

3D Garment Segmentation Based on Semi-supervised Learning Method^{*}

Mian Huang^a, Li Liu^{a,b,*}, Ruomei Wang^b, Xiaodong Fu^a, Lijun Liu^a

^a*Yunnan Provincial Key Laboratory of Computer Technology Application, Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China*

^b*National Engineering Research Center of Digital Life, State-Province Joint Laboratory of Digital Home Interactive Applications, School of Data and Computer Science, Sun Yat-sen University Guangzhou 510006, China*

Abstract

In this paper, we propose a semi-supervised learning method to simultaneous segmentation and labeling of parts in 3D garments. The key idea in this work is to analyze 3D garments using semi-supervised learning method which can label parts in various 3D garments. We first develop an objective function based on Conditional Random Field (CRF) model to learn the prior knowledge of garment components from a set of training examples. Then, we exploit an effective training method that utilizes JointBoost classifiers based on the co-analysis for garments. And we modify the JointBoost to automatically cluster the segmented components without requiring manual parameter tuning. The purpose of our method is to relieve the manual segmentation and labeling of components in 3D garment collections. Finally, the experimental results show the performance of our proposed method is effective.

Keywords: Semi-supervised; Segmentation; Co-analysis; Conditional Random Field; 3D Garments

1 Introduction

With the requirements of various garments in computer graphics modeling and garment customization, how to rapidly create types of 3D garments is a challenging problem. Compared with other complicated 3D models, 3D garments can be regarded as the composition of some prototype components, the changing and blending to some components can directly create a new garment. The prior knowledge learned from contextual component analogies in 3D garments can partition an existing garment model into meaningful components automatically.

^{*}Project supported by the National Nature Science Foundation of China (No. 61462051, No. 61272192, No. 61462056), the Applied Fundamental Research Project of Yunnan Province (No. 2014FB133), the Applied Fundamental Research Key Project of Yunnan Province (No. 2014FA028), the Science Research Funded Project of Kunming University of Science and Technology (No. KKS201403119), and the Open Research Program Key Lab of Computer Technologies Application of Yunnan Province.

^{*}Corresponding author.

Email address: kmust_mary@163.com (Li Liu).

Recently, data-driven supervised segmentation approaches and unsupervised methods are proposed to segment and label components in various 3D models. But supervised approach [1] needs a substantial number of manually labeled training shapes. The performance drops dramatically when the number of training shapes decreases. Without prior knowledge of label information, the unsupervised segmentation [2, 3] results are inferior to the supervised one, especially when the meshes are complicated. Based on these observation, we present a semi-supervised learning method that can generate better garment segmentation results than both labeled and unlabeled methods. And the method overcomes the difficulties of requiring a large amount of labeled meshes and the inability to use unlabeled meshes. Different from these methods which were just designed for man-made objects [4], an important observation is that many garments share prototyping parts in spite of various fashions. Thus, we propose a simple but effective garment segmentation method that utilizes semi-supervised learning method. Fig. 1. illustrates the overview of our method. The contributions of our work are as follows:

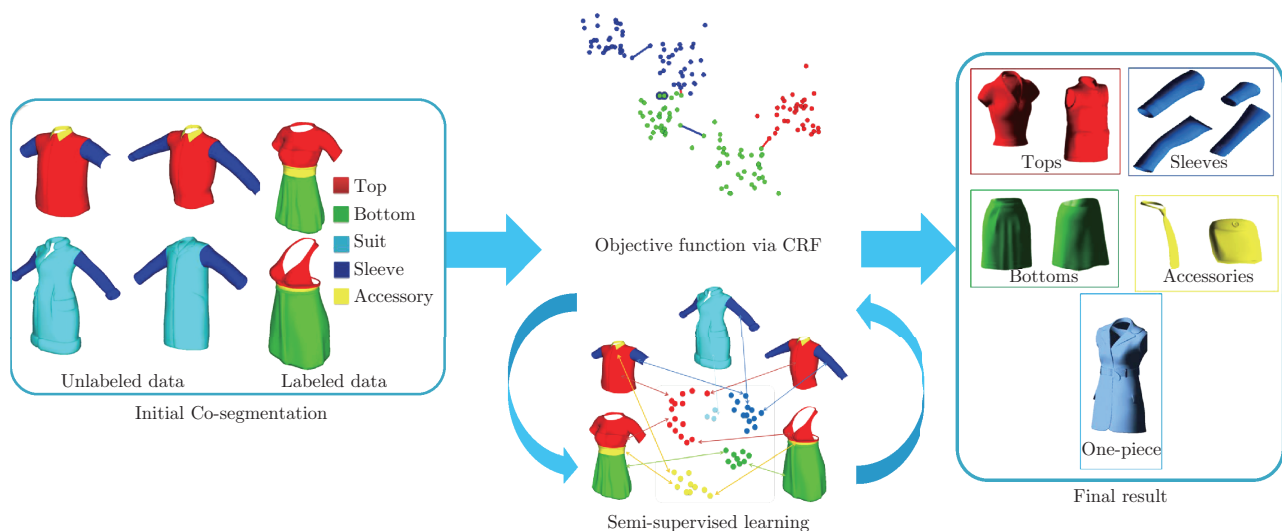


Fig. 1: The pipeline of our semi-supervised learning method consists three main parts: initial co-segmentation, objective function via CRF, and semi-supervised learning

1. Propose a semi-supervised learning method to simultaneous segmentation of components in 3D garments and exploit an effective training method that utilizes JointBoost classifiers.
2. Develop a semi-supervised garment segmentation approach using Conditional Random Field (CRF) model to learn the prior knowledge of garment parts from a set of training examples.
3. Demonstrate the advantages of our 3D garment segmentation method and illustrate the applications for 3D garments, which outperforms robust to mislabeling of training data.

2 Related Work

As an important research topic, shape segmentation has attracted much attention in years, which a comprehensive survey can be referred to [5]. In this section, we briefly review the related work in this area. Mesh segmentation is an optimization problem, including geometric based and semantic based approaches. Earlier approaches focused on finding low-level geometric features

to form meaningful segments, with examples including region merging, hierarchical clustering, iterative clustering, and spectral analysis. It is difficult to get a meaningful segmented results.

In a recent evaluation, many tasks in 3D modeling benefit from automatic segmentation and no segmentation algorithm performed well across all tested datasets [6]. The supervised approaches [1, 7] utilize data-driven techniques to segment a given shape with a substantial number of manually labeled training shapes. Without enough labeled meshes, the unsupervised approaches are proposed such as [2, 3]. However, the results are inferior to the supervised one, especially when the meshes are complicated. To evaluate the effectiveness of these presented approaches, our method is related to the semi-supervised mesh segmentation and labeling developed by Huang et al. [4], which generates better results than both labeled and unlabeled methods. And the method overcomes the difficulties of requiring a large amount of labeled meshes.

One of the advantages of data-driven approach is to utilize prior knowledge from labeled data or a shape collection, with examples including supervised learning method [1, 7], unsupervised co-segmentation with analyzing a set of shapes together [3], semi-supervised segmentation by active learning [8], and projective analysis for 3D shape segmentation [9]. Semi-supervised learning addresses the case where the labeled data is sparse. Wang et al. [8] proposed a semi-supervised learning method to improve the quality of segmentations among a shape collection using a sparse set of user specified constraints, i.e., the label sets. However, none of these approaches consider a shape collection of garments which contains semantic structure except for the recent literature [10]. But our work focuses on semi-supervised learning method for garment segmentation.

3 Objective Function via CRF

The training dataset which incorporates the prior knowledge of clothing attributes contains labeled meshes and unlabeled meshes.

As shown in Fig. 2, we first define each mesh face $s \in F$. Then, let x_s be a vector of features X in each face, which includes descriptors of local surface geometry and context, such as curvatures, shape diameter, and shape context. Let l_s be a random variable over corresponding labeled sequences L , where L is a predefined set of possible garment components, such as “sleeve”, “top”, “accessory”. And Y is the output label by the Conditional Random Field (CRF) model [11]. Assume we have a set of labeled examples $E^\ell = [(x_s^1, l_s^1), \dots, (x_s^N, l_s^N)]$ and unlabeled examples $E^u = [x_s^{N+1}, \dots, x_s^M]$, then all of mesh labels would like to build the following conditional

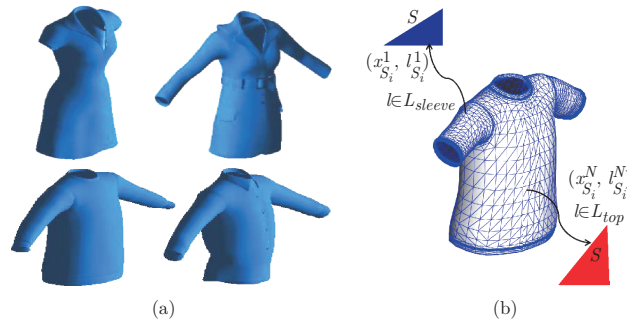


Fig. 2: (a) Input initial garments. (b) One of the triangular mesh s for T-shirt in initial garment models

probability function with CRF model.

$$P_\theta(Y|X, L) = \frac{1}{Z_\theta(X, Y, L)} \exp \left(\Psi_\theta^{E^\ell}(X, L) + \Phi_\theta^{E^u}(X, Y) \right), \quad (1)$$

where $Z_\theta(X, Y, L)$ is a normalization function, and $\exp(\cdot)$ is a statistical distribution which accord with normal distribution. $\Psi(X, L)$ is the penalized log conditional likelihood of the labeled data under the CRF model and is defined as:

$$\Psi_\theta^{E^\ell}(X, L) = \sum_{s=1}^N \log p_\theta(l_s|x_s). \quad (2)$$

The term $\Phi(Y, L)$ is the negative conditional entropy of the CRF on the unlabeled data, and it has the following form:

$$\Phi_\theta^{E^u}(X, Y) = \sum_{s=N+1}^M \sum_Y p_\theta(y_s|x_s) \log p_\theta(y_s|x_s), \quad (3)$$

and $Z_\theta(X, Y, L)$ is a normalization factor. The nature of semi-supervised methods determines that our method can work well even if the labeled meshes are just partially labeled, which may result from missing data or human errors. So this semi-supervised mesh segmentation method can be applied to this kind of meshes and is robust to mislabeling of training data.

For a semi-supervised CRF, the mesh segmentation is achieved by the following maximization of the objective function in the form of entropy regularization:

$$o(\theta) = \sum_{s=1}^n \log p_\theta(l_s|x_s) + \lambda \sum_{s=N+1}^M \sum_Y p_\theta(y_s|x_s) \log p_\theta(y_s|x_s), \quad (4)$$

where λ is a tradeoff parameter that controls the influence of the unlabeled data. The motivation is that minimizing conditional entropy over unlabeled data encourages the algorithm to find putative labels for the unlabeled data that are mutually reinforcing with the supervised labels. To derive an efficient iterative ascent procedure, we need to compute gradient of Eq. (4) with respect to the parameter θ . In our implementation, we use the compatible segmentation technique of Kalogerakis et al. [1], assisted by manual labeling.

4 Semi-supervised Learning

We construct shape set $S := \{S_i | S_i \in M, i = 1, \dots, n\}$, where M denotes the garment library. Each shape S_i represents a triangular mesh. And the training dataset which incorporates the prior knowledge of clothing attributes contains labeled meshes and unlabeled meshes. Our method segments and labels a 3D garment according to the garment shape analysis from the training dataset S . The data-driven approach utilized the prior knowledge to accomplish co-segmentation. Given $S := \{S_i | S_i \in M, i = 1, \dots, n\}$, we compute each segment $G_i = \{m_j^i | m_j^i \subset S_i; \bigcup m_j^i = S_i; m_j^i \cap m_k^i \neq \emptyset, j \neq k\}$ from S_i , where the sub-cluster of G_i is m_j^i . The clustering for each

sub-cluster G_i is defined as $C = \{\Phi_K\}_{K=1,\dots,M}$. For each cluster Φ_K for a garment shape, it has a sub-cluster $m_j^i \subset \Phi_K$. Then, the co-segmentation from garment dataset S is $\Omega = \{S_i\}_{i=1,\dots,N}$.

Learning the parameters of the CRF model is the second procedure of semantic segmentation and clustering. We use JointBoost learning algorithm [12] to train CRF. JointBoost is a boosting algorithm that has many appealing properties: it performs automatic feature selection, each classifier only uses a subset of the provided features. Given the input training pairs (\mathbf{v}_i, c_i) from a 3D garment collection M , where \mathbf{v}_i is a feature vector and c_i is the corresponding class label for the feature. And $\omega_{i,c}$ is assigned weight for each training pair (\mathbf{v}_i, c_i) . We use JointBoost algorithm which proposed by Torralba et al., the weighted multiclass exponential loss over the exemplar set is minimized as follows:

$$J = \sum_{i=1}^M \sum_{l \in C} \omega_{i,c} \exp(-I(c_i, l)H(\mathbf{v}_i, l)), \quad (5)$$

here, C is the set of possible class labels, and $I(c, c')$ is an indicator function. when $c = c'$, this function is 1, otherwise is -1 . And $H(\mathbf{v}, l) = \sum_j h(\mathbf{v}, l; \varphi_j)$, φ_j is the parameters for each decision stumps. For the unary term, the weight $\omega_{i,c}$ is the area of face i . For the pairwise term, $\omega_{i,c}$ is used to adjust the boundary edges. Let N_B and N_{NB} be the number of boundary and non-boundary edges, then $\omega_{i,c} = \ell N_B$ is calculated for non-boundary edges and $\omega_{i,c} = \ell N_{NB}$ for boundary edges, respectively, where ℓ is the corresponding edge length.

The algorithm proceeds iteratively with the weights $\tilde{\omega}_{i,c}$ are initialized to the weights $\omega_{i,c}$. At each iteration, one decision stumps is added to the classifier. The JointBoost decision stump can be written as follows:

$$h(v, l; \varphi) = \begin{cases} a & v_f > \tau \text{ and } l \in C_S, \\ b & v_f \leq \tau \text{ and } l \in C_S, \\ k_l & l \notin C_S. \end{cases} \quad (6)$$

The parameters φ_j of the stump at each iteration j can be computed by the following weighted least-squares optimization:

$$J_{wse}(\varphi_j) = \sum_{l \in C} \sum_{i=1}^M \tilde{\omega}_{i,l} (I(c_i, l) - h(\mathbf{v}_i, l; \varphi_j))^2. \quad (7)$$

According to the method proposed by Torralba [12], the optimal a, b, k_l are computed in closed-form, and a, b, k_l are computed by method of exhaustion. when the parameters of φ_j are determined, the weights can be updated by the following form,

$$\tilde{\omega}_{i,c} \leftarrow \tilde{\omega}_{i,c} \exp(-I(c_i, l)h(\mathbf{v}_i, l; \varphi_j)). \quad (8)$$

Then, we need to define an error function by which to evaluate classification results. And it can be measured by the accurate percentage of the mesh surface area, which refer to as the classification error.

$$E = (\sum_i a_i (I(c_i, c_i^*) + 1)/2) / (\sum_i a_i), \quad (9)$$

where a_i and c_i is the area and label of face i respectively. $c_i^* = \arg \max P(c|\mathbf{x}_i)$ is the output of the classifier for face i . However, when training against this error, the algorithm tends to mostly refine boundaries between larger parts but skip cuts that generate small parts, producing

noticeable errors in the results without incurring much penalty. Instead, we optimize with respect to the Segment-Weighted Error which weighs each segment equally:

$$E_S = \sum_i \frac{a_i}{A_{c_i}} (I(c_i, c_i^*) + 1) / 2, \quad (10)$$

where A_{c_i} is the total area of all faces within the segment that has ground-truth label c_i . These parameters are optimized in two steps. First, the Segment-Weighted Error is minimized over a coarse grid in parameter space by brute-force search. Second, starting from the minimal point in the grid, optimization continues using MATLAB implementation of Preconditioned Conjugate Gradient with numerically-estimated gradients.

Given a set of segmented and labeled components, we further cluster them based on the clothing style by applying the method of Gaussian mixture model and Bayesian network. For each category, the components are partitioned into a set of clusters with similar geometric feature vectors based on geometric style of clothing. An example of this clustering is illustrated in Fig. 3, where sleeves with a obvious shape appearance are grouped into one cluster, while tops are grouped into another. The clustering allows the probabilistic model to operate not only on the semantic labels, but also on the geometric appearance of components.

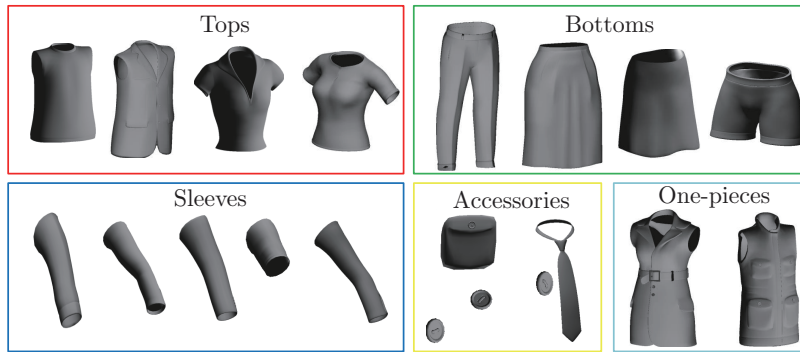


Fig. 3: Garment component library. The garment components are clustered into the garment components library with five categories including tops, bottoms, sleeves, accessories and one-pieces

5 Experimental Results and Discussion

In this section, we evaluate our semi-supervised segmentation approach and present qualitative and quantitative results. We have implemented our proposed garment modeling method and applied it to generate various types of garment models. All the tests were performed on a PC with an Intel Core Duo 3.1 GHz CPU, 4 GB DDR3 Ram, Visual Studio 2010 software.

We evaluate our semantic segmentation method and discuss the feature selection for the analysis of the qualitative and quantitative results. We also make comparisons to both the supervised method of [1] and the semi-supervised approach [4]. The robustness of our method to inconsistently labeled training meshes, such as human mislabeling and label missing, is also illustrated.

We use accuracy to measure the result performance, as [1] and [4] show that different measuring criteria yield consistent results. The accuracy is calculated as the percentage of the mesh surface area that is correctly labeled by $Seg.(A) = \sum_i a_i \frac{(I(k_i, k_i^*) + 1)}{2} / \sum_i a_i$, where a_i is the area of face i ,

Table 1: The comparative results of supervised method [1] and semi-supervised method in our garment segmentation with 6 unlabeled meshes and 16 unlabeled meshes (Unit: %)

Dataset	Accuracy		Robutness		
	Sup. [1]	Semi (6)	Semi (16)	Sup. [1]	Semi (16)
Shirt	92.1	96.6	97.1	84.5	89.6
T-shirt	93.3	94.8	95.6	85.6	90.2
Coat	91.4	92.3	96.2	87.2	91.3
Dress	94.1	98.6	98.8	93.3	96.8
Jumpsuit	92.4	94.3	95.3	86.3	90.4
Long coat	91.6	93.4	95.1	86.7	89.7
Blazer	91.3	94.4	95.3	88.4	94.1
Sweater	90.7	92.5	93.4	89.1	92.7
Average	92.0	94.6	95.9	87.6	91.9

k_i is the ground-truth label for face i , $k_i^* = \arg \max P(k|\mathbf{x}_i)$, $P(k|\mathbf{x}_i)$ is the output of the classifier for face i , $I(k, k^*) = 1$ when $k = k^*$, $I(k, k^*) = -1$ when $k \neq k^*$.

To determine the effect of the number of unlabeled training meshes in our semi-supervised method, we repeated the experiment by using 8 unlabeled meshes and 16 unlabeled meshes. Table 1 shows the accuracy and robustness of our segmentation results compared with the supervised approach of [1]. For the supervised method and ours with the same labeled training sets and features, the average accuracy of our semi-supervised method is 95.2%, which is approximately 3.2% better than the supervised approach. And we also tested on some training dataset with label missing and label error. The experimental results demonstrate that the semi-supervised method with 16 unlabeled meshes show relatively robustness to inconsistently labeled data.

Fig. 4 (a) presents the distribution of the number of diverse components used as test scenario in

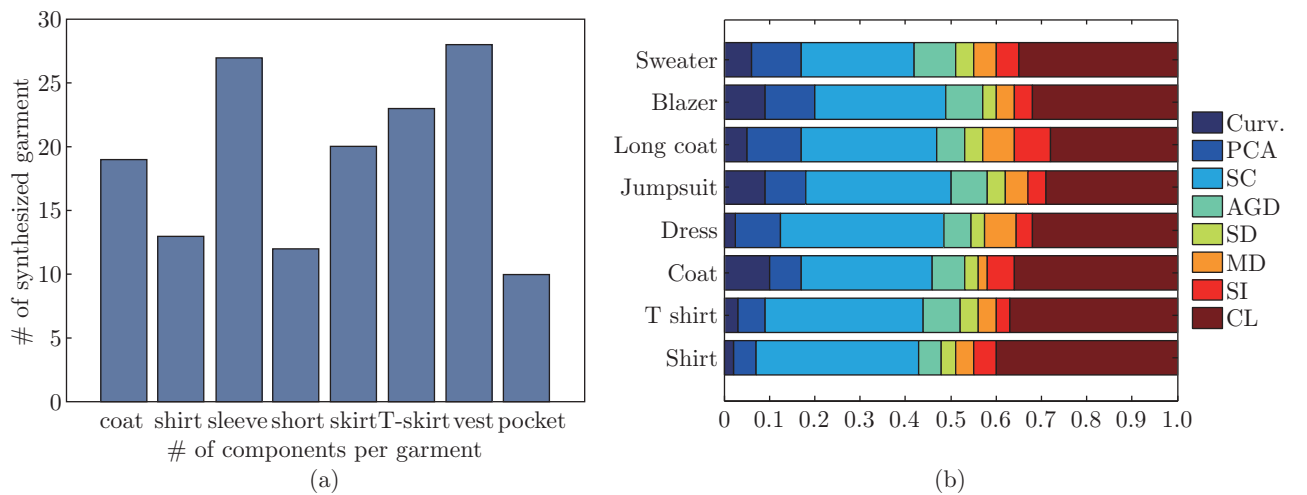


Fig. 4: Quantitative evaluation of experimental data. (a) The distribution of the number of diverse components used as test data in the correspondence analysis. (b) Percentages of features used by JointBoost for different types of garment models

the correspondence analysis. Fig. 4 (b) gives the analysis of feature selection used by JointBoost classifiers in our garment decomposition. The percentage of each feature used in the unary term is visualized for each type of garment models in our training dataset. Curv.=curvature, PCA=PCA singular values, SC=shape contexts, AGD=average geodesic distances, SD=shape diameter, MD=distance from medial surface, SI=Spin Images, CL=contextual label features.

By contrast to the approach [1,4,10], our major differences include the following aspects. Firstly, our method can work well even if the labeled meshes are just partially labeled, which may result from missing data or human errors. Secondly, only dihedral angles which are used for pairwise term can make the boundaries in garment meshes jagged. Thirdly, it is the first approach in mesh segmentation that uses both labeled and unlabeled meshes.

6 Conclusion

Our approach can be applied to many applications, such as digital game, garment prototype modeling, and 3D garment model retrieval [13]. Although the proposed method can successfully generate segmentation results using both labeled and unlabeled meshes, it still has a few limitations that need to be addressed in the future. Firstly, our method only considers the fundamental garment component prototype and is not able to segment the garments with complicated fashions. Secondly, this method is not suitable for professional clothing design or garment examples of high fashion in equation. Finally, the segmentation results are still data-dependent. Besides solving above limitations, we would like to improve our semi-supervised CRF model and yield better results than the supervised one in the future work.

References

- [1] E. Kalogerakis, A. Hertzmann, K. Singh, Learning 3D mesh segmentation and labeling, *ACM Transactions on Graphics (TOG)*, Vol. 29, No. 4, 2010, 102: 1-12
- [2] Q. Huang, V. Koltun, L. Guibas, Joint shape segmentation with linear programming, *ACM Transactions on Graphics (TOG)*, Vol. 30, 2011, 125: 1-12
- [3] O. Sidi, O. van Kaick, Y. Kleiman, H. Zhang, D. Cohen-Or, Unsupervised co-segmentation of a set of shapes via descriptor-space spectral clustering, *ACM Transactions on Graphics (TOG)*, Vol. 30, No. 6, 2011, 126: 1-10
- [4] J. Lv, X. Chen, J. Huang, H. Bao, Semi-supervised mesh segmentation and labeling, *Computer Graphics Forum*, Vol. 31, No. 7, 2012, 2241-2248
- [5] A. Shamir, A survey on mesh segmentation techniques, *Computer Graphics Forum*, Vol. 27, No. 6, 2008, 1539-1556
- [6] X. Chen, A. Golovinskiy, T. Funkhouser, A benchmark for 3D mesh segmentation, *ACM Transactions on Graphics (TOG)*, Vol. 28, No. 3, 2009, 73: 1-12
- [7] O. Van Kaick, A. Tagliasacchi, O. Sidi, H. Zhang, D. Cohen-Or, L. Wolf, G. Hamarneh, Prior knowledge for part correspondence, *Computer Graphics Forum*, Vol. 30, No. 2, 2011, 553-562
- [8] Y. Wang, S. Asafi, O. van Kaick, H. Zhang, D. Cohen-Or, B. Chen, Active co-analysis of a set of shapes, *ACM Transactions on Graphics (TOG)*, Vol. 31, No. 6, 2012, 165: 1-10
- [9] Y. Wang, M. Gong, T. Wang, D. Cohen-Or, H. Zhang, B. Chen, Projective analysis for 3D shape segmentation, *ACM Transactions on Graphics (TOG)*, Vol. 32, No. 6, 2013, 192: 1-12

- [10] L. Liu, R. Wang, F. Zhou, Z. Su, X. Fu, Semantic segmentation and labeling of 3D garments, in 2014 5th International Conference on Digital Home (ICDH), 2014, 299-304
- [11] J. Lafferty, A. McCallum, F. C. Pereira, Conditional random fields: Probabilistic models for segmenting and labeling sequence data, in Proceedings of the 18th International Conference on Machine Learning 2001 (ICML 2001), 2001, 282-289
- [12] A. Torralba, K. Murphy, W. Freeman, Sharing visual features for multiclass and multiview object detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 5, 2007, 854-869
- [13] Y. Oua, R. Wang, F. Zhou, L. Liu, A sketch-based 3D garment model retrieval algorithm, Journal of Fiber Bioengineering and Informatics, Vol. 6, No. 2, 2013, 217-224