

Modeling the Human Body in 3D: Data Registration and Human Shape Representation

by

Aggeliki Tsoli

B. Sc., University of Thessaly, 2005

Sc. M., Brown University, 2008

A Dissertation submitted in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy
in the Computer Science at Brown University

Providence, Rhode Island

May 2014

© Copyright 2014 by Aggeliki Tsoli

This dissertation by Aggeliki Tsoli is accepted in its present form
by the Computer Science as satisfying the
dissertation requirement for the degree of Doctor of Philosophy.

Date _____

Michael J. Black, Director
Max Planck Institute for Intelligent Systems, Germany

Recommended to the Graduate Council

Date _____

Brian D. Corner, Reader
US Army Natick Soldier RDEC, MA

Date _____

James Hays, Reader
Department of Computer Science, Brown University

Approved by the Graduate Council

Date _____

Peter M. Weber, Dean of the Graduate School

Published, Submitted and Draft Papers

- [1] Aggeliki Tsoli, Naureen Mahmood, and Michael J. Black. “Breathing Life into Shape: Capturing, Modeling and Animating 3D Human Breathing”, ACM SIGGRAPH, 2014.
- [2] Aggeliki Tsoli, Matthew Loper, and Michael J. Black. “Model-based Anthropometry: Predicting Measurements from 3D Human Scans in Multiple Poses”. IEEE Winter Conference on Applications of Computer Vision (WACV), 2014.
- [3] David A. Hirshberg, Matthew Loper, Eric Rachlin, Aggeliki Tsoli, Alexander Weiss, Brian Corner, and Michael J. Black. “Evaluating the Automated Alignment of 3D Human Body Scans”. International Conference on 3D Body Scanning Technologies (3DBST), 2011.
- [4] Aggeliki Tsoli, and Michael J. Black. “Shape- and Pose-invariant Correspondences using Probabilistic Geodesic Surface Embedding”. Annual Symposium of the German Association for Pattern Recognition (DAGM), 2011.
- [5] Aggeliki Tsoli, and Michael J. Black. “Constraint-based Human Body Shape Analysis”. Draft.

Acknowledgements

In the long trip of the PhD, I have been fortunate to have interacted with very creative and intellectually stimulating people. First and foremost, I would like to thank my advisor, Michael Black. Michael will always be a role model for me for his approach to research and fair play. He is a deep thinker and intellectually curious in almost every science problem. At the same time, he is open to other opinions. His attention to detail has definitely refined my way of approaching problems and presenting my work. His most admirable attribute is that he teaches by example. I have been very positively surprised to see him staying up late with his students before paper submission deadlines. Along with Michael, I would like to thank Lee, Michael's wife, who has made the "Friday beer" evenings at MPI less nerdy and for being a loyal participant in data collection trials.

I would like to thank my committee members, Brian Corner and James Hayes, for useful comments on shaping and refining the presented work.

This work would not have been the same without the support of the computer vision lab at Brown and MPI. I am especially grateful to Matt Loper, Javier Romero, Eric Rachlin, David Hirshberg, Alex Weiss, Naureen Mahmood, Peter Gehler for maintaining a very collaborative environment in the lab. Thanks for all the research discussions, code, and the support during deadlines. I would also like to thank Emma-Jayne Holderness for her patience and dedication during the long 3D scanning sessions and all the friends that got scanned in order to provide data for my work. The assistance of the technical staff and secretaries at Brown and MPI with dealing with non-research related issues has been invaluable.

Sharing my PhD years between USA and Germany, I got to enjoy the company and receive the generous support of many friends. I would like to thank Stavroula for being a "partner in crime", Athina, whose long-distance friendship has definitely passed the test of time, my international family at Brown, Jean, Aparna, Wenjin, Misha, Radu, Doria, and the Greek communities at Brown and MPI. I would also like to thank the students at the Brown University club tennis team and the MPI students in the "Friday beer" for making my life more balanced and fun.

Last but not least, I would like to thank my family for their endless support at any time and at any place. My parents, Christos, Charikleia, have always been there for me. I thank my brothers, Elias, Kostas, Dimitris, Tasos, for playing a big role in making my childhood happy and for being friends I can depend on up to now.

This work was supported in part by the Office of Naval Research, Contract W911QY-10-C-0172.

Dedicated to my family

Contents

Published, Submitted and Draft Papers	iv
Acknowledgements	v
Dedication	vi
List of Tables	x
List of Figures	xi
Chapter 1. Introduction	1
1. Problem and Thesis Statement	1
2. Motivation	2
3. Modeling the Human Body in 3D	4
3.1. 3D Articulated Deformable Object Representation	5
3.2. Shape-, Pose- and Resolution- Invariant Correspondences between 3D Non-rigid Articulated Objects	6
3.3. Measuring the Human Body in 3D: Correlation of the 3D Human Body Shape with Anthropometric Measurements	8
3.4. Data-driven Modeling and Animation of Breathing	9
4. Challenges	9
4.1. Correspondences between 3D non-rigid articulated objects varying in pose, shape, and resolution.	9
4.2. Anthropometric measurement prediction from 3D human bodies.	11
4.3. Modeling human shape and pose change due to breathing.	12
5. Contributions of the Thesis	12
Chapter 2. Shape- and Pose-Invariant 3D Correspondences	14
1. Related Work	15
2. Probabilistic Geodesic Surface Embedding	16
2.1. Probabilistic Model	16

2.2. Inference	19
2.3. Sampling	19
2.4. Geodesic Distance calculation	19
3. Results	21
3.1. Data	21
3.2. Evaluation	21
4. Conclusions	26
Chapter 3. Constraint-based Human Body Shape Analysis	27
1. Previous work	28
2. Methodology	30
3. Experiments	31
3.1. Compact shape representation for anthropometric measurement prediction	31
3.2. Qualitative analysis of the nullspace	34
4. Conclusions	36
Chapter 4. Anthropometric Measurement Prediction from 3D Human Scans in Multiple Poses	37
1. Related Work	39
2. Measurements from 3D scans	40
2.1. Registration	41
2.2. Feature Extraction	42
2.3. Feature Type Selection per Measurement	44
2.4. Measurement Prediction	44
3. Results	47
3.1. Method Evaluation	47
3.2. Prediction from Multiple Poses	49
3.3. Features for Measurement Prediction	49
4. Conclusions	51
Chapter 5. Breathing Life into Shape: Capturing, Modeling and Animating 3D Human Breathing	53
1. Related Work	55
2. Breath Taking (Data Capture)	57
2.1. Data processing	58
3. Breathing Space (Shape Model)	59
3.1. Adding Breathing	61
3.2. Extracting the Breathing Deformations and Pose	61

4. Statistics of Breathing	63
4.1. Breathing Shape Statistics	64
4.2. Breathing Pose Statistics	65
5. Breathing Model	67
5.1. Shape change during breathing	67
5.2. Breathing types	67
5.3. Breathing model	68
5.4. Pose model	71
5.5. Influence of shape and pose on animation realism	71
6. Breathing Animation	72
6.1. Trajectory editing	72
6.2. Breath acting	74
7. Results	74
7.1. Breathing in Action: Poses and Motion	74
8. The Last Breath (Conclusions)	75
Chapter 6. Conclusions and Future Work	83
1. Shape Matching between 3D meshes	83
2. 1D Measurements and Attributes from 3D Scans	84
3. Data-driven animation of subtle body deformations	86
4. The Challenge: Data-driven whole human body modeling in 3D	87
Appendix A. Detailed Performance of Measurement Prediction from 3D Scans	88
1. Measurement prediction from the standing pose	92
2. Measurement prediction from the seated pose	96
Bibliography	100

List of Tables

1	Effect of pose on measurement prediction.	49
2	Effect of our global and local features on measurement prediction.	51

List of Figures

1.1 Graphics techniques for generating moving human bodies.	3
1.2 Virtual try-on scenario.	4
1.3 A 3D scan as a snapshot of the human body shape in time.	5
1.4 Triangular 3D meshes that differ in mesh topology.	6
1.5 Human body pose.	6
1.6 Human body shape.	7
1.7 Pairwise correspondences between 3D meshes.	7
1.8 Model-based extraction of anthropometric measurements in 3D.	8
1.9 Correlating human body shape variation with measurements.	9
1.10 Illustration of various breathing types.	10
2.1 Example of local optimum in optimization for finding 3D correspondences.	15
2.2 Conditional Random Field model for finding correspondences between two triangular meshes.	17
2.3 PGSE pairwise geodesic potential.	18
2.4 PGSE geodesic signature potential.	18
2.5 PGSE inference procedure.	20
2.6 Voronoi-based vs. geodesics-based error metric.	22
2.7 Comparative visual 3D correspondences between PGSE and GMDS.	23
2.8 Comparative Voronoi error between PGSE and GMDS.	24
2.9 3D correspondences between meshes with different resolution.	25
3.1 Factoring of human body shape variation into a measurement space that is constrained to predict anthropometric measurements and a null space of body shape variation that are unrelated to measurements.	28
3.2 Root Mean Squared Error (RMSE) on measurement prediction for PCA, Hasler et al. [44] and our method for $k = 3$ basis vectors.	32

3.3	Root Mean Squared Error (RMSE) on measurement prediction for PCA, Hasler et al. [44] and our method for $k = 9$, and $k = 9$ basis vectors.	33
3.4	Average Root Mean Squared Error for different MPCA formulations relative to previous work.	34
3.5	Human body shape variation captured by methods that take into account measurements relative to methods that are ignorant of measurements.	35
3.6	The nullspace of body shape.	36
4.1	Registered meshes of a person in two poses after fitting a human body model.	38
4.2	Pipeline for measurement prediction from 3D scans.	41
4.3	Model-based alignment for effective hole filling.	42
4.4	Low-dimensional PCA representation of human shape.	43
4.5	Local features for measurement prediction.	45
4.6	Examples of feature selection for measurement prediction.	46
4.7	Example anthropometric measurements from the CAESAR dataset.	48
4.8	Comparative evaluation of measurement prediction on CAESAR.	50
5.1	Full inhale and full exhale for various breathing types.	54
5.2	Example 3D scans for modeling breathing.	57
5.3	Example of 3D scan registration.	59
5.4	SCAPE+Breathing model.	60
5.5	Pose change during breathing.	62
5.6	Shape change due to breathing.	63
5.7	Principal components of shape change due to breathing.	64
5.8	Correlation between coarse human body shape and shape change due to breathing.	65
5.9	Principal components of pose change due to breathing.	66
5.10	Volume change vs. shape change for a single subject.	67
5.11	Linear models of shape change during breathing for various breathing types.	68
5.12	Shape change during breathing as a function of breathing type and volume.	69
5.13	Examples of “abdominal” and “chest” breathing.	70
5.14	Pose model per gender.	71
5.15	Visualization of various breathing types using a Moiré-like effect.	73

5.16	“Breath acting”.	73
5.17	Spirometer-driven animation.	75
5.18	Recorded vs. synthesized breathing animations.	76
5.19	Breathing in standing vs. seated pose.	76
5.20	Simulated breathing while running.	77
5.21	Simulation of different breathing types (Part I).	79
5.22	Simulation of different breathing types (Part II).	80
5.23	Front and back views of simulated breathing (Part I).	81
5.24	Front and back views of simulated breathing (Part II).	82
A.1	Measurement prediction on females in CAESAR (part I).	88
A.2	Measurement prediction on females in CAESAR (part II).	89
A.3	Measurement prediction on males in CAESAR (part I).	90
A.4	Measurement prediction on males in CAESAR (part II).	91
A.5	Measurement prediction on females in CAESAR using only data in the standing pose (part I).	92
A.6	Measurement prediction on females in CAESAR using only data in the standing pose (part II).	93
A.7	Measurement prediction on males in CAESAR using only data in the standing pose (part I).	94
A.8	Measurement prediction on males in CAESAR using only data in the standing pose (part II).	95
A.9	Measurement prediction on females in CAESAR using only data in the seated pose (part I).	96
A.10	Measurement prediction on females in CAESAR using only data in the seated pose (part II).	97
A.11	Measurement prediction on males in CAESAR using only data in the seated pose (part I).	98
A.12	Measurement prediction on males in CAESAR using only data in the seated pose (part II).	99

Abstract

Realistic modeling of the human body in 3D has many applications ranging from fashion to the production of movies and video games. However, leveraging data coming from state-of-the-art 3D acquisition systems poses a set of problems. A 3D scan of a person contains holes and thousands of unordered points. In addition, a 3D scan is a single snapshot of the human body in time, while the shape of the human body changes with motion, breathing, aging, *etc.* We pose that modeling the 3D human shape in a data-driven, surface-based way using multiple instances of captured geometry can lead to (a) accurate estimation of human attributes, e.g. anthropometric measurements and (b) realistic representation of human body deformations.

In this thesis, we tackle parts of the whole process of modeling the human body in 3D. First, we describe a method for finding sparse pairwise correspondences between 3D triangular meshes of articulated objects, such as humans, in various shapes and poses. We propose a search framework that effectively explores the space of possible correspondences and is more robust to local optima than previous work. Central to this framework are features based on mesh surface paths that are invariant to shape, pose, and resolution.

Second, we present an approach for extracting standard anthropometric measurements from 3D human scans. Our approach relies on fitting a deformable 3D body model to scan data; this model-based fitting is robust to scan noise. In addition, it brings a query scan into registration with a database of registered body scans, which facilitates statistical analysis. We show that combining multiple poses yields optimal measurement prediction. For the first time, we perform an extensive evaluation of existing commercial and research systems using the CAESAR dataset.

Third, we propose a flexible optimization framework to describe variation in 3D human body shape using orthogonal shape subspaces related to different linear or non-linear constraints. In particular, we find directions of body shape variation that are directly related to human measurements. We also define a “null space” that captures how body shape varies in ways that are orthogonal to the measurements. The null space reveals interesting properties of body shape that are not captured by standard anthropometric measurements such as inter- and intra-person posture variation, “bow legs”, and body shape asymmetries.

Fourth, we propose a method for capturing and modeling the non-rigid intrinsic shape variation of the human body during breathing. We learn a detailed model of body shape deformations due to breathing for different breathing types and provide simple animation controls to render lifelike breathing regardless of body shape. We also develop a novel interface for breathing animation using a *spirometer*, which measures the breathing volume of a “breath actor.” Our approach generates fine-scale body shape deformations due to breathing with greater ease and realism than previously achieved.

In this thesis, we introduce application-dependent 3D shape representations that describe realistically the details of the human shape, capture variation in shape and are tied to the anthropometry of the human body.

Introduction

1. Problem and Thesis Statement

In this thesis, we address the problem of modeling the human body in three dimensions (3D). We focus mainly on modeling the shape of the human body. Sometimes that refers to the observed shape of the human, other times we consider the intrinsic human shape after factoring out any deformations that depend on the human pose. We find that the representation of 3D human shape is application-dependent. For instance, for predicting tailoring measurements from 3D scans, local features such as circumferences on the human body across various poses are quite informative. On the other hand, modeling body deformations due to breathing requires features that account for shape change of the same individual. Thus, we take into account multiple scans of a person in a single pose and focus mainly on the torso area. As a common theme, we aspire to model the human body in 3D in such a way that the proposed shape representation

- describes realistically the details of the human body shape
- is able to capture variation in human body shape
- is accurate in the sense that it is tied to the anthropometry of the human body.

We pose that *data-driven, surface-based modeling of 3D human shape using multiple high-resolution 3D scans can lead to (a) accurate estimation of human attributes, e.g. anthropometric measurements and (b) realistic representation of human body deformations.*

We take four steps towards modeling accurately the human body in 3D:

- (1) We propose a method for sparse pointwise correspondences between 3D meshes. Methods for finding sparse correspondences can be used to automate 3D scan registration.
- (2) We describe a framework for measuring the human body in 3D. We extract anthropometric measurements using a new method of model-based anthropometry that leverages multiple poses and shape statistics.
- (3) We associate semantics with shape variation. Here we go beyond previous work to explore the idea of the “null space” of body shape variations not captured by standard anthropometric measurements.
- (4) We present a method to capture and model how body shape varies during breathing. Previously, body shape due to breathing has not been modeled in detail.

2. Motivation

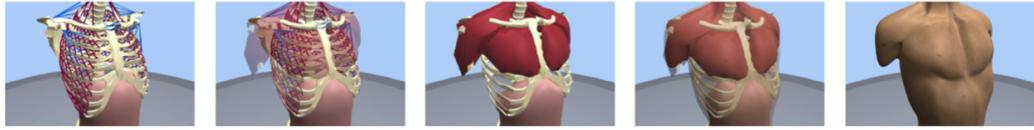
Studying humans has always been a big focus in the fields of Computer Vision and Graphics. Earlier attempts focused on detecting people in images or videos [28, 95, 100]. Subsequent research efforts ranged from estimating the pose and shape [11, 40, 64] of people to modeling hair [38, 60] and clothing [36, 37]. Arguably, modeling the human body is essential for making applications about and for humans. However, the human body shape offers unique research challenges compared to other 3D objects in our world. Although the human body consists of many parts that move in a rigid way, there are also many ways in which the human shape deforms non-rigidly; e.g. different muscles are flexed when a person is standing compared to when the same person is sitting, deformations due to breathing lead to a constantly changing body shape even when the person remains seated, *etc.* Additionally, the human body shape is a rich source of information about human attributes, ranging from age, gender to the emotional state of a person.

Several graphics approaches for modeling and generating human characters have taken an inside-out approach and relied on complex anatomically inspired, physically based models of the human body [57, 61]. However, humans are surprisingly good at deriving conclusions based on what they observe; e.g. even by looking once at a stranger we can tell whether he is happy, tired, how old he is, if he exercises regularly *etc.* All we observe is surfaces embedded in a 3D space! One of the first attempts to model full-body skin deformations using surface data employed a marker-based capture system with a large number of markers distributed over the body of the subject [70]. The results were promising, but the approach suffered from time consuming instrumentation of the subject and could not capture deformations of the whole surface of the human body. Nowadays, we are presented with a set of alternatives. High-resolution 3D scanners, image-based 3D reconstruction systems, low-resolution 3D capture systems offer us the ability to capture the whole surface of the body, potentially across time, and generalize easily to capturing different human shapes. Examples of the abovementioned approaches and systems are shown in Figure 1.1. Modeling the human body in 3D using data can lead to realistic animation of virtual humans in movies, video games, medical simulation environments, *etc.*

With the large amount of 3D data of humans that is being gathered lately and the increasing popularity of 3D virtual world environments, there is a motivation to establish a connection between a person and the 3D representation of that person. For instance, in a *virtual try-on* scenario (Figure 1.2), users could potentially upload their 3D avatar and get personalized prediction of their size as well as personalized suggestions on how different products may fit them. In the fashion industry, each human body is identified by its size in a coarse scale and its tailor measurements in a finer scale. Although tailoring measurements convey less

¹<http://123kinect.com/first-kinect-2-depth-map-image-2/36985/>

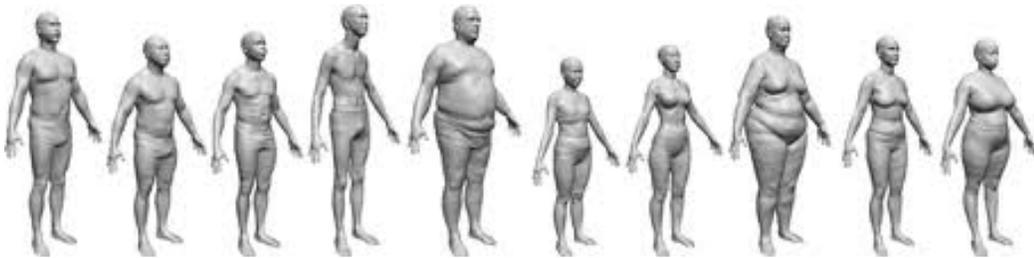
²<http://www.enmalayalam.com/site/english/topic/entertainment/category/variety/2011/10/3307-article>



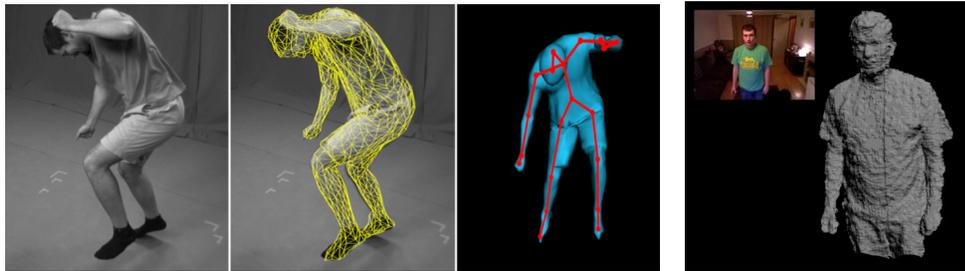
(a)



(b)



(c)



(d)

(e)

FIGURE 1.1. Graphics techniques for generating moving human bodies: (a) anatomically inspired, physically based model [108], (b) marker-based capture of skin deformations [70] (left: animated body, right: subject instrumentation). 3D Reconstruction of humans from (c) a high-resolution laser scanner (CAESAR dataset [4, 78]), (d) an image-based reconstruction system [33] (left to right: captured image, reconstructed surface, estimated skeleton), (e) Microsoft Kinect [55] ¹

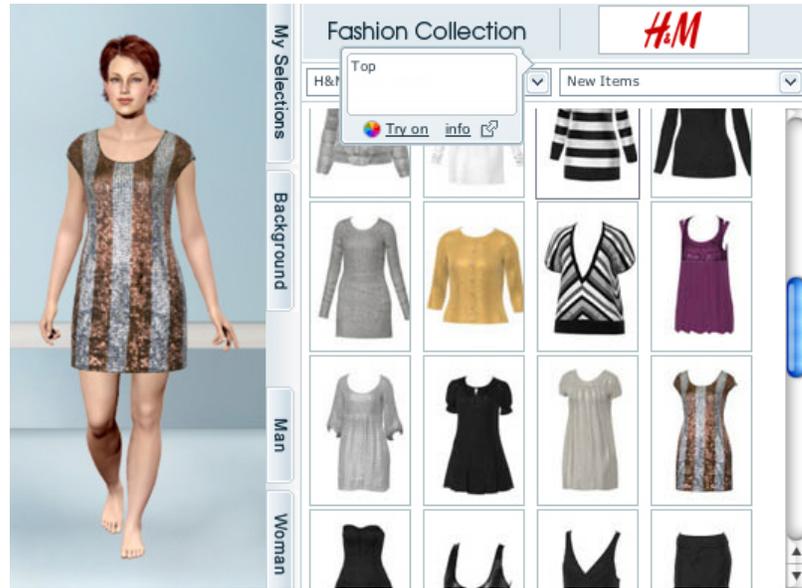


FIGURE 1.2. Virtual try-on scenario ²: Current online dressing rooms can be enhanced with realistic 3D avatars of actual users and personalized size recommendations.

information than the whole body surface of a human, measurement prediction from a 3D mesh representing a person would facilitate greatly applications such as virtual try-on.

3. Modeling the Human Body in 3D

Although there are many systems for capturing bodies in 3D, the acquired data cannot be readily used as a human body model. There are two main problems. The first problem is that a scan of a person may contain holes and thousands of unordered points. In addition, different scans of the same or of different people have different numbers of points. The second problem is that the shape of the human body is a time-varying quantity. A 3D scan is essentially a single snapshot of the human body in time (Figure 1.3) while the shape of the human body changes with motion, breathing, aging, *etc.*

In this work, we focus on three key problems of modeling the human body in 3D. We start with describing an approach for finding sparse correspondences between 3D models of humans (or, in general, non-rigid articulated objects) varying in pose, shape and resolution. This approach can be used in applications, such as object recognition and retrieval, *etc.* We proceed by making the connection between 3D models of the human body and standard anthropometric measurements. This could be useful for applications, such as *on-line shopping* and *virtual try-on* described above. Finally, we describe our approach on capturing and modeling intra-person variation in human shape from involuntary activities such as breathing. Capturing these shape deformations offers the potential to enhance the realism of animated human characters.

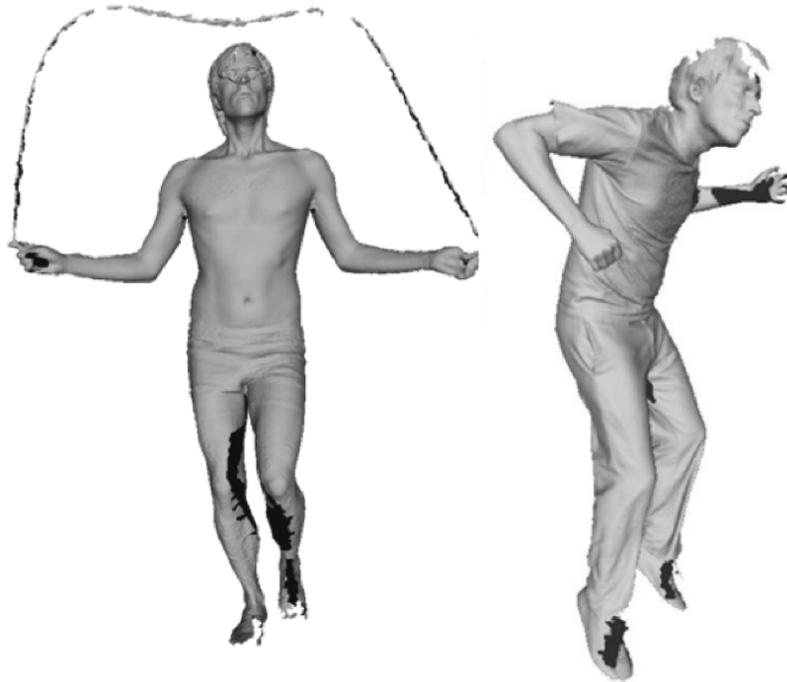


FIGURE 1.3. A 3D scan is only a snapshot of the human body shape in time.

3.1. 3D Articulated Deformable Object Representation. A common way to represent an articulated 3D deformable object is using a 3D triangular mesh (Figure 1.4). The observed shape of the articulated 3D deformable object is determined mainly by three factors: pose, intrinsic shape, resolution. Although we are not limited to 3D human bodies, we will focus mainly on 3D meshes of humans.

- **Resolution** refers to the number of vertices in the 3D mesh representing an articulated 3D object. Figure 1.4 shows two triangular 3D meshes with different resolution.
- Even when not explicitly modeled, it is often assumed that there is a skeleton associated with the 3D mesh. **Pose** refers to the relative joint rotations in the underlying skeleton or the rotation and translation of groups of triangles (i.e. parts). Figure 1.5a shows an example of parts in a 3D human mesh displayed using different colors. Figure 1.5b shows an example of two 3D bodies that differ only in pose. Notice that apart from change in observed shape due to the rigid rotation of the body parts, pose induces non-rigid change in shape because of muscle contraction, body fat, *etc.* For instance, our stomach is more flat when we are standing than when we are squatting or sitting.
- **Intrinsic shape** typically refers to the pose-independent shape, such as shape related to height, weight, waist size etc. For instance, as we see in Figure 1.6a, a fat and a thin person will have different intrinsic shape. People that look very similar but are scaled versions of each other will have different intrinsic shape as well (Figure 1.6b). The intrinsic shape is some sort of identity for the specific person.

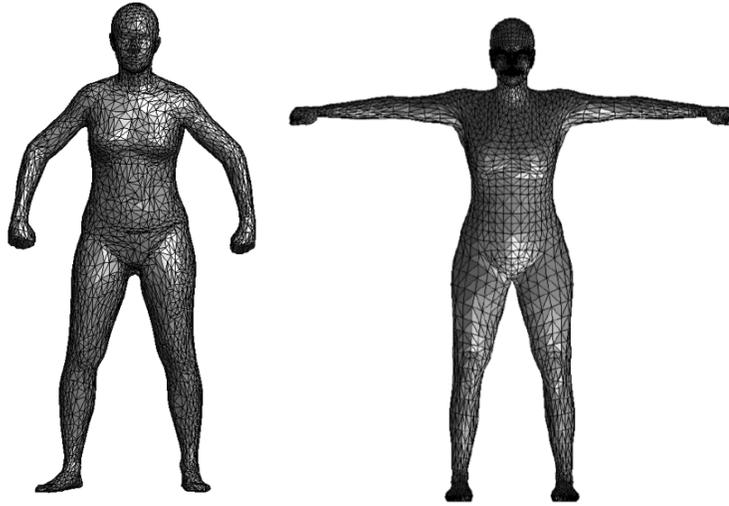


FIGURE 1.4. Triangular 3D meshes that differ in mesh topology.

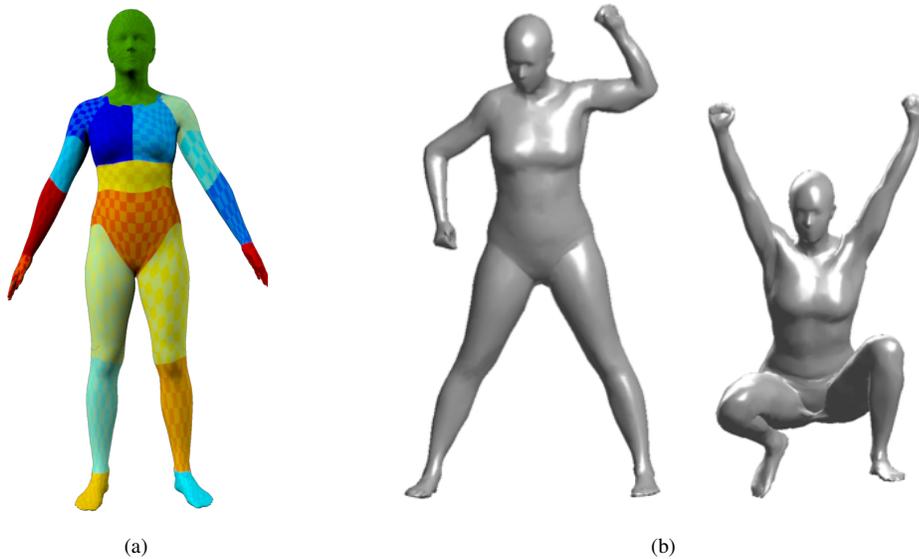


FIGURE 1.5. (a) Human body pose refers to the relative rotation of body parts (displayed with different colors). (b) Example of changing the pose of a 3D mesh.

However, this definition of the observed 3D shape is just an approximation. In reality, the observed 3D shape of a human depends also on the motion of the body, involuntary activities such as breathing, *etc.* In the last part of the thesis, we will introduce an additional component of the observed 3D shape that corresponds to shape due to breathing. For simplicity, from now on we will refer to the intrinsic shape using the term shape.

3.2. Shape-, Pose- and Resolution- Invariant Correspondences between 3D Non-rigid Articulated Objects. Finding correspondences between 3D articulated non-rigidly deformable objects is a critical task

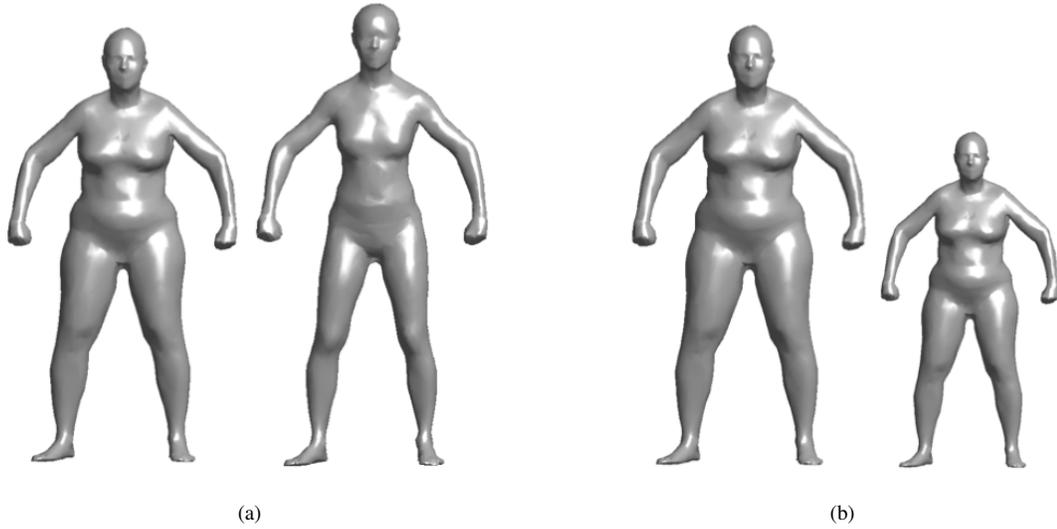


FIGURE 1.6. (a) Example of 3D meshes with different intrinsic shape. (b) 3D meshes that look very similar but are scaled versions of each other have different intrinsic shape as well.

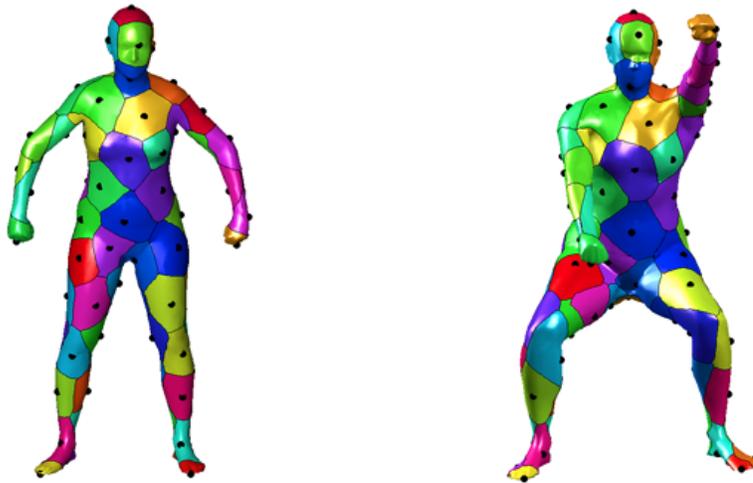


FIGURE 1.7. **Pairwise Correspondences between 3D meshes:** Example of correspondences obtained with our method (PGSE) for two 3D meshes of humans varying in pose and shape. Corresponding areas are denoted with the same color.

for many applications, such as object recognition and retrieval, shape deformation and morphing, 3D surface registration, *etc* [90, 92]. The term correspondences refers to pairs of points between the surfaces of two 3D meshes that are “similar” based on some notion of similarity; e.g. the nose tips in two 3D meshes of humans. Our work is based on the common assumption that deformations due to pose are isometric, which means that all geodesic distances between surface points P in one mesh are similar to the geodesic distances between all the corresponding points of P in another mesh. In addition to preserving geodesic distances, we use local

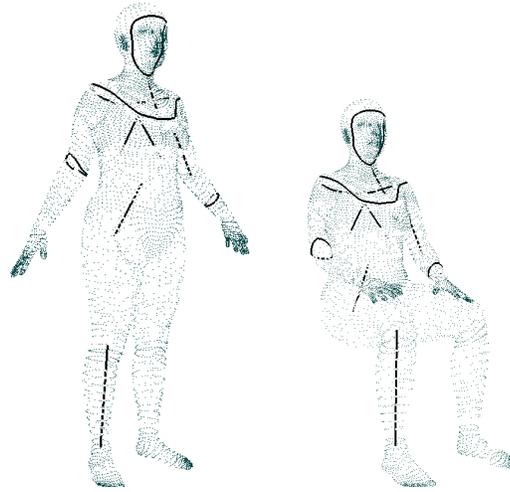


FIGURE 1.8. **Model-based extraction of anthropometric measurements in 3D.** Registered meshes of a person in 2 poses after fitting a human body model [6, 45] (the dots correspond to mesh vertices). Measurement prediction is based on local and global shape features. Local shape features comprise circumferences on the mesh surface (shown as curves) and limb lengths (shown as straight lines). Global shape features (not shown here) describe statistics of shape in a database of registered 3D meshes.

surface descriptors that are invariant to intrinsic shape, pose, and resolution variations. An example outcome of our method is shown in Figure 1.7.

3.3. Measuring the Human Body in 3D: Correlation of the 3D Human Body Shape with Anthropometric Measurements. Studying how human shape varies across the population is useful for many applications, ranging from designing clothes and furniture to coming up with ways to assess obesity. The standard way to quantify human body shape has been anthropometric measurements. What is the optimal way for measuring our 3D human body? Here, we make the connection between 3D human shape and a set of standard anthropometric measurements. We propose a set of local and global features representing human shape in 3D and learn the optimal features for predicting each anthropometric measurement (Figure 1.8). We envision using this work for personalized sizing and measurement extracting in on-line shopping applications.

Although anthropometry is a well established means for measuring the human body, the dimensions correspond to a limited set of lengths and paths on the human body surface. Are they able to capture the rich shape information that human bodies in 3D convey? In a novel application, we extend Principal Component Analysis (PCA) to constrain the main directions of shape variation to be predictive of standard anthropometric measurements. The remaining shape variation provides insights into what standard anthropometric measurements do not capture (Figure 1.9). This “nullspace” of human body shapes is a new intriguing concept with implications for human factors and engineering anthropometry.

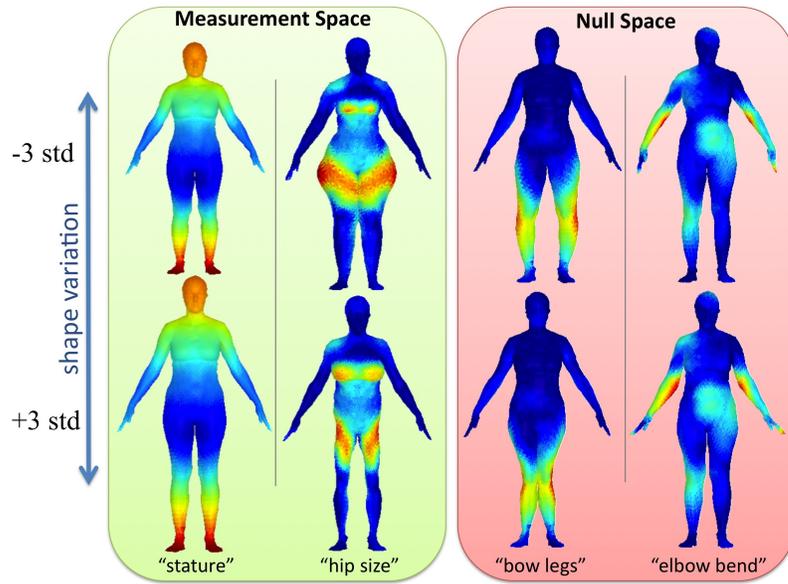


FIGURE 1.9. **Correlating human body shape variation with measurements:** We factor human body shape variation into a measurement space that is constrained to predict anthropometric measurements like height and hip girth and a *null space* of body shape variation that are unrelated to these measurements such as bow-legs and pose/posture variations.

3.4. Data-driven Modeling and Animation of Breathing. Modeling how the human body deforms during breathing is important for life-like animation of 3D avatars. Most procedural animation lacks the nuance of real breathing while physical simulation remains too complex to apply to a wide range of bodies and breathing types. In contrast, we learn a detailed model of body shape deformations due to breathing across different breathing types (Figure 1.10) and provide simple animation controls to render lifelike breathing. We capture and align high-resolution 3D scans of 58 human subjects. We compute deviations from each subject’s mean shape during breathing, enabling us, for the first time, to study the statistics of such shape changes for different genders, body shapes, and breathing types. We use the volume of the registered scans as a proxy for lung volume and learn a novel non-linear model relating this volume and breathing type to 3D shape deformations and pose changes. We then augment a SCAPE body model [6] so that body shape is determined by identity, pose, and the parameters of the breathing model.

4. Challenges

In this section, we describe in more detail the challenges involved in modeling the human body in 3D.

4.1. Correspondences between 3D non-rigid articulated objects varying in pose, shape, and resolution. Although the problem of establishing correspondences among rigid objects has been addressed in the literature adequately [89], finding correspondences between non-rigid deformable articulated objects is



FIGURE 1.10. **Animating breathing types.** Respiration induces changes in torso shape and posture. We learn a model of how 3D breathing deformations relate to lung volume and breathing type and use it to animate bodies of varying shape and pose. Here we show the maximal inhale and exhale shapes overlaid for three different bodies breathing mainly with the stomach (left), mainly with the chest (right), or using a combination of chest and stomach (middle).

still a challenge. Variations in pose and intrinsic shape change the local geometry of the objects surface, increasing the likelihood of a false match. In addition, matching two objects entails solving a combinatorial problem in the exponential space of possible pairwise correspondences. Such an optimization may get stuck in local optima resulting in non-meaningful correspondences. More specifically, the difficulties in finding correspondences between 3D meshes arise from the following:

- **Pose-dependent geometry.** The geometry of the human body changes with the pose of the human; e.g. the area around the knee deforms in a different way when someone is standing compared to when he is jumping.
- **Shape-dependent geometry.** Shape variation across the human population is not a product of scaling a “canonical body shape” along three dimensions. There are many subtle variations of human body shape; e.g. two people with mostly the same body shape, but one of them having a bigger belly than the other.
- **Mesh resolution.** Mesh resolution refers to the number and distribution of vertices along the surface of an object represented by a mesh. In that sense, a mesh is an approximation of the object’s surface. Due to this subsampling of the object’s surface, operations on meshes with different resolution may yield different results; e.g. shortest-paths on the surface may be slightly different than shortest paths on a graph formed by the vertices/edges of the mesh.

- **Exponential number of correspondences.** Given two meshes with M and N vertices respectively, the number of possible vertex correspondences from the first to the second mesh is N^M . Assumptions about deformations help to restrict the number of candidate correspondences. For instance, the common assumption of isometric deformations states that shortest path distances on a 3D mesh of an articulated object are nearly invariant relative to pose (e.g. human standing vs. human sitting) [19].

4.2. Anthropometric measurement prediction from 3D human bodies. A main challenge in predicting measurements from 3D data of human bodies is to accommodate data that are noisy or incomplete. More specifically, measurement prediction inevitably contains a set of errors in the following order: real bodies are captured (acquisition error), the captured data are potentially registered to a template mesh (registration error) and the resulting meshes are used to predict measurements on the initial real bodies (computational measurement prediction error).

- **Data acquisition.** Apart from the acquisition error of the capturing device (e.g. holes and noise in a scan), capturing the intrinsic shape of a human is a hard task. The observed shape of a human is a time varying quantity; i.e. it changes with pose as well as during involuntary activities, such as breathing. In that sense, a single scan per person contains only partial information about his intrinsic shape. In order to model the human body realistically we need to take a set of representative samples of an individual's shape variation.
- **Registration.** The data that we get from high-resolution scanners usually contain a large number of points ($\sim 80K$), holes and noise. In order to make use of this data, it is common practice to fit an existing mesh with desired properties (e.g. symmetric mesh with varying resolution based on mesh curvature) to the captured data. During the registration procedure there is a tradeoff between explaining every detail of the captured geometry and producing a locally smooth fit of the captured data.
- **Shape Descriptors.** Generic 3D shape descriptors have proved useful for classification tasks, e.g. object retrieval, but they are not very discriminative between shapes within a single class. Given objects of a specific category, a common way to characterize variation in shape is using dimensionality reduction techniques. However, for the case of anthropometric measurement prediction, the resulting components of global shape variation correlate poorly with measurements on the human body. For instance, measurements that do not dramatically impact the whole body (e.g. wrist circumference) will be only slightly correlated with the major directions of shape variation found from the abovementioned techniques. Local shape descriptors can potentially enhance the performance of global shape descriptors, but it is not obvious what type of local descriptors should be used and how local they should be.

4.3. Modeling human shape and pose change due to breathing. Although body deformations due to breathing are subtle, they are essential for perceiving “life” in an animated 3D character. Yet they are so complex that hand-tuned animation of breathing is laborious and limited in realism [1, 2]. Previously proposed anatomy-based models of breathing are more realistic than hand-tuned animation, but are computationally expensive and do not generalize easily to new shapes and styles. Data-driven approaches are very promising for overcoming the limitations of previous work using data captured from real humans. Nevertheless, there are still many challenges to overcome.

- **Data acquisition.** In order to model the fine deformations of the human body due to breathing, we need high-resolution data. At the same time, breathing is a temporal process which means that multiple scans per subject need to be captured at a relatively high frame rate. However, state-of-art systems for 3D acquisition (Figure 1.1) capture data in either high temporal and low spatial resolution or vice versa. That explains the limited previous work on data-driven approaches for modeling breathing. In addition to capturing detailed data in terms of spatial and temporal resolution, registering the data in order to perform statistical analysis needs to be performed in an accurate way.
- **Extraction of breathing deformations.** Breathing is an effect that cannot be directly observed in a single time instance. Given a single 3D human scan, it is ambiguous which part of the observed shape is due to breathing and which part corresponds to the intrinsic shape of the human. It is even ambiguous which part of the observed shape is due to the intrinsic shape or pose of the subject. How can we factor out all the causes of observed shape apart from breathing? Extracting breathing deformations inevitably means that multiple scans of the same subject need to be examined. In each scan we need to at least discern between deformations due to pose, shape, breathing. Breathing itself induces both pose and shape change.
- **Breathing type.** Deformations of the human body due to breathing vary across people and contexts [98]. Different people breath in a different way. Different activities induce different patterns of breathing. At rest breathing seems to be an almost periodic function of time, in other scenarios it is just a timeseries depending on other factors, such as activity, pose, *etc.*
- **Animation controls.** To make animation of breathing practical a compact set of intuitive, but expressive, controls needs to be derived.

5. Contributions of the Thesis

In this thesis, we present different aspects of modeling the human body in 3D ranging from low-level processing of 3D data representing humans, such as finding correspondences, to higher-level processing, such as estimating anthropometric measurements. Our main contributions can be summarized as follows:

- (1) We describe a method for finding surface vertex correspondences of a non-rigid object represented as a 3D mesh undergoing significant deformation due to *pose* and *shape* variation. The presented method is also suitable for finding surface point correspondences between objects differing in global/local *resolution* and triangulation, containing up to a small proportion of holes. It incorporates a search procedure that explores effectively the space of possible correspondences and is more robust to local optima than previous work. It relies on a discriminative probabilistic model that preserves properties related to geodesic distances and uses Loopy Belief propagation (LBP) [103] for inference.
- (2) We present a model-based approach for predicting anthropometric measurements from registered high-resolution 3D human scans. Based on point-to-point correspondences among the registered scans, we propose a set of features that describe global and local human shape variation and are, additionally, predictive of measurements. We select the optimal features per measurement automatically using regularized linear regression. In most of the cases, the optimal local features per measurement are indicative of the corresponding measurement; e.g. circumference features near the desired “circumference measurement” or limb length features related with the desired “height measurement”. Finally, we show that using 3D scans of the same person in multiple poses increases the accuracy of predicting measurements for the specific person.
- (3) We propose a constraint optimization framework with a flexible objective function that yields human shape representations that correlate variation in human shape with different factors. These factors are used to constrain directions of shape variation coming from common dimensionality reduction techniques. The proposed objective function is general enough to be used with different constraints, such as anthropometric measurements, age, gender, income etc. Here, we correlate human shape variation with standard anthropometric measurements and find a subspace of shape variation that is predictive of measurements. The remaining shape variation, that we term “nullspace”, sheds light onto what kind of shape variation standard anthropometric measurements do not capture.
- (4) We introduce a non-linear model for synthesizing breathing deformations for animated 3D human characters. Our model is parameterized by the breathing volume and type (% of chest breathing). We augment a SCAPE body model so that body shape is determined by identity, pose, and the parameters of the breathing model. These parameters provide an intuitive interface with which animators can synthesize 3D human avatars and their breathing motion. We show results of interactive animation from arbitrary sequences of breathing volumes coming from real humans, such as sequences captured by a spirometer. Our approach generates fine-scale body shape deformations due to breathing with greater ease and realism than previously achieved. Furthermore, we give insights on how human shape and pose during breathing correlates with gender and intrinsic human shape.

Shape- and Pose-Invariant 3D Correspondences

Finding correspondences between non-rigid 3D deformable objects is a critical task for many applications. Examples include object recognition and retrieval, shape deformation and morphing, 3D surface registration, *etc.* [90, 92]. By defining correspondences using a structure preservation criterion, we can assess the similarity between two objects based on the amount of structure distortion. For applications involving search for similar 3D object models, it may be critical to have a measure of similarity that is invariant to common variations within a class (e.g. body pose and identity variation). Additionally, mesh alignment, for example of laser scans of human bodies, typically employs surface registration methods like ICP [13], [81] which require an initial set of correspondences. Here we describe a fully automated method for obtaining such correspondences between meshes that vary in shape, pose, and resolution.

Although the problem of establishing correspondences among rigid objects has been addressed in the literature adequately, finding correspondences between non-rigid deformable objects is still a challenge [89]. Variations in pose and shape change the local geometry of the object’s surface increasing the likelihood of a false match. In addition, matching two objects entails solving a combinatorial problem in the exponential space of possible pairwise correspondences. Such an optimization may get stuck in local optima resulting in non-meaningful correspondences. Figure 2.1 shows an example of non-meaningful correspondences produced by related work, Generalized Multi-Dimensional Scaling (GMDS) [19], where the chest is mapped to the back of the human model and vice versa. This effect is significantly diminished using our method, Probabilistic Geodesic Surface Embedding (PGSE).

The main contributions of this work can be summarized as follows:

- A method for finding surface point correspondences of a non-rigid object undergoing significant deformation due to *pose* and *shape* variation.
- A method for finding surface point correspondences between objects differing in global/local *resolution* and triangulation, containing up to a small proportion of holes.
- Correspondence search that effectively explores the space of possible correspondences and is more robust to local optima than previous work. It relies on a discriminative probabilistic model that preserves properties related to geodesic distances and uses loopy belief propagation (LBP) for inference.

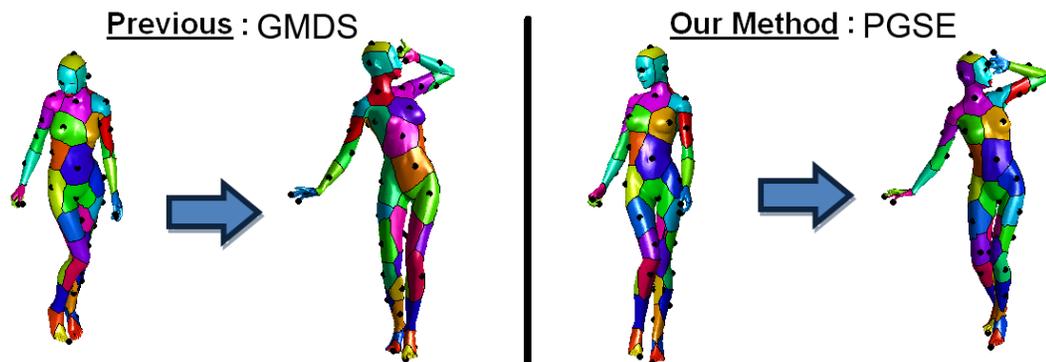


FIGURE 2.1. Local optima in a combinatorial optimization problem for matching objects varying in pose and shape. Previous work, Generalized Multi-Dimensional Scaling (GMDS) [19], relies only on the preservation of geodesic distances and can yield non-meaningful correspondences; e.g. the chest of the body in the left pose is mapped to the back of the body in the right pose (corresponding regions are shown with the same color). Our method, Probabilistic Geodesic Surface Embedding (PGSE), achieves more intuitive results by combining geodesic distances with local surface descriptors in a coarse-to-fine probabilistic optimization framework.

1. Related Work

Surface registration is often used as a necessary step in various applications, such as content-based shape retrieval, deformation transfer, model training from data in correspondence, pose estimation etc. [6, 73, 92]. There is a plethora of methods and descriptors for rigid registration [67, 80, 89]; we will focus on non-rigid registration of articulated deformable objects. Previous methods for matching nonrigid deformable objects with significant variation in pose aim at providing global consistency of correspondences by preserving intrinsic properties of the objects. Usually these methods find deformation-invariant representations of the objects and match the objects in the representation domain. Examples include the use of geodesic distances [19], diffusion distances [20] or representations in the Möbius domain [59].

Although preservation of the intrinsic properties of the objects may be sufficient to assess their similarity, intrinsic-only matching criteria are oblivious to object self-symmetries and may yield non-meaningful correspondences. To overcome this weakness, previous work has explored the use of local surface properties and/or costs of surface deformation. Previous local surfaces properties are either geometric or based on the intrinsic characteristics of the shape or both. For instance, the work in [7] uses oriented histograms describing the distribution of points in local neighborhoods along the object surface (spin images [50]). Dubrovina *et al.* [29] use a local surface descriptor based on the eigenvalues of the Laplace-Beltrami operator which is related to the flow in the mesh representation of the object. Aubry *et al.* [8] derived a descriptor considering the

Schrödinger equation governing the temporal evolution of quantum mechanical particles. Wang *et al.* [96] use descriptors based on curvature and surface normals targeted towards a specific class of surfaces (brain surfaces). Efforts that also take into account object deformation include [47], [107].

Most previous work considers pose variations of the same object. To the best of our knowledge, only the work in [107] considers variations in shape, but the objects to be matched do not have significant differences in pose. We are concerned with finding correspondences among objects of the same category varying in shape, pose, and resolution. Extending previous approaches for global matching, we rely on preserving normalized geodesic distances to account for the additional variation in shape. We also employ a probabilistic framework for optimization similar to the one in [7]. We enforce stricter geodesic preservation constraints and use alternative local surface descriptors that are invariant to shape, pose, and resolution variations.

2. Probabilistic Geodesic Surface Embedding

We consider the problem of finding correspondences between two triangular meshes, a model mesh X and a data mesh Z . The *model mesh* $X = (V^X, E^X)$ is a complete surface consisting of a set of vertices $V^X = (x_1, \dots, x_{N^X})$ and a set of edges E^X . The *data mesh* $Z = (V^Z, E^Z)$ may contain a modest number of holes (missing data); the vertices and edges are $V^Z = (z_1, \dots, z_{N^Z})$ and E^Z respectively. Typically the data and model meshes differ in shape, pose, and resolution. Each data mesh vertex $z_k, k = 1, \dots, N^Z$ is associated with a *correspondence variable* $c_k \in \{1, \dots, N^X\}$ that specifies the model mesh vertex it corresponds to. The task of finding correspondences is one of estimating the most likely set of all correspondence variables $C = (c_1, \dots, c_{N^Z})$ given a specific pair of model and data meshes X, Z .

2.1. Probabilistic Model. We cast the problem of finding correspondences as one of finding the most likely embedding of the data mesh Z into the model mesh X encoded as an assignment to all correspondence variables $C = (c_1, \dots, c_{N^Z})$. More specifically we take a discriminative approach where our goal is to find a configuration of C that maximizes the distribution $p(C|X, Z)$ over all correspondence variables conditioned on a pair of mesh instances X, Z . Writing this distribution as an undirected graphical model, we get the Conditional Random Field (CRF) model [56] depicted in Figure 2.2. Each latent variable node in the model denotes the correspondence variable c_k of vertex $z_k, k = 1, \dots, N^Z$, in the data mesh. The observed variable is a pair of model and data meshes X, Z .

We *approximate* the conditional distribution of the correspondence variables using potential functions, ψ , linking all *pairs* of latent variables and unary potentials, ϕ , linking each latent variable with the data. Formally we approximate the conditional distribution as: $p(C|X, Z) \propto \prod_k \phi(c_k, X, Z) \prod_{k,l} \psi(c_k, c_l, X, Z)$.

The main idea behind our approach is that the geodesic distances between points in the data mesh Z should be the same as the geodesic distances between the corresponding points in the model mesh X . Our method searches for correspondences that satisfy this property. At the same time we want to preserve in the

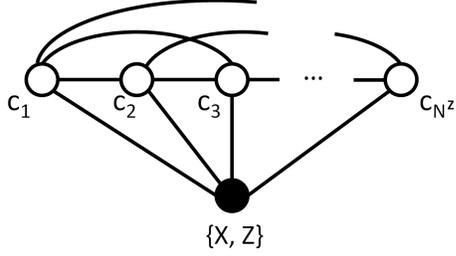


FIGURE 2.2. Conditional Random Field (CRF) model for finding correspondences. The observed variable in the model is a pair of a model mesh $X = (V^X, E^X)$ and a data mesh $Z = (V^Z, E^Z)$. The latent variables are the correspondence variables $C = (c_1, \dots, c_{N^z})$ of all data mesh vertices. Edges in the model between latent and observed variables favor correspondences that preserve the intrinsic properties of the data mesh vertices. Geodesic constraints between all possible pairs of correspondence variables are enforced through the edges between latent variables.

embedding the intrinsic geodesic properties (geodesic signature) of the data mesh vertices. All the above-mentioned constraints are enforced using the potentials described below.

Pairwise geodesic potential $\psi(c_k, c_l, X, Z)$: We consider normalized geodesic distances as the invariant used to match meshes that deform non-rigidly due to changes in shape and pose. We calculate exact geodesic distances using the the Fast Marching method described in [54]. For each pair of data mesh vertices z_k, z_l , we define a potential function $\psi(c_k, c_l, X, Z)$ that constrains the pair of correspondences c_k, c_l in the model mesh X to be geodesically consistent with vertices z_k, z_l in the data mesh Z (Figure 2.3). Let $M = (V, E)$ be a mesh with vertices V and edges E and $h : V \times V \rightarrow \mathfrak{R}$ be a geodesic distance function. Then $h(j, m; M)$ represents the normalized geodesic distance between two vertices j and m in mesh M . The normalization is done by dividing the geodesic distance by the maximum geodesic distance over all pairs of vertices in M . The geodesic potential between a pair of data mesh vertices z_k, z_l is defined as

$$(1) \quad \psi(c_k, c_l, X, Z) = N(h(c_k, c_l; X); h(k, l; Z), \sigma_{kl}^2)$$

where σ_{kl} is a user defined parameter; here $\sigma_{kl} = 0.1 \cdot h(k, l; Z)$.

Geodesic signature potential $\phi(c_k, X, Z)$: We encode a potential that enforces that corresponding vertices c_k in the model mesh have similar intrinsic properties as those in the data mesh z_k (Figure 2.4). Our goal is to distinguish spatially different areas in the model and data meshes as much as possible. The intrinsic property we use is the mean normalized geodesic distance of vertex z_k over all possible vertices in the data mesh (geodesic signature). The resulting potential can be written as

$$(2) \quad \phi(c_k, X, Z) = N(g(c_k; X); g(k; Z), \sigma_k^2)$$

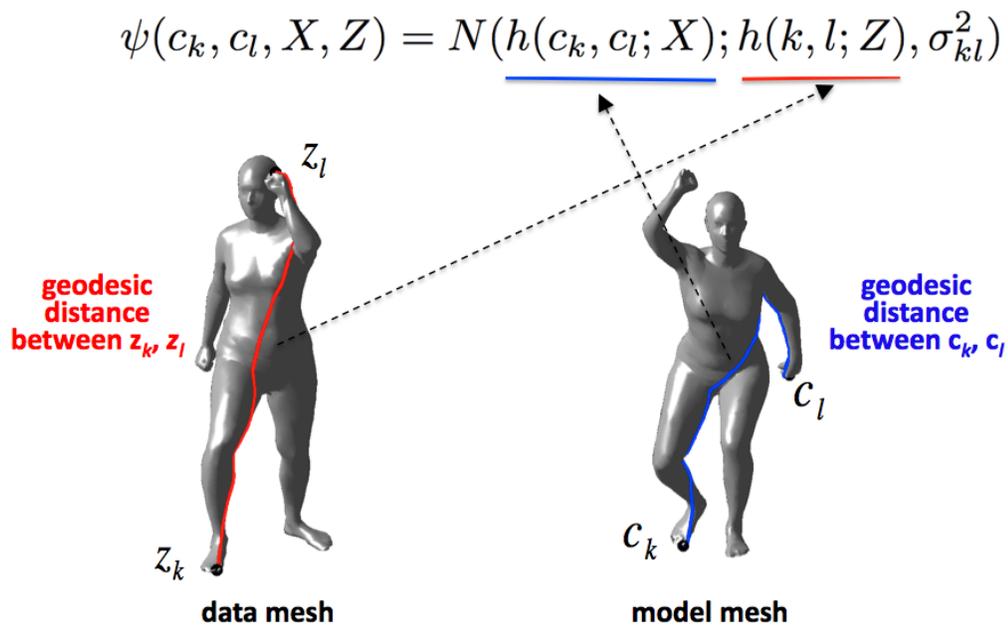


FIGURE 2.3. The pairwise geodesic potential enforces preservation of geodesic distances between pairs of vertices (z_k, z_l) and their correspondences (c_k, c_l) .

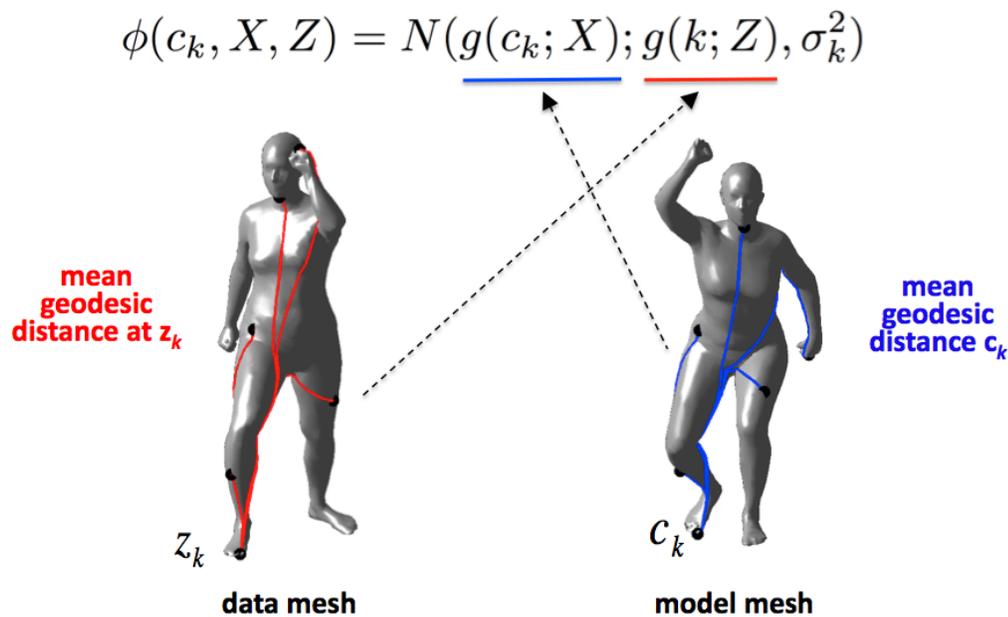


FIGURE 2.4. The geodesic signature potential matches vertices in the data and model meshes with similar average geodesic distances to the rest of the mesh vertices. Only a subset of geodesic paths is displayed.

where $g(j; M = (V, E)) = \frac{1}{|V|} \sum_{m \in V} h(j, m; M)$ is the mean normalized geodesic distance from j to all other vertices m in the mesh M and σ_k is a user defined parameter. The use of geodesic signatures is important because it biases the embedding of the data mesh to the model mesh to match spatially similar areas between the meshes. In practice we observe that this also improves convergence of the optimization procedure described below.

2.2. Inference. Our goal is to find an assignment of the correspondence variables that maximizes the probability $p(C|X, Z)$ as represented by the graphical model. Exact inference is computationally infeasible due to the large number of variables and loops in the graph. Instead we use max-product *loopy belief propagation* (LBP) [103] for approximate inference. Running LBP until convergence yields a set of probabilities over model mesh vertices for each correspondence variable c_k . We compute the optimal correspondence for each data mesh vertex z_k as the model mesh vertex that maximizes the probability distribution of the correspondence variable c_k .

Our inference scheme is performed in two rounds as shown in Figure 2.5. In the first round, the data mesh is sampled at a coarse level (Figure 2.5 (c)) using the farthest point sampling method [32]. In a similar way, the model mesh is sampled at a coarse level (Figure 2.5 (d)) and an initial set of correspondences is obtained using LBP. In the second round, the initial correspondences are refined by restricting the domain for each correspondence variable to be geodesically close to the solution of the first round of inference (Figure 2.5 (f)). Here we restrict the search to vertices with a geodesic distance up to $1/2$ the average geodesic distance between nearby samples in the model mesh. The complexity of each round is $O(K^2L^2)$ where K is the number of samples in the data mesh and L the number of corresponding samples in the model mesh. To speed up convergence we start the message passing procedure from data mesh samples having maximum geodesic signature and then move to samples that are more and more geodesically further. Scan samples with maximum geodesic signatures are the most discriminative in terms of their potential correspondences in the model mesh. A few more iterations of random message passing follow.

2.3. Sampling. Samples in the data and model meshes were obtained using the Farthest Point Sampling ([32]) method. It’s a greedy algorithm very good at selecting samples containing the extremities of the object. Although our samples space is not limited to the extremities of the object, extremities are stable under bending and stretching which makes them suitable to use for finding correspondences.

2.4. Geodesic Distance calculation. The approach we use for calculating the geodesic distances on the meshes is the Fast Marching method described in [54]. Contrary to graph-based shortest path algorithms (e.g. Dijkstra’s) the shortest path in our case is not restricted to the graph edges. Shortest paths can pass through the triangular faces of the mesh yielding much more accurate calculations of geodesic distances.

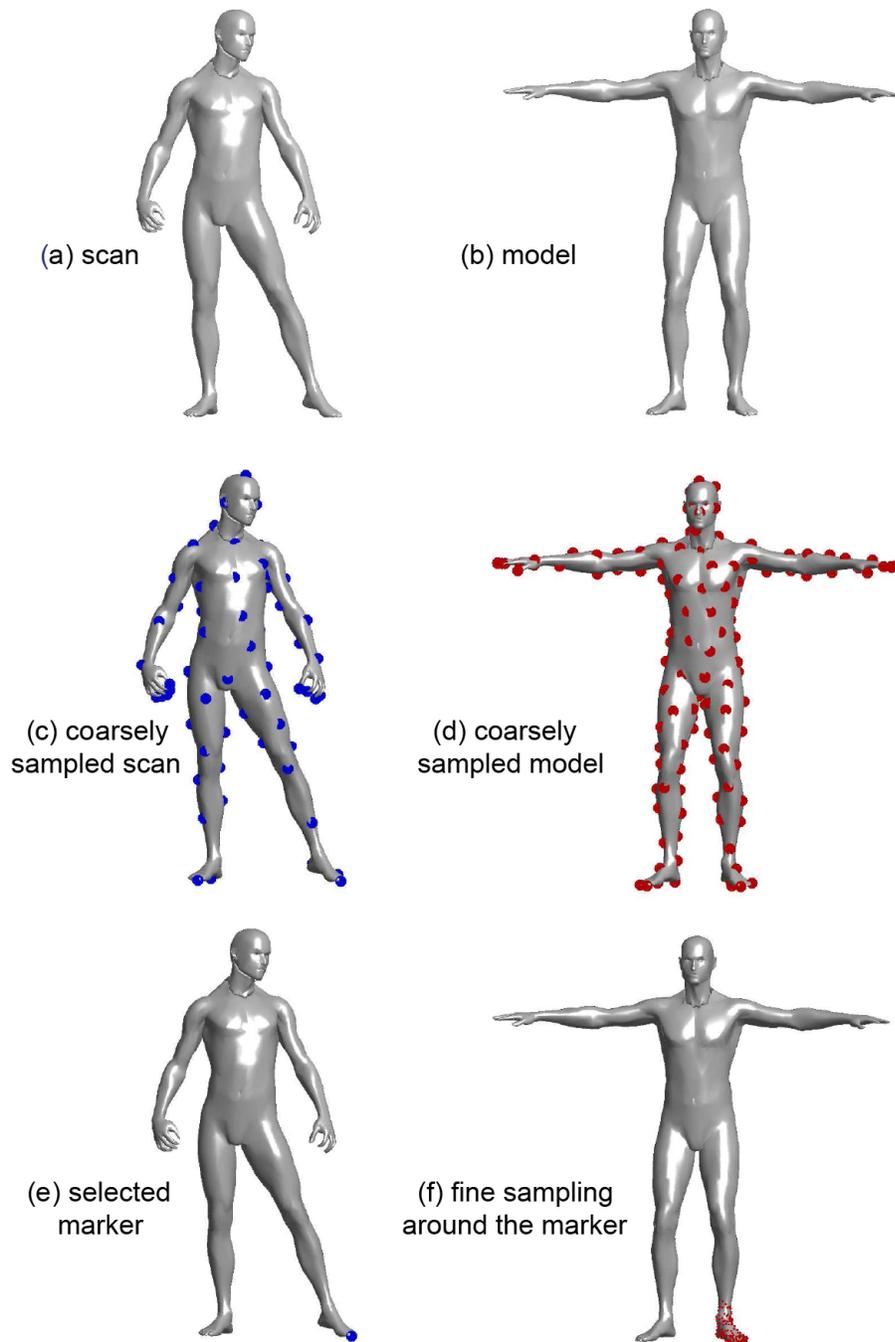


FIGURE 2.5. Illustration of the sampling process during the inference procedure. In the first round, a data mesh (a) and a model mesh (b) are sampled at a coarse level. A coarse sampling of the data mesh to e.g. 75 markers (c) and a coarse sampling of the model mesh to e.g. 150 samples (d) produce an initial set of correspondences. In the second round, for each individual marker in the data mesh (e), the domain of possible correspondences is obtained from finer sampling around the solution found in the first round (f).

3. Results

3.1. Data. We evaluate our algorithm on triangular meshes from the TOSCA nonrigid world database [18] and human bodies generated using the SCAPE model [6]. All the objects are represented as closed triangular meshes and they are simplified to have 2000-4000 vertices to aid comparison with related work. For each pair of meshes we find correspondences of 75-100 surface points. For the following experiments our method requires around 5GB of RAM per pair of meshes. The running time is approximately 1h on a 2.66GHz Intel Xeon processor.

3.2. Evaluation. The meshes we use do not come with any ground truth information about correspondences between their vertices. Typical error metrics in this case measure the degree that geodesic distances are preserved between the data mesh and the model mesh. Let us consider markers i, j in the data mesh and their optimal correspondences c_i^*, c_j^* in the model mesh. The following error metric, T_g , represents the average change in geodesic distances between all pairs of markers and their correspondences.

$$(3) \quad T_g = \frac{2}{|U|(|U|-1)} \sum_{i,j \in U} |h(i,j) - h(c_i^*, c_j^*)|$$

where U is the set of markers in the data mesh.

However, preservation of geodesic distances does not ensure that the correspondences are qualitatively meaningful. The smaller the number of markers used and the larger the number of self-symmetries in the object, the larger the number of possible correspondence configurations with geodesic distances similar to the geodesic distances between data mesh markers. We find that comparing Voronoi regions around the markers and their optimal correspondences provides a more intuitive measure than comparing the degree in which geodesic distances have been preserved (Figure 2.6). Similar Voronoi regions between the data and model meshes also lead to similar geodesic distances among markers and their optimal correspondences. The opposite is not necessarily true. Comparing Voronoi regions does not only include how well the geodesic distances are preserved, but also how similar the neighborhoods around markers and their optimal correspondences are.

Let $v_s(i)$ be the area of the Voronoi region around marker i and $v_m(c_i^*)$ the area of the Voronoi region around the optimal correspondence c_i^* of marker i in the model. We define the following error metric, T_e , representing the average change in the Voronoi area over all markers and their correspondences.

$$(4) \quad T_e = \frac{1}{|U|} \sum_{i \in U} \left| \frac{v_s(i) - v_m(c_i^*)}{v_s(i)} \right|$$

where U is the set of markers in the data mesh. In order to deal with shape variation, Voronoi areas are calculated as a percentage of the whole area of each object.

3.2.1. Correspondences in meshes with same topology. We compare our method, PGSE, to the GMDS method presented in [19] using triangular meshes of the same topology. For each object in the TOSCA nonrigid world database, we find correspondences to the canonical object of the category it belongs to. For

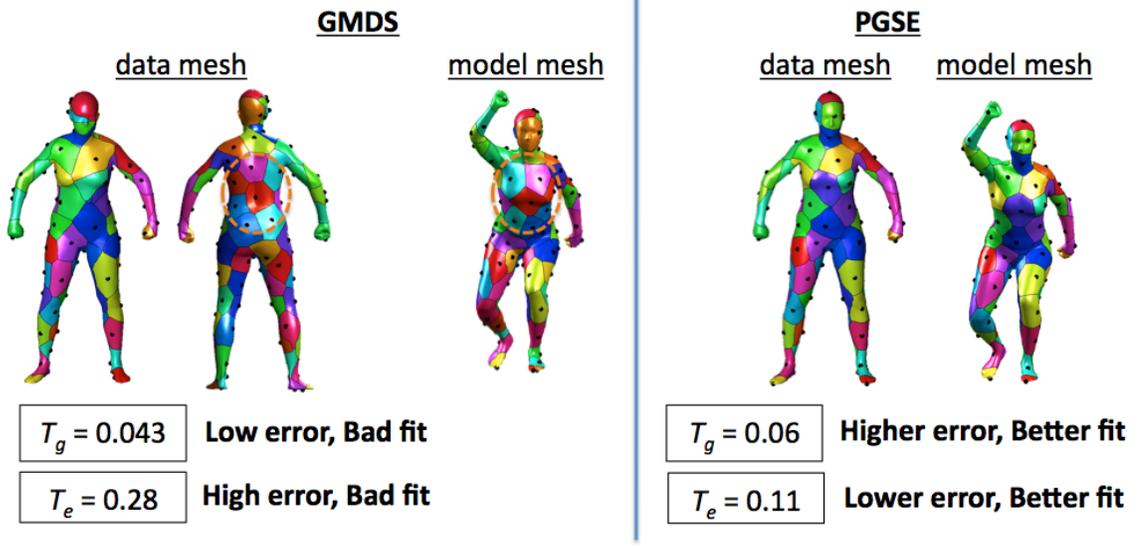


FIGURE 2.6. Comparison of error metrics in a case of non-meaningful correspondences. In the case of GMDS, vertices in the back are mapped to vertices in the chest and vice versa. Although correspondences from PGSE are more meaningful anatomically, they result in a higher geodesics-based error T_g . The Voronoi-based error T_e is more consistent with how a human would assess the quality of correspondences.

the SCAPE bodies, we find correspondences between the mean SCAPE body in the canonical pose as defined in the CAESAR dataset [78] and SCAPE bodies varying in pose, shape, and pose and shape together.

Figure 2.7 illustrates the correspondences found with GMDS and PGSE. Evaluating the correspondences using the error metric defined above, we get the error plots shown in Figures 2.8(a, b, c). For the parameterized bodies generated using the SCAPE model, we sort the results based on pose or shape variation. Pose variation is measured as the average joint angle deviation from the joint angle configuration in the canonical pose. It is weighted by the percentage of mesh vertices each joint controls and it is measured in radians. Shape variation is measured based on the L2-norm of the shape coefficients in the SCAPE model. Given the variety of categories in the TOSCA nonrigid world database, we present only summary statistics of the error over the database. For the case of PGSE, the average T_e error is 0.1410 with standard deviation 0.1059. For the case of GMDS, the average T_e error is 0.2799 with standard deviation 0.1564. Statistical significance was assessed using the Wilcoxon signed rank test [76]. The Wilcoxon signed rank test does not make any assumptions about the underlying distribution of the data, e.g. normal distribution. This attribute makes the test a reasonable choice given the great variation in the error as we change the pose, shape and resolution parameters.

In all cases we see that the error increases as we vary the pose or the shape. Although not reported with error metrics, GMDS performs better on average at preserving geodesic distances; this is not surprising

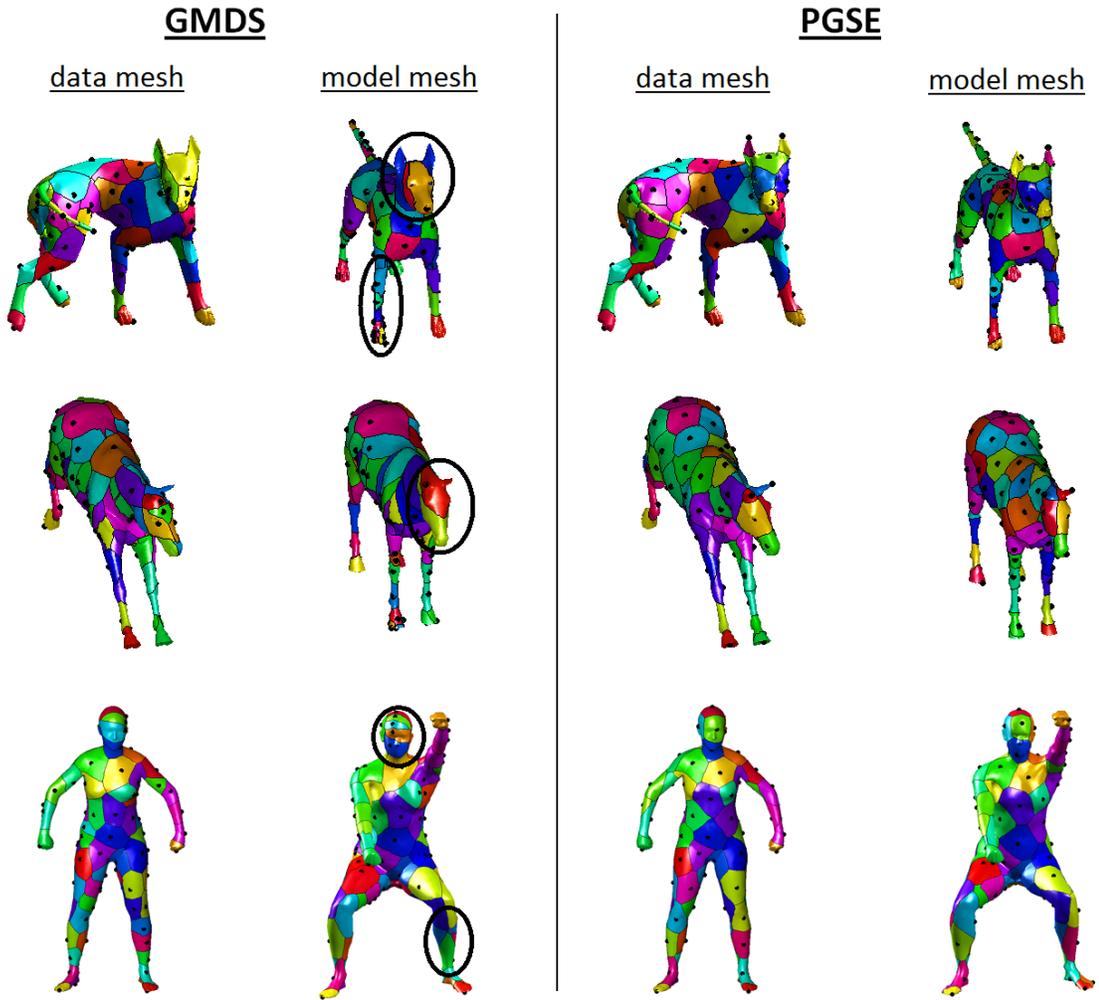
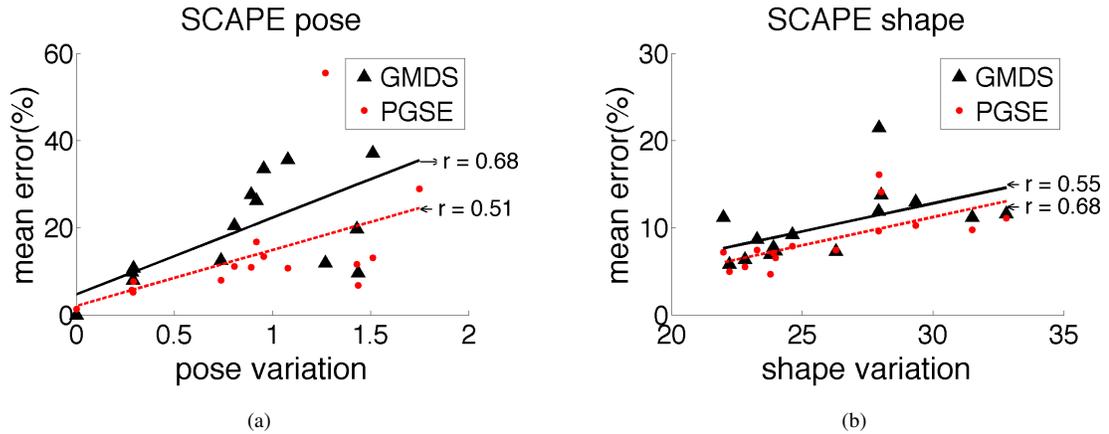


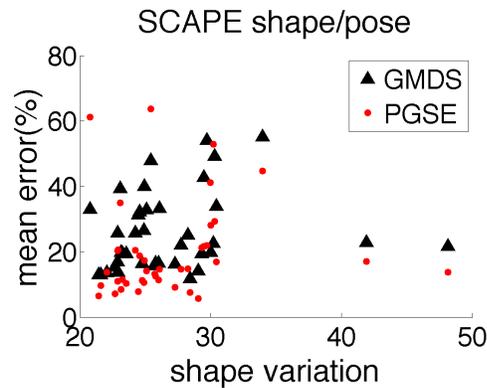
FIGURE 2.7. Visual correspondences between meshes in the TOSCA nonrigid world database and SCAPE bodies varying in pose and/or shape. Corresponding areas are shown with the same color. Areas where our method, PGSE, performs better than GMDS are circled. Note that correspondences are defined up to intrinsic symmetries (left-right) in the meshes.

as the GMDS method minimizes exactly this error. In contrast, our method combines the preservation of geodesic distances with local shape matching constraints. Our approach, PGSE, performs better in terms of the maximum discrepancy in geodesic distances between pairs of markers and their correspondences. Evaluating the correspondences using the T_e error (Figure 2.8), we see that PGSE performs better in all cases. Statistical significance values for the errors per dataset are shown in Table 2.8(d). Changing the pose yields a bigger increase in the mean error than changing the shape. Changing both shape and pose yields the biggest increase in error as expected. Note that our approach is still sensitive to local minima in the space of possible correspondences (outliers in Figure 2.8), but, overall, it yields more meaningful correspondences than GMDS.



(a)

(b)



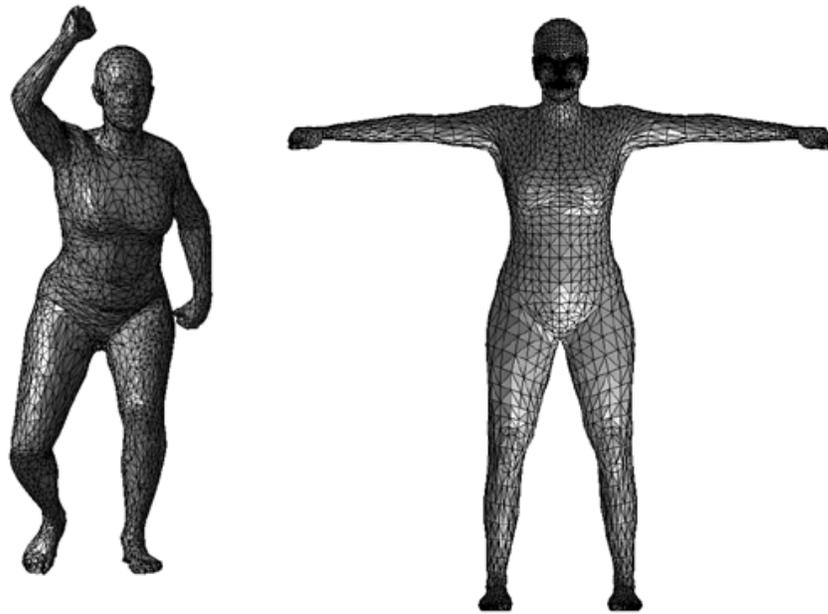
(c)

	TOSCA nonrigid 3D world	SCAPE pose	SCAPE shape	SCAPE shape/pose
T_e	$3 \cdot 10^{-16}$	0.013	$3 \cdot 10^{-4}$	$3.15 \cdot 10^{-4}$

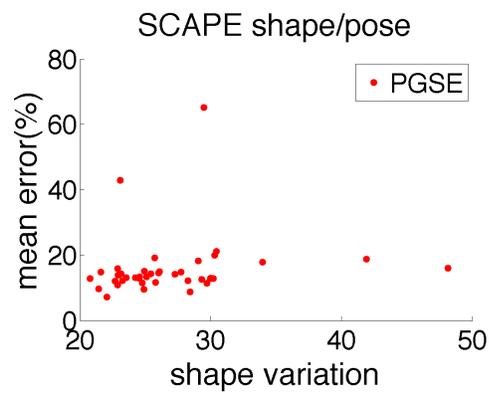
(d)

FIGURE 2.8. Mean Voronoi error plot for correspondences between SCAPE bodies varying in (a) pose, (b) shape, and (c) pose and shape. The data points in figures (b,c) are ordered based on shape variation. Table (d) shows the results of the Wilcoxon signed rank test on the errors induced by the GMDS, PGSE correspondences. GMDS, PGSE are evaluated using data synthesized with the SCAPE model and the TOSCA dataset. All the p -values displayed in the table are below the default significance level of 5%.

3.2.2. *Correspondences in meshes with different topology.* Next we evaluate the effects of changing the global and local resolution of the triangulated meshes used above. We use QSLIM [34] to change the global resolution of the meshes generated based on the SCAPE model and we observe an almost uniform



(a)



(b)

FIGURE 2.9. (a) An example pair of meshes with significant differences in local resolution and mesh topology: a SCAPE body and our template mesh. (b) Mean Voronoi-based error for correspondences between the SCAPE bodies varying in shape & pose and the template. To simplify visualization the SCAPE bodies are ordered only based on shape variation. A Voronoi-based error cannot be defined for the case of GMDS due to markers collapsing at the same vertex.

reduction in resolution across the surface of the SCAPE bodies. In this case, we find no significant difference in performance between GMDS and PGSE as a function of mesh resolution.

Often one wants to align an artist-generated template mesh with higher-resolution meshes created by a laser scanner or other structured light system. In this case the meshes have very different topology and resolution. Consequently we find correspondences between the SCAPE bodies varying in shape and pose as above and a custom made template mesh shown as the right mesh in Figure 2.9 (a). This template mesh exhibits significant differences in local resolution and topology compared with the SCAPE bodies. We are unable to quantitatively evaluate GMDS because in most cases the markers collapse to the same vertex on the data mesh surface resulting in Voronoi regions with zero area. In contrast, we observe that even large differences in local resolution between the surface of the data and model meshes does not influence the performance of our algorithm (the error in Figure 2.9 (b) is similar to the error in Figure 2.8 (c)).

4. Conclusions

We presented a method that finds correspondences between non-rigid articulated objects varying in pose, shape, and global or local resolution. Our method preserves pairwise normalized geodesic distances between a pair of objects as well as local surface properties also based on geodesic distances. We showed improved correspondence over previous work on widely varying mesh models. Additionally using the SCAPE model we were able to separately evaluate accuracy as a function of pose, shape, and resolution variation. We also defined a Voronoi-based error measure that better measures correspondences that are intuitively “good.” Future work involves making our method robust to noisy surfaces as well as surfaces with missing information. Learning the parameters of our CRF model from training data is another direction for future work.

Constraint-based Human Body Shape Analysis

Modeling the human body can be useful in many domains ranging from fashion to health. Realistic human body models aid garment design and facilitate garment visualization for personalized on-line shopping. Virtual 3D human characters are central in the media industry, such as in the production of movies or computer games as well as in virtual environments. Digital human bodies are used in ergonomics to design both equipment and work environments to meet the needs of human operators. Correlating variation in human body shape to certain diseases might provide new methods for disease diagnosis and prevention. However, there is neither a definition of human body shape nor a 3D human body shape representation that is widely accepted.

Historically, body shape is measured with a tape measure and calipers [35]. Today 3D scanners are increasingly used to capture detailed point clouds or meshes; analyzing such meshes however is still limited. Some methods focus on detecting landmarks on the scans and then simulating a tailor's measurement process using geodesic distances on the surface. Several recent approaches take the scans, perform Principal Component Analysis (PCA) and then relate the principal components to quantities of interest.

Yet still, "body shape" is not well defined. There are many causes of body shape variation such as age, weight, or fitness. Some variance may be related to body pose or posture. Still others are transitory such as changes due to respiration. Standard anthropometric (or tailoring) measurements are well established, but provide a limited view of body shape (Figure 5.20). It would be useful to know what variations in shape are captured by such measurements and what variations are not.

We put forward a principled formulation of this question as an objective function (or a family of related functions). Our constraint-based optimization framework for modeling 3D human body shape factors shape into a set of orthogonal shape subspaces related to different constraints (Figure 5.20). Our goal is to capture overall human body shape variation, while at the same time produce shape subspaces that are tied to specific constraints that relate shape to specific applications. Such an approach can produce more compact shape representations, improve performance on regression problems, aid visualization and regularize solutions with limited data. It also provides insight for engineers or designers to visualize what body shapes might be excluded by focusing on a given set of measurements.

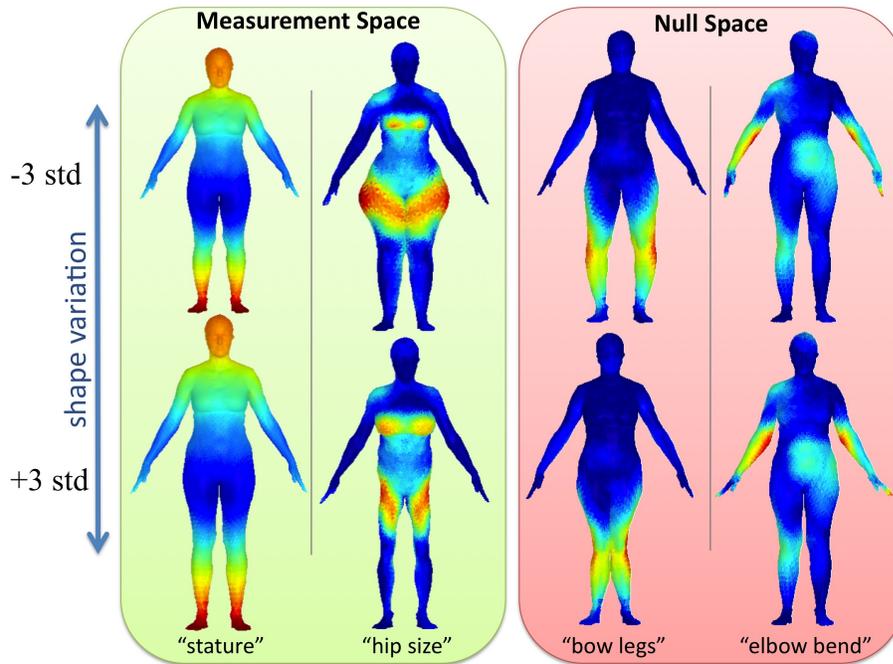


FIGURE 3.1. We factor human body shape variation into a measurement space that is constrained to predict anthropometric measurements like height and hip girth and a *null space* of body shape variation that are unrelated to measurements such as bow-legs and pose/posture variations.

Specifically, we start with the standard formulation of PCA as a least squares optimization problem. We then modify this objective function in several ways: 1) we relate the PCA coefficients linearly to anthropometric measurements such as height, inseam, thigh circumference, etc.; 2) we split body shape into two orthogonal subspaces: one related to measurements and a *null space* of body shape variations that illustrates aspects of body shape that standard tailoring measurements fail to capture; 3) we add sparseness priors on the shape coefficients in both the measurement and null space and find that the resulting model is better able to predict body measurements from 3D scans.

We quantitatively compare our different objective functions with each other and with previous approaches including a standard PCA approach and the method of Hasler et al. [44]. We find that our approach produces more accurate predictions of body shape and results in basis vectors that reveal interesting properties of body shape.

1. Previous work

Three-dimensional shape modeling has a long history which we will not review here. Many recent approaches have focused on developing effective descriptors of 3D shapes for assessing similarity or matching 3D shapes in applications such as content-based 3D shape retrieval. Examples include spin images, spherical

harmonics, mesh HoG, heat kernel signatures; these are summarised in the following overview papers [21, 48, 90]. These generic 3D shape descriptors are useful for classification, but are less useful for detailed analysis of shape within a class. Analysis of 3D human shape is a relatively young but growing field spurred by the accessibility of 3D body scanners. A common way to characterize human body variation in shape is to use dimensionality reduction techniques. We analyze 3D human body shape by extending PCA [51].

Principal component analysis is a well-established technique for dimensionality reduction with various applications including data compression, visualization, and exploratory data analysis. PCA was initially formulated as an unsupervised dimensionality reduction technique. However, there have been numerous extensions of PCA that constrain the principal components to increase their discriminative power or aid understanding of the data. Previous efforts to constrain PCA have focused on two types of constraints: constraints that are related to labels on the data and constraints that are related only to the structure of the resulting principal components. Methods that belong in the first category are supervised dimensionality reduction techniques such as Partial Least Squares (PLS) [99] and Canonical Correlation Analysis (CCA) [42]. Previous work for the second type has used constraints such as sparseness, nonnegativity, priors with different assumptions about the data. Examples of such methods include sparse PCA [110], Sparse Probabilistic Principal Component Analysis [41], nonnegative sparse PCA [105], and constrained subspace modelling [94]. We propose a framework for constrained PCA where both types of constraints can be applied to the principal components. Additionally, we divide the principal components into multiple orthogonal subspaces, each of them capturing a different set of constraints. The formulation in terms of objective functions is similar in spirit to that in [26].

The analysis of 3D body shape can be divided into two paradigms. The first analyzes the statistics of *vertices* [4, 85, 102] (or similarly voxels [12]) while the second analyzes *deformations* of triangles [6]. While the former is far simpler, previous methods have been limited because scans of the human body contain pose variations, posture, breathing, etc. which vary the body shape in ways that are unrelated to things like measurements. The latter approach, particularly [6], makes it possible to factor out some variations, e.g. due to pose, but involves additional complexity (e.g. SCAPE requires a least squares optimization to construct a valid mesh). We work directly in the space of vertices here but get the benefits normally found with deformations. Specifically by constraining shape variations to be related to anthropometric measurements, we find body shape variations that factor out pose. Thus our method extends the value of vertex-based methods for anthropometric applications.

Application of PCA to describe 3D human body shape has shown encouraging results on generating and describing 3D human bodies [6]. Weiss et al. [97] used PCA coefficients to predict anthropometric measurements of 3D human models from Kinect data by combining information from multiple poses. We pose that exploring the structure of 3D human body shapes using dimensionality reduction in conjunction with alternative descriptions of the human body (anthropometry) will give more realistic compact descriptors

of the human body as well as enhance the accuracy of the resulting descriptor with respect to predicting measurements. Early work on this problem [4, 14, 75] performed PCA on body scans and related the PCA coefficients linearly to standard tailoring measurements. More recent work took standard PCA basis vectors and rotated these to better predict body measurements [44]. Alternatively Guan et al. [39] constrained body shape variations related to a particular measurement and excluded these from the orthogonal subspace. Our approach is different in that our objective function formulation provides a simple and intuitive way of exploring body shape by varying the terms in the optimization. It is more expressive in that it can model non-linear constraints on the resulting subspace(s).

2. Methodology

Our data consists of 3D meshes representing human bodies. We assume that these meshes are watertight and their vertices are in correspondence. Let $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_P] \in R^{D \times P}$ represent the meshes in the training set where P is the total number of meshes and D is $3 \times$ the number of vertices in a mesh.

In the case of the standard PCA model, we can represent the training data using a set of K basis vectors where $K \ll P$. In order to do that we want to minimize the following energy function:

$$(5) \quad E_{\text{PCA}}(\boldsymbol{\mu}, \mathbf{B}, \mathbf{C}) = \sum_{p=1}^P \|\mathbf{d}_p - \boldsymbol{\mu} - \mathbf{B}\mathbf{c}_p\|_2^2$$

where $\mathbf{d}_p \in R^D$ represents the coordinates of the vertices of the p -th mesh and $\boldsymbol{\mu} \in R^D$ is the mean over all aligned meshes. $\mathbf{B} \in R^{D \times K}$ is the matrix with the shape principal components and $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_P] \in R^{K \times P}$ is the matrix with the shape coefficients for each body in the training set. To enforce orthonormality of the basis vectors, we optimize subject to the constraint that $\mathbf{B}^T \mathbf{B} = I$.

Measurement-constrained PCA (MPCA). In the first case, our goal is to extract a compact representation of human body shape that is predictive of anthropometric measurements by minimizing the following energy function:

$$(6) \quad E_{\text{MPCA}}(\mathbf{B}, \mathbf{C}, \mathbf{M}) = \sum_{p=1}^P (\|\mathbf{d}_p - \boldsymbol{\mu} - \mathbf{B}\mathbf{c}_p\|_2^2 + \lambda_1 \|\mathbf{z}_p - \mathbf{M}\mathbf{c}_p\|_2^2)$$

where $\mathbf{z}_p \in R^M$ is a vector of M body measurements for the p -th body in the training set and $\mathbf{M} \in R^{M \times P}$ is a linear transformation matrix from shape coefficients to measurements. λ_1 controls the relative weight of the two terms. Note that, again, we optimize subject to the constraint that the \mathbf{B} are orthonormal.

This formulation has the effect of focusing the basis vectors on shape variations that are linearly related to measurements. There is a problem, however, as there is a tension (balanced by λ_1) between faithfully reconstructing body shape data \mathbf{d}_p and fitting the measurements.

Null space of body shape. To address this trade off, we introduce a second set of basis vectors, \mathbf{G} , that represents shape variation that is not captured by anthropometric measurements; this is what we call the null space of body shapes. The energy function we minimize is the following:

$$(7) \quad E_{\text{null}}(\mathbf{B}, \mathbf{C}, \mathbf{M}, \mathbf{G}, \mathbf{V}, \boldsymbol{\mu}) = \sum_{p=1}^P (\| \mathbf{d}_p - \boldsymbol{\mu} - \mathbf{B}\mathbf{c}_p - \mathbf{G}\mathbf{v}_p \|^2 + \lambda_1 \| \mathbf{z}_p - \mathbf{M}\mathbf{c}_p \|_2^2)$$

where we have introduced a new basis, $\mathbf{G} \in R^{D \times G}$, where G is the number of null-space basis vectors used, and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_P] \in R^{G \times P}$ is the matrix of shape coefficients that are not constrained to be related to body measurements. Here the optimization is performed as before but with the additional constraints that $\mathbf{G}^T \mathbf{G} = I$ and $\mathbf{B}^T \mathbf{G} = 0$.

The idea is that \mathbf{G} can capture body shape not captured by \mathbf{B} .

Sparseness. In the third case, we impose sparsity on both the basis vectors related to measurements as well as to the basis vectors in the null space to aid interpretation of the data and increase robustness:

$$(8) \quad E_{\text{sparse}}(\mathbf{B}, \mathbf{C}, \mathbf{M}, \mathbf{G}, \mathbf{V}, \boldsymbol{\mu}) = \sum_{p=1}^P (\| \mathbf{d}_p - \boldsymbol{\mu} - \mathbf{B}\mathbf{c}_p - \mathbf{G}\mathbf{v}_p \|_2^2 + \lambda_1 \| \mathbf{z}_p - \mathbf{M}\mathbf{c}_p \|_2^2) + \lambda_2 \| \mathbf{B} \|_1 + \lambda_3 \| \mathbf{G} \|_1$$

where we add an L1 norm on the elements of the two bases to penalize non-zero elements in the basis vectors. This can be thought of as a regularization while also focusing the basis vectors on what is most important to capture shape and measurements. As above, we maintain orthonormality of the two spaces which are mutually orthogonal; i.e. $\mathbf{B}^T \mathbf{B} = I$, $\mathbf{G}^T \mathbf{G} = I$, $\mathbf{B}^T \mathbf{G} = 0$

3. Experiments

The data we used for training and testing were aligned 3D scans of a subject in a standing pose and anthropometric measurements from the CAESAR dataset [78]. The scans were registered using a procedure similar to [45]. Optimization was performed using the Interior Point algorithm in the constrained optimization toolbox of Matlab.

3.1. Compact shape representation for anthropometric measurement prediction. We used the objective functions in Eq. 6, 7, 8 to generate a compact representation of human body shape for accurate prediction of anthropometric measurements. We have estimated basis vectors that are related to 44 measurements of 800 female bodies from the CAESAR dataset and predicted measurements on a different set of 271 female bodies. We projected the test data into the optimized principal components \mathbf{B} and used linear regression to predict measurements from the low-dimensional shape coefficients of the test data. The Root Mean Squared Error between hand measurements (in mm) and predicted measurements for different configurations of our objective function relative to previous work is shown in Figures 3.2, 3.3, 3.4.

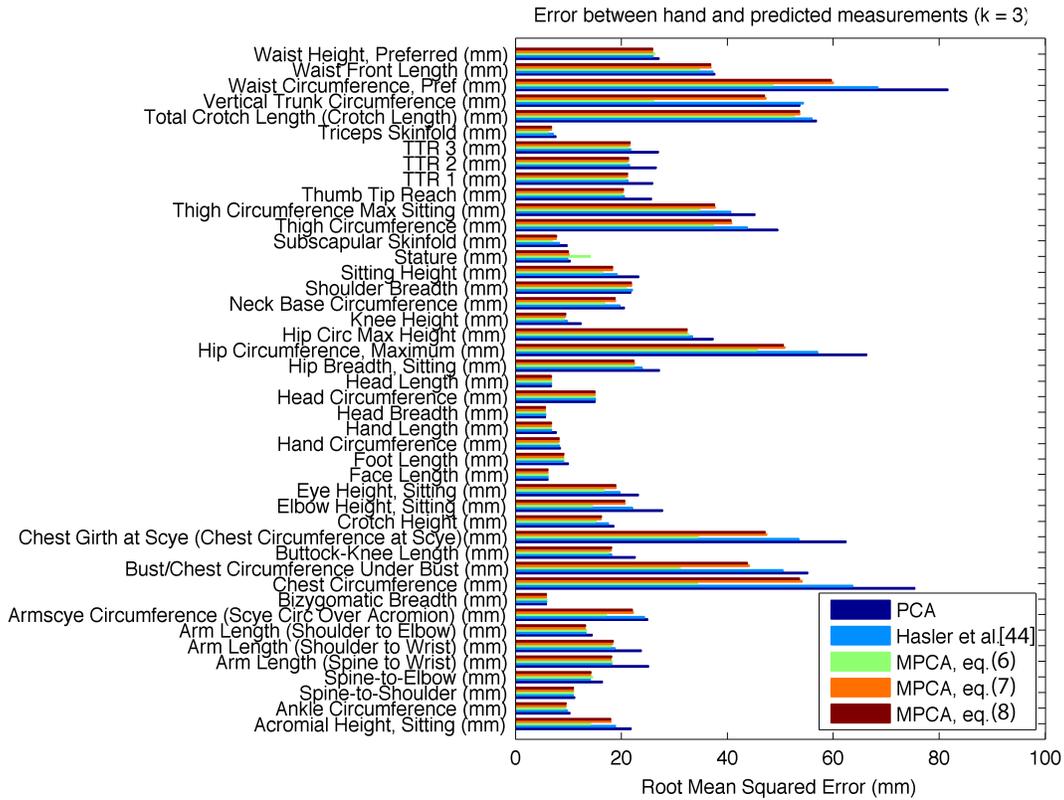
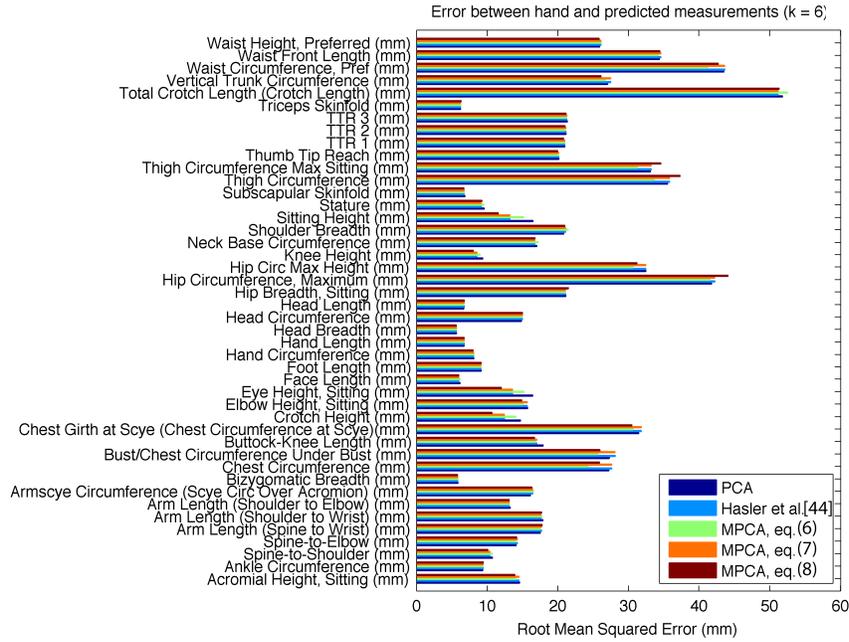
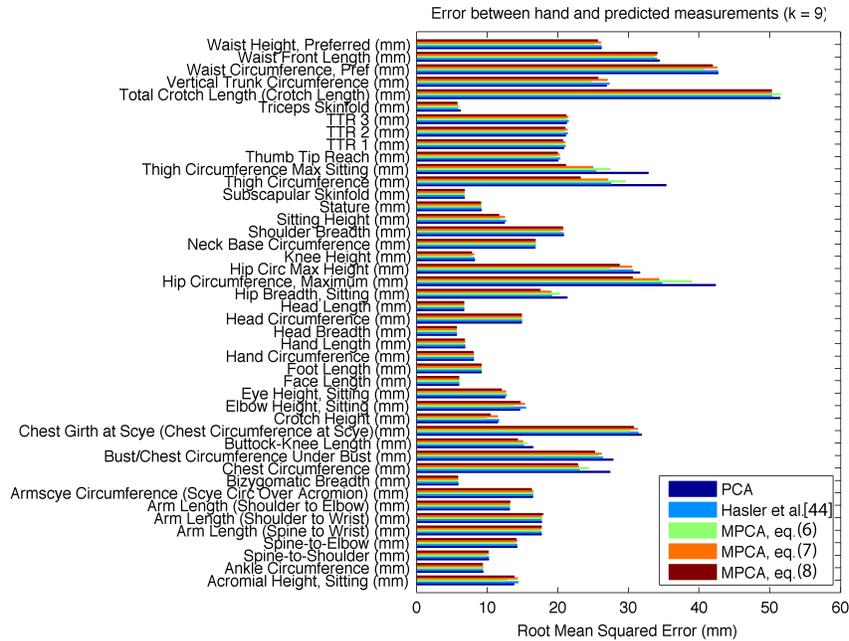


FIGURE 3.2. Root Mean Squared Error (RMSE) on measurement prediction for PCA, Hasler et al. [44] and our method using the objective functions in Equations 6, 7, 8. The error is calculated for $k = 3$ basis vectors.

We can observe that our formulation leads to a more compact representation of human body shape with regards to measurements giving smaller error on predicting measurements. Measurement prediction improves for all the methods as the number of basis vectors considered is increased. In most of the cases, our improved measurement prediction is statistically significant (p -values smaller than 0.05 using the Wilcoxon signed rank test) as shown in Figure 3.4. A comparative visualization of the resulting principal components using PCA, Hasler et al. [44] and the MPCA formulation using Equation 8 is shown in 3.5.



(a)



(b)

FIGURE 3.3. Root Mean Squared Error (RMSE) on measurement prediction for PCA, Hasler et al. [44] and our method using the objective functions in Equations 6, 7, 8. The error is calculated for (a) $k = 6$, and (b) $k = 9$ basis vectors.

	k = 3	k = 6	k = 9
PCA	27.1	19.20	18.79
Hasler et al. [44]	24.09	19.02	18.05
MPCA, eq. 6	19.99	18.78	17.88
MPCA, eq. 7	22.60	19.01	18.65
MPCA, eq. 8	22.53	18.74	17.41

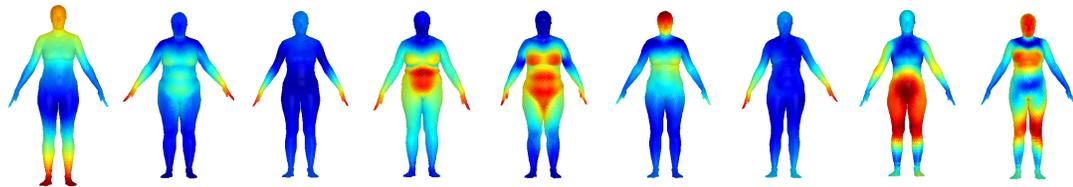
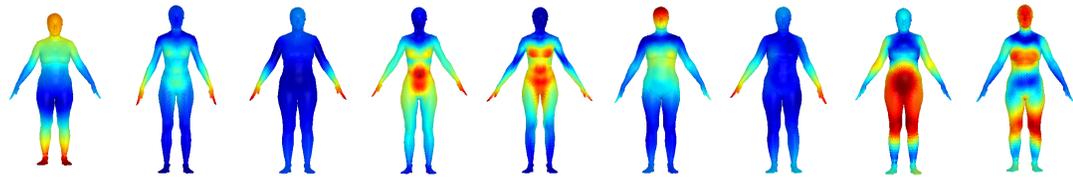
(a) Root Mean Squared Error averaged over all measurements (mm)

	k = 3	k = 6	k = 9
MPCA, eq. 6	$2.9 \cdot 10^{-6}$	10^{-7}	10^{-7}
MPCA, eq. 7	0.17	$8 \cdot 10^{-7}$	$7.5 \cdot 10^{-3}$
MPCA, eq. 8	0.12	0.88	$1.7 \cdot 10^{-6}$

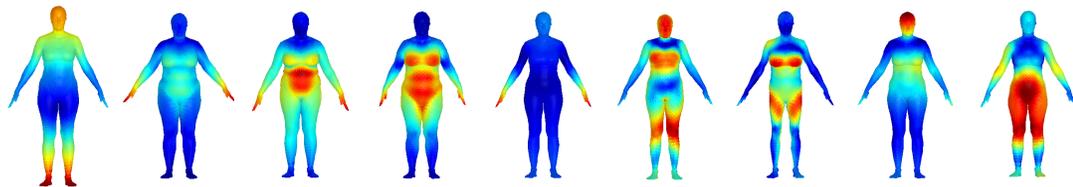
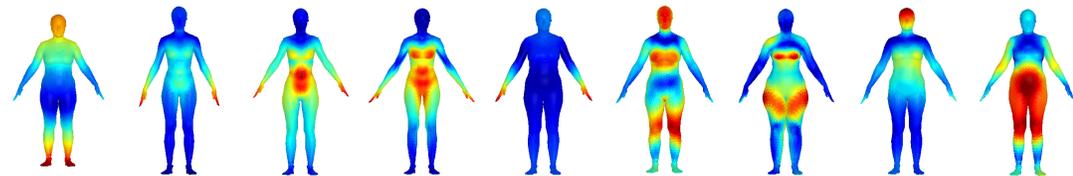
(b) Statistical significance p -values with respect to Hasler et al. [44]

FIGURE 3.4. (a) Root Mean Squared Error (RMSE) for different MPCA formulations relative to previous work. The RMSE is additionally averaged over all measurements in the CAESAR dataset [78]. (b) Statistical significance p -values with respect to Hasler et al. [44]

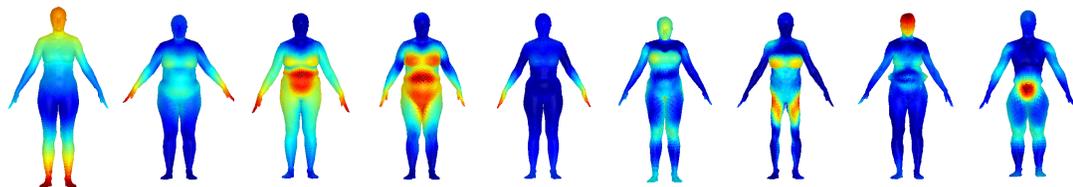
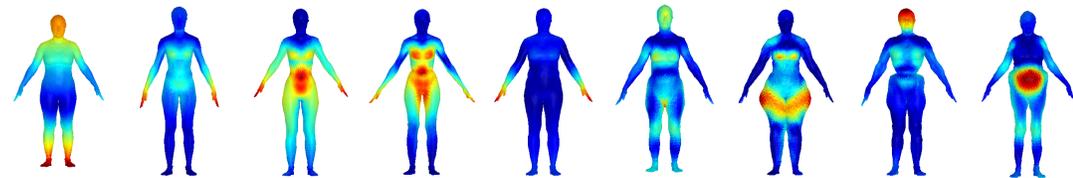
3.2. Qualitative analysis of the nullspace. We used the objective function in Eq. 8 to generate 9 basis vectors \mathbf{B} predictive of anthropometric measurements and 11 shape basis vectors \mathbf{G} , orthogonal to \mathbf{B} , that are not constrained by the anthropometric measurements. The resulting \mathbf{B} and \mathbf{G} basis vectors are shown in Figure 3.6. We observe qualitatively that the subspace \mathbf{G} includes shape variation such as inter- and intra-person posture variation, “bow legs”, body shape asymmetries.



PCA basis vectors (1-9): top/bottom = ± 3 std

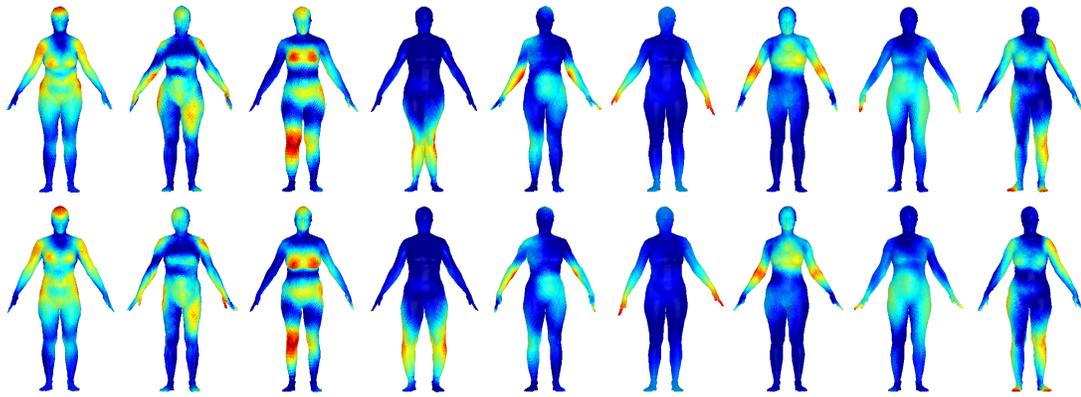


Hasler et al. [44]: top/bottom ± 3 std



MPCA: top/bottom ± 3 std

FIGURE 3.5. Human body shape variation captured by PCA, Hasler et al. [1], MPCA. The i -th column represents the i -th principal component. Principal components are ordered in terms of significance. No ordering for the case of MPCA. Each component has a big effect on vertices in red areas and small effect on vertices in blue areas.



Null space components (MPCA): top/bottom +/- 3 std

FIGURE 3.6. Null space of body shape: The first nine components of human body shape variation not captured by anthropometric measurements. These are orthogonal to the MPCA measurement vectors. Interesting effects: pose variation, body asymmetries, bow-legs, *etc.*

4. Conclusions

We have shown that simple modifications of the PCA objective function can lead to new ways of analyzing 3D body shape that 1) improve the prediction of anthropometric measurements and 2) shed light on how these traditional measurements constrain body shape and how they do not. This approach is more general than the previous methods such as Hasler et al. [44] and achieves statistically significant improvements in measurement prediction. Since the objective function approach is very flexible, it is straightforward to modify our formulation to add additional constraints, including non-linear constraints relating body shape and measurements. This would likely lead to even larger differences between traditional methods and the MPCA approach. Note that our approach leads to a factoring of body shape into pose-dependent and pose-independent shape variation. If we had access to pose information, we could further add a pose-dependent shape space that could be linearly (or non-linearly) related to pose. While we have applied our method to vertices here, it can also be applied to triangle deformations. While we plan to explore this in future work, one advantage of our method is that we get the benefits of a factored model while working with vertices which are simpler than deformations. Additionally, we plan to use this framework to study other properties of body shape variation due, for example, to respiration.

Anthropometric Measurement Prediction from 3D Human Scans in Multiple Poses

Measuring the human body from 3D data is gaining increasing importance in applications such as virtual try-on and online shopping. Extracting tailoring measurements directly from 3D scans of people could accelerate the tedious and time consuming process of custom tailoring. That, in turn, offers the potential of personalized sizing in online shopping and a decreased return-rate for web purchases. These applications are currently of interest due to the emergence of low-cost scanning devices such as the Microsoft Kinect [55]. The promise of such methods is that they will be even more accurate than humans at measuring the body.

Here we revisit the problem of predicting measurements from 3D scans introduced in Chapter 3. We develop a method for *model-based anthropometry* that accurately predicts measurements; the approach is summarized in Fig. 4.1. Model-based anthropometry has several components: 1) a statistical model of body shape variation across a population; 2) a deformable 3D body model and a method to fit it reliably to a scan; 3) a method to extract a variety of local and global features from the model; and 4) a method to predict 1D measurements from the features. A significant novelty of our approach is that it allows us to integrate information from multiple scans of a person in different poses. Contrary to the method in Chapter 3, here we propose a discriminative framework for measurement prediction. It provides greater flexibility on the shape features that are used and incorporates information from scans in multiple poses. We show that this approach is more accurate than existing methods.

Exactly how accurate are current methods and are they accurate enough for custom tailoring applications? While there have been large studies of the accuracy of anthropometrists [35], there have been no large published studies of automatic methods for deriving measurements from scans of real humans. This is despite the fact that there are relatively large collections of laser body scans with associated hand measurement data [16, 77, 91]. For the first time, we perform an extensive evaluation of existing commercial and research systems using the CAESAR dataset [77].

Previous work on extracting anthropometric measurements from 3D human scans is based on either measuring directly on the raw scan or using a database of registered scans to correlate human shape variation with measurements. In the first case, measurements are extracted by locating anthropometric landmarks on the scan’s surface; this simulates and automates the process of acquiring measurements as performed by

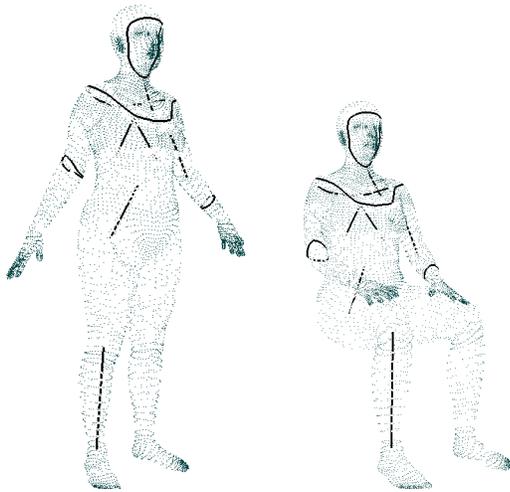


FIGURE 4.1. Registered meshes of a person in two poses after fitting a human body model [6, 45]. Measurement prediction is based on local and global shape features. Local shape features comprise circumferences on the mesh surface (shown as curves) and limb lengths (shown as straight lines). Global shape features (not shown here) describe statistics of shape in a database of registered 3D meshes. For visualization clarity, the 3D meshes are displayed as point clouds.

an anthropometrist. Measurements typically correspond to straight lines (heights) starting and/or ending on landmarks as well as circumferences along planar slices of the scan based on landmark locations. These approaches are sensitive to acquisition noise and missing data on the scan’s surface that distort the shape of the captured body. They also require high-resolution (e.g. laser scans), making such approaches impractical with today’s low-resolution home scanning systems based on Kinect. Additionally, this need for accurate localization of anthropometric landmarks in 3D typically limits such systems to scans captured in a single canonical pose. We show that different poses are optimal for different measurements and our model-based approach is able to integrate information from multiple poses. Finally, existing commercial solutions are limited in the kinds of measurements they make (linear and circumferential), whereas our model-based approach can regress body shape to any measurement (e.g. weight or even age).

Our model-based approach addresses the problems of previous methods. First, given a database of registered 3D scans of humans, together with their measurements, we build a statistical model of shape variation in the population. Features corresponding to 3D shape variation between individuals in the database are correlated with their measurements. Then, given a 3D scan of a new subject, we *register* the model with the scan by deforming the model to match the scan data. Shape features are easily computed from the model and measurements are then predicted from these shape features. A model-based approach could also be used to predict measurements from low-resolution scans [97]. Most previous efforts on correlating

measurements with human shape variation assume the human body shape is represented by a single pose ([97] is an exception in that they use several poses). Moreover, previous evaluations are limited either in the set of measurements considered [44] or the number of subjects used [97, 102].

Our contributions are the following: 1) We introduce a set of surface-based shape features that are predictive of standard anthropometric measurements. 2) We optimize over a wide range of features to find the ones most predictive of measurements. 3) We introduce model-based anthropometry for predicting anthropometric measurements from various poses and demonstrate more accurate measurement prediction than the state-of-the-art. 4) We present a comprehensive comparative study between our model-based approach and state-of-the-art commercial and research efforts for measurement prediction. We consider a wide range of standard anthropometric measurements and a large number of subjects using the CAESAR dataset. This evaluation provides a solid foundation for evaluation of commercial and research work in this area.

1. Related Work

There are several studies comparing the performance of commercial 3D scanning systems relative to measurements obtained using standard manual anthropometry [17, 62, 68]. In particular, [62, 68] provide an evaluation of commercial 3D scanning systems in terms of predicting anthropometric measurements on or around the torso. The measurements were extracted from rigid torso dressforms. Rigid mannequins, however, do not exhibit the things that make real bodies a challenge to measure: pose/posture variation, breathing, soft tissue deformation, body fat obscuring anatomical structures, and general ambiguity about where to measure.

Rather than evaluate automated systems, previous work has focused on the accuracy of humans at measuring the body. Standard anthropometry remains the gold standard for measurement but there is variability between measurers and by the same measurer over time. The ANSUR study was designed to measure the accuracy of such human measurements [35]. Without ground truth, ANSUR focused on the variability of measurement and used this variability to define a allowed error for each measurement.

A main study of automated scan measurement analysis is that of Bradtmiller and Gross [17]. Compared to the abovementioned studies, they report a broader set of automatically extracted measurements from real human subjects. They found that these measurements were generally sufficient for garment fitting, but the prediction error was larger than the ANSUR allowable error.

To evaluate measurement prediction, we use the CAESAR dataset [77], which represents the largest publicly available dataset of 3D body scans with associated measurements; here we use the US dataset of approximately 1000 men and 1000 women in both seated and standing poses. While in wide use, to our knowledge the accuracy of measurement prediction from CAESAR scans has not been evaluated, and nobody has attempted to predict measurements from seated poses. Robinette and Daanen [79] measured the variance

of estimates extracted semi-automatically from CAESAR but did not evaluate the error with respect to manual measurements.

The model-based anthropometry approach we introduce relies on human shape features extracted from a database of registered 3D human bodies. There are a wide range of 3D shape descriptors for assessing similarity or matching 3D shapes in applications such as content-based 3D shape retrieval. Examples include spin images, spherical harmonics, mesh HoG, heat kernel signatures; these are summarized in the following overview papers [21, 48, 90]. These generic 3D shape descriptors are useful for classification but are less useful for detailed analysis of shape within a class. Given objects of a specific category, a common way to characterize variation in shape is using dimensionality reduction techniques. Employing Principal Component Analysis (PCA) [51] to describe 3D human body shape has shown encouraging results on generating and describing 3D human bodies [6]. Weiss et al. [97] use PCA coefficients to predict anthropometric measurements of 3D human models from Kinect data by combining information from multiple poses. More recent work takes standard PCA basis vectors and rotates these to better predict body measurements [44]. Alternatively Guan et al. [39] constrain body shape variations related to a particular measurement and exclude these from the orthogonal subspace. The work in [9] correlates body shape variation with measurements by performing PCA on the joint space of bodies and measurements. We derive global features of shape variation using PCA and augment them with features such as circumferences around limbs and limb lengths to represent local shape details.

Although we are interested in extracting measurements from 3D scans of humans, there has also been work on measuring synthetically generated 3D human bodies. Wuhrer et al. [102] present a hybrid of the above mentioned paradigms by measuring, in a consistent way, paths along the surface of 3D human meshes with the same topology. However, the goal of the authors was mainly to synthesize 3D human bodies that conform to a set of input measurements. In addition, [24] reports measurement prediction on 3D bodies generated from photos of real humans and measurements on these photos; they evaluate, however, a very limited set of four measurements.

2. Measurements from 3D scans

Our approach for measurement prediction consists of a training and a testing stage. In the training stage, we register a set of high-resolution 3D human scans to a common 3D template mesh, learn a statistical model of shape deformations in the training set, extract shape features for each registered scan, and learn the optimal shape features for measurement prediction. Figure 4.2 provides an illustration of the training stage. In the testing stage, given a new high-resolution 3D scan, we register it with a common template mesh using the learned statistical shape model, derive shape features and use them for predicting standard anthropometric measurements.

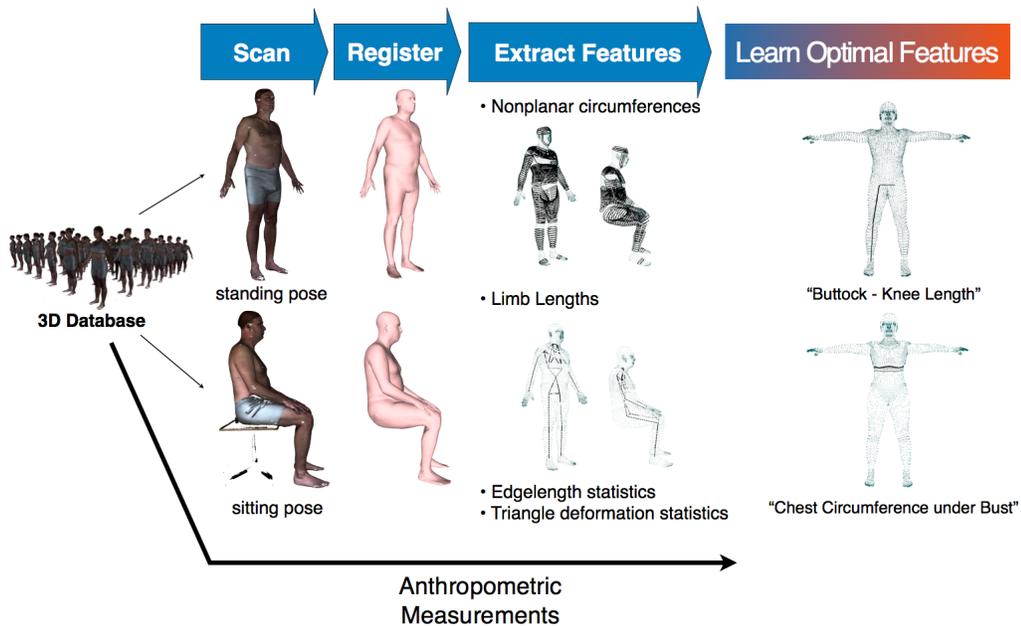


FIGURE 4.2. The training stage of our method. We start with a database of 3D scans in multiple poses (standing, seated) with corresponding anthropometric measurements [77]. Initially, we register the scans using prior knowledge about human body shape. Then, we extract shape features. We consider local features, such as body circumferences and limb lengths, as well as global features, such as statistics on edge lengths and triangle deformations of the registered meshes. Finally, we learn optimal features for predicting each measurement.

2.1. Registration. Registration refers to the fitting of a template body mesh to a scan. This brings the scan into alignment with a database of pre-aligned meshes and our statistical body shape model. We only briefly summarize the mesh registration process as it is not the main contribution and has been described elsewhere [45]. The registration procedure serves two goals: (a) it provides a hole-free mesh that accurately captures the shape in the scan; and (b) it provides correspondences between 3D meshes, which facilitates statistical analysis.

Figure 4.3 shows an example of holes in the armpit area of a 3D scan. In this scenario, a generic surface reconstruction approach, such as [52], creates an unrealistic human shape by lowering the height of the reconstructed armpit. Our approach, which takes into account prior knowledge about the human shape and articulation gets much closer to the true shape of the scanned human subject.

Our registration energy and procedures are similar to those in [45]; as in that work, a BlendSCAPE body model is used (whose form was heavily influenced by SCAPE [6]). As in [45], the data term encourages the template to match the scan and the prior term encourages deformations that are consistent with a learned

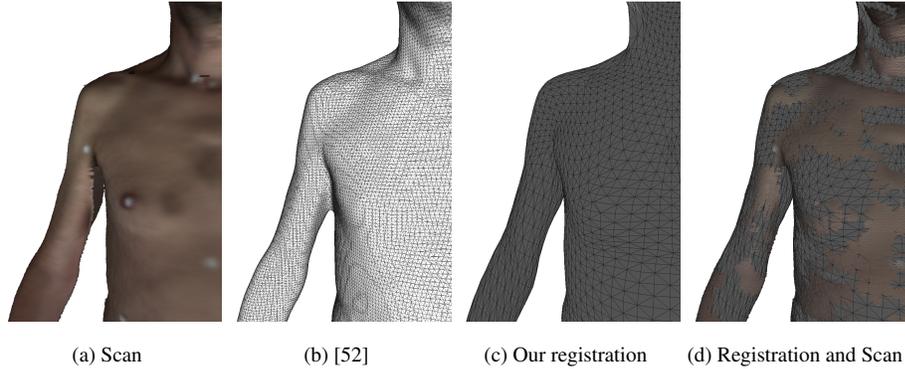


FIGURE 4.3. Model-based alignment effectively deals with holes. (a) A 3D scan showing a hole in the armpit area. Generic surface reconstruction approaches, such as (b) [52], create “webbing” effects and unrealistic human shapes. Our model-based approach (c) is robust to holes and captures the body shape. Subfigure (d) shows an overlay of our registration outcome on the initial 3D scan. Scan and registered mesh interleave with high frequency which indicates that the two surfaces are very close to each other.

statistical body shape model. The shape space of our model was trained from approximately 800 aligned CAESAR scans. While the registration process used 73 landmarks (part of CAESAR) for initialization, we observed that a Gaussian prior on body pose parameters worked equally well.

While not necessary, here we assume the subject is scanned in both standing and seated poses, which means that the registration procedure produces two registered meshes per subject. Registered meshes across subjects are in correspondence by construction.

2.2. Feature Extraction. We extract global and local features of shape variation from a set of registered 3D scans. Let $M_{ia} = (V_{ia}, E_{ia})$, $i = 1, \dots, N$, $a = 1, 2$ denote the registered scan of the i -th human subject in pose a where N is the total number of human subjects in our dataset. $a = 1$ corresponds to the standing pose and $a = 2$ to the seated pose. Each registered 3D scan is represented as a mesh with vertices V_{ia} and edges E_{ia} .

Triangle deformations from a template mesh provide a common representation of 3D shape [6, 88]. Previous authors have shown the triangle deformations carry information about measurements [44, 97]. In particular, the coefficients of a low-dimensional PCA representation can be used to predict linear measurements. An illustration of the first 3 PCA principal components related to shape in the standing pose expressed using triangle deformations is shown in Figure 4.4. Here we go further and consider deformations from three scenarios: 1) from a standing pose only; 2) from a seated pose only; and 3) using deformations from both poses.

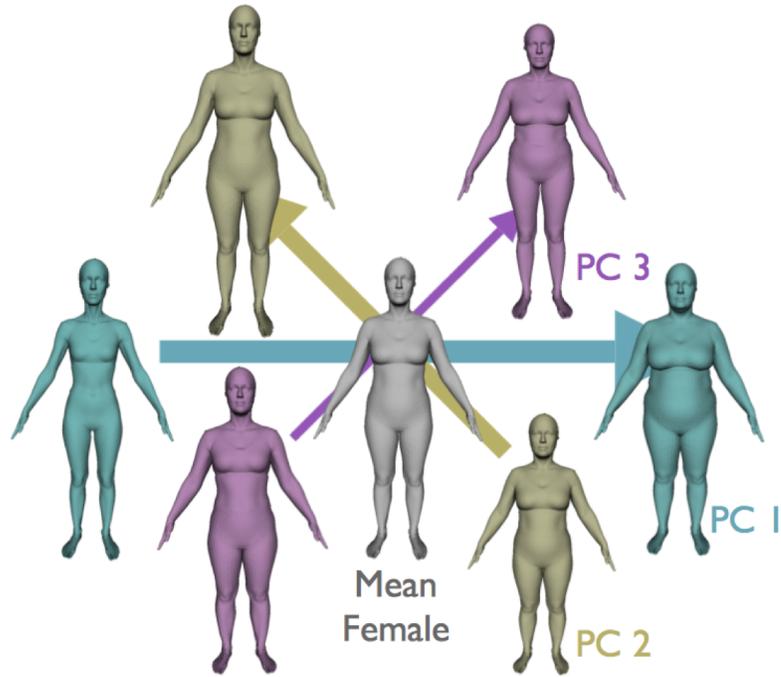


FIGURE 4.4. Low-dimensional PCA representation of human shape. For training, we used around 2000 bodies in the standing pose from the CAESAR dataset [77] represented as a set of triangle deformations with respect to a reference mesh. As show in the photo, the first three principal components are roughly related to body mass, height, and waist shape respectively. © Eric Rachlin.

Triangle deformations, however, are non-linearly related to geodesic distances on the body, and geodesics are similar to many standard tailoring measurements. Consequently we also consider the length of mesh edges, which are directly related to 3D lengths, as the foundation for additional shape features.

Given triangle deformations and edge lengths of template meshes registered with the training dataset, we compute low dimensional representations for each using PCA. Given a new registered test mesh, it is projected onto these low-D spaces, yielding a set of coefficients that characterize the shape; in both bases we use 300 principal components. Let \mathbf{t}_{ia} , $a = 1, \dots, 3$ and \mathbf{d}_{ia} , $a = 1, 2$ denote respectively the coefficients of the triangle deformations and the edge lengths for the meshes of the i -th test mesh across different poses. Here we have extended our notation of a to account for standing and seated poses together ($a = 3$). Considering both poses together is possible because the triangle deformations are taken with respect to the intrinsic shape of the human subject. According to [6], the intrinsic shape is abstracted from effects due to pose, such as muscle bulging, etc.

Low-dimensional representations of the body necessarily remove fine shape details. Additionally these global linear shape models capture correlations in the population and an individual may differ from this.

We address this by adding extra local features that we observe on the surface of the registered bodies. We hypothesize that these features may be more directly related to tailoring measurements. More specifically, we consider circumferences around limbs and the trunk as well as limb lengths. A circumference feature is calculated as a piecewise linear path over the edges of the mesh. Limb lengths are defined as Euclidean distances between neighboring joint locations, where joint locations are defined as a linear combination of pre-defined mesh vertices. Let c_{ia}, l_{ia} respectively be the features corresponding to circumferences and limb lengths associated with the i -th mesh in pose a , $a = 1, 2$.

Summarizing, we consider nine types of features, global and local, per body with $\mathbf{o}_i = \{\mathbf{t}_{i1}, \mathbf{t}_{i2}, \mathbf{t}_{i3}, \mathbf{d}_{i1}, \mathbf{d}_{i2}, \mathbf{c}_{i1}, \mathbf{c}_{i2}, \mathbf{l}_{i1}, \mathbf{l}_{i2}\}$ being the features of the i -th body in the dataset.

2.3. Feature Type Selection per Measurement. To find the most predictive feature types for each measurement, we learn the relationship between shape features and each measurement using Elastic Net linear regression [109]. For computational efficiency, we examine only unary and pairwise combinations of feature types.

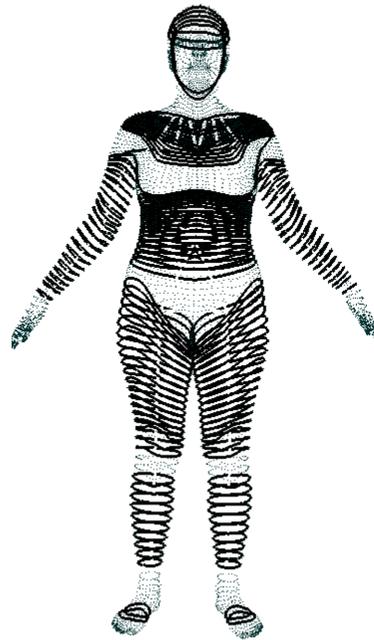
Let \mathcal{S} denote the set of unary and pairwise combinations of the feature types described above. Let also $\mathbf{z} = \{z_{qi}\}$, $q = 1, \dots, Q$, $i = 1, \dots, N$ be the set of anthropometric measurements for all subjects in the dataset where Q is the total number of measurements. We select the optimal combination of feature types \mathbf{s}^q for the q -th measurement as

$$(9) \quad \mathbf{s}^q = \arg \min_{\mathbf{s} \in \mathcal{S}} \sum_{i=1}^N |f_{\mathbf{s}}(\mathbf{o}_i^{\mathbf{s}}) - z_{qi}|$$

where $f_{\mathbf{s}}$ is an Elastic Net regression function for predicting measurements trained on the subset of features \mathbf{s} . $\mathbf{o}_i^{\mathbf{s}}$ denotes the feature values of the i -th human subject from the subset of features \mathbf{s} .

2.4. Measurement Prediction. Measurement prediction is performed independently for each measurement. Let $\mathbf{o}^{\mathbf{s}^q}$ be the values of the optimal features for predicting the q -th measurement (Sec. 2.3) extracted after registering the input scan. The predicted measurements for the human subject is the set $\{f_{\mathbf{s}^q}(\mathbf{o}^{\mathbf{s}^q})\}, q = 1, \dots, Q$.

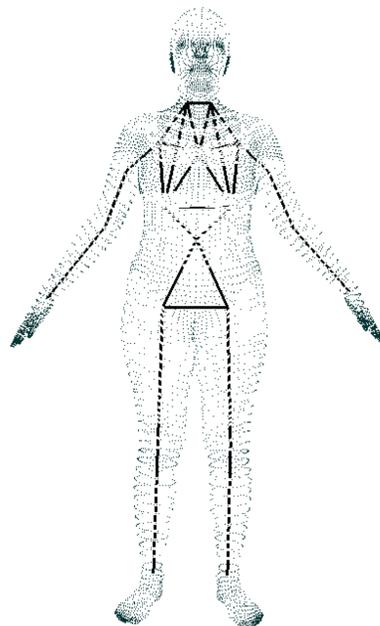
An alternative approach for predicting measurements would be to predefine manually (i.e. by an anthropometer) curves or lines on the reference template mesh for each measurement. Measuring their length directly after the registration procedure would yield an estimate of the corresponding measurement. However, given a reference mesh with a predefined topology, it is not certain that there is a sequence of edges or vertices that match exactly the measurement we are interested in. We believe that using a learning framework that correlates shape features with measurements will introduce to some degree invariance to the topology of the registered meshes. As a by-product, we are able to predict attributes of human shape, such as weight, that cannot be measured directly from a 3D scan, but could potentially be useful in virtual try-on applications.



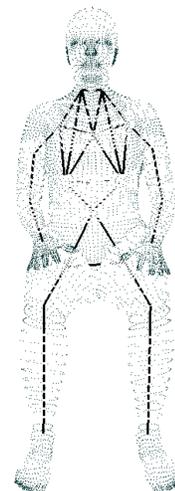
(a) Circumferences, standing pose



(b) Circumferences, seated pose

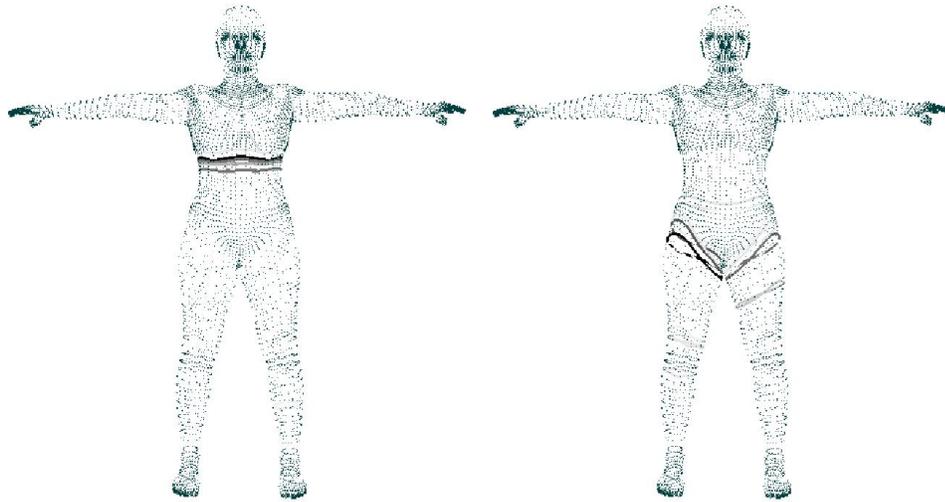


(c) Limb lengths, standing pose



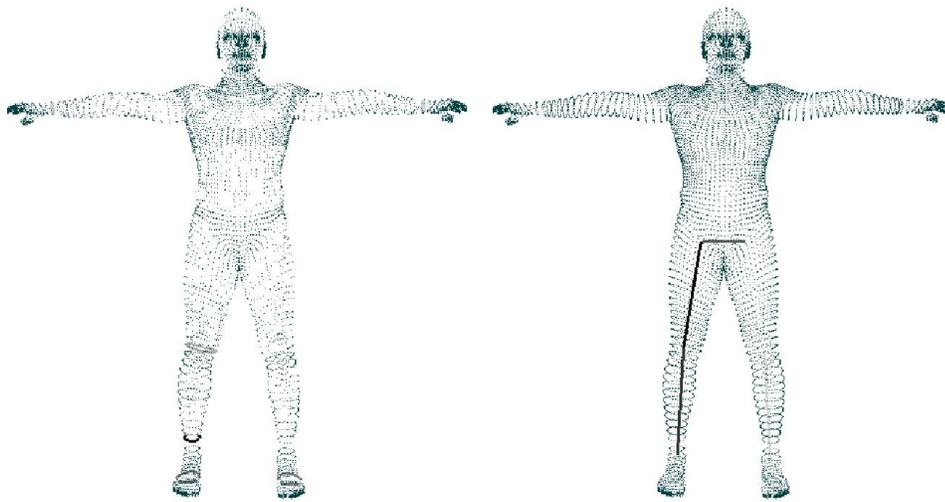
(d) Limb lengths, seated pose

FIGURE 4.5. Local features. (a, b) Circumferences in the standing and seated pose. (c, d) Limb lengths in the standing and seated pose.



(a) Chest Circumference under Bust

(b) Hip Circumference, Maximum



(c) Ankle Circumference

(d) Buttock Knee Length

FIGURE 4.6. Examples of feature selection in females and males for the following measurements: (a) Chest Circumference under Bust, (b) Hip Circumference, Maximum, (c) Ankle Circumference, (d) Buttock Knee Length. The higher the importance of each feature, the darker its color. Features are displayed on the average female or male shape from our database posed in the T-pose. Optimal local features usually correspond to circumferences near the desired circumference measurement or to a set of limb lengths related with the desired height measurement.

3. Results

3.1. Method Evaluation. We compare our method with the commercial software Anthroscan [3] as well as with [44]. Anthroscan predicts measurements directly from a 3D scan in the standing pose (only). Hasler et al. [44] performs registration-based measurement prediction. The data we use for training and testing are registered high-resolution 3D scans and the 40 anthropometric measurements from the CAESAR dataset. Example measurements in CAESAR are shown in Figure 4.7. We use approximately 800 subjects per gender for training and 200 for testing. The error metrics that we use are the Mean Absolute Difference (MAD) between extracted measurements and direct measurements in CAESAR, e_{MAD}^q , for each measurement q as well as the Average Mean Absolute Difference (AMAD), e_{AMAD} , over all measurements:

$$(10) \quad e_{\text{MAD}}^q = \frac{1}{N_s} \sum_{i=1}^{N_s} |f_{s^q}(\mathbf{o}_i^s) - z_{qi}|$$

$$(11) \quad e_{\text{AMAD}} = \frac{1}{Q} \sum_{i=1}^Q e_{\text{MAD}}^q$$

where N_s is the number of subjects in the test set.

Optimal groups of features per measurement are derived using 20-fold cross validation in the training set of 800 subjects. For the 11 common anthropometric measurements between Anthroscan and CAESAR, we compute a linear correction for each Anthroscan measurement using leave-one-out cross validation. This effectively adapts the Anthroscan measurements to the CAESAR measurements. We compare with [44] by using their proposed features and Elastic Net regression. Elastic Net regression compares favorably to the linear prediction framework presented in [44]. We additionally compare our performance with the ANSUR inter-observer error [35]. Recall that ANSUR reports the median absolute deviation between measurements made by experts rather than measurement error from the survey data.

Figure 4.8a shows aggregate statistics of the performance of our method relative to [3] and [44]. More specifically, we report the e_{AMAD} both in terms of absolute values (millimeters) as well as expressed relative to the ANSUR allowable error (AE). Our e_{AMAD} is around 1 cm, which translates to 1.2 to 1.3 times the AE. The error using the features from [44] is around 10% higher than our error. In 50% to 65% of the cases our improved prediction accuracy is also statistically significant. Statistical significance was assessed per measurement using a paired t-test. Comparing our method with the Anthroscan software for the measurements that are common between CAESAR and Anthroscan, we see that the e_{AMAD} error for Anthroscan is 10-15% higher than the error in our case. For the measurements that are common between CAESAR and Anthroscan we come very close to the allowable error. In 65% to 80% of the cases our improved prediction accuracy is also statistically significant.



(a) Hip Circumference, Maximum



(b) Ankle Circumference



(c) Waist Front Length



(d) Shoulder-to-elbow length



(e) Height



(f) Bizygomatic Breadth

FIGURE 4.7. Example anthropometric measurements in the CAESAR dataset reproduced from [77]. We could roughly categorize them into *circumferences*, such as (a) maximum hip circumference and (b) ankle circumference, *surface lengths* on the body, such as (c) waist front length and (d) shoulder-to-elbow length, *Euclidean lengths*, such as (e) height and (f) bizygomatic breadth.

Figure 4.8b shows the performance of the above mentioned approaches for the subset of CAESAR measurements that is common between CAESAR and Anthroscan for the female test set. A detailed overview of performance for females and males is shown in Appendix A. We observe that for most of the measurements,

	Standing	Seated	Stand+Sit
Male	10.21 mm (1.3 x AE)	10.8 mm (1.38 x AE)	10.09 mm (1.28 x AE)
Female	10.24 mm (1.27 x AE)	11.44 mm (1.37 x AE)	10.02 mm (1.23 x AE)

TABLE 1. Effect of **pose** on measurement prediction. We report the Average Mean Absolute Difference (AMAD) error over the 40 anthropometric measurements in CAESAR [77]. AE denotes the allowable error based on ANSUR [35]. The error using the standing pose is lower than the error using the seated pose. Best results are obtained combining both poses.

our approach performs better than previous work and the errors are close to the allowable error for each measurement. Most of the differences between our measurement predictions and predictions from previous work are statistically significant. Measurement predictions with no statistical significance between our method and at least one of the other approaches are denoted with small font size in Figure 4.8b.

3.2. Prediction from Multiple Poses. Most of previous work has focused on predicting measurements from a single standing pose. However, it is unclear which scanning pose is the optimal one or whether different poses would give different prediction results. Muscles bulge, soft tissue deforms, and joints vary in different ways depending on the pose, which is why anthropometric or tailoring measurements are typically acquired using multiple poses. Table 1 shows a comparison of measurement predictions from two poses that were available in CAESAR. The AMAD error over the 40 anthropometric measurements that come with CAESAR is around 1 cm for each single pose which translates to around 1.3 or 1.4 times the variance of the measurements that expert measurers would report based on ANSUR. The measurement prediction error using the standing pose is lower than the error using the seated pose; this is expected given that most of the CAESAR measurements were taken in the standing pose. Combining the two poses by selecting automatically the pose that gives the best prediction, using cross-validation on the MAD error, results in lower measurement prediction error. This confirms our hypothesis that combining information from multiple poses is beneficial.

3.3. Features for Measurement Prediction. In Sec. 2.2 we proposed a set of global and local features for measurement prediction. On one hand, we proposed generative global features, coefficients of PCA components that can be used to generate new bodies. On the other hand, we proposed local features that resemble

	[44]		Our Method			[3]		Our Method	
Male	11.11 mm (1.41 x AE)	0%	10.09 mm (1.28 x AE)	51%	Male	12.65 mm (1.51 x AE)	9%	10.78 mm (1.15 x AE)	64%
Female	11.25 mm (1.42 x AE)	0%	10.02 mm (1.23 x AE)	65%	Female	11.11 mm (1.24 x AE)	9%	10.28 mm (1.06 x AE)	82%

(a) Average Mean Absolute Difference (AMAD) Error

Measurement	[3]	[44]	Our method			AE [35]
Ankle Circumference	7.55	6.59	6.19	deformations, stand	edges, stand	4
Arm Length Shoulder - Elbow	11.26	8.42	6.65	limbs, stand	edges, stand	6
Arm Length Shoulder - Wrist	11.67	10.42	10.05	limbs, stand	edges, stand	
Arm Length Spine - Wrist	13.19	13.40	11.87	girths, stand	limbs, stand	
Chest Circumference	12.43	13.02	12.73	girths, stand	edges, stand	15
Crotch Height	7.45	7.53	5.50	limbs, stand	deformations, stand	10
Head Circumference	7.44	7.45	5.87	girths, sit	limbs, stand	5
Hip Circ Max Height	17.05	18.96	18.59	girths, stand	limbs, stand	
Hip Circumference, Maximum	7.47	16.15	12.35	girths, stand	edges, stand	12
Neck Base Circumference	21.13	16.96	15.79	limbs, sit		11
Stature	5.60	10.21	7.51	girths, stand	limbs, stand	10

(b) Mean Absolute Difference (MAD) Error per Measurement in Females

FIGURE 4.8. Comparative evaluation. (a-left) Comparison with [44] on 40 CAESAR measurements for 200 test of each gender. (a-right) Comparison with Anthroscan on the measurements that Anthroscan and CAESAR have in common. In both cases we report Average Mean Absolute Difference (AMAD) error between extracted and direct measurements. AE denotes the allowable error based on the ANSUR study [35]. Percentages represent the fraction of measurements where the predictions of the best performing method are statistically significant. (b) Mean Absolute Difference (MAD) between extracted and direct measurements on females for the common measurements between CAESAR and Anthroscan (in mm). Bold is best. Results that are not statistically significant are denoted with smaller font size. For our method, we additionally show the optimal types of features (deformations, edges, girths, limbs) and pose (stand, sit) selected for each measurement.

	Global	Local	Global+Local
Male	10.29 mm (1.3 x AE)	11.47 mm (1.46 x AE)	10.09 mm (1.28 x AE)
Female	10.34 mm (1.27 x AE)	10.98 mm (1.40 x AE)	10.02 mm (1.23 x AE)

TABLE 2. Effect of our **global and local features** on measurement prediction. We report the Average Mean Absolute Difference (AMAD) error over the 40 anthropometric measurements in CAESAR [77]. The error using only local features is almost 10% higher than using global features. AE denotes the allowable error based on ANSUR [35].

measurements. Do we really need both and which set of features is more influential? To answer these questions we predicted measurements only from global and only from local features. The results are summarized in Table 2. The AMAD error using only local features is almost 10% higher than using global features. This is interesting because it shows that global shape features are good for predicting linear measurements. But, as with multiple poses, we find that the combination of global and local features yields the best results.

Figure 4.6 shows the most influential local features for a representative sample of measurements. It is interesting to see that the most influential local features for each measurement are in areas of the body close to where a tailor would choose to take the specific measurement. Importantly, these features are automatically discovered. In Fig. 4.8b, we see the features selected by our approach for a subset of measurements. More detailed results are presented in the supplemental material. For most of the measurements, optimal measurement prediction is achieved through a combination of global and local features.

4. Conclusions

In this chapter, we presented an alignment-based approach for extracting anthropometric measurements from high-resolution 3D human body scans. Representing the 3D scan using a reference mesh deformed based on a human body model allowed us to capture effectively the shape of the scanned subject. Additionally, it provided correspondences to a database of 3D humans scans (already in correspondence) varying in intrinsic shape and pose. Shape features derived consistently across the registered scan and all the bodies in the database were fused using a regularized linear learning framework that led to more accurate measurement prediction than state-of-art approaches. Our approach generalizes easily to scans in arbitrary poses (as long as registration with a reference mesh can be performed). We performed an extensive evaluation and found that our method significantly outperforms the state-of-the-art.

This work, for the first time, establishes baseline accuracy on a widely used dataset (CAESAR). This makes it possible for anyone using CAESAR to compare with the results here and this should encourage new approaches. The accuracy numbers reported here are critical for industrial applications in which clothing manufacturers must know how accurate automated methods can be. Our accuracy was significantly better than existing methods, including commercial solutions. The accuracy of current commercial methods has never before been demonstrated. While our errors were above the inter-observer errors in ANSUR, this does not mean they are insufficient for real applications. Inter-observer variance ignores observer bias and does not directly quantify measurement accuracy.

Future work includes experimenting with 3D human scans of lower resolution or partial 3D scans of humans. Because our method can provide correspondences between 3D scans, it facilitates the use of a great variety of 3D shape descriptors and learning methods. With that in mind, we are interested in extending our approach to predict more intrinsic attributes of the human body, such as age, muscularity, *etc.*

Breathing Life into Shape: Capturing, Modeling and Animating 3D Human Breathing

In this work, we describe a method to animate realistic breathing in virtual humans with a simple intuitive interface. Realistic human models and avatars are common in movies and video games. While 3D body scanning technology produces realistic looking 3D body meshes, making them look “alive” requires that they breathe. Moreover, breathing is part of body language and is essential in order to convey specific emotions. Apart from visually pleasing animations in the film or video game industry, realistic animation of breathing is also essential in the medical domain (e.g. for planning radiation therapy). Given the importance of breathing, there are surprisingly few techniques that produce realistic breathing motions, across a range of body shapes, without extensive animation by hand.

Modeling breathing in a realistic, lifelike, way is challenging. First, it entails modeling subtle, yet complex, deformations of the human body that vary across time and context. Second, breathing has a time-varying global effect on the human body; it induces shape change mainly in the torso, but also posture changes over the whole body. Previous work on animating breathing 3D avatars has been either limited in realism or does not generalize easily to new shapes and breathing types [70, 71, 74, 82, 93, 108].

We propose a new approach for modeling body deformations due to breathing using high-resolution 3D human scans, a statistical model of the human body, and physiological parameters related to respiration. An example of animating the breathing of a running character is shown in Figure 1. To capture the subtle and complex deformations of the human body shape due to breathing, we scan 58 human subjects at multiple time instants during their breathing activity. High resolution triangulated meshes are captured using 22 pairs of stereo cameras and a random projected texture pattern together with 22 color cameras and a white-light flash system; this gives high quality 3D meshes with registered texture. To elicit a range of deformations, subjects were asked to breathe naturally, with the chest, and with the stomach. To separate breathing-induced shape change in our data from pose-induced shape change, we register all scans to a statistical model of body shape and pose variation. We compute a mean shape for each subject and the deviations from this due primarily to breathing.

We perform principal component analysis (PCA) on the estimated breathing deformations to produce a low-dimensional model of breathing variation. The PCA space has distinct components for “chest breathing”



FIGURE 5.1. Animating breathing types. Respiration induces changes in torso shape and posture. We learn a model of how 3D breathing deformations relate to lung volume and breathing type and use it to animate bodies of varying shape and pose. Here we show the maximal inhale and exhale shapes overlaid for three different bodies breathing mainly with the stomach (left), mainly with the chest (right), or using a combination of chest and stomach (middle).

and “stomach breathing”. For instance, the shape change during chest breathing is as much up and down as in and out; this is quite different from the shape changes used in simple animations. We found postural changes that were significantly correlated with breathing and that differed between men and women. We also found that the dominant breathing deformations were independent of body shape but that body shape is correlated with fine-scale differences of shape change due to breathing.

To animate breathing we need natural controls that are related to the statistics of pose and shape deformations. To that end, we compute the difference in volume between each 3D body and the mean shape of the subject. We take this change in volume as a proxy for change in lung volume. This allows us to model breathing deformations as a function of volume and to use volume as a simple, and physiologically relevant, control for animation. We also define different *types* of breathing as illustrated in Figure 5.1. For a given breathing type, we find that body shape varies linearly with volume. This linear relationship, however varies non-linearly with breathing type. We learn a novel mathematical model of body shape deformation and pose change as a function of volume and type. We also extend the SCAPE body model [6] to include body shape deformations predicted by our breathing model. These deformations are combined with identity and pose deformations to produce realistic breathing for bodies of any shape and any pose (Figs. 1 and 5.1).

We describe an intuitive interface for creating breathing cycles and for changing breathing types. To more easily capture realistic lung volume sequences for animation we use a device called a spirometer. This

makes it possible to “act out” a particular breathing sequence to correspond with the desired action or emotional content. The recorded changes in lung volume drive the animated character using the learned shape deformation model. This provides an easy and novel way to achieve realistic breathing animation.

While our shape model is built from subjects in a static standing pose, we show that the learned model applies to other poses. We animate a 3D body model and use breath acting to recover the corresponding breathing sequence. The animated sequences with breathing look more natural than sequences without breathing.

1. Related Work

In feature films, characters may have sophisticated controls to hand animate breathing while in lower-quality animations and games, characters may have very simple controls or not breathe at all. In contrast, we learn a detailed and realistic model of how body shape changes with breathing and provide simple controls to make animation easy.

Hand animation. Breathing is a strong indication of life and realistic characters in feature films often have many parameters for hand animation of breathing; the animation is labor intensive. For simpler characters (e.g. in video games) fairly primitive models may be used that capture the gist of breathing through changes in posture (rocking back and forth) or simple cyclic expansion of the chest. Basic breathing controls like these are sometimes used for idle motion generation [30, 31]. In this case breathing is seen as a cause of idling motion, rather than something to model on its own. What is missing is a realistic model of breathing, with simple animation controls, that can be applied to many body shapes in motion.

Anatomy- and physics-based modeling. There is extensive work on anatomy- and physics-based modeling of the human body; see [57, 61] for reviews. For breathing, prior work focuses on modeling the torso [74, 93, 108]. Zordan et al. [108] propose an anatomically motivated model of the human torso that consists of rigid parts (bones) and deformable parts (muscles). Animation requires physical simulation. Veltkamp et al. [93] introduce a similar model that combines better control over abdominal and chest breathing using two independent breathing systems. Lee et al. [58] present a comprehensive biomechanical model of the upper human body with a proof-of-concept demonstration of synthesized breathing motions. In the medical domain, breathing models focus primarily on representing lung shape [22, 65, 84].

Although anatomy- and physics-based body models offer the potential for high detail, they do not generalize easily to new subjects. Synthesizing new human bodies as well as tuning the parameters to generate specific types of breathing is not straightforward. Since these breathing models focus on the torso, they do not model whole-body posture variation during breathing. These issues, combined with the computational expense of physics simulation, mean that such methods are difficult to use in practice. In contrast, our model is learned from data, generalizes easily to new subjects, models whole-body posture variation during breathing,

and provides intuitive controls for synthesizing breathing animations. In previous anatomy-based models, volume change over time is the observed outcome that is used to evaluate whether an animation is realistic. In our case, breathing volume is the input that drives the animation and we can animate arbitrary breathing sequences using volume.

Statistical human body models. Previous work on synthesizing breathing in a data-driven way is limited to replaying recorded breathing motions for 3D shapes similar to the shape of the recorded subjects [71, 82]. However, statistical body models have been used successfully in the past to model the observed body shape across the human population [5, 6, 23, 44]. In these models the observed shape is conceptually decomposed to the intrinsic shape of the subject and deformations that change based on the pose of the subject, such as muscle bulging etc. They do not model breathing deformations or, in fact, other deformations not due to identity or pose. In this work we extend the SCAPE model [6] to include breathing deformations and define controls to easily animate these deformations.

Dynamic shape capture and modeling. The modeling of breathing shape deformations has been limited by a lack of data. High-resolution 3D body scanners typically require several seconds between consecutive scans meaning that fine temporal resolution is lost. Depth maps with high temporal resolution are available from range sensors [72] but these are noisy and have low spatial resolution. Despite progress on tracking complex surfaces such as human clothing in video sequences [87], capturing accurate subtle deformations of the human body remains a challenge. Low spatial resolution and high temporal resolution is available from tracked markers but, with standard marker sets, breathing is not readily visible [70]. Larger marker sets can capture breathing motions of individuals [70] but not populations. To analyze breathing across the population, we need breathing deformations that are in correspondence across people; high-resolution meshes facilitate this. Computer vision methods, with texture painted on the skin provide a possible solution [66] but have not been used to model breathing. In contrast to previous approaches, we acquire a dense reconstruction of the human subject’s shape using a high-resolution 3D scanner. We acquire multiple scans of each subject at unknown time instants in the breathing cycle and register them by taking into account both the geometry and appearance of the 3D scans.

Animation of breathing motions from marker data has been limited to transferring sparse pre-recorded deformations of subjects breathing intensely to subjects of similar shape [70]. Here we go further to learn a model of breathing deformations from examples that can be applied to any body shape and different poses. The model is parameterized by lung volume and breathing type using concepts from the physiology of respiration [63, 98].

Controls for animation. The motion of markers on the chest has been used to drive an anatomy-inspired model [82]. Other controls for breathing animation include audio [25, 27] and parameters related to human physiology. Animation from physiology-related input (including a stretch sensor on the chest,



FIGURE 5.2. Example scans. 58 subjects were scanned in an “A” pose while breathing. Subjects wore tight fitting clothing so that shape changes during breathing were evident. They were asked perform different “types” of breathing: normal, breathing with the chest, and breathing with the stomach. The full dataset consists of 2807 3D meshes with associated texture.

EKG, pulse, skin temperature) has been limited to the anatomically-based models described above and lacks visual realism. Our approach is more similar in concept to [53], where human body surface deformations are correlated with recorded physiology data related to the level of fatigue. In our case, we link the observed surface deformations with the lung volume during breathing. In addition, we animate 3D human characters using spirometer data (lung volume measurements) recorded by “breath actors.”

2. Breath Taking (Data Capture)

To model deformations of the human body due to breathing as realistically as possible we capture high-resolution 3D full-body scans of 58 subjects (28 men and 30 women); Figure 5.2 shows a few representative scans. These scans were captured with a custom multi-camera stereo-based system (3dMD LLC, Atlanta, GA) using flashed texture patterns (for stereo) and white light flashes (for texture capture). Shape capture happens in about 34ms and, since it is flash-based, there is no motion blur. There is a recovery time between captures meaning that we can only capture discrete instants during breathing. Subjects wore minimal tight-fitting clothing (bike-shorts style bottoms for both men and women and a sports-bra style top for women) as shown in Figure 5.2; this made shape changes during respiration readily apparent. To make later registration

of scans with a common template more accurate, some of the subjects were painted in a multi-colored pattern using a water-based paint [15]. Subjects were a mix of professional models (with a modeling contract) and volunteers. Scanning involves standard cameras and lighting, posing no risk to the participants. Before a capture session, each subject gave their informed, written, consent for the analysis and publication of their 3D scan data including images and scans of their faces.

We focus on normal breathing of the upright body in an “A pose” (Figure 5.2); that is, we do not consider different activities or pose-dependent changes in breathing. Subjects were informed that the study was about breathing and were instructed to breathe at what they considered a normal pace. The physiology of respiration [98] leads to two main types of breathing: chest and abdominal breathing which correspond to different motions of the diaphragm. We initially asked the subjects to breathe normally. Then we explicitly asked them to focus on breathing with the chest or with the stomach. Additionally, to be able to represent the extremes of the breathing deformation, we recorded the subject shape during complete inhale and complete exhale. We also captured a small set of scans where the subjects were instructed to breathe strongly/intensely. In total we captured and analyzed 2807 full body scans.

Although breathing is naturally a time evolving process, current high-resolution 3D body scanning systems can give us only sparse samples of this temporal process. Thus, our data consist of static 3D scans that were taken at unknown time instances of the subject’s breathing activity. We address this limitation below.

2.1. Data processing. Our first step is to bring all the 3D scans into correspondence by registering (aligning) them to a 3D body template represented as a triangulated mesh (10,777 vertices, 21,550 triangles) as illustrated in Figure 5.3. The detailed process is described elsewhere [15]; the result is that all 2807 meshes are in correspondence with the template. Shapes are represented as triangle deformations from a template shape. Behind this process is a 3D parametric shape model similar to SCAPE [6] in that it factors body shape changes due to identity from those due to pose. We normalize all registered scans to a common pose and save the pose parameters. For each subject we compute the mean shape and, for each scan, we then compute the residual shape deformation from the mean. This constitutes our shape training data. Additionally we have the pose of each aligned scan and this is used as pose training data.

The shape and pose change during respiration is directly related to the volume of air in the lungs and the motion of the diaphragm. Consequently lung volume and diaphragm motion would provide natural controls for breathing animation. Unfortunately, neither is directly observable from the scans. What is observable, however, is mesh volume, which is easily computed from the aligned meshes using signed volumes of tetrahedra as described in [83]. We assume that mesh volume changes result exclusively from changes in lung volume and consequently take mesh volume (and change in volume) as a proxy for actual lung volume.

According to the physiology of respiration [63, 98], there are two main types of breathing: chest and abdominal breathing. In practice, however, people breathe in a variety of ways with varying amounts of chest

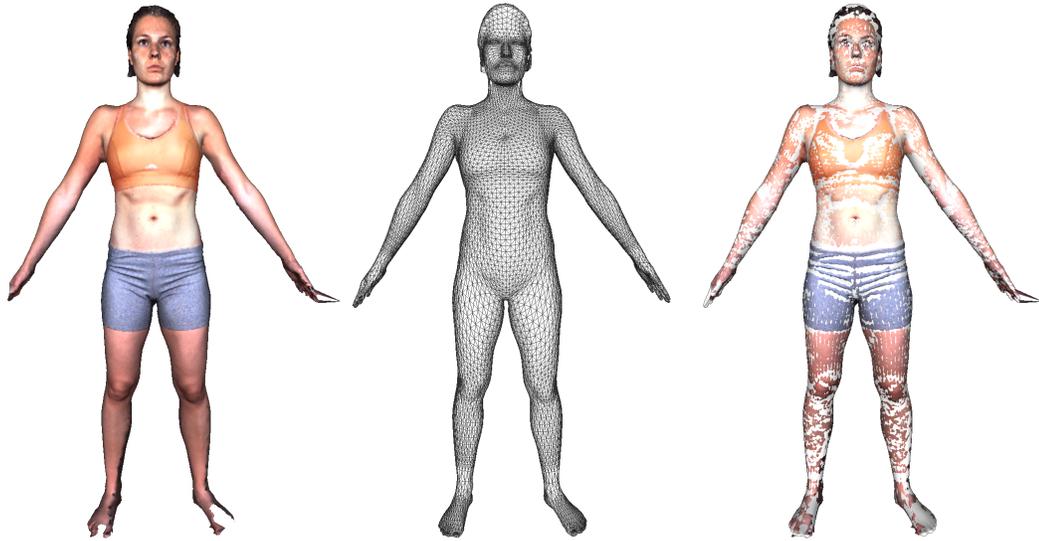


FIGURE 5.3. Example of 3D scan registration. A template mesh is warped to match a high-resolution 3D scan (a). The warped template mesh at the end of the registration procedure is shown in (b). Overlaying the warped template on the scan (c), we see that the two surfaces are very close to each other (they interleave with high frequency). We followed the registration procedure described in [Anonymous 2014]. No landmarks have been used for initialization.

and stomach deformation. While we cannot observe the diaphragm’s motion, we can observe its affect on body shape. To define the type of breathing we segment the torso into an upper and lower segment of roughly equal volume. At maximal inhale we compute the difference in volume of each segment from that of the mean segment volume. The ratio of chest volume change over the total volume change defines the percentage of “chest breathing”, which we refer to as the “type” of breathing.

3. Breathing Space (Shape Model)

Given a single 3D scan of a subject it is not well defined what part of the observed shape is due to breathing and what is due to the intrinsic shape of the person; e.g. do they have a large chest or are they inhaling deeply? However, given multiple scans of the same subject at different time instants in the breathing cycle, we can extract the shape and pose variations due to breathing. After registering the initial 3D scans (above), our data consist of aligned 3D meshes of multiple subjects at unknown time instants in their breathing cycle. Given a set of K 3D meshes in correspondence, $X_{ij}, i = 1, \dots, K$, for a subject j , we extract their intrinsic shape, D^j , as well as the shape deformations due to breathing B^{ij} for each X_{ij} by extending a SCAPE body model [6].

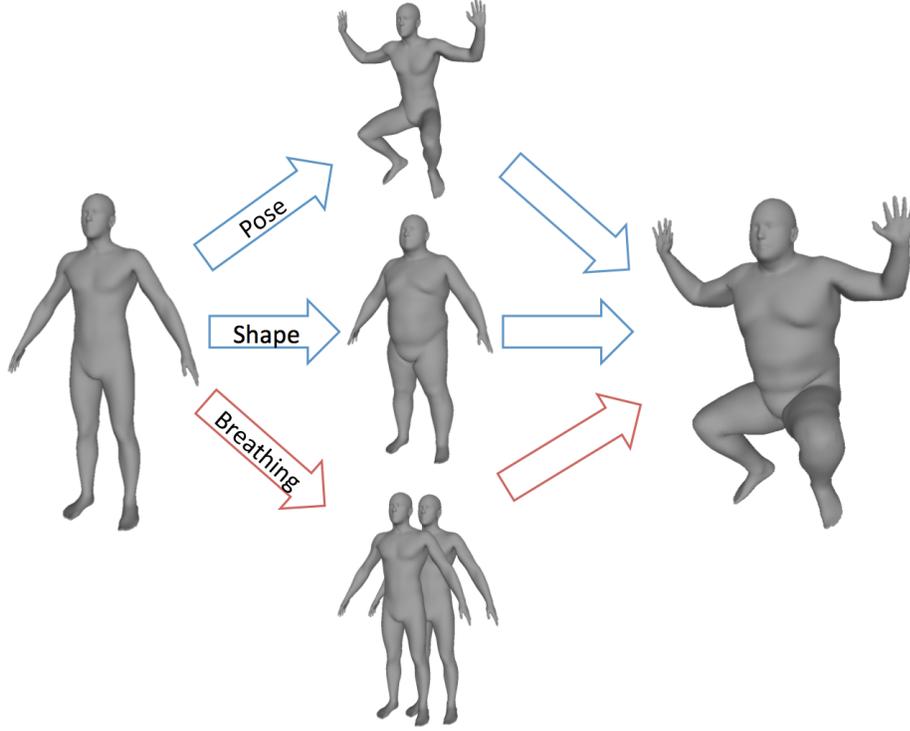


FIGURE 5.4. **SCAPE+Breathing**. The standard SCAPE [6] factors body shape into intrinsic shape and pose-dependent shape. We add a new shape deformation for breathing and combine all three into a model with separate controls for breathing.

SCAPE represents body shape as a deformation from a template mesh to an instance mesh using deformation gradients [88]; the basic idea is summarized in Figure 5.4. The deformation gradients in SCAPE are linear transformations that align corresponding triangles t between a template mesh T and an instance mesh X_{ij} . Since we have aligned the template with all the scans, T and X_{ij} have the same topology. To reconstruct mesh X_{ij} using the SCAPE model, three types of deformation gradients are applied to the triangles t of a template mesh T : pose-dependent transformations, Q_t^{ij} , identity-dependent transformations, S_t^{ij} , and rigid part rotations $R_{l[t]}^{ij}$. More specifically, given the edges $\hat{v}_{t,e}, e = 0, 1$ of each triangle t on the template, we compute the edges $v_{t,e}^{ij}, e = 0, 1$ of triangle t belonging to the i -th mesh of subject j as

$$(12) \quad v_{t,e}^{ij} = R_{l[t]}^{ij} S_t^{ij} Q_t^{ij} \hat{v}_{t,e}$$

where $l[t]$ denotes the body part to which triangle t belongs. The template mesh is segmented into distinct parts and all the triangles of the part undergo the same rotation $R_{l[t]}$; the part segmentation is illustrated in Figure 5.5.

There is one extra step to SCAPE. The above equation acts on every triangle in the mesh independently, resulting in a collection of triangles that do not necessarily form a valid mesh. SCAPE adds an extra step of solving for the valid mesh with triangle deformations that best match those above.

Additionally, the identity dependent deformations for a population of people can be approximated as a linear combination of basis deformations learned using principal component analysis (PCA). We do not discuss here the details of how to learn a SCAPE model in detail. We have trained our SCAPE model using approximately 4000 laser scans of men and women in roughly the same pose [77] and our own dataset of approximately 1800 scans of people in a wide variety of poses. This later dataset is used to learn the non-rigid deformations, Q_t^{ij} , which are a function of the part rotations (see [6]).

3.1. Adding Breathing. The deformation matrices R, S, Q mentioned above are functions of either pose parameters, \mathbf{r} , or shape parameters, \mathbf{u} , corresponding to linear coefficients in the PCA space; that is,

$$(13) \quad v_{t,e}^{ij} = R_{l[t]}(\mathbf{r}^{ij})S_t(\mathbf{u}^{ij})Q_t(\mathbf{r}^{ij})\hat{v}_{t,e}.$$

These parameters provide the animator controls to create a body shape \mathbf{u} in pose \mathbf{r} .

One of our key contributions is to extend SCAPE by separating the identity-dependent deformations S into two parts: one due to the intrinsic shape of the person, D , and one due to breathing, B (Figure 5.4). The functions D and B depend on intrinsic shape parameters, \mathbf{d} , and the shape parameters related to breathing, \mathbf{b} . Additionally, we separate the pose into static pose, \mathbf{a} , and, optionally, pose due to breathing, \mathbf{c} . Our proposed model, B-SCAPE, takes the following form:

$$(14) \quad v_{t,e}^{ij} = R_{l[t]}(\mathbf{a}^{ij} + \mathbf{c}^{ij})(D_t(\mathbf{d}^j) + B_t(\mathbf{b}^{ij}))Q_t(\mathbf{a}^{ij} + \mathbf{c}^{ij})\hat{v}_{t,e}.$$

To describe pose we use an axis-angle representation. In this representation it is meaningful to add pose parameters as long as self-intersection constraints and joint limits are not violated. Previous SCAPE models (and related models) ignore breathing deformations. Here we make them explicit. Below we show how to learn and then parameterize these by breathing type s , volume v and gender g . We end up with a model of the following form:

$$(15) \quad v_{t,e} = R_{l[t]}(\mathbf{a} + \mathbf{E}(g, v))(D_t(\mathbf{d}) + B_t(\mathbf{F}(s, v)))Q_t(\mathbf{a} + \mathbf{E}(g, v))\hat{v}_{t,e}.$$

3.2. Extracting the Breathing Deformations and Pose. Given multiple scans from the subjects in our training set, our goal is to extract the intrinsic shape, D^j , of each subject as well as the shape deformations due to breathing, B^{ij} . Recall that all scans are in correspondence with the template (and hence the SCAPE model).

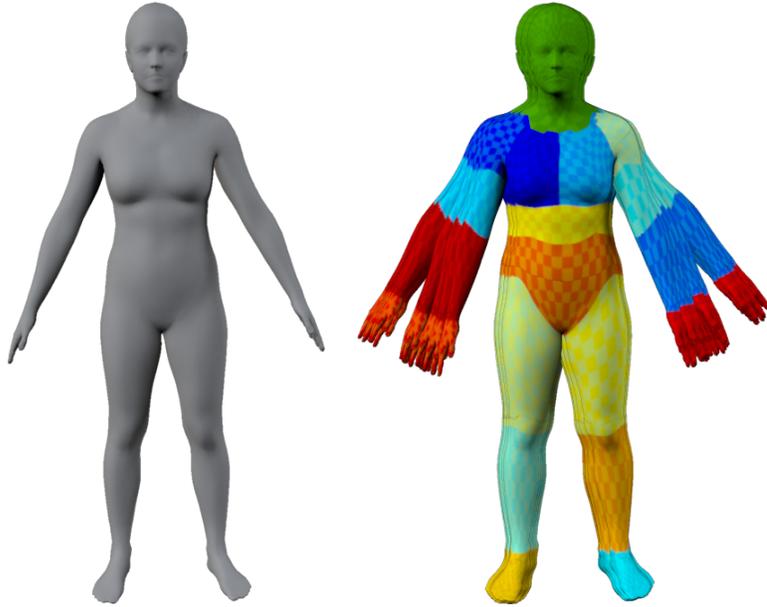


FIGURE 5.5. Pose change during breathing. Left: Mean pose of a subject. Right: Posture variation of this subject while breathing in the standing A-pose. Each part is color coded to show the body segmentation.

Consider one subject j with K aligned meshes X_{ij} ; we seek to extract the breathing-related deformations B^{ij} . This means we want to effectively factor out pose, pose-dependent deformations, and identity to focus on what is left. This remainder should be due to breathing.

To recover the deformations for X_{ij} , we first solve for the shape deformations S^{ij} by minimizing (see [6])

$$(16) \quad \arg \min_{S^{ij}} \sum_t \sum_{e=0,1} \|R_{l[t]}^{ij} S^{ij} Q_t^{ij} \hat{v}_{t,e} - v_{t,e}^{ij}\|_F^2 + \beta \sum_{t_1, t_2 \text{ adj}} \|S_{t_1}^{ij} - S_{t_2}^{ij}\|_F^2.$$

The first term minimizes the reconstruction error between the vertices of the captured meshes and their mesh representation based on deformations gradients. The second term enforces smooth deformations between adjacent triangles that represent the shape component of the mesh.

Given that our meshes are in correspondence and segmented, it is easy to estimate the rigid rotation matrices $R_{l[t]}^{ij}$ between corresponding body parts in the aligned mesh. We convert the rotation matrix per body part to an axis-angle representation of pose relative to the template mesh consisting of 3 parameters. That amounts to a vector, \mathbf{r}^{ij} , of 57 pose parameters per mesh (3 parameters, 19 body parts). We approximate the static pose with the average pose parameters over all meshes per subject, $\mathbf{a}^j = \frac{1}{K} \sum_i \mathbf{r}^{ij}$, and the dynamic pose with the residual pose parameters $\mathbf{c}^{ij} = \mathbf{r}^{ij} - \mathbf{a}^j$. Figure 5.5 shows the mean and breathing-related pose for one subject.

Each subject was scanned multiple times at unknown time instants in their breathing cycle. After we estimate S^{ij} , $i = 1, \dots, K$, we approximate the intrinsic shape (average) of the subject as the average of

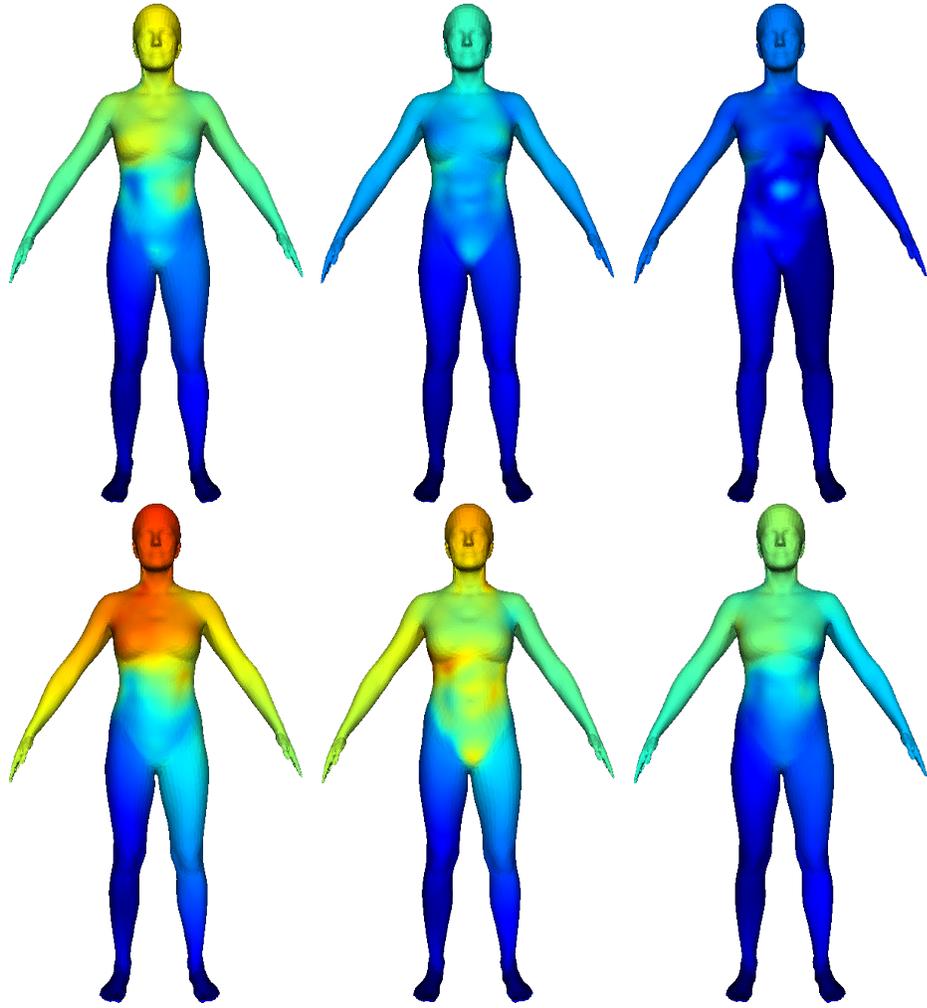


FIGURE 5.6. Shape change during breathing. Several examples of fitted meshes and how they deviate from the mean shape of the subject. Hot colors indicate greater distance.

the deformations, $D^j = \frac{1}{K} \sum_i S^{ij}$. Figure 5.6 shows example meshes and how they deviate from the mean. We found that as few as $K = 20$ scans were sufficient to extract a reasonable representation of a subject's intrinsic shape. The residual shape deformation due to breathing then simply is $B^{ij} = S^{ij} - D^j$. We do this for all subjects in our dataset and use this below to learn a model of breathing.

4. Statistics of Breathing

Respiration induces change in body shape and pose. In this section, we study the statistics of body deformations and posture variation due to breathing. In addition, we examine correlations with intrinsic attributes of humans, such as gender and identity shape.

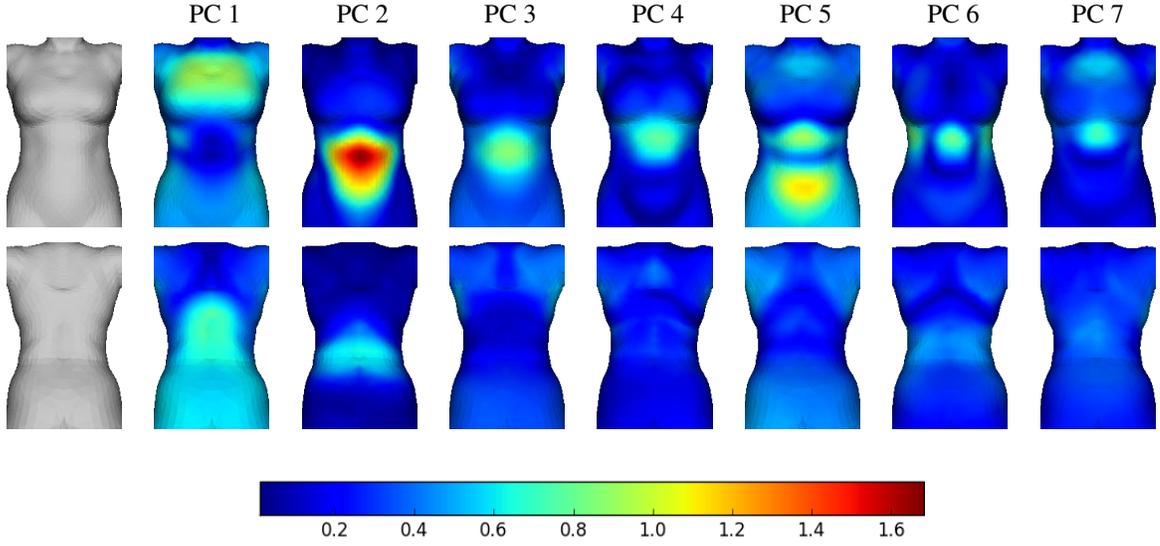


FIGURE 5.7. **Principal components of breathing.** (Gray) Mean female body. (Color) Ordered principal components shown at +5 standard deviations. Each body is color coded based on the Euclidean distance (in cm) between corresponding vertices in the visualized body and the mean shape. Top row corresponds to front view, bottom row to back view.

4.1. Breathing Shape Statistics. After estimating S^{ij} over all subjects, we end up with a very high-dimensional representation of the shape of each mesh. The dimensionality of S^{ij} is $9 \times F$, where F is the number of mesh triangles and 9 is the number of parameters of the 3×3 deformation gradient per triangle. Intuitively, the shape deformations due to breathing can be expressed with a much smaller number of parameters. Similar in concept to SCAPE, we learn a low-dimensional representation of shape change during breathing expressed as a linear combination of basis vectors, $G^m \in \mathbb{R}^{9F}$, $m = 1, \dots, M$, $M \ll F$. We learn the basis vectors of breathing by computing the principal components (PCs) of the breathing deformations B^{ij} of all subjects. We then represent a breathing shape deformation using a small number of linear coefficients, \mathbf{b}^{ij} ; these are our breathing animation controls. Breathing deformations can be approximated using the basis vectors and the linear coefficients as

$$(17) \quad \hat{B}^{ij} = \sum_m b_m^{ij} G^m.$$

Figure 5.7 illustrates the principal components of breathing. Conceptually, the first two components correspond mostly to motion of the chest and the stomach, respectively. The remaining components represent higher-frequency variation of shape in the torso area. In our experiments, we have used $N = 20$ PCs which account for 76% of the variance in the data. The number of components was selected empirically; using more components does not noticeably improve the realism of the synthesized breathing animations.

We evaluated whether breathing deformations were correlated with body shape. In general the correlation is quite weak (correlation coefficients are below 0.5). In particular the first few principal components of body shape are not strongly correlated with breathing shape deformations. For higher-order shape components, capturing finer details of the body (e.g. rolls of fat), we did find some correlation with breathing deformations. To examine the dependence of the shape change due to breathing to the intrinsic shape of a human, we trained a SCAPE model using 4000 human scans of various shapes from the CAESAR dataset [77]. Figure 5.8 shows the correlation coefficients matrix for the first 5 PCs related to body shape variation and the 10 PCs of breathing shape. We observe that coarse breathing shape, described by the first few PCs of breathing deformations, is not very correlated with coarse body shape. The coarse human shape is described using the first 5 PCs of a SCAPE model trained from CAESAR. However, we do observe correlation between higher-order components of breathing and intrinsic shape with maximum correlation coefficient around 0.5. Intuitively, we would expect dependence between intrinsic shape and breathing shape in areas of the body where there are prominent skin folds and muscles. Since the effect is weak, we ignore this in our model.

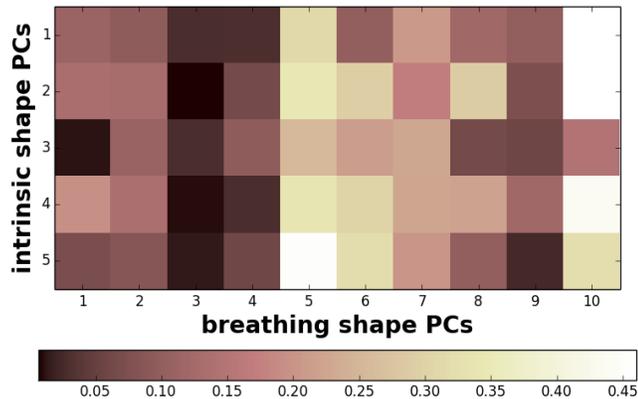


FIGURE 5.8. Linear correlation coefficients (unsigned) between coarse human body shape and shape change due to breathing. The coarse human shape is described using the first 5 PCs of a SCAPE model trained from CAESAR. 10 PCs from breathing shape.

4.2. Breathing Pose Statistics. As with breathing shape, we extract a low-dimensional representation of breathing pose variation, \mathbf{c}^{ij} , using PCA. The low-dimensional pose representation can be expressed as $p_n^{ij} = P^n T \mathbf{c}^{ij}, n = 1, \dots, N$ where $P^n \in \mathbb{R}^{57}$ are the principal components; here we use $N = 4$ components. This results in a low-dimensional description of breathing pose, \mathbf{p}^{ij} . We found, however, that not all components were correlated with breathing. The subjects were allowed to “relax” between consecutive scans (20 sec) and adjusted their pose and moved their feet slightly. Consequently we discarded pose components

that were not strongly correlated with breathing (i.e. volume). Figure 5.9 shows the 3 most informative principal components of pose change during breathing. As expected, they are related to spine and shoulder/neck motion during breathing. Examining the low-dimensional pose space, we did not find strong correlations with the intrinsic shape of the subject, but we did find correlation with gender. In particular, women show a more pronounced forward/backward rocking of the upper body during breathing. Consequently, we build a separate model of pose variation for men and women.

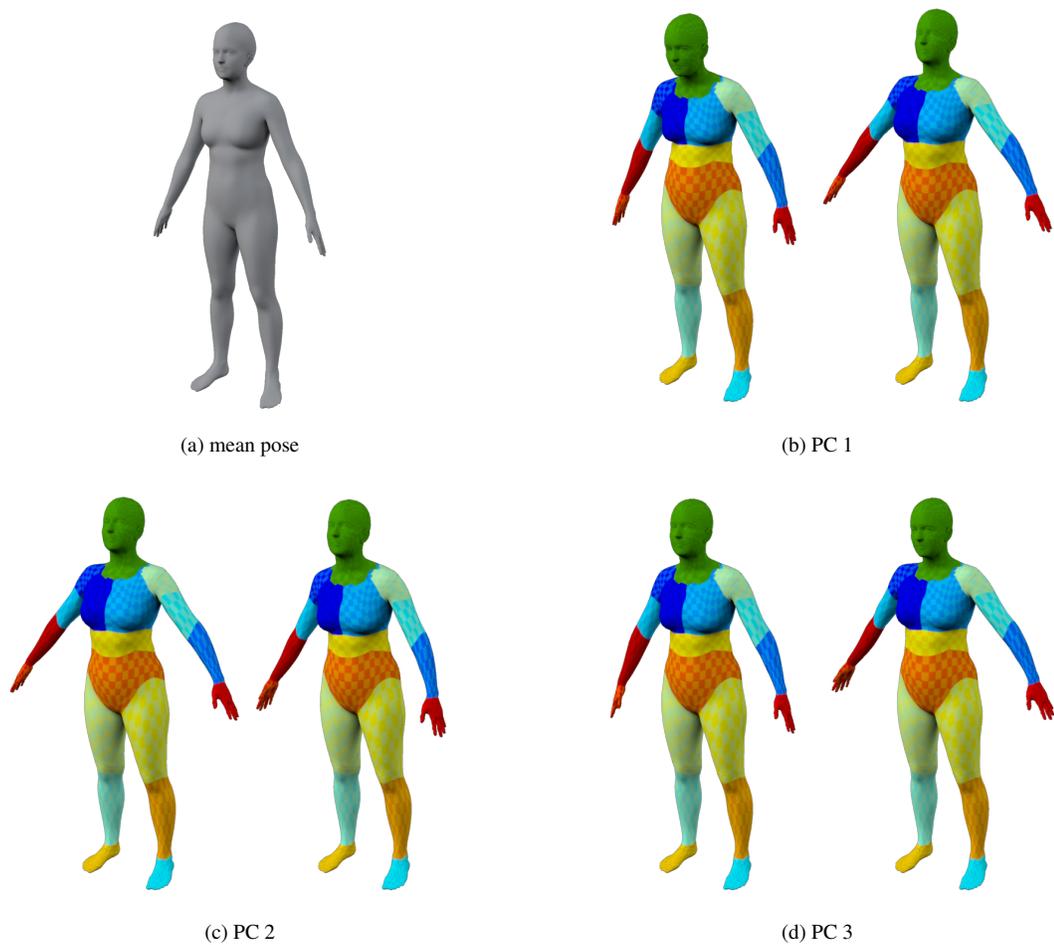


FIGURE 5.9. Pose change during breathing. (a) Mean female body in the “A pose”. (b,c,d) The 3 pose principal components most correlated with breathing (volume) displayed at ± 6 standard deviations.

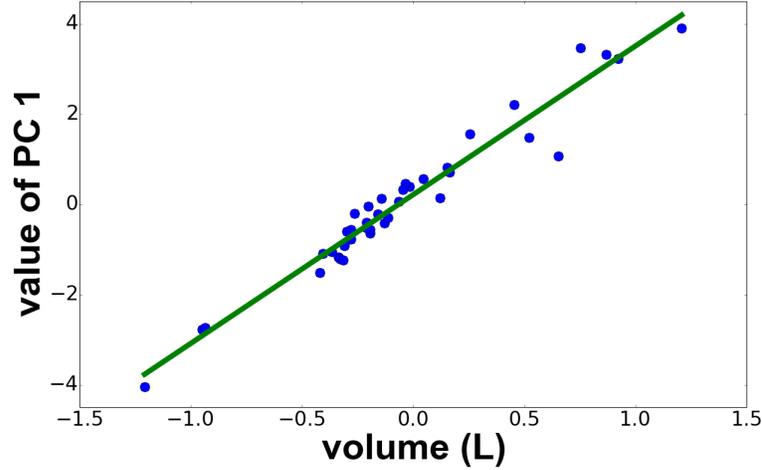


FIGURE 5.10. Volume change versus shape change. For a specific subject, with a particular type of breathing, we find there is a linear relationship between the breathing shape coefficients and changes in mesh volume. Here we see this for the first principal component.

5. Breathing Model

The statistics of breathing shape and pose change do not provide a model for animation. What we need is a *model* that relates these changes in pose and shape to physiological parameters like lung volume over time. We develop our model in stages.

5.1. Shape change during breathing. Subjects were instructed to breath in three different ways: normally, with the chest, and with the stomach. Scans from each of these conditions were treated as separate trials. Using mesh volume as a proxy for lung volume, we express shape change of a subject within a trial as a function of changes in mesh volume from the mean subject mesh. We find a largely linear relationship between the coefficients of breathing shape and mesh volume change (Figure 5.10). Let $\mathbf{Z}_j \in \mathbb{R}^{K \times 2}$ be a matrix containing a column with ones and a column with the volume differences, v_{ij} , between the i^{th} mesh of subject j and the mesh corresponding to their mean shape; K is the number of meshes in the trial. Let $\mathbf{Y}_j \in \mathbb{R}^{K \times M}$ be a matrix containing the low-dimensional breathing shape coefficients, \mathbf{b}^{ij} , representing the breathing shape deformations of the training meshes (Sec. 4.1). For each trial, we learn a subject-specific linear model, \mathbf{W}_j , relating changes in breathing volume to shape deformation coefficients

$$(18) \quad \arg \min_{\mathbf{W}_j \in \mathbb{R}^{2 \times M}} \|\mathbf{Z}_j \mathbf{W}_j - \mathbf{Y}_j\|_F^2.$$

5.2. Breathing types. In the linear model above we assumed that the subject performed the same type of breathing throughout each of his 3 scanning sessions (normal, chest, stomach breathing). The type of

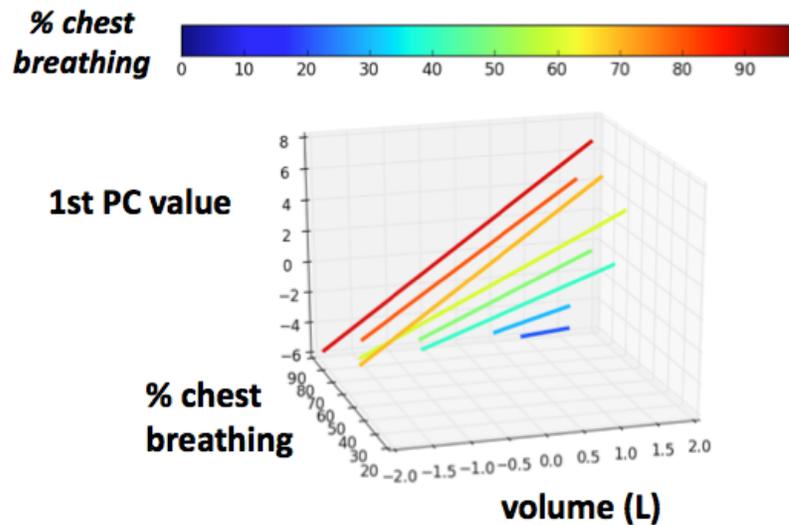


FIGURE 5.11. Shape as a function of volume and type of breathing. Linear models of shape change during breathing for various breathing types (percent chest breathing) considering only the 1st PC. Color coding is based on breathing type.

breathing plays an important role in animation. However, the trial classification above provides only a crude classification of the type of breathing. To more precisely classify the type of breathing performed in a trial we used the linear function and the maximum inhale volume to predict the shape of the body at maximum volume. Using the segmentation of the torso into chest and stomach regions (Sec.), we computed the ratio of chest volume change of this mesh from the mean to total volume change of chest and stomach. This gives a value s_j for each trial, indicating the percentage of chest breathing present in that trial.

5.3. Breathing model. Finally, we have what we need to learn a function, $F(s, v)$, that takes as input the breathing type s and volume difference v and returns the corresponding linear shape deformation and pose coefficients. Given the classification of breathing type above, we divide the trials into 10 categories corresponding to 0%-10%, 10%-20%, . . . , and 90%-100% chest breathing. Within each category we combine all the individual linear models into an aggregate linear model relating each shape coefficient to change in volume. This aggregate model can be thought of as the average linear relationship predicting shape change from volume change.

Figure 5.11 shows what this looks like for the first principal component. Each colored line is an aggregate linear model for a specific value of breathing type, s . Note that the slope of each line is different. Recall that the first principal component captures mostly chest deformation (Figure 5.7). The higher the value of s , the more the chest is involved, and the greater the correlation of the first component with changes in volume. Note further that this results in a function that is non-linear in s and v .

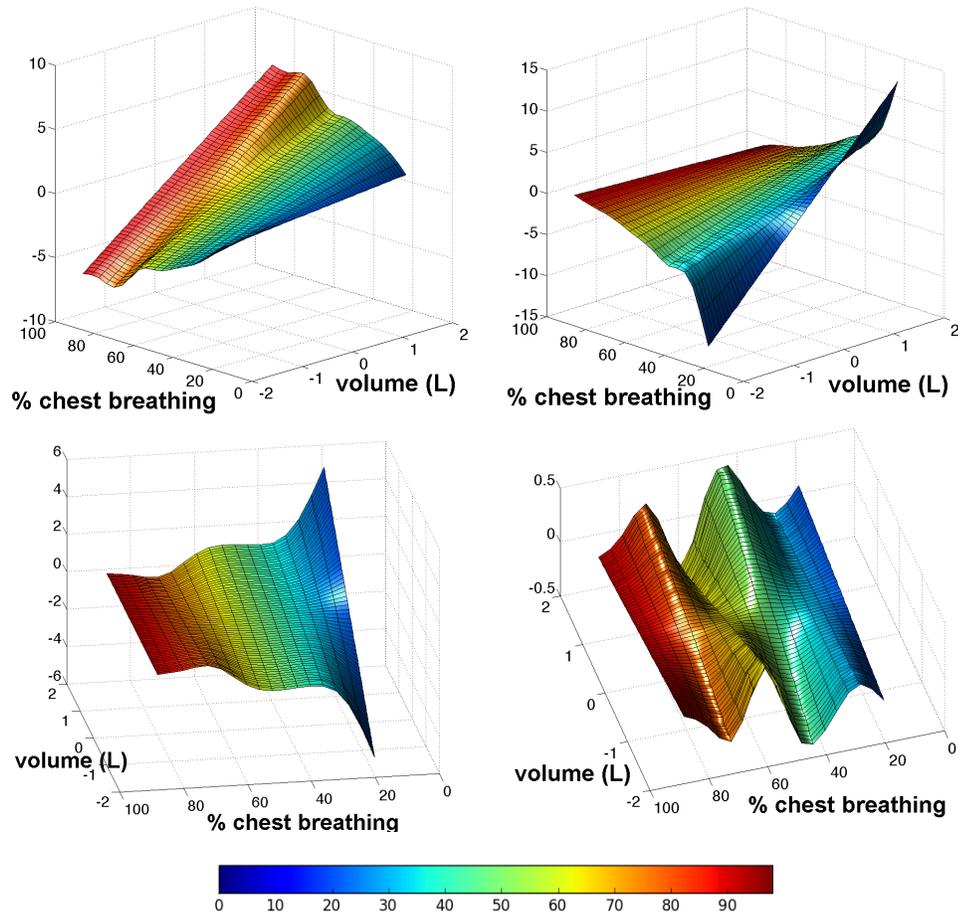
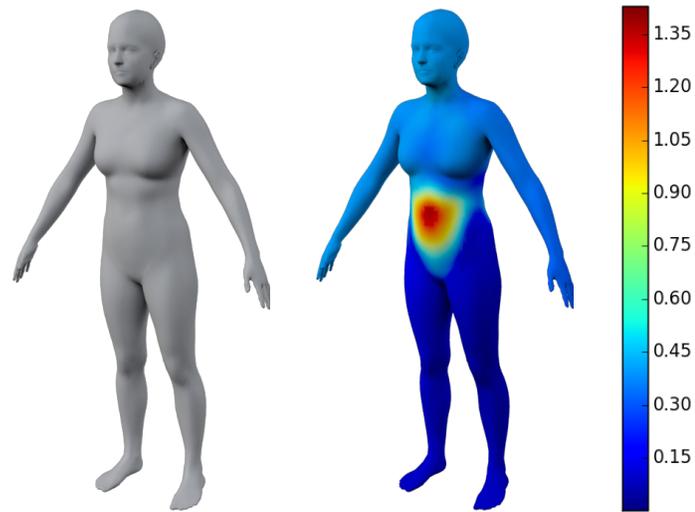


FIGURE 5.12. Shape change during breathing as a function of breathing type and volume. The breathing type is expressed as the percentage of chest relative to stomach breathing. The first two PCA coefficients of shape change (first two plots from the left) are correlated with "chest-breathing" and "abdominal-breathing" types respectively. The third and fourth PCA coefficients (two rightmost plots) are generated from higher frequency functions. Color coding is based on breathing type (% of chest breathing).

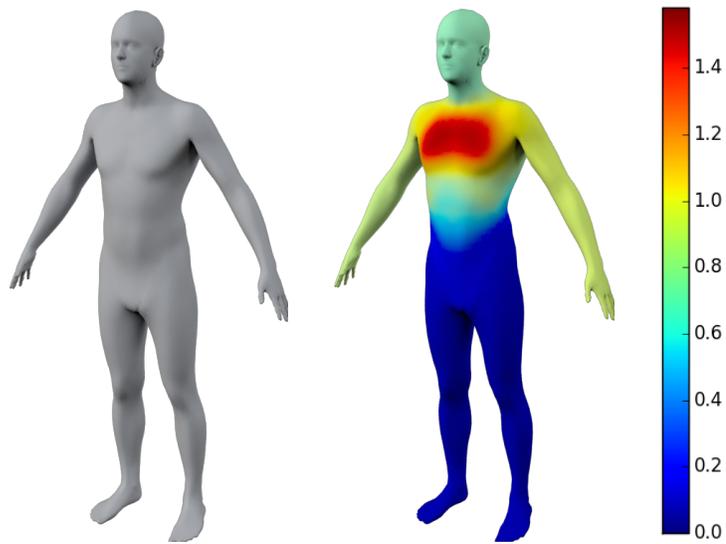
We want a model of breathing that is continuous in s and v and we achieve this by fitting a surface to the changing regression functions using cubic interpolation. Figure 5.12 shows some examples of the resulting functions $w_m(s, v)$. As we saw before, the first two principal components are very correlated with chest and stomach breathing respectively. This is evident in the corresponding weight functions (top two subplots in Figure 5.12).

In the final breathing shape model then, we weight the principal components, G^m , by a non-linear functions $w_m(s, v)$

$$(19) \quad F(s, v) = \sum_m w_m(s, v) G^m.$$



(a) “abdominal-breathing”



(b) “chest-breathing”

FIGURE 5.13. Examples of (a) “abdominal-breathing” and (b) “chest-breathing”. The gray bodies represent the mean shape of the depicted humans. Shape change due to breathing is color coded based on the Euclidean distance (in cm) between every mesh vertex and the corresponding vertex of the mean shape.

Figure 5.13 shows two example meshes at maximum inhale: a female breathing with the stomach and a male breathing with the chest.

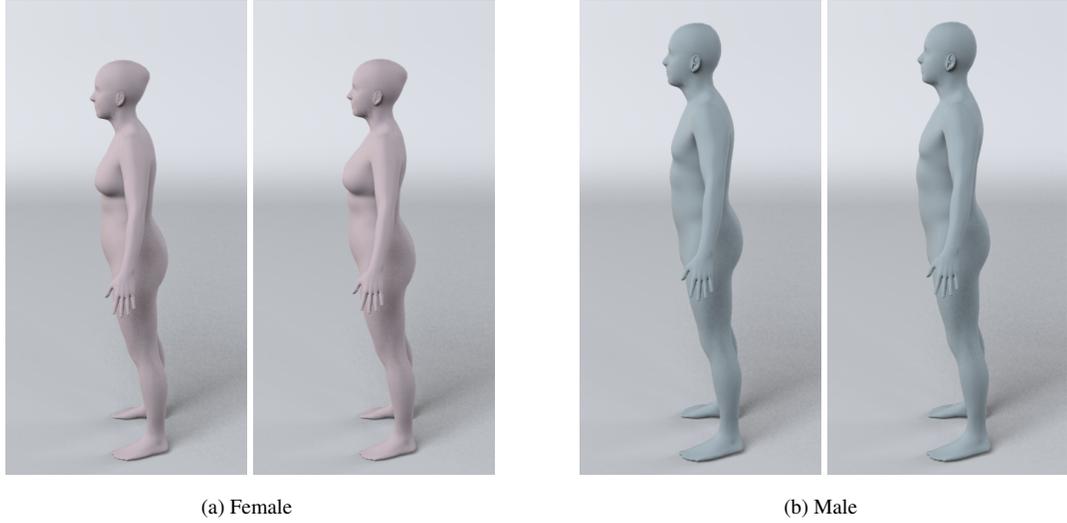


FIGURE 5.14. Pose model evaluated at “breathing in” and “breathing out” for (a) females, and (b) males considering the first 3 principal components.

5.4. Pose model. Based on the insights from Sec. 4.2, we derive a generic model of pose change per gender g , $E(g, v)$, parameterized additionally by the breathing volume v . Let $\mathbf{O}_g \in \mathbb{R}^{K_g \times 2}$ be a matrix containing a column with ones and a column with the volume differences, v_{ij} , over all subjects j of gender $g = \{\text{male, female}\}$. Let $\mathbf{H}_g \in \mathbb{R}^{K_g \times N}$ be a matrix containing the PCA projections of pose, \mathbf{p}^{ij} , corresponding to the training meshes as described in Sec. 4.2. For each gender, we define a linear model for predicting breathing induced pose deformations using linear least squares regression:

$$(20) \quad \arg \min_{\mathbf{L}_g \in \mathbb{R}^{2 \times N}} \|\mathbf{O}_g \mathbf{L}_g - \mathbf{H}_g\|_F^2.$$

The pose model per gender is shown in Figure 5.14. The upper body of females moves more than the upper body of males. Interestingly, shape deformations were the same for men and women. The pose model is not parameterized by the breathing type. Note also that the pose model does not need to be used for animation; for example, when animating the breathing of a moving character, we do not use the pose model.

5.5. Influence of shape and pose on animation realism. Previous work on fine-tuning animations of 3D avatars to create the effect of breathing has focused on varying either the shape or the pose of the animated character. Given our proposed breathing model describing both shape and pose change, we want to find out which of the two is more influential at conveying the feeling of breathing. We conducted a perceptual study where users were asked to select the most realistic animation out of three: breathing animation by changing only the pose, only the shape, pose and shape together. The animations were synthesized using the personalized shape model described in Section 5.1 and personalized pose models derived in a similar way as in Section 5.4, but considering the data of each subject separately. The three animations were displayed at

the user simultaneously (side to side) and in a randomized order. Each animation was rated by 9 users using Amazon Mechanical Turk. For evaluation, we used 53 subjects that were captured after being instructed to breath “naturally”. We excluded subjects that were explicitly instructed to breathe with the chest or stomach because focusing on breathing with a specific type led many participants to keep their arms stiff and still. To ensure that the users did not select an animation randomly, they were asked to justify their choice. After gathering the votes of the users, we count how many times each configuration is selected as the configuration that yields the most realistic animation. In case there is a tie among the votes for a subject, we split the credits among the first place configurations. In 47% of the cases, the Mechanical Turk users selected the animations synthesized by combining pose and shape change as the most realistic ones. The animations synthesized by varying only the shape were ranked first 48% of the times and pose-only breathing deformations were selected as the most realistic ones in only 5% of the cases. We can see that the users favored shape change as the element that yields greatest realism in animating breathing. Adding pose change on top of breathing shape change did not create a noticeable difference in the perceived realism.

To reinforce this statement, we conducted a follow-up perceptual experiment in order to investigate whether the users are able to distinguish the difference between the animations generated by varying shape and varying pose and shape together. The users were presented with a pair of animations using the same subjects as in the previous experiment. Each animation was synthesized by changing either only the shape of the 3D character or both the pose and shape (the two animations could also be the same). The users were asked whether they thought the animations were the same or different. In 85% of the cases, they were able to give the correct answer which denotes that they do not have a strong preference over adding pose change for animating breathing.

6. Breathing Animation

Respiration is time varying. In particular, as air moves in and out of the lungs, their volume changes. To animate breathing using the model defined above, we need a way to vary lung volume over time.

6.1. Trajectory editing. We developed a Maya tool to create and edit realistic 3D body shapes; that is similar to previous work on body shape modeling [4, 6, 44, 49]. We do not describe it further here. The tool also allows an animator to edit the temporal pattern of breathing.

Our breathing model takes two inputs: the breathing type and volume difference. Our interface includes a slider with which the animator selects the percentage of chest breathing enabling them to achieve different “styles.” A common assumption in the physiology of respiration [63, 98] is that air flow in lungs during breathing at rest pose is a sinusoidal function of time. Thus we provide an interface for controlling the parameters of a sinusoid function of volume over time. The intensity of the pose change can be adjusted separately from shape deformation. The amplitude and frequency of the sinusoid can be varied using sliders.

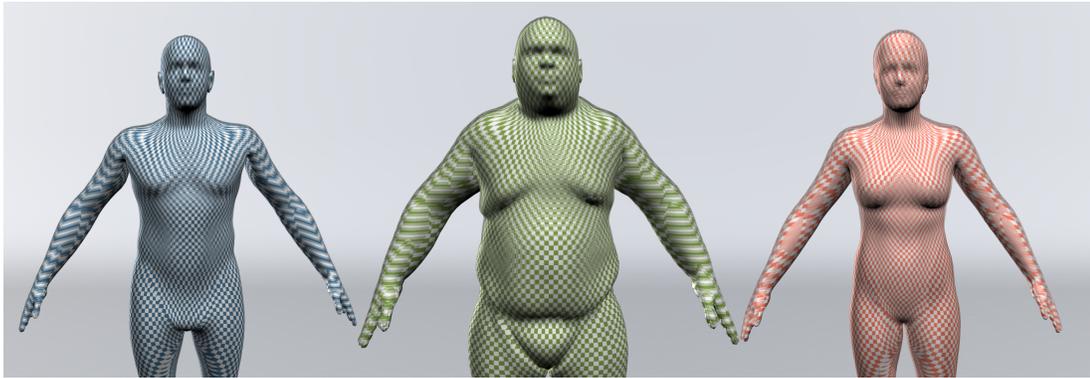


FIGURE 5.15. Visualization of breathing animation using a Moiré-like effect. We show the same three bodies and breathing types as in Figure 5.1, but here the maximum inhale and exhale shapes are overlaid with a transparent checkerboard that shows deformations of the surface (due to shape deformation only). When there is shape change, the checker pattern is misaligned and looks blurry. For example, focusing on the stomach of the middle character, one may have the sense of an expanding stomach.

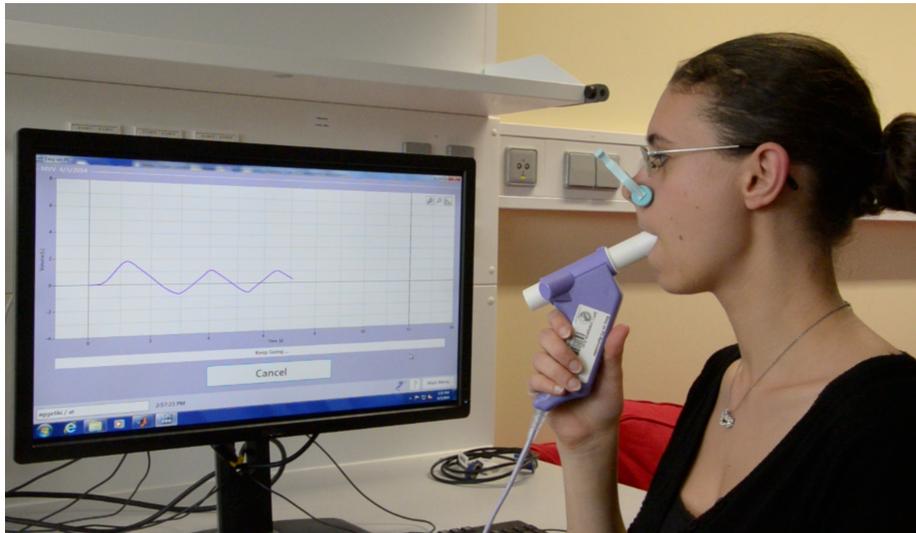


FIGURE 5.16. Breath acting. An actor breathes into a spirometer to convey the action and emotional content of a character. The changing volume of the lungs is recorded and used to animate breathing.

Figures 5.15, 5.21, 5.22, 5.23, 5.24 show examples of breathing animations based on the abovementioned assumption at full inhale and exhale.

6.2. Breath acting. Breathing in real life does not always follow a pure sinusoidal function. It varies with activity and emotion and plays a role in telling a story. We use a device call a spirometer (NDD MEDICAL, Zurich, Switzerland), which measures change in air flow, to capture the breathing pattern of a “breath actor” (Figure 5.16). Like a voice actor, the breath actor observes an animation and acts out the breathing that fits the action. We then use the recorded changes in volume to produce deformations (and possibly pose changes) and to animate a 3D avatar. We manually adjust the breathing type (chest or stomach) based on the action and emotions in a scene. This provides a simple and intuitive interface to produce realistic and compelling breathing animations. Figure 5.17 shows an example of synthesized breathing that matches the “surprise” emotion.

7. Results

To evaluate the realism of our model, we capture reference video material of a subject breathing. We then had a breath actor (different from the subject) watch the video and imitate the breathing using the spirometer. We used a 3D scan of the subject to create an avatar for their body shape, selected the amount of chest breathing manually, and then animated the body in the style of the subject. Note the we did not capture the pose of the actor during breathing and did not attempt to match the pose. Focusing on the breathing deformations, however, we find a good qualitative match between reference and animation (Figure 5.18).

7.1. Breathing in Action: Poses and Motion. Our model of respiration is trained using body scans of people in a standing “A” pose. While the pose variation model may be quite specific to this pose, the shape deformation model can be easily applied to other poses with realistic results. Figure 5.19 shows a body in a seated pose and a standing pose with the same breathing model applied. Notice that the breathing deformations are, in fact, different because the mesh is in a different pose.

We also animate the breathing of characters in motion. Figure 5.20 shows frames from a running sequence. Here again, the breathing deformation is color coded in terms of distance from the average shape. In this case a breath actor observed the animated body without breathing and simulated the breathing to go with it. While the running motion makes it harder to see the breathing animation, one can readily tell the difference between sequences animated with and without breathing.

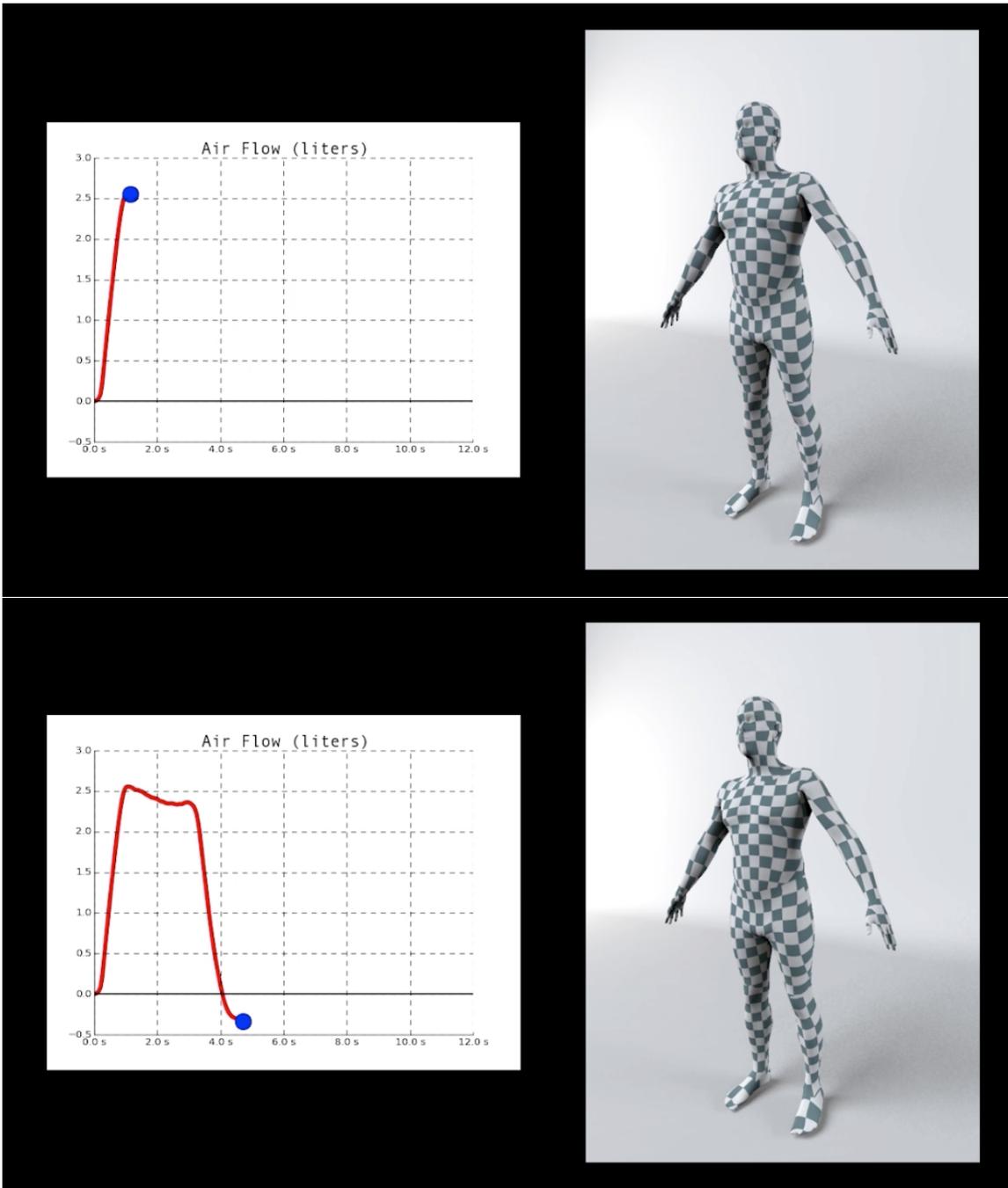


FIGURE 5.17. Spirometer-driven animation. Example snapshots of an animation where the “breath actor” was instructed to breathe while acting surprised.

8. The Last Breath (Conclusions)

We describe a model for realistic breathing animation. A key novelty of our approach is the use of high-resolution 3D scans in combination with a human body model to capture pose and shape change during



FIGURE 5.18. Reference video. Example frames from a reference video of a subject breathing with different types. On the left of each image pair we see the recorded motion. On the right, we show a roughly corresponding frame from our animation.

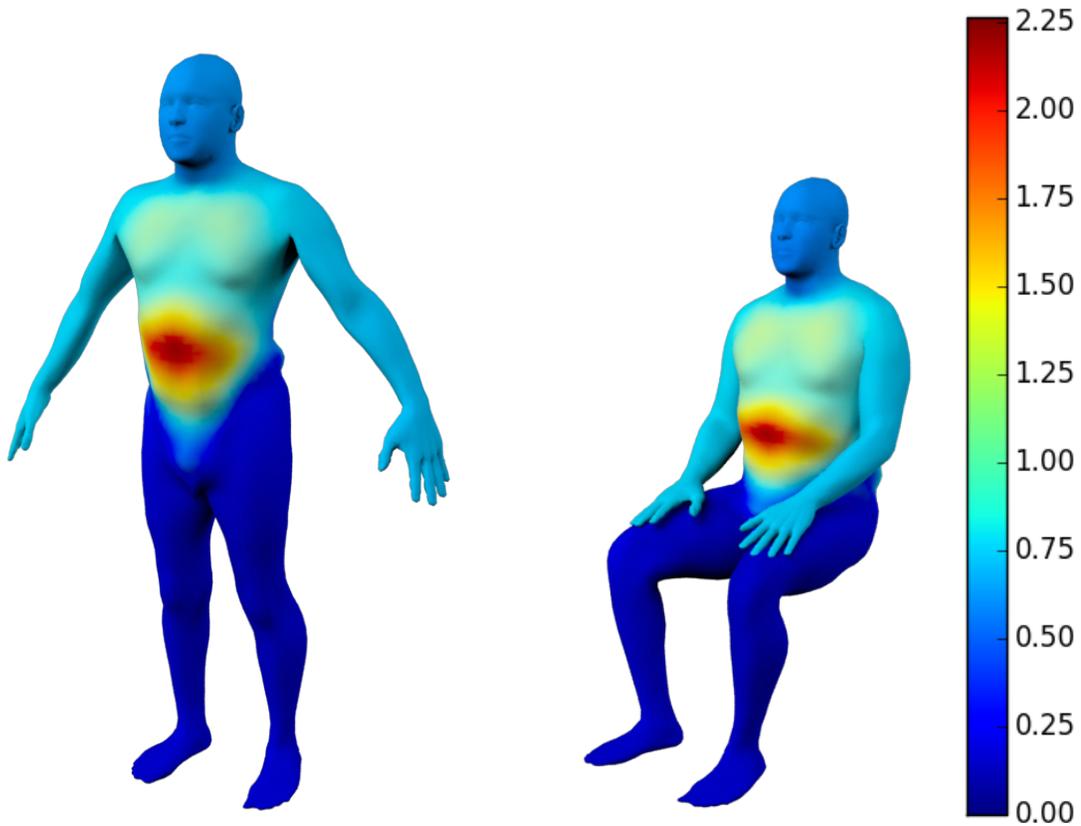


FIGURE 5.19. Pose and motion. Applying the breathing model with 40% chest breathing to a standing and seated pose. Color coding corresponds to the distance (in cm) between the vertices of the meshes at full inhale and the mean shape of the human.

breathing. We use mesh volume as a proxy for lung volume, which allows us relate breathing shape changes to a simple, physiological, control parameter.

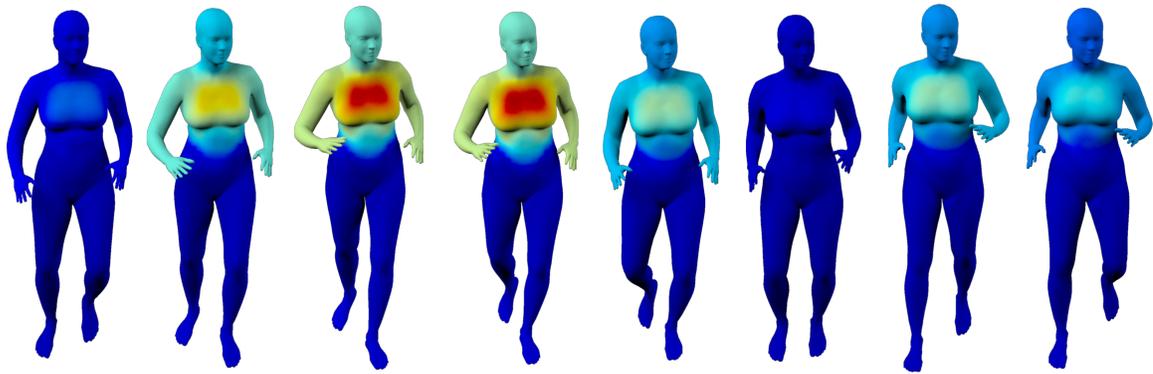


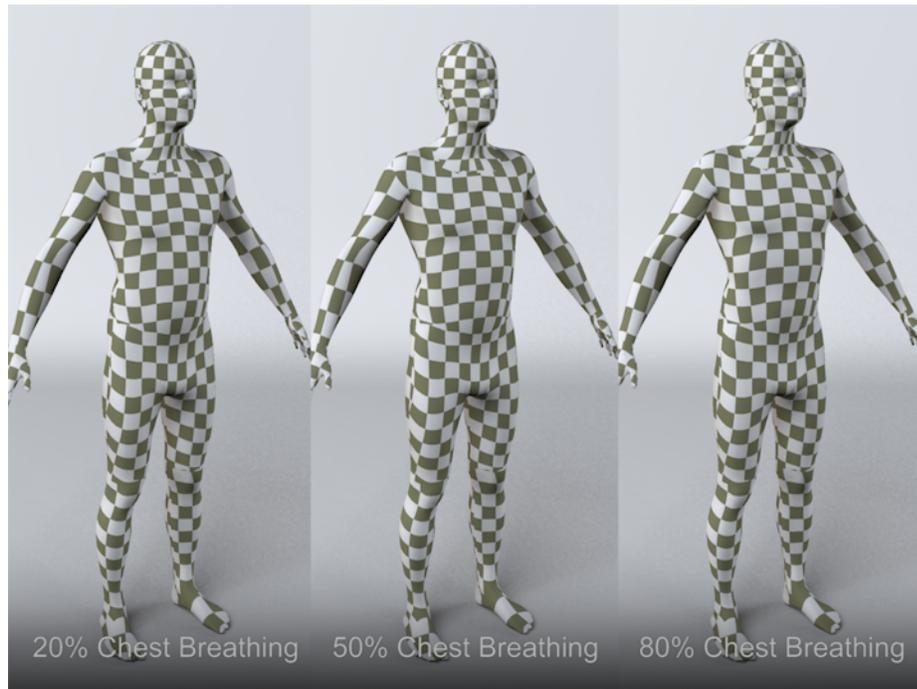
FIGURE 5.20. Simulated breathing while running. We took a 3D model of a person running and added realistic breathing shape deformations. The Euclidean distances between vertices of the breathing and non-breathing model are color coded (red is more distant while dark blue means no difference). Here the runner is breathing mostly with the chest and the temporal pattern of breathing was created by a “breath actor.”

By analyzing the statistics of breathing shape changes we found that: 1) there are statistically significant changes in whole-body posture and shape during breathing, 2) the differences in breathing shape between men and women are statistically insignificant but that there are some significant postural differences; 3) the dominant breathing shape changes are independent of body shape but more detailed changes are correlated with body shape; 4) people can perform different types of breathing (chest and stomach) and these are clearly reflected in the principal components of breathing shape; 5) for a particular type, lung volume is linearly related to these principal components; 6) this linear relationship varies with type, resulting in a non-linear model.

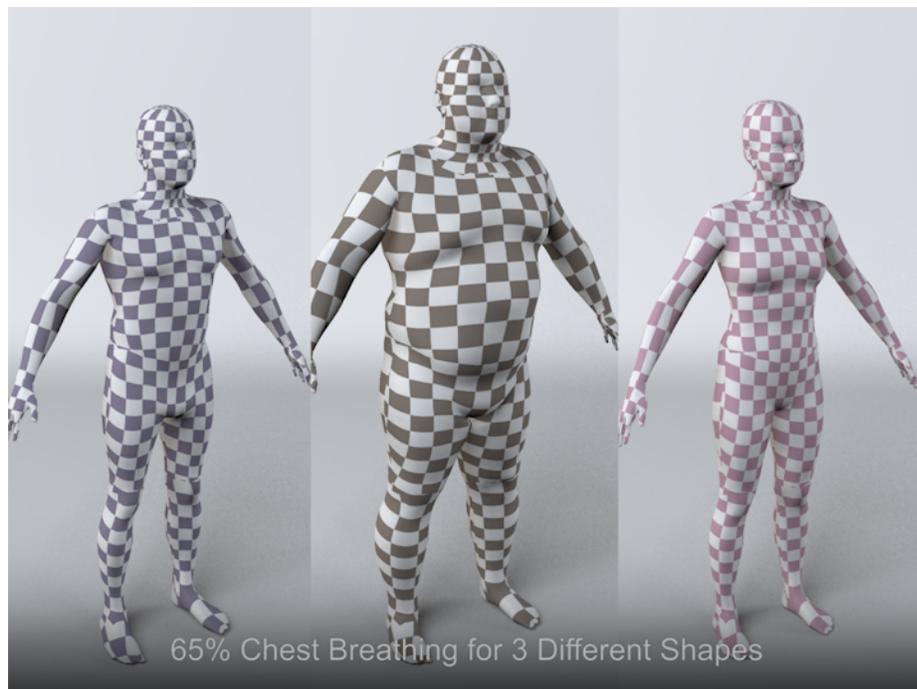
Based on this statistical analysis, we learn a model of breathing shape from 3D scan data. The parameters of the model are the lung volume and breathing type. We extend the SCAPE body model to include deformations due to breathing. This enables easy animation of arbitrary body shapes in any pose. Finally, we provide animators with a new breathing animation tool that uses a spirometer to capture changes in lung volume. By acting out a the respiration of a character by breathing into the spirometer, one can create novel and realistic animations that convey action and emotional content.

While usable today by an animator, there are ways to improve and extend our method. We capture breathing in a fixed pose but clearly shape changes will be influenced by posture (e.g. lying down). Future work should study how pose affects breathing deformations. Breathing shape is likely also correlated with activity and it would be good to build a temporal model of breathing *dynamics* as it relates to pose changes during activity. We used mesh volume to measure lung volume but it would be interesting to synchronize the output of a spirometer directly with the 3D scanning process. We focused on the two dominant types of

breathing described in the literature but we would like to capture a much wider range of scenarios including other actions like puff, pant, blow, gasp, wheeze, sigh, huff. Our methods could be use to give an animator the ability to select among these styles. We have focused on the body but it would be interesting to simultaneously analyze facial motions, which are also influenced by breathing. Finally, it would be interesting to explore example-based methods that could reconstruct coherent breathing from our scans; such methods are popular with motion capture data [46, 69] but have not been applied to 3D breathing shape.

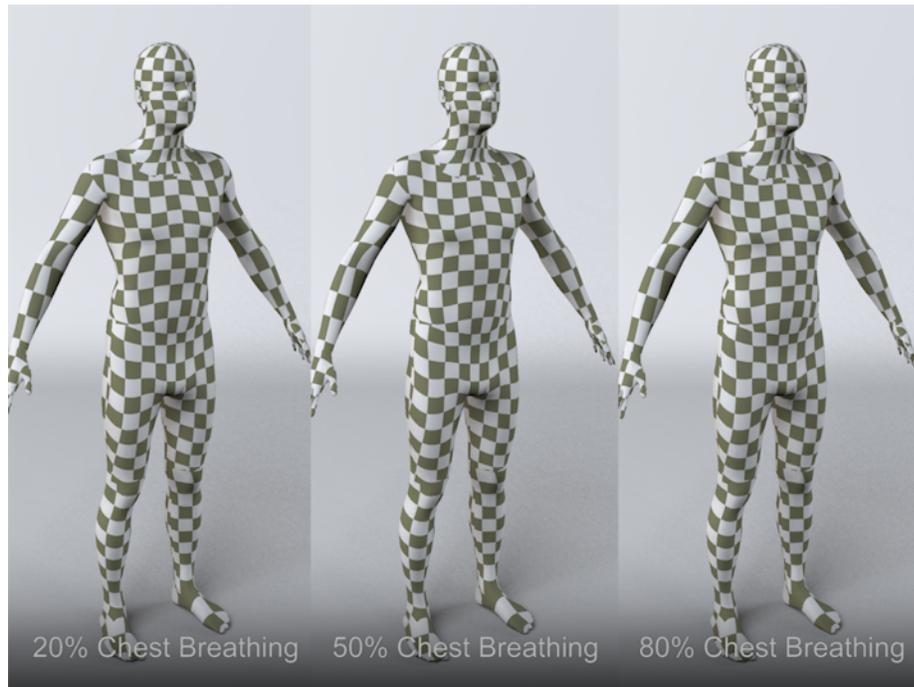


(a) same shape, different breathing type

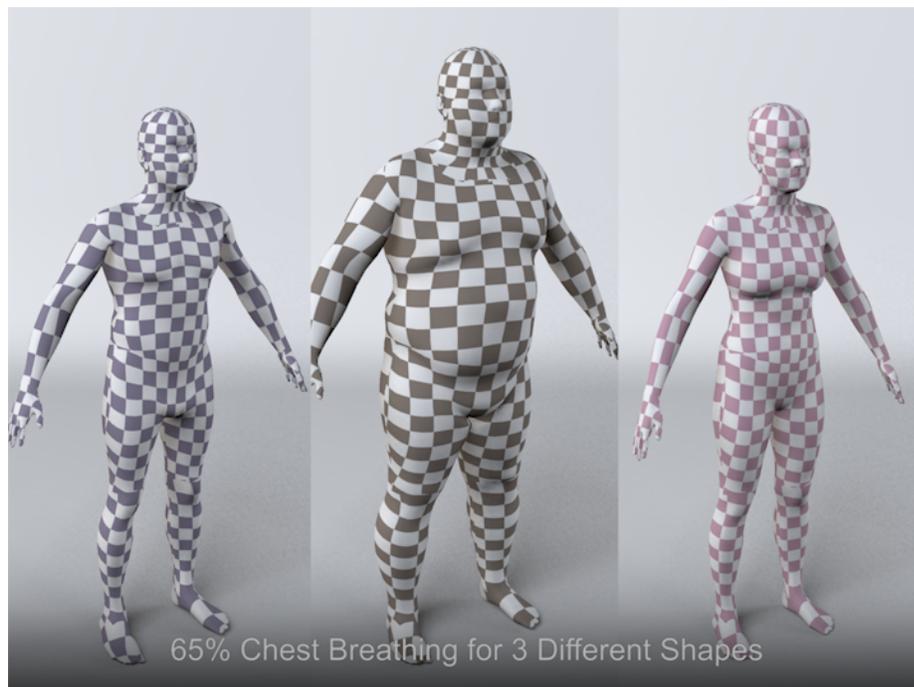


(b) different shape, same breathing type

FIGURE 5.21. Special for the electronic edition (switch back and forth with the next page to see the animation). (a) Different types of breathing applied to the same human. (b) Same breathing type applied to humans of different shape. Bodies are displayed at full inhale.

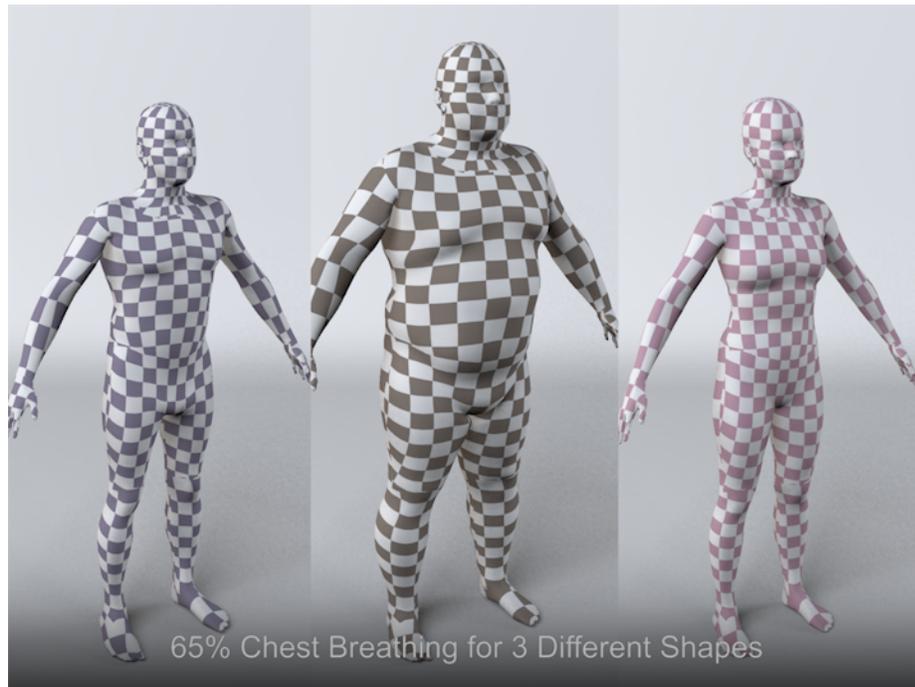


(a) same shape, different breathing type



(b) different shape, same breathing type

FIGURE 5.22. Special for the electronic edition (switch back and forth with the next page to see the animation). (a) Different types of breathing applied to the same human. (b) Same breathing type applied to humans of different shape. Bodies are displayed at full exhale.

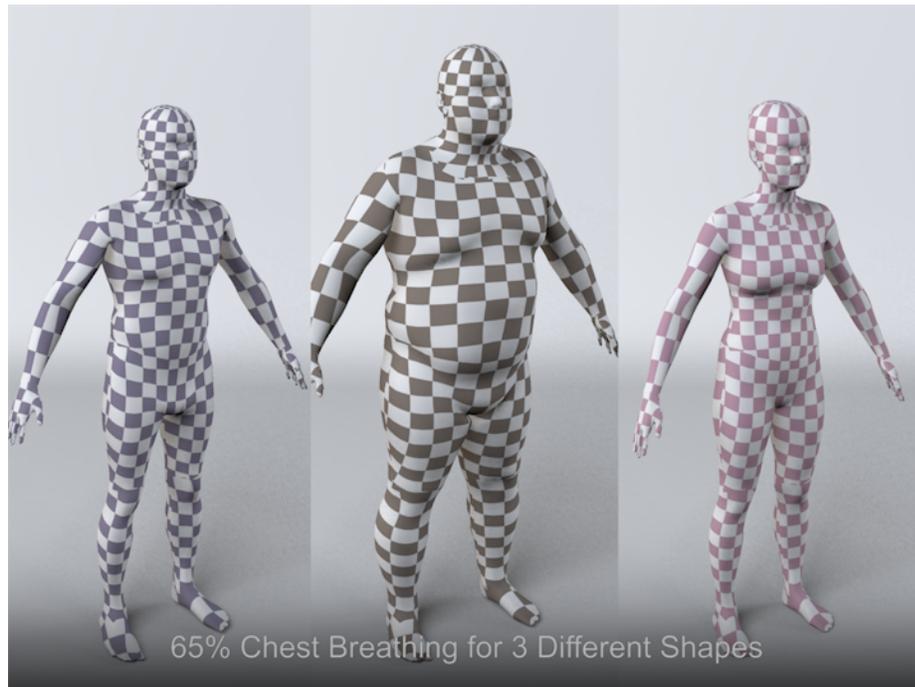


(a) front view



(b) back view

FIGURE 5.23. Special for the electronic edition (switch back and forth with the next page to see the animation). Bodies of different shape and gender at full inhale.



(a) front view



(b) back view

FIGURE 5.24. Special for the electronic edition (switch back and forth with the previous page to see the animation). Bodies of different shape and gender at full exhale.

Conclusions and Future Work

This thesis has focused on several aspects of modeling the shape of the human body in 3D. In this chapter, we summarize our contributions and discuss limitations and directions for future work.

1. Shape Matching between 3D meshes

Registering two arbitrary 3D meshes, no matter if they belong to the same object class, is a difficult task. Even for a human the smoother the meshes, the harder it is to define meaningful correspondences. In Chapter 2, we presented a method, PGSE, that yields reasonable sparse correspondences between 3D meshes. More specifically, we showed how geodesic distances can be used in a pairwise probabilistic graph framework to provide correspondences that are invariant to the pose, shape and resolution of two meshes. First, we proposed a local shape descriptor relying on geodesic distances that is invariant to pose, shape and mesh resolution. Geodesic distances, shortest paths on the mesh, are almost pose invariant. Avoiding protruding areas of the human body, such as belly, they are also almost shape invariant when normalized. Shortest paths that traverse the triangles of a 3D mesh are also to some degree independent to the mesh topology (as long as there are enough triangles to represent the underlying geometry). The proposed descriptor has high discriminative power on limbs of the body, but lower discriminative power in the rest of the body. To overcome this limitation it was used in a probabilistic framework in conjunction with pairwise geodesic constraints among possible correspondences. The problem of finding sparse correspondences was formulated as an inference problem on a Conditional Random Field (CRF) model. A CRF provides a plug-and-play framework for introducing constraints and there are also well established methods for approximate inference that work well in practice. In the absence of ground truth correspondences between two meshes, we proposed an error metric that is more consistent with how humans evaluate correspondences than previous metrics.

We have described a method that can be used for finding only sparse (up to ~ 200) correspondences between 3D meshes. An extension to finding dense correspondences could be achieved by rerunning our method selectively in areas spanned by a subset of the sparse correspondences found and joining the resulting new correspondences. In its current form, our method is oblivious of the class of the input 3D meshes. We expect the use of prior knowledge to enhance its performance. For instance, it could potentially be combined with anthropometric landmark detectors, e.g. nose detector. Such an approach could also eliminate false front-to-back matches (e.g. the back of a man with a hunch is mapped to the chest of a woman). An

alternative for incorporating prior knowledge is to learn the types of deformations for meshes of a specific class. Geodesic distances assume that deformations caused by different poses of articulated objects (or face expressions) are isometric; i.e. they do not change as the pose changes. However, this is not true for all parts of the body. Recent work [106] dealt with learning the expression deformations among faces and pose deformations between synthetic 3D human meshes varying in pose. To model richer deformations high-order constraints between corresponding points needed to be specified. However, learning the types of deformations is a chicken-and-egg problem. Deformation models lead to more accurate correspondences, but these models are learnt using a set of data already in correspondence! Interestingly enough there is very little work done on generating benchmarks with ground truth correspondences coming from real data, such as 3D scans. We expect such benchmarks to play an important role at building more accurate algorithms for finding correspondences as well as evaluating their performance. Matching articulated objects, such as humans, in arbitrary poses may inevitably lead to left-to-right mirroring of the correspondences due to the high degree of symmetry in the human body. This issue will be eliminated if there is temporal information among the 3D scans, such as in a human tracking scenario. In this work we have taken into account only the geometry of the 3D scans. However, we believe that appearance information can be used to refine the resulting correspondences.

Correspondences are at the core of many applications. Sparse correspondences can be used as an initialization to full registration techniques, 3D content-based retrieval applications, even pose estimation. Previous work has focused mainly on finding correspondences when the pose of articulated objects changes. We expect future work to deal more with shape and resolution changes. In this work, we have focused on finding correspondences among articulated objects of the same class. However, correspondences can be defined between objects of different classes as well; e.g. for the case of deformation transfer in 3D animation. In such scenarios, it may be useful to focus on making the connection between surface correspondences and perceptual correspondences. Most of the approaches so far for finding correspondences rely on relatively clean, watertight 3D meshes. We expect that as 3D data become more and more common, there will be an increased interest in finding correspondences between real, noisy 3D scans with missing data as well as matching data with self-intersections; e.g. a 3D scan of a human with his legs in contact.

2. 1D Measurements and Attributes from 3D Scans

In Chapter 4, we described a method for predicting 1D measurements from 3D scans. The key insight in this work is that we can get increased accuracy in predicting measurements, if instead of working with noisy 3D scans containing holes, we register the 3D scans with a template mesh by taking into account prior knowledge about the shape of the human body. The registration procedure fills in the holes in a meaningful way as well as brings all the captured 3D scans into correspondence. That in turn allows the use of machine

learning techniques for measurement prediction. We expect that this learning framework can reveal correlations of body shape with non-surface attributes of the human body, such as age, muscularity, *etc.* In addition, contrary to existing commercial systems, this combination of registration and learning allows measurement prediction from arbitrary poses (as long as the registration can be performed and the specific pose is in our training set). For the first time, we take into account features from multiple poses for measurement prediction. As expected, measurements are predicted better using a 3D scan in the pose where the measurement was taken. Moreover, we perform an extensive evaluation against state-of-art. We expect this study to serve as a baseline for measurement prediction from 3D data, particularly relevant nowadays that 3D acquisition devices become cheaper to acquire.

The main limitation of our method is that we rely heavily on the registration procedure mentioned above. Aiming at setting a form of a baseline when it comes to measurement prediction from 3D scans, we have used only high-resolution 3D scans and landmarks provided by the CAESAR dataset. However, further studies on measurement prediction using low-resolution noisy data coming from e.g. Microsoft Kinect [55], or commodity webcams, should be performed. We have, additionally, introduced a set of shape features that work well enough for predicting measurements. These shape features were initially defined manually based on our intuition and we only perform feature selection. We do not claim that they are the best shape features to use. We could potentially have used as an input to the feature selection framework generic features, such as mesh HOG [104], *etc.* Alternatively, we could have learnt the shape features to use for prediction, but that would have required a large quantity of training data. Finally, we assumed standard anthropometry to be our ground truth. Although direct measurements in CAESAR were consistent among different measurers, the measurers may have also been consistently biased in the way they measure.

What is a good representation of human body shape and what can the body shape in 3D tell us about the human? Body shape is a temporally evolving quantity, thus, measurements also change over time; e.g. during breathing, as we age, *etc.* We believe that new representations of human shape will need to be employed in the future. Potentially these shape representations should be parameterized by factors such as activity, *etc.* The first attempt to describe human shape as a temporally evolving quantity is described in Chapter 5. In the present work, we provide some form of backward compatibility of 3D representations of human body shape with conventional 1D measurements. We additionally regard our work as a small step towards predicting human attributes from 3D scans. Previous work has tried to infer human attributes from photos or human motion data [86]. A 3D scan or a sequence of 3D scans, with both the geometry and appearance taken into account, should be an even richer source of information about the emotional state, health and individual characteristics of a human. So far, the human bodies that we used for predicting measurements were in tight fitting clothing. Future work should address predicting measurements and attributes in scenarios that resemble more the real life; i.e. prediction directly from dressed human bodies [10, 43].

3. Data-driven animation of subtle body deformations

In Chapter 5, we described a data-driven approach for modeling subtle deformations of the human body due to breathing. Key elements of the presented approach are the use of high-resolution 3D scans in combination with a human body model to model pose and shape change during breathing. The human body model additionally facilitates synthesis of new human shapes, which is hard for anatomically inspired, physics-based models. Although previous work has reproduced breathing motions that involve posture change, we are the first ones to explicitly model posture change during breathing. The high-resolution 3D scans are essential for modeling the shape of the human body in detail. Because breathing is a temporally evolving process, we capture multiple 3D scans of each subject at various unknown instances of their breathing cycle. We show that supplementing these scans with concepts from the physiology of respiration about the timing of breathing yields reasonable animations of breathing 3D avatars. Expressing the timing of breathing as a function of lung volume allows us to reproduce arbitrary breathing patterns captured from real humans using a spirometer. Finally, we provide animators with compact and intuitive controls for generating various breathing types as a combination of chest and abdominal breathing, varying the speed and the intensity of the breathing cycle. Our user interface for animation allows, additionally, animation from arbitrary sequences of lung volumes applied on new 3D human bodies synthesized in real-time.

In this work we have undertaken the difficult task of modeling the temporal aspect of human body shape without having any temporal data! We expect that modeling the human body over time using data from a 4D whole body scanner will give more accurate shape representations. Our method relies heavily on the use of the SCAPE model. We have assumed that SCAPE models the human body accurately enough and that the residual shape deformations come only from breathing. That is a reasonable assumption for a person standing in the “A pose” with small posture variation during breathing. We have shown breathing animations driven from arbitrary changes in lung volume. However, we have implicitly made an assumption that breathing is symmetric, i.e. for the same volume inside the human body, the breathing deformations while breathing in are the same as the deformation while breathing out. Our goal was to model subtleties of the human body. Breathing is by definition a subtle deformation, but there are even more subtle ones; e.g. the deformation of the veins during heartbeats, swallowing, *etc.* In terms of geometry, we cannot capture that degree of detail. This effect might be achieved through appearance cues (texture). In this work, we focus mainly on modeling how the human shape changes over time in terms of geometry. Change in appearance; e.g. change in face color during breathing just after running [53, 101], is part of future work.

Modeling shape deformations that bring 3D avatars into life, such as breathing, sneezing, motion of veins while talking *etc.* is still in its infancy. We focused on modeling the shape change of the torso as well as the whole body posture change in the standing pose. However, breathing is a much more complex motion. Breathing should be synchronized with human body motion; e.g. running. In addition, breathing involves

changes on the face as well. Not only the facial expression changes while breathing, but also the timing of breathing should be synchronized with the speech of the animated character. Our current breathing model assumes human bodies in tight fitting clothing. For increased realism, future work should additionally model breathing under clothing. Apart from modeling temporal shape change due to vital deformations, another interesting direction of future work is modeling temporal shape changes such as aging, shape change during a weight loss diet, *etc.*

4. The Challenge: Data-driven whole human body modeling in 3D

Data-driven modeling of the human body in 3D promises to convert lifeless 3D human scans to lively and expressive digital avatars. Admittedly, a lot of progress has been made in the last few decades. Initial human body models were coarse, consisting of 3D geometric shape primitives. Early data-driven approaches for modeling the human body could only leverage 2D images and motion capture data. Nowadays, it is much easier to acquire 3D data, even using our personal mobile devices, or find synthetic 3D data online. The focus has shifted to modeling in detail subtle deformations in specific areas of the body as well as real-time user-driven animation. In addition, modeling is performed in terms of both 3D geometry and appearance.

However, modeling the human body in 3D is still a big endeavour. Despite the large amount of available data, we are limited to what aspects of human shape can be captured. Given current state-of-art technologies, whole body real-time range data acquisition is not option. However, it is through that type of acquisition that we can capture shape deformations that bring 3D scans into life, such as sneezing, coughing, laughing. Additionally, data capture remains challenging in real life scenarios, such as capturing dressed people in natural environments. In terms of modeling, we envision that future work will take a holistic approach in modeling the human body. Realistic modeling of humans should take into account interaction with the objects in the scene; e.g. body deformations when a person is sitting, as well as interaction with other humans.

APPENDIX A

Detailed Performance of Measurement Prediction from 3D Scans

Measurement	[3]	[44]	Our method			Allowable Error [35]
Acromial Height, Sit		8.99	7.75	edges, sit	limbs, stand	9
Ankle Circumference	7.55	6.59	6.19	deformations, stand	edges, stand	4
Arm Length Shoulder - Elbow	11.26	8.42	6.65	limbs, stand	edges, stand	6
Arm Length Shoulder - Wrist	11.67	10.42	10.05	limbs, stand	edges, stand	
Arm Length Spine - Wrist	13.19	13.40	11.87	girths, stand	limbs, stand	
Armscye Circumference		12.57	12.64	girths, sit	girths, stand	13
Bizygomatic Breadth		3.18	2.75	limbs, stand	deformations, stand	2
Chest Circumference under Bust		17.49	15.40	girths, stand	limbs, stand	16
Buttock-Knee Length		7.17	7.04	girths, sit	limbs, sit	6
Chest Circumference	12.43	13.02	12.73	girths, stand	edges, stand	15
Chest Girth at Scye		27.19	25.65	girths, stand	deformations, both	15
Crotch Height	7.45	7.53	5.50	limbs, stand	deformations, stand	10
Elbow Height, Sit		9.80	8.21	deformations, both	limbs, stand	10
Eye Height, Sit		9.81	8.52	limbs, sit	deformations, both	8
Face Length		3.37	3.20	deformations, both	deformations, stand	
Foot Length		4.74	4.05	limbs, stand	edges, stand	3
Hand Circumference		5.22	4.95	deformations, stand	edges, stand	4
Hand Length		4.23	4.17	limbs, stand	edges, stand	3

FIGURE A.1. Measurement prediction on females (part I). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold.

Measurement	[3]	[44]	Our method			Allowable Error [35]
Head Breadth		3.80	3.05	girths, stand	deformations, stand	2
Head Circumference	7.44	7.45	5.87	girths, sit	limbs, stand	5
Head Length		4.66	3.87	edges, sit	girths, sit	3
Hip Breadth, Sit		11.42	10.22	edges, sit	girths, sit	6
Hip Circ Max Height	17.05	18.96	18.59	girths, stand	limbs, stand	
Hip Circumference, Maximum	7.47	16.15	12.35	girths, stand	edges, stand	12
Knee Height		7.42	7.01	girths, stand	limbs, stand	6
Neck Base Circumference	21.03	16.96	15.79	limbs, sit		11
Shoulder Breadth		16.57	17.37	girths, stand	limbs, stand	
Height, Sit		8.87	6.75	deformations, both	limbs, stand	6
Spine-to-Elbow		8.41	8.10	girths, stand	limbs, stand	
Spine-to-Shoulder		7.60	6.59	edges, sit	edges, stand	
Stature	5.60	10.21	7.51	girths, stand	limbs, stand	10
Subscapular Skinfold		4.30	4.06	girths, stand	deformations, stand	
Thigh Circumference		10.63	8.17	girths, stand	edges, stand	6
Thigh Circumference Max, Sit		14.19	12.27	girths, stand	limbs, stand	
Thumb Tip Reach		18.18	16.70	deformations, both	limbs, stand	20
Total Crotch Length		24.14	19.34	girths, stand	edges, stand	
Triceps Skinfold		4.61	4.37	edges, sit	girths, stand	
Vertical Trunk Circumference		17.34	15.12	limbs, stand	edges, stand	23
Waist Circumference, Preferred		24.73	22.79	girths, stand	edges, stand	
Waist Front Length		19.41	17.24	limbs, stand	edges, stand	5
Waist Height, Preferred		12.28	10.71	limbs, stand	edges, stand	
Weight		1.24	1.08	edges, sit	girths, stand	0.3

FIGURE A.2. Measurement prediction on females (part II). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold.

Measurement			Our method			Allowable Error [35]
	[3]	[44]				
Acromial Height, Sit		9.93	10.45	limbs, sit	limbs, stand	9
Ankle Circumference	13.66	5.72	5.56	edges, stand	girths, stand	4
Arm Length Shoulder - Elbow	13.99	12.66	13.32	edges, stand	limbs, stand	6
Arm Length Shoulder - Wrist	14.49	13.76	12.66	edges, stand	limbs, stand	
Arm Length Spine - Wrist	14.71	11.81	10.40	edges, stand	limbs, stand	
Armscye Circumference		9.86	8.88	edges, stand	girths, stand	13
Bizygomatic Breadth		3.25	3.02	girths, stand	deformations, both	2
Buttock-Knee Length		8.66	8.17	limbs, stand	deformations, both	6
Chest Circumference	13.96	15.21	13.19	girths, stand	deformations, both	15
Chest Girth at Sycye		17.57	16.92	girths, stand	deformations, both	15
Crotch Height	11.01	9.77	8.36	deformations, stand	limbs, stand	10
Elbow Height, Sit		14.50	12.33	girths, sit	limbs, stand	10
Eye Height, Sit		11.86	11.31	limbs, sit	deformations, both	8
Face Length		3.72	3.63	deformations, stand	girths, sit	
Foot Length		4.60	4.64	edges, stand	edges, sit	3
Hand Circumference		4.76	4.51	girths, stand	deformations, stand	4
Hand Length		3.98	3.55	edges, stand	limbs, stand	3
Head Breadth		3.73	3.25	edges, sit	deformations, stand	2
Head Circumference	5.51	6.51	5.42	girths, sit		5
Head Length		3.68	3.36	edges, stand	girths, stand	3
Hip Breadth, Sit		8.54	6.82	edges, sit	limbs, stand	6
Hip Circ Max Height	16.50	18.89	19.05	edges, stand	limbs, stand	
Hip Circumference, Maximum	7.90	12.57	10.66	edges, stand	girths, stand	12

FIGURE A.3. Measurement prediction on males (part I). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold.

Measurement	[3]	[44]	Our method			Allowable Error [35]
Knee Height		5.79	4.42	edges, stand	limbs, stand	6
Neck Base Circumference	21.57	13.33	13.47	deformations, stand	limbs, stand	11
Shoulder Breadth		9.76	8.94	edges, stand	girths, stand	
Height, Sit		10.34	8.37	limbs, stand	deformations, both	6
Spine-to-Elbow		12.22	10.35	edges, stand	limbs, stand	
Spine-to-Shoulder		8.24	7.72	edges, stand	girths, stand	
Stature	5.86	7.98	6.53	edges, stand	limbs, stand	10
Subscapular Skinfold		5.89	5.65	deformations, stand	girths, stand	
Thigh Circumference		11.72	8.61	edges, stand	girths, stand	6
Thigh Circumference Max, Sit		11.56	9.57	edges, stand	girths, stand	
Thumb Tip Reach		18.63	17.99	girths, stand	limbs, stand	20
Total Crotch Length		26.79	24.13	edges, stand	girths, stand	
Triceps Skinfold		3.78	3.59	deformations, stand	deformations, both	
Vertical Trunk Circumference		25.46	23.18	girths, stand	limbs, stand	23
Waist Circumference, Preferred		22.90	21.92	girths, stand	deformations, both	
Waist Front Length		20.63	18.97	edges, stand	limbs, stand	5
Waist Height, Preferred		13.72	10.73	edges, stand	limbs, stand	
Weight		1.87	1.91	girths, stand	limbs, stand	0.3

FIGURE A.4. Measurement prediction on males (part II). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold.

1. Measurement prediction from the standing pose

Measurement			Our method			Allowable Error [35]
	[3]	[44]				
Acromial Height, Sit		10.29	10.04	girths, stand	limbs, stand	9
Ankle Circumference	7.55	6.59	6.19	deformations, stand	edges, stand	4
Arm Length Shoulder - Elbow	11.26	8.42	6.65	limbs, stand	edges, stand	6
Arm Length Shoulder - Wrist	11.67	10.42	10.05	limbs, stand	edges, stand	
Arm Length Spine - Wrist	13.19	13.40	11.87	girths, stand	limbs, stand	
Armscye Circumference		12.57	12.22	girths, stand		13
Bizygomatic Breadth		3.18	2.75	limbs, stand	deformations, stand	2
Chest Circumference under Bust		17.49	15.40	girths, stand	limbs, stand	16
Buttock-Knee Length		9.02	8.52	girths, stand	limbs, stand	6
Chest Circumference	12.43	13.02	12.73	girths, stand	edges, stand	15
Chest Girth at Scye		27.19	27.31	girths, stand	edges, stand	15
Crotch Height	7.45	7.53	5.50	limbs, stand	deformations, stand	10
Elbow Height, Sit		10.29	10.06	girths, stand	limbs, stand	10
Eye Height, Sit		9.81	9.96	girths, stand	limbs, stand	8
Face Length		3.37	2.97	girths, stand	deformations, stand	
Foot Length		4.74	4.05	limbs, stand	edges, stand	3
Hand Circumference		5.22	4.95	deformations, stand	edges, stand	4
Hand Length		4.23	4.17	limbs, stand	edges, stand	3

FIGURE A.5. Measurement prediction on females using only data in the standing pose (part I). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold.

Measurement	[3]	[44]	Our method			Allowable Error [35]
Head Breadth		3.80	3.05	girths, stand	deformations, stand	2
Head Circumference	7.44	7.45	5.91	girths, stand	edges, stand	5
Head Length		4.23	4.17	deformations, stand	edges, stand	3
Hip Breadth, Sit		14.07	12.28	edges, stand	girths, stand	6
Hip Circ Max Height	17.05	18.96	18.59	girths, stand	limbs, stand	
Hip Circumference, Maximum	7.47	16.15	12.35	girths, stand	edges, stand	12
Knee Height		7.42	7.01	girths, stand	limbs, stand	6
Neck Base Circumference	21.03	16.35	15.43	limbs, stand		11
Shoulder Breadth		16.57	17.37	girths, stand	limbs, stand	
Height, Sit		8.87	8.52	girths, stand	limbs, stand	6
Spine-to-Elbow		8.41	8.10	girths, stand	limbs, stand	
Spine-to-Shoulder		7.60	6.60	limbs, stand	edges, stand	
Stature	5.60	10.21	7.51	girths, stand	limbs, stand	10
Subscapular Skinfold		4.30	4.06	girths, stand	deformations, stand	
Thigh Circumference		10.63	8.17	girths, stand	edges, stand	6
Thigh Circumference Max, Sit		14.19	12.27	girths, stand	limbs, stand	
Thumb Tip Reach		18.18	15.56	deformations, stand	limbs, stand	20
Total Crotch Length		24.14	19.34	girths, stand	edges, stand	
Triceps Skinfold		4.61	4.58	limbs, stand	girths, stand	
Vertical Trunk Circumference		17.34	15.12	limbs, stand	edges, stand	23
Waist Circumference, Preferred		24.73	22.79	girths, stand	edges, stand	
Waist Front Length		19.41	17.24	limbs, stand	edges, stand	5
Waist Height, Preferred		12.28	10.71	limbs, stand	edges, stand	
Weight		1.24	1.15	edges, stand	girths, stand	0.3

FIGURE A.6. Measurement prediction on females using only data in the standing pose (part II). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold.

Measurement			Our method			Allowable Error [35]
	[3]	[44]				
Acromial Height, Sit		11.63	11.25	edges, stand	limbs, stand	9
Ankle Circumference	13.66	5.72	5.56	edges, stand	girths, stand	4
Arm Length Shoulder - Elbow	13.99	12.66	13.32	edges, stand	limbs, stand	6
Arm Length Shoulder - Wrist	14.49	13.76	12.66	edges, stand	limbs, stand	
Arm Length Spine - Wrist	14.71	11.81	10.40	edges, stand	limbs, stand	
Armscye Circumference		9.86	8.88	edges, stand	girths, stand	13
Bizygomatic Breadth		3.75	3.02	girths, stand	deformations, stand	2
Buttock-Knee Length		8.91	9.25	limbs, stand	girths, stand	6
Chest Circumference	13.96	15.21	13.02	girths, stand	deformations, stand	15
Chest Girth at Seye		17.57	16.92	girths, stand	deformations, both	15
Crotch Height	11.01	9.77	8.36	deformations, stand	limbs, stand	10
Elbow Height, Sit		14.50	13.82	girths, stand	limbs, stand	10
Eye Height, Sit		12.51	11.88	limbs, stand	deformations, stand	8
Face Length		3.72	3.72	deformations, stand	girths, stand	
Foot Length		4.60	4.57	edges, stand	limbs, stand	3
Hand Circumference		4.76	4.51	girths, stand	deformations, stand	4
Hand Length		3.98	3.55	edges, stand	limbs, stand	3
Head Breadth		3.73	3.30	edges, stand	deformations, stand	2
Head Circumference	5.51	7.46	5.59	girths, stand	edges, stand	5
Head Length		3.68	3.36	edges, stand	girths, stand	3
Hip Breadth, Sit		8.68	7.11	edges, stand	girths, stand	6
Hip Circ Max Height	16.50	18.89	19.05	edges, stand	limbs, stand	
Hip Circumference, Maximum	7.90	12.57	10.66	edges, stand	girths, stand	12

FIGURE A.7. Measurement prediction on males using only data in the standing pose (part I). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold.

Measurement	[3]	[44]	Our method			Allowable Error [35]
Knee Height		5.79	4.42	edges, stand	limbs, stand	6
Neck Base Circumference	21.57	13.33	13.47	deformations, stand	limbs, stand	11
Shoulder Breadth		9.76	8.94	edges, stand	girths, stand	
Height, Sit		10.34	9.18	limbs, stand	deformations, stand	6
Spine-to-Elbow		12.22	10.35	edges, stand	limbs, stand	
Spine-to-Shoulder		8.24	7.72	edges, stand	girths, stand	
Stature	5.86	7.98	6.53	edges, stand	limbs, stand	10
Subscapular Skinfold		5.89	5.65	deformations, stand	girths, stand	
Thigh Circumference		11.72	8.61	edges, stand	girths, stand	6
Thigh Circumference Max, Sit		11.56	9.57	edges, stand	girths, stand	
Thumb Tip Reach		18.63	17.99	girths, stand	limbs, stand	20
Total Crotch Length		26.79	24.13	edges, stand	girths, stand	
Triceps Skinfold		3.78	3.60	deformations, stand	girths, stand	
Vertical Trunk Circumference		25.46	23.18	girths, stand	limbs, stand	23
Waist Circumference, Preferred		22.90	23.33	girths, stand	edges, stand	
Waist Front Length		20.63	18.97	edges, stand	limbs, stand	5
Waist Height, Preferred		13.72	10.73	edges, stand	limbs, stand	
Weight		1.87	1.91	girths, stand	limbs, stand	0.3

FIGURE A.8. Measurement prediction on males using only data in the standing pose (part II). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold.

2. Measurement prediction from the seated pose

Measurement	[44]	Our method			Allowable Error [35]
Acromial Height, Sit	8.99	8.03	edges, sit	limbs, sit	9
Ankle Circumference	7.00	6.36	edges, sit	limbs, sit	4
Arm Length Shoulder - Elbow	8.09	6.66	edges, sit	limbs, sit	6
Arm Length Shoulder - Wrist	10.89	10.49	edges, sit	limbs, sit	
Arm Length Spine - Wrist	13.16	12.34	edges, sit	limbs, sit	
Armscye Circumference	13.12	13.30	girths, sit		13
Bizygomatic Breadth	3.18	2.79	edges, sit	limbs, sit	2
Chest Circumference under Bust	19.64	18.22	girths, sit	limbs, sit	16
Buttock-Knee Length	7.17	7.04	girths, sit	limbs, sit	6
Chest Circumference	13.48	13.05	girths, sit	edges, sit	15
Chest Girth at Scye	28.45	28.39	girths, sit	edges, sit	15
Crotch Height	11.03	9.66	edges, sit	limbs, sit	10
Elbow Height, Sit	9.80	8.76	edges, sit	limbs, sit	10
Eye Height, Sit	8.87	8.77	edges, sit	girths, sit	8
Face Length	3.74	3.54	edges, sit	limbs, sit	
Foot Length	5.55	5.35	edges, sit		3
Hand Circumference	5.91	5.84	edges, sit	limbs, sit	4
Hand Length	5.33	5.05	edges, sit	limbs, sit	3

FIGURE A.9. Measurement prediction on females using only data in the seated pose (part I). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold. Anthroscan [3] cannot predict measurements from scans in the seated pose.

Measurement	[44]	Our method		Allowable Error [35]
Head Breadth	3.76	3.08	edges, sit limbs, sit	2
Head Circumference	7.45	5.67	girths, sit limbs, sit	5
Head Length	4.66	3.87	edges, sit girths, sit	3
Hip Breadth, Sit	11.42	10.22	edges, sit girths, sit	6
Hip Circ Max Height	23.72	22.26	girths, sit limbs, sit	
Hip Circumference, Maximum	17.34	15.48	girths, sit edges, sit	12
Knee Height	7.47	6.25	edges, sit limbs, sit	6
Neck Base Circumference	16.96	15.79	limbs, sit	11
Shoulder Breadth	17.60	17.52	edges, sit girths, sit	
Height, Sit	7.70	6.17	edges, sit limbs, sit	6
Spine-to-Elbow	8.41	7.80	girths, sit limbs, sit	
Spine-to-Shoulder	7.65	6.56	limbs, sit edges, sit	
Stature	11.40	10.28	edges, sit limbs, sit	10
Subscapular Skinfold	4.30	3.90	edges, sit girths, sit	
Thigh Circumference	12.56	12.28	girths, sit edges, sit	6
Thigh Circumference Max, Sit	15.23	14.58	girths, sit edges, sit	
Thumb Tip Reach	17.57	17.21	edges, sit limbs, sit	20
Total Crotch Length	30.87	30.63	limbs, sit edges, sit	
Triceps Skinfold	4.49	4.40	edges, sit girths, sit	
Vertical Trunk Circumference	19.16	17.44	limbs, sit edges, sit	23
Waist Circumference, Preferred	29.59	27.36	girths, sit edges, sit	
Waist Front Length	21.13	20.49	limbs, sit edges, sit	5
Waist Height, Preferred	16.27	15.82	limbs, sit edges, sit	
Weight	1.40	1.14	edges, sit girths, sit	0.3

FIGURE A.10. Measurement prediction on females using only data in the seated pose (part II). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold. Anthroscan [3] cannot predict measurements from scans in the seated pose.

Measurement	[44]	Our method		Allowable Error [35]
Acromial Height, Sit	9.93	9.82	edges, sit limbs, sit	9
Ankle Circumference	5.79	5.22	edges, sit girths, sit	4
Arm Length Shoulder - Elbow	13.83	13.07	edges, sit limbs, sit	6
Arm Length Shoulder - Wrist	12.75	12.29	edges, sit limbs, sit	
Arm Length Spine - Wrist	11.61	11.26	edges, sit limbs, sit	
Armscye Circumference	10.58	9.25	edges, sit girths, sit	13
Bizygomatic Breadth	3.25	3.14	edges, sit girths, sit	2
Buttock-Knee Length	8.66	8.42	girths, sit limbs, sit	6
Chest Circumference	15.40	13.86	edges, sit girths, sit	15
Chest Girth at Scye	17.34	16.49	edges, sit girths, sit	15
Crotch Height	14.19	14.54	edges, sit limbs, sit	10
Elbow Height, Sit	11.60	11.48	edges, sit limbs, sit	10
Eye Height, Sit	11.86	11.11	edges, sit limbs, sit	8
Face Length	4.28	4.55	edges, sit girths, sit	
Foot Length	4.66	4.28	edges, sit limbs, sit	3
Hand Circumference	5.48	5.34	girths, sit limbs, sit	4
Hand Length	4.86	4.60	edges, sit limbs, sit	3
Head Breadth	3.95	3.16	edges, sit girths, sit	2
Head Circumference	6.51	5.42	girths, sit	5
Head Length	3.68	3.17	edges, sit	3
Hip Breadth, Sit	8.54	6.74	edges, sit girths, sit	6
Hip Circ Max Height	21.41	20.84	edges, sit limbs, sit	
Hip Circumference, Maximum	13.84	13.68	edges, sit girths, sit	12

FIGURE A.11. Measurement prediction on males using only data in the seated pose (part I). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Anthroscan [3], Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold. Anthroscan [3] cannot predict measurements from scans in the seated pose.

Measurement			Our method		Allowable Error [35]
	[44]				
Knee Height	5.41	5.11	edges, sit	girths, sit	6
Neck Base Circumference	13.33	13.47	girths, sit	limbs, sit	11
Shoulder Breadth	13.33	12.14	girths, sit	limbs, sit	
Height, Sit	9.27	8.50	edges, sit	limbs, sit	6
Spine-to-Elbow	12.22	11.04	edges, sit	limbs, sit	
Spine-to-Shoulder	7.78	7.24	girths, sit	limbs, sit	
Stature	10.59	8.99	edges, sit	limbs, sit	10
Subscapular Skinfold	5.81	5.52	edges, sit	girths, sit	
Thigh Circumference	15.41	13.36	edges, sit	girths, sit	6
Thigh Circumference Max, Sit	13.10	12.34	edges, sit	girths, sit	
Thumb Tip Reach	18.63	19.35	girths, sit	limbs, sit	20
Total Crotch Length	35.28	31.80	edges, sit	girths, sit	
Triceps Skinfold	3.72	3.45	girths, sit		
Vertical Trunk Circumference	26.18	27.86	edges, sit	girths, sit	23
Waist Circumference, Preferred	24.96	24.20	edges, sit	girths, sit	
Waist Front Length	24.09	20.79	edges, sit	limbs, sit	5
Waist Height, Preferred	17.38	15.69	edges, sit	limbs, sit	
Weight	1.72	1.72	edges, sit	girths, sit	0.3

FIGURE A.12. Measurement prediction on males using only data in the seated pose (part II). Mean Absolute Difference (MAD) between extracted and direct measurements evaluated on a test set of 200 bodies from CAESAR [77] respectively. We compare Hasler et al. [44] and our method against the allowable inter-observer error per measurement as reported in the ANSUR study [35]. For our method, we additionally show the optimal groups of features per measurement. The reported values are in mm. Measurement prediction with no statistical significance between our method and at least one of the other approaches is denoted with smaller font size. The best performance per measurement is shown in bold. Anthroscan [3] cannot predict measurements from scans in the seated pose.

Bibliography

- [1] <http://www.museoffire.com/Lyceum/Dojo/Tutorials/Breathing/t3dab6.htm>.
- [2] <http://www.packtpub.com/article/blender-25-rigging-torso>.
- [3] Human Solutions GmbH. Anthroscan. <http://www.human-solutions.com/fashion/>.
- [4] B. Allen, B. Curless, and Z. Popović. The space of human body shapes: Reconstruction and parameterization from range scans. *ACM Transactions on Graphics (TOG)*, 22(3):587–594, 2003.
- [5] B. Allen, B. Curless, Z. Popović, and A. Hertzmann. Learning a correlated model of identity and pose-dependent body shape variation for real-time synthesis. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA)*, pages 147–156, 2006.
- [6] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis. SCAPE: Shape Completion and Animation of PEople. *ACM Transactions on Graphics (TOG)*, 24(3):408–416, 2005.
- [7] D. Anguelov, P. Srinivasan, H. Pang, D. Koller, S. Thrun, and J. Davis. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *Neural Information Processing Systems (NIPS)*, pages 33–40. The MIT Press, 2005.
- [8] M. Aubry, U. Schlickewei, and D. Cremers. The Wave Kernel Signature: A quantum mechanical approach to shape analysis. In *International Conference on Computer Vision (ICCV) Workshops*, pages 1626–1633. IEEE, 2011.
- [9] S. Baek and K. Lee. Parametric human body shape modeling framework for human-centered product design. *Computer-Aided Design (CAD)*, 44(1):56–67, 2012.
- [10] A. O. Bălan and M. J. Black. The naked truth: Estimating body shape under clothing. In *European Conference on Computer Vision (ECCV)*, pages 15–29. Springer, 2008.
- [11] A. O. Balan, L. Sigal, M. J. Black, J. E. Davis, and H. W. Haussecker. Detailed human shape and pose from images. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE, 2007.
- [12] Z. Ben Azouz, M. Rioux, C. Shu, and R. Lepage. Characterizing human shape variation using 3-D anthropometric data. *The Visual Computer: International Journal of Computer Graphics*, 22(5):302–314, 2005.

- [13] P. Besl and N. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, pages 239–256, 1992.
- [14] M. J. Black, A. Balan, A. Weiss, L. Sigal, M. Loper, and T. St Clair. Method and apparatus for estimating body shape. US (12/541,898) and PCT patent application, August 2009.
- [15] F. Bogo, J. Romero, and M. J. Black. FAUST: Dataset and evaluation for 3D mesh registration. In *International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [16] J. Bougourd and P. Treleaven. Capturing the shape of a nation: size UK. In *International Federation of the Fashion Technology Institute (IFFTI)*. American Standards for Testing and Materials, 2002.
- [17] B. Bradtmiller, M. Gross, B. Bradtmiller-Anthrotech, and M. Gross-Anthrotech. 3D whole body scans: Measurement extraction software validation. In *International Conference on Digital Human Modeling for Design and Engineering (DHM)*, 1999.
- [18] A. Bronstein, M. Bronstein, M. Bronstein, and R. Kimmel. *Numerical geometry of non-rigid shapes*. Springer-Verlag, 2008.
- [19] A. Bronstein, M. Bronstein, and R. Kimmel. Generalized Multidimensional Scaling: a framework for isometry-invariant partial surface matching. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 103(5):1168, 2006.
- [20] A. Bronstein, M. Bronstein, R. Kimmel, M. Mahmoudi, and G. Sapiro. A Gromov-Hausdorff framework with diffusion geometry for topologically-robust non-rigid shape matching. *International Journal of Computer Vision (IJCV)*, 89(2):266–286, 2010.
- [21] A. M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov. Shape Google: Geometric words and expressions for invariant shape retrieval. *ACM Transactions on Graphics (TOG)*, 30(1):1:1–1:20, February 2011.
- [22] M. Cavagnaro, E. Pittella, and S. Pisa. Anatomical models of breathing subjects for absorption and scattering analysis. In *International Symposium on Electromagnetic Compatibility (EMC EUROPE)*, pages 149–153. IEEE, 2013.
- [23] Y. Chen, Z. Liu, and Z. Zhang. Tensor-based human body modeling. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 105–112. IEEE, 2013.
- [24] Y. Chen, D. Robertson, and R. Cipolla. A practical system for modelling body shapes from single view measurements. In *British Machine Vision Conference (BMVC)*, pages 82–91, 2011.
- [25] D. Cosker and J. Edge. Laughing, crying, sneezing and yawning: Automatic voice driven animation of non-speech articulations. Citeseer, 2009.
- [26] F. De la Torre and M. J. Black. A framework for robust subspace learning. *International Journal of Computer Vision (IJCV)*, 54(1–3):117–142, Aug.–Oct. 2003.

- [27] P. C. DiLorenzo, V. B. Zordan, and B. L. Sanders. Laughing out loud: Control for modeling anatomically inspired laughter using audio. *ACM Transactions on Graphics (TOG)*, 27(5):125–133, 2008.
- [28] P. Dollar, C. Wojek, B. Schiele, and P. Perona. Pedestrian detection: An evaluation of the state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 34(4):743–761, 2012.
- [29] A. Dubrovina and R. Kimmel. Matching shapes by eigendecomposition of the Laplace-Beltrami operator. In *International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, volume 2, 2010.
- [30] A. Egges, T. Giacomo, and N. Thalmann. Synthesis of realistic idle motion for interactive characters. *Game Programming Gems*, 6:3–2, 2006.
- [31] A. Egges, T. Molet, and N. Magnenat-Thalmann. Personalised real-time idle motion synthesis. In *Pacific Conference on Computer Graphics and Applications (PG)*, pages 121–130. IEEE, 2004.
- [32] Y. Eldar, M. Lindenbaum, M. Porat, and Y. Zeevi. The farthest point strategy for progressive image sampling. *Transactions on Image Processing*, 6(9):1305–1315, 2002.
- [33] J. Gall, C. Stoll, E. De Aguiar, C. Theobalt, B. Rosenhahn, and H. Seidel. Motion capture using joint skeleton tracking and surface estimation. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1746–1753. IEEE, 2009.
- [34] M. Garland and P. Heckbert. Surface simplification using quadric error metrics. In *Annual Conference on Computer Graphics and Interactive Techniques*, pages 209–216. ACM Press/Addison-Wesley Publishing Co., 1997.
- [35] C. Gordon, T. Churchill, C. Clauser, B. Bradtmiller, and J. McConville. Anthropometric survey of US Army personnel: Methods and summary statistics 1988. Technical report, DTIC Document, 1989.
- [36] P. Guan, O. Freifeld, and M. J. Black. A 2D human body model dressed in eigen clothing. In *European Conference on Computer Vision (ECCV)*, pages 285–298. Springer, 2010.
- [37] P. Guan, L. Reiss, D. A. Hirshberg, A. Weiss, and M. J. Black. DRAPE: Dressing any person. *ACM Transactions on Graphics (TOG)*, 31(4):35, 2012.
- [38] P. Guan, L. Sigal, V. Reznitskaya, and J. K. Hodgins. Multi-linear data-driven dynamic hair model with efficient hair-body collision handling. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA)*, pages 295–304. Eurographics Association, 2012.
- [39] P. Guan, A. Weiss, A. Balan, and M. J. Black. Estimating human shape and pose from a single image. In *International Conference on Computer Vision (ICCV)*, pages 1381–1388, February 2009.
- [40] P. Guan, A. Weiss, A. O. Balan, and M. J. Black. Estimating human shape and pose from a single image. In *International Conference on Computer Vision (ICCV)*, pages 1381–1388. IEEE, 2009.
- [41] Y. Guan and J. Dy. Sparse Probabilistic Principal Component Analysis. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 5, pages 185–192, 2009.

- [42] D. R. Hardoon, S. Szedmak, and J. Shawe-Taylor. Canonical Correlation Analysis: An overview with application to learning methods. *Neural Computation*, 16(12):2639–2664, 2004.
- [43] N. Hasler, H. Ackermann, B. Rosenhahn, T. Thormahlen, and H.-P. Seidel. Multilinear pose and body shape estimation of dressed subjects from image sets. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1823–1830. IEEE, 2010.
- [44] N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn, and H. Seidel. A statistical model of human pose and body shape. *Computer Graphics Forum*, 28(2):337–346, 2009.
- [45] D. Hirshberg, M. Loper, E. Rachlin, and M. Black. Coregistration: Simultaneous alignment and modeling of articulated 3D shape. In *European Conference on Computer Vision (ECCV)*, volume 7577, pages 242–255. Springer, 2012.
- [46] E. Hsu, S. Gentry, and J. Popović. Example-based Control of Human Motion. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA)*, pages 69–77. Eurographics Association, 2004.
- [47] Q. Huang, B. Adams, M. Wicke, and L. Guibas. Non-rigid registration under isometric deformations. In *Proceedings of the Symposium on Geometry Processing*, pages 1449–1457. Eurographics Association, 2008.
- [48] N. Iyer, S. Jayanti, K. Lou, Y. Kalyanaraman, and K. Ramani. Three-dimensional shape searching: state-of-the-art review and future trends. *Computer Aided Design*, 37(5):509–530, 2005.
- [49] A. Jain, T. Thormählen, H.-P. Seidel, and C. Theobalt. MovieReshape: Tracking and reshaping of humans in videos. *ACM Transactions on Graphics (TOG)*, 29(5), 2010.
- [50] A. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 21(5):433–449, 2002.
- [51] I. Jolliffe. *Principal Component Analysis*. Wiley Online Library, 2005.
- [52] M. Kazhdan, M. Bolitho, and H. Hoppe. Poisson surface reconstruction. In *Eurographics Symposium on Geometry Processing*, 2006.
- [53] J. Kider, K. Pollock, and A. Safonova. A data-driven appearance model for human fatigue. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA)*, pages 119–128. ACM, 2011.
- [54] R. Kimmel and J. Sethian. Computing geodesic paths on manifolds. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 95(15):8431, 1998.
- [55] KINECT. Microsoft Corporation. <http://www.xbox.com/kinect>.
- [56] J. Lafferty, A. McCallum, and F. C. Pereira. Conditional Random Fields: Probabilistic models for segmenting and labeling sequence data. 2001.

- [57] D. Lee, M. Glueck, A. Khan, E. Fiume, and K. Jackson. Modeling and simulation of skeletal muscle for computer graphics: A survey. *Foundations and Trends in Computer Graphics and Vision*, 7(4):229–276, 2012.
- [58] S.-H. Lee, E. Sifakis, and D. Terzopoulos. Comprehensive biomechanical modeling and simulation of the upper body. *ACM Transactions on Graphics (TOG)*, 28(4):99, 2009.
- [59] Y. Lipman and T. Funkhouser. Möbius voting for surface correspondence. *ACM Transactions on Graphics (TOG)*, 28(3):1–12, 2009.
- [60] L. Luo, H. Li, and S. Rusinkiewicz. Structure-aware hair capture. *ACM Transactions on Graphics (TOG)*, 32(4):76, 2013.
- [61] N. Magnenat-Thalmann, J. J. Zhang, and D. D. Feng. *Recent Advances in the 3D Physiological Human*. Springer, 2009.
- [62] L. Mckinnon and C. Istook. Comparative analysis of the Image Twin system and the 3T6 body scanner. *Journal of Textile and Apparel, Technology and Management*, 1(2):1–7, 2001.
- [63] A. H. Mines. *Respiratory physiology*. Raven Press, 1981.
- [64] T. B. Moeslund, A. Hilton, and V. Krüger. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2):90–126, 2006.
- [65] A. Moreno, S. Chambon, A. P. Santhanam, R. Brocardo, P. Kupelian, J. P. Rolland, E. Angelini, and I. Bloch. Thoracic CT-PET registration using a 3D breathing model. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 626–633. Springer, 2007.
- [66] T. Neumann, K. Varanasi, N. Hasler, M. Wacker, M. Magnor, and C. Theobalt. Capture and statistical modeling of arm-muscle deformations. *Computer Graphics Forum*, 32:285–294, 2013.
- [67] E. Paquet, M. Rioux, A. Murching, T. Naveen, and A. Tabatabai. Description of shape information for 2-D and 3-D objects. *Signal Processing: Image Communication*, 16(1):103–122, 2000.
- [68] S. Paquette, J. Brantley, B. Corner, P. Li, and T. Oliver. Automated extraction of anthropometric data from 3D images. In *Human Factors and Ergonomics Society Annual Meeting*, volume 44, pages 727–730. SAGE Publications, 2000.
- [69] M. J. Park and S. Y. Shin. Example-based motion cloning. *Computer Animation and Virtual Worlds*, 15(3-4):245–257, 2004.
- [70] S. I. Park and J. K. Hodgins. Capturing and animating skin deformation in human motion. *ACM Transactions on Graphics (TOG)*, 25(3):881–889, 2006.
- [71] S. I. Park and J. K. Hodgins. Data-driven modeling of skin and muscle deformation. *ACM Transactions on Graphics (TOG)*, 27(3):96, 2008.
- [72] J. Penne, C. Schaller, J. Hornegger, and T. Kuwert. Robust real-time 3D respiratory motion detection using Time-of-Flight cameras. *International Journal of Computer Assisted Radiology and Surgery*,

- 3(5):427–431, 2008.
- [73] G. Pons-Moll, J. Taylor, J. Shotton, A. Hertzmann, and A. Fitzgibbon. Metric regression forests for human pose estimation. In *British Machine Vision Conference (BMVC)*, September 2013.
- [74] E. Promayon, P. Baconnier, and C. Puech. Physically-based model for simulating the human trunk respiration movements. In *Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery (CVRMed-MRCAS)*, pages 379–388. Springer, 1997.
- [75] M. P. Reed and M. B. Parkinson. Modeling variability in torso shape for chair and seat design. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC/CIE)*, pages 561–569. American Society of Mechanical Engineers, 2008.
- [76] D. Rey and M. Neuhäuser. Wilcoxon-signed-rank test. In *International Encyclopedia of Statistical Science*, pages 1658–1659. Springer, 2011.
- [77] K. Robinette, S. Blackwell, H. Daanen, M. Boehmer, S. Fleming, T. Brill, D. Hoeflerlin, and D. Burnsides. Civilian American and European Surface Anthropometry Resource (CAESAR) Final Report. Technical Report AFRL-HE-WP-TR-2002-0169, US Air Force Research Laboratory, 2002.
- [78] K. Robinette, H. Daanen, and E. Paquet. The CAESAR Project: a 3-D Surface Anthropometry Survey. In *International Conference on 3-D Digital Imaging and Modeling*, pages 380–386. IEEE, 1999.
- [79] K. M. Robinette and H. A. Daanen. Precision of the CAESAR scan-extracted measurements. *Applied Ergonomics*, 37(3):259–265, 2006.
- [80] F. J. Rohlf and D. Slice. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Biology*, 39(1):40–59, 1990.
- [81] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *International Conference on 3D Digital Imaging and Modeling (3DIM)*, page 145. IEEE, 2001.
- [82] B. Sanders, P. DiLorenzo, V. Zordan, and D. Bakal. Toward anatomical simulation for breath training in mind/body medicine. In *Recent Advances in the 3D Physiological Human*, pages 105–119. Springer, 2009.
- [83] Z. Sánta and Z. Kato. Correspondence-less non-rigid registration of triangular surface meshes. In *International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2275–2282. IEEE, June 2013.
- [84] A. Santhanam, S. Pattanaik, J. Rolland, C. Imielinska, and J. Norfleet. Physiologically-based modeling and visualization of deformable lungs. In *Pacific Conference on Computer Graphics and Applications (PG)*, pages 507–511, Oct. 2003.
- [85] H. Seo, Y. I. Yeo, and K. Wohn. 3D body reconstruction from photos based on range scan. In *Technologies for E-Learning and Digital Entertainment*, pages 849–860, 2006.

- [86] L. Sigal, D. J. Fleet, N. F. Troje, and M. Livne. Human attributes from 3D pose tracking. In *European Conference on Computer Vision (ECCV)*, pages 243–257. Springer, 2010.
- [87] C. Stoll, J. Gall, E. De Aguiar, S. Thrun, and C. Theobalt. Video-based reconstruction of animatable human characters. *ACM Transactions on Graphics (TOG)*, 29(6):139, 2010.
- [88] R. W. Sumner and J. Popović. Deformation transfer for triangle meshes. *ACM Transactions on Graphics (TOG)*, 23(3):399–405, 2004.
- [89] G. K. Tam, Z.-Q. Cheng, Y.-K. Lai, F. C. Langbein, Y. Liu, D. Marshall, R. R. Martin, X.-F. Sun, and P. L. Rosin. Registration of 3D point clouds and meshes: A survey from rigid to nonrigid. *Transactions on Visualization and Computer Graphics*, 19(7):1199–1217, 2013.
- [90] J. W. Tangelder and R. C. Veltkamp. A survey of content based 3D shape retrieval methods. *Multimedia Tools and Applications*, 39(3):441–471, 2008.
- [91] TC2. SizeUSA Sizing Survey. <http://www.sizeusa.com/>.
- [92] O. Van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or. A survey on shape correspondence. In *Computer Graphics Forum*, volume 30, pages 1681–1707. Wiley Online Library, 2011.
- [93] R. C. Veltkamp and B. Piest. A physiological torso model for realistic breathing simulation. In *Modelling the Physiological Human*, pages 84–94. Springer, 2009.
- [94] J. Vermaak and P. Pérez. Constrained subspace modeling. In *International Conference on Computer Vision (CVPR)*, volume 2. IEEE, 2003.
- [95] P. Viola, M. J. Jones, and D. Snow. Detecting pedestrians using patterns of motion and appearance. *International Journal of Computer Vision (IJCV)*, 63(2):153–161, 2005.
- [96] Y. Wang, B. Peterson, and L. Staib. 3D brain surface matching based on geodesics and local geometry. *Computer Vision and Image Understanding*, 89(2-3):252–271, 2003.
- [97] A. Weiss, D. A. Hirshberg, and M. J. Black. Home 3D body scans from noisy image and range data. In *International Conference on Computer Vision (ICCV)*, pages 1951–1958, 2011.
- [98] J. West. *Respiratory physiology*. Wolters Kluwer Health, 2012.
- [99] H. Wold. Partial least squares. *International Journal of Cardiology*, 147(2):581–591, 1985.
- [100] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland. Pfunder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 19(7):780–785, 1997.
- [101] H.-Y. Wu, M. Rubinstein, E. Shih, J. Gutttag, F. Durand, and W. T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (TOG)*, 31(4), 2012.
- [102] S. Wuhrer and C. Shu. Estimating 3D human shapes from measurements. *Machine Vision and Applications*, 24(6):1133–1147, 2013.

- [103] J. Yedidia, W. Freeman, and Y. Weiss. Understanding Belief Propagation and its generalizations. volume 8, pages 236–239. Citeseer, 2003.
- [104] A. Zaharescu, E. Boyer, K. Varanasi, and R. Horaud. Surface feature detection and description with applications to mesh matching. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 373–380. IEEE, 2009.
- [105] R. Zass and A. Shashua. Nonnegative sparse PCA. 2007.
- [106] Y. Zeng, C. Wang, X. Gu, D. Samaras, N. Paragios, and N. Komodakis. A generic deformation model for dense non-rigid surface registration: a higher-order MRF-based approach. In *International Conference on Computer Vision (ICCV)*. IEEE, 2013.
- [107] H. Zhang, A. Sheffer, D. Cohen-Or, Q. Zhou, O. Van Kaick, and A. Tagliasacchi. Deformation-driven shape correspondence. In *Computer Graphics Forum*, volume 27, pages 1431–1439. Wiley Online Library, 2008.
- [108] V. B. Zordan, B. Celly, B. Chiu, and P. C. DiLorenzo. Breathe Easy: model and control of simulated respiration for animation. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA)*, pages 29–37. Eurographics Association, 2004.
- [109] H. Zou and T. Hastie. Regularization and variable selection via the Elastic Net. *Journal of the Royal Statistical Society, Series B*, 67:301–320, 2005.
- [110] H. Zou, T. Hastie, and R. Tibshirani. Sparse Principal Component Analysis. *Journal of Computational and Graphical Statistics*, 15(2):265–286, 2006.