J. Vis. Commun. Image R. 33 (2015) 20–30

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

A site entropy rate and degree centrality based algorithm for image co-segmentation ☆

Souradeep Chakraborty [a,*], Pabitra Mitra [b]

[a] Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur 721302, India
[b] Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur 721302, India

ARTICLE INFO

Article history:
Received 24 April 2015
Accepted 31 August 2015
Available online 7 September 2015

Keywords:
Site entropy rate
Degree centrality
Co-segmentation
Co-saliency
Markov chain
Object proposal
k-partite graph
Stationary distribution

ABSTRACT

In this paper, we propose a graph based algorithm that efficiently segments common objects from multiple images. We first generate a number of object proposals from each image. Then, an undirected graph is constructed based on proposal similarities and co-saliency maps. Two different methods are followed to extract the proposals containing common objects. They are: (1) degree centrality of nodes obtained after graph thresholding and (2) site entropy rate of nodes calculated on the stationary distribution of Markov chain constructed on the graph. Finally, we obtain the co-segmented image region by selecting the more salient of the object proposals obtained by the two methods, for each image. Multiple instances of the common object are also segmented efficiently. The proposed method has been compared with many existing co-segmentation methods on three standard co-segmentation datasets. Experimental results show its effectiveness in co-segmentation, with larger IoU values as compared to other co-segmentation methods.

1. Introduction

Co-segmentation has been an active research topic in the area of image processing. Many practical object segmentation methods are based on generating object priors through human interaction [1,2]. Users are here asked to provide segmentation cues manually [3,4]. When the number of target images is high, users face a huge workload of providing manual segmentation cues. The principle of co-segmentation is to exploit the availability of multiple images that contain instances of the same "object" classes to supplement detailed supervisory information. This reduces the user workload significantly. As opposed to single image segmentation, co-segmentation aggregates information from multiple images (which contain objects with similar features) to improve the segmentation of individual images. Co-segmentation methods which can handle large numbers of images and object classes find potential applications in many fields such as automated image retrieval, object tracking, and object recognition.

Many of the existing methods for co-segmentation are modeled on the Markov random field (MRF)-based optimization procedures [7,9–13] and other graph theoretic methods [19,36]. The MRF based methods mostly formulate energy functions based on foreground or background consistency constraints and optimize such functions to obtain the segmentations. The graph based methods model image regions as nodes of a graph (which represent object proposals) and then perform graph processing operations such as node clustering and shortest path finding, to extract the strongly connected nodes representing co-segmented image regions. However, these methods do not utilize the essential entropy information furnished from the graphs, such as the rate entropy of the stationary distribution on a Markov chain constructed on the graph, or the edge-weight threshold entropy information. Furthermore, if there are a large number of original images in an image group, the problem becomes more expensive to compute. Development of fast algorithms which avoid time complex optimization procedures as in MRF-based methods, as well as utilize entropy information in constructed graphs, still remains a challenge.

In this paper, we present a simple and effective co-segmentation model, which integrates the notion of degree centrality of nodes, and their site entropy rate information (obtained from the stationary distribution of the constructed Markov chain), to co-segment multiple similar images. The proposed model consists of four main steps. The first step is to segment the original images into a number of local semantic regions, which is achieved by applying object proposal generation algorithm as described in [5]. Out of the several object proposals suggested by the algorithm, we select only those

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author.
E-mail addresses: sourachakra@gmail.com (S. Chakraborty), pabitra@gmail.com (P. Mitra).

http://dx.doi.org/10.1016/j.jvcir.2015.08.016
1047-3203/

proposals whose mean saliency value is greater than the Otsu threshold calculated on the respective images [6]. In the second step, we construct a graph to represent the local region similarities according to the feature distance and the co-saliency maps generated by using the method in [7]. As the third step, we apply two graph based algorithms on the constructed graph to extract the object proposal (which contains the common object) from each image in the image group. They are: (i) Degree centrality based node selection and (ii) Site entropy rate based node selection. Finally, we select the object proposal with maximum mean saliency value among the two object proposals computed by the two graph based methods, as the co-segmented image region for an image. We evaluate our method on many groups of images. The experimental results demonstrate the effectiveness of our method.

The remainder of this paper is organized as follows. A brief review of the related works in the field of co-segmentation is given in Section 2. We explain the proposed method in Section 3, describing the two sub-methods: site entropy rate based and degree centrality based, in details. Experimental results are provided in Section 4 to support the efficiency of the proposed algorithm. Finally, in Section 5, we draw conclusions with future research issues.

2. Related work

The task of co-segmentation was first introduced by Rother et al. [8], where co-segmentation was modeled as an optimization problem in which a Markov random field (MRF)-based method was proposed to extract the objects from image pairs by adding the constraint of foreground similarity (measured by L1-norm) to traditional MRF-based procedures. Trust region graph cuts (TRGC) method was employed for energy function optimization. Many other methods were proposed [9–12] following this MRF-based optimization framework. Mukherjee et al. replaced the L1-norm by L2-norm in [9]. Pseudo-Boolean optimization method was used for the energy function optimization. Hochbaum and Singh [10] rewarded foreground similarity instead of penalizing foreground difference and that simplified the energy function optimization. In [13], Vicente et al. extended the foreground similarity measurement by employing dual decomposition for the energy function optimization. Chang et al. [12] used the graph-cut algorithm to optimize a global energy term which considered both foreground similarity and background consistency.

Several other co-segmentation models were proposed apart from MRF-based methods. Joulin et al. [14] combined discriminative clustering and spectral clustering methods to perform co-segmentation of multiple classes and for a significantly large number of images. To exploit priors about image more directly, an interactive co-segmentation method was proposed by Batra et al. in [15], which segments common objects through human interaction guided by an automatic recommendation system. Mukherjee et al. [16] put forward a scale-invariant method of co-segmentation with the requirement that the rank of the matrix corresponding to foreground regions should equal one. Vicente et al. [17] proposed a model which emphasizes interesting objects co-segmentation by selecting useful features from a total of 33 features through random forest regressor. In [18], Kimet al. followed a distributed co-segmentation approach via sub-modular optimization on anisotropic diffusion for a highly variable large-scale image collection. Meng et al. [19] designed a digraph to represent the local region similarities according to the feature distance and the saliency map, and formulated the co-segmentation problem as a shortest path problem. In this paper, we adopt a similar graph based approach. Unlike the layered digraph as followed in [19], we construct a k-partite graph and then implement two methods which make use of entropy information: degree centrality based and site entropy rate based, and then select the object proposal which most accurately segments the common object in the group.

Rubio et al. [20] proposed a multiple-scale multiple-image generative model, which jointly estimated the foreground and background appearance distributions from many images. Meng et al. [21] proposed a model which integrates active contours method and rewarding strategy. They generate a new energy function with two conflicting goals: foreground similarity among the images and background consistency in each image, and then use a mutual evolution approach to minimize the energy function value. In a more recent method of Tao et al. [22], object co-segmentation method based on shape conformability is put forward. It focuses on the shape consistency of the foreground objects in image set. The common shape pattern is extracted if the foreground objects are varied in appearance but share similar shape structures.

There have been many works which have utilized saliency information in the process of segmentation as well as co-segmentation. [23] used saliency to automate the selection of foreground object and background seeds, needed for image segmentation. In a similar work [24], a co-saliency prior has been used as a hint about possible foreground locations for image co-segmentation task. Besides image segmentation, saliency/co-saliency has also been utilized in similar and related applications including image classification [25], ranking [26] and de-blurring [27].

3. Proposed co-segmentation method

The flowchart of the main steps of the proposed method is shown in Fig. 1. It consists of four major steps. In subsequent sections, we describe the object proposal generation method, the approach of graph construction, and the graph based co-segmentation methods used to extract common object from each image.

3.1. Object proposal generation and salient proposal selection

We explain the steps followed to generate object proposals and co-saliency maps for each image, followed by subsequent salient object proposal selection.

Step 1: The object proposal generation procedure in [5] is implemented to segment the original image into a number of local regions by object proposal generation. In this approach, both local and global search procedures are combined in the space of sets of superpixels, to obtain accurate segmentations for all objects of an image. Assume that $I = \{I_1, I_2, \ldots, I_m\}$ denotes the original image set, of size m. We first segment each image I_i into a set of overlapping objects proposals $P^i = \left\{P^i_1, \ldots, P^i_{N_i}\right\}$, where N_i is the number of the object proposals in image I_i. The set of all initial objects proposals is denoted as $P_{initial} = \left\{P^1_{initial}, \ldots, P^m_{initial}\right\}$.

Step 2: To extract the salient object proposals from the set of proposals obtained (as described in Step 1), we first generate co-saliency maps of all the images in an image set by the self-adaptively weighted co-saliency detection technique recently proposed method by Cao et al. [7]. This method exploits the relationship of multiple saliency cues and obtains the self-adaptive weight to generate the co-saliency maps. In our experiment, we obtain saliency maps from the methods in [28–30], and get co-saliency maps using this self-adaptively weighted co-saliency detection method.

Step 3: After the co-saliency maps are obtained, we select only those object proposals whose mean co-saliency value is greater than a threshold. Following [19], for an object proposal P^i_j from an image I_i, we measure its mean co-saliency value s^i_j as:

$$s^i_j = \left(\frac{\sum_{(k,l)\in P^i_j} S_i(k,l)}{m^i_j}\right) \cdot \left(\frac{m'_{ij}}{M'_i}\right), \tag{1}$$