BLSM: A Bone-Level Skinned Model of the Human Mesh - Supplementary Material

Haoyang Wang[1,2], Riza Alp Güler[1,2], Iasonas Kokkinos[2],
George Papandreou[2], and Stefanos Zafeiriou[1,2]

[1] Imperial College London
{haoyang.wang15, r.guler, s.zafeiriou}@imperial.ac.uk
[2] Ariel AI, London, UK
{hwang,alpguler,iasonas,gpapan,szafeiriou}@arielai.com

1 Visualisation of the Model

Here we provide visualisation of our model. Section 1.1 shows samples of our proposed kinematically feasible motions. Full visualisation of the motions is available in the supplementary video. In section 1.2 we visualise samples from our linear and graph convolutional shape models. For linear model, the samples are draw by varying each of the shape coefficients from -3σ to 3σ. For graph convolutional modelling, we vary each dimension of the latent space from -3 to 3.
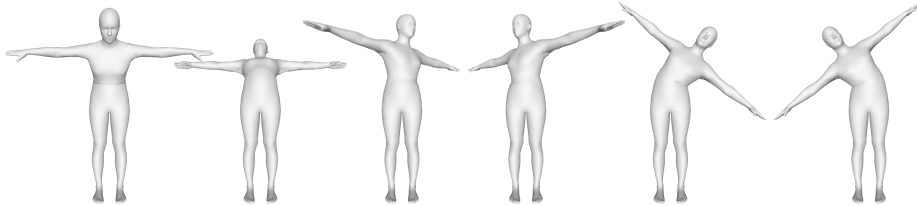
1.1 Kinematically Feasible Body Motions

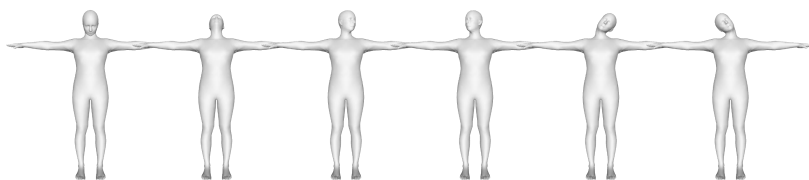Fig. 1: Samples of Pose Atoms - Spine.
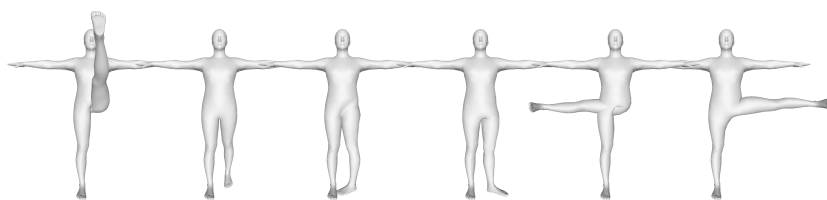
Fig. 2: Samples of Pose Atoms - Head.

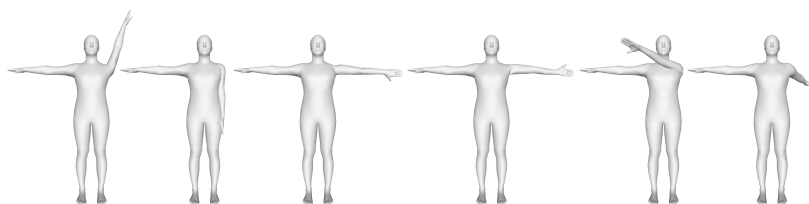Fig. 3: Samples of Pose Atoms - Left Upper Leg.

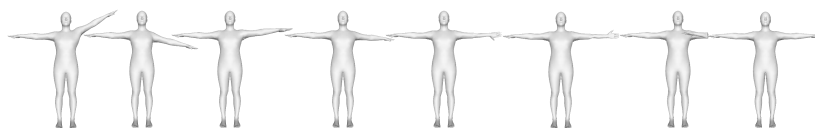Fig. 4: Samples of Pose Atoms - Left Arm.

Fig. 5: Samples of Pose Atoms - Left Shoulder and Left Forearm.
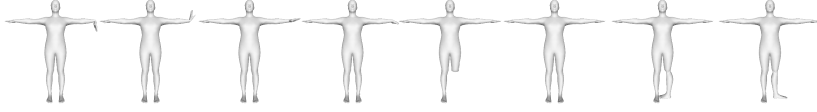
Fig. 6: Samples of Pose Atoms - Left Hand and Left Leg.
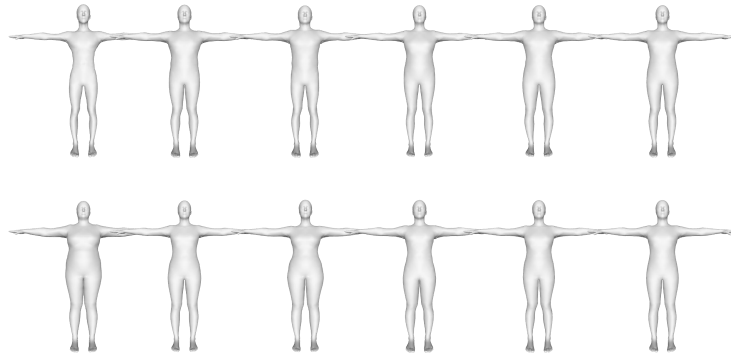
1.2   Shape Modelling

Fig. 7: Linear modelling of shape variations. Top: -3σ, Bottom: 3σ
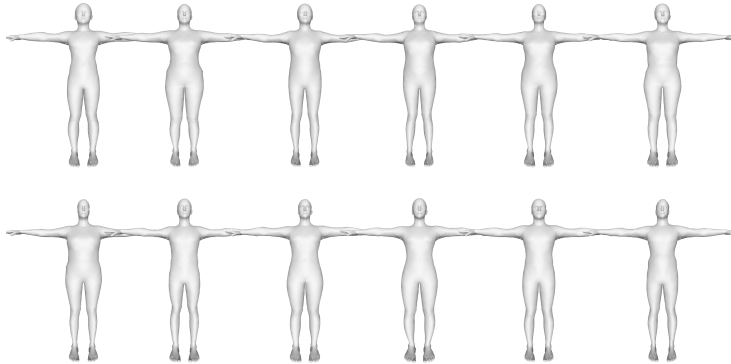

Fig. 8: Graph Convolutional modelling of shape variations.

2   Reconstruction Analysis

2.1   Comparison to original SMPL implementation

Here we also compare to the publicly available version of SMPL up to 10 shape bases. In order to evaluate the model on our CAESAR registrations and In-house testset, we transferred the SMPL learnt parameters, namely the template, shape blendshapes, joint regressors and blending weights to our template topology with barycentric correspondences between our template and the SMPL template, while keeping the original SMPL kinematic structures. Poseblend shapes are excluded in this comparison. Tab. 1 shows the errors when 10 and 32 coefficients are used, as well as the area under curve for the cumulative per-vertex error distribution. Fig. 9 shows the mean absolute error plot from Fig. 4 in the main paper, with evaluation of SMPL that is transferred to our topology. We observe that SMPL has slightly higher errors than our SMPL-reimpl, and we believe that this is due to the topological difference and registration method difference presented in the SMPL training set and our training set. Other than this difference, our SMPL-reimpl performs in par with the publicly available SMPL.

No. Bases Method	10		32	
	error (mm)	AUC	error (mm)	AUC
SMPL	5.45 ± 3.51	0.695	-	-
SMPL-reimpl	4.98 ± 3.64	0.686	4.09 ± 3.24	0.769
BLSM-linear	$\mathbf{3.63 \pm 2.33}$	**0.794**	3.46 ± 2.86	0.803
BLSM-spiral	-	-	$\mathbf{2.74 \pm 2.03}$	**0.819**

Table 1: Mean absolute generalisation error and Area Under Curve for cumulative error distribution of evaluated methods on our in-house testset.
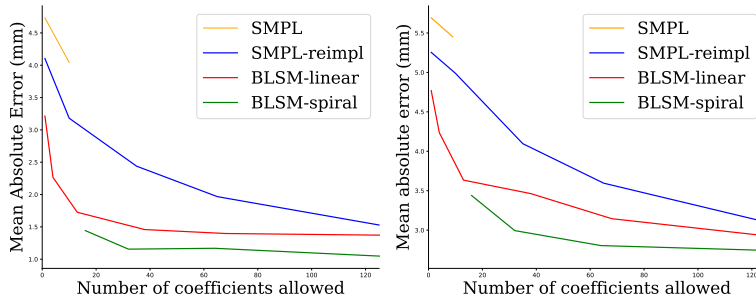

Fig. 9: Mean absolute vertex error on the CAESAR dataset (left) and our in-house testset (right) against number of shape coefficients compared to the original SMPL implementation.

2.2   Reconstruction Results

Here we show reconstruction results on samples of our test set. For each evaluated model, heatmaps of the per-vertex absolute error are visualised.
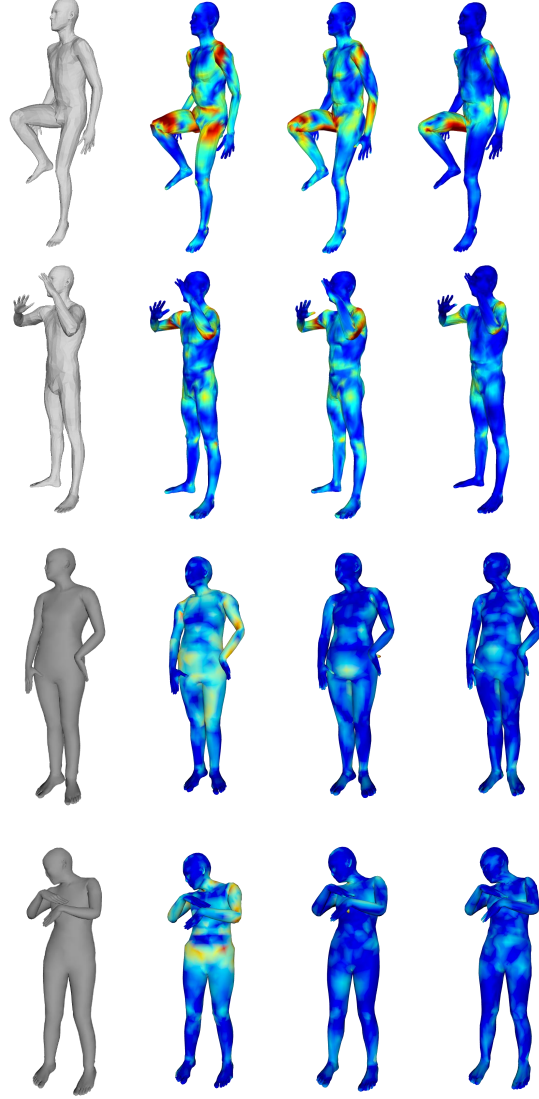

Fig. 10: Samples from reconstructions of D-Faust and our testset. Left to right: ground truth, SMPL-reimpl (125), BLSM-lienar (125), BLSM-spiral (128)

2.3   Reconstruction Performance Analysis on CAESAR

In this section we evaluate the reconstruction performance of our proposed linear and graph convolutional model on different demographic groups in the CAESAR dataset. We observe that the bias in our proposed model reflects the biases in the dataset, specifically related to body type variations. For example in the training set, only 0.9% of the subjects has BMI over 40, while our models have significantly higher reconstruction error on this group of subjects. For other attributes (e.g. gender) that are less biased, our model performs similar on all the groups while slightly better on the more common group (53.7% female vs 46.3% male).

Model / Group	BLSM-linear-125 error (mm)	BLSM-spiral-128 error (mm)
BMI $\leq$ 40	1.224 ± 0.337	1.071 ± 0.426
BMI $>$ 40	2.385 ± 0.429	1.369 ± 0.240

Model / Group	BLSM-linear-125 error (mm)	BLSM-spiral-128 error (mm)
Height $\leq$ 200 cm	1.221 ± 0.338	1.064 ± 0.419
Height $>$ 200 cm	2.160 ± 0.384	1.642 ± 0.399

Model / Group	BLSM-linear-125 error (mm)	BLSM-spiral-128 error (mm)
Weight $\leq$ 120 kg	1.217 ± 0.329	1.065 ± 0.422
Weight $>$ 120 kg	2.355 ± 0.377	1.534 ± 0.341

Model / Group	BLSM-linear-125 error (mm)	BLSM-spiral-128 error (mm)
Female	1.283 ± 0.346	0.982 ± 0.385
Male	1.399 ± 0.332	1.204 ± 0.445

Table 2: Reconstruction error on groups of subjects with different BMI value, height, weight and gender

3   Regargeting

In this section we show results of image-driven character animation with pixel-level accuracy with our model. We rig two characters from [1] using our bone structure, and apply animation on videos from [2].

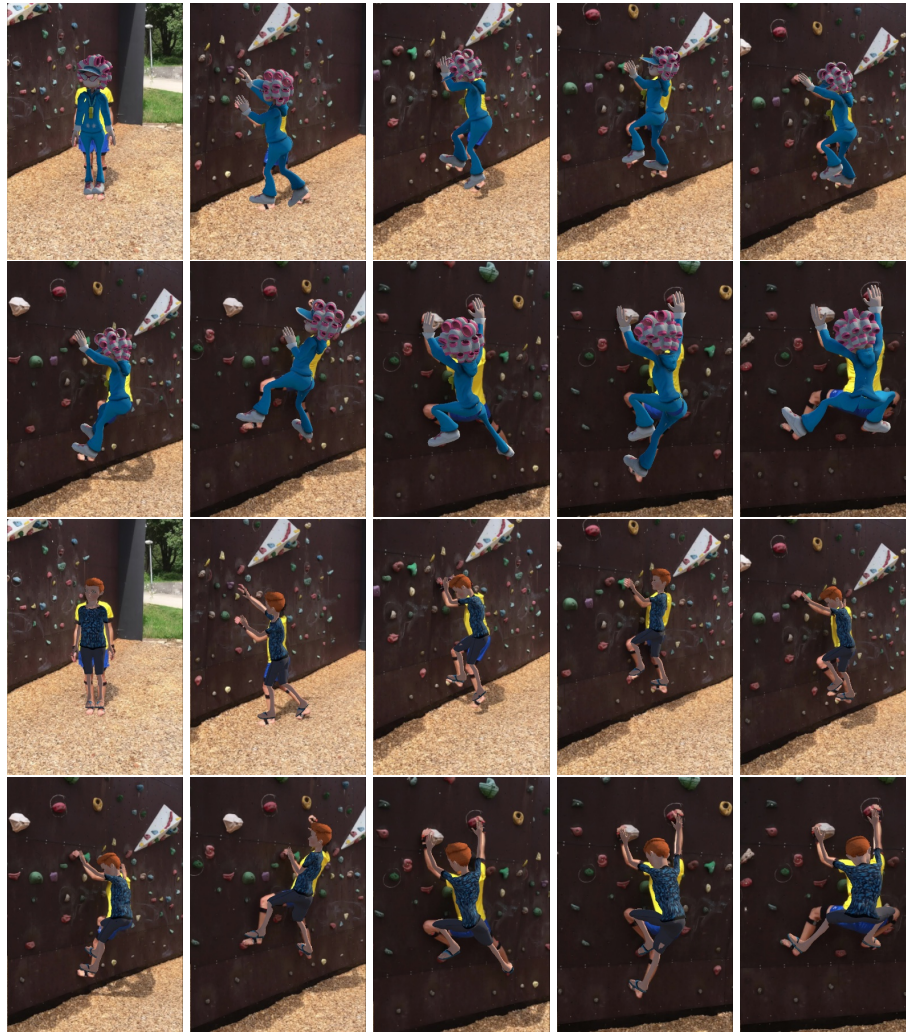Fig. 11: Pixel-accurate, image-driven character animation in-the-wild.

Fig. 12: Pixel-accurate, image-driven character animation in-the-wild (continued).

References

1. Mixamo. https://www.mixamo.com (2019)
2. von Marcard, T., Henschel, R., Black, M., Rosenhahn, B., Pons-Moll, G.: Recovering accurate 3d human pose in the wild using imus and a moving camera. In: European Conference on Computer Vision (ECCV) (sep 2018)