Information Sciences 521 (2020) 422–434

Contents lists available at ScienceDirect

Information Sciences

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

Extensible image object co-segmentation with sparse cooperative relations

Kunqian Li[a], Shengbo Qi[a], Hua Yang[b], Liqin Zhou[a,*], Dalei Song[a,c]

[a] College of Engineering, Ocean University of China, Qingdao 266100, China
[b] College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China
[c] Institute for Advanced Ocean Study, Ocean University of China, Qingdao 266100, China

ARTICLE INFO

Article history:
Received 19 April 2019
Revised 15 February 2020
Accepted 18 February 2020
Available online 19 February 2020

Keywords:
Image segmentation
Object co-segmentation
Graph theory
Extensible sparse cooperative graph

ABSTRACT

In the co-segmentation problems of sparse related image group, how to mine correlations between common objects has been one of the most significant steps. In this paper, we first prioritize the processing efficiency and construct sparse cooperative graph for related images in feature space, instead of examining the full correlations of common foregrounds which lead to high computation cost and great risk of false connection. And more importantly, we design a flexible and efficient extending strategy to reuse the cooperative graph for incremental image data. Then, with the sparse correlations of the image group, we propose a unilateral proposal co-selection method for pair-wise region level co-segmentation. Finally, we have achieved the segmentation results with further pixel level refinement. Our experimental results on publicly available datasets show that, compared with the approaches using dense cooperative constraints, the proposed method can achieve more competitive results with extremely sparse correlative constraints, which shows its bright application prospects for sparse correlated image groups with incremental data due to its high efficiency and flexible extensibility.

1. Introduction

As a basic problem in computer vision field, image segmentation plays an important role in high-level machine visual perception task [1–3]. Since a single image can only provide few information and vision patterns always contain the intrinsic ambiguity, segmentation problem of independent images is difficult [4,5]. In terms of the photographed subjects, images taken from similar scenes are always related actually. To cooperatively segment these related images, image co-segmentation has also become a popular research area in recent years [6].

To thoroughly investigate image co-segmentation for related images, we define the facts that vision units containing the same contents as implying cooperative relations or correspondences. Such cooperative relations allow us to utilize the corresponding information in mutual ways for better segmentation performance on each single image. Actually, the cooperative relations widely exist in image groups. The early studies of image co-segmentation focus on image pairs or image groups with dense correspondences. That is to say, image pairs in group are highly correlated in content by default. However, more and more researchers found that the cooperative relations are sparse in practice [7]. Take the images from FlickrMFC

* Corresponding author.
E-mail addresses: likunqian@ouc.edu.cn (K. Li), qsb@ouc.edu.cn (S. Qi), hyang@ouc.edu.cn (H. Yang), zlq@ouc.edu.cn (L. Zhou), songdalei@ouc.edu.cn (D. Song).

https://doi.org/10.1016/j.ins.2020.02.055
0020-0255/

Fig. 1. An example image group with sparse correlations. Images are from *Apple+picking* group of FilckrMFC dataset [7].

dataset[1] as an example. As Fig. 1 presented, even though all the images are taken from the same scene, each image only captures parts of all the targets. For this case, performing cooperative constraint on each image pair is obviously unreasonable.

For image object co-segmentation problem, mining appropriate correlations is the key step [8] and the correlative graph for an image group is always sparse. False correlation constraints may even degrade the co-segmentation performance. Moreover, for a relatively larger image group, where efficiency is preferred, performing cooperative constraint on a very sparse correlation graph will be much more appropriate.

Predictably, with rapid popularization of consumer digital camera, the size of visual data has become larger and larger. Another concern of image co-segmentation is how to make the existing cooperative graph extensible. With established cooperative graph and the new related data, we would like to directly attach the new data to the graph rather than reconstruct the whole graph.

With all the above concerns, we propose a new image object co-segmentation method, which can handle sparse correlated images efficiently with high extensibility for new image data. We further demonstrate that our method can deliver more competitive segmentation results even with very sparse constraint. In summary, our contributions are as follows:

1. We introduce the Sparse Cooperative Graph (SCG) to sparsely and effectively represent the content correlations of the given images.
2. We design a flexible graph attaching approach to make the SCG extensible for incremental images, which improves the processing efficiency for new images.
3. For the object co-segmentation task, we design a multi-level co-segmentation framework, which takes full use of the pair-wise sparse cooperative relations and performs flexible co-segmentation operation on both region and pixel levels.

We organize the rest of the paper as follows. We first present a review of the related literatures in Section 2, and then problem formulation and the overview of the proposed algorithm are described in Section 3. The proposed extensible sparse cooperative graph and its constructing method are introduced in Section 4. In Section 5, we introduce an image object co-segmentation method with the sparse cooperative relations. The experimental results on public datasets and qualitative/quantitative analysis are given in Section 6. Finally, we conclude our paper in Section 7.

2. Related works

Image co-segmentation problem was first raised by Rother et al. [9] in 2006. In his work, co-segmentation was defined as simultaneously segmenting two images which contain the same target. From then on, image co-segmentation problem has been given a much broader definition. By reviewing the developments of the co-segmentation problem definition, we find that it follows the changes of correlative assumptions for image group. In image-pair co-segmentation problems, as Fig. 2(a) presented, it assumes that the common foregrounds are identical targets in different background. Accordingly, most

[1] http://vision.snu.ac.kr/~gunhee/software/mfc/FlickrMFCdataset.zip.

Fig. 2. Images from typical co-segmentation datasets. (a) image pair dataset; (b) iCoseg dataset; (c) MSRC dataset.

co-segmentation methods for this problem force the foreground histogram to be coherent and jointly solve the labelling problem [9,10]. iCoseg dataset collected by Batra et al. [11] further promotes the development of co-segmentation.

iCoseg dataset is much larger, which contains 643 images from 38 groups. From Fig. 2(b), it can be noticed that the common foreground targets are much more complicated. For this dataset, the correlative assumption is that a group of images are taken from the similar scenes and particularly the common foreground targets have similar appearance. However, the variations of target pose, illumination and integrity exist in almost all the image groups. Obviously, this problem is much more challenging.

It is worth noting that the appearance variation of foreground targets from pair-images and iCoseg dataset are still very limited. As presented in Fig. 2, more significant intra-category variation shows up in MSRC-v2 dataset [12]. Here the correlative assumption turns to be that the images contain common targets from the same semantic category. Significant appearance variation makes this task much more difficult. Many researchers view this task as a multi-class co-segmentation problem [13,14].

To address the co-segmentation problem implied in the above co-segmentation datasets, many classic approaches were proposed. Segmentation models for single image were extended to adapt to co-segmentation formulation, such as MRF/CRF [9,10,15–18], active contour model [19,20], random walk model [21,22], et al. Besides, new models and strategies also achieved good performance, such as interactive methods [11,23,24], saliency/co-saliency based methods [15,25,26], proposal selection based methods [8,27–30] and Convolutional Neural Networks (CNN) based methods [31–33].

However, for pair-images, iCoseg, MSRC dataset and the instance co-segmentation dataset defined by [33], there exists a common hidden assumption, i.e., all the images contain the same category targets. This assumption is the foundation of constructing dense cooperative relations. Even though dense cooperative relations provide a wealth of mutual common information, its computational cost is very large and it is unsuitable for extensible image groups. On the other hand, for images taken from the real associated scenes, such as the images of FlickrMFC dataset (Fig. 1), common targets do not appear in every single image. Kim et al. [7] defined this task as a Multiple Foreground Cosegmentation (MFC) problem, whose correlative assumption is much more relaxed.

The early MFC algorithms mainly focus on mining the cooperative relations and discovering the consistency. Kim et al. [7] formulated the algorithm as an alternative optimizing procedure, which performs foreground modelling and region assignment iteratively. This method aims to first discover highly repeated object regions and then establish discriminate models for different object. Meng et al. [34,35] proposed to generate cooperative prior via directed graph clustering and then propagate it for foreground maps. Li et al. [36] proposed to use an ensemble clustering scheme for the discovery of object-like proposals, which will be used to derive unary and pairwise potentials across images for α-expansion optimization. Chang et al. [37] proposed to first separate the MFC problem into three sub-tasks, i.e. image segmentation, segment matching, and figure/ground assignment. Then, the final co-segmentation result will be achieved by a foreground object discovering algorithm. However, modelling or clustering the common targets with significant appearance variations is truly difficult. And the mixed model for multiple instances of the same target in different images may even degrade the descriptive ability. More specifically, for weak cooperative relations, such as several independently related image pairs in a large image group, performing clustering is unrealistic. Therefore, digging the most informative pair-wise cooperative relations seems to be a more efficient and universally applicable strategy.

Besides, Liu et al. [38] proposed to construct a fully linked 3D co-diffusion graph for target correlations. Liu et al. [39] proposed to use deformable part model detection to generate potential foreground maps, then generate refined maps for co-segmentation by further introducing cooperative terms and nonrigid mapping terms. In nonrigid mapping step, it needs extra training source pools for neighbouring regions retrieving. Besides, like most other co-segmentation approaches, the dense correspondence for related images makes algorithm more time-consuming to reconstruct the correlations when new images comes. It will be better to make the correlations extensible for incremental data.

Jointly consider the above concerns, we believe that the promising developing trend of co-segmentation algorithm should be improving extensibility and efficiency. In the following sections, we will introduce a new co-segmentation method which takes both the two demands into account and also receives good segmenting performance.

3. Problem formulation and algorithm overview

Given a group of related images $G_{img} = \{I_1, I_2, \ldots, I_N\}$, whose size N could be still increasing. Specially, the foreground objects do not appear in every single image. The object co-segmentation problem aims to generate the foreground segmenta-

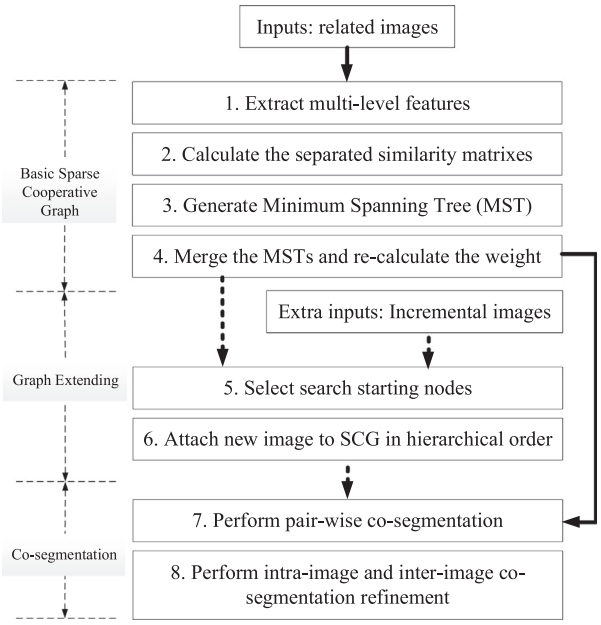

Fig. 3. The overall diagram of the proposed co-segmentation method. The graph extending (steps denoted with dash arrows) is optional, which will be executed when there are incremental images need to be segmented.

tion $G_{seg} = \{S_1, S_2, \ldots, S_N\}$ cooperatively within G_{img}. Considering the sparsity of the image correlations and the compatibility demand for incremental images, we design the Extensible Sparse Cooperative Graph (ESCG) for related images connection and propose a new co-segmentation method accordingly. The overview of our algorithm is given in Fig. 3. We decompose the extensible co-segmentation task into three components, i.e., constructing the basic sparse cooperative graph, extending graph with incremental images and performing the co-segmentation according to the extended sparse cooperative graph.

In the first component, we propose to construct sparse cooperative graph to link the limited visual data, which reflects the most reliable cooperative relations. In the second component, to get the most close correlative information for the extra data with low computational cost, we propose to hierarchically search the attaching node to extend the basic sparse graph. This component is designed for incremental data, which is optional in real applications. In the third component, we follow the correlation connections to perform co-segmentation and multi-level refinement. The whole algorithm covers both correlation mining and co-segmentation application, which provides a general framework for the cooperative processing of sparse correlated data. In the following sections, we will give more detailed description for each module.

4. Extensible sparse cooperative graph for images

For completely disordered visual data, the correlations are sparse and unknown beforehand. Discover the implied sparse correlations properly is the key step for cooperative processing. To this end, we propose a simple but efficient Sparse Cooperative Graph (SCG) constructing method. More importantly, we creatively introduce the extensibility into the SCG, which makes the new-coming data can fully utilize the already established graph.

4.1. Basic sparse cooperative graph

For $G_{img} = \{I_1, I_2, \ldots, I_N\}$, we define the sparse cooperative graph

$$M_s = \{m_{i,j}\}, 1 \le i, j \le N \tag{1}$$

as a sparse matrix, whose element $m_{i,j}$ is the correlation of image pair $\{I_i, I_j\}$. To give a comprehensive measurement of the image correlations, we first extract the deep feature from the final convolutional layer of pre-trained VGG19 model [40] and the Bag-of-Feature (BoF) [41] for each image. Particularly, we use max pooling operation on each dimension of deep feature, which finally generates a 512-D feature. Deep feature and BoF feature describe the image in different levels. Deep feature gives the high-level global semantic description and BoF feature describes the low-level visual elements. For each feature description, we calculate the cosine similarity of every image pair. The similarity matrix of deep feature and BoF are denoted as M_d and M_b, respectively. Taking deep feature similarity for example, to construct a sparse but closely related graph, we

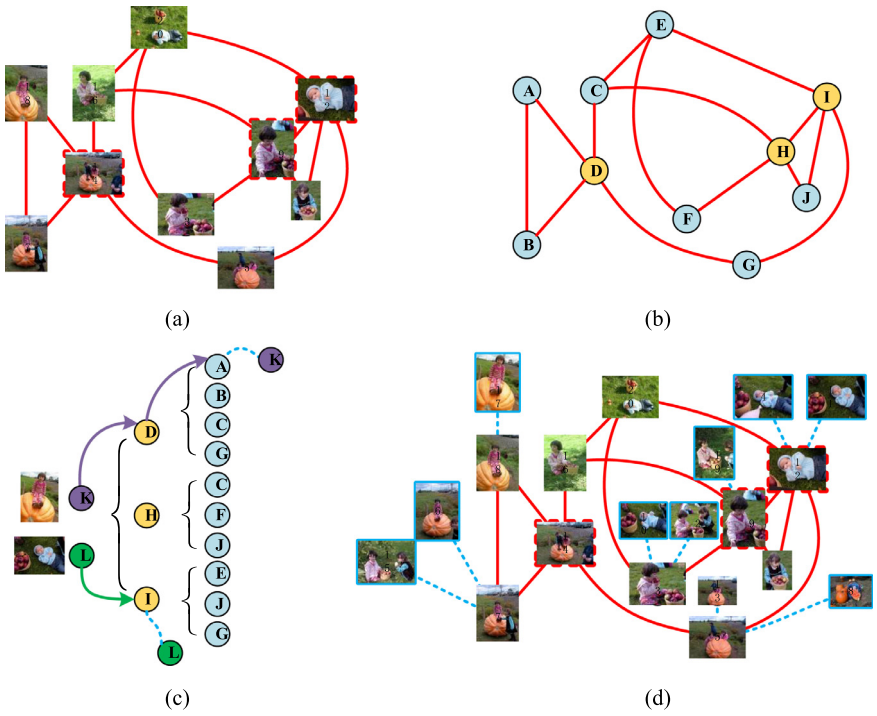

Fig. 4. An example of the extensible sparse cooperative graph. (a) the basic sparse cooperative graph, the images with red dash boundary are the start searching nodes for graph extending; (b) the graph skeleton of (a); (c) the extending procedure of hierarchical graph extending; (d) the extended sparse cooperative graph with incremental image data (images with blue boundary).

can formulate a binary optimizing problem as follows,

$$\begin{aligned} &\boldsymbol{x}_d = \arg\max \sum_{1\le i,j\le N} x_{ij}^d \cdot M_d(i,j), \\ &\text{s.t.} \quad \forall i, \sum_{j=1}^{N} x_{ij}^d \ge 1, \sum_{i=1}^{N}\sum_{j=1}^{N} x_{ij}^d = N-1, \end{aligned} \tag{2}$$

where $x_{ij}^d \in \boldsymbol{x}_d$ is the binary variable to optimize and $x_{ij}^d = 1$ means to keep the edge linking image i and image j. The constraint condition means that we only keep necessary number of edges to make all the images linked. This optimizing problem can be solved with the Minimum Spanning Tree (MST) algorithm [42] by transforming deep feature similarity matrix M_d into distance matrix $D_d = 1 - M_d$, accordingly. Then, the form of objective function turns into

$$\boldsymbol{x}_d = \arg\min \sum_{1\le i,j\le N} x_{ij}^d \cdot D_d(i,j). \tag{3}$$

Analogously, the sparse correlative graph with BoF measurement can be constructed by solving $x_{ij}^b \in \boldsymbol{x}_b$ in the same way.

To construct comprehensive sparse cooperative graph for the visual data, we merge the two MSTs into a single sparse graph as follows

$$\begin{aligned} &M_s = \left\{m_{i,j}\right\}, 1 \le i, j \le N, \\ &m_{i,j} = M_s(i,j) = \left\{\max\left(x_{ij}^d \cdot M_d(i,j), x_{ij}^b \cdot M_b(i,j)\right)\right\}. \end{aligned} \tag{4}$$

In Fig. 4(a), we give an example of the sparse graph, where the example images are from the *apple+picking* group of FlickrMFC dataset. It can be noticed that the sparse graph presents the correlation well, and accordingly the common foregrounds can be provided with proper constraints in the following cooperative processing steps.

4.2. Hierarchical graph extending for incremental data

As Fig. 4(a) shown, the basic sparse graph has established a relatively complete structure by fully measuring the relations of the input visual data. However, in many real-world applications, the data size is very large and is even increasing. Obviously, starting from scratch to reconstruct the graph is very inefficient. In this section, we combine the features of the sparse graph and design a hierarchical graph extending strategy. This extensible sparse cooperative graph will provide more flexibility for the incremental data.

We denote the incremental images as $G_{inc} = \{\tilde{I}_1, \tilde{I}_2, \ldots, \tilde{I}_{\tilde{N}}\}$. We plan to attach the new images to our basic graph, i.e. extending the $N \times N$ graph M_s to a $(N+\tilde{N}) \times (N+\tilde{N})$ sparse graph $\tilde{M}_s$. An intuitive attaching strategy is to traverse the whole basic graph and select the most closely related node (image). The efficiency of this strategy is obviously very low. Considering the specificity of graph structure, we design a hierarchical graph extending strategy by searching the attaching point in a multi-branch searching form.

In Fig. 4(b) and (c), we give two graph extending examples. For clarity, we simplify the Fig. 4(a) as a skeleton graph and the images are alphabetically denoted with nodes. The key motivation of our design is to fully utilize the graph structure and lighten the computational cost. Here, we first select the searching starting nodes $P = \{p_1, p_2, \ldots, p_k\}$, which satisfies

$$\#\left\{\bigcup_{i=1}^{k} C(p_i)\right\} \geq T_c \cdot N, \tag{5}$$

where $C(p_i)$ returns the linking nodes set of p_i, $\#(\cdot)$ returns the element number of $C(p_i)$ and T_c is the element covering threshold. We expect that the chosen start nodes can link most of the graph nodes, i.e. cover at least $T_c \cdot N$ nodes.

The designed hierarchical multi-branch searching is presented in Fig. 4(c). The searching starts from the starting nodes $P = \{p_1, p_2, \ldots, p_k\}$. We first chose the current best attaching node $p_a \in P$ which satisfies

$$\chi(p_0, p_a) = \min_{p_i \in P} \chi(p_0, p_i), \tag{6}$$

where $\chi(p_i, p_j)$ measures the image distance of p_i and p_j in feature space. Here, we equally weight and combine the distance of both deep feature and BoF feature. Secondly, with p_a as the current attaching node, we iteratively traverse its linked nodes to find the next attaching node which is more related with the new image node. We define a new operation

$$\phi(p_0, p_a) = \min\left(\min_{p_j \in C(p_a)} \chi\left(p_0, p_j\right), \chi(p_0, p_a)\right) \tag{7}$$

as the generalized distance between the new node p_0 and the current attaching node p_a. Accordingly, the update criterion for the new attaching node is as follows

$$p_a = \begin{cases} p_a, & \text{if } \phi(p_0, p_a) = \chi(p_0, p_a), \\ p_{a'}, & \text{if } \phi(p_0, p_a) = \chi(p_0, p_{a'} \in C(p_a)). \end{cases} \tag{8}$$

When it satisfies the first condition, the searching will be ended automatically. In Fig. 4(c), we give two examples of the searching and attaching procedure. The branch structure shows all the searching paths and the arrows show the changing of attaching node. Take node "L" for example, our strategy only needs comparing 6 times. Compared with the traversing strategy, the computation cost of our method decreases by 40 percent. For larger sparse graph, this advantage will be more significant.

5. Co-segmentation via sparse cooperative relations

5.1. Region level based unilateral proposal co-selection

With the sparse cooperative graph, we can correlate the visual data properly. To fully utilize the correlated information and achieve the co-segmentation of the linked images, we design a multi-level co-segmentation strategy based on the pair-wise relations of the sparse graph. We perform the co-segmentation in both object region level and pixel level.

Given the sparse cooperative graph M_s, for $M_s(i, j) \neq 0$, which means I_i and I_j are closely correlated, we first perform object region level co-segmentation by optimizing the following objective function

$$\begin{aligned} &\left(\boldsymbol{y}_i, \boldsymbol{y}_j\right) = \arg\max\left(\boldsymbol{y}_i R_{ij}^{\mathrm{T}} + \boldsymbol{y}_j R_{ij}^{\mathrm{T}}\right), \\ &\text{s.t.} \quad \forall v, \sum_{u=1}^{n_i} \boldsymbol{y}_i(u, v) \leq 1, \sum_{v=1}^{n_j} \boldsymbol{y}_i(u, v) = T_{cnt}(i, j); \\ &\qquad \forall u, \sum_{v=1}^{n_j} \boldsymbol{y}_j(u, v) \leq 1, \sum_{u=1}^{n_i} \boldsymbol{y}_j(u, v) = T_{cnt}(i, j). \end{aligned} \tag{9}$$

Optimizing this objective function aims to select the proposals which are both outstanding and correlated with each other. Here, R_{ij} is the region proposal correlative matrix of image I_i and I_j, whose element is defined as

$$R_{ij}(u, v) = \gamma(r_u) + \gamma(r_v) + \varphi(r_u, r_v). \tag{10}$$

$r_u \in I_i, u = 1, 2, \ldots, n_i$ and $r_v \in I_j, v = 1, 2, \ldots, n_j$ are the region proposals generated with the method introduced in [43], $\gamma(r)$ is the object ranking score of proposal r. $\varphi(r_u, r_v)$ is the similarity between the two proposals, which is comprehensively measured according to their color, texture, BoF and deep feature. The variables $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ we are going to optimize are the proposal selection label matrixes for image I_i and I_j, which have the same size with R_{ij}.

As formulated in Eq. (9), we design the region level co-segmentation as a unilateral common proposal co-selection problem. To select proper common proposals, we specially design the above constraint conditions. Firstly, we set unilateral once-only selection strategy to select the most similar regions in the pair-correlated image. The unilateral once-only selection means we set every image as the reference image in turn and the regions from the reference image can be only selected once. Particularly, we separately optimize $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$ by selecting I_i and I_j as reference image, respectively. When select I_i as the reference image, $\boldsymbol{y}_i$ will be optimized and its element $\boldsymbol{y}_i(u, v) = 1$ means proposal $r_u \in I_i$ and $r_v \in I_j$ are co-selected as foregrounds. This unilateral once-only selection makes our method to select more proper proposals. Secondly, we automatically determine the co-selection time of each unilateral setting. The time of co-selection $T_{cnt}(i, j)$ is intuitively determined by the degree of their correlation distance $\chi(I_i, I_j)$, i.e.,

$$T_{cnt}(i, j) = \left\lceil 1/\left(-\log\left(1 - \chi\left(I_i, I_j\right)\right)\right)\right\rceil. \tag{11}$$

If the images are more related, accordingly we will allow more proposal pairs to be selected as reward.

5.2. Pixel level based co-segmentation refinement

By merging the co-selected proposal regions, we can get the region level co-segmentation result with proposal co-selection strategy. However, the object proposals are usually coarse, then how to further refine the co-segmentation results in pixel level is another important problem. In our method, we perform both intra-image and inter-image co-segmentation refinement to fully utilize the shared common information.

Let $\boldsymbol{S} = \{S_1, S_2, \ldots, S_{\tilde{N}}\}$ be the region level co-segmentation results, we first perform intra-image refinement within the GrabCut framework [44]. For each image I_i, we take $S_i \in \boldsymbol{S}$ as the soft initialization for foreground/background modelling and perform the refinement twice. Specially, in the first refinement, we fix the image edge region as the hard background for robust segmentation. Here, we denote the intra-image refinement results as $\boldsymbol{S}' = \{S_1', S_2', \ldots, S_{\tilde{N}}'\}$. The intra-image refinement can be viewed as the multi-targets co-segmentation within the single image, which can discover more targets and improve the segmentation details.

In the inter-image refinement, we would like to further utilize the shared common information of two correlated images for reliable inference. Therefore, in the inter-image refinement procedure, we will combine the two images for further refinement. For the linked image pair $\{(i, j)|\tilde{M}_s(i, j) = 1\}$, we take $S_i \bigcap S_i'$ and $S_j \bigcap S_j'$ as the initial foreground modelling maps. This setting aims to provide the refinement with more reliable foreground prior information. After resizing and splicing the two images as a single one, we utilize the same refine strategy for intra-image.

Responding to the trait of graph connecting structure, for the nodes which have multiple connections, we choose label voting strategy for multiple refined maps merging. In the voting step, for each pixel, we choose the major label in multiple maps as the final label, which forms the final segmentation of the proposed method.

6. Experiments and analysis

To verify the performance of the proposed method, we compare our method with other typical co-segmentation approaches on publicly available datasets, i.e. FlickrMFC dataset [7], iCoseg dataset[2] [11] and MSRC dataset[3] [12]. In this section, after a brief description for the parameter settings, both qualitative and quantitative comparison on publicly available datasets will be presented. Besides, we will give necessary analysis and discussion for each comparison and algorithm complexity.

6.1. Parameter setting

The parameter settings are fixed throughout the experiments on different datasets. Some parameters have been introduced above, and for clarity, here we give a comprehensive summary for those parameters. In our experiment, for each image group which consists about 20-30 images, we evenly select 10 images as the initial data for basic SCG construction and the others are viewed as the incremental images. In the graph extending step for incremental data, the covering threshold T_c is fixed to 0.8. In co-segmentation step, we select the top 10 proposals of each image as the co-selection candidate pool. The similarity of proposals $\varphi(r_u, r_v)$ is measured with color, texture, BoF and deep feature, whose weight coefficients are 0.1, 0.1, 0.4 and 0.4, respectively.

6.2. Experiment on image groups with sparse correlations

We first test our method on FlickrMFC dataset. FlickrMFC has 14 image groups and the images in the same group are collected in the related scenes. As we mentioned above, FlickrMFC is a challenging dataset since the correlation of images is relatively sparse, how to construct such sparse correlations effectively and utilize the shared information are of great

[2] http://chenlab.ece.cornell.edu/projects/touch-coseg/.
[3] http://research.microsoft.com/en-us/projects/ObjectClassRecognition/.

Table 1
Quantitative comparisons on FlickrMFC Dataset [7]. Jaccard score is considered, a larger value indicates a better performance. The best in each row is shown in red and the second best is shown in blue.

	[7]	[37]	[39]	[35]	[14]	[36]	[38]	Ours
apple+picking	0.657	0.713	0.661	**0.779**	0.470	**0.740**	0.527	0.735
baseball+kids	0.401	0.421	0.655	0.609	0.500	**0.680**	**0.690**	0.606
butterfly+blossom	0.406	0.722	0.641	0.723	0.550	0.710	**0.735**	**0.736**
cheetah+safari	0.337	0.585	**0.683**	0.568	0.620	0.550	0.626	**0.725**
cow+pasture	0.549	**0.648**	0.586	0.577	0.390	0.630	0.401	**0.677**
dog+park	0.375	**0.594**	0.570	0.528	0.540	0.510	0.456	**0.641**
dolphin+aquarium	0.341	0.506	**0.619**	0.544	0.610	0.530	0.616	**0.635**
fishing+alaska	0.304	0.404	0.450	**0.471**	0.470	0.350	0.390	**0.518**
gorilla+zoo	0.319	0.451	**0.609**	0.399	0.480	0.450	0.343	**0.525**
liberty+statue	0.427	0.530	0.563	0.520	**0.580**	**0.580**	0.370	0.506
parrot+zoo	0.436	0.467	**0.590**	0.485	0.540	0.550	0.563	**0.567**
stonehenge	0.531	**0.773**	0.476	0.466	0.550	0.700	0.683	**0.702**
swan+zoo	0.312	0.514	0.504	0.497	0.470	0.450	**0.565**	**0.581**
thinker+Rodin	0.425	0.474	0.642	0.491	**0.690**	0.550	**0.760**	0.643
Averages	0.416	0.557	**0.589**	0.547	0.533	0.570	0.552	**0.628**

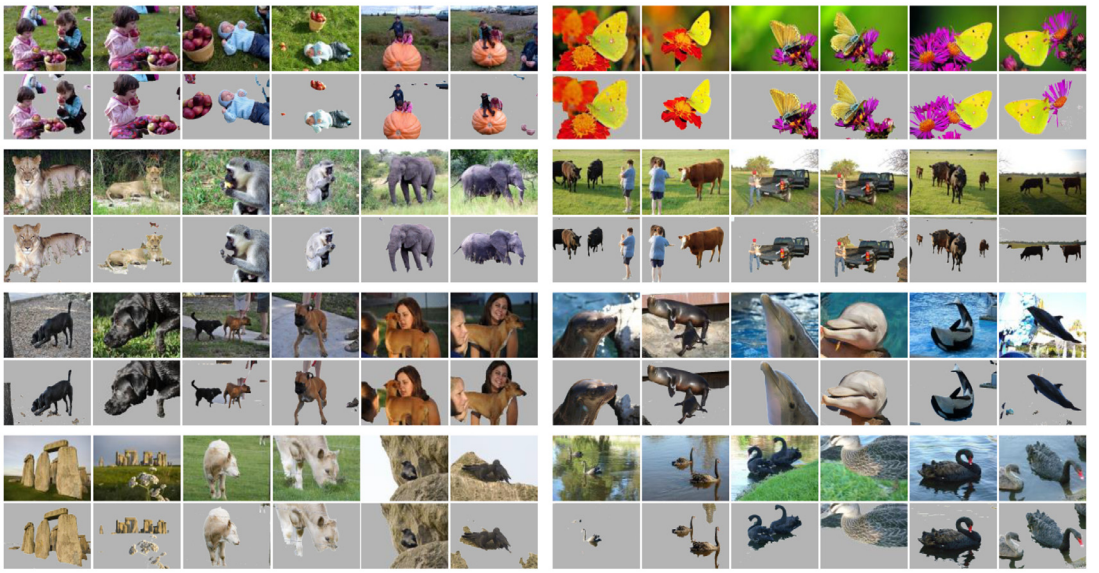

Fig. 5. The co-segmentation results of the proposed method on FlickrMFC dataset.

importance. We compare our method with typical multiple foreground co-segmentation algorithms [7,14,35–39], which are designed or suitable for multiple foreground co-segmentation problems with sparse correlation. Therefore, this direct performance comparison on FlickrMFC is appropriate.

We test our method on all the 14 image groups and present the quantitative evaluation with the Jaccard score in Table 1. For fair comparison, we select the evaluation results reported on their literatures or re-evaluate them with projects provided by the authors. It can be noticed that from Table 1, our method outperforms all the comparison methods in terms of average evaluation, and it achieves top 2 performance on most image groups. The segmentation examples of the proposed method on FlickrMFC dataset are presented in Fig. 5. The visualization results also show that our method performs well on FlickrMFC dataset. Such competitive performance benefits from the effective discovery and the efficient utilization of the limited sparse correlations. Besides, the designed region level unilateral proposal co-selection strategy and pixel-level refinement method jointly promote the segmentation results.

To further verify the performance of the proposed method, we give the comparison of the module performance in Fig. 6. Performance comparison after co-selection, intra-image refinement and inter-image refinement are presented. After unilateral proposal co-selection, we achieve 0.560 average performance, which is only behind the best competitor by 2.8% and outperforms all the others. This result demonstrates the effectiveness and good performance of our unilateral proposal co-selection strategy. With intra-image refinement, the performance has an apparent promotion and this is because the

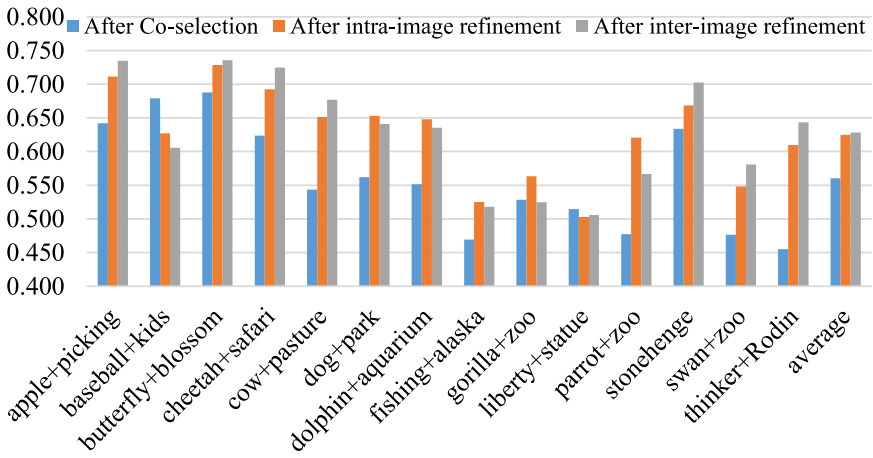

Fig. 6. The performance with intra-image and inter-image refinement.

intra-image refinement can repair the poor foreground proposals and discover the missing common targets with the foreground models. As to the inter-image refinement, the segmentation of 8 image groups receive further improvement, which demonstrates the effectiveness of the cross-refinement in pixel level.

A noteworthy phenomenon is that our approach is based on very sparse correlations, i.e. the completeness of the correlation mining is not our pursuit, but the competitors aim to fully discover the correlations with clustering strategies [7,35,36,39] or partial consistency matching/linking strategies [14,37,38]. It means that we discover and utilize less correlative information but achieve better performance. By analysing the different segmentation strategies and performance, we find that in many cases pair-wise co-segmentation constraints are good enough for segmenting the common foregrounds. Introducing insignificant co-operative information or inappropriately applying co-operative constraint may bring more ambiguity and also reduce the efficiency. Actually, in the real application, how to efficiently construct the necessary sparse connections for large dataset is an attractive problem. The proposed extensible sparse cooperative graph provides a new strategy to address this problem.

6.3. Experiment on image groups with dense correlations

In last section, we test our method on image groups with sparse correlations. For most traditional co-segmentation problems, the correlations are always dense, i.e. every image contains common foreground. On image groups with dense correlations, we would like to further compare our method with typical approaches which try to utilize complete cooperative correlations, so as to verify that whether the sparse cooperative relations can bring enough abundant constraints for good segmentation performance. As we mentioned above, iCoseg and MSRC are typical co-segmentation datasets including image groups with dense correlations. MSRC is more complicated and difficult to segment since more intra-category variations are involved. Particularly, for fair comparison, co-segmentation approaches for multiple foregrounds [7,13,35,37] are presented for quantitative comparison on iCoseg dataset. On MSRC dataset, considering its particularity, we compare our method with co-segmentation methods using invariant feature [28,46,47] or multi-class co-segmentation approaches [13,14,38,45]. The quantitative comparison results on iCoseg and MSRC are given in Tables 2 and 3, respectively. Visualization results are presented in Figs. 7 and 8, respectively.

From Tables 2 and 3, we can find that on most image groups, our method achieves top 2 performance among all the approaches. It demonstrates that our approach built on very sparse image correlations can also receive good performance on dense related images. More importantly, our co-segmentation strategy with sparse correlations are more suitable for the cases where the common foregrounds have significant variation. Take MSRC dataset for example, the objects belonging to the same semantic category are viewed as the common foregrounds. Obviously, the intra-class variation brings difficulties and confusions to the common foreground modelling and the following constraining steps. An intuitive observation is that an image contributes the most to its nearest neighbours in the co-segmentation problem, where the common foregrounds have more close feature description. Actually, besides providing sufficient correlative information, the sparse cooperative graph also avoids introducing those complicated distracting information brought by intra-class variation. It can explain that why our method performs well on MSRC dataset. It is worth noting that the proposed method is only slightly behind the 3D co-diffusion method [38], which focus on constructing complete sparse correlation with a global co-diffusion framework. However, the efficiency and extensibility of [38] are extremely limited since the co-diffusion is a global operation on all the images, which can be hardly applied to large dataset. It shows bright prospect that our method achieves close performance to [38] with efficient extensibility.

Table 2
Quantitative comparisons on iCoseg Dataset [11]. Jaccard score is considered, a larger value indicates a better performance. The best in each row is shown in red and the second best is shown in blue.

	[7]	[13]	[37]	[35]	Ours		[7]	[13]	[37]	[35]	Ours
Alaskan	0.231	0.453	**0.539**	0.534	**0.548**	Red Sox	0.511	0.541	**0.698**	0.508	**0.704**
Stonehenge-1	0.161	0.386	**0.856**	0.378	**0.870**	Stonehenge-2	0.712	0.581	**0.784**	0.538	**0.944**
Liverpool	0.279	0.413	**0.583**	0.435	**0.532**	Ferrari	0.539	0.527	**0.780**	**0.614**	0.607
AgraTaj-1	0.392	**0.734**	**0.475**	0.469	0.415	AgraTaj-2	0.233	0.161	0.269	**0.305**	**0.408**
Pyramids	0.143	**0.590**	0.500	0.397	**0.609**	Elephants	0.248	0.454	**0.599**	0.400	**0.497**
Goose	0.613	0.697	**0.913**	0.547	**0.799**	Pandas-1	0.419	0.658	**0.813**	0.710	**0.771**
Helicopter	0.191	0.283	**0.816**	**0.765**	0.757	Planes	0.148	0.223	**0.483**	0.304	**0.599**
Airshows	0.150	0.086	0.315	**0.449**	**0.388**	Cheetah	0.372	0.541	0.692	**0.769**	**0.828**
Pandas-2	0.556	0.396	0.403	**0.641**	**0.685**	Brighton	0.367	0.519	**0.883**	0.723	**0.744**
Kitekid	0.406	0.223	0.302	**0.496**	**0.699**	Margate	0.348	0.678	**0.821**	0.571	**0.877**
Colt	**0.492**	0.282	**0.728**	0.475	0.457	Gymnastic-1	**0.569**	0.307	0.350	**0.594**	0.550
Gymnastic-2	0.312	0.234	0.503	**0.512**	**0.608**	Gymnastic-3	0.399	0.525	0.225	**0.728**	**0.797**
Skating-1	0.213	0.610	0.453	**0.772**	**0.909**	Skating-2	0.424	0.488	**0.917**	0.660	**0.750**
Soccer-Red	0.592	0.396	0.588	**0.614**	**0.661**	Soccer-White	0.065	**0.514**	**0.541**	0.484	**0.514**
Monks	0.504	0.649	**0.769**	0.726	**0.745**	Balloons	0.468	0.228	**0.972**	0.557	**0.639**
Liberty	0.484	0.860	0.798	**0.926**	**0.930**	Christ	0.378	0.324	**0.789**	0.764	**0.824**
Speedskating	0.186	0.112	**0.237**	0.202	**0.288**	Track	0.153	0.169	0.063	**0.422**	**0.567**
Windmill	0.334	0.262	**0.801**	0.523	**0.566**	Kendo-1	0.297	0.755	**0.931**	0.762	**0.847**
Kendo-2	0.485	0.844	**0.927**	0.717	**0.908**	brown bear	0.405	0.764	**0.817**	0.562	**0.901**
						Averages	*0.363*	*0.460*	***0.630***	*0.567*	***0.678***

Table 3
Quantitative comparisons on MSRC-v2 Dataset [12]. Jaccard score is considered, a larger value indicates a better performance. The best in each row is shown in red and the second best is shown in blue.

	[45]	[13]	[46]	[14]	[47]	[38]	[28]	Ours
Bike	0.299	0.433	0.511	0.512	**0.539**	**0.606**	0.323	0.383
Bird	0.299	0.477	0.512	0.557	0.455	**0.795**	0.581	**0.611**
Car	0.371	0.597	0.637	**0.729**	0.571	**0.672**	0.472	0.593
Cat	0.244	0.319	0.610	**0.659**	0.553	0.550	0.576	**0.637**
Chair	0.287	0.396	**0.561**	0.465	0.544	0.440	0.417	**0.559**
Cow	0.335	0.527	**0.699**	0.684	0.663	0.661	0.683	**0.771**
Dog	0.330	0.418	**0.638**	0.558	0.479	**0.679**	0.491	0.637
Face	0.332	**0.700**	0.556	0.609	0.538	0.569	0.524	**0.577**
Flower	0.402	0.519	0.688	0.672	0.621	**0.748**	0.647	**0.728**
House	0.322	0.510	**0.708**	0.566	0.618	0.691	0.413	**0.746**
Plane	0.251	0.216	0.465	**0.522**	0.480	0.321	0.466	**0.525**
Sheep	0.608	0.663	0.752	0.722	0.701	0.768	0.702	**0.774**
Sign	0.432	0.589	0.733	0.591	0.651	**0.835**	0.691	**0.784**
Tree	0.612	0.670	**0.743**	0.620	0.702	**0.804**	0.593	0.651
Averages	0.366	0.502	0.630	0.605	0.580	**0.653**	0.537	**0.641**

Another noteworthy issue is that [28] is the first co-segmentation work which utilizes sparse correlations. It places proposals of each image in a column where only proposals in adjacent columns are linked with correlation weights. It creatively solves the co-segmentation problem by selecting the shortest path which is consists of the most similar proposals. Since the sparse correlation is unknown beforehand, such graph constructing strategy has great blindness which will inevitably hurt the performance. By contrast, our method avoids such blindness with MST searching and hierarchically graph extending to some extent. From Table 3, we can see that our method shows obvious performance advantage when compared with [28].

6.4. Discussion

6.4.1. Time complexity analysis

As we mentioned above, by introducing extensible sparse cooperative graph, our algorithm aims to pursuit high efficiency and flexibility for image group with large incremental data while keeping competitive segmentation performance. In previous subsections, the experimental comparison on public datasets demonstrates the competitive performance. In this subsection, we will give a discussion on the time complexity of the proposed method.

Fig. 7. The co-segmentation results of the proposed method on iCoseg dataset.

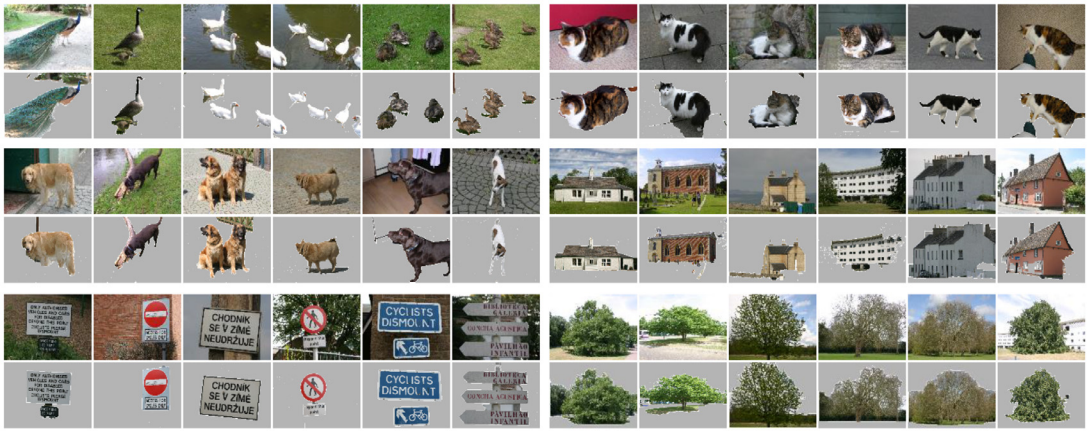

Fig. 8. The co-segmentation results of the proposed method on MSRC dataset.

In our algorithm setting, the image group we are going to co-segment is consist of two parts, i.e., the basic images $G_{img} = \{I_1, I_2, \ldots, I_N\}$ and the incremental images $G_{inc} = \{\tilde{I}_1, \tilde{I}_2, \ldots, \tilde{I}_{\tilde{N}}\}$. Our algorithm includes four main steps for this co-segmentation task: construction of the basic sparse cooperative graph, co-segmentation for the basic images, sparse cooperative graph extending and co-segmentation for the incremental images. For the construction of basic sparse cooperative graph, the submodule time complexity is $O(N^2)$ since we should compute dense pair-wise similarity for MST construction. While the time complexity of co-segmentation for the basic images is $O(N)$ since only sparse pair-wise co-segment operations need to be conducted. The proposed hierarchical graph extending strategy provides efficient correlation method for incremental images, whose time complexity is $O\big(\tilde{N}\log(N)\big)$. Accordingly, the time complexity of co-segmentation for incremental images is $O(\tilde{N})$. Therefore, the total time complexity of the proposed method is

$$O\big(N^2 + N + \tilde{N}\log(N) + \tilde{N}\big) = O\big(N^2 + \tilde{N}\log(N)\big). \tag{12}$$

It can be easily noticed that given a fixed basic sparse cooperative graph, the additional computation caused by incremental images grows linearly, which is very efficient in co-segmentation problems. Especially for a sizable basic sparse cooperative graph which has been well constructed, our hierarchically extending strategy can help the new data to find close correlations efficiently.

6.4.2. Limitation and future works

In the proposed method, we construct the basic sparse cooperative graph with only image-level features, i.e., the deep feature extracted from the pre-trained deep model and BoF. Considering that every single image may contain different

multiple objects, image-level features may not be detailed enough for these cases, we suggest to introduce instance-level correlations in the construction of the sparse cooperative graph in the future works.

Additionally, the time complexity for basic sparse cooperative graph construction on image set $G_{img} = \{I_1, I_2, \ldots, I_N\}$ is $O(N^2)$, which is still inefficient for a large N. An improved strategy is to selectively pick out the typical targets from G_{img} to construct a compact but powerful basic sparse cooperative graph beforehand.

7. Conclusion

This paper focuses on the co-segmentation task of extensible image set with sparse relations, which has never been completely explored. Our solution to construct a sparse cooperative graph provides necessary consistency constraints and makes it reusable for incremental data. Rather than repeatedly constructing the sparse cooperative graph for new related images, performing the proposed hierarchical graph extending is much more efficient. Given the established basic sparse graph, the additional computation brought by incremental images grows linearly. Besides, by comparing the proposed method with full correlation constraints based co-segmentation algorithms, the introduced extremely sparse correlation is demonstrated to be able to provide effective enough constraints for the co-segmentation task. In the future works, introducing instance-level correlations into the construction of basic sparse cooperative graph will help to give more detailed cooperative constraints. Additionally, proposing a more efficient graph construction strategy for a relative large basic image set is also a promising topic for the further improvement.

Declaration of competing interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

CRediT authorship contribution statement

Kunqian Li: Conceptualization, Methodology, Software, Writing - original draft. **Shengbo Qi:** Resources, Writing - review & editing. **Hua Yang:** Writing - original draft, Writing - review & editing. **Liqin Zhou:** Conceptualization, Project administration, Validation, Writing - review & editing. **Dalei Song:** Supervision, Writing - review & editing.

Acknowledgment

The research has been supported by the National Natural Science Foundation of China under Grant 61906177, in part by the Natural Science Foundation of Shandong Province under Grant ZR2019BF034, in part by Fundamental Research Funds for the Central Universities under Grants 201813022 and 201964013.

References

[1] J. Miao, T.-Z. Huang, X. Zhou, Y. Wang, J. Liu, Image segmentation based on an active contour model of partial image restoration with local cosine fitting energy, Inform. Sci. 447 (2018) 52–71.
[2] S. Han, W. Tao, X. Wu, Texture segmentation using independent-scale component-wise riemannian-covariance gaussian mixture model in kl measure based multi-scale nonlinear structure tensor space, Patt. Recog. 44 (3) (2011) 503–518, doi:10.1016/j.patcog.2010.09.006.
[3] W. Tao, H. Jin, J. Liu, Unified mean shift segmentation and graph region merging algorithm for infrared ship target segmentation, Opt. Eng. 46 (12) (2007) 1–7, doi:10.1117/1.2823159.
[4] W. Tao, F. Chang, L. Liu, H. Jin, T. Wang, Interactively multiphase image segmentation based on variational formulation and graph cuts, Patt. Recog. 43 (10) (2010) 3208–3218, doi:10.1016/j.patcog.2010.04.014.
[5] W. Tao, H. Jin, Y. Zhang, Color image segmentation based on mean shift and normalized cuts, IEEE Trans. Syst. Man Cybernet. Part B 37 (5) (2007) 1382–1389.
[6] H. Zhu, F. Meng, J. Cai, S. Lu, Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation, J. Vis. Commun. Image Represent. 34 (2016) 12–27.
[7] G. Kim, E.P. Xing, On multiple foreground cosegmentation, in: IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 837–844.
[8] K. Li, J. Zhang, W. Tao, Unsupervised co-segmentation for indefinite number of common foreground objects, IEEE Trans. Image Process. 25 (4) (2016) 1898–1909.
[9] C. Rother, T. Minka, A. Blake, V. Kolmogorov, Cosegmentation of image pairs by histogram matching-incorporating a global constraint into mrfs, in: IEEE Conference on Computer Vision and Pattern Recognition, 1, 2006, pp. 993–1000.
[10] S. Vicente, V. Kolmogorov, C. Rother, Cosegmentation revisited: models and optimization, in: Proc. Eur. Conf. Computer Vision, 6312, 2010, pp. 465–479.
[11] D. Batra, A. Kowdle, D. Parikh, J. Luo, T. Chen, icoseg: interactive co-segmentation with intelligent scribble guidance, in: IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 3169–3176.
[12] J. Shotton, J. Winn, C. Rother, A. Criminisi, Textonboost: joint appearance, shape and context modeling for multi-class object recognition and segmentation, in: European Conference on Computer Vision, 2006, pp. 1–15.
[13] A. Joulin, F. Bach, J. Ponce, Multi-class cosegmentation, in: IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 542–549.
[14] F. Wang, Q. Huang, M. Ovsjanikov, L. Guibas, Unsupervised multi-class joint image segmentation, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2014, pp. 3142–3149.
[15] K.-Y. Chang, T.-L. Liu, S.-H. Lai, From co-saliency to co-segmentation: An efficient and fully unsupervised energy minimization model, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, Providence, RI, USA, 2011, pp. 2129–2136.
[16] W. Tao, K. Li, K. Sun, Sacoseg: Object cosegmentation by shape conformability, IEEE Trans. Image Process. 24 (3) (2015) 943–955.
[17] L. Liu, W. Tao, H. Liu, Complementary saliency driven co-segmentation with region searching and hierarchical constraint, Inform. Sci. 372 (2016) 72–83.
[18] H. Zhu, J. Lu, J. Cai, J. Zheng, S. Lu, N.M. Thalmann, Multiple human identification and cosegmentation: A human-oriented crf approach with poselets, IEEE Trans. Multim. 18 (8) (2016) 1516–1530.

[19] F. Meng, H. Li, G. Liu, K.N. Ngan, Image cosegmentation by incorporating color reward strategy and active contour model, IEEE Trans. Cybernet. 43 (2) (2013) 725–737.
[20] F. Meng, H. Li, G. Liu, Image co-segmentation via active contours, in: IEEE International Symposium on Circuits and Systems, 2012, pp. 2773–2776.
[21] M. Collins, J. Xu, L. Grady, V. Singh, Random walks based multi-image segmentation: Quasiconvexity results and gpu-based solutions, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2012, pp. 1656–1663.
[22] C. Lee, W. Jang, J. Sim, C. Kim, Multiple random walkers and their application to image cosegmentation, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2015, pp. 3837–3845.
[23] X. Dong, J. Shen, L. Shao, M. Yang, Interactive cosegmentation using global and local energy optimization, IEEE Trans. Image Process. 24 (11) (2015) 3966–3977.
[24] W. Wang, J. Shen, Higher-order image co-segmentation, IEEE Trans. Multim. 18 (6) (2016) 1011–1021.
[25] H. Li, F. Meng, K.N. Ngan, Co-salient object detection from multiple images, IEEE Trans. Multim. 15 (8) (2013) 1896–1909.
[26] K.R. Jerripothula, J. Cai, J. Yuan, Image co-segmentation via saliency co-fusion, IEEE Trans. Multim. 18 (9) (2016) 1896–1909.
[27] S. Vicente, C. Rother, V. Kolmogorov, Object cosegmentation, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, Providence, RI, USA, 2011, pp. 2217–2224.
[28] F. Meng, H. Li, G. Liu, K.N. Ngan, Object co-segmentation based on shortest path algorithm and saliency model, Multim. IEEE Trans. on 14 (5) (2012) 1429–1441.
[29] F. Meng, H. Li, K.N. Ngan, L. Zeng, Q. Wu, Feature adaptive co-segmentation by complexity awareness, Image Process. IEEE Trans. on 22 (12) (2013) 4809–4824.
[30] A. Faktor, M. Irani, Co-segmentation by composition, in: 2013 IEEE International Conference on Computer Vision, 2013, pp. 1297–1304.
[31] Z. Yuan, T. Lu, Y. Wu, Deep-dense conditional random fields for object co-segmentation, in: Proceedings of the International Joint Conference on Artificial Intelligence, 2017, pp. 3371–3377.
[32] K.-J. Hsu, Y.-Y. Lin, Y.-Y. Chuang, Co-attention cnns for unsupervised object co-segmentation, in: Proceedings of the International Joint Conference on Artificial Intelligence, 2018, pp. 748–756.
[33] K.-J. Hsu, Y.-Y. Lin, Y.-Y. Chuang, Deepco3: Deep instance co-segmentation by co-peak search and co-saliency detection, in: Computer Vision and Pattern Recognition, IEEE Conference on, 2019.
[34] F. Meng, B. Luo, C. Huang, Object co-segmentation based on directed graph clustering, in: IEEE Visual Communications and Image Processing, 2013, pp. 1–5.
[35] F. Meng, H. Li, S. Zhu, B. Luo, C. Huang, B. Zeng, M. Gabbouj, Constrained directed graph clustering and segmentation propagation for multiple foregrounds cosegmentation, IEEE Trans. Circ. Syst. Video Technol. 25 (11) (2015) 1735–1748.
[36] H. Li, F. Meng, Q. Wu, B. Luo, Unsupervised multiclass region cosegmentation via ensemble clustering and energy minimization, IEEE Trans. Circ. Syst. Video Technol. 24 (5) (2014) 789–801.
[37] H.-S. Chang, Y.-C.F. Wang, Optimizing the decomposition for multiple foreground cosegmentation, Comput. Vis. Image Understand. 141 (2015) 18–27.
[38] L. Liu, K. Li, X. Liao, Image co-segmentation by co-diffusion, Circ. Syst. Signal Process. 36 (11) (2017) 4423–4440.
[39] Z. Liu, J. Zhu, J. Bu, C. Chen, Object cosegmentation by nonrigid mapping, Neurocomputing 135 (2014) 107–116.
[40] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: International Conference on Learning Representations, 2015.
[41] L. Fei-Fei, P. Perona, A bayesian hierarchical model for learning natural scene categories, in: Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2, 2005, pp. 524–531.
[42] M. Held, R.M. Karp, The traveling-salesman problem and minimum spanning trees, Oper. Res. 18 (6) (1970) 1138–1162.
[43] I. Endres, D. Hoiem, Category-independent object proposals with diverse ranking, Patt. Analy. Mach. Intelli. IEEE Trans. on 36 (2) (2014) 222–234.
[44] C. Rother, V. Kolmogorov, A. Blake, Grabcut : interactive foreground extraction using iterated graph cuts, ACM Trans. Graph 23 (2004) 309–314.
[45] G. Kim, E. Xing, L. Fei-Fei, T. Kanade, Distributed cosegmentation via submodular optimization on anisotropic diffusion, in: Proc. Int. Conf. Computer Vision, Barcelona, Spain, 2011, pp. 169–176.
[46] J. Dai, Y.N. Wu, J. Zhou, S. Zhu, Cosegmentation and cosketch by unsupervised learning, in: International Conference on Computer Vision, 2013, pp. 1305–1312, doi:10.1109/ICCV.2013.165.
[47] Y. Li, J. Liu, Z. Li, H. Lu, S. Ma, Object co-segmentation via salient and common regions discovery, Neurocomputing 172 (2016) 225–234.