Analysis of Human Shape Variation using Volumetric Techniques
Ben Azouz, Zouhour; Rioux, Marc; Shu, Chang; Lepage, R.

National Research Council Canada
Conseil national de recherches Canada

Institute for Information Technology
Institut de technologie de l'information

NRC · CNRC

*Analysis of Human Shape Variation using Volumetric Techniques **

Ben Azouz, Z., Rioux, M., Shu, C., Lepage, R.
July 2004

* published in The 17th Annual Conference on Computer Animation and Social Agents (CASA2004). Geneva, Switzerland. July 7-9, 2004. NRC 47147.

Canada

Analysis of Human Shape Variation Using Volumetric Techniques

Z. Ben Azouz (1,2)
Zouhour.Benazouz@nrc-cnrc.ca

M. Rioux (1)
Marc.Rioux@nrc-cnrc.ca

C. Shu (1)
Chang.Shu@nrc-cnrc.ca

R. Lepage (2)
Richard.Lepage@etsmtl.ca

(1) National Research Council of Canada
(2) École de Technologie Supérieure de Montréal

Abstract

Analyzing human shape variability is important for human system engineering and for animation applications where synthesizing realistic virtual humans is needed to re-create social human activities in historical and cultural context. In this paper, we present a new method for extracting main modes of variations of the human shape from a 3-D anthropometric database. Previous approaches rely on anatomical landmarks. Using a volumetric representation, we show that human shape analysis can be performed despite the lack of such information. We first introduce a technique for repairing the 3-D models from the original scans. Principal components analysis is then applied to the volumetric description of a set of human models to extract dominant components of shape variability for a target population. We demonstrate good reconstructions of the original models from a reduced number of components. We provide a visualization of the main modes of human shape variation.

Keywords: {3-D Anthropometry, volumetric description, human body modeling}

1 Introduction

Characterizing and understanding human shape variation is traditionally the subject of anthropometry – the study of human body measurement. It is essential for better ergonomic design of any products with which people interact such as clothing, automobiles, and work stations. Recently, researchers in the computer animation community have increasingly realized the importance of generating realistic actors by capturing real human shapes [1].

The traditional anthropometry is based on a set of measurements corresponding to linear distances between anatomical landmarks and circumference values at predefined locations. These measurements provide limited information about the human body shape [2]. Furthermore, the traditional measurement, performed by hand, is a long and tedious process, requiring about 30 minutes for each subject. This presents a severe limitation for surveying a large number of subjects. The advances in surface digitization technology led to the appearance of 3-D surface anthropometry where whole body scanners are used to generate detailed human models in a few seconds.

Whole body scanners generate verbose data that cannot be used directly for shape analysis. Therefore, it is necessary to convert 3-D scans to some form of useful representation. The key problem in analyzing human shapes from scanning data is in establishing correspondences between models. In this paper, we obtain a correspondence by converting the polygonal mesh description of human models to a volumetric representation. We align 3-D scans inside a volume of fixed dimensions which is sampled to a set of voxels. A human models is then characterized by an array of signed distances between the voxels and their nearest point on the body surface. Correspondence is achieved by comparing for each voxel the signed distances attributed to different models. Surfaces are reconstructed from the volumetric description using the marching cube algorithm [3]. The advantage of the volumetric representation is that correspondence between different models is achieved without using feature markers.

Principal components analysis is applied to the volumetric representation to extract a reduced number of components representing main modes of variation of the human body. These components represent an orthogonal basis of the shape space and thus human models can be characterized compactly by their projections onto this basis. We visualize the shape variation induced by the main components. This visualization is useful for understanding main modes of the human shape variation.

The reminder of this paper is organized as follows. In the next section we present a brief description of the CAESAR database. Then we review related work in human shape analysis and the related work in model repairing and voxelization. Next, we describe our method to develop a representation that allows a correspondence between different models. Then, we present the extraction of main modes of variation in the human shape space using PCA. Finally, we discuss the results of our approach.

2 CAESAR Data Base

CAESAR(Civilian American and European Surface Anthropometry Resource) is the first 3-D surface anthropometry survey performed in the USA and Europe [4].height=4.86cm During this project, body measurements were taken from about 6000 civilians between the age of 18 and 65 in the USA, the Netherlands and Italy. Subjects were scanned in three postures wearing tight clothes and hair coverings. The Cyberware WB4 [5] and Vitronic [6] full-body scanners were used. Each generated 3-D models contains around 300,000 triangles. A set of 74 white markers were placed at anatomical landmarks prior to scanning. The 3D locations of these landmarks were extracted from the range data using a semi-automatic approach [7].

Our experiments are based on a set of 3-D human models from the CAESAR database.

3 Related Work

In this section, we review related work in the areas of analyzing shape variability from 3-D human scans, converting surfaces to volumetric representations (voxelization), and model repairing techniques.

The major problems in using 3-D human scans for shape analysis is the large number of measured points and the lack of consistent representations between different models. A compact representation of human models can be found in [8] for applications related to the navigation and the visualization of 3-D anthropometric databases. The proposed description is compact but it does not allow the reconstruction of the original shape. Extended hyperquadrics were used in [9] to model human scans. Fitting extended hyperquadrics to 3-D objects is time consuming and a large number of coefficients are required to reconstruct detailed surfaces. Principal components analysis (PCA) has been widely used for shape analysis. For instance, in face recognition, PCA is used to extract a reduced number of component known as *eigenfaces* to represent the space of faces [10]. Applying the PCA to 3-D anthropometric data requires bringing all the models in correspondence to each other. The most popular approach to establish this correspondence is based on fitting template models to measured scans such as in [11] and [12]. Anatomical landmarks are used for guiding the deformation of template surfaces to fit the scanned data. Most of the landmarks are difficult to detect without palpating the body and placing markers on the subjects prior to scanning. Unfortunately, this operation is time consuming and thus will not be done in future anthropometric surveys. In this paper we focus on human shape analysis without relying on anatomical landmarks. We propose to establish a correspondence between different models by converting their surfaces to a volumetric representation and analyze how the same volume is occupied by different models.

Voxelization is concerned with converting geometric objects from their continuous representations into a set of voxels. Early voxelization algorithms were binary, assigning 1 to the occupied voxels and 0 to the unoccupied [13, 14]. Rendered images using binary voxelization suffer from aliasing. Lately, alias-free voxelization techniques were proposed. They can be classified into two categories. The first one concerns filtering techniques where the problem of aliasing is solved by low-pass filtering [14, 15]. The second category concerns distance field techniques that assign to each voxel of the distance to its nearest surface point [16]. This category is adequate for our application since it allows a description of the volume with a continuous function that can be used for a principal components analysis. Here, the cost of computing the signed-distance is a major concern.

Models from the CAESAR database are characterized by incomplete meshes due to occlusions and low grazing angles, as shown in figure 3. In order to compute an accurate distance field, it is necessary to repair these models. Several research treated the problem of hole-filling in 3D models [17, 18]. These methods provide smooth hole filling which is not adequate for the sole of the feet where bulbous shape can be generated. Moreover, these techniques can produce bridges between the two legs and in the areas under the arms.

Since general hole-filling techniques can not deal with the geometry of the human body, specific model repairing techniques have been developed. In [19], a toolbox for identifying holes in 3D human body scans is proposed. The basic idea is to detect and classify holes according to body segments. A potential use of the classification is to fit templates of body parts to missing areas. A library of body parts is then required. Moreover, fitting these parts to missing data is not straightforward. In the same spirit, the approach proposed in [12] solves efficiently the problem of hole filling by fitting a complete template surface to the CAESAR data. This technique is based on anatomical landmarks.

We propose to repair the CAESAR models by estimating missing data from measured information using a slice based method. The human models are sliced horizontally and hole-filling is performed by transforming each slice to a set of closed curves. This step is followed by generating a volumetric representation. The complete surface of a human body is extracted as an isosurface from the volumetric representation.

4 Volumetric Representation

Converting the digitized human surface to a volumetric representation is combined with the hole-filling operation in the method we propose.

4.1 Hole-Filling

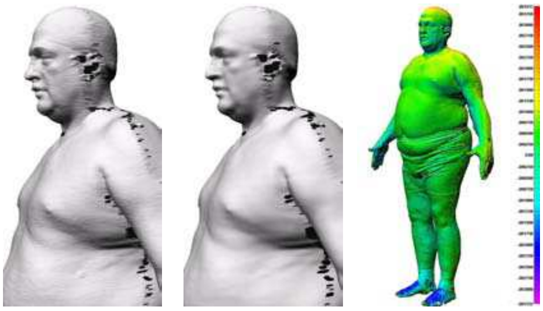
The hole-filling technique we developed is based on slicing horizontally the human models and transforming each slice to a set of closed curves as shown in figure 2.

During hole-filling, the normals to the surface at each vertex around holes boundaries are used. Since the surface is particularly noisy in these areas, the models are smoothed using the Taubin filter [20] before hole-filling. This filter eliminates the noise while minimizing distortion of the original geometry. Surface smoothing of a 3-D human model using the Taubin filter is illustrated in figure 1. The distribution of the distance between the original and the filtered models, computed using the PolyWorks [21] software, proves that there is no significant distortion after surface smoothing.

After filtering, repairing human models consists of the following steps:

3-D scans alignment: During the scanning process, subjects are asked to keep the same posture. Also, the platform of the used whole body scanners have marks where subjects should stand. Despite these precautions, 3-D scans are not sufficiently aligned. This represents a source of noise that affects the shape analysis.

We minimize the misalignment by translating



(a) (b) (c)

Figure 1: Surface smoothing of a 3-D human scan using Taubin filter. (a) Original model. (b) Filtered model. (c) Distribution of the displacement between the original and the filtered model: the mean value is 0.4 mm and the standard deviation is 0.52 mm.

the human models in order that their centers of gravity are vertically aligned. The 3-D scans are then rotated so that the coordinate system correspond to the principal axes of their tensor of inertia. We define X, Y, and Z axes as the direction of the thickness, width and height of the human body, respectively.

Segmentation: We distinguish hands, the left foot and the right foot from the rest of the body. This segmentation is useful for two reasons. First, the CAESAR models present a severe lack of data in the hands. Therefore, scanning hands separately with an adequate resolution is required for shape analysis. Thus, for our experiments we eliminate hands. Second, for the hole filling, slices are processed differently according to the body segment they belong to. In our experiments human models in a standing posture are used. We define boxes of equal dimensions for all the models to border hands and feet. To locate the boxes for the hands, we identify tips of the middle fingers in both hands. These features correspond to extremities in the y direction as shown in figure 3(b). The boxes corresponding to both feet are located using the point of coordinates (0, 0, minimum(z)).

Converting polygonal surface to horizontal slices: The models are sliced horizontally by intersecting planes (parallel to the XY plane) with the edges of the surface triangles. Each slice is a set of intersection points where points belonging to the same triangle are connected with a segment. Hole boundaries are identified as points that are connected to at most one other point. Hole-filling is achieved by linking these points.

Connecting boundary points: After identifying the points to be connected, we establish criteria to find the optimal way to connect them. Two points are most likely to be connected if they are close to each other and their corresponding normals to the surface have similar orientations. These criteria are expressed in the form of a cost function that has to be minimized in order to find the optimal connection. The cost of connecting two points p_1 and p_2 from the feet is

$$cost(p_1, p_2) = \|p_1 - p_2\| + (1 - \cos(\vec{n}_1, \vec{n}_2))$$

and the cost of connecting two points from the rest of the body is defined as

$$cost(p_1, p_2) = \|p_1 - p_2\| * (1 - \cos(\vec{n}_1, \vec{n}_2)).$$

The second function gives more importance to the normal orientation in order to minimizes bridges in areas such as the under-arms and between the legs.

Additional conditions are used to differentiate between valid and non-valid connections. A maximum distance is set between two connected points. The value of the maximum distance depends on the body segments. Also, in a valid configuration, a connection should not intersect other segments from the slice unless they have a common vertex. After determining the optimal connections, points are linked with second-order Bezier curves to preserve the curvature.

4.2 Voxelization

Voxelization based on a signed distance map consists of two steps. The first step classifies each voxel as lying in the interior or the exterior of the body. The second step computes the distance from each voxel to its nearest surface point.

For manifold polygonal models that are watertight, a voxel can be classified by counting

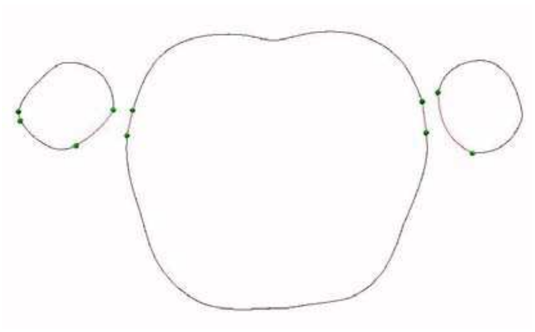


Figure 2: Hole-filling of a slice. Black curves correspond to original data and red curves correspond to estimated data.

the number of times that a ray with its origin at the center of the voxel intersects polygons of the model (parity count). An odd number of intersections corresponds to an interior voxel and an even number corresponds to an exterior voxel. In our implementation, voxelization is performed after slicing the surface. Therefore rays are intersected with slices rather than with the model surface.

For some slices where the data is extremely noisy, the criteria used in our hole-filling approach are not sufficient to connect appropriately boundary points. Thus, voxels can be misclassified. To overcome this limitation, we use different directions for scan-converting models as proposed in [22]. Each direction votes on the classification of a voxel and the majority vote is the voxel's final classification.

Computing the accurate distance from the center of each voxel to its nearest surface point is time consuming. We used the following methods to accelerate the process:

Reducing the number of voxels for which the distance has to be computed: A fast approximation of the distance field $\widetilde{dist}$ is first determined using the Danielson algorithm [23]. Accurate distance is then computed only for voxels that are within a certain distance from the body surface. The Danielson algorithm starts from a binary volumetric representation where voxels are classified either as feature or non-feature voxels. The algorithm computes the distance of each non-feature voxel to the nearest feature voxel by propagating the distances in a small neighborhood. In order to compute distances of both interior and exterior voxels, the Danielson algorithm is applied twice to generate an estimate distance for each voxel. During the first pass, feature voxels correspond to interior and surface voxels while in the second pass feature voxels correspond to exterior and surface voxels.

Computing distance to the nearest segment of curve rather than the nearest polygon of a surface: Before voxelization, polygonal models are converted to slices. Each slice is a set of connected segments. This fact is useful since computing the distance to a segment is faster than computing the distance to a triangle.

Reducing the number of segments which the distance has to be computed with: For each voxel V we start by identifying the distance ds to the closest segments from the same slice. Only segments that are inside a square around V and having a diagonal equal to 2 $\widetilde{dist}$ are considered. After identifying ds, segments from other slices that are inside a cube around V and having a diagonal equal to $2ds$ are considered.

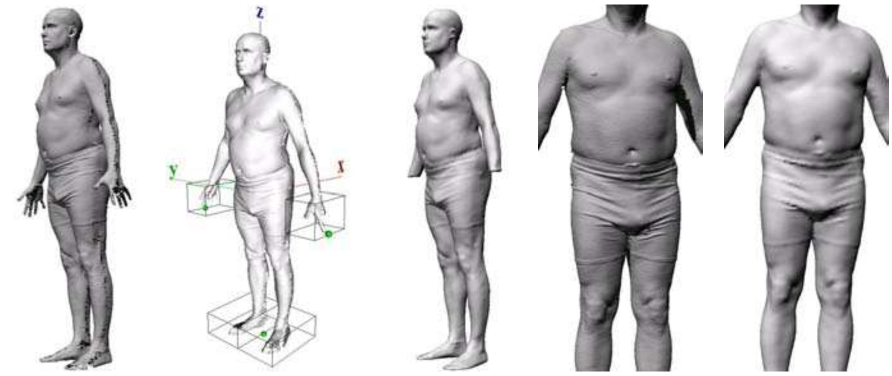
Positive distances are assigned to interior voxels and negative distances are assigned to exterior voxels. A complete surface model is then extracted from the volumetric description using the marching cube algorithm [3].

Figure 3 illustrates a human model after hole-filling. The surface is extracted from a volumetric representation with a sampling rate of 8mm. The result shows that most of the holes in the original models are properly filled except for areas such as the ears where most of the information is missing.

5 Principal Component Analysis

Principal component analysis (PCA) is used to extract a reduced number of components that span a high variability percentage in the human shape space. These components correspond to the main modes of shape variation within a given population.

In order to apply the PCA, human models are converted to a vector form $\vec{\Psi}$ where each element represent the signed distance of a voxel to the model's surface. The average over a set of



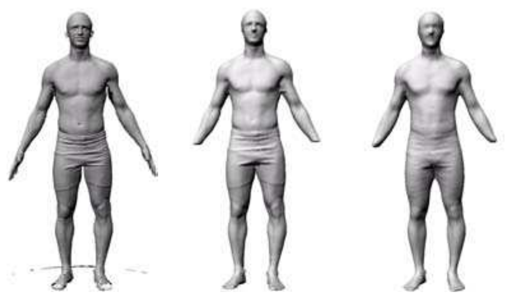
(a) (b) (c) (d) (e)

Figure 3: Hole-filling of the CAESAR models. (a) Original model. (b) Segmentation of the hands and the feet of a human model. (c) Human model after hole filling and eliminating the hands. (d) Details of holes in the original model. (e) Details of repairing some of the difficult areas such as the underarm and between the legs.

N models is given by: $\bar{\Psi} = \frac{1}{N}\sum_{i=1}^{N}\vec{\Psi}_i$. Deviation vectors $\vec{\Phi}_i = \vec{\Psi}_i - \bar{\Psi}$ are arranged in a matrix $A = [\vec{\Phi}_1\vec{\Phi}_2....\vec{\Phi}_N]$. The PCA of the matrix A generates a set of non-correlated eigenvectors $\vec{u}_i$ and their corresponding variances λ_i. The eignevectors are sorted in decreasing order of their variances. Any vector $\vec{\Phi}_i$ can be approximated as follows: $\vec{\Phi}_i \approx \sum_{j=1}^{M} c_{ij}\vec{u}_j$ where $0 \leq M \leq N$ and $c_{ij} = \vec{\Phi}_i \cdot \vec{u}_j$ are the coefficients of similarity of the model with each eigenvector. The quality of reconstruction can be evaluated by the fraction $\sum_{i=1}^{M}\lambda_i / \sum_{i=1}^{N}\lambda_i$ representing the percentage of variance spanned by the set of eigenvectors used in the reconstruction.

We applied the PCA on a set of 300 scans of males in a standing posture. The result shows that the first 40 eigenvectors span 92% of the variability of the studied data set. Figure 4 illustrates that the shape of a human body is globally reconstructed from only 40 main components. Thus the coefficients (c_{ij}) of similarities of 3-D scans with a set of main components is a compact representation that contains most of the information about the original shape. This representation can be considered as a generalisation of the concept proposed by Sheldon [24] where three components, endomorphy(soft roundness in the body), mesomorphy (hardness and muscularity) and ectomorphy(linearity and skinnines) are proposed to characterize human body.

Visualizing the shape variation induced by the main components associates an intuitive meaning of these components which provides a better understanding of shape variability within a population. The models illustrated in Figure 5 are synthesized using the projection of a captured model into the basis of the extracted eigenvectors. For each row, models are generated by varying only the coefficient of similarity of the original model with the corresponding component. The main modes of variation are visualized in decreasing order of the percentage of variability they reperesent in the population under study.



(a) (b) (c)

Figure 4: Reconstruction of a human model from a reduced number of components. (a) Original model. (b) Model obtained after hole filling. (c) Approximation of the original model using 40 main components.

The first component is obviously correlated to the height and the second one is highly correlated to the weight. This result is consistent with traditional anthropometry knowledge where the height and weight are considered main measurements to differentiate between human bodies. Some of the components, such as the third one, reflect posture variation. This observation could be interpreted as a random posture variation among subjects being digitized. But, it could also be interpreted as the natural neutral posture of the subject which reveals state of mind and character. The fourth component reflects a variability of muscularity which is consistent with the qualitative observation made by Sheldon using photographs. The extracted component reveals though, all the subtle changes related to this variability such as the arcing of the arms and the legs.

The visual interpretation of the main components is useful for generating virtual humans from real ones. For instance, increasing the coefficient of similarity with the fourth component will generate a virtual model that is more muscular than the original model.

The synthesized models have for the most part a realistic appearance. Artifacts that appear mainly in the head and the arms are related to two different reasons. The first reason is the use of signed distance. In fact a linear combination of surfaces characterized by a distance field description is only an approximation of the actual linear combination of these surfaces. The approximation is more accurate if there is alignment between different models. Artifacts are also due to the fact that some of the generated models lie outside the space of human shape spanned by the set of human scans used to extract the main components.

In order to isolate shape variation from height variation, we re-applied the PCA on the same set of models after normalizing their heights. Using the signed distance after normalization provides a better correspondence between different models. Figure 6 visualizes the main modes of shape variation extracted after height normalization. The first component is highly correlated to the weight. We can see that the third component, related to the variation of mass distribution and muscularity, is consistent with the fourth component generated using human models without height normalization.

6 Conclusion

In this paper, we propose the use of volumetric representation of 3-D human models based on signed-distance for shape analysis. This representation provides a correspondence between different models without using landmarks. Principal component analysis is applied to the volumetric representation of a set of 3-D scans from the CAESAR database. The PCA generates a reduced number of components that form an orthonormal basis of the human shape space. Characterizing human models by their coefficients of similarity to these components provides a compact representation that encloses most of the shape information. This description can be used to generate virtual humans that are for the most part realistic except for the head and the hands where additional processing is required. The compact description generated by the PCA is useful for anthropometric applications to extract representative cases from a given population.

References

[1] N. Magnenat-Thalmann and H. Seo. Automatic Modeling of Animatable Virtual Humans- A Survey. *Fourth International Conference on 3-D Digital Imaging and Modeling, IEEE Computer Society*, pages 2–10, October 2003.

[2] K.M. Robinette, M.W. Vannier, M. Rioux, and P.R.M. Jones. 3-D surface anthropometry: Review of technologies = L'anthropométrie de surface en troi dimensions: examen des technologies. In *Neuilly-sur-Seine: North Atlantic Treaty Organization Advisory Group for Aerospace Rearch & Development, Aerospace Medical Panel*, 1997.

[3] W.E. Lorenson and H.E. Cline. Marching Cubes: A High Resolution 3-D Surface Construction Algorithm. *Computer Graphics*, 21(3), July 1987.

[4] K.M. Robinette, H. Daanen, and E.Paquet. The Caesar Project: A 3-D Surface Anthropometry Survey. *Second International*

Conference on 3-D Digital Imaging and Modeling, pages 380–386, 1999.

[5] Cyberware Inc. `http://www.cyberware.com`.

[6] Vitronic Inc. `http://www.vitronic.com`.

[7] D. Burnsides, M. Boehmer, and K.M. Robinette. 3-D Landmark Detection and Identification in the Caesar Project. *Third International Conference on 3-D Digital Imaging and Modeling*, pages 393–398, 2001.

[8] E. Paquet, K.M. Robinette, and M. Rioux. Management of Three-dimentional and Anthropometric Databases: Alexandria and Cleopatra. *Journal of Electronic Imaging*, 9(4), 2000.

[9] Z. Ben Azouz, M. Rioux, R. Lepage, and C. Shu. Compact Description of Human Body Shape Using Extended Hyperquadrics. *International Ergonomics Association XVth Triennial Congress*, August 2003.

[10] M. Kirby and L. Sirovich. Application of the Karhunen-Loeve Procedure for the Characterisation of Human Faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):103–108, 1990.

[11] H. Seo and N. Magnenat-Thalmann. An Automatic Modeling of Human Bodies from Sizing Parameters. *SIGGRAPH Symposium on Interactive 3D Graphics*, pages 19–26, July 2003.

[12] B. Allen, B. Curless, and Z. popović. The Space of Human Body Shapes: Reconstruction and Parametrisation from Range Scans. *ACM SIGGRAPH'2003*, July 2003.

[13] A. Kaufman. An Algorithm for 3D Scan Conversion of Polygons. *EUROGRAPHICS'87*, pages 197–208, August 2000.

[14] A. Kaufman. *Volume Visualization*. IEEE Computer Society Press, 2000.

[15] S.W. Wang and A. Kaufman. Volume Sampled Voxelization of Geometric Primitives. *Visualisation'93*, pages 78–84, October 1993.

[16] M.W. Jones. The Production of Volume Data from Triangular Meshes Using Voxelization. *Computer Graphics Forum*, 15(5):311–318, December 1996.

[17] J. Davis, S.R. Marschner, M. Garr, and M. Levoy. Filling Holes in Complex Surfaces Using Volumetric Diffusion. *Proc. First International Symposium on 3D Data Processing, Visualization, Transmission*, 2002.

[18] J.C. Carr, R.K. Beatson, J.B. Cherrie, T.J. Mitchell, W.R. Fright, B.C. McCallum, and T.R. Evans. Reconstruction and Representation of 3D Objects with Radial Basis Functions. *Computer Graphics Proceedings, Annual Conference Series(SIGGRAPH'2001*, pages 67–76, August 2001.

[19] M.V. Stralen, H.A.M. Daanen, and J.W.H. Tangelder. A Tool Box to Identify Holes in 3D Human Body Scans. *International Ergonomics Association XVth Triennial Congress*, August 2003.

[20] G. Taubin. Geometric Signal Processing on Polygonal Meshes. *EUROGRAPHICS'2000*, 2000.

[21] InnovMetric Inc. `http://www.innovmetric.com`.

[22] F.S. Nooruddin and G. Turk. The Production of Volume Data from Triangular Meshes Using Voxelization. *IEEE Transactions on Visualization and Computer Graphics*, 9(2):191–205, June 2003.

[23] P-ES. Danielsson. Euclidean Distance Mapping. *Workshop on Volume Visualization*, 14:227–248, July 1980.

[24] W.H. Sheldon, S.S. Stevens, and W.B. Tucker. *The Varieties of Human Physique*. Harper and Brothers Publishers, New York, 1940.

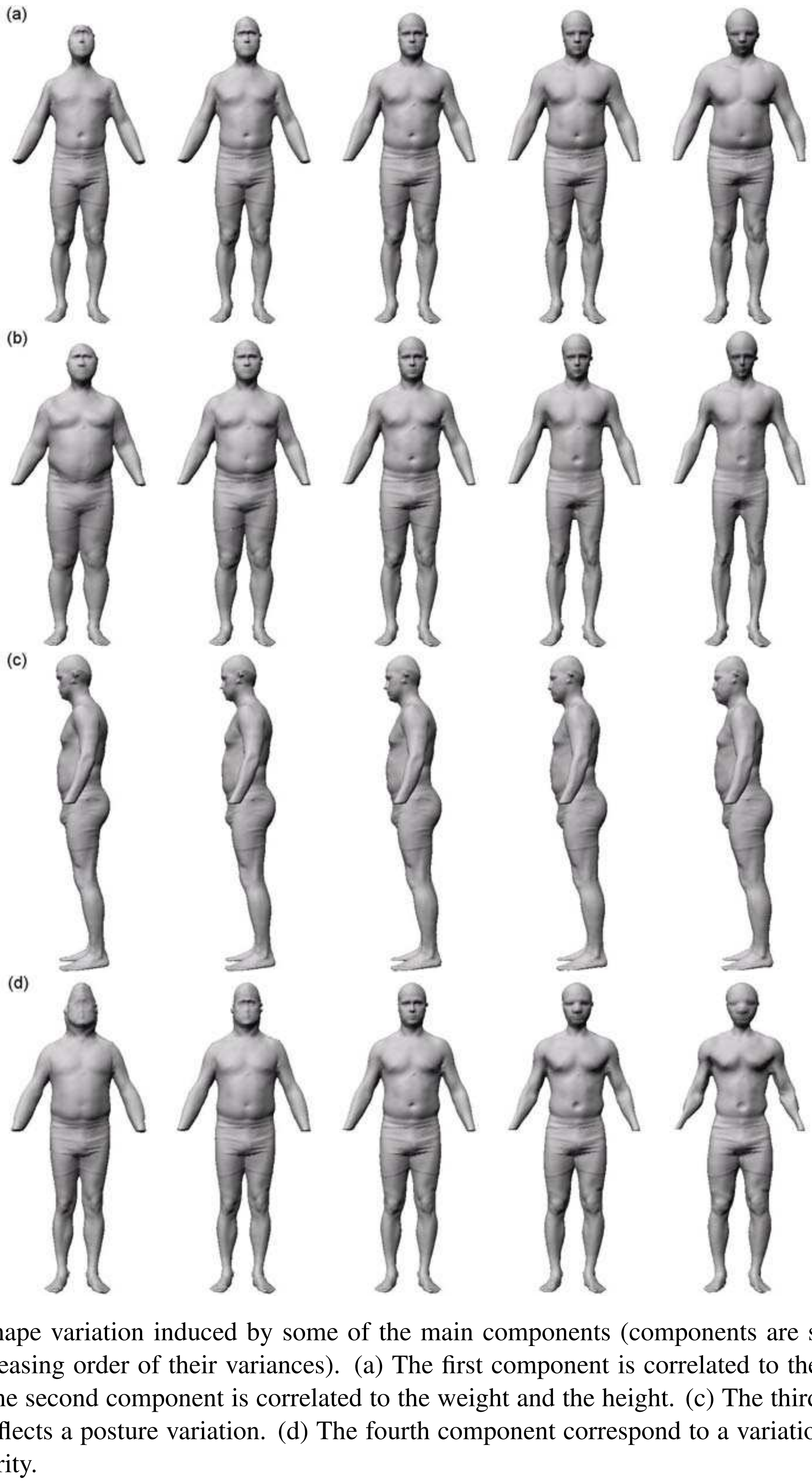


Figure 5: Shape variation induced by some of the main components (components are sorted in decreasing order of their variances). (a) The first component is correlated to the height. (b) The second component is correlated to the weight and the height. (c) The third component reflects a posture variation. (d) The fourth component correspond to a variation of muscularity.

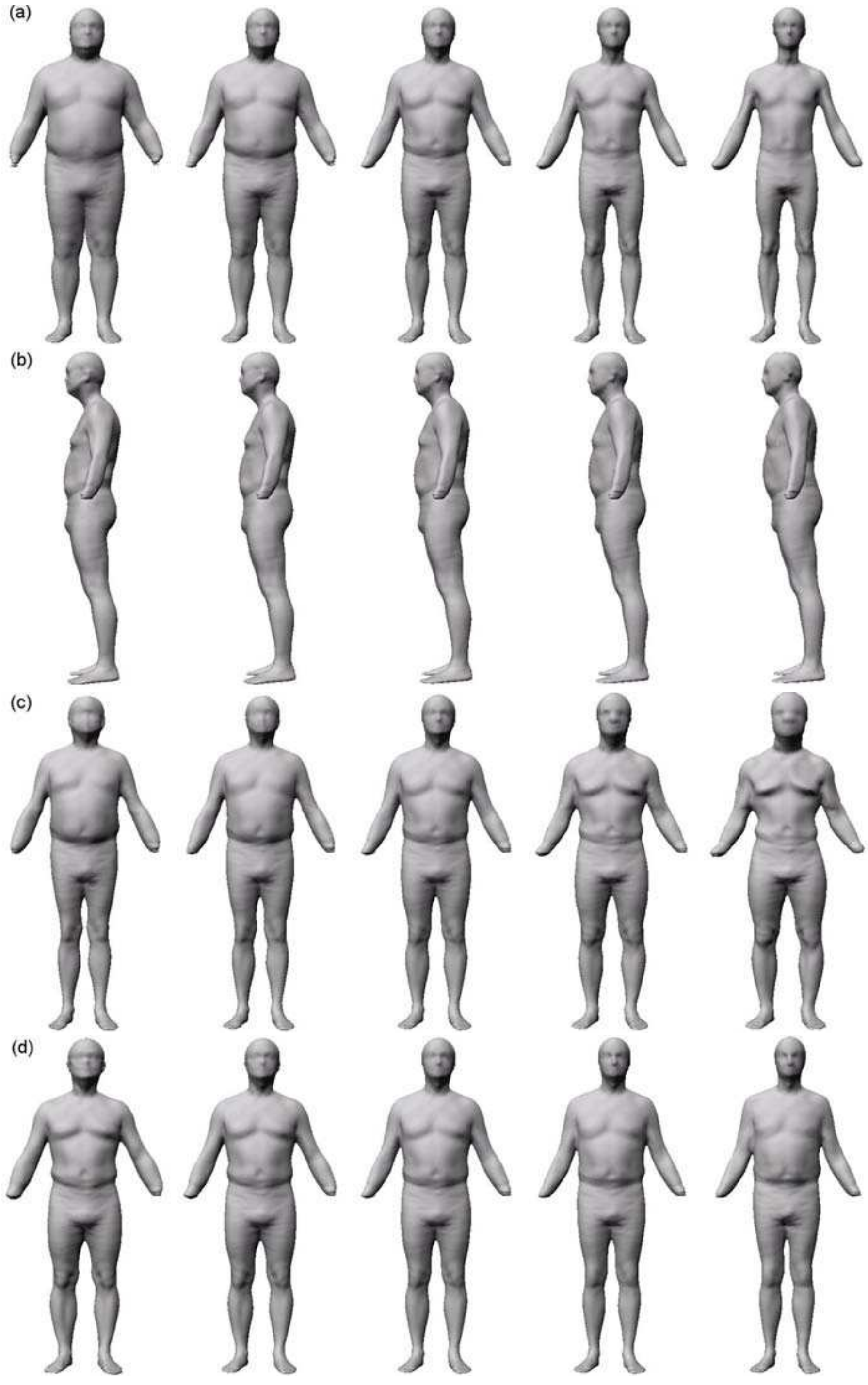


Figure 6: Shape variation induced by some of the main components after height normalization (components are sorted in decreasing order of their variances). (a) The first component is correlated to the weight. (b) The second component reflects a posture variation. (c) The third component correspond to a variation of muscularity. (d) The fourth component correspond to a variation of mass distribution between the torso and the legs and a variation in the arms position.