Shape from Selfies : Human Body Shape Estimation using CCA Regression Forests

Endri Dibra[1] Cengiz Öztireli[1] Remo Ziegler[2] Markus Gross[1]

[1]Department of Computer Science, ETH Zürich [2]Vizrt
{edibra,cengizo,grossm}@inf.ethz.ch rziegler@vizrt.com

1 Supplementary Material

In this supplementary material, we present new quantitative results, shown in Table 1, illustrations of the running and walking poses in Fig. 1 and enclose a video illustrating one of the practical applications of our method.

1.1 Quantitative Results

In the paper, we presented quantitative results of mean error and standard deviation for 16 body measurements performed on the mesh. There we compared to state-of-the-art methods that work under more restrictive assumptions, one of which being known distance from the camera. In contrast, we made no assumption on the absolute scale of the silhouette from which we estimate the body shape. Here, we present results of an additional experiment performed under the assumption that the absolute scale of the silhouette is known. The experiment is performed on Dataset 1, both with (CCA-RF-S-1) and without (RF-S-1) projecting the features onto the CCA bases. As can be observed in Table 1, we get significant reductions in error for many measurements as compared to the case with no known absolute scale (RF-1), especially for the height. These errors are close to the ground truth (GT) error, which is the lowest error possible with the body shape model we use. Additionally better predictive results are noticed when the CCA is applied to the extracted features.

1.2 Qualitative Results

As part of our qualitative results, we show an application of our shape estimation to free-viewpoint video. Even though the video should be self contained, here we highlight the important points for further reference. In addition, in Fig. 2, we show the raw silhouette extraction results for two random frames, using the Gaussian mixture model for background modelling and graphcut.

- The shape of the subject in the video is estimated using a monocular silhouette set to a fixed scale, as in the paper, and different from the images used for texture projection.

Measurement	RF-1	RF-S-1	CCA-RF-S-1	GT
A. Head circumference	16±13	14±11	13±11	13±9
B. Neck circumference	13±10	8±7	7±8	6±6
C. Shoulder-blade/crotch length	31±24	18±16	17±16	14±11
D. Chest circumference	38±31	28±25	25±23	24±24
E. Waist circumference	35±28	25±23	23±23	16±14
F. Pelvis circumference	33±26	19±17	18±17	14±12
G. Wrist circumference	10±8	6±6	6±6	5±5
H. Bicep circumference	16±13	10±11	10±10	9±10
I. Forearm circumference	14±11	10±9	10±8	8±8
J. Arm length	19±14	14±12	13±12	8±8
K. Inside leg length	26±19	18±15	16±13	9±9
L. Thigh circumference	22± 18	16±15	15±14	11±11
M. Calf circumference	18±13	11±9	11±9	7±8
N. Ankle circumference	10±7	7±7	7±7	5±5
O. Overall height	60±45	36±29	29±25	14±11
P. Shoulder breadth	15±14	13±15	13±13	12±11

Table 1: Comparisons of the complementary results via various measurements. The measurements are illustrated in Fig. 3 (middle) in the paper. Errors represent Mean±Std. Dev and are expressed in millimeters. From left to right : Our results (as from the paper) without applying CCA, the new results under the known scale assumption, the same with CCA applied to the features, the ground truth error defined as the error between the original model and its projection to the shape space spanned by the 20 parameters we utilize.

- The estimated height of the subject is 172 cm as compared to the real height, which is 176 cm.
- The mesh is posed using linear blend skinning, and an automatic tool for simple rigging [1]. Texture is projected onto the mesh from two camera views.
- The rendering of the estimated mesh is compared to that of the mean mesh, which is the mean of the meshes that we use to learn the body shape model.
- No further deformation (e.g. scaling or stretching to fit to the silhouette) is applied to the estimated mesh except for pose deformation, hence the mesh is used as-it-is.
- The free-viewpoint video is merely shown to highlight the quality and accuracy of the estimated and fitted mesh, rather than to show perfect rendering.
- The methods we compare to, to the best of our knowledge, have not been used for such an application before.

In summary, we have shown in a real scenario that we can estimate body shape parameters quite accurately, without any further information. Height estimation, in particular, is a measure that has not been reported before using other methods. The reason for a good height estimate is due to the high correlation that height has with other shape parameters, some of which can be estimated without actually knowing the scale. This in turn is due to our global and local features, their projection to the CCA bases, and also to the statistical

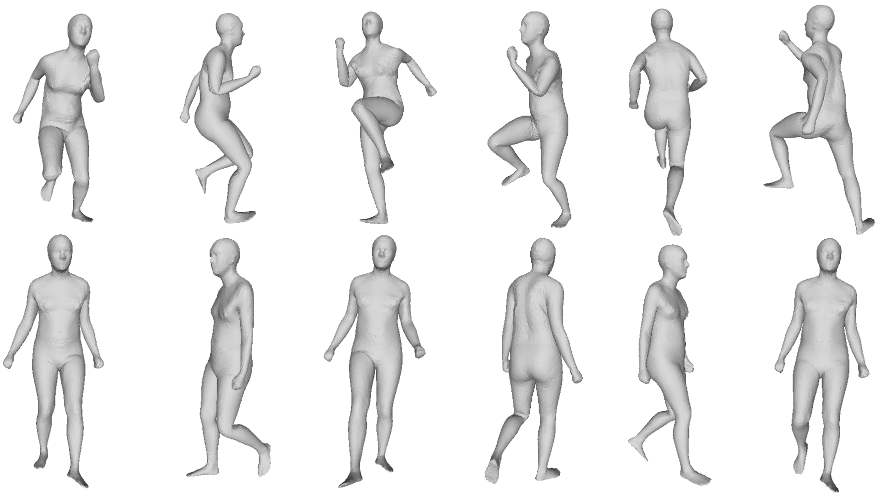

Fig. 1: (top) Meshes in running poses. (bottom) Meshes in walking poses.

human shape model, that discards impossible variations and combinations of those parameters altogether.

Fig. 2: Raw silhouettes extracted using Gaussian mixture model and graphcut for two random frames from the video. No further post-processing is applied.

References

1. Baran, I., Popovic, J.: Automatic rigging and animation of 3d characters. ACM Trans. Graph. (2007)