Interactive shape co-segmentation via label propagation

Zizhao Wu a, Ruyang Shou a, Yunhai Wang b, Xinguo Liu a

Q1 State Key Lab of CAD & CG, Zhejiang University, China
Q1 Shenzhen VisaCA Key Lab/SIAT, China

KEYWORDS
Shape co-segmentation
Interactive segmentation
Label propagation

1. Introduction

In recent years, there have been increasing interests in shape co-analysis, i.e., simultaneously analyzing a set of shapes. One of the most fundamental problems in this field is co-segmentation. Different from the traditional segmentation tools which treat shapes individually, co-segmentation approaches process shapes from an input set in a batch, and generate segmentations carrying consistent semantics across the shapes. The consistent segmentation has demonstrated great utility in modeling [1,2], shape retrieval [3,4], texturing [5], etc.

Previous attempts for solving this problem can be classified into three categories as supervised, semi-supervised and unsupervised. The supervised ones [5,6] take advantages of manually labeled training sets to generate consistent segmentation results. However, the accuracy of the results relies on the training sets, and not surprisingly, the training process is tedious and time consuming. The unsupervised methods [7,8] generally build their approaches on the patch-level. These methods have superior performance, but the results hinge upon the in-sample data.

Recently, Wang et al. [9] presented a semi-supervised learning method with the aid of constrained clustering, where the user can actively assist in the co-segmentation process by assigning pairwise constraints like must-link and cannot-link. This approach can generate error-free results with a sparse set of constraints. However, as some authors [10] mentioned, pairwise constraints are not expressive to the users. In addition, their approach is a transductive algorithm which does not handle with the out-of-sample data, i.e., given a new datum, it needs performing the algorithm over the whole pipeline, which is inefficient.

In this paper, we address the above issues by introducing an interactive shape co-segmentation method. Our motivation drives from label propagation which propagates labels through the dataset along high density areas defined by unlabeled data. Our method allows the users to participate in the co-segmentation procedure, and is built upon the patch-level, which guarantees the high speed. Specifically, starting from over-segmenting a set of shapes into primitive patches, we allow the users to assign labels to some patches, and propagate label information from these patches to the unlabeled ones. We iterate the last two steps until the error-free consistent segmentations are obtained. Additionally, we provide an inductive extension of our framework, which effectively addresses the out-of-sample data. The experimental results demonstrate the effectiveness of our approach.

In this paper, we present an interactive approach for shape co-segmentation via label propagation. Our intuitive approach is able to produce error-free results and is very effective at handling out-of-sample data. Specifically, we start by over-segmenting a set of shapes into primitive patches. Then, we allow the users to assign labels to some patches and propagate label information from these patches to the unlabeled ones. We iterate the last two steps until the error-free consistent segmentations are obtained. Additionally, we provide an inductive extension of our framework, which effectively addresses the out-of-sample data. The experimental results demonstrate the effectiveness of our approach.

In this paper, we address the above issues by introducing an interactive shape co-segmentation method. Our motivation drives from label propagation which propagates labels through the dataset along high density areas defined by unlabeled data. Our method allows the users to participate in the co-segmentation procedure, and is built upon the patch-level, which guarantees the high speed. Specifically, starting from over-segmenting a set of shapes into primitive patches, we allow the users to assign labels to some patches, and then propagate the labels from these patches to the unlabeled ones. We iterate the last two steps until the error-free consistent segmentations are obtained. Additionally, we provide an inductive extension of our framework, which effectively addresses the out-of-sample data. The experimental results demonstrate the effectiveness of our approach.
results on benchmark datasets in Section 4, followed by conclusions and future work in Section 5.

2. Related work

In this section, we provide a brief review of the existing work on shape co-segmentation, interactive segmentation and label propagation.

**Shape co-segmentation:** Shape co-segmentation refers to simultaneously segmenting a set of shapes into meaningful parts and building their correspondence. The existing co-segmentation methods can be classified into three categories: the unsupervised, the supervised, and the semi-supervised.

In the unsupervised setting, the early work reported by Golovinskiy and Funkhouser [11] builds reliable correspondences across segments of shapes using rigid shape alignment. However, their approach cannot handle shapes with large variations. Xu et al. [12] factor out the scale variation in the shape segments by clustering the shapes into different styles, depending on the scales of the shape parts. Still, their approaches are limited to the shapes that can be properly aligned.

To overcome this limitation, Huang et al. [13] introduce an optimization strategy for simultaneously optimizing the saliency of each segmentation as well as consistency between segmentations. However, due to the computational complexity, this approach does not scale well for large datasets.

Sidi et al. [7] present a descriptor-based method that employs multiple feature descriptors to measure the similarities of the segments and poses co-segmentation as a clustering problem in a concatenated descriptor space. Because the descriptors are independent of the pose and location of the shapes, this method can handle shapes with rich variations in part composition and geometry. Instead of concatenating the different feature descriptors into one vector, Hu et al. [8] propose a feature fusion method to co-segment a set of shapes via subspace clustering. However, these unsupervised techniques hinge upon the in-sample data.

Kalogerakis et al. [5] present a supervised learning method to simultaneously segment and label shapes. Their approach needs prior knowledge learned from the training dataset, and has demonstrated a labeling high accuracy on a broad class of shapes. van Kaick et al. [6] optimize the previous method by incorporating the prior knowledge to train a classifier. However, the above supervised methods require a substantial number of manually labeled training shapes, and the training set has a large impact on the segmentation performance.

Very recently, Wang et al. [9] propose a semi-supervised method where the user can actively assist in the learning process by interactively providing inputs. The input consists of a sparse set of pairwise constraints, which are marked as must-link and cannot-link constraints. The authors show that a sparse set of constraints can quickly converge toward an error-free result. However, the pairwise constraints are not clearly expressed to the users. In addition, their approach is a transductive algorithm that is ineffective at handling out-of-sample data.

**Interactive segmentation:** Interactive shape segmentation approaches are simple and intuitively help users express their intentions. Consequently, they have received significant attention [14].

Many interactive techniques have been proposed. Some of them require the user to specify a few points on the desired cutting contour and then employ the geometric snake [15], scissoring [16,17], graph cut [18] or some other method [19] to find the final cutting boundaries. These methods are called boundary-based approaches.

In the last few years, a series of region-based approaches [20–22] have been proposed, which take regional information as the input and require a much smaller amount of user effort to complete the labeling process for all of the unlabeled faces of a shape.

In this paper, rather than segmenting an individual shape, we present an interactive region-based technique to simultaneously segment a set of shapes in a consistent manner.

**Label propagation:** Label propagation was first introduced by Zhu and Ghahramani [23]. This technique propagates the labels through dense unlabeled regions and locates data with properties that are similar to those of the labeled data. Their approach is graph-based, which can be constructed straightforwardly by computing pairwise similarities among all of the data. Due to its simplicity and robustness, it has been used in processes such as patch labeling [24], image segmentation [25], and image annotation [26].

Some authors [27,28] have tried to optimize the original label propagation. Among them, Wang and Zhang [29] propose approximating the graph with a set of overlapped linear neighborhood patches (LNPs) and computing the edge weights in each patch using the neighborhood linear projection. Our work is directly inspired by the LNP. We apply this algorithm to our interactive shape co-segmentation setting.

3. Algorithm

3.1. Overview

Define a set of shapes \( S = \{s_1, s_2, \ldots, s_M\} \), where \( s_i \) represents the \( i \)-th shape and \( N \) is the total number of shapes. Our algorithm simultaneously produces segments of the set of shapes \( S \) and builds their correspondences across these segments.

The pipeline of our approach is illustrated in Fig. 2. First, the algorithm pre-processes the set of shapes by partitioning the dataset into primitive patches and building a graph that represents the geometric similarities across them. Then, the user interactively labels some patches, which are used as initial seeds that guide the iterative propagation to find labels for the others.

Our algorithm is an iterative approach. Each iteration includes two steps: user interaction and label propagation based on the user input. These steps repeat until satisfactory results are obtained. Additionally, we apply an extension to the pipeline to handle out-of-sample data.

We discuss the preprocessing step in the next section, the label propagation in Section 3.3, and the inductive extension in Section 3.4.

3.2. Preprocessing

In this step, we start by over-segmenting the input shapes, where normalized cuts [30] are employed to decompose each shape \( s_i \) into primitive patches. In our settings, the number of patches per shape is set to 30. Let \( P = \{p_1, p_2, \ldots, p_M\} \) be the set of patches from all of the shapes; it is clear that \( M = 30N \). Fig. 1 gives an example of our over-segmentation results.

Our approach associates the representation of relations between the patches with graphs. We represent this graph in matrix form, i.e., by constructing an affinity matrix \( W \) whose entries \( w_{ij} \) carry the similarities of \( p_i \) and \( p_j \). Thus, to measure the similarities among patches, we first choose five robust and discriminative shape descriptors to extract extrinsic geometric information about the patches; these can be informatively represented based on these data. These widely recognized descriptors are the Shape Diameter Function (SDF) [31], the Conformal Factor (CF) [32], the Shape Contexts (SCs) [33,34], the Average Geodesic Distance (AGD) [35], and the geodesic distance to the base of the shape (GB) [7]. The descriptors are all defined on the mesh faces, so no additional conversions are required to make them mutually representationally compatible. Then, to describe the patches using the descriptors of the faces within them, we incorporate histograms. Specifically, for patch \( p_i \) we first build for each descriptor...
that the local similarities are more reliable than the distant ones, which has been widely accepted in many research communities [36].

3.3. Labeling and propagation

Our algorithm allows the user to actively assign some labels to a small subset of patches as the seeds for propagation. In this section, we discuss the iterative process, which repeatedly asks the user to specify some seedling labels and provides the user with the segmentation results via label propagation until satisfactory results are achieved. For convenience, we denote the labeled patches as $X_l \subset P$ and the unlabeled ones as $X_u$. We also define $L = \{l_1, l_2, \ldots, l_C\}$ as the label set and $y_l \in L$ as the label of $p_i$.

The goal of this step is to propagate the labeled set $X_l$ to the unlabeled set $X_u$, based on the affinity matrix $W$. Let $F$ be the $M \times C$ labeling matrix that associates each patch $p_i$, $i \in M$ with each label $l_j$, $j \leq C$, of which the element $f_{ij} \in F$ denotes the probability that $p_i$ corresponds to the $j$-th label. After propagation, each unlabeled patch has a probability of belonging to each label. We assign to $y_l$ the label that is the maximum in the set $\{f_{ij} \}_{j \leq C}$, i.e., $y_l = \arg \max_{j \leq C} f_{ij}$.

We use an iterative convergence method to propagate labels to the unlabeled patches. In each iteration, the patches both absorb the label information from their neighbors and retain some fractions of their own initial label stations. This procedure repeats until none of the patches’ labels changes. Let $t$ be a time stamp and $F^t$ be the labeling function at time $t$. The label propagation can be written as

$$F^{t+1} = \alpha W F^t + (1-\alpha) F^0,$$

where $\alpha \in (0, 1)$ is the weight, based on the label information inherited from the neighbors and the initial possibilities. $F^0$ is constructed according to $X_l$ and $X_u$; that is, $f_{ij}^0 = 1$ if $p_i$ is labeled as $l_j$ by the results of a previous iteration or assignments from a user which can overwrite the former if applicable, otherwise $f_{ij}^0 = 0$. Our algorithm iterates Eq. (2) until convergence. Let $F^t$ be the limit of $F^t$. Finally, we output the labels $Y$ with $y_l = \arg \max_{j \leq C} f_{ij}^t$ for the unlabeled patches. This procedure is summarized as Algorithm 1.

The convergence analysis of the label propagation algorithm is presented in Appendix A, where we refer to the LNP [29].

3.4. Inductive extension

To enhance the scalability of our algorithm and handle out-of-sample computations, we propose an inductive extension. Given a new input shape, we first pre-process it using the method mentioned in Section 3.2, obtaining small patches with the corresponding histograms. Denoting $h_k$ as the histogram of patch $p_k$ of the new shape, we then assign the weights to its linear neighbors by minimizing the following objective function:

$$\arg \min_{w_{kj}} \sum_{l \in L(h_k)} \sum_{i \in M} w_{kj} f_{ij}^t,$$

where $\Omega(h_k)$ denotes the linear neighborhoods of $h_k$. After computing the weights, we can directly calculate $y_l$ through the
following interpolation:

\[ y_k = \arg \max_{j \leq c} \sum_{i_h \epsilon \Omega_{h(j)}} \sum_{j} w_{i_h,j} f_{i,j}. \]  

(3)

here \( y_k \) is the final label for \( p_k \).

Algorithm 1. Label propagation algorithm.

**Input:** The set of over-segmented patches \( P \); The set of labels \( L = \{l_1, l_2, \ldots, l_c\}; X_k \subseteq P \), where patches are assigned to labels in \( L \) and \( X_k = P - X_i \); The balancing parameter \( \alpha \in (0, 1) \).

**Output:** The labels of patches \( Y = \{y_1, y_2, \ldots, y_c\} \).

1. Employ descriptors to measure the patch and then construct the linear neighborhood graph \( W \) by solving Eq. (1) based on the histograms of the patches.
2. Construct the labeling matrix \( P^T \) based on the label assignment of \( X_k \).
3. Iterate Eq. (2) until convergence. Let \( f^* \) be the limit of \( F \).
4. Output the labels \( Y \) with \( y_i = \arg \min f_{i,j}^* \) for the unlabeled patch \( X_i \subseteq P \).
5. Iterate Steps 2 and 3 by interactively assigning labels to the patches until satisfactory results are obtained.

4. Experimental results

We evaluate our approach on the shape co-segmentation benchmark (COSEG) [9]. We set \( \alpha = 0.5 \) in all our experiments. Fig. 5 shows the visual results of our approach, which show that we can obtain close to error-free results with the help of user interaction and label propagation. More specifically, we use the 7, 15, 6, 12, 10, 21, 12 label assignments on the Lamps, Candelabra, Goblets, Irons, Guitars, Vases and Chairs sets of COSEG to generate the results. Note that other alternative assignments may produce the same results; the user can add as many constraints as desired to refine the results, depending on the user’s intention.

It would be challenging for our algorithm to label large datasets because a small modification of the labeling patches may substantially change the final results. In our experiments, we asked three users to label these large datasets until satisfactory results were gained. After obtaining these results, we choose the results with minimal labeling effort for each set. In our results, we have taken 113 assignments, 63 assignments and 97 assignments for the Large chairs, the Large vases and the Large tele-aliens sets, respectively. We note that it is a good choice to divide the large set into several smaller sets and then label these sets separately. Taking the large tele-aliens as an example, we divided the set into four groups; each group has 50 shapes. Then, we performed our algorithm to obtain the error-free results for each group. As a result, we produced 17, 19, 25, and 16 assignments, and the total number of assignments is 92, which is more effective than considering the large set as a whole.

Fig. 6 illustrates some co-segmentation results of the unsupervised approaches [7,8] on the COSEG. These results contain many mislabeled patches because the geometry alone cannot fully convey the semantics of parts. Instead, by incorporating the user’s guidance, our approach can effectively alleviate this problem.

We also demonstrate the effectiveness of label propagation in Figs. 3 and 4. Given a dataset, by assigning some labels on some points, our approach propagates the label information from the labeled points to the unlabeled points. The labeling status of the whole dataset highly depends on the initial labeling assignments.

**User study:** We conducted a user study to compare our method and the method of Wang et al. [9]. Ten participants were invited to perform the experiments on several small datasets (Candelabra, Irons) and large datasets (Large vases, Large tele-aliens) from the COSEG. We required the participants to use the two interfaces of these semi-supervised leaning approaches for each dataset, where some fixed numbers for the constraints are given. Then, we measured the accuracy of co-segmentation results for each method under the given constraints. The method with a higher accuracy was considered to be more effective. In our setting, we used the average accuracy as the final score of a dataset for each method, and the users were given four numbers (5, 10, 15, 20) for the small datasets and four numbers (20, 40, 60, 80) for the large datasets. Finally, because the users wanted to achieve satisfying results easily and quickly, we also required the users to judge which method is more intuitive.

Please cite this article as: Wu Z, et al. Interactive shape co-segmentation via label propagation. Comput Graph (2013), http://dx.doi.org/10.1016/j.cag.2013.11.009
We used the ground truth segmentation in [9] and their accuracy measure as our quality measure:

\[
\text{Accuracy}(l, t) = \frac{\sum_i a_i \delta(l_i = t_i)}{\sum_i a_i}
\]

where \(a_i\) is the area of the face \(i\), \(l\) is the label computed by the co-segmentation, \(t\) is the ground-truth labeling, and \(\delta(x = y)\) is 1 if and only if \(x\) equals \(y\). We average the accuracies for all of the shapes in the set as the accuracy of the co-segmentation method.

Fig. 7 shows the average co-segmentation accuracies achieved by the participants in our experiments. We can see that our system helps the users to perform better than the other method [9]. In addition, we provide much faster feedback when updating the constraints, which we discuss later. As for the survey results, 7 users thought that our interface was more intuitive, while the other 3 users stated that our method was inconvenient due to the lack of active learning.

**Inductive evaluation:** We tested the inductive extension of our algorithm on the Large chair dataset, where we randomly selected 200 shapes as in-sample data and used the remaining samples as out-of-sample data. After we obtained the error-free results of the in-sample data, the inductive extension was conducted to predict the labels of the out-of-sample data. Fig. 8 shows an example where a new datum has been pre-processed and labeled. As a result, we achieved an average of 88.7% accuracy for the out-of-sample shapes, which is associated with \(O(n)\) time complexity (less than 1 s for each shape in our case). In comparison, the method of Wang et al. [9] requires \(O(n^3)\) (20 min). This result demonstrates the effectiveness of our inductive algorithm.
Performance: We implemented our approach in C++ and MATLAB and evaluated the performance on a 2.83 GHz Intel Core™ 2 Quad processor with 4 GB of RAM.

The pre-processing of our pipeline requires considerable time, particularly the computation of geodesic distance and shape context. The label propagation is very fast; the time complexity is $O(n^2)$, where $n$ is the total number of patches. Consider a small dataset that contains 20 shapes, where each shape is decomposed into 30 patches. Therefore, the linear neighborhood matrix $W$ has the dimension of $600 \times 600$, and our algorithm needs approximately 100 iterations to converge, which takes approximately 7 s for one update of its labels. For a large dataset, we need approximately 8 min to complete an update.

5. Conclusions

We presented a novel algorithm to interactively co-segment a set of shapes via label propagation. Given a set of shapes with a small number of user-defined labels on the over-segmented patches, our method automatically propagates those labels to the remaining unlabeled patches through an iterative procedure. Moreover, we provide an out-of-sample extension of our approach. The experiments demonstrated the effectiveness of the proposed approach.

Limitations and future work: There are many limitations in our algorithm, which suggests many avenues for future work. First, our approach suffers from a lack of suggestions when labeling samples...
to improve the co-segmentation results with less effort. In the future, we plan to integrate our labeling process with the active learning setting [38]. Second, our approach is graph-based. How to better measure the similarities preserved in the graph plays an important role in the following steps. A more rigorous treatment of this problem is appreciated. Finally, our approach builds upon the patch-level, so the final co-segmentation cuts are relying on the initial over-segmentation boundaries, which encourages us to exploit a more robust over-segmentation method to refine the cutting boundaries for the applications.

Acknowledgment

Many thanks to the anonymous reviewers for their valuable comments. This work was partially supported by NSFC (Nos. 6137 9068 and 61202222), Shenzhen Science and Innovation Program (CXB201604220029A). [44]

Appendix A. Convergence analysis

In this section, we present the convergence analysis of Algorithm 1. We can see from the Algorithm 1 that our approach needs to iterate Eq. (2) until it converges with the given input α, W and P0. By substituting variable F in Eq. (2), we have

\[ F_{t+1} = \alpha W f_t P_t + (1 - \alpha) \sum_{n=0}^{t} (\alpha W)P_0. \]  

Because we have constrained \( \sum w_{ij} = 1 \) and \( w_{ij} \geq 0 \), according to the theorem of Perron–Frobenius [37], we can conclude that the spectral radius of \( W \) is no larger than one. In addition, \( 0 \leq \alpha \leq 1 \). We have

\[ \lim_{t \to \infty} \alpha W f_t = 0 \quad \text{and} \quad \lim_{t \to \infty} \sum_{n=0}^{t} (\alpha W)^n = (I - \alpha W)^{-1}. \]

where \( I \) is the identity matrix. From these, Eq. (2) will converge to

\[ F^* = \lim_{t \to \infty} F_t = (1 - \alpha)(I - \alpha W)^{-1}P_0. \]  

In other words, we can predict the labeling results in one step.

References


Please cite this article as: Wu Z, et al. Interactive shape co-segmentation via label propagation. Comput Graph (2013), http://dx.doi.org/10.1016/j.cag.2013.11.009