

Animatable Human Body Model Reconstruction from 3D Scan Data using Templates

Laurent Moccozet, Fabien Dellas, Nadia Magnenat-Thalmann
MIRALab, University of Geneva, Switzerland
[moccozet|fabien|thalmann]@miralab.unige.ch

Silvia Biasotti, Michela Mortara, Bianca Falcidieno
CNR-IMATI, Genova Italy
[silvia|michela|bianca]@ima.ge.cnr.it

Patrick Min, Remco Veltkamp
Dept. of Computer Science, Utrecht University, the Netherlands
[min|remcov]@cs.uu.nl

Abstract

Many detailed 3D datasets of human bodies are available, and with current scanning technology, new ones are relatively easy to produce. As a result, more realistic looking animations become possible, but only after considerable processing: holes in the data, (too) high resolutions, and the laborious task of fitting an animation skeleton to the data are the main problems to overcome. In this paper, we propose a reconstruction pipeline which solves these three problems by fitting a template with surface and animation information to the scanned data. A correspondence between the scanned data and the template is set manually, with the aid of a tool that identifies and visualises *landmarks* (characteristic points) on the mesh. Holes are handled using a "vertex confidence weight" scheme, with low weights assigned to vertices in or near holes (implying higher weights to template vertices, and vice versa). Initial tests on two detailed 3D human body scans produced high-quality closed surface meshes, which are directly animatable.

Keywords: human body modelling, skeleton-based animation, shape feature extraction, template-based modelling, 3D range scanning

1 Introduction

Nowadays many different 3D scanning methods are available for digitizing 3D human body models. Because the source of these body models is real human data, they are suitable for producing realistic looking animations.

However, many processing steps are required before data resulting from a 3D scanning process can be readily used by an animator. Scanned data typically is noisy, contains many holes, and requires significant simplification. Scanned models of the human body are particularly prone to having holes, because it is usually not possible to produce sufficient scans covering every body surface area. Moreover, before the body model is ready for animation, a correspondence between an articulated skeleton and the model surface has to be established, which is a time-consuming process.

Lee *et al.* [1] repair the scanned surface of human faces by locally interpolating surrounding points. Global interpolation methods as presented by Carr *et al.* [2] have several advantages over local methods, such as the smoothness of the resulting surface. They directly reconstruct a hole-free surface that closely approximates the scanned target surface. Algorithms making use of *domain knowledge* can be more efficient because they prevent the surface from erroneously being smoothed out in areas with insufficient scan data. Examples are template-model based

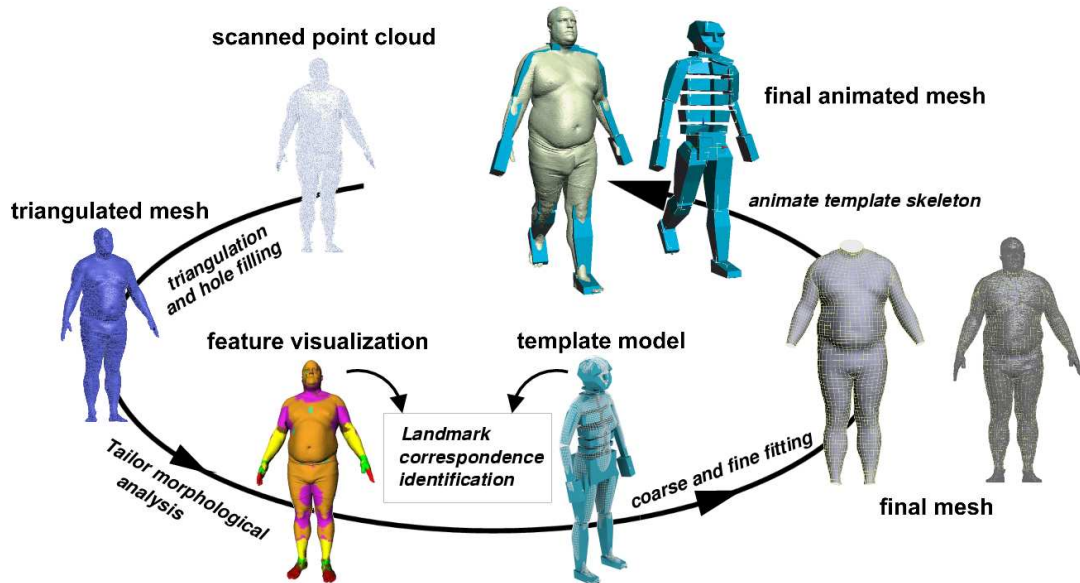


Figure 1: High-level overview of the reconstruction pipeline

fitting strategies (where holes are filled with geometry from the template surface), as presented by Kähler *et al.* [3] (for faces, with the template-to-data correspondence found using easily identifiable landmarks) and Allen *et al.* [4, 5] (for bodies, with correspondence markers attached to the scanned people).

Our objective is to reconstruct the body shape (with no assumptions about markers on the input data) *and* fit the animation control structure in order to directly obtain a fully animatable body model. To this end we propose a full reconstruction pipeline that produces a close approximation of the scanned data of a human body. It is based on fitting a human template model (which includes both the skin surface and the animation control information) to the scanned data. The template we use was developed by Hyewon Seo and is described in detail in [6]. It captures both the variability of individual human bodies and their common structure. The correspondence between the template and the scanned data is set manually, aided by a tool called *Tailor*, which is based on a multi-scale morphological analysis method [7]. This method decomposes the surface into meaningful shape features, like tips, tubular protrusions, concave regions, sharp points, etc. A visualisation of these shape features helps the user localise many of the required landmarks.

By fitting to this particular type of template, we simultaneously solve many problems related to reconstruction from scanned data:

- **level of detail**
can now be easily set to the desired value
- **holes**
are filled in a semantically correct way through the fitting process
- **animation**
the resulting model can be directly animated

The remainder of this paper is organised as follows. The next section presents a high-level overview of our reconstruction pipeline. Sections 3, 4, and 5 describe each step in the pipeline in more detail. Section 6 has some preliminary results, followed by conclusions and a discussion of current work on our method.

2 Method Overview

The complete pipeline, going from range-scanned point clouds to an animatable closed-surface model, consists of the following steps:

1. fill the holes of the scanned model with any straightforward method

2. assign confidence weights to mesh vertices (with vertices in a filled hole having a low weight)
3. establish correspondence between particular landmarks on the scanned model and the template
4. fit the template to the model (consisting of a coarse and fine fitting step)

The entire process is illustrated in Figure 1. The following sections will describe each step in more detail.

3 Hole Filling and Assigning Vertex Confidence Weights

The first step deals with the often unavoidable holes in the scanned data. These holes can be of many different sizes and are usually located in areas that were hard to “see” by the scanner.

In order to arrive at a reasonable template fit at the holes, we assign “confidence weights” to the vertices of the template mesh. First the holes in the scan data are filled using any available straightforward method. Next, the vertices of the template model are assigned confidence weights: low weights to vertices in untrusted regions (i.e. in and near holes), high weights in trusted regions of the mesh. Figure 2 shows a screenshot of the user interface for setting these weights. The white dashed circle helps to locate a template vertex (the red dot in the armpit), the black dashed circle visualises a “weight falloff” area. The colors correspond to weight values (continuously changing between blue (0%) and orange (100%)).

What follows are two fitting stages, a “skeleton fitting” and a “fine fitting” stage (described in the next section). The confidence weights are used during this second “fine fitting” stage: they determine how much a template vertex is attracted towards the scanned data.

4 Correspondence between Scanned Data and Template using Body Landmarks

A set of characteristic points (called *landmarks*) on the surface are used to establish an initial

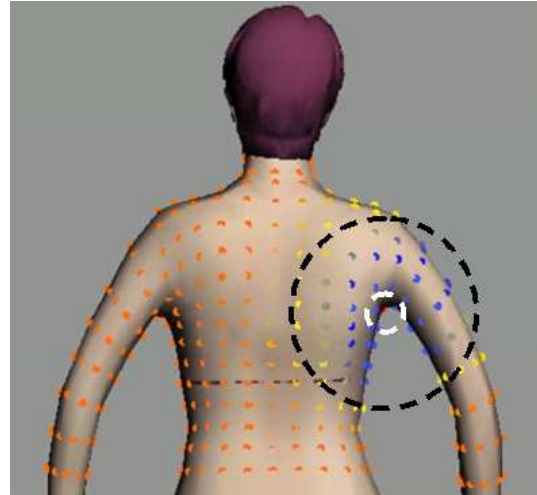


Figure 2: The user interface to set the template vertex confidence weights at holes. The white dashed circle helps to localise a template vertex (the red dot in the armpit). The black dashed circle visualises a “weight falloff” area

correspondence between the scanned model and the template model. The landmark concept was first introduced in biology and used to compare changes in biological shape [8]: a landmark is a feature point on an object (e.g. a bone, skull, or body) that can be identified within and between populations. Figure 3 shows some landmarks defined by the Humanoid Animation Working Group (from the H-Anim 200x specification [9]).

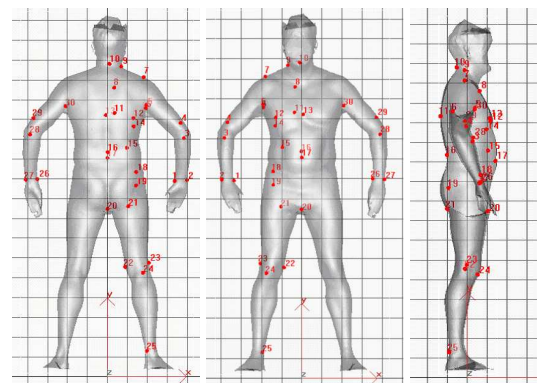


Figure 3: H-Anim-based landmarks topology for the human body [9]

Even though landmarks are often positioned at geometrically salient features (e.g. juxtapositions of parts, points of maximum curvature, extremal points), it is not always possible to find them using local geometric information alone.

Therefore, often the location of landmarks on surfaces (or images) relies on manual registration before [4, 5] or after [3, 6] the data capture.

Similarly, in our pipeline the user has to select a number of landmarks on both the scanned surface and the template model. To aid the user in finding these landmarks, we use a tool called *Tailor*, which is based on a multi-scale morphological analysis method [7]. This method decomposes the surface into meaningful shape features, like tips, tubular protrusions, concave regions, sharp points, etc. A visualisation of these shape features helps the user localise many of the required landmarks. We now briefly sketch the key points of the algorithm (see [7] for more detail).

The main idea of the Tailor method is to evaluate the curvature and other geometric descriptors over vertex neighbourhoods of variable sizes. This multi-scale analysis is achieved by using a set of levels of detail which define increasingly larger neighbourhoods of each surface vertex (see Figure 4). Both the geometry and topology of these neighbourhoods are used for identifying the type of local feature.

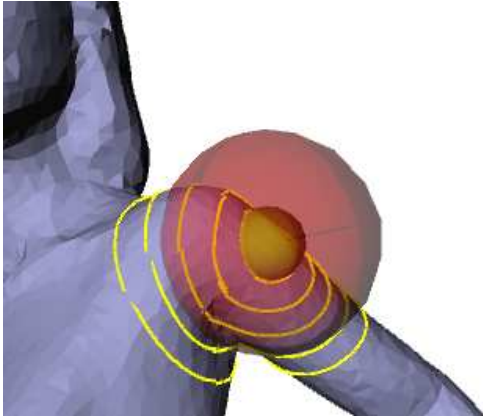


Figure 4: Evolution of the intersection curves between the input surface and a set of spheres with the same centre and increasing radius

More formally, given a set of radii $\{r_i\} \wedge i = 1 \dots n$, for each vertex v of the surface mesh M and radius r_i we consider the components of the boundary γ of the surface region containing v and delimited by the intersection between M and the sphere $S(v, r_i)$ with centre v and radius r_i . Other regions of intersection may occur but will be discarded. The following cases can be identified:

- one boundary component: the surface around v is equivalent to a disk (Fig. 5 (a))
- two boundary components: the surface around v is tubular-shaped (Figure 5 (b))
- three or more boundary components: in a neighbourhood of v a branching of the surface occurs (Figure 5 (c))

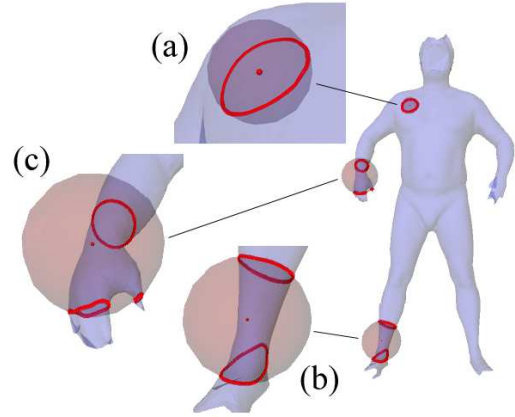


Figure 5: Intersection between the input surface and spheres of the same radius r centred at different vertices of M . The number of intersection curves changes depending on the local shape of M around the vertex

If the intersection curve γ has one component, the curvature value at v with respect to the level of detail r is approximated by the ratio $G_r(v) = \text{length}(\gamma)/r$. In this case, v is classified as

1. *planar* if $G_r(v) \approx 2\pi$
2. *sharp* if $G_r(v) < 2\pi$
3. *blend* if $G_r(v) > 2\pi$

If there are two intersection curves, then more feature shape information can be derived from the relative length of the curves. For example, if their lengths are similar, then v is classified as *cylindrical* (at scale r).

Given a set of radii, each vertex of the mesh is associated with an n -dimensional vector of morphological labels, each label corresponding to the vertex type at the related scale. Shape features of the mesh are then identified by connected regions of vertices with the same label at a given scale. The defined model allows us to analyze a shape at different scales, and to derive

information about the persistence of a shape feature across the scale range.

Several different landmarks can now be detected at centroids of certain morphological regions (see Figure 6 for some examples).

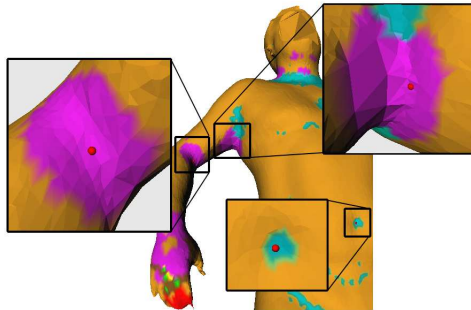


Figure 6: Examples of landmarks found in centres of morphological regions (the armpit, inner elbow, and umbilical region)

Figure 7 highlights how the landmark detection depends on the scale chosen. In fact, when the human model proposed in Figure 3 is characterised with respect to a small scale (see Figure 7(a)), small landmarks like the umbilical region are recognised, while at a larger scale (see Figure 7(b)), a bigger number of characteristic points is detected and the region classification becomes numerically more stable.

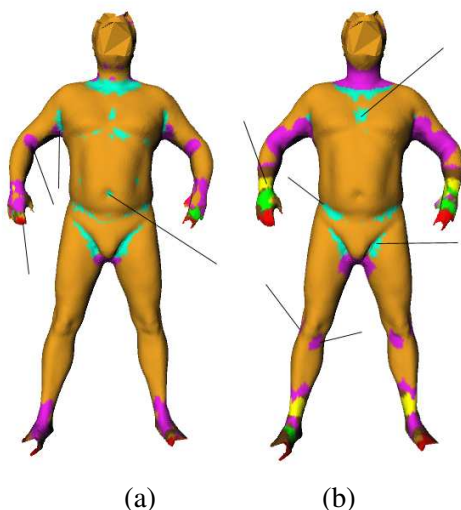


Figure 7: Shape characterisation at different scales (the lines point to identified landmarks)

A promising attribute of the method is that it appears to produce a consistent decomposition of several different body models.

5 Template Model Fitting

With the given correspondence between the scanned data and template, and weight settings for the template vertices at holes in the data, the pipeline now proceeds to two fitting stages. The first is a coarse “skeleton fitting” stage, the second a more accurate “fine fitting” stage. Both are described in detail in [6].

The first “skeleton fitting” stage finds the relative position, proportion and the posture of the scanned model by appropriately fitting the template model to it (see Figure 8).

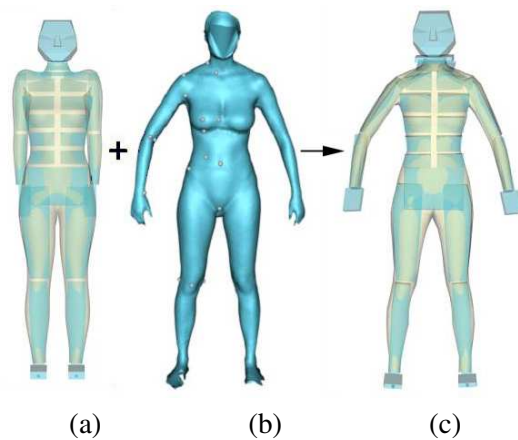


Figure 8: **Skeleton fitting** The template model in (a) is fitted to the scanned data in (b), resulting in (c) a linear approximation of the scanned data

The second “fine fitting” stage starts with the coarsely fitted template model found in the previous stage and iteratively improves the fitting accuracy by minimizing the shape difference between the template and the scan model.

The end result is a 3D mesh without holes, containing areas of varying “trustworthiness”: trusted regions corresponding to the scan data, and less trusted regions corresponding to the filled holes.

6 Preliminary Results

Figure 10 shows two example reconstructed models. A traditionally difficult area in reconstruction of scanned human shapes, the armpit, has been highlighted.

Because in the reconstructed model the correspondence of the surface with the underlying template is known, it is directly animatable (see Figure 11).

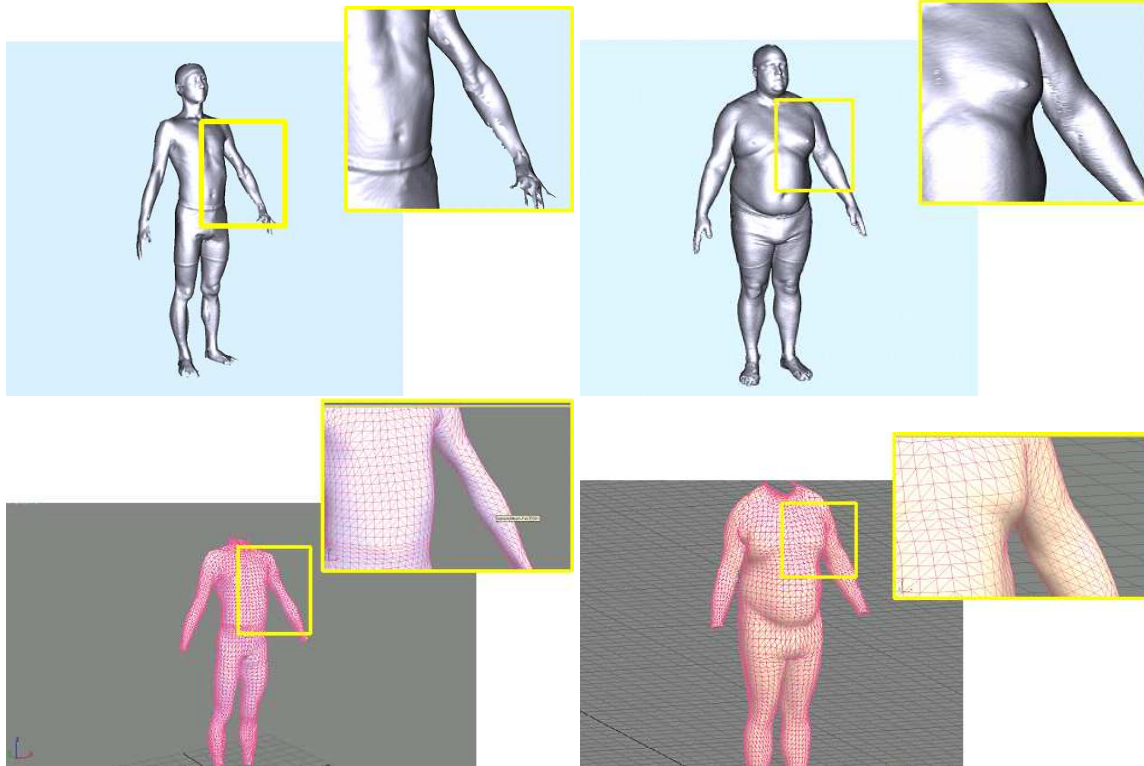


Figure 10: Examples of scanned human body models with an area with holes highlighted (top row) and the resulting reconstructed models (bottom row)

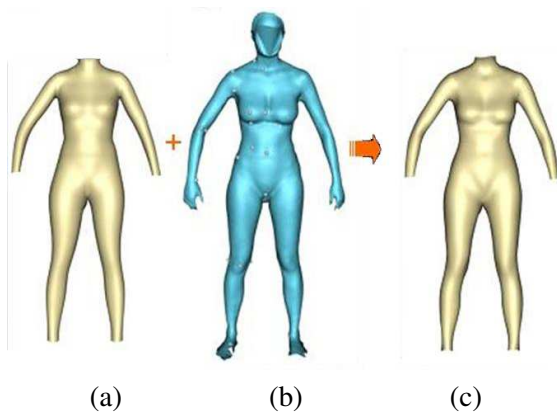


Figure 9: **Fine fitting** From (a) a coarsely fitted model from the previous phase and (b) the scanned model, the fine fitting produces (c) the final mesh

7 Conclusions and Work in Progress

In this paper we have outlined the design of a method to create fully animatable closed 3D meshes from scanned data of human bodies. It relies on fitting a template model (containing skeletal, surface and animation information) to

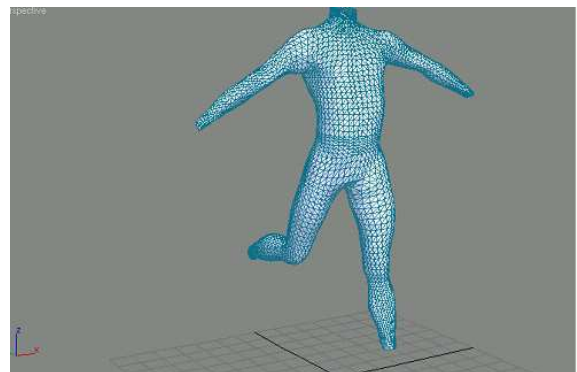


Figure 11: Reconstructed models are directly animatable

the scanned data, after an initial correspondence is established by the user by localizing *landmarks* (characteristic points) on both.

Currently we are working on making the process fully automatic, which implies that these landmarks have to be identified automatically. Some landmarks can already be found automatically, such as the armpits, the umbilical region, and others. Initial tests using the “Tailor” tool (a multi-scale morphological analysis tool for identifying local features [7]) show that it is

likely we can locate a number of “morphologically meaningful” regions, and that these appear to be consistent across different data sets. We expect to derive anthropometric rules based on the relative location of these regions with respect to the whole surface allowing us to reliably assign regions to landmarks.

In other future work we intend to use the Tailor tool at coarser scales to be able to identify body postures.

Acknowledgements

All authors of this paper were supported in part by the AIM@SHAPE Network of Excellence grant 506766 by the European Commission. We thank TNO Human Factors in Soesterberg, the Netherlands, and the CAESAR (Civilian American and European Surface Anthropometry Resource) project for the use of their scanned body models. Finally, thanks to Michela Spagnuolo and Giuseppe Patané for their helpful suggestions.

References

- [1] Y. Lee, D. Terzopoulos, and K. Waters. Realistic modeling for facial animation. In *Proc. SIGGRAPH 1995*, pages 55–62.
- [2] J.C. Carr, R.K. Beatson, J.B. Cherrie, T.J. Mitchell, W.R. Fright, B.C. McCallum, and T.R. Evans. Reconstruction and representation of 3D objects with radial basis functions. In *Proc. SIGGRAPH 2001*, pages 67–76, Los Angeles.
- [3] K. Kähler, J. Haber, H. Yamauchi, and H.-P. Seidel. Head shop: Generating animated head models with anatomical structure. In *Proc. ACM SIGGRAPH Symposium on Computer Animation*, pages 55–64, 2002.
- [4] B. Allen, B. Curless, and Z. Popović. The space of human body shapes: Reconstruction and parameterization from range scans. In *Proc. SIGGRAPH 2003*, pages 587–594. Addison-Wesley.
- [5] B. Allen, B. Curless, and Z. Popović. Exploring the space of human body shapes: Data-driven synthesis under anthropometric control. In *Proc. Digital Human Modeling for Design and Engineering Conf.*, Rochester, MI, June 2004. SEA International.
- [6] H. Seo and N. Magnenat-Thalmann. An automatic modeling of human bodies from sizing parameters. In *Proc. ACM SIGGRAPH Symposium on Interactive 3D Graphics*, pages 19–26. ACM Press, 2003.
- [7] M. Mortara, G. Patané, M. Spagnuolo, B. Falcidieno, and J. Rossignac. Blowing bubbles for multi-scale analysis and decomposition of triangle meshes. *Algorithmica*, 38(1):227–248, 2004.
- [8] F.L. Bookstein. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge Press, 1997.
- [9] ISO/IEC FCD 19774:200x, humanoid animation, annex B, feature points for the human body. <http://www.h-anim.org/>.