Supplementary Material: Learning to Transfer Texture from Clothing Images to 3D Humans

1. Non-Rigid Garment Fitting

In all our experiments, we use the neutral gendered SMPL model. For each garment class we only use a subset of angles in the input axis-angle representation of pose as our optimization variables. For shirts, we use only the pose parameters corresponding to the shoulder joints - numbered $13, 14, 16, 17$ in the canonical ordering. For shorts, joint corresponding to the hip joint - joints 1 and 2 are used. Joints $1, 2, 3, 4$ which correspond to the hip and knee joints are used while registering pants.

The weights used in the optimization based registration are given below. For registration the front of T-shirts, we set $w_{\boldsymbol{\beta}} = 3.0, w_s = 100, w_{\boldsymbol{\theta}} = 1.0, w'_s = 2.0, w'_c = 0.3, w'_b = 2.0, w'_l = 30.0, w'_e = 0$. The same weights when registration the back are: $w_{\boldsymbol{\beta}} = 5.0, w_s = 100, w_{\boldsymbol{\theta}} = 0, w'_s = 2.0, w'_c = 0.3, w'_b = 2.0, w'_l = 30.0, w'_e = 0$.

For the front view of shorts, we set $w_{\boldsymbol{\beta}} = 4.0, w_s = 100$ $w_{\boldsymbol{\theta}} = 30.0$ in the first stage when shape is not being optimized and $w_{\boldsymbol{\theta}} = 20$ when it is. $w'_s = 100, w'_c = 4.0, w'_b = 2.0, w'_l = 45.0, w'_e = 0.001$. For the back $w_{\boldsymbol{\beta}} = 3.0, w_s = 160.0$ $w_{\boldsymbol{\theta}} = 40.0$ when shape is not an optimization variable and $w_{\boldsymbol{\theta}} = 30$ when it is. $w'_s = 100, w'_c = 3.8, w'_b = 6.0, w'_l = 75.0, w'_e = 0.01$.

For both front and back of pants, we set $w_{\boldsymbol{\beta}} = 4.0, w_s = 100$ $w_{\boldsymbol{\theta}} = 40.0$ in the stage when just the pose is optimized $w_{\boldsymbol{\theta}} = 5$ when both shape and pose are optimization variables. $w'_s = 100, w'_c = 0.5, w'_b = 2.0, w'_l = 45.0, w'_e = 0.1$.

2. Learning Automatic Texture Transfer

For all our experiments, we use an Adam optimizer with $\beta 1 = 0.5$ and $\beta 2 = 0.999$. We use a constant learning rate of 0.0001. All mapping networks are trained for 200 epochs and all segmentation networks are trained for 20 epochs. During the training of the segmentation networks we use heavy colour jittering to prevent the network from overfitting to the textures in the training data. The weights are set as $\lambda_{reg} = \lambda_{recon} = \lambda_{perc} = 1$

3. Baseline Implementation

For the SC-matching TPS based baseline, all parameters are set to their default values. The pix2pix baseline is im-
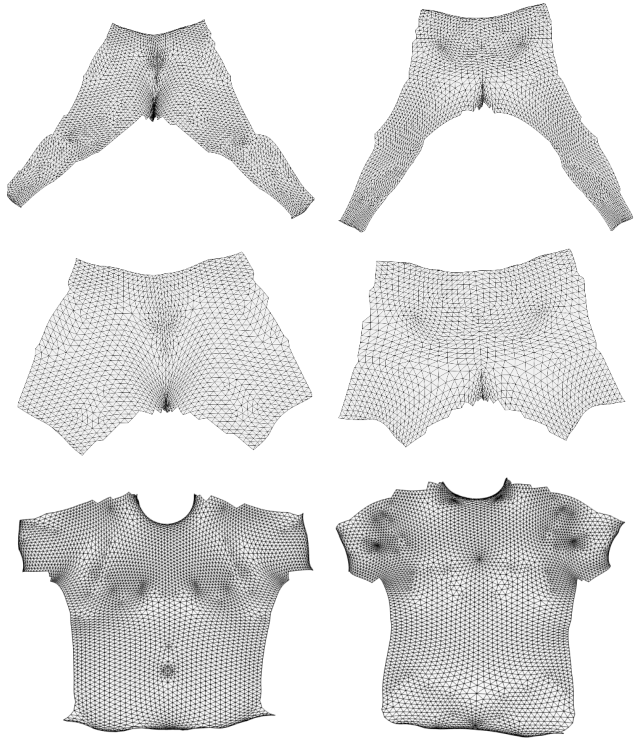

Figure 1. Custom UV maps

plemented with the weight $\lambda_{gan} = 1$. We observed that increasing the weight of the adversarial loss worsens performance.

4. Custom UV Maps

All the UV maps used are created in blender. We obtained best performance when the shape of the UV maps are similar to the shapes of the input garments. The exact UV maps used can be seen in Fig. 1

5. Results

In Fig. 2 and Fig. 3 more textures mapped by Pix2Surf are displayed atop SMPL.

Once the garment textures are mapped to their corresponding texture maps, they can be rendered atop SMPL in different shapes and poses as seen in Fig. 4.

Figure 2. Textures mapped by Pix2Surf rendered atop SMPL

Figure 3. Textures mapped by Pix2Surf rendered atop SMPL

Figure 4. Once textures are mapped to their corresponding texture maps, they can be rendered atop SMPL in different shapes and poses