Garment Recovery with Shape and Deformation Priors

Supplementary Material

6. Data Processing

In this section, we introduce the processing steps to generate the data required by the training of our models.

6.1. Transfer to the Neutral Body

The training of ISP requires the rest garment mesh draped on the neutral body ($\beta = \mathbf{0}$). However, the rest garments in CLOTH3D [4] are draped on bodies with varying shapes. To transfer these garments onto the neutral body, we use the diffused SMPL skinning procedure [47] to make the garment fit to the neutral body:

$$v = v - \frac{1}{N} \sum_{v_n \sim \mathcal{N}(v, \mathbf{d})} B_\beta(\phi(v_n)) \ , \tag{14}$$

where $\phi(v_n)$ computes the closest SMPL body point for v_n, $\mathbf{d}$ the distance from v to the body surface, $B_\beta(\cdot)$ returns the shape displacement of SMPL, $\mathcal{N}$ is the normal distribution and $N = 1000$. While Eq. 14 produces a garment that fits the neutral body, its surface may be inconsistent with the original mesh, as illustrated in Fig. 8(b). To address this, we perform an optimization process using the following equation:

$$v^* = \underset{v}{\operatorname{argmin}} \, L_{style} + L_{strain} + L_{bend} + L_{col} \ . \tag{15}$$

L_{strain}, L_{bend} and L_{col} are the strain energy, the bending energy and the collision loss introduced in Eq. 13. L_{style} is the style preserving loss inspired by [7]

$$L_{style} = \lambda_s L_{shape} + \lambda_e L_{edge} + \lambda_r L_{rl} + \lambda_f L_{fit} \ , \tag{16}$$

$$L_{shape} = \sum_f ||n_f - n_f^o||_2^2 \ , \tag{17}$$

$$L_{edge} = \sum_e 1 - cos(e, e^o) \ , \tag{18}$$

where n_f and n_f^o are the normals of face f in the transferred mesh and the original mesh respectively, and e and e^o are the vectors of the boundary edges. L_{rl} and L_{fit} are the loss terms that enforce relative location preservation and fit in tight regions as defined in [7].The balancing weights λ_s, λ_e, λ_r and λ_f are set to 10, 0.5, 1 and 1, respectively. The optimization of Eq. 15 produces the result that preserves the shape and style of the original mesh as illustrated in Fig. 8(c).

6.2. Cutting

We cut the shirt, trousers and skirt using the following rules:

- Shirt: the faces that cross the plane of $z = -0.03$ are cut, resulting in 2 surfaces (front, back) for the shirt without opening as shown in Fig. 3(b) or 3 surfaces (front-left, front-right, back) for the shirt with opening as shown in Fig. 9(b).
- Trousers: the faces that cross the plane of $x = 0$ or the plane of $z = 0.07y + 0.03$ are cut, resulting in 4 surfaces (front-left, front-right, back-left, back-right) as shown in Fig. 10(b).
- Skirt: the faces that cross the plane of $z = 0.03$ are cut, resulting in 2 surfaces (front, back) as shown in Fig. 11(b).

6.3. Flattening

Once the piece of garment is subdivided into meshes with disk topology, we parameterize them in 2D using as-rigid-as-possible (ARAP) flattening [33], as implemented in [42]. This method minimizes area distortion and produces realistic patterns which can be used for garment fabrication. Given an initial solution computed with LSCM [27], we first determine the ideal rigid transformation for each triangle that maps its vertices from 3D to 2D using the Procrustes method [49]. Then, given a mesh of triangles $\mathcal{T}$ in 3D and its current 2D parameterization $\mathcal{T}'$, the ARAP energy is defined as:

$$E_{\text{rigid}} = \sum_{t \in \mathcal{T}} \sum_{e \in t} (u_e - u_{e'})^2 + (v_e - v_{e'})^2$$

where e is an edge of triangle $t \in \mathcal{T}$, e' is the corresponding edge in $\mathcal{T}'$, u_e and v_e are the lengths of edge e along the first and second axis in 2D, respectively. We minimize this energy using a sparse least-squares solver.

Minimizing E_{rigid} naively would be impractical, however, as the resulting patterns would have arbitrary orientation and lack consistency. Our application requires consistent patterns across the dataset so that ISP can learn to represent them jointly. To this end, we add further constraints in the optimization of $\mathcal{T}'$ depending on the category of garment. Specifically:

- In trousers, vertices along the waist line are constrained to have $y = 0$, and the vertex in the center is set to $x = 0$.
- In upper-body garments, if there is an opening in the center, the vertices on the central boundary are constrained to $x = 0.04$ for the right side of the garment, and $x = -0.04$ on the left. Additionally, for each vertex $V = (x, y, z)$ on the upper boundary of a sleeve, we set its 2D position to $v = (x - 0.1, y)$ if it is on the left side, or $v = (x + 0.1, y)$ on the right side.

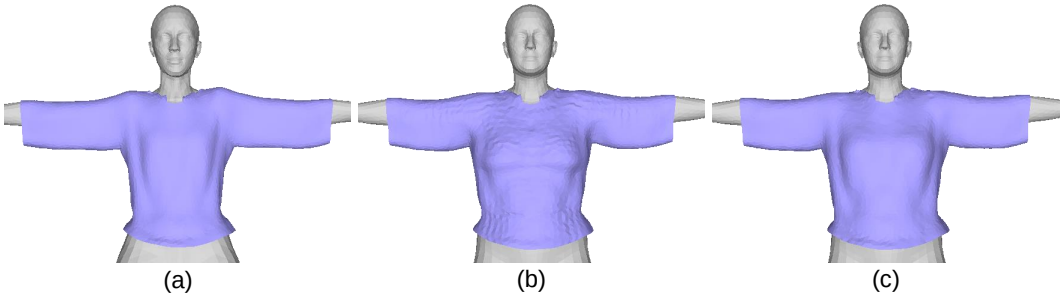

Figure 8. **Transfer to the neutral body.** (a) The original shirt on β-shape body. (b) The shirt transferred by Eq. 14 on the neural body. (c) The shirt refined by Eq. 15.

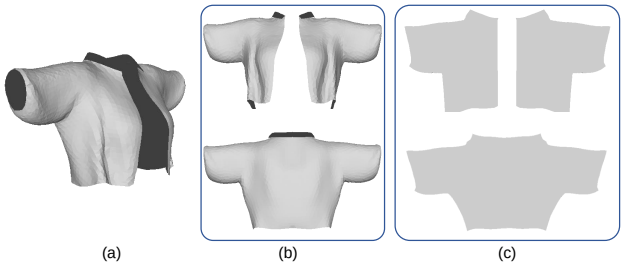

Figure 9. **The cutting and flattening for open shirt.** (a) The original mesh. (b) The cut surfaces. (c) The flattened panels.

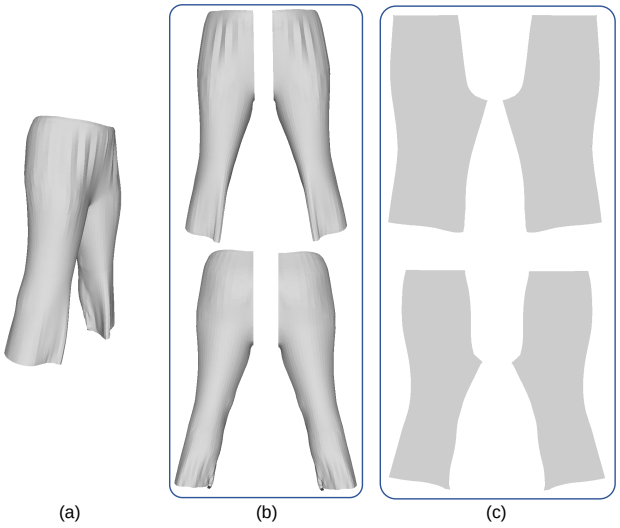

Figure 10. **The cutting and flattening for trousers.** (a) The original mesh. (b) The cut surfaces. (c) The flattened panels.

- In skirts, we observed that no additional constraint was necessary to produce consistent patterns, other than aligning them with a rotation and a transformation.

For trousers and shirt openings, we combine left and right side patterns with a constant translation between them to form the front and the back panels. We have found this setting to produce patterns closest to the ones commonly used in the garment fabrication industry. The resulting patterns for each category are illustrated in Figures 9(c), 10(c) and 11(c).

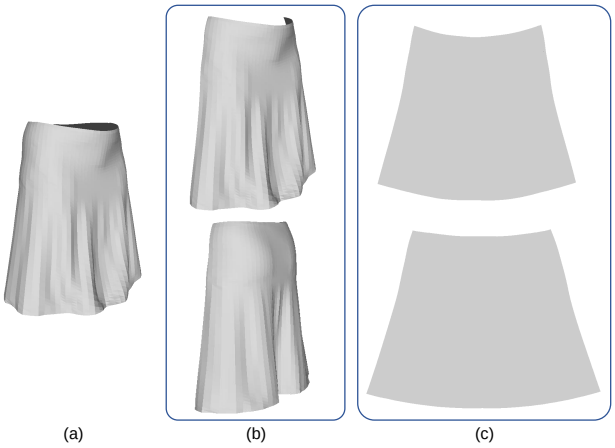

Figure 11. **The cutting and flattening for skirt.** (a) The original mesh. (b) The cut surfaces. (c) The flattened panels.

7. Comparison with ECON

In Fig. 12, we show the comparison between the reconstruction of our method and ECON [53]. Our method achieves more realistic reconstructions, which is particularly evident in the second example. Thanks to the designed fitting process that utilizes the prior captured by the deformation model and the physics-based loss, we are able to recover plausible deformations that extend to both the visible and occluded portions of the clothing. Furthermore, our reconstructions yield separated garment layers that are animatable, whereas ECON does not offer this capability.

8. Recovered Patterns

In Fig. 13, we show a comparison between the occupancy maps estimated by the deformation model $\mathcal{D}$ and those recovered through the optimization process described by Eq.

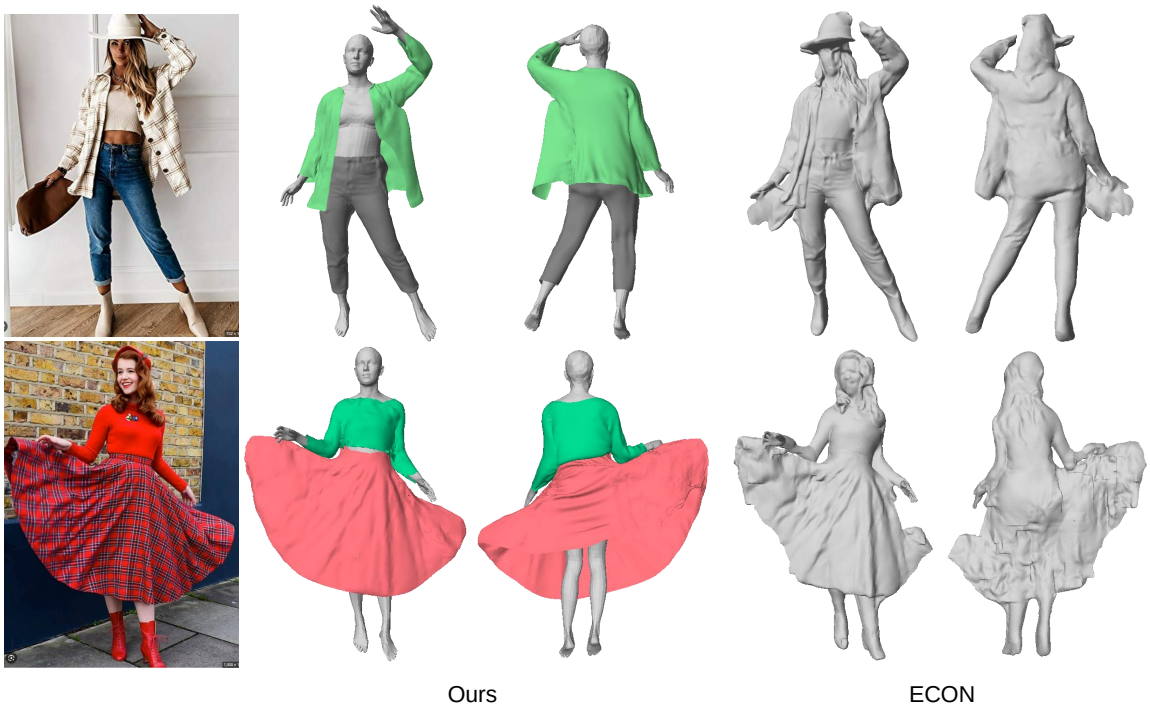

Figure 12. **Comparison with ECON [53].** Our method can recover realistic separated open surfaces for garments, while ECON generates closed meshes that encompass both the garments and the body.

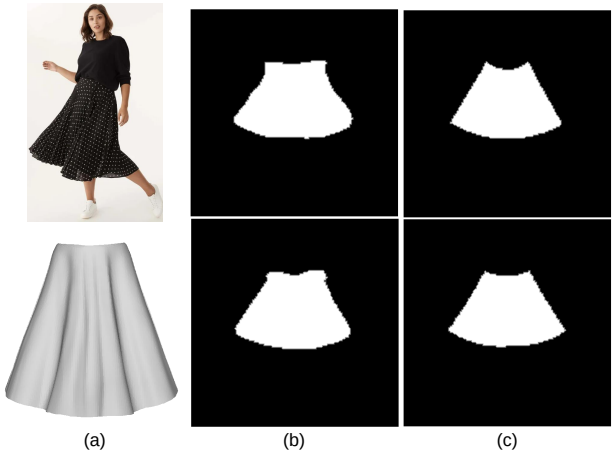

Figure 13. (a) The input image (top) and the corresponding rest mesh recovered by ISP (bottom). (b) The occupancy maps predicted by the deformation model $\mathcal{D}$. (c) The occupancy maps recovered by the optimization of the latent code of ISP (Eq. 9).

9. While the estimated occupancy maps may contain noise, we are able to recover the garment latent code from them with ISP, producing maps that best matches the estimation.

9. Hyperparameters

In this section, we introduce the hyperparameters for the training and the optimization.

- For the training of the deformation network $\mathcal{D}$, we set $\lambda = 0.05$ in Eq. 8.
- For the optimization of Eq. 9, we set $\lambda_{\mathbf{z}} = 0.04$ and run 1000 iterations with the Adam optimizer and the learning rate of 10^{-3}.
- For the optimization of Eq. 10 with respect to $\mathcal{D}$, we set $\lambda_C = 0.2$, $\lambda_n = 1$, $\lambda_p = 1$, and run 300 iterations with the Adam optimizer and the learning rate of 10^{-4}. For the optimization with respect to the coordinates of the garment mesh vertices, we run 400 iterations with the same hyperparameters.
- The resolution of the UV maps, the occupancy maps and the position maps is 128×128.
- For each garment category, i.e. shirts, skirts and trousers, we train one separate set of ISP and the deformation model $\mathcal{D}$.

10. Discussion

Since our method relies on the estimation of SMPL body parameters and garment normals, inaccuracies in these estimations can lead to incorrect garment reconstructions. As illustrated in Fig. 14, a failure case arises from the erroneous estimation of SMPL parameters, where the recon-

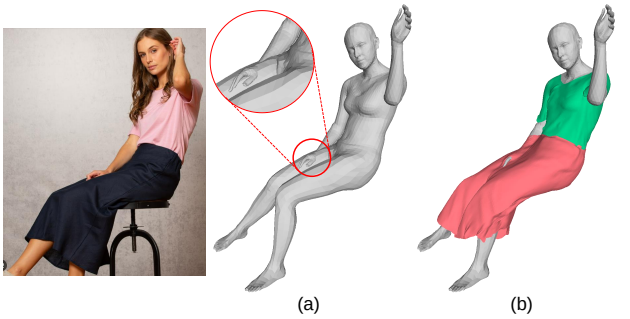

Figure 14. (a) The self-interpenetration of the estimated SMPL body leads to (b) the intersection between the reconstructed skirt and the hand of the body.

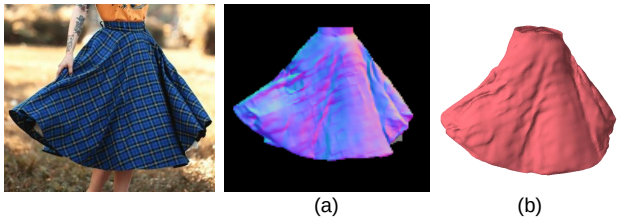

Figure 15. (a) The inaccurate normal estimation of the skirt leads to (b) the artifacts on the reconstructed surface.

structed skirt intersects with the hand due to the interpenetration of the body's hand and leg. In Fig. 15, we show the result with incorrect normal estimation. The garment's texture was misclassified as the shadow by the algorithm of [53], resulting in surface artifacts. Due to the normal consistency loss in the optimization of Eq. 10, our reconstruction retains this artifact.