DOI: 10.1111/cgf.14804

COMPUTER GRAPHICS *forum*
Volume 42 (2023), number 6, e14804

Garment Model Extraction from Clothed Mannequin Scan

Qiqi Gao and Takafumi Taketomi

CyberAgent, AI Lab, Tokyo, Japan
{gao_qiqi, taketomi_takafumi}@cyberagent.co.jp

Abstract

Modelling garments with rich details require enormous time and expertise of artists. Recent works re-construct garments through segmentation of clothed human scan. However, existing methods rely on certain human body templates and do not perform as well on loose garments such as skirts. This paper presents a two-stage pipeline for extracting high-fidelity garments from static scan data of clothed mannequins. Our key contribution is a novel method for tracking both tight and loose boundaries between garments and mannequin skin. Our algorithm enables the modelling of off-the-shelf clothing with fine details. It is independent of human template models and requires only minimal mannequin priors. The effectiveness of our method is validated through quantitative and qualitative comparison with the baseline method. The results demonstrate that our method can accurately extract both tight and loose garments within reasonable time.

Keywords: cloth modelling, digital geometry processing, mesh segmentation, modelling

CCS Concepts: • Computing methodologies → Shape modelling

1. Introduction

3D modelling and authoring are important research topics in augmented reality (AR) and virtual reality (VR). Especially, 3D modelling of real-world objects can be used to construct photo-realistic virtual environments. Recently, 3D scanning technologies such as Smartphones' Lidar cameras and photogrammetry (structure-from-motion and multi-view stereo) have been widely adopted to digitize real-world objects.

Specifically, virtual garments have received much attention as they are playing an increasingly significant role in AR/VR applications. 3D modelling of highly realistic clothing is a pre-requisite for dressing and animating clothed avatars [XPB*21]. E-commerce platforms are beginning to embrace virtual try-on for better shopping experience and lower product return rates [GCH*12]. 3D virtual garments can be created from 2D sewing patterns with CAD software such as Marvelous Designer. However, manipulating such tools generally requires a considerable amount of expertise and effort from CG artists or fashion designers, especially when one tries to model garments with rich details.

In order to find a simple yet effective way to model 3D garments, a variety of research has been carried out over the last decade. Some image-based methods re-construct 3D virtual garments from single-view images [ZCF*13, JHK15, DDÖ*17, YPA*18] or a few images [XYS*19] from different views. Many existing work of this category employ certain garment templates and recover the garments in the input images by deforming the pre-defined garment templates [CZL*15, DDÖ*17, YPA*18, XYS*19, SWY*22]. Others take 2D sewing patterns [BGK*13, SLL20] or sometimes sketches [CXJL21] as input to generate 3D garments. While input images for these methods can be easily acquired and prepared, image-based methods capture limited details and are generally unable to generate high-fidelity garment models.

On the other hand, thanks to rapidly evolving scanning techniques, 3D scans nowadays are much more easier to acquire than before and are thus gaining popularity. In both the CG industry and the apparel business, it has become increasingly common for people to scan clothed mannequins for the purpose of digitizing real-world clothing. Virtual garments are then obtained via segmentation of these clothed mannequin scans. In this work, we aim to acquire high-fidelity virtual garments from real-world clothing.

Despite the popularity of this setup, segmentation of clothed mannequin scans generally requires tedious manual labour. With 3D modelling software such as RealityCapture [Cap23], one can extract the target object from a re-constructed scene using a

Figure 1: *Our method can handle both tight and loose garments. Given a 3D scan of a clothed mannequin, our system extracts a garment mesh by colour-based boundary tracking, and then automatically categorizes the garment into tight or loose. For a loose garment, geometry-based boundary tracking is executed to refine the extracted mesh.*

bounding box as shown in Figure 2. This kind of interface/tools can be used for extracting the entire shape of the target object from the scene. Nonetheless, extracting a part of a scanned object, such as extracting clothing from a scanned human body, is often a laborious process. Developing automatic solutions to this problem remains a challenging task, and much effort has been made in this area. Most existing methods [PMPHB17, BTTPM19, BKL21] of this category try to extract garments from clothed human scans in various poses. To handle pose variety, they rely on a certain body template such as the SMPL model [LMR*15] and thus fail to perform as well on loose garments whose topologies differs greatly from the body. On the other hand, in the case of high-fidelity garment re-construction using a mannequin, we deem it unnecessary to consider pose variety and thus unnecessary to resort to rich priors such as SMPL.

In this study, we propose a simple yet practical solution to the problem of garment extraction from clothed mannequin scans, which employs minimal mannequin priors and works for both tight and loose garments. Similar to Bang *et al*. [BKL21], our approach also focuses on the boundaries between garment and skin rather than per-node classification. While Bang *et al*. [BKL21] uses only colour information in its exploration of such boundaries, our approach incorporates geometric information into this process. An example of garment extraction with our proposed method is illustrated in Figure 1. Our proposed method is capable of producing highly realistic textured garment models.

Our main contributions are three-fold:

- We present an approach to acquire high-fidelity 3D models of off-the-shelf clothing using a mannequin with simple set-up.
- We propose a novel metric for estimating the looseness of a garment.
- Our key contribution is that we propose a novel method for tracking boundaries of garments, which takes both colour and geometric information into account.

Our approach does not rely on any form of body template and leverages both the colour information and the geometric information. This not only frees our approach from the need for time-consuming template registration, but also allows for higher accuracy when extracting loose garments whose topologies differ greatly from the body.

2. Related Work

In this section, we discuss various garment modelling methods. Existing methods can be classified into two categories based on their input: 2D images or 3D scan data.

2.1. Garment capture from 2D images

Yang *et al*. [YPA*18] dress a parametric human body model with 2D garment patterns and optimizes 2D garment shape parameters based on the re-constructed human body and image information from a single-view image. A similar work [SWY*22] takes it a step further to generate garments with realistic dynamics by fitting a simulated garment model to the garment observed in the input video frames. Nevertheless, the range of garments recoverable through these methods is limited by the garment templates available.

During the last decade, there has been a noticeable trend towards leveraging the learning power of neural networks for the task of garment capture from images. Xu *et al*. [XYS*19] propose a multi-task learning network that can identify the garment landmarks and segment the garment semantically at the same time. A pre-defined garment template mesh is then deformed according to predicted landmarks to re-construct the 3D garment model. This method requires a minimum of two images and can work on images of both garments on human body and laid on a flat surface. Zhu *et al*. [ZCJ*20] provide a data-set of 2078 models

re-constructed from real garments, for evaluation of image-based garment re-construction. Although this proves to be an interesting area of research, it is theoretically difficult to re-construct high-fidelity garment models from single-view or a relatively small number of images.

2.2. Garment modelling from 3D scan

An early work [BPS*08] of this category recovers a garment mesh with a multi-view stereo re-construction algorithm from a relatively small number of images. This mesh has many holes and boundaries with missing geometry. A smooth garment mesh is then acquired through template-based completion. As the first marker-free approach to capture off-the-shelf garments, this method tends to remove the fine details in the geometry. Chen *et al*. [CZL*15] first obtain a rough garment mesh by KinectFusion and interactive cutting. It then synthesizes a garment by detecting garment components on a reference image and assembling pre-defined corresponding garment components in a database. The assembled garment is fit to the initial garment mesh to acquire a final mesh. This method has a notable limitation that instead of recovering the target mesh itself, it can only reproduce a garment that resembles the target garment in structure.

Some more recent works try to solve the problem of mesh segmentation on 3D scans of clothed human. Currently, graph-based methods [PMPHB17, BTTPM19] are the most popular among them. They usually utilize weak garment prior and perform vertex-level mesh segmentation by solving the Markov Random Field (MRF). Specifically, Pons-Moll *et al*. [PMPHB17] use a sequence of 3D scans (4D scan) to improve the accuracy of vertex classification; Bhatnagar *et al*. [BTTPM19] lift image-based semantic segmentation to the 3D domain by incorporating a CNN-based data term when solving an MRF on the UV-map of the SMPL model. Despite all the improvement, mis-classification still often occurs within the garment region.

Different from the popular graph-based segmentation approach, Bang *et al*. [BKL21] managed to explicitly find the boundaries between garment and skin or between garment parts through curve optimization, which outperformed previous methods by generating much smoother boundaries. However, similar to Refs. [PMPHB17, BTTPM19], we observed that their method does not work as well on loose garments as on tight ones. Besides, Bang *et al*. [BKL21] also suffers from occlusion, so that invisible area of a garment piece, *e.g*. upper part of a pair of pants under a T-shirt, can never be recovered. As a result, garment models recovered from clothed human scan of arbitrary poses may be incomplete. Goto and Umetani [GU21] used neural network to estimate garment patterns from a 3D scan of a clothed human. Their method essentially worked on multi-view projected images of the scan by learning to predict the probability of each pixel belonging to certain type of patterns, *e.g*. sleeve, front or back. The classification was of limited accuracy, and it was not clear how well it could perform on clothing other than T-shirts.

In contrast, our mannequin takes T-pose and is clothed with a single garment each time, so that both occlusion and overlapping of different layers of garments are negligible, *i.e*. our method is more capable of extracting the complete geometry of a garment than previous methods. More importantly, our method does not require any form of body template and, thus, has the potential for performing better at modelling loose garments.

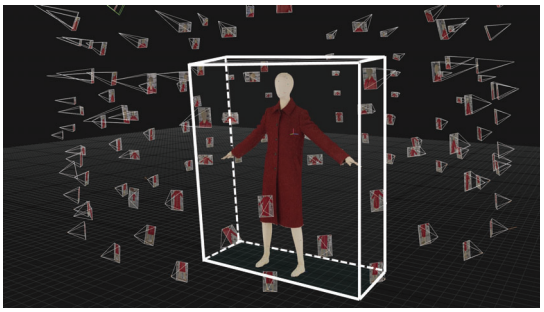

Figure 2: *An example of photogrammetry-based re-construction using RealityCapture.*

3. Method

Following the common practice of the CG industry and the apparel business, we put a target garment on a mannequin with a fixed pose and acquire the 3D mesh of the clothed mannequin (see Figure 2). After scanning, we extract the target garment mesh region from the scanned mannequin model by detecting boundary edges between the mannequin and garment regions. While colour information alone may be enough for extraction of tight clothing and geometric information alone may be enough for extraction of certain loose clothing (*e.g*. through mesh boolean operation), yet in order to deal with both tight and loose clothing with a single pipeline, we incorporate both colour and geometric information into the extraction process. As shown in Figure 3, our method consists of two stages: tight and loose garments extraction stages. We employ the coarse-to-fine strategy to extract tight and loose garments. First, our method extracts boundary candidates using colour information. Extracted boundary candidates are then filtered and refined using the mannequin priors. During the extraction process, the input garment is automatically categorized into tight or loose. If the input garment is loose, we apply the additional refinement process to get more accurate boundary edges. The details of our algorithm are explained in the following sections.

3.1. Initial boundary curve detection

First, we roughly categorize vertices of the input scanned mesh into mannequin skin and garment using colour information. We then acquire the initial boundary curves as shown in Figure 4. Note that actual boundaries exist between mannequin skin and garment triangles, which are represented as blue dotted curves in Figure 4. In this study, we aim to extract all the garment meshes (yellow regions in Figure 4) with minimal loss. Therefore, target boundary curves pass through the red edges in Figure 4. Concretely, we define the

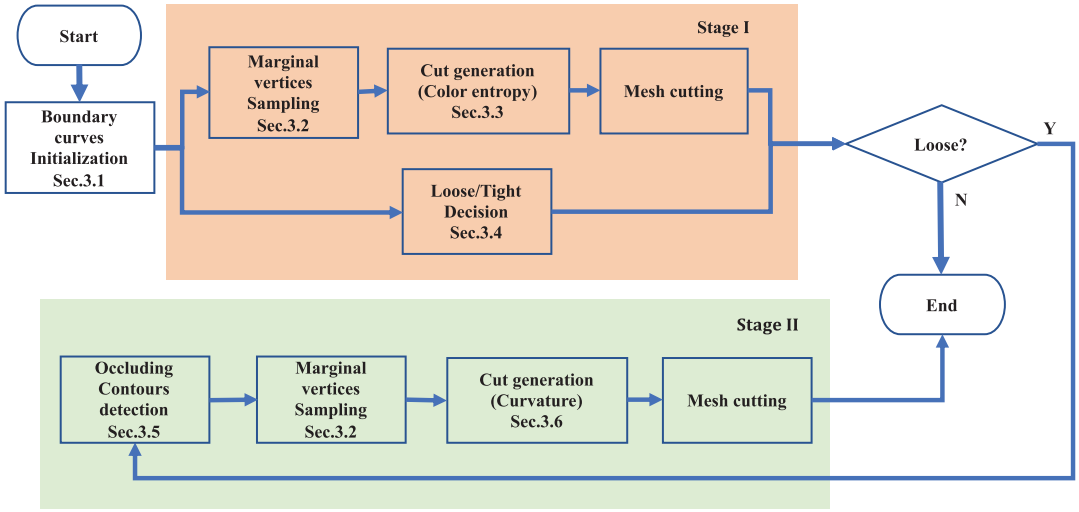

Figure 3: *Overview of the proposed method. Our framework is composed of two stages. Tight garments can be extracted in the stage I. The stage II is executed for handling loose garments.*

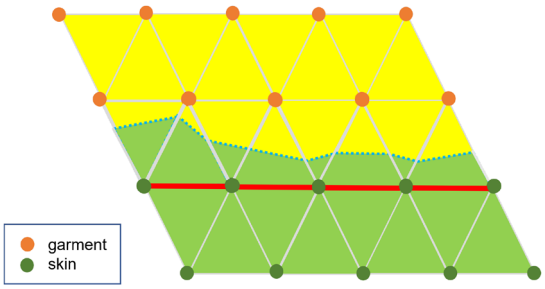

Figure 4: *An example of an initial boundary curve between garment and skin. It consists of mesh edges highlighted in red. The corresponding actual boundary is illustrated with the dotted blue curve.*

boundary curves between mannequin and garment as sequences of mesh edges which satisfy both of the following conditions:

1 Both of the two end vertices of the edge are labelled as mannequin skin.
2 Two vertices opposite to the edge belong to different classes.

To classify vertices, similar to Pons-Moll *et al*. [PMPHB17], we employ HSV colour space instead of RGB to gain more robustness against illumination changes. The classification of vertices is based on simple K-Means clustering. For an input scan of clothed mannequin V, each of its vertex $\boldsymbol{v}_i$ ($i \in \{0, 1, \ldots, N_V\}$, where N_V is the number of vertices in $\boldsymbol{V}$) is assigned a label $l_i \in \{0, 1\}$, where 0 and 1 correspond to mannequin skin and garment, respectively. Figure 6b illustrates an example of the result of classification.

Ideally, each initial boundary curve should correspond to one opening of the garment such as the neckline and the hemline. However, extracted initial boundaries may include unwanted curves when the input garment has texture patterns such as logos. As a result, it is difficult to obtain accurate initial boundaries only from colour information. To solve this problem, we use the mannequin priors to filter out unwanted boundaries. These mannequin priors are defined using pre-defined landmark positions on the naked mannequin (see Figure 5).

We first locate the corresponding body part for each initial boundary curve by comparing the centre position of curve vertices and landmark positions. The curves which are located higher than the jaw position or lower than the ankle position are then removed. In addition, we exclude the initial boundary curves shorter than a threshold th_{curve}. We set the value of th_{curve} for each initial boundary curve based on the perimeter of its corresponding body part. For example, th_{curve} for an initial boundary curve around a leg is set to the perimeter of an ankle. The example result of filtering is shown in Figure 6d.

After landmark-based and length-based filtering, there may still be unwanted curves due to texture patterns or noise within the garment region. We then detect initial boundary curves which satisfy both of the following conditions:

(i) It should be an extreme (*i.e*. uppermost, down-most, leftmost or rightmost) initial boundary curve. As indicated in Bradley *et al*. [BPS*08] and Zhu *et al*. [ZCJ*20], the number and the position of boundary loops differ across garment types. In our pipeline, a garment is first categorized into either upper-wear/one-piece or bottom-wear based on the existence of neck-related initial boundary curves. Next, among multiple initial boundary curves around the same body part, we locate the most extreme one based on central position of the curves and the category of the garment (see Table 1). Specifically, among waist-related curves, we select the down-most one for an upper-wear/one-piece, and the upper-most one for a bottom-wear; among initial boundary curves around the torso or the legs, we always select the down-most ones; among

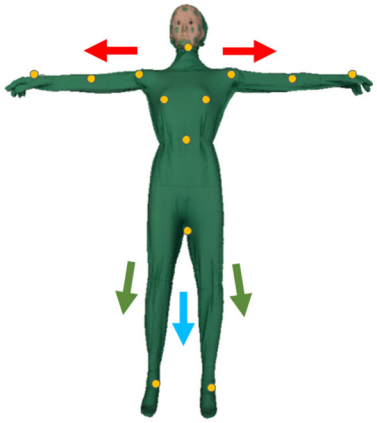

Figure 5: *The naked mannequin with pre-defined landmarks is highlighted in yellow. The directions of the arms, the torso and the legs, which are calculated from landmarks, are illustrated with arrows in red, blue and green, respectively.*

initial boundary curves around the left/right arm, we select the leftmost/rightmost one for an upper-wear/one-piece garment.

(ii) Its orientation should match the direction of its corresponding body part. First, we calculate the orientation of the curve through principal component analysis. The eigenvector which corresponds to the minimal eigenvalue is regarded as the direction of the curve. To verify curves, we define the body-part directions of the mannequin in advance (see Figure 5). The curve direction and each body-part direction are compared and the body part with minimal direction error with the curve is selected. We then check if the body part with minimal direction error matches the location of the curve.

We select the curve satisfying both of the above-mentioned conditions as a final candidate for openings as shown in Figure 6e.

3.2. Marginal vertices sampling

Initial boundary curves are jaggy due to miss-classification near actual boundaries. It is especially difficult to classify concave regions on the mesh accurately because these regions are not observed during the scanning process (see Figure 7). Both their shape and colour are interpolated from their surroundings. To obtain smoother and more reliable boundary curves, we sample 'marginal' vertices from initial boundary curves through geometric analysis.

To specify marginal vertices, we project an initial boundary curve $\boldsymbol{c}$ onto the plane which is perpendicular to the direction of the corresponding body part. Figure 8 illustrates $\boldsymbol{c}'$, the 2D profile of the projected initial boundary curve. On the projection plane, we create a polar coordinate system, whose origin O is the projection of the centre of $\boldsymbol{c}$. We further divide the polar angles of the projection plane into M intervals ($M = 32$). Vertices of $\boldsymbol{c}'$ are then distributed into one of the M intervals based on their polar angles. Subsequently, marginal vertices are sampled based on their radial distance. In each interval $m_i (i \in 1, 2, \ldots, M)$, we sample the outermost vertex as the

Figure 6: *(a) The input scan of a clothed mannequin. (b) The result of vertex classification. Vertices labelled as garment and mannequin skin are represented in yellow and green, respectively. (c) Extracted initial boundary curves based on vertex labels. (d) Initial boundary curves after filtering using landmark positions. (e) Boundary curves chosen as final candidates for garment openings. They are the uppermost, down-most, leftmost and rightmost initial boundary curves.*

Table 1: *Body parts corresponding to extreme boundary curves for commonly seen garment types.*

Body part	Upper-wear	One-piece (*e.g.* coat, dress)	Skirt	Trousers
Uppermost	Neck	Neck	Waist	Waist
Down-most	Waist	Torso	Torso	Left leg and right leg
Leftmost	Left arm	Left arm	—	—
Rightmost	Right arm	Right arm	—	—

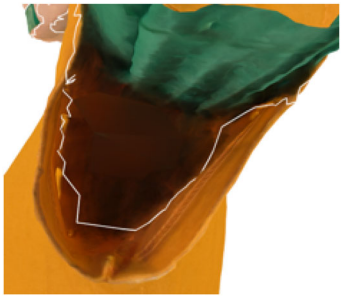

Figure 7: *An example of an initial boundary curve along the sleeve, which is too noisy to approximate the actual boundary.*

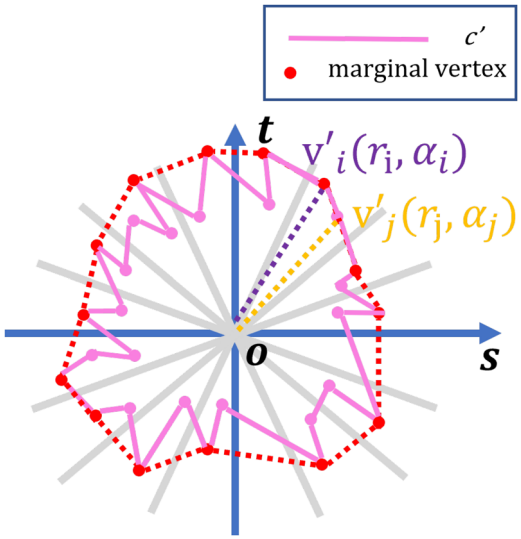

Figure 8: *An example of marginal vertices sampled with our strategy.* c', *the projection of an initial boundary curve* c *on the plane perpendicular to the body, is illustrated with the pink solid curve. The sampled marginal vertices are highlighted in red. Note that* $\mathbf{v'}_i$ *is sampled as a marginal vertex instead of* $\mathbf{v'}_j$*. While* $\mathbf{v'}_i$ *and* $\mathbf{v'}_j$ *fall into the same interval based on their polar angles,* $\mathbf{v'}_i$ *has a larger radial distance than* $\mathbf{v'}_j$ *(*$i, j \in \{0..N_V\}$*,* N_V *is the number of vertices in the input scan).*

marginal vertex. These marginal vertices are used as control points in the boundary tracking process. As shown in Figure 9 (left), for computational efficiency, we create a bounding box for vertices of the initial boundary curve and only utilize meshes within this bounding box in the following process.

3.3. Colour-based boundary tracking

Using marginal vertices as control points, we detect actual boundary vertices through colour-based boundary tracking. In general, a vertex is likely to lie on the actual boundary if half of its neighbouring vertices are labelled as garment and the other half as the mannequin skin.

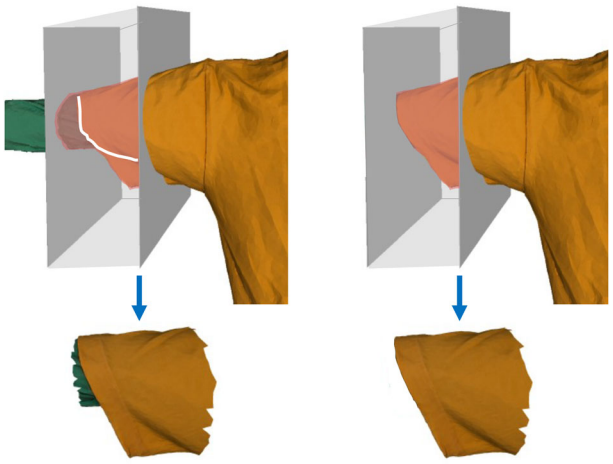

Figure 9: *(Left) For an initial boundary curve (the same curve shown in Figure 7) along a sleeve opening, highlighted in white, only meshes within the bounding box are utilized in colour-based boundary tracking. One pair of opposite faces of the bounding box are perpendicular to the arm. (Right) The same bounding box is re-used to locate meshes required for occluding contours detection and loose-boundary tracking.*

In order to find the weighted shortest path on the mesh patch between marginal vertices, for each vertex of the mesh, we define an entropy-like quantity as follows:

$$\phi(v_i) = -(p \log_2(1-p) + (1-p)\log_2 p) \tag{1}$$

where p is the probability for a vertex in v_i's one-ring neighbourhood to take the label of garment.

The weight w_{colour} of a mesh edge $\boldsymbol{e} = [v_i, v_j]$ is defined as

$$w_{colour}(i, j) = \frac{1}{\phi(v_i) + \phi(v_j) + \epsilon} \tag{2}$$

where $i, j \in \{0..N_V\}$, N_V is the number of vertices in the input scan, ϵ is a small positive number ($\epsilon = 0.0001$).

The weighted shortest path between sampled vertices can then be computed with Dijkstra's algorithm [Dij59]. The garment is extracted by cutting the input scan along these shortest paths.

3.4. Garment looseness evaluation

During the re-construction of a clothed mannequin, some photogrammetry software solutions automatically enclose the volume to create a 'watertight' mesh. As a result, since the topologies of loose garments vary greatly from the body, the input scan may contain gap-filling triangles between the body and the garment. These regions are not observed during the scanning process and such gap-filling triangles are often with meaningless colour. Loose garments, therefore, should be handled with a strategy different from tight garments.

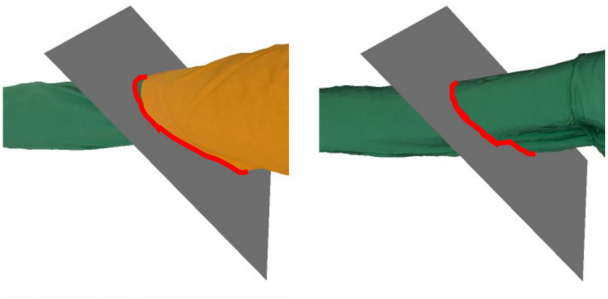

Figure 10: *An example of tight-or-loose decision for a boundary curve along a sleeve opening. The intersection curve between the best-fitting plane of the boundary curve and the clothed mannequin is highlighted in red (left). The intersection curve between the best-fitting plane of the boundary curve and the naked mannequin is highlighted in red (right).*

In this paper, by 'loose garments', we refer to garments with at least one loose boundary curve. Now we describe the procedure of tight-or-loose decision for each boundary curve.

For each boundary curve, we find its best-fitting plane and intersect the plane with both the naked and the clothed mannequin to acquire two sets of closed intersection curves as shown in Figure 10.

We can then define the looseness, η , of a boundary curve as follows:

$$\eta = \frac{\Omega_{clothed}}{\Omega_{naked}} \tag{3}$$

where $\Omega_{clothed}$ and Ω_{naked} are the unsigned area enclosed by the intersection curves between the fitted plane and the static scan of the clothed and the naked mannequin, respectively. A boundary curve between mannequin and garment that satisfies the following condition is assumed to be loose:

$$\eta > \lambda \tag{4}$$

where λ is a pre-defined threshold for differentiating loose and tight boundary curves. During the experiments, we set $\lambda = 1.20$. For simplicity, convex hulls of these two intersection curves are used instead of the curves themselves in our implementation.

Note that since necklines may involve complex structures such as collars and hoods, definition of necklines may become ambiguous. We, therefore, assume that all boundary curves around the neck are tight boundaries regardless of their looseness.

3.5. Occluding contours detection

After cutting the input scan along colour-based boundaries, loose garments may still contain gap-filling triangles near the cuts as shown in Figure 11 (left). In stage II of our pipeline, loose garments are to be refined by taking occluding contours into account.

Suppose, we place an anchor near the centre of the opening, as shown in Figure 11 (right), the occluding contours viewed from the anchor turn out to be a better approximation of the actual boundary

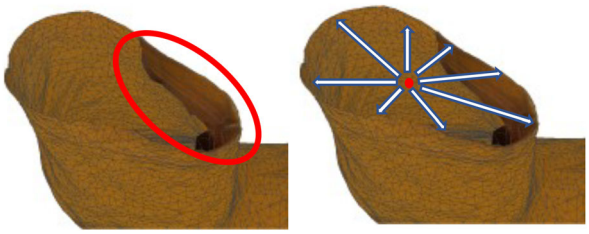

Figure 11: *The gap-filling triangles (left) are expected to be removed if the mesh is cut again, along occluding contours viewed from an anchor, which is highlighted in red (right).*

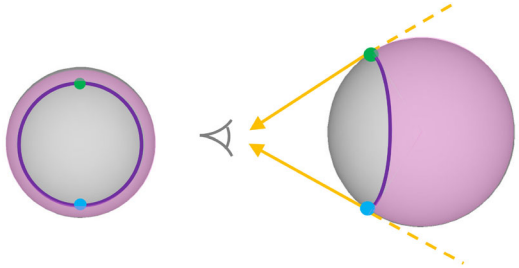

Figure 12: *A front view of an occluding contour on a sphere (left). A side view of the scene (right). The occluding contour is painted in purple. Two points on the occluding contour from a certain viewpoint are highlighted in green and blue, respectively. They are where depth discontinuity occurs. The occluded region on the sphere is painted in pink.*

at the garment opening than the original colour-based cut. There are a few ways to define the occluding contours or, simply contours [RDF05, BH19]. As illustrated in Figure 12, occluding contours are the locations where depth discontinuity occur when viewed from a certain viewpoint.

Intuitively, such occluding contours only need to be detected from the mesh patch around the garment opening. We locate the mesh patch with the same bounding box defined in stage I (see Figure 9 (right)). Similarly, we initialize the anchor with the centre of the initial boundary curve defined in stage I.

We now describe the procedure for detecting occluding contours. After anchor set-up, we create a plane Ψ whose normal is $\mathbf{n}$. Ψ passes through the anchor and is perpendicular to the corresponding body part as shown in Figure 13 (left). For a series of azimuthal angles $\alpha = \frac{2\pi i}{L}$ on Ψ ($i \in 0, 1, \ldots, L-1$, L is set to two times the number of vertices along the colour-based cut defined in stage I in our implementation), we find the point on the mesh patch with the highest elevation angle along each direction $(\cos\alpha, \sin\alpha)$. The loci of such points form the occluding contours. Nonetheless, since our objective is to find the sequences of mesh edges that best approximate garment openings, we search for the mesh edges where such points are located. Specifically, for each sampled azimuthal angle α, we compute the intersections between mesh edges in the patch and a plane Γ. The plane Γ is determined by $(\cos\alpha, \sin\alpha)$ and $\boldsymbol{n}$ as shown in Figure 13 (right). Among these intersections, we locate the intersection $\boldsymbol{Q}$ with the highest elevation angle θ and the mesh edge

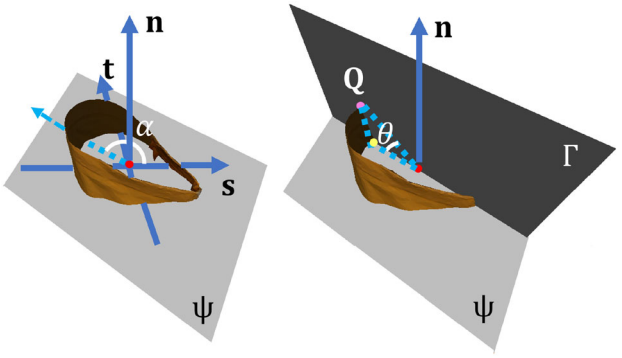

Figure 13: *(Left) A direction* $(\cos\alpha, \sin\alpha)$ *is sampled on the plane* Ψ. Ψ*, whose normal is* **n***, is perpendicular to the body part. (Right) A plane* Γ *is determined by* $(\cos\alpha, \sin\alpha)$ *and* **n***. Among all intersections between* Γ *and mesh edges on the mesh patch, Q has the highest elevation angle. The anchor is highlighted in red.*

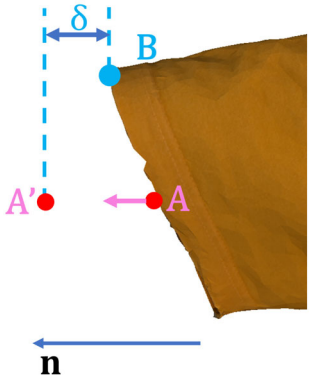

Figure 14: *During the anchor fine-tuning process, we search for the optimal anchor position on the line segment between its initial position A and its upper-bound A′ by shifting the anchor a small step each time along* **n** *and evaluating the goodness of the position.* **n** *is the direction of the body part. The anchor points are highlighted in red.*

where $\boldsymbol{Q}$ lies. By projecting the collection of end vertices of such edges onto plane Ψ, we can sample the marginal vertices from them with the same strategy described in Section 3.2. These marginal vertices are the control points in our loose-boundary tracking process.

Note that occluding contours are view-dependent. We introduce an additional anchor fine-tuning process to gain robustness against noisy boundaries. To that end, we measure the goodness of an anchor position based on how much more information we can gain about the location of actual garment opening, other than information already provided by the colour-based cut. In other words, we evaluate an anchor position based on the number of unique vertices which are on the occluding contours detected from the anchor but not on the colour-based cut.

As illustrated in Figure 14, from the initial position A, we shift the anchor by a small step (the step size is set to half of the average edge length of the input mesh) each time along \boldmath n, the direction of the body part, to find the optimal position between A and A' based on its goodness. A' is the upper-bound of A, determined by B, the most extreme vertex on the mesh patch along \boldmath n, and a small displacement δ. For each garment opening, δ is defined as follows:

$$\begin{aligned} A' &= A + \mathbf{n} * \delta \\ \delta &= k * r \\ r &= \sqrt{\frac{\Omega_{clothed}}{\pi}} \end{aligned} \tag{5}$$

where r is the estimated 'radius' of the boundary, deduced from $\Omega_{clothed}$ (see Equation 3); k is a parameter controlling the impact of the radius. During the experiments, we set $k = 0.7$ because it enables successful garment extraction for all our test data. However, we confirmed that our method is insensitive to the value of k as long as k is within a reasonable range. The impact of k on the accuracy of extraction is discussed in detail in Section 6.1.

3.6. Loose-boundary tracking

Given marginal vertices sampled from occluding contours, we track loose boundaries by finding the weighted shortest path between them.

Based on our observation that the input mesh usually bends more near loose boundaries, we use the notion of curvature in definition of such weights. For a curve, curvature is the reciprocal of the radius of the circle that best approximates the curve locally. For a surface, intuitively, it is the amount by which the surface deviates from being a plane. The weight $w_{geometric}$ of a mesh edge $\boldsymbol{e} = [v_i, v_j]$ is defined as follows:

$$w_{geometric}(i, j) = \frac{1}{|H_i| + |H_j| + \xi} \tag{6}$$

where H_i and H_j are the mean curvature at v_i and v_j, respectively; $i, j \in \{0..N_V\}$, N_V is the number of vertices in the input scan, ξ is a small positive number ($\xi = 0.0001$).

4. Experiments

We prepared our test data by scanning a clothed mannequin in different types of clothing and evaluated our method both quantitatively and qualitatively against the baseline method [BKL21]. We implemented our algorithm using Python and C++. Libigl [JP*18], CGAL [The22] and OpenMesh [BSBK02] are used for geometry processing.

4.1. Data preparation

We test our method with a mannequin taking T-pose. A green tight suit is put on the mannequin before scanning to better differentiate skin and garment. The mannequin is scanned with a multi-view scanning system consisting of 134 cameras. The cameras are set up in a ring configuration.

Note that although our extraction technique takes the scan data of a clothed mannequin as input, it is totally independent of the data

acquisition method/system (*e.g*. photogrammetry or laser scanning). The large-scale camera rig we utilized can be replaced by multi-view images taken with a single camera.

Similarly, although a fixed mannequin and a colour-tight-suit are prerequisites for our extraction technique, they can be easily purchased nowadays. Moreover, sophisticated commercial/open-source photogrammetry software has become widely available recently. We hence believe that one has little difficulty preparing the input scanned data.

In total, the mannequin was scanned with 14 different pieces of garments, with 12 out of 14 being loose garments based on our looseness criterion. Each scan of clothed mannequin is with a single piece of garment so that no overlapping of multiple layers of clothes occurs. During the entire scanning process, the mannequin is kept roughly in a fixed pose and position so that the same set of mannequin priors can be re-used for all the clothed mannequin scans.

High-resolution 3D meshes of the clothed and naked mannequin are then re-constructed with RealityCapture [Cap23]. After scanned meshes are acquired, we prepare mannequin priors. These priors include: pre-defined landmarks on the scanned naked mannequin corresponding to joints; the body-part directions derived from landmark positions (see Figure 5). Among all our test data, each clothed mannequin scan has about 200K polygons.

4.2. Evaluation metric

We use precision and recall for evaluation of extraction accuracy. After locating true positive (TP), *i.e*. triangles on the extracted mesh which overlap the ground truth mesh, precision and recall of extraction can be computed as follows:

$$\begin{aligned} precision &= \frac{\Omega_{TP}}{\Omega_{extracted}} \\ recall &= \frac{\Omega_{TP}}{\Omega_{GT}} \end{aligned} \tag{7}$$

where $\Omega_{extracted}$ and Ω_{GT} are the area of extracted garment mesh and ground truth garment mesh, respectively. The ground truth meshes were created by manually extracting the garments from the scanned data.

5. Results

The accuracy of extraction with baseline method [BKL21] and our proposed method are reported in Table 2. Among 12 loose garments extracted with our proposed method, 11 and nine garments demonstrate accuracy higher than baseline method in terms of precision and recall, respectively. Some qualitative comparison results are illustrated in Figure 15. Please refer to Figures S1–S3 for complete comparison results. Note that some of the garments in our test data are categorized into 'loose' by definition while they are partially 'tight'. For example, in the case shown in Figure 15k, the boundary curves passing through the armpits are 'tight' while the boundary curve that corresponds to the hemline is 'loose'.

The baseline method [BKL21] extracts boundaries by locating zero-isolines of an implicit function defined over the input scan.

Table 2: *Extraction accuracy with baseline method and our proposed method for our test data. The looseness column refers to the maximal looseness of the garment's all openings (except neckline). Bold values indicate better results between the baseline and our proposed method.*

		Precision (%)		Recall (%)	
Garment	Looseness	Baseline	Ours	Baseline	Ours
(a) Tight-pants	1.142	99.999	**100.000**	98.715	**99.752**
(b) Hoodie	1.117	98.877	**99.940**	**99.999**	99.882
(c) Shirt	1.251	99.198	**99.748**	**99.098**	98.796
(d) Polo-Shirt	1.340	96.279	**99.867**	**99.907**	99.698
(e) Long-Tshirt	1.436	99.256	**99.892**	98.980	**99.829**
(f) Knit	1.453	98.485	**99.919**	99.046	**99.490**
(g) Tshirt	1.503	94.087	**99.949**	**100.000**	99.391
(h) Jacket	1.530	96.981	**99.601**	98.343	**99.875**
(i) Down-coat	1.871	100.000	**99.917**	97.961	**98.850**
(j) Cardigan	2.511	99.085	**99.948**	99.300	**99.371**
(k) Tanktop	2.684	98.587	**99.636**	95.984	**99.568**
(l) Wide-pants	2.891	99.084	**99.801**	99.006	**99.731**
(m) Trench-coat	3.467	98.366	**99.688**	98.722	**99.547**
(n) Long-skirt	5.181	99.581	**99.892**	96.084	**99.876**

Since these isolines usually traverse across mesh edges, instead of detecting sequences of existing mesh edges, the baseline method sub-divides the original triangles near detected isolines. It can, therefore, generate relatively smoother boundaries than our proposed method, and performs generally well on most of our scan data.

However, such smooth isolines turned out to be incapable of approximating curvy boundaries, usually leaving mannequin skin mesh unremoved around such boundaries (see the unremoved meshes along the neckline in Figure 15f,h,j). For certain garments, this poses a problem as the optimal boundary curves determined by their curve optimization algorithm may deviate significantly from actual boundaries, causing loss of useful structural information of the target garment. As illustrated in Figure 15k, the baseline method failed to preserve the shoulder-straps during skin-cloth segmentation.

Another key difference between the baseline method and our method is that their boundary optimization strategy uses a displacement function that is based only on colour information, whereas ours incorporates geometric information into the extraction of loose garments. As highlighted by the different extraction results around the opening at the sleeves in Figure 15g,m, our method achieves notably higher accuracy for loose garments with noisy gap-filling triangles by incorporating geometric information (occluding contours and curvature) into the pipeline. Furthermore, boundary initialization in the baseline method requires multiple sets of labelled region points, splines and sign points, involving complicated projection of these sign points and splines from SMPL template to the input scan. Meanwhile, the only mannequin priors required by our proposed method are the landmark positions of joints and perimeters of limbs.

Last but not the least, the baseline method relies on the accuracy of SMPL fitting [ZPBPM17], and the accuracy of the initial pose estimated with OpenPose library [CHS*19]. During the experiments, we confirmed that SMPL fitting of input scans made reasonably good estimation of pose/shape ('under clothing'). Nevertheless, this

Figure 15: *Examples of cases in which our proposed method performs noticeably better than the baseline method. Loss in garment meshes and unremoved skin meshes resulted from the baseline method and corresponding results generated with proposed method are highlighted with boxes in orange, lime and red, respectively.*

remains a tedious and time-consuming procedure. To the contrary, different from previous methods, our approach does not rely on any form of body template. This not only frees our approach from the need for time-consuming template registration, but also allows for more robustness when extracting loose garments whose topologies differ greatly from the body. Table 3 shows computation time for the baseline method and our proposed method. The results prove that our method can extract garments more efficiently than the baseline method. Note that computation time for the baseline method does not include the processing time for SMPL fitting. Therefore, the actual computation time of the baseline method is expected to be longer than it appears in the table.

Table 3: *Comparison of computation time for baseline method (boundary optimization) and our proposed method (garment extraction). Bold values indicate shorter computation time between the baseline and our proposed method.*

	Time (s)	
Garment	Baseline	Ours
(a) Tight-pants	529	**67**
(b) Hoodie	463	**78**
(c) Shirt	289	**141**
(d) Polo-Shirt	306	**112**
(e) Long-Tshirt	711	**118**
(f) Knit	316	**139**
(g) Tshirt	299	**109**
(h) Jacket	392	**206**
(i) Down-coat	330	**229**
(j) Cardigan	325	**140**
(k) Tanktop	305	**171**
(l) Wide-pants	157	**124**
(m) Trench-coat	315	**154**
(n) Long-skirt	456	**113**

6. Discussion

6.1. Anchor positioning

The extent to which detected occluding contours match the actual garment opening depends on the position of the anchor. The occluding contours detected by our method are not guaranteed to be optimal for all garments, especially when the input scan has noisy gap-filling triangles.

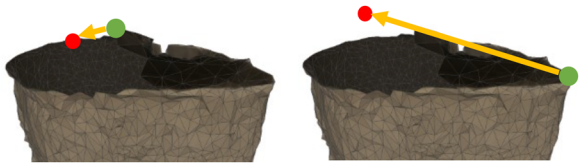

Figure 16: *Occluding contours are view-dependent: (left) occluding contours detected from the anchor correspond to colour-based cut generated in stage I; (right) occluding contours detected from the anchor correspond to actual garment opening. The anchors and points on occluding contours viewed from the anchors are highlighted in red and lime, respectively.*

Figure 17: *F1-score for garment extraction on our test data when k takes equidistant values ranging from 0.0 to 2.0. For most of our test data, when k fluctuates within such a reasonably wide range, accuracy of extraction remains relatively stable despite minor oscillation.*

As illustrated in Figure 16, when some of the gap-filling triangles are on the exterior of the garment opening, the anchor may needs to be shifted a little away from the boundary curve centre, towards exterior of the garment opening. By doing so, the occluding contours detected correspond to the garment opening instead of colour-based cut generated in stage I.

We have hence introduced an additional anchor fine-tuning process as described in Section 3.5. According to Equation (5), the range of anchor position is determined by a parameter k. We have tested the impact of k on the accuracy of extraction. As illustrated by Figure 17, the accuracy of our method, in terms of F1-score, is insensitive to the value of k when k is within a reasonable range. F1-score is defined as below:

$$F1 = 2 * \frac{precision * recall}{precision + recall} \tag{8}$$

Although a common value of k generates reasonably good results for all garments in our test data, as indicated by Figure 17, different garments may have slightly different optimal k values. Optimal occluding contours are more likely to be detected if a user can manually position the anchor properly with the aid of a real-time visualization tool for occluding contours. This is a topic out of the scope of this paper and remains to be investigated for future work.

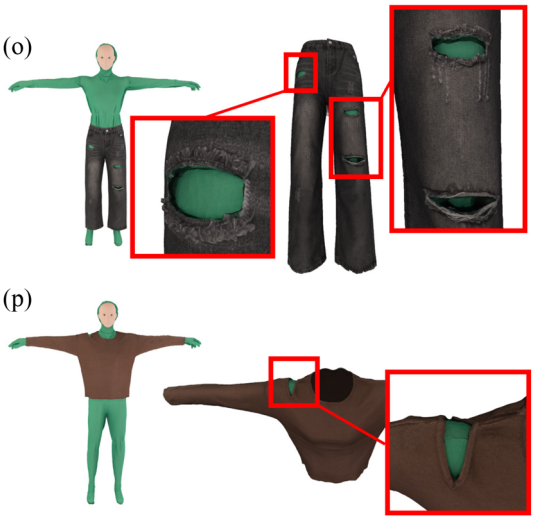

Figure 18: *Examples of failure cases for our proposed method: clothing with ornamental holes on the legs (o) and the shoulder (p). The skin between the cuts, which is extracted as a part of the garment, is highlighted with boxes in red.*

6.2. Limitations

Although performing generally well on our test data, including clothing with complex patterns such as stripes (see Figure 15f), our proposed method still has certain limitations. First of all, while our method takes a clothed mannequin scan as input, the re-construction of the clothed mannequin itself is out of the scope of this work. Our method assumes that the garments can be scanned properly and is not intended for garments that may fail to meet the requirement, *e.g.* garments made from semi-transparent materials. Another obvious limitation of our method is that it is targeted at regular clothing with specific numbers of openings, *e.g.* upper-wear/one-piece with four openings, pants with three openings and skirts with two openings. It cannot handle clothing with ornamental holes such as ripped jeans: the skin between the cuts would be extracted as a part of the garment (see Figure 18). A possible solution to surpass this limitation would be to enable the selection of boundary curves among candidates by users through a user interface, which we shall leave for future exploration.

7. Conclusion

In this work, we have introduced an easy-to-use two-stage pipeline for extracting garments from scan data of clothed mannequins, which enables the modelling of off-the-shelf clothing with fine details. Specifically, we have proposed a novel method for tracking both tight and loose boundaries with a single pipeline. Our algorithm is independent of both the data acquisition method and human template models such as the SMPL model and requires only minimal mannequin priors. We have demonstrated the effectiveness

Figure 19: *Examples of simulated garments using garment meshes extracted with our proposed method: (m) (left); (e) and (n) (middle); (g) and (a) (right). Note that both tight and loose garments deform following different poses of the avatars. Animation is created with Marvelous Designer[CLO23].*

of our method through quantitative and qualitative comparison with the baseline method.

It is noteworthy that while we have only confirmed our method using the set-up of a mannequin taking T-pose, our methodology can be easily applied to the setup of a mannequin taking A-pose.

As for application, assets of virtual garments extracted with our proposed method can be utilized directly in animation of clothed avatars (see Figure 19). Note that the garment mesh extracted with our method retains the geometric quality of the input mesh because it simply cuts the input mesh along its original edges. The quality of animation can be further improved after cleaning-up and retopology of the garment meshes.

We noticed that for certain garments, *e.g*. Figure 19 (left), some undesirable wrinkles are formed near the shoulders after reposing. Leveraging an A-pose mannequin is expected to help produce garments without such wrinkles, while being more likely to introduce more occlusion into regions such as armpits. A strategy to balance this trade-off between occlusion and wrinkles is yet to be devised. Such tasks, together with rigging, and skinning of garments are promising directions to explore in the future for realistic cloth simulation and real-time animation.

Acknowledgements

We would like to thank Seungbae Bang, Maria Korosteleva, and Sung-Hee Lee for generously sharing the code for the baseline method, and the anonymous reviewers for their extremely valuable comments and suggestions.

References

[BGK*13] Berthouzoz F., Garg A., Kaufman D. M., Grinspun E., Agrawala M.: Parsing sewing patterns into 3D garments. *ACM Transactions on Graphics (TOG) 32*, (2013), 1–12. https://doi.org/10.1145/2461912.2461975

[BH19] Bénard P., Hertzmann A.: Line drawings from 3D models: A tutorial. *Foundations and Trends®in Computer Graphics and Vision 11*, 1-2 (2019), 1–159. https://doi.org/10.1561/0600000075

[BKL21] Bang S., Korosteleva M., Lee S.-H.: Estimating garment patterns from static scan data. *Computer Graphics Forum 40*, (2021), 273–287. https://doi.org/10.1111/cgf.14272

[BPS*08] Bradley D., Popa T., Sheffer A., Heidrich W., Boubekeur T.: Markerless garment capture. *ACM Transactions on Graphics (TOG) 27*, 3 (2008), 1–9. https://doi.org/10.1145/1360612.1360698

[BTTPM19] Bhatnagar B. L., Tiwari G., Theobalt C., Pons-Moll G.: Multi-garment net: Learning to dress 3D people from images. In *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (2019), pp. 5419–5429. https://doi.org/10.1109/ICCV.2019.00552

[Cap23] Capture Reality Community: RealityCapture. https://www.capturingreality.com/realitycapture (2023). Accessed: 2022-06-07.

[CHS*19] Cao Z., Hidalgo Martinez G., Simon T., Wei S., Sheikh Y. A.: OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence 43*(1), (2021), 172–186. https://doi.org/10.1109/TPAMI.2019.2929257

[CLO23] CLO Virtual Fashion, Inc.: MarvelousDesigner. https://www.marvelousdesigner.com/ (2023). Accessed: 2023-02-12.

[CXJL21] Chen Y., Xian C., Jin S., Li G.: 3D Shape-adapted garment generation with sketches. In *Advances in Computer Graphics: 38th Computer Graphics International Conference, CGI 2021, Virtual Event, September 6–10, 2021, Proceedings*. Springer, Cham (2021), pp. 125–136. https://doi.org/10.1007/978-3-030-89029-2_10

[CZL*15] Chen X., Zhou B., Lu F.-X., Wang L., Bi L., Tan P.: Garment modeling with a depth camera. *ACM Transactions on Graphics (TOG) 34*, (2015), 1–12. https://doi.org/10.1145/2816795.2818059

[DDÖ*17] Danerek R., Dibra E., Öztireli A. C., Ziegler R., Gross M. H.: DeepGarment: 3D garment shape estimation from a single image. *Computer Graphics Forum 36*, (2017), 269–280. https://doi.org/10.1111/cgf.13125

[Dij59] Dijkstra E. W.: A note on two problems in connexion with graphs. *Numerische Mathematik 1*, 1 (1959), 269–271.

[GCH*12] Giovanni S., Choi Y. C., Huang J., Tat K. E., Yin K.: Virtual try-on using kinect and HD camera. In *Motion in Games: 5th International Conference* Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34710-8_6

[GU21] Goto C., Umetani N.: Data-driven Garment Pattern Estimation from 3D Geometries. In *Proceedings of the 42nd Annual*

Conference of the European Association for Computer Graphics, Eurographics 2021 - Short Papers, Vienna, Austria, May 3-7, 2021 (2021), Eurographics Association, pp. 17–20. https://doi.org/10.2312/EGS.20211013

[JHK15] Jeong M.-H., Han D.-H., Ko H.: Garment capture from a photograph. *Computer Animation and Virtual Worlds 26*, (2015), 291–300. https://doi.org/10.1002/cav.1653

[JP*18] Jacobson A., Panozzo D., et al.: libigl: A simple C++ geometry processing library. https://libigl.github.io/, (2018). Accessed: 2023-02-13.

[BSBK02] Botsch M., Steinberg S., Bischoff S., Kobbelt L.: OpenMesh: A generic and efficient polygon mesh data structure. https://www.graphics.rwth-aachen.de/software/openmesh/, (2002). Accessed: 2023-02-13.

[LMR*15] Loper M., Mahmood N., Romero J., Pons-Moll G., Black M. J.: SMPL: A skinned multi-person linear model. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia) 34*, 6 (Oct. 2015), 248:1–248:16.

[PMPHB17] Pons-Moll G., Pujades S., Hu S., Black M. J.: ClothCap: seamless 4D clothing capture and retargeting. *ACM Transactions on Graphics 36*, (2017), 73:1–73:15. https://doi.org/10.1145/3072959.3073711

[RDF05] Rusinkiewicz S., DeCarlo D., Finkelstein A.: SIGGRAPH 2005 Course 7: Line drawings from 3D models. https://gfx.cs.princeton.edu/proj/sg05lines/ (2005). Accessed: 2022-06-02.

[SLL20] Shen Y., Liang J., Lin M. C.: GAN-based garment generation using sewing pattern images. In *Computer Vision - ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII*. Springer, Berlin, Heidelberg (2020). https://doi.org/10.1007/978-3-030-58523-5_14

[SWY*22] Su Z., Wan W., Yu T., Liu L., Fang L., Wang W., Liu Y.: MulayCap: Multi-layer human performance capture using a monocular video camera. *IEEE Transactions on Visualization and Computer Graphics 28*, 4 (2022), 1862–1879. https://doi.org/10.1109/TVCG.2020.3027763

[The22] The CGAL Project: *CGAL User and Reference Manual*. https://doc.cgal.org/5.4/Manual/packages.html, 2022. Accessed: 2023-02-13.

[XPB*21] Xiang D., Prada F., Bagautdinov T. M., Xu W., Dong Y., Wen H., Hodgins J., Wu C.: Modeling clothing as a separate layer for an animatable human avatar. *ACM Transactions on Graphics (TOG) 40* (2021), 1–15. https://doi.org/10.1145/3478513.3480545

[XYS*19] Xu Y., Yang S., Sun W., Tan L., Li K., Zhou H.: 3D virtual garment modeling from RGB images. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2019), pp. 37–45. https://doi.org/10.1109/ISMAR.2019.00-28

[YPA*18] Yang S., Pan Z., Amert T., Wang K., Yu L., Berg T. L., Lin M. C.: Physics-inspired garment recovery from a single-view image. *ACM Transactions on Graphics (TOG) 37*, (2018), 1–14. https://doi.org/10.1145/3026479

[ZCF*13] Zhou B., Chen X., Fu Q., Guo K., Tan P.: Garment modeling from a single image. *Computer Graphics Forum 32*, (2013), 85–91.

[ZCJ*20] Zhu H., Cao Y., Jin H., Chen W., Du D., Wang Z., Cui S., Han X.: Deep Fashion3D: A dataset and benchmark for 3D garment reconstruction from single images. In *Computer Vision - ECCV 2020* (Cham, 2020), Springer International Publishing, pp. 512–530. https://doi.org/10.1007/978-3-030-58452-8_30

[ZPBPM17] Zhang C., Pujades S., Black M., Pons-Moll G.: Detailed, accurate, human shape estimation from clothed 3D scan sequences. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Washington, DC, USA, July 2017), IEEE Computer Society, pp. 5484–5493. Spotlight. https://doi.org/10.1109/CVPR.2017.582

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information