VIDEO CO-SEGMENTATION FOR MULTIPLE FOREGROUND OBJECTS FROM MULTIPLE VIDEOS

Mrs. M.PUSHPALATHA, Assistant Professor, PG & Research Department of Computer Science & Computer Applications, Padmavani Arts & Science College for Women, Salem-11

C.VANMATHI, M.Phil, Research Scholar, Department of Computer Science & Computer Applications, Padmavani Arts & Science College for Women, Salem-11

ABSTRACT

We present a video co-segmentation method that uses non-categorized object proposals as a base element and can extract multiple foreground objects in a video collection. The use of feature objects overcomes limitations of low-level representations when separating complex foregrounds and backgrounds. We devise co-segmentation in which foreground-like properties are taken, as well as accounting for intra-video and inter-video foreground coherence. To treat several objects in the foreground, we extend the model of the co-selection graph to a proposed multi-state selection graph model (MSG), which optimizes the segmentation of different objects together. Not only can this extension into the MSG be applied to our co-selection diagram, but it can also be used to turn any standard graph model into a multi-state selection solution that can be directly optimized by existing energy minimization techniques. Our experiments show that our multi-foreground video co-segmentation method compares well with related techniques in both single and various foreground cases.

Keywords: *Video co-segmentation, Multistate selection graph, multiple foreground object extraction, energy minimization.*

1. INTRODUCTION

The goal of video foreground segmentation is common extract the main object from a series of videos. In contrast to the unsupervised foreground segmentation problem for a single video [16]. In this article we present a general technique for video Co-segmentation that is formulated with object suggestions as the basic element of processing, and that can easily work around single or multiple foreground objects individually or in multiple videos. Our multiple foreground video co-segmentation method for extraction of foreground objects is being developed by two most important technical contributions. The first is an object-based framework in which a co-selection graph is constructed to connect each foreground candidate in multiple videos. The foreground candidates in each frame are category independent Object suggestions that probably represent regions include an object after structured learning Method of [6]. The second technical contribution is a method of enlargement the graph models like the selection mentioned above diagram to allow selection of multiple states in each node. In the context of video co-segmentation, we turn to each other this method extends the co-selection diagram to a multistate Selection chart (MSG) in the multiple foreground objects can be handled in our object-based framework. The MSG is additionally able to handle the cases of a single foreground and / or a single video and can be optimized through existing energy minimization techniques. Our method yields results that exceed co-segmentation related techniques.

2. RELATED WORK

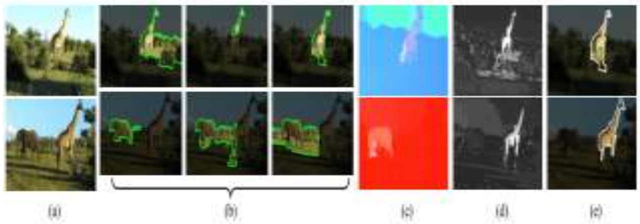
Video Co-segmentation: Only a few methods have been proposed for video co-segmentation, and they all base their processing on low-level features. Chenetal. [2] identified regions with coherent motion in the videos and then find a common foreground based on similar chroma and texture feature distributions. Rubioetal. [19] presented an iterative process for foreground/background separation based on feature matching among video frame regions and spatiotemporal tubes. The low-level appearance models in these methods, however, are often not discriminative enough to accurately distinguish complex foregrounds and backgrounds. Guoetal. [9] employed trajectory co-saliency to match the action form the video pair. However, this method only focuses on the common action extraction rather than the foreground object segmentation. In [3], the Bag-ofWords representation was used within a multi-class video co-segmentation method based on distant-dependent Chinese Restaurant Processes. While BoW features provide more discriminative ability than basic color and texture features, they may not be robust to appearance variations of a foreground object in different videos, due to factors such as pose change. Fig. Object-based Segmentation: In contrast to the methods based on low-level descriptors, object-based techniques make use of a mid-level representation that aims to delineate an object's entirety. Vicenteetal. [21] introduced the use of object proposals for co-segmentation of images. Mengetal. [17] employed the shortest path algorithm to select a common foreground from object proposals in multiple images. Leeetal. [14] utilized object proposal regions as foreground candidates in the context of single video segmentation, with the objectness measure used in ranking foreground hypotheses. More recent works [16], approach and incorporated a common constraint that the foreground should appear in every frame. This constraint is formulated within a weighted graph model, with the solution optimized via maximum weight cliques [16], shortest path algorithm [22], or dynamic programming [23]. As these single video segmentation methods do not address the co-segmentation problem, they do not account for the information within other videos. Moreover, they do not present a way to deal with multiple foreground objects. In our work, we present a more general co-selection graph to formulate correspondences between different videos, and extend this Framework to handle both single and multiple foreground objects using the MSG model. Multiple foreground co-segmentation: Some co-segmentation methods can handle multiple objects. Kimetal. [11] employed an anisotropic diffusion method to find out multiple object classes from multiple images. They also presented a different approach for multiple foreground co-segmentation in images [12], which builds on an iterative framework that alternates between foreground modelling and region assignment. Joulinetal. [10] proposed an energy-based image co-segmentation method that combines spectral and discriminative clustering terms. Mukherjeeet al. [18] segmented multiple objects from image collections, by analyzing and exploiting their shared subspace structure. The video co-segmentation method in [3] can also deal with multiple foreground extraction, which uses a nonparametric Bayesian model to learn a global appearance model that connects the segments of the same class. However, all of these methods are based on low-level feature representations for clustering the foregrounds into classes. On the other hand, object-based techniques operate on a midlevel representation of object proposals but lack an effective way to deal with multiple foregrounds. In our work, we extend the object-based co-segmentation approach to handle multiple foregrounds using the MSG model, where multiple foreground objects can be segmented jointly in multiple videos via the existing energy minimization method.

3. OUR METHOD

We present our object-based video co-segmentation algorithm by describing it first in the case of a single foreground Object, and then handle this approach multiple foreground objects with the MSG model.

3.1. Single Object Co-segmentation

We refer to the amount of videos as fV g, where everyone Video V n consists of Tn1; :::; V Frames labeled f FNG. In each frame, a set of object candidates is obtained to use the category-independent object suggestion method [6] from which the generated candidates may originate some overlapping areas. Identify the foreground object in each frame we consider different object properties s, which point to foregrounds, taking into account intra-video coherence of foreground and foreground Coherence between the different videos. We formulate this problem as a selection graphic in the form of a Conditional Random Field (CRF). As in Coward. Corresponding to him. By concatenating the selected candidates all frames of the video set we get a candidate series u = fu nt jn = 1; :::; N; t = 1; :::; Tnnt G. For every video intra-video edges are placed between the nodes of adjacent ones frame. The nodes of different videos are fully connected with each other through inter-video edges.



For this co-selection graph, we express its energy function

Ecs E (u) as follows:

$$E_{cs}(\mathbf{u}) = \sum_{n=1}^{N}\sum_{t=1}^{T_n}\left[\Psi(u_t^n) + \Phi_\alpha(u_t^n, u_{t+1}^n)\right] + \sum_{\substack{n,m=1,\\ n\neq m}}^{N}\;\sum_{t=1}^{T_n}\sum_{s=1}^{T_m}\Phi_\beta(u_t^n, u_s^m), \qquad (1)$$

Each frame of a video is a node, and the foreground object candidates of the frame are the states a node can take. The nodes (frames) from different videos are fully connected by inter-video terms. Within a given video, only adjacent nodes (frames) are connected by intra-video terms.

In contrast to the directed graph used in our co selection graph is a cycle graph that connects candidates among multiple videos. Optimizing a cycle graph is a NP hard problem. We employ TRW-S [13] to obtain a good approximated solution as in [5]. Since object candidates generated by [6] are only roughly segmented, we refine the results as in [14] with a pixel-level spatiotemporal graph-based segmentation.

3.2. Multiple foreground co-segmentation

We extend our single object video co-segmentation approach to handle multiple foregrounds using a multi-state selection graph model (MSG). With MSG, multiple foregrounds can be solved jointly in the multiple videos via existing energy minimization methods.

3.2.1. Multiple foreground selection energy

For the case of multiple foregrounds, K different candidates are to be found in each frame. We refer to the set of selected candidates throughout the videos for the k foreground object as the candidate series u (k). In solving for the multiple foreground co-segmentation, we account for the independent co-segmentation energies Ecs (u) of each of the K candidate series. In addition, it must be ensured that the K candidate regions have minimal overlap throughout the videos, since an area in a video frame cannot belong to two or more foreground objects.

3.2.2. Multi-state selection graph model

To optimize the multiple foreground selection energy ,we propose the multi-state selection graph model (MSG). In MSG, the co-selection graph for single object co-segmentation is replicated K1 times to produce K sub graphs in total, one for each candidate series. We observe that the candidate overlap penalty can be treated as edges between corresponding nodes in the sub graphs.

4. EXPERIMENTS

The proposed method is general enough to handle single/multiple videos and single/multiple foreground segmentation. In our experiments, we test our method in the two video co-segmentation cases, with a single foreground and with multiple foregrounds. We employ two metrics for the evaluation. The first is the average per-frame pixel error [20]

4.1. Single foreground video co-segmentation

In evaluating for the single foreground case, we employ the MOViCS dataset [3], which includes four video sets in total with five frames of each video labeled with the ground truth. The foregrounds in these video sets are taken to be the primary objects, namely the Chicken, Giraffe, Lion and Tiger. Using the codes obtained from the corresponding authors, we compare our ObMiC algorithm to six state-of-the-art methods that are the most closely related works published in recent years:

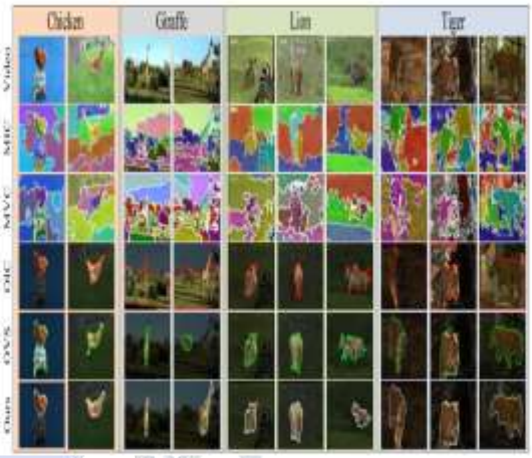


Figure 2. Single object segmentation results on the MOViCS dataset, where the displayed video frames are from different videos. From top to bottom: input videos, MIC [10], MVC [3], OIC [17], OVS [23], and our ObMiC method.

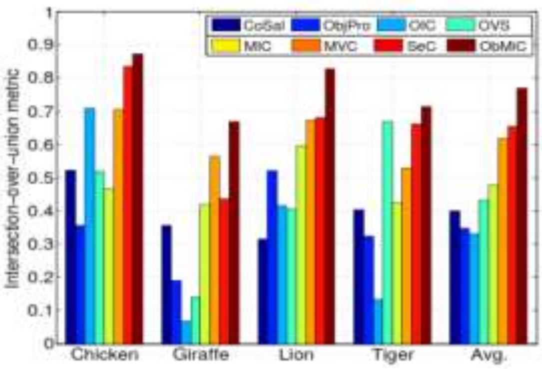


Figure 3. The intersection-over-union metric on MOViCS dataset.

4.2. Multiple foreground video co-segmentation

Since there are no datasets for multiple foreground video co-segmentation, we have collected our own, consisting of four sets, each with a video pair and two foreground objects in common. The dataset includes ground truth manually obtained for each frame. With these videos, we compare our method to two multi-class co-segmentation methods: MIC [10] and MVC [3]. It also classifies pixels based on a low-level representation without an objectness constraint, which may result in wrongly merged object classes from the foreground and background. For example, the black dog in the first video of the Dog set is wrongly classi fied together with the background tree shadows in the second video. Also, for the complex foreground (e.g, the bigger monster) in the Monster set, MIC produces a fragmentary segmentation from the low-level features.

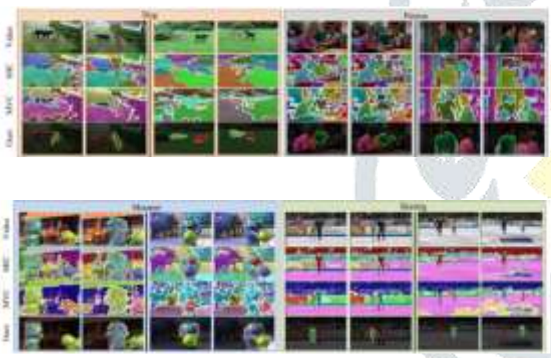


Figure 3. Segmentation results on our newly collected multiple foreground video dataset, where different videos in a set are separated by a line.

5. CONCLUSION

We proposed an object-based multiple foreground video co-segmentation method, whose key components are the use of object proposals as the basic element of processing, with a corresponding co-selection graph that places constraints among objects in the videos, and the multistate selection graph for addressing the problem of multiple foreground objects. Our MSG, which can handle single/multiple videos with single/multiple foregrounds, provides a general and global framework that can be used to extend any standard graph model to handle multi- state selection while still allowing optimization by existing energy minimization techniques.

6. REFERENCES

[1] M. Bergtholdt, J. Kappes, S.Schmidtand C. Schnorr. A study of parts-based object class detection using complete graphs. IJCV, 2010.

[2] D. Chen, H. Chen, and L.Chang. Video object cosegmentation. in ACM MM, 2012.

[3] W. Chiu and M. Fritz. Multi-class video co-segmentation with a generative multi-video model. In CVPR, 2013.

[4] N. Dalal and B. Triggs.

Histograms of oriented gradients for human detection. In CVPR, 2005.

[5] T. Deselaers, B. Alexe, and V. Ferrari. Weakly supervised localization and learning with generic knowledge. IJCV, 2012.

[6] I. Endres and D. Hoiem. Category independent object proposals. in ECCV, 2010.

[7] P. Felzenszwalb and D. Huttenlocher. Efficient belief propagation for early vision. IJCV, 2006.

[8] H. Fu, X. Cao, and Z. Tu. Cluster-based co-saliency detection. TIP, 2013.

[9] J. Guo, Z. Li, L. Cheong, and S. Zhou. Video co-segmentation for meaningful action extraction. In ICCV, 2013.

[10] A. Joulin, F. Bach, and J. Ponce. Multi-class cosegmentation. In CVPR, 2012.

[11] G. Kim, E. Xing, L. Fei-Fei, and T. Kanade. Distributed cosegmentation via submodular optimization on anisotropic diffusion. In ICCV, 2011.

[12] G. Kim and E. P. Xing. On multiple foreground cosegmentation in CVPR, 2012.

[13] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. TPAMI, 2006.

[14] Y. Lee, J. Kim, and K. Grauman. Key-segments for video object segmentation. In ICCV, 2011.

[15] C. Liu. Beyond pixels: exploring new representations and applications for motion analysis. PhD thesis, MIT, 2009.

[16] T. Ma and L. Latecki.Maximum weight cliques with mutex constraints for video object segmentation. In CVPR, 2012.

[17] F. Meng, H. Li, G. Liu, and K. Ngan. Object co-segmentation based on shortest path algorithm and saliency model. TMM, 2012.

[18] L. Mukherjee, V. Singh, J. Xu, and M. D.Collins. Analyzing the subspace structure of related images: Concurrent segmentation of image sets. In ECCV, 2012.

[19] J. Rubio, J. Serrat, and A. Lopez. Video co-segmentation. In ACCV, 2012.

[20] D. Tsai, M. Flagg, and J. Rehg. Motion coherent tracking with multilabel MRF optimization. In BMVC, 2010.

[21] S. Vicente, C. Rother, and V. Kolmogorov. Object cosegmentation. In CVPR, 2011.

[22] B. Zhang, H. Zhao, and X. Cao. Video object segmentation with shortest path. In ACM MM, 2012.

[23] D. Zhang, O. Javed, and M. Shah. Video object segmentation through spatially accurate and temporally dense extraction of primary object regions. In CVPR, 2013.