Computer-Aided Design 45 (2013) 312–320

Contents lists available at SciVerse ScienceDirect

Computer-Aided Design

journal homepage: www.elsevier.com/locate/cad

Unsupervised co-segmentation for 3D shapes using iterative multi-label optimization

Min Meng [a], Jiazhi Xia [b,c,*], Jun Luo [a], Ying He [a]

[a] School of Computer Engineering, Nanyang Technological University, Singapore

[b] BeingThere Center, Institute for Media Innovation, Nanyang Technological University, Singapore

[c] School of Information Science and Engineering, Central South University, China

ARTICLE INFO

Keywords:
Co-segmentation
Unsupervised algorithm
Shape correspondence
Multi-label optimization

ABSTRACT

This paper presents an unsupervised algorithm for co-segmentation of a set of 3D shapes of the same family. Taking the over-segmentation results as input, our approach clusters the primitive patches to generate an initial guess. Then, it iteratively builds a statistical model to describe each cluster of parts from the previous estimation, and employs the multi-label optimization to improve the co-segmentation results. In contrast to the existing "one-shot" algorithms, our method is superior in that it can improve the co-segmentation results automatically. The experimental results on the Princeton Segmentation Benchmark demonstrate that our approach is able to co-segment 3D shapes with significant variability and achieves comparable performance to the existing supervised algorithms and better performance than the unsupervised ones.

1. Introduction

The past few years has witnessed an increasing research trend towards the high-level analysis of 3D shapes to derive structural and semantic shape information. This research heavily relies on low-level geometric properties, in particular, the semantic knowledge based shape segmentation [1,2]. Recently, researchers observed that segmenting a set of 3D shapes as a whole into consistent parts can infer more knowledge than from an individual shape, which leads to a new research problem of co-segmentation. The existing co-segmentation algorithms can be classified into two categories, i.e., supervised and unsupervised. Taking advantage of the training sets, the supervised ones [2,3] are able to generate consistent results, however, the training process usually requires a large amount of manual labeling on the training models, which is tedious and time consuming. The unsupervised algorithms [4–7], on the other hand, are flexible and effective in segmenting shapes in a heterogenous shape database, however, these approaches either depend on accurate alignment of input shapes or cannot guarantee the consistency of the final consistency of the final co-segmentations across the whole set.

In this paper, we propose a new framework for unsupervised co-segmentation of a set of 3D shapes. Our algorithm first segments all input shapes into the primitive patches, which are then clustered into corresponding parts to generate an initial estimation of co-segmentation. Next, the initial guess is automatically improved by an iterative optimization scheme, which generates a statistical model for the estimated clusters of all shapes and performs multi-labeling optimization for individual shapes respectively.

It is well known that shape descriptors play an important role in analyzing the surface geometry and a classical application of surface descriptors is shape segmentation [8]. While two perceived corresponding parts of models may differ in some features significantly. Our method adopts multiple shape descriptors such that they are rich enough to classify the faces belonging to different clusters. Our framework distinguishes the intrinsic and extrinsic shape descriptors in the two phases. Specifically, in phase one, it focuses on the intrinsic shape descriptors, e.g., average geodesic distance (AGD) [9,10], to generate a rough estimation of co-segmentation by grouping the primitive patches. Then the extrinsic descriptors, e.g., shape diameter function (SDF) [11], play a more important role in the subsequent optimization procedure aiming to classify the faces into their corresponding labels.

The novelty of our approach is twofold. First, the iterative optimization is able to improve the initial rough co-segmentation results automatically and allows us to handle a large shape library without prior correspondences or alignments. Second, our method distinguishes the intrinsic and extrinsic shape descriptors at different levels of 3D shape co-analysis. In particular, our method relies more on the intrinsic descriptors in the initial grouping

* Corresponding author at: BeingThere Center, Institute for Media Innovation, Nanyang Technological University, Singapore.

E-mail addresses: mengmin@ntu.edu.sg (M. Meng), jzxia@ntu.edu.sg (J. Xia), junluo@ntu.edu.sg (J. Luo), yhe@ntu.edu.sg (Y. He).

0010-4485/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cad.2012.10.014

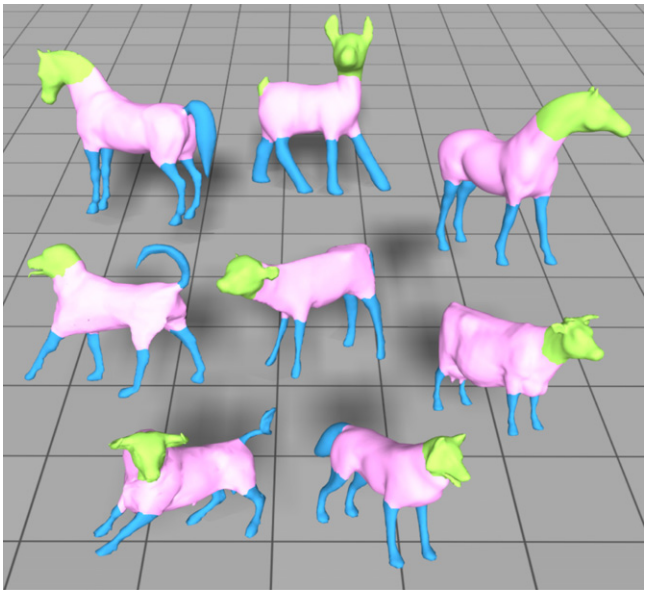


Fig. 1. Unsupervised co-segmentation results of our approach. Corresponding segments are shown with the same color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of primitive patches, while it utilizes more extrinsic ones in the subsequent classifying of faces at the semantic level. This feature allows us to consistently segment a set of shapes with significant diversity.

We evaluate the proposed approach on the Princeton Segmentation Benchmark and make comparisons with state-of-the-art techniques. Fig. 1 shows our co-segmentation results on the class of four-leg animals. The experimental results demonstrate that our approach is able to achieve comparable performance to the supervised approach and produces better results than the unsupervised ones.

The remainder of the paper is organized as follows. Section 2 reviews the related work in shape segmentation and shape descriptors. Section 3 presents the overview of our co-segmentation framework followed by the detailed description of the initial co-segmentation in Section 4 and the iterative optimization in Section 5. Then, we document the experimental results and make a comparison with state-of-the-art techniques in Section 6. Finally, Section 7 concludes this paper.

2. Related work

2.1. Shape segmentation

Shape segmentation plays an important role in high-level shape analysis. Despite much effort devoted to segment single shapes into meaningful parts [8], a recent evaluation shows that no segmentation algorithm performs well for all models because individual shapes may not provide sufficient geometric cues to distinguish all parts that would be perceived as meaningful to a human observer [12].

To tackle the challenge in individual shape segmentation, consistently segmenting a set of shapes from the same family into semantic parts, or co-segmentation, has received increased attention [13,4,14,7]. Golovinskiy and Funkhouser formulate a graph clustering problem [4], which takes a global rigid alignment between matching parts and then finds the connection between corresponding parts using the iterative closest point. Due to the lack of shape semantics, their approach works only for limited model types. To handle the non-homogenous part stretching of shapes, Xu et al. [7] firstly group the shapes based on their styles and then perform part correspondences in each style group. However their approach is computationally expensive within the group generation process.

There are also data-driven techniques that utilize information from multiple shapes. Kalogerakis et al. [2] propose a supervised approach to segmentation, demonstrating significant improvement over single shape segmentation algorithms. van Kaick et al. [3] incorporate prior knowledge imparted by a set of pre-segmented and labeled models with content-driven analysis for performing part correspondence. Although consistent segmentation may be established based on the knowledge observed from multiple examples from the same family, these supervised approaches still require a large number of manually segmented training shapes to learn from.

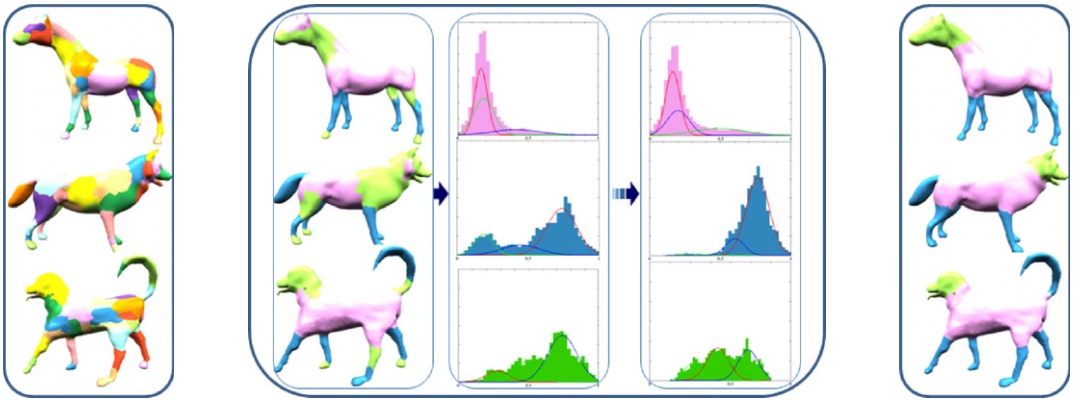
Very recently, two promising unsupervised techniques [5,6] have been proposed for co-segmentation. Aided by the co-analysis in a descriptor space, Sidi et al.'s algorithm [5] can co-segment a set of shapes with large variability, revealing the semantic shape parts and establishing their correspondences. However, due to the initial segmentations required for co-analysis, Sidi et al.'s algorithm may fail if the single-shape segmentation is poor. Huang et al. [6] present a novel linear programming approach to jointly segment the shapes in a heterogeneous shape library, producing comparable results to the supervised approaches on the Princeton Segmentation Benchmark. They observed that the segmentations produced in the multi-way joint condition are generally better than those produced in the pairwise joint condition, as the variability can be further exploited from the database in the former case. This variation-dependent feature, however, may reduce the algorithm to single-shape segmentation if the input models are lacking of variability. Moreover, this approach produces only mutually consistent segmentation but cannot guarantee the consistent segmentations across a shape class.

In a concurrent effort, Hu et al. [15] present an unsupervised co-segmentation algorithm that generates the segmentations by grouping the primitive patches of the shapes directly and obtains their correspondences simultaneously. However, unlike our approach, this technique does not guarantee that all the underlying segmentation can be captured well, since the final segmentations of each shape are generated from the initially computed patches.

2.2. Shape descriptor

Shape descriptors, characterizing the geometric features of input shapes, are widely used in digital geometry processing. Gatzke et al. [16] construct the curvature map signature for shape matching relying on geodesic distance. Lai et al. [17] propose a feature-sensitive metric by combing geodesic and isophotic distances for sampling, remeshing and multi-scale feature selection. Other metrics, such as shape distribution [18] and average geodesic field [19], have also been used to derive global shape properties in shape retrieval.

Shape segmentation has to face the problem of defining a part or part boundary. As one of the best known rules, the minimal rule [20] inducing part boundaries along negative curvature minima is typically realized via curvature measurements. On the other hand, volumetric considerations for part analysis appear to be more closely linked to skeletal shape representations, e.g. shape diameter function [11], and volumetric shape image [21]. Specific to shape segmentation [22–24], it is critical to find the proper metric which can capture the essence of the semantic components. Kalogerakis et al.'s algorithm automatically selects the most informative features from a given set of shape descriptors to train the JointBoost classifier [2]. Inspired by their work, our algorithm also adopts multiple shape descriptors to distinguish the faces at different analysis levels.



(a) Over-segmentation. (b) Iterative optimization. (c) Co-segmentation results.

Fig. 2. Overview of our approach. (a) The over-segmentation is computed for each shape. (b) In place of one-shot labeling, we iteratively improve the segmentation of shapes using multi-label optimization from the initial guess of co-segmentation. (c) The final co-segmentation of the set is obtained until convergence.

3. Overview

Our co-segmentation approach takes as input a set of 3D meshes from the same class and provides a consistent segmentation of these meshes. Our approach contains three stages, as illustrated in Fig. 2. In the first stage, we compute a set of primitive patches for each input mesh independently. In the second stage, we perform a clustering in the common space of all patches and obtain an initial estimate of co-segmentation. Finally, in the third stage, we iteratively build the statistical model to describe each part cluster from previous estimation, and adopt the multi-label optimization to improve the co-segmentation quality.

Over-segmentation. Inspired by the computation of superpixels for image segmentation [25], we first decompose each shape into primitive patches to generate an over-segmentation [6]. To this end, we employ the normalized cuts [26,27] to compute the primitive patches for each shape, and further align the patch boundaries with shape features by fuzzy cuts [28].

Initial co-segmentation. To generate the initial estimate of co-segmentation of 3D shapes, we perform a clustering in a common space of primitive patches for all shapes. To this end, we construct an affinity matrix based on the similarities between pairs of primitive patches, which can be computed by the distances between shape descriptors. By carrying out the normalized cuts [26], we obtain an initial co-segmentation in which a single cluster potentially represents a certain class of semantic parts. As this stage was performed at the primitive patch level, we emphasize the shape descriptors with intrinsic properties, e.g. average geodesic distance (AGD), and conformal factor (CF) [29], to measure the similarities of primitive patches.

Iterative multi-label optimization. Given the initial guess of co-segmentation of shapes, we build a statistical model to describe each cluster of parts, and employ the multi-label optimization scheme to produce the improved segmentation of input meshes. With the new estimate of co-segmentation, we can iteratively improve the segmentation of all shapes in the set. In place of a one-shot algorithm, our iterative scheme has the advantage of allowing automatic refinement of the co-segmentation. Specifically, the statistical model can be driven in terms of Gaussian mixture models (GMM) which are learned from the clusters via the shape descriptors. In contrast to the initial stage, we put more emphasis on the extrinsic shape descriptors, e.g. shape diameter function (SDF), to define the properties of each face for the labeling at a semantic level. As newly labeled faces for each cluster are provided, we can refine the GMM parameters to further improve the co-segmentation of the shapes in the set. We terminate the algorithm when the optimization converges or the accuracy of the co-segmentation result is good enough (e.g. $\geq$0.90).

4. Initial co-segmentation

Starting with the primitive patches computed individually for each shape, we compute an initial estimate of co-segmentation by clustering them into the potential classes of parts. As a large variability may exist within the shape library, a simple clustering will not group the primitive patches into proper classes. Rather than focusing on local features, we aim at extracting the global similarities of the patches. To further explore the similarities of patches, we first compute the pairwise affinities between patches by a set of shape descriptors, then apply the Normalized Cuts [26] method to produce a more accurate grouping of patches.

Descriptor space. Aiming to classify the primitive patches, we prefer to collect a set of shape descriptors exploiting geometric features. Observing the study on feature selection in [2], we chose four shape descriptors in this paper, average geodesic distance (AGD), conformal factor (CF), shape diameter function (SDF) and shape contexts (SC) [30]. All shape descriptors are defined and computed on the normalized shapes for overall scale. Hence for each face, we define a feature vector obtained from all $D(D = 4)$ shape descriptors,

$$S(f) = (\ldots, w_i d_i(f), \ldots) \quad i = 1, \ldots, D. \tag{1}$$

By adjusting the weight value w for the corresponding descriptor, we can perform co-analysis of the shapes for different purposes.

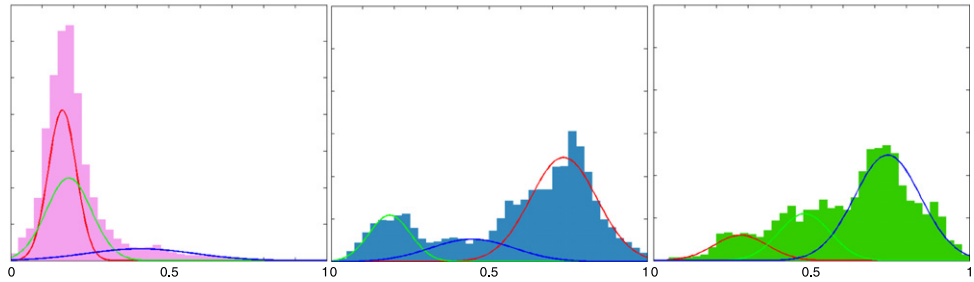
In order to perform clustering of primitive patches, we need to construct the affinity matrix of the patches reflecting their pair-wise similarities. Considering the variability of the shape library, in this stage we prefer to emphasize the intrinsic shape descriptors, e.g. AGD, and CF, by enlarging the weights w of those descriptors. Based on the weighted face-level descriptors, we formulate the patch-level descriptors as the histograms that capture the distribution of each descriptor for all faces within the primitive patch. For each face-level descriptor d, we compute the histogram as the patch-level descriptor for patch p_i, denoted as h_i^d. As a result, every patch is associated with a set of histograms. Thence, we define the distance between two patches p_i and p_j as follows:

$$d(p_i, p_j) = \sqrt{\sum_d EMD^2(h_i^d, h_j^d)} \tag{2}$$

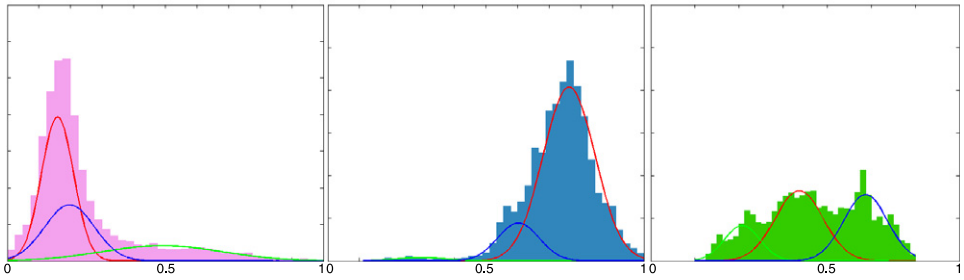
where $EMD()$ is the earthmover's distance, commonly used to measure the dissimilarity between two probability distributions.

With the distance measure between patches, we apply a Gaussian kernel to the distance aiming to construct the affinity matrix A, with elements

$$A_{ij} = \exp(-d(p_i, p_j)/2\sigma^2). \tag{3}$$



(a) Initial GMM.



(b) After optimization.

Fig. 3. The GMM for each cluster before and after the optimization. Clearly, the ranges of GMMs are narrowed down after optimization, as the face labeling has become more accurate.

Clustering. After obtaining the affinity matrix, we then apply the Normalized Cuts [26] technique based on a generalized eigenvalue problem to group the patches into the potential clusters of parts that exist in the shape set. The user could specify the number of clusters, approximately corresponding to the number of semantic parts that constitute the shapes.

The clustering results serve as the initial guess of the co-segmentation of shapes, and can be improved iteratively in the next stage. Various initial segmentations have been tried in our framework. For the initial segmentations, it is more important to be meaningful than to be precise. Thus distinct from the methods which proceed on the facet level, we perform the clustering directly from the patches, allowing us to handle more kinds of categories.

5. Iterative multi-label optimization

This section describes the iterative scheme for co-segmentation: iterative estimation and multi-label optimization.

Cluster modeling. For each cluster of parts, we can construct a statistical model, e.g. Gaussian mixture model (GMM) in our approach, based on the shape descriptors. In contrast to the initial stage, we focus more on the extrinsic descriptors, e.g. SDF, and SC, to define the properties of each face for the labeling at a semantic level. For each cluster, the feature vector obtained from all shape descriptors for all faces in the cluster are collected to estimate the multi-dimensional Gaussian mixture model. Each GMM is taken to be a full-covariance Gaussian mixture with K components (typically $K = 3$ in our approach). Therefore, the statistical model for the cluster c_i is as follows,

$$p(f|c_i) = \sum_{k=1}^{K} w_k g(f|\mu_k, \Sigma_k), \tag{4}$$

where f is a D-dimensional vector computed by the shape descriptors, and w_k and $g(f|\mu_k, \Sigma_k)$ are the mixture weight and the Gaussian density for the k-th component. The mixture weights satisfy the constraint that $\sum_{k=1}^{K} w_k = 1$. All GMM parameters $\lambda = w_k, \mu_k, \Sigma_k, k = 1, \ldots, K$ for each cluster are estimated using the iterative expectation-maximization algorithm.

Co-segmentation. Once the statistical model for all clusters is constructed, we treat the co-segmentation as a multi-labeling optimization, solved for each shape in the set individually. Consider the dual graph of mesh $G = \{\mathcal{V}, \mathcal{E}\}$ with $\mathcal{V}$ denoting the set of faces and $\mathcal{E}$ denoting the set of edges between neighboring faces. We formulate the segmentation for each shape as a face-labeling problem that minimizes the energy,

$$E(L) = \sum_{v \in \mathcal{V}} E_d(l_v) + \omega \sum_{(v,u) \in \mathcal{E}} E_s(l_v, l_u), \tag{5}$$

where the data term $E_d(l_v)$ describes the penalty of assigning a label l_v to a face, the smoothness term E_s depicts the penalty for assigning different labels to two adjacent faces, and ω is the weight to regulate the influence of the smoothness term in the total energy.

In our setting, we define the data term as

$$E_d(l_v) = -ln(p(c_{l_v}|v) + \varepsilon), \tag{6}$$

where $p(c_{l_v}|v)$ is the probability measured by the statistical model of cluster c_{l_v}, and ε ($\varepsilon = 1e-6$) is a small threshold to avoid zero value in the logarithm function. Similar to [5,14], we define the smoothness term as follows,

$$E_s(l_v, l_u) = \begin{cases} 0 & \text{if } l_u = l_v \\ -ln(\theta_{vu}/\pi) \cdot e_{vu} & \text{otherwise} \end{cases} \tag{7}$$

where θ_{vu} denotes the dihedral angle between the faces v and u, and e_{vu} denotes the length of the edge between the two faces.

Finally, we employ the multi-label optimization technique [31] to minimize the energy $E(L)$. Therefore, the co-segmentation of the set is composed of the labeling results for each shape.

Iterative optimization. Rather than one-shot labeling, the energy minimization scheme proceeds iteratively in our approach. This has the advantage of allowing automatic refinement of the co-segmentation, as the updated clusters are used to fit the GMM parameters λ. The iterative optimization scheme is given as follows:

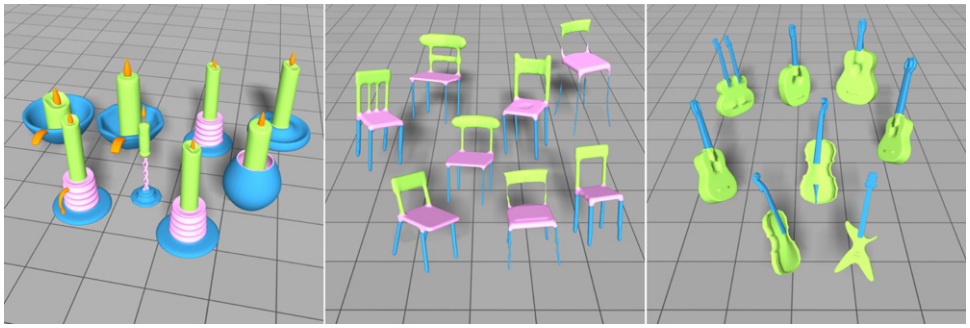


Fig. 4. Co-segmentation results on the categories: Candle, Chair and Guitar. Corresponding segments in each class are shown with the same color. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

1. Learn GMM parameters λ from the shape descriptors for all faces in each cluster.
2. Compute the probability p that an unknown face f belongs to each cluster c.
3. Estimate the improved co-segmentation of the shapes using multi-label optimization algorithm.
4. Repeat from step 1 until convergence.

In general, the above procedure converges within a few iterations. In our implementation, we terminate the algorithm when the accuracy is above 0.9 or the number of iterations is 5.

Fig. 3 illustrates that, for the four-leg model (shown in Fig. 1) the iterative optimization in our approach can increase the classification of the GMMs, relative to the one-shot labeling. It is obvious that the GMMs are better separated after iteration.

6. Results and discussions

In this section we describe experimental results and demonstrate the performance of our unsupervised co-segmentation approach.

Data set. To evaluate the performance of our proposed co-segmentation algorithm, we have collected an experimental dataset with 23 different object categories, of which 4 categories (candelabra, goblets, guitars and lamps) are from [5] and the other 19 categories are from the Princeton Segmentation Benchmark (PSB) [12]. The ground truth of the categories selected from [5] is provided by the authors. For the categories selected from the PSB, we use the ground truth based on the labeled database [2]. Specifically for the human category from the PSB, only 15 models are chosen in our experiments since the other human models are not suitable for co-segmentation owing to inseparable parts from the bodies.

Furthermore the ground truth associated with the experimental dataset is also employed for evaluation. To assist the comparison between the co-segmentation results and ground truth, we have manually merged the labels for the categories with too detailed labels, e.g. airplane, armadillo and hand categories, according to a specific labeling scheme.

Experimental results. The proposed approach is entirely unsupervised. As mentioned in Section 4, we combine the intrinsic and extrinsic descriptors into a blended metric which can perform co-analysis of the shapes at different levels by adjusting the weight value w for specific descriptors. Specifically, in the initial stage we focus more on the intrinsic descriptors by setting the weight vector w to (0.7, 0.7, 0.3, 0.3) for the shape descriptors AGD, CF, SDF, and SC respectively to suppress the outliers in the initial estimation. The weights are then changed to (0.3, 0.3, 0.7, 0.7) in the subsequent optimization procedure with the aim to enhance the ability to classify the faces into their corresponding labels. Without requiring any manual parameter tuning, the fixed parameters work well for all experimental results in this paper.

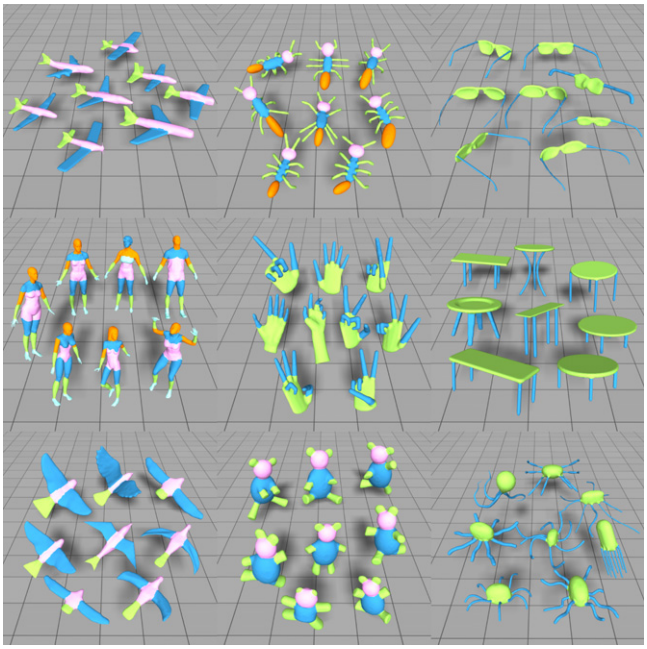


Fig. 5. Co-segmentation results on the representative categories from the Princeton Segmentation Benchmark [12]. (Top: Airplane, Ant, Glass; Middle: Human, Hand, Table; Bottom: Bird, Teddy, Octopus.)

Fig. 4 shows the co-segmentation results for the shapes of categories selected from [5]. The resulting co-segmentation for the candelabra category demonstrates that outlier segments are allowed in our approach, as the detection of holders only on those candles that contain them in the category. For the chair models, our method is able to identify their main consistent parts with different topologies. Noting the similarities of geometric features between the legs and middle parts of the backs, our approach can still extract the common parts and yield a coherent labeling owing to the power of our iterative optimization procedure.

We have also tested our proposed approach on the Princeton Segmentation Benchmark. Fig. 5 shows the co-segmentation results for some presentative categories. Our method successfully segmented these models into corresponding parts, demonstrating its insensitivity to poses and shape variations, especially by the co-segmentation results of the human, hand, and ant categories. From these experimental results, we can see that our algorithm works well for a large variety of shape sets, including organic models, articulated models and man-made models as shown in Figs. 1, 4 and 5.

Like previous work, our algorithm tends to produce better results when the input data set contains more models, but does

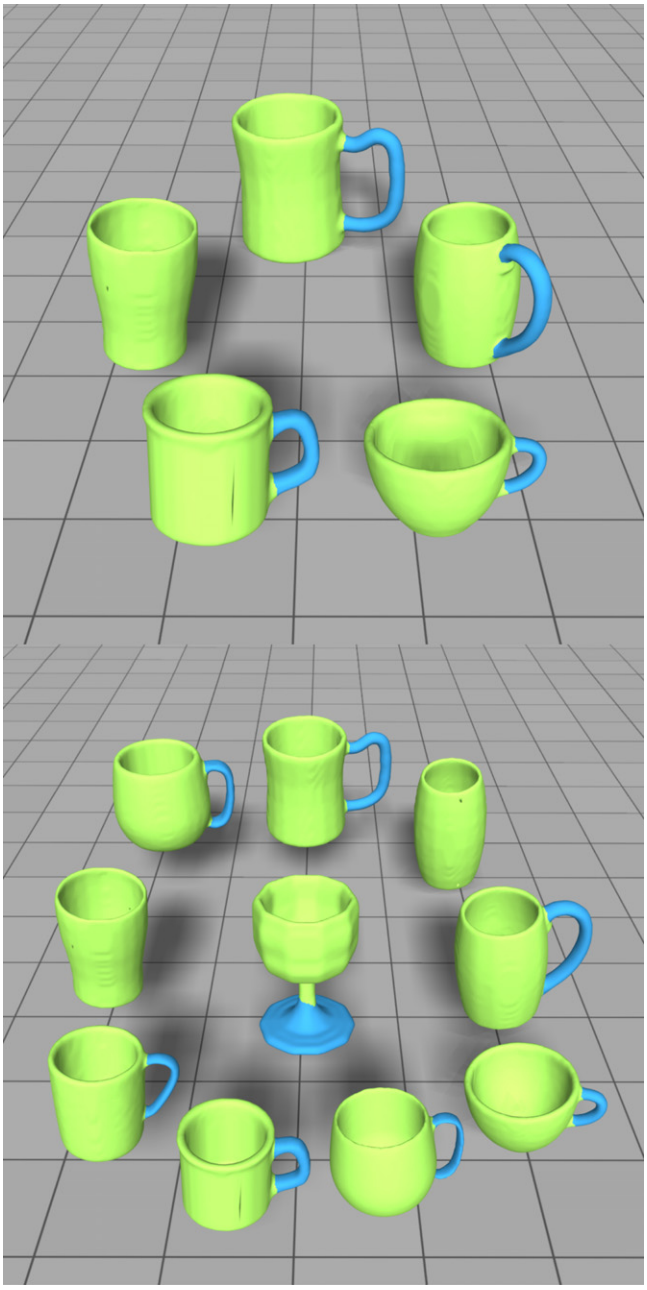


Fig. 6. Consistent segmentation results can be obtained by applying our approach on different subsets of Cup models, e.g. 5 and 10 models for the top set and bottom set respectively.

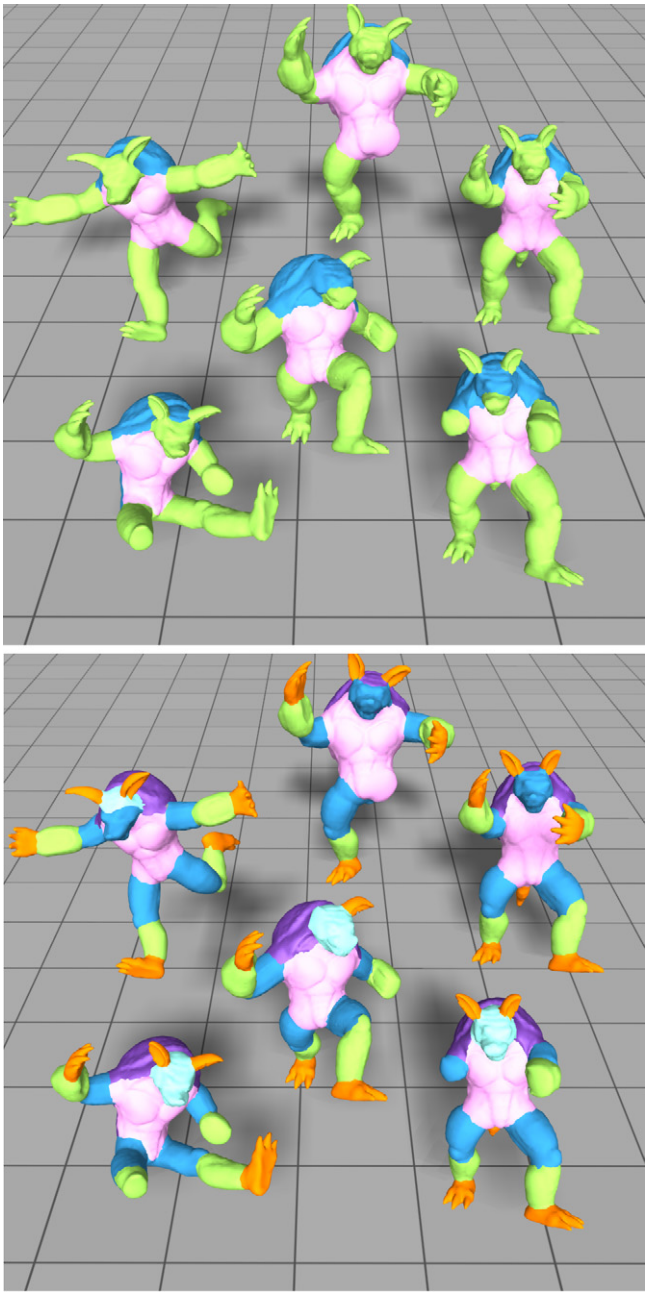


Fig. 7. Multi-level co-segmentation of the Armadillos.

not require a large number of models in the input set. Fig. 6 shows the co-segmentation results of two subsets of the Cup category. We can see that our algorithm still works well for these subsets with few models and can generate consistent segmentation results for different subsets on the common category.

Thanks to the combined shape descriptors and iterative optimization scheme, our algorithm exhibits a substantial degree of flexibility when applying to a set of shapes with significant variability. As shown in Fig. 7, our algorithm is able to co-segment the Armadillos at two levels, producing the corresponding parts according to the desired cluster number, which is specified by the user.

To further access the performance of our algorithm in a quantitative manner, we need to define an error function which measures the quality of the experimental results. Similar to [2,5], we adopt the accuracy metric to measure the amount of area of the given shape that is labeled correctly,

$$Accuracy(l, t) = \left(\sum_i a_i \delta(l_i - t_i)\right) \Big/ \left(\sum_i a_i\right) \tag{8}$$

where a_i is the area of face i, l and t are the labeling computed by the co-segmentation and corresponding ground truth respectively, and the Dirac delta function $\delta(x)$ is 1 only if $x = 0$. We then average the labeling accuracies over all shapes from the same category. The statistical evaluation of our co-segmentation results for the categories in the dataset is shown in Table 1. Our algorithm has obtained an average accuracy of 91.6% over all categories in our dataset. We notice that the accuracies are at least 83.2% or higher except the Bust category for which it is difficult to extract meaningful parts even for the data-driven techniques.

Fig. 8 shows the accuracy improvement by using our iterative optimization. Starting from the rough initial guess, our method

Table 1
Average accuracies of our co-segmentation results.

Category	Accuracy	Category	Accuracy
Human	83.2	Bird	86.1
Cup	99.2	Armadillo	90.8
Glasses	93.6	Bust	63.6
Airplane	92.7	Mech	97.6
Ant	96.3	Bearing	89.8
Chair	90.4	Vase	86.1
Octopus	99.1	Four-leg	90.7
Table	97.2	Candle	94.7
Teddy	99.2	Guitar	98.6
Hand	92.7	Goblet	99.2
Plier	89.5	Lamp	88.3
Fish	87.2	**Average**	**91.6**

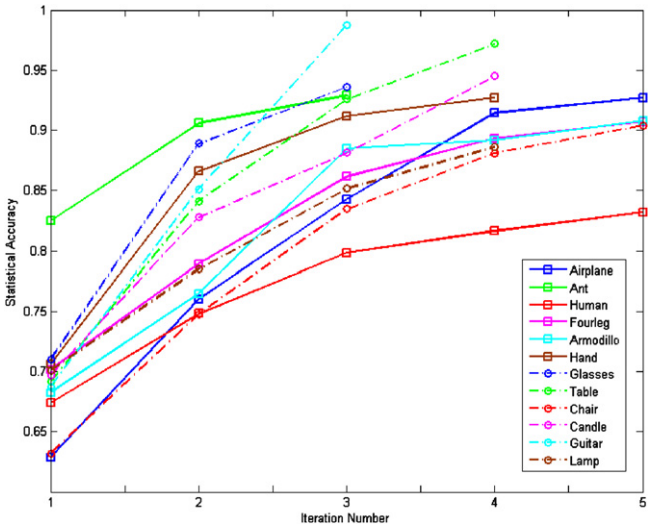


Fig. 8. Our iterative optimization scheme is very effective to improve the accuracy. The horizontal axis shows the iteration number. The accuracy of the co-segmentation results declare a visible trend of improvement for the co-segmentation.

gradually improves the accuracy of the co-segmentation for the categories shown in the paper. The general trend of the accuracy curves clearly demonstrates that the co-segmentation accuracy is improved as more iterations are performed. In place of one-shot labeling, the iterative scheme can automatically refine the co-segmentation, providing more accurate labeling for all shapes.

Comparison. We have compared our approach to the data-driven approach [2] on the Princeton Segmentation Benchmark. The comparison is carried out between our co-segmentation results and their leave-one-out-error experimental results, the most accurate correspondence results shared by the author. The comparison is shown in Fig. 9(bottom). The average accuracies for [2] on the whole benchmark are 93.8%, which is better than ours, 90.8%. However, the supervised approach requires substantial work for preparing the training set, which is tedious and time-consuming. It is also worth noting that their results are guided by the training data where the knowledge can be automatically inferred by our unsupervised method.

In addition to the supervised algorithm, we have compared the presented technique to the unsupervised algorithms of Sidi et al. [5] and Huang et al. [6]. The comparison with the unsupervised algorithm [5] was performed on six categories of shapes provided by the authors of [5]. The average accuracies for [5] and ours are 88.2% and 92.6% respectively. As shown in Fig. 9(top), our results get higher accuracy than theirs for most cases, except for the Lamp category as it contains corresponding parts which are very difficult to classify with purely geometry

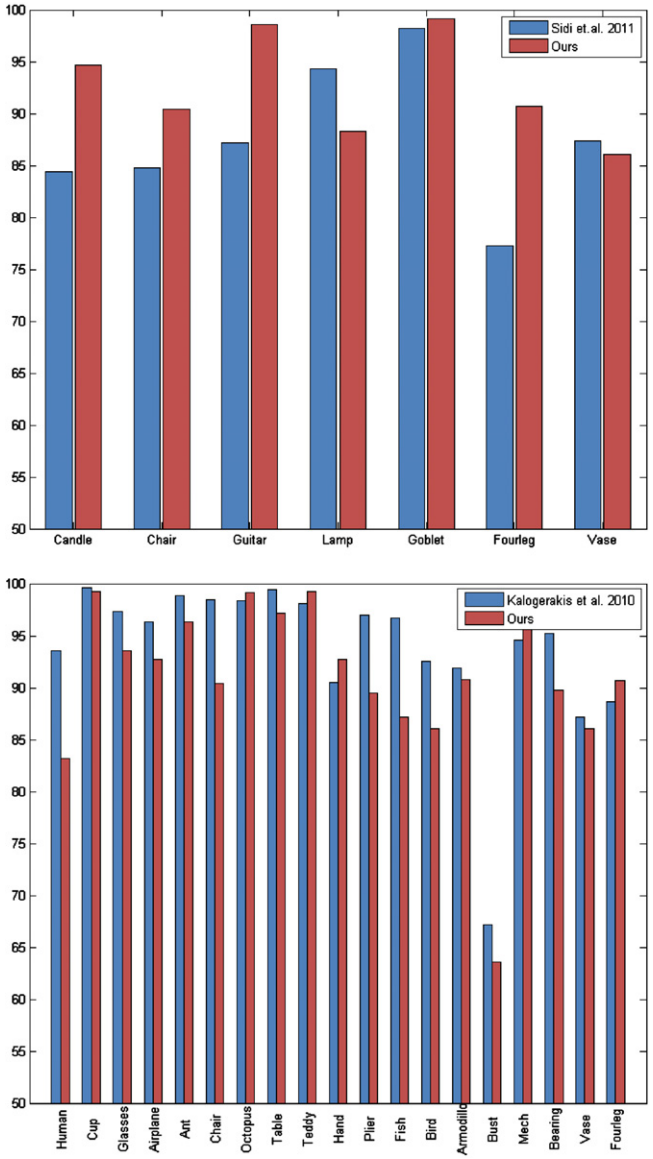


Fig. 9. Comparison with the state-of-the-art techniques [5] and [2] using accuracy measure. Higher values indicate closer similarity to human-generated ground truth.

features. Furthermore, it is worth noting that their approach relies on good initial per-object segmentations which are difficult to obtain for the organic models, e.g. four-leg models. Therefore, poor initial segmentation may result in unsatisfactory co-segmentation in their method. With the proposed iterative optimization scheme, our method can gradually improve the co-segmentation accuracy, which significantly reduces the dependence on the initial guess.

We also compared to the unsupervised algorithm [6] on the Princeton Segmentation Benchmark. By means of the experimental results provided in the paper [6], we use the Rand index measure [12] to carry out the comparison between our co-segmentation results and their segmentation results produced in the JointAll condition, the generally better results provided in their paper. The average Rand index scores for [6] and ours are 10.1% and 10.8% respectively, lower values indicate better closer similarity to human-generated ground truth. As shown in Fig. 10, the overall performance of these two approaches is similar, although their behavior on individual categories is quite different. The algorithm [6] performs generally better than our approach in the categories with large variation, e.g. the Human category. On the other hand, our technique outperforms their algorithm

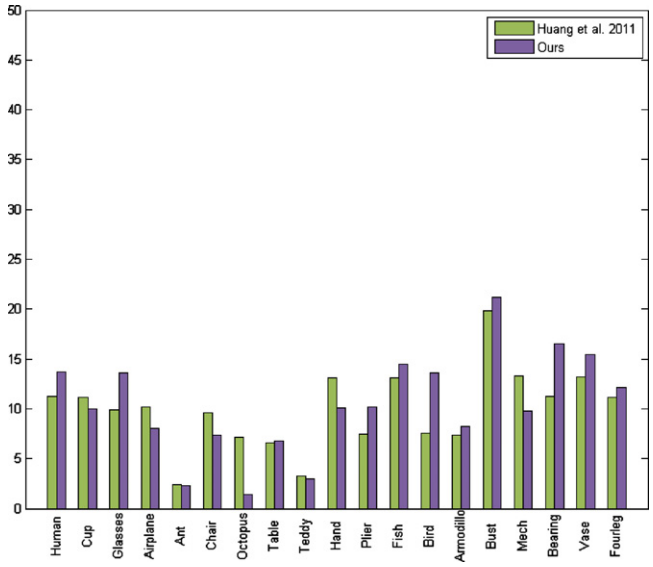


Fig. 10. Comparison with the state-of-the-art technique [6] using Rand index scores. Lower values indicate closer similarity to human-generated ground truth.

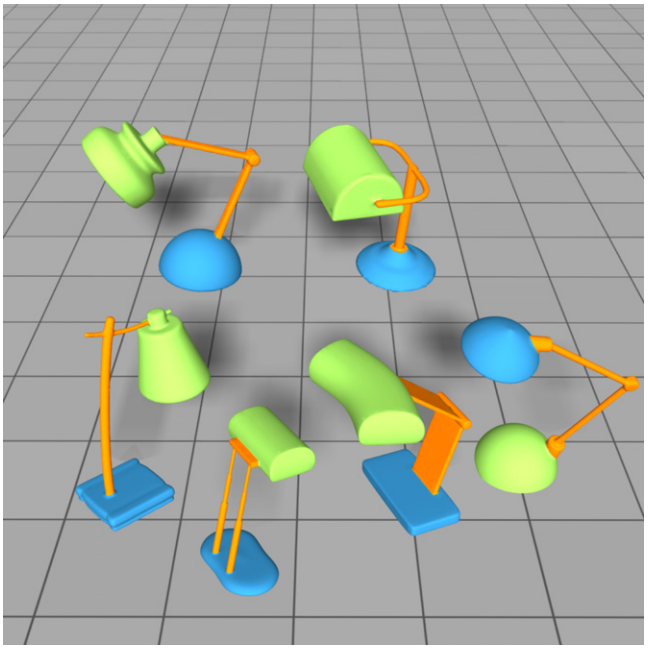


Fig. 11. Our approach groups the upper part of one lamp (lower right) into the lower part, as they are very similar in geometry.

when the variation in the category is small, e.g. the Airplane category which contains many highly similar shapes. However, this variation-dependent feature may reduce their algorithm to single-shape segmentation when the input category lacks variability. Moreover, their algorithm cannot guarantee the consistency of the final segmentations across a shape class while our approach can generate consistent co-segmentations by classifying the similar parts of all the shapes into the same cluster.

Performance. All experiments were conducted on a workstation with an Intel(R) Dual-core 2.67 GHz CPU and 12 GB RAM. Our algorithm takes around 8 min for a category with 20 medium-sized shapes. It is worth noting that our algorithm performs most of the analysis at the patch or part level rather than face level and it does not require any data-training or pre-alignment of shapes. Thus, our method is more efficient than [2,5].

Limitation. Purely driven by the geometric features, our algorithm does not employ the pre-alignment of the shapes or any prior knowledge. Thus, it may fail for some ambiguous cases where the corresponding parts are very difficult to classify based on purely geometry features, such as the Lamp model shown in Fig. 11.

7. Conclusion

In this paper we have developed an approach for co-segmenting a set of shapes from a common family. By performing clustering in the space of primitive patches and exploiting the power of iterative multi-label optimization, we are able to consistently co-segment the shapes from the set exhibiting significant variability. We have evaluated our approach on the Princeton Segmentation Benchmark [12] which consists of multiple shape categories, and compared to the state-of-the-art techniques of co-segmentation. Experimental results have demonstrated that the proposed algorithm can properly extract consistent parts across the model set, achieving comparable performance to the data-driven method.

Acknowledgments

We thank the anonymous reviewers for their valuable comments and suggestions. Thanks also go to Prof. Hao Zhang for providing their co-segmentation results. This project is supported by NRF2008IDM-IDM-004-006, AcRF 32/09 and NRF IDM BeingThere Project.

References

[1] Simari PD, Nowrouzezahrai D, Kalogerakis E, Singh K. Multi-objective shape segmentation and labeling. Computer Graphics Forum (Proc. SGP) 2009;28:1415–25.
[2] Kalogerakis E, Hertzmann A, Singh K. Learning 3d mesh segmentation and labeling. ACM Transactions on Graphics 2010;29(3):1–11.
[3] van Kaick O, Tagliasacchi A, Sidi O, Zhang H, Cohen-Or D, Wolf L, Hamarneh G. Prior knowledge for part correspondence. Computer Graphics Forum (Proc. Eurographics) 2011;30:553–62.
[4] Golovinskiy A, Funkhouser T. Consistent segmentation of 3D models. Computers and Graphics (Shape Modeling International 09) 2009;33(3):262–9.
[5] Sidi O, van Kaick O, Kleiman Y, Zhang H, Cohen-Or D. Unsupervised co-segmentation of a set of shapes via descriptor-space spectral clustering. ACM Transactions on Graphics, (Proc. of SIGGRAPH Asia) 2011;30(6): p. 126:1–10.
[6] Huan Q, Koltun V, Guibas L. Joint shape segmentation with linear programming. ACM Transactions on Graphics, (Proc. of SIGGRAPH Asia) 2011;30(6): p. 125:1–12.
[7] Xu K, Li H, Zhang H, Cohen-Or D, Xiong Y, Cheng Z. Style-content separation by anisotropic part scales. ACM Transactions on Graphics, (Proc. of SIGGRAPH Asia) 2010;29(5):1–9.
[8] Shamir A. A survey on mesh segmentation techniques. Computer Graphics Forum 2008;27(6):1539–56.
[9] Hilaga M, Shinagawa Y, Kohmura T, Kunii TL. Topology matching for fully automatic similarity estimation of 3d shapes. In: Proceedings of the 28th annual conference on Computer graphics and interactive techniques, SIGGRAPH'01. 2001. p. 203–12.
[10] Zhang E, Mischaikow K, Turk G. Feature-based surface parameterization and texture mapping. ACM Transactions on Graphics 2005;24(1):1–27.
[11] Shapira L, Shamir A, Cohen-Or D. Consistent mesh partitioning and skeletonisation using the shape diameter function. Visual Computer 2008;24(4):249–59.
[12] Chen X, Golovinskiy A, Funkhouser T. A benchmark for 3D mesh segmentation. ACM Transactions on Graphics (Proc. SIGGRAPH) 2009;28(3):1–12.
[13] Kreavoy V, Julius D, Sheffer A. Model composition from interchangeable components. 2007. p. 129–38.
[14] Shapira L, Shalom S, Shamir A, Zhang RH, Cohen-Or D. Contextual part analogies in 3d objects. International Journal of Computer Vision 2010;89(2–3):309–26.
[15] Hu R, Fan L, Liu L. Co-segmentation of 3d shapes via subspace clustering. Computer Graphics Forum (Proc. SGP) 2012;31(5).
[16] Gatzke T, Grimm C, Garland M, Zelinka S. Curvature maps for local shape comparison. In: In shape modeling international. 2005. p. 244–56.
[17] Lai Y-K, Zhou Q-Y, Hu S-M, Wallner J, Pottmann H. Robust feature classification and editing. IEEE Transactions on Visualization and Computer Graphics 2007;13(1):34–45.

[18] Funkhouser T, Min P, Kazhdan M, Chen J, Halderman A, Dobkin D, Jacobs D. A search engine for 3D models. ACM Transactions on Graphics 2003;22(1): 83–105.
[19] Gal R, Shamir A, Cohen-Or D. Pose-oblivious shape signature. IEEE Transactions on Visualization and Computer Graphics 2007;13:261–71.
[20] Hoffman D, Richards W. Parts of recognition. Cognition 1983;18:65–96.
[21] Liu R, Zhang H, Shamir A, Cohen-Or D. A part-aware surface metric for shape analysis. Computer Graphics Forum (Special Issue of Eurographics 2009) 2009; 28(2):397–406.
[22] Lai Y-K, Hu S-M, Martin RR, Rosin PL. Rapid and effective segmentation of 3d models using random walks. Computer Aided Geometric Design 2009;26(6): 665–79.
[23] Fan L, Liu L, Liu K. Paint mesh cutting. Computer Graphic Forum (Proceedings of Eurographics) 2011;30(2):603–11.
[24] Du P, Ip HHS, Hua B, Feng J. Using surface variability characteristics for segmentation of deformable 3d objects with application to piecewise statistical deformable model. Visual Computer 2012;28(5):493–509.
[25] Ren X, Malik J. Learning a classification model for segmentation. Proc. IEEE international conference on computer vision, vol. 1. 2003. p. 10–7.
[26] Shi J, Malik J. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 2000;22(8):888–905.
[27] Golovinskiy A, Funkhouser T. Randomized cuts for 3D mesh analysis. ACM Transactions on Graphics (Proc. SIGGRAPH ASIA) 2008;27(5).
[28] Katz S, Tal A. Hierarchical mesh decomposition using fuzzy clustering and cuts. ACM Transactions on Graphics (Proc. SIGGRAPH) 2003;22(3):954–61.
[29] Ben-Chen M, Gotsman C. Characterizing shape using conformal factors. In: 3DOR. 2008. p. 1–8.
[30] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002;24(4):509–22.
[31] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) 2001;23(11):1222–39.