nose, ankles, blends on armpits, and so on are directly and automatically detected by our method without requiring a further user interaction.

The tube descriptors can be used also for identifying human body models in a biological database: in fact, the ratio between upper/lower limbs defines the "*intermembral distance*" and has been long studied in biology as a discriminant factor between bipedi and quadrupedi. From experimental results, it is known that for human beings this ratio is nearly 72, and can be easily applied to digital models in a database to discriminate biological species using our approach (see Figure 15).

Finally, the extracted skeleton can serve as a ba-

Table 2
Shape parameters of the features detected in examples of Figure 10: the first row corresponds to examples (a) and (b), and the second raw to examples (c) and (d); the first column contains the feature identifier, the second the tube length, the third the α turning value and finally the fourth the average section length, as defined in Section 4.

N.	Tube Length	Max. α	Aver. Sect. Length	N.	Tube Length	Max. α	Aver. Sect. Length
2	7.07	-0.94	4.94	2	5.27	-0.95	5.21
3	5.50	-0,93	3.63	3	5.51	-0.93	3.61
4	5.51	-0.95	3.39	4	5.52	-0.95	3.39
5	7.07	-0.94	4.94	5	3.65	-0.94	4.16
6	5.47	-0.95	4.02	6	7.13	-0.93	4.76
7	5.50	-0.95	3.40	7	5.48	-0.93	4.08
8	39.91	-0.91	23.26	8	5.47	-0.93	3.95
9	39.95	-0.95	23.17	9	42.91	-0.39	26.60
10	17.84	-0.94	67.34	10	39.58	-0.61	23.73
11	82.17	-0.43	39.22	11	17.61	-0.92	67.65
12	93.14	-0.79	43.98	12	91.30	-0.74	43.86
				13	74.86	-0.73	36.97

N.	Tube Length	Max. α	Aver. Sect. Length	N.	Tube Length	Max. α	Aver. Sect. Length
2	5.30	-0.95	5.22	2	7.15	-0.94	4.95
3	5.56	-0.96	3.21	3	3.63	-0.96	4.64
4	5.55	-0.94	3.33	4	5.57	-0.94	3.26
5	5.30	-0.96	5.25	5	7.15	-0.94	4.95
6	3.61	-0.96	4.71	6	3.63	-0.92	4.69
7	5.55	-0.94	3.28	7	5.57	-0.94	3.27
8	39.87	-0.69	23.60	8	40.30	-0.97	23.28
9	48.23	-0.76	28.21	9	40.07	-0.72	23.46
10	17.70	-0.96	69.48	10	18.03	-0.95	69.23
11	92.23	-0.92	44.29	11	89.31	-0.50	42.53
12	89.61	-0.10	41.99	12	96.30	-0.53	44.07

sis for building the animation control skeleton because many of the joints and segments found by our method correspond to segments of the control skeleton. Our method could be especially useful for animating digital model of real humans, because in this case it is much more difficult to associate automatically a skeleton to the model [15]. In this context, it is worth to say that the Semantic Body Model is also compliant with the requirements of the H-ANIM standard for representing animatable human body in virtual environments [2].

6. Conclusions and future work

In this paper, we have proposed a framework for the automatic segmentation of human body models and their annotation with shape measures, based on a multi-scale geometric and structural analysis. The proposed approach is flexible, produces good results in the context of virtual and real human bodies and supports a variety of anthropometric analysis. The method has however a wider applicability and can be used to characterize any object with tubular features, as demonstrated in an application to smart object characterization [1].

Our research on semantic annotation of body models with anthropometric measures is currently focused

Table 3
Shape parameters of tubular and conical features related to the examples in Figure 14.

| N.	Tube Length	Max. α	Max Sec. Length	Min Sec. Length	Aver. Sect. Length
2	51.13	43.04	47.42	15.51	26.58
3	53.97	42.31	50.26	16.02	28.29
5	66.63	80.21	56.58	23.29	39.67
6	85.88	111.84	56.25	16.08	35.97
7	16.70	-33.07	23.88	23.29	23.58
8	23.66	52.77	81.88	55.69	65.33
N.	Tube Length	Max. α	Max Sec. Length	Min Sec. Length	Aver. Sect. Length
2	40.99	34.99	39.96	17.65	27.62
3	39.45	36.08	44.24	16.63	28.48
4	66.72	63.43	56.55	22.27	39.63
5	62.75	63.85	53.25	23.08	38.02
6	11.09	16.54	17.65	11.37	14.51
7	17.74	53.49	83.56	58.29	69.27
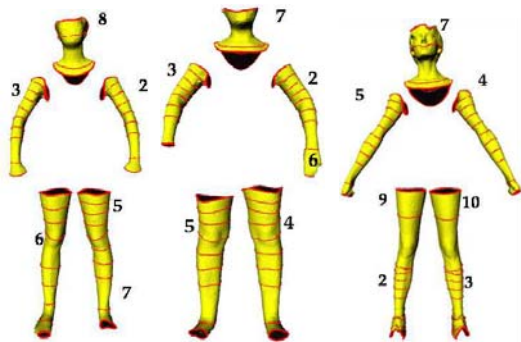
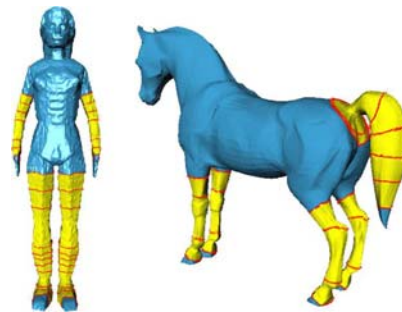Figure 14. Shape segmentation; parameters are reported in Table 3.

Figure 15. Tubular features of a human body model (left) and a horse (right). Note the different length of upper and lower limbs in the two cases; this measure can be used to discriminate between biological species in a database.

on methods to augment and optimize the quality of the descriptors, exploiting the property of our framework of being posture-independent. In the recent anthropometric survey supported by the CAESAR project, the scanning of human body has been conducted using three postures for every individual: the standing posture and two sitting postures. The standing and one of the sitting are mainly aimed at gathering as many data as possible for fully reconstructing the body, while the second sitting position is used to measure data on

how the subjects are really sitting in a comfortable and natural position. These data are very important for ergonomic studies. Using the graph-based representation jointly with a graph-matching technique, it is possible to match parts of the semantic body model of the same individual in the three different postures [3,6]. This would allow in turn to devise geometric editing of the feature shape and descriptors in order

to optimize and augment the measuring capability of the proposed framework.

REFERENCES

1. T. Abaci, M. Mortara, G. Patanè, M. Spagnuolo, F. Vexo, and D. Thalmann. Bridging geometry and semantics for object manipulation and grasping. In *Proc. of SVE2005 Workshop towards Semantic Virtual Environments*, pages 110–119, 2005.

2. C. Babski and D. Thalmann. A seamless shape for hanim compliant bodies. In *VRML '99: Proceedings of the fourth symposium on Virtual reality modeling language*, pages 21–28, New York, NY, USA, 1999. ACM Press.

3. T. K. Dey, J. Giesen, and S. Goswami. Shape segmentation and matching with flow discretization. In *WADS*, pages 25–36, 2003.

4. V. Guillemin and A. Pollack. *Differential Topology*. Englewood Cliffs, NJ, Prentice-Hall, 1974.

5. P. R. M. Jones and M. Rioux. Three-dimensional surface anthropometry: Applications to the human body. *Optics and Lasers in Engineering*, 28(2):89–117, September 1997.

6. S. Marini, M. Spagnuolo, and B. Falcidieno. From exact to approximate maximum common subgraph. In *GbRPR*, pages 263–272, 2005.

7. M. Mortara, G. Patanè, M. Spagnuolo, B. Falcidieno, and J. Rossignac. Blowing bubbles for the multi-scale analysis and decomposition of triangle meshes. *Algorithmica, Special Issues on Shape Algorithms*, 38(2):227–248, 2004.

8. M. Mortara, G. Patanè, M. Spagnuolo, B. Falcidieno, and J. Rossignac. Plumber: A multi-scale decomposition of 3D shapes into tubular primitives and bodies. In *Proc. of Solid Modeling and Applications*, pages 139–158, 2004.

9. J. H. Nurre. Locating landmarks on human body scan data. In *Proceedings., International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 289 – 295, May 1997.

10. J. H. Nurre, J. Connor, E. Lewark, and J. S. Collier. On segmenting the three-dimensional scan data of a human body. *IEEE Transaction on Medical Imaging*, 19(8):787–797, August 2000.

11. G. Patanè, M. Spagnuolo, and B. Falcidieno. Para-graph: graph-based parameterization of triangle meshes with arbitrary genus. *Computer Graphics Forum*, 23(4):783–797, 2004.

12. L. Peng and P.R.M. Jones. Automatic editing and curve-fitting of 3-d surface scan data of the human body. In *International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 296 – 301, May 1997.

13. K.M. Robinette, H. Daanen, and E. Paquet. The caesar project: a 3-d surface anthropometry survey. In *Proceedings of the Second International Conference on 3-D Digital Imaging and Modeling*, 1999.

14. R. Suikerbuik, J.W.H. Tangelder, H.A.M. Daanen, and A. Oudenhuijzen. Feature detection in 3D human body scans. In *SAE Digital human modeling for design and engineering conference*, June 15-17 2004.

15. D. Thalmann, J. Shen, and E. Chauvineau. From fast human body deformations for animation and vr applications. In *Proc. Computer Graphics International 96, IEEE Computer Society Press*, pages 166–174, 1996.

16. C.C.L. Wand, T.K.K. Chang, and M.M.F. Yuen. From laser-scanned data to feature human model: a system based on fuzzy logic concept. *Computer-Aided Desing*, 35:241–253, 2003.

17. Y. Xiao, P. Siebert, and N. Werghi. A discrete Reeb graph approach for the segmentation of human body scans. In *Proceedings of the Fourth International Conference on 3-D Digital Imaging and Modeling,*, pages 378 – 385. IEEE Computer Society Press, October 2003.