Unsupervised Image Co-segmentation Based on Cooperative Game

Bo-Chen Lin, Ding-Jie Chen, and Long-Wen Chang

Department of Computer Science, National Tsing Hua University, Taiwan

Abstract. In computer vision, co-segmentation is defined as the task of jointly segmenting the common objects in a given set of images. Most proposed co-segmentation algorithms have the assumptions that the common objects are singletons or with the similar size. In addition, they might assume that the background features are simple or discriminative. This paper presents a cooperative co-segmentation without these assumptions. In the proposed cooperative co-segmentation algorithm, each image is treated as a player. By using the cooperative game, heat diffusion, and image saliency, we design a constrained utility function for each player. This constrained utility function push all players, with the instinct to maximize their self-utility, to cooperatively define the common-object labels. We then use cooperative cut to segment the common objects according to the common-object labels. Experimental results demonstrate that the proposed method outperforms the state-of-the-art co-segmentation methods in the segmentation accuracy of the common objects in the images.

1 Introduction

Image segmentation is a fundamental problem in computer vision. Segmentation partitions an image into several regions that each region shares certain similar appearances. The goal of segmentation is to simplify the representation of an image for locating the objects. An important issue of image segmentation is that the regions found by a typical image segmentation algorithm usually tend to be fragmented or lack semantic meanings. That is, it is difficult to locate the objects from a single image. Therefore, Rother et al. [1] proposed the idea of co-segmentation that one additional image is provided to segment both images together to increase the accuracy of the object segmentation.

Recently, co-segmentation has been widely studied in computer vision. The goal of image co-segmentation refers to segment the similar regions from two or more images. Although the authors [1–9] proposed some methods to solve this problem, there are still some restrictions as follows:

1) Some algorithms are supervised.
2) The given images have only one instance of the common object.
3) The backgrounds of the given images are discriminative.

Heat diffusion framework [9–11] is a successful technique in image processing and computer vision. It can be applied in image segmentation [10], and optical flow estimation [11]. Here we adopt the heat-gain of this framework to measure the segmentation confidence [9]. In order to deal with images with similar back-ground, we use the

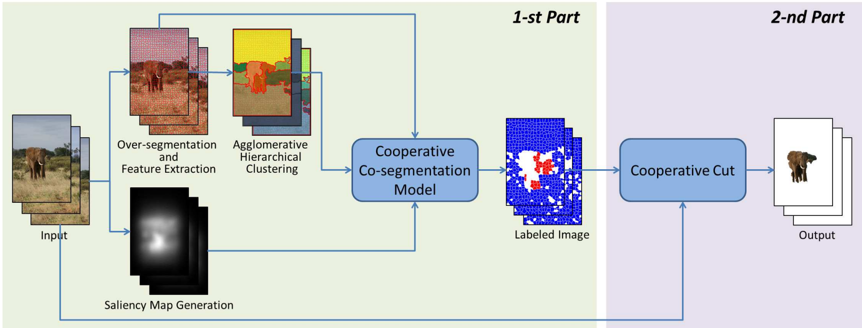


Fig. 1. The block diagram of the proposed image co-segmentation method.

saliency map [12] in our algorithm to provide an initial guess about the object positions. Game theory [13] has been widely used as a powerful method to solve problems in social science, biology, economics, computer science [14, 15], etc. It is the study of the rational choice of strategies by interacting agents called players. In this paper, we model the labeling problem as a cooperative game to constrain the labeling procedure.

The co-segmentation problem implies cooperative feasibility for jointly segmenting the common objects among the images. In this paper, we base on the cooperative game, heat diffusion, and image saliency to propose a cooperative co-segmentation framework (see Fig. 1) for overcoming the aforementioned restrictions. In the proposed method, each image is treated as a player in the heat diffusion system, and her heat gain is treated as her utility. In the first part, the superpixels with very low saliency value are labeled as background. Next, all players cooperatively define the common-object superpixels with their constrained utility function. The remaining unlabeled superpixels are treated as neutral. In the second part, we apply the cooperative cut [16] with the aid of the labeled superpixels and thus generate the pixel-level segmentation.

There are three advantages of the proposed cooperative co-segmentation method. First, the proposed method can discover multiple instances of the common objects. Secondly, our method is capable of handling images whose backgrounds are similar. Thirdly, we can segment the common objects with different scales and achieve higher accuracy than other methods [6, 9].

1.1 Related Work

Existing image co-segmentation works [1, 5, 7, 8] formulated the co-segmentation problem as a binary labeling problem. Their objectives are minimization of an energy function with a histogram difference term which derived from the input image pairs. The histogram difference term penalizes the difference between the foreground histograms calculated from the input images. Since the histogram difference term is computed between any two images, the energy function is computationally intractable as the number of the input images increase. In addition, these methods implicitly assume that only one object appears in each image.

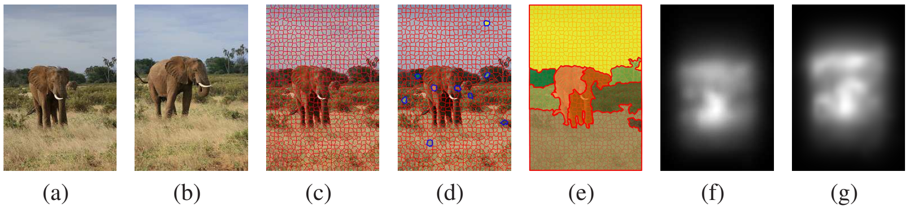


Fig. 2. (a) and (b) are Elephants-safari from the CMU-Cornell iCoseg dataset [2]. (c) the over-segmentation of (a). (d) the representative superpixels set of (c). (e) The yellow sky region corresponding to the yellow representative superpixel of (d). (f) and (g) are the corresponding saliency maps of (a) and (b).

Joulin et al. [6] proposed a discriminative clustering based image co-segmentation. The main idea in this method is to train a supervised classifier for maximal separation of the foreground and the background. Although it can solve the co-segmentation problem for up to dozens of images, the segmentation results are not satisfactory for the number of the input images less than a certain number. Kim et al. [9] proposed the distributed co-segmentation algorithm based on temperature maximization on anisotropic heat diffusion. The approach can deal with a large number of input images. Chu et al. [4] proposed a method that has the ability to segment multiple objects that repeatedly appear among input images. It incorporates a common pattern discovery algorithm, that using the SIFT descriptor, with an energy function. The method can achieve high accuracy of segmentation when the common objects have high texture complexity.

However, the above unsupervised co-segmentation methods [1, 4–9] require the testing dataset to be chosen carefully, because their methods would fail for images with similar background. For example, in Fig. 2(a) and Fig. 2(b), the backgrounds of images are so similar that the common objects cannot be cut out via unsupervised co-segmentation methods. Due to this problem, some interactive co-segmentation techniques were proposed [2, 3], which allow users to decide where the foreground or background is, and then users can guide the output of the co-segmentation algorithm toward it via scribbles.

2 Cooperative Image Co-segmentation

Given a set of input images, the co-segmentation goal is to segment the common objects among these images. The block diagram of our method is given in Fig. 1, which is divided into two parts. The first part consists of four stages, namely *over-segmentation and feature extraction*, *agglomerative hierarchical clustering*, *saliency map generation*, and *cooperative co-segmentation model*. In this part, the goal is to label each input image. We label the images with common-object-label, background-label, and neutral-label. The neutral-label is just used to denote the unsure regions. In the second part, we use cooperative cut [16] with the labeled image-regions to obtain the final result.

2.1 Over-segmentation and Feature Extraction

In order to reduce the computational loading in the following heat diffusion system, we represent each image as a set of superpixels. We adopt the over-segmentation method [17] to obtain the regular-size superpixels. Fig. 2(c) shows an example of the over-segmentation [17] that applied on Fig. 2(a).

After over-segmentation, we need some descriptors to describe each superpixel. Color and texture descriptors are commonly used in computer vision. The color of a superpixel can be represented in terms of average color or color histogram. Texture descriptor is used to describe the superpixel in terms of its texture property. Here we simply represent each superpixel as a 3-dimensional average color vector.

2.2 Agglomerative Hierarchical Clustering

To select the representative superpixels as the candidate heat sources in the following heat diffusion system, we apply the agglomerative hierarchical clustering [9] to find out some representative superpixels. Precisely, a set of superpixels with the similar features will be represented as one representative superpixel. Fig. 2(c) is the superpixel representation of Fig. 2(a). A region, for example, the sky in Fig. 2(e), consists of a set of superpixels of similar features and represented with the yellow representative superpixel in Fig. 2(d).

2.3 Saliency Map Generation

We assume that the objects in the foreground usually have higher saliency value than those in background. The saliency detection methods usually focus on identifying the fixation points that human viewer would focus on at the first glance. Harel et al. [12] proposed a method of computing bottom-up saliency maps which shows a remarkable consistency with the deployment of attention of human subjects.

For an input image, the method [12] extracts three kinds of features of each pixel, thus generate three kinds of feature maps. For each kind of feature map, the method obtains corresponding activation map by computing Markov chains. Finally, the method normalizes and averages the three kinds of activation maps to generate the saliency map. The saliency maps of Fig. 2(a)-(b) are shown as Fig. 2(f)-(g).

2.4 Cooperative Co-segmentation Model

With a heat diffusion system, each image can evaluate the segmentation confidence of each region with the value of heat gain [9]. In addition, the heat gain is proportional to this segmentation confidence in a heat diffusion system. The image co-segmentation goal is to segment the common object region. Intuitively, the common object region should has as high segmentation confidence, i.e., heat gain, as possible. However, we observed that a representative superpixel with high heat gain could be the foreground or the background. In order to label the representative superpixels right on the foreground as the common object, here we consider the image saliency.

In image co-segmentation scenario, a superpixel is considered as the common object candidate if it has similar appearance with the superpixels from other images. This means that an image should selects its superpixels as the common object with considering what are the selected superpixels of other images. This kind of consideration is like the scenario in game theory, that is, each player choose her best strategy according to the strategies chosen by other players.

In game theory [13], one assumption is that each player is rational. Namely, each player maximizes her utility, given the adopted strategies of other players. In cooperative game, each player still need to maximize her utility. However, the designed utility functions also trigger them to maximize the coalition utility. In this paper, we design a utility function constrained on the other players' strategies, self heat gain, and self image saliency. To begin with labeling, the superpixels with very low saliency value are regarded as background. Then, all players cooperatively define the common-object-label via maximizing their constrained utility functions. The remaining unlabeled superpixels are assigned the neutral-label. Finally, we use the cooperative cut [16] with the label information to finely segment the images.

Heat Diffusion System Each image corresponds to a heat diffusion system while the images correspond to those systems are coupled together. In each system, there are K_i heat sources, i.e., representative superpixels.

Given an input image set I. For each input image $I_i \in I$ $(i = 1, ..., N)$ consider a graph $\mathcal{G}_i = (\mathcal{V}_i, \mathcal{E}_i)$ where the node set $\mathcal{V}_i$ is the set of superpixels of I_i, and the edge set $\mathcal{E}_i$ connect all pairs of adjacent superpixels in $\mathcal{V}_i$. The heat diffusion system [9] has the following definitions:

1) Each input image I_i is an insulated heat diffusion system T_i. A heat diffusion system T_i contains: (i). a temperature function u_i. (ii). an environment node g_i with zero temperate, denoted by $u_i(g_i) = 0$. (iii). a heat source node h_i with constant temperate, denoted by $u_i(h_i) = 1$.
2) Each node v_x in T_i diffuses heat to its neighbors and is connected to an environment node with constant diffusivity of z_{v_x}. The diffusivity between any two nodes $v_x, v_y \in \mathcal{V}_i$ are defined by their Gaussian similarity:
$$d_{v_x,v_y} = \begin{cases} \exp(-\gamma\|f^c(v_x) - f^c(v_y)\|^2), & \text{if } (v_x, v_y) \in \mathcal{E}_i \\ 0, & \text{otherwise} \end{cases}, \tag{1}$$
where $f^c(v_x)$ is the average color of the pixels in v_x, γ is a constant parameter.
3) The diffusion equation for $v_x \in \mathcal{V}_i$ is defined as follows:
$$u_i(v_x) = \frac{1}{a_{v_x}} \sum_{(v_x,v_y)\in\mathcal{E}_i} d_{v_y v_x} u_i(v_y)\,, \tag{2}$$
where $a_{v_x} = \Sigma_{(v_x,v_y)\in\mathcal{E}_i} d_{v_y v_x} + z_{v_x}$ is a normalization factor.
4) Assume that the system temperature is zero before putting the heat source h_i on v_x. Once putting a heat source h_i on v_x, the corresponding heat gain δ_i is computed by
$$\delta_i(v_x) = \sum_{(v_x,v_y)\in\mathcal{E}_i} u_i(v_y). \tag{3}$$

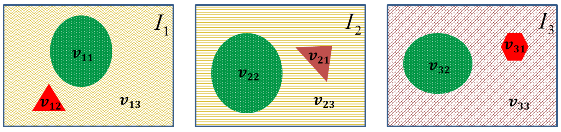


Fig. 3. An example to explain the cooperative behavior conditioned on function π. Each image I_i has three different regions, and each region can be represented as the corresponding representative superpixels v_{i_1}, v_{i_2}, and v_{i_3}. Since the objects usually have higher saliency value than background, the strategies $v_{11}, v_{12}, v_{21}, v_{22}, v_{31}, v_{32}$ can pass the saliency function ϕ in π. Since any pair from the two strategy sets $\{v_{11}, v_{22}, v_{32}\}$ and $\{v_{12}, v_{21}, v_{31}\}$ have similar color features, the pairs from these two sets can obtain high values of the similarity function ψ. Thus, the strategy profiles which can pass the π function will belong to $\{v_{11}, v_{12}\} \times \{v_{21}, v_{22}\} \times \{v_{31}, v_{32}\}$. However, the optimal strategy profile is (v_{11}, v_{22}, v_{32}) because it has the largest sum of the pairwise feature similarities according to the constrained utility function.

Cooperative Label Generation Model We propose a cooperative game model to assign the common-object-label to each image, which is configured as follows:

Players: Given an input image set I, each image I_i $(i = 1, ..., N)$ is regarded as a player in the game. The superpixels of I_i are separated into K_i clusters with the aforementioned agglomerative hierarchical clustering. The collection of these representative superpixels of I_i are denoted by R_i.

Strategies: The strategy set of each player I_i is $R_i = \{v_{i_1}, v_{i_2}, ..., v_{i_{K_i}}\}$. Each player choose one strategy $v_i \in R_i$ to put her heat source in one diffusion process. We denote the strategy profile **v** of all players as $(v_1, v_2, ..., v_N) \in R_1 \times R_2 \times ... \times R_N$.

Preference: We treat the image set I with the common objects as a coalition. In a cooperative game, each player should takes the strategy with considering what are the adopted strategies of other players. Thus we define the preference of each player I_i is represented by the constrained utility function U_i as follows:

$$U_i(v_i|\mathbf{v}) = \pi(v_i|\mathbf{v})\delta_i(v_i)(\frac{1}{|N-1|}\sum_{j\in -i}\psi(v_i, v_j))\,, \tag{4}$$

where $-i$ denotes all players except i. Precisely, the similarity function ψ of any two representative superpixels $v_i \in R_i, v_j \in R_j$ is defined as the Gaussian similarity:

$$\psi(v_i, v_j) = \exp(-\gamma||f^c(v_i) - f^c(v_j)||^2)\,, \tag{5}$$

where $f^c(v_i)$ is the average color of the superpixels v_i, and γ is a constant parameter. Considering with the image saliency, the candidate strategy s_i of player I_i obeys the following function:

$$\pi(v_i|\mathbf{v}) = \begin{cases} 1, \text{ if } \psi(v_i, v_{-i}) > \alpha \text{ and } \phi(v_i) > \beta \\ 0, \text{ otherwise} \end{cases}\,, \tag{6}$$

where α and β are threshold parameters, $\phi(v_i)$ is the average saliency value of all pixels in v_i. Eq. (6) shows that we only concern the superpixel v_i with high saliency value and with feature similar to other players' strategies. Fig. 3 shows a simple example to explain the cooperative behavior conditioned on function π.

The goal of the cooperative model is to find the optimal strategy profile **v*** that maximizes the coalition utility. That is,

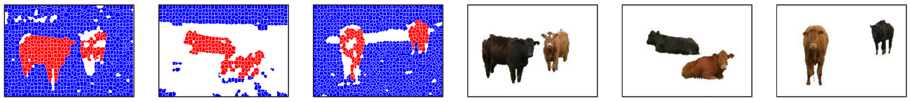
$$\mathbf{v}^* = \underset{\mathbf{v}}{\operatorname{argmax}}\left(\sum_{i=1}^{N} U_i(v_i|\mathbf{v})\right), \tag{7}$$

where $\mathbf{v} = (v_1, v_2, ..., v_N) \in R_1 \times R_2 \times ... \times R_N$. In the designed utility function, the best response of each player I_i is conditioned on not only the heat gain δ of herself but also the feature similarity ψ comparing with other players' strategies. It is hard to calculate the exact optimal strategy profile $\mathbf{v}^* \in R_1 \times R_2 \times ... \times R_N$. In practice, we use loopy belief propagation (LBP) [18] to approximate the optimal strategy profile.

The common objects usually represented as several representative superpixels (see Fig. 2(d)). That is, we need to find more than one optimal strategy profiles **v*** as the common-object-label. We summarize the complete label generation as algorithm 1 which based on the greedy method and LBP. In algorithm 1, before assigning the common-object-label, the background-label is assigned to the representative superpixels which have very low saliency values. After assigning the common-object-label, the neutral-label is assigned to the remaining unlabeled representative superpixels. Notice that, once a representative superpixel v_i is labeled, all the superpixel represented by it will get the same label.

Algorithm 1 Label Generation

Input: N players: image set $\{I_1, I_2, \cdots, I_N\}$; Strategy set: each I_i has strategy set $R_i = \{v_{i_1}, v_{i_2}, \cdots, v_{i_{K_i}}\}$; Parameter set: $\{\alpha, \beta\}$.
Output: Background-label set: $\mathbf{V}_B$; Common-object-label set: $\mathbf{V}_C$; Neutral-label set: $\mathbf{V}_N$;
1: $\mathbf{V}_B = \emptyset; \mathbf{V}_C = \emptyset; \mathbf{V}_N = \emptyset$;
2: For all $v_i \in R_i$, if $\phi(v_i) < \beta$ then remove v_i from R_i and $\mathbf{V}_B = \mathbf{V}_B \cup v_i$;
3: Construct graph G with node $\{R_1, R_2, \cdots, R_N\}$ and edge $\{(v_i, v_j)|v_i \in R_i, v_j \in R_{-i}\}$;
4: **for all** $v_i \in R_i, v_j \in R_{-i}$ **do**
5:     **if** $\psi(v_i, v_{-i}) > \alpha$ **then**
6:         define edge weight $w(v_i, v_j) = \frac{\delta_i(v_i)\psi(v_i, v_j)}{|N-1|}$;
7:     **else**
8:         define edge weight $w(v_i, v_j) = 0$;
9:     **end if**
10: **end for**
11: Iteration $t = 1$;
12: **while** each R_i is non-empty **do**
13:     $\mathbf{v}^*_t \leftarrow LBP(G)$; /* state set of image I_i is R_i */
14:     **for all** R_i **do**
15:         for all $v_i \in R_i$, if $v_i \in \mathbf{v}^*_t$ then remove v_i from R_i;
16:     **end for**
17:     $\mathbf{V}_C = \mathbf{V}_C \cup \mathbf{v}^*_t; t = t + 1$;
18:     reconstruct graph G with updated R_i as line 3 to line 10;
19: **end while**
20: **return** $\mathbf{V}_B, \mathbf{V}_C, \mathbf{V}_N = \{R_i|i = 1, ..., N\}$; /* $\mathbf{V}_N$: the remaining unlabeled strategies */



(a) (b) (c) (d) (e) (f)

Fig. 4. The co-segmentation result of the proposed method. Fig. 5(a)-(c) are the input images. (a)-(c) show the labeled results generated by the proposed cooperative game. The red parts are the common object labels, the blue parts are the background labels, and the white parts are the neutral labels. (d)-(f) show the segmentation results after the cooperative cut.

2.5 Cooperative Cut

For the input images, the proposed cooperative co-segmentation generate the corresponding labeled images such as Fig. 4(a)-(c). The remaining problem is to label the neutral-label regions as the common-object or the background, and thus yielding the final segmentation results such as Fig. 4(d)-(f). Given some labeled image regions, a cut-algorithm [16, 19–21] is used to label the remaining unlabeled image regions. In practice, graph cuts [19–21] is known to shortcut elongated boundaries, especially in low contrast or shaded region. Thus, Jegelka et al. proposed the cooperative cut [16] to utilize edge cooperation to selectively reward global features of true boundaries in the image. It has ability to segment fine structured objects and objects with shading variation.

In our experiments, we use the cooperative cut [16] with the given common-object regions and background regions to label the neutral-label regions in the pixel-level. The parameter setting is the same as [16].

3 Experimental Results

We discuss the experimental results on several image sets for evaluating the performance of the proposed cooperative co-segmentation method. The test images are collected from various database such as CMU-Cornell iCoseg dataset [2], MSRC dataset [22], and ImageNet [23]. We present qualitative and quantitative results of our algorithm. The segmentation accuracy of a given image is measured by the intersection-over-union metric. The metric defined as $Acc_i = \frac{GT_i \cap S_i}{GT_i \cup S_i}$, where GT is the ground truth segment, S is the segment obtained by the co-segmentation algorithm.

3.1 Cooperative Behavior

In the proposed cooperative co-segmentation, we designed the constrained utility function. The effects of the constrained utility function are shown in Fig. 5. Fig. 5(a)-(c) are the input images. Fig. 5(d)-(f) show the common-object-label results of noncooperative behavior. That is, the utility function only consider the heat gain and all edges have the same weight. We can find that each player just chooses the strategy to maximize his own heat gain. Fig. 5(g)-(i) shows the common-object-label results of the constrained utility function without considering the saliency maps. Fig. 5(j)-(l) shows the common-object-label results of the constrained utility function with considering the saliency maps.

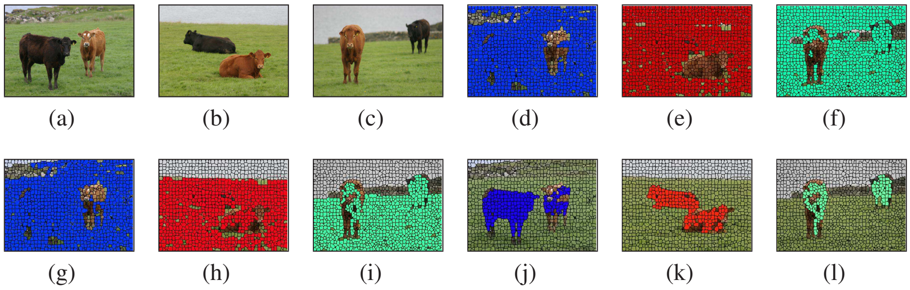


Fig. 5. Effect of the constrained utility function. (a)-(c) are the input images. (d)-(f) represent the labeled results of noncooperative behavior. (g)-(i) show the labeled results for cooperative players without considering the saliency maps. (j)-(l) show the labeled results with considering the saliency maps.

3.2 Comparison

The parameter α and β, which represent the similarity and saliency threshold respectively, are the only two free parameters in our method. We usually set $\alpha = 0.5$, $\beta = 0.25$ in general condition, and $\alpha = 0.8$, $\beta = 0.5$ for high-variability images. Note that we use the default parameters to generate the results of [6] and [9]. Precisely, we use the sixth output image, i.e. the output of function $disp\_draw\_imgs\_clust\_cut$ of [9] for comparison.

Fig. 6 to Fig. 9 illustrate some results obtained by the proposed method on a set of images in different conditions. 1). The images with multiple common objects. 2). The images with similar backgrounds. 3). The common objects for different scales. 4). The images with complex backgrounds.

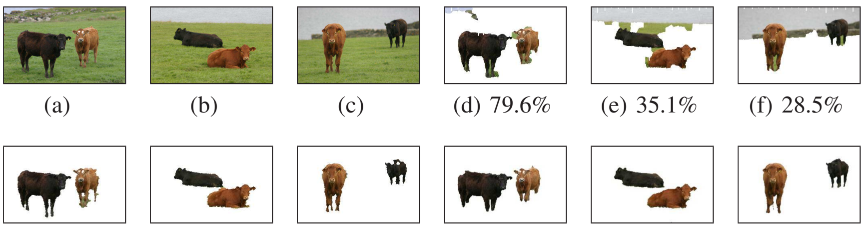
We first evaluated the proposed method on multiple common objects. The result is shown in Fig. 6. In Fig. 7, these sets of images (iCoseg dataset) are particularly difficult to segment due to the high similarity on the image background. Thanks to the saliency map, our result performs much better than [6] and [9]. Fig. 8 shows comparative results on MSRC dataset. Our method outperforms state-of-the-art co-segmentation methods [6] and [9]. When the objects among images are with different scales. We can observe that even if there is enormous size differences among objects, the proposed method still achieve high accuracy up to 93.7%. Another difficult problem of co-segmentation is shown in Fig. 9. There is only one common object (i.e. the yellow lemon) among images but backgrounds are complicated. As shown in the figure, both [6] and [9] cannot recognize the common object, lead to extremely low accuracy. On the contrary, our method produces the satisfactory results with more than 96% averaged segmentation accuracy in these images. For more general comparisons, Table 1 shows the comparative results on the iCoseg dataset. Since the dataset contains images with similar background, our method can reach higher average accuracy than [6] and [9].

4 Conclusion

We proposed a cooperative co-segmentation algorithm by using the concepts of cooperative game, heat diffusion, and image saliency. Our method takes advantage of a cooperative game model, which enables us to detect the common objects unsupervisedly and accurately. We treat images as players in the cooperative game model, and define the constrained utility function to promote the cooperation on the label estimation. After generating the labeled image, we apply the cooperative cut to precisely segment each labeled image independently. Compared to other co-segmentation methods, our method can solve those co-segmentation problems for images with similar or complex background, or images with objects of different scales or numbers. Experimental results demonstrate that our method outperforms the state-of-the-art co-segmentation algorithms.

Table 1. Co-segmentation results on the iCoseg dataset

iCoseg Dataset	Ours	Joulin [6]	Kim [9]
Elephants	83.6	19.0	50.2
Kite	75.0	29.2	47.1
Kite panda	85.0	37.9	46.1
Gymnastics1	90.9	47.0	41.5
Gymnastics2	83.9	39.2	41.6
Gymnastics3	86.4	51.8	59.0
Taj Mahal	76.0	30.4	28.4
Stonehenge	70.4	71.9	40.5
Liberty Statue	79.2	45.5	64.5
Skating	86.8	12.6	23.9
Livepool FC	78.2	40.7	36.5
Helicopter	79.6	55.1	6.2
mean accuracy	81.3	40.0	40.5



(a) (b) (c) (d) 79.6% (e) 35.1% (f) 28.5%

(g) 87.1% (h) 90.1% (i) 84.0% (j) 86.7% (k) 90.0% (l) 89.6%

Fig. 6. Multiple common objects. The percentage under each image denotes the segmentation accuracy. (a)-(c) are the input images. (d)-(f) show the co-segmentation of [6]. (g)-(i) show the co-segmentation of [9]. (j)-(l) show the proposed cooperative co-segmentation.

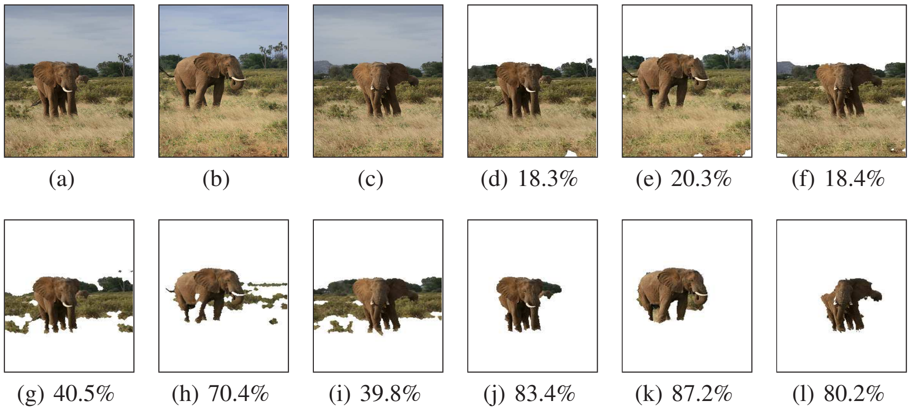


Fig. 7. Images with similar backgrounds. The percentage under each image denotes the segmentation accuracy. (a)-(c) are the input images. (d)-(f) show the co-segmentation of [6]. (g)-(i) show the co-segmentation of [9]. (j)-(l) show the proposed cooperative co-segmentation.

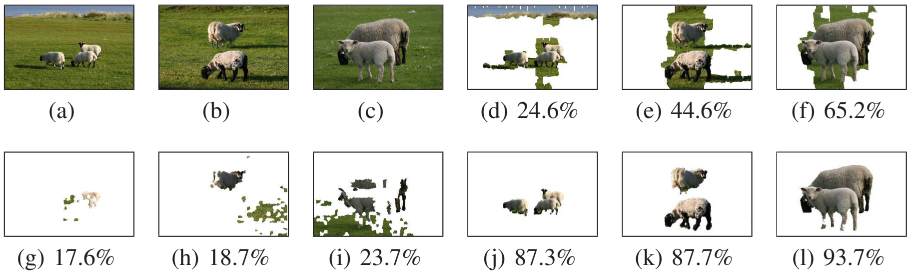


Fig. 8. Different-scale common objects. The percentage under each image denotes the segmentation accuracy. (a)-(c) are the input images. (d)-(f) show the co-segmentation of [6]. (g)-(i) show the co-segmentation of [9]. (j)-(l) show the proposed cooperative co-segmentation.

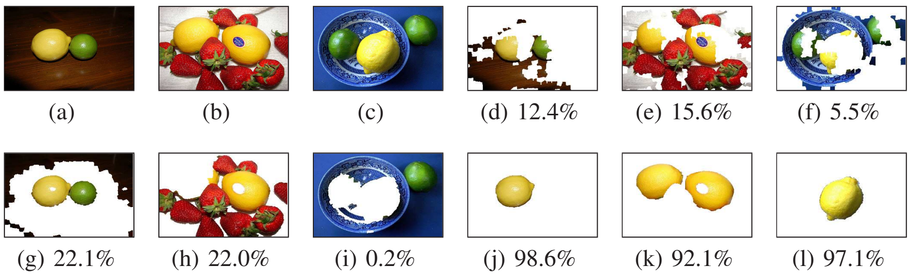


Fig. 9. Common objects with complicated backgrounds. The percentage under each image denotes the segmentation accuracy. (a)-(c) are the input images. (d)-(f) show the co-segmentation of [6]. (g)-(i) show the co-segmentation of [9]. (j)-(l) show the proposed cooperative co-segmentation.

References

1. Rother, C., Minka, T.P., Blake, A., Kolmogorov, V.: Cosegmentation of image pairs by histogram matching - incorporating a global constraint into mrfs. In: CVPR (1). (2006) 993–1000
2. Batra, D., Kowdle, A., Parikh, D., Luo, J., Chen, T.: icoseg: Interactive co-segmentation with intelligent scribble guidance. In: CVPR. (2010) 3169–3176
3. Batra, D., Kowdle, A., Parikh, D., Luo, J., Chen, T.: Interactively co-segmentating topically related images with intelligent scribble guidance. International Journal of Computer Vision **93** (2011) 273–292
4. Chu, W.S., Chen, C.P., Chen, C.S.: Momi-cosegmentation: Simultaneous segmentation of multiple objects among multiple images. In: ACCV (1). (2010) 355–368
5. Hochbaum, D.S., Singh, V.: An efficient algorithm for co-segmentation. In: ICCV. (2009) 269–276
6. Joulin, A., Bach, F.R., Ponce, J.: Discriminative clustering for image co-segmentation. In: CVPR. (2010) 1943–1950
7. Mukherjee, L., Singh, V., Dyer, C.R.: Half-integrality based algorithms for cosegmentation of images. In: CVPR. (2009) 2028–2035
8. Vicente, S., Kolmogorov, V., Rother, C.: Cosegmentation revisited: Models and optimization. In: ECCV (2). (2010) 465–479
9. Kim, G., Xing, E.P., Li, F.F., Kanade, T.: Distributed cosegmentation via submodular optimization on anisotropic diffusion. In: ICCV. (2011) 169–176
10. Zhang, J., Zheng, J., Cai, J.: A diffusion approach to seeded image segmentation. In: CVPR. (2010) 2125–2132
11. Bruhn, A., Weickert, J., Schnörr, C.: Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. International Journal of Computer Vision **61** (2005) 211–231
12. Harel, J., Koch, C., Perona, P.: Graph-based visual saliency. In: NIPS. (2006) 545–552
13. Osborne, M.: An introduction to game theory. Oxford Univ. Press (2004)
14. Chen, Y., Wang, B., Lin, W.S., Wu, Y., Liu, K.J.R.: Cooperative peer-to-peer streaming: An evolutionary game-theoretic approach. IEEE Trans. Circuits Syst. Video Techn. **20** (2010) 1346–1357
15. Hsiao, P.C., Chang, L.W.: Image denoising with dominant sets by a coalitional game approach. IEEE Transactions on Image Processing **22** (2013) 724–738
16. Jegelka, S., Bilmes, J.: Submodularity beyond submodular energies: Coupling edges in graph cuts. In: CVPR. (2011) 1897–1904
17. Levinshtein, A., Stere, A., Kutulakos, K.N., Fleet, D.J., Dickinson, S.J., Siddiqi, K.: Turbopixels: Fast superpixels using geometric flows. IEEE Trans. Pattern Anal. Mach. Intell. **31** (2009) 2290–2297
18. Murphy, K.P., Weiss, Y., Jordan, M.I.: Loopy belief propagation for approximate inference: An empirical study. In: UAI. (1999) 467–475
19. Rother, C., Kolmogorov, V., Blake, A.: ”grabcut”: interactive foreground extraction using iterated graph cuts. ACM Trans. Graph. **23** (2004) 309–314
20. Vicente, S., Kolmogorov, V., Rother, C.: Graph cut based image segmentation with connectivity priors. In: CVPR. (2008)
21. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. In: ICCV. (2001) 105–112
22. Winn, J.M., Criminisi, A., Minka, T.P.: Object categorization by learned universal visual dictionary. In: ICCV. (2005) 1800–1807
23. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Li, F.F.: Imagenet: A large-scale hierarchical image database. In: CVPR. (2009) 248–255