ENGINEERING IN ADVANCED RESEARCH SCIENCE AND TECHNOLOGY

ISSN 2278-2566
Vol.02, Issue.03
June -2018
Pages: 482-488

BANDWIDTH OPTIMIZED EFFICIENT IRRELEVANT FRAME DETECTION

1. K.KARTHIK,2. N.B.S.NAVEEN

1. M. Tech Student, Dept. of ECE, AKRG College of Engineering & Technology, Nallajerla, A.P
2. Assistant Professor, Dept. of ECE, AKRG College of Engineering & Technology, Nallajerla, A.P

ABSTRACT:

Video object co-segmentation refers to the problem of simultaneously segmenting a common category of objects from multiple videos. Most existing video co-segmentation methods assume that all frames from all videos contain the target objects. Despite the fact that there has been a lot of past work on video segmentation technique, it is still a difficult task to extricate the video objects precisely without interactions, particularly for those videos which contain irrelevant frames (frames containing no normal targets). In this paper, a novel multi-video object co-segmentation technique is raised to co-segment normal or comparable objects of pertinent casings in various videos, which incorporates three stages: 1) object proposal generation and clustering within each video; 2) weighted graph construction and common objects selection; and 3) irrelevant frames detection and pixel-level segmentation refinement. Further, this project is enhanced by using HAAR wavelet transformations to reduce band width utilization.

KEYWORDS: segmentation, Frames, Clustering, pertinent, Normal target, HAAR, Wavelet, Compression, Bandwidth.

INTRODUCTION:

Video object segmentation has played an important role in senior computer vision tasks, such as activity recognition, content-based retrieval, and object tracking. However, due to the complicated spatial-temporal correlations, large data, and high scene complexity, video object segmentation still faces serious challenges Recently, to make full use of the additional information across multiple videos, video co-segmentation [9][17] has become another research hotspot. Researchers propose to use the information across multiple videos to realize the segmentation of the common targets. The methods in [9][11] enforce the cooperative constraint with a global appearance model for the common targets. In [12] and [13], the trajectory cosaliency constraint and the coherent moving constraint for the foreground local parts are introduced into video co-segmentation, respectively. And more recently, the object-based approaches in [14][17] discover the common targets by mining the consistency information among the segmentation proposals. The cooperative constraint brings more information, while a neglected problem is that the interference information brought by the noisy frames, which contain no common targets, makes the video cosegmentation more complicated. Actually, these irrelevant frames involved videos are very common, such as a target is out of the view or occluded by the background as presented in Figure. In this letter, a more direct and unsupervised method is proposed, which is based on common cluster selection and adaptive recognition for the irrelative frames.

LITERATURE SURVEY:

The methods in [9]–[11] enforce the cooperative constraint with a global appearance model for the common targets. In [12] and [3], the trajectory cosaliency constraint and the coherent moving constraint for the foreground local parts are introduced into video cosegmentation, respectively. And more recently, the object-based approaches in [4]–[7] discover the common targets bymining the consistency information among the segmentation proposals. The cooperative constraint brings more information, while a neglected problem is that the interference information brought by the noisy frames, which contain no common targets, makes the video cosegmentation more complicated. Actually, these irrelevant frames involved videos are very common, such as a target is out of the view or occluded by the background as presented Wang *et al.* [8] creatively design a video cosegmentation framework for the irrelevant frames involved circumstance.

SEGMENTATION:

Segmentation partitions an image into distinct regions containing each pixels with similar attributes. To be meaningful and useful for image analysis and interpretation, the regions should strongly relate to depicted objects or features of interest. Meaningful segmentation is the first step from low-level image processing transforming a greyscale or colour image into one or more other images to high-level image description in terms of features, objects, and scenes. The success of image analysis depends on reliability of segmentation, but an accurate partitioning of an image is generally a very challenging problem.

EXISTING TECHNIQUE:

overview of the proposed approach is in Fig. (a). And accordingly, we will introduce the proposed method in this section from the following three stages. The first is the generation and clustering of object proposal within each video, the second step is the construction of weighted graph and selecting common objects, and the last is the irrelevant frames discrimination and pixel-level refinement.

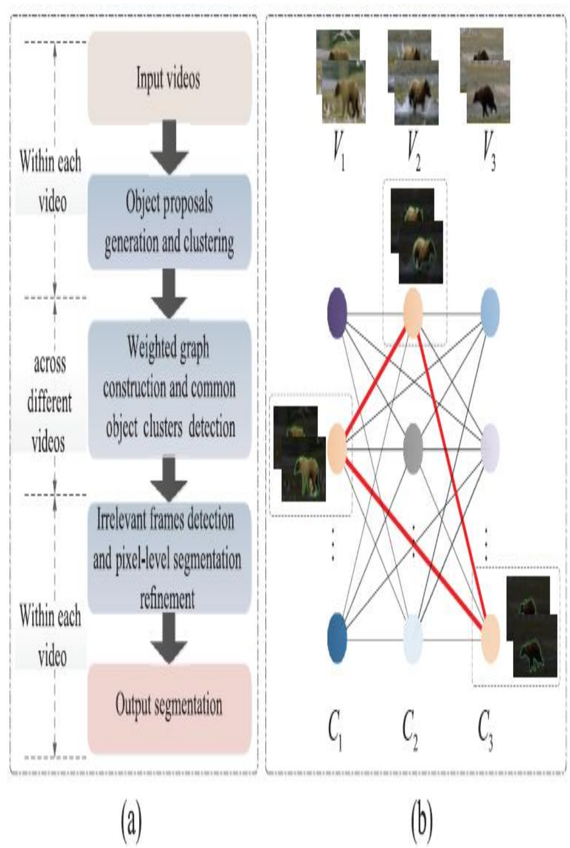


Fig. Algorithm flowchart and the key steps. (a) Overall framework of the proposed video cosegmentation method. (b) Weighted graph construction with proposal clusters from multiple videos and common object clusters' detection. The proposal clusters (denoted with solid nodes) of the same video are arranged in the same column.

INTRAVIDEO OBJECT PROPOSAL GENERATING AND CLUSTERING:

Let $V = \{Vi\}$, $i = 1, 2, \ldots, N$ be the original video set, where video $Vi = \{fj\ I\}$, $j = 1, 2, \ldots, Mi$, Mi is the number of the frames contained in video Vi. We first generate an object proposal set $Pi = \{p1\ i,1, p1\ i,2, \ldots, p1\ i,K1, p2\ i,1, p2\ i,2, \ldots, p2\ i,K2, \ldots, pMi\ i,1, pMi\ i,2, \ldots, pMi\ i,KMi\}$ with [20] for each video Vi, where Kj is the number of proposals of $j$thframe. To measure the possibilities of those proposals to be real objects, we utilize a score function which combines appearance and motion information $S(p) = Sa(p) + Sm(p)$ (1) where p is the object proposal to be measured, $Sa(p)$ is the appearance score that measures the boundary occlusion of proposal as well as the appearance difference from its nearby pixels. The appearance score $Sa(p)$ is computed as defined in [20], which is achieved with a pretrained ranker. Additionally, to measure the probability of a region belonging to a moving object, we also compute the motion score as defined in [5] $Sm(p) = 1 - \exp\_ -\chi 2$ flow$(p, \bar{}p)$ _ . (2) It measures the $\chi 2$ distance of optical flow histograms for proposal region p and its surrounding pixels $\bar{}p$ within the loosely fitting bounding box. According to $S(p)$, top 20 proposals of each frame are selected as the target candidates, which are most likely to be the true targets. For each video Vi, to mine the consistency implied in different frames as well as to establish associations between the candidate proposals, we apply affinity propagation (AP) clustering [21] to divide the proposal set Pi into multiple clusters, i.e., $Ci = \{c1i, c2i, \ldots, c\ NCi\ I\}$, where NCi is the cluster numbers for Vi. Here, we take the color histogram as the clustering feature and construct the affinity matrix by calculating the histogram $\chi 2$ distances for all the proposals. The cluster parameter (reference value) is fixed to the median of the similarity matrix.

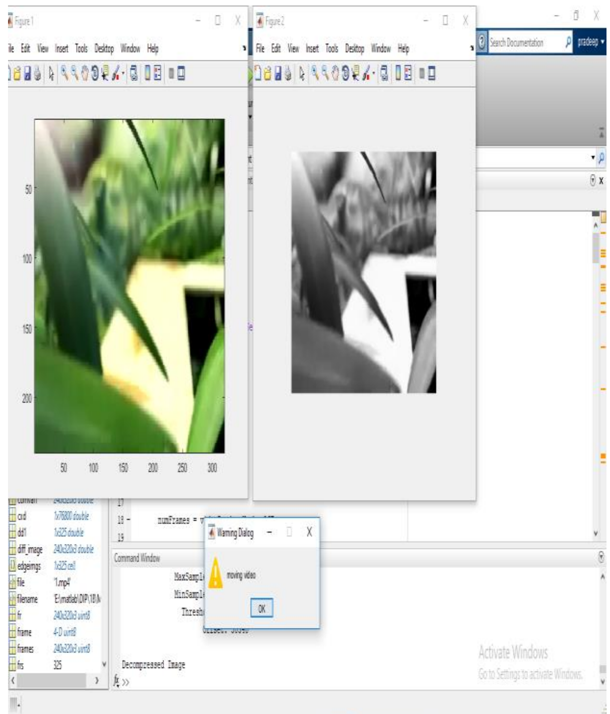
PROPOSED TECHNIQUE:

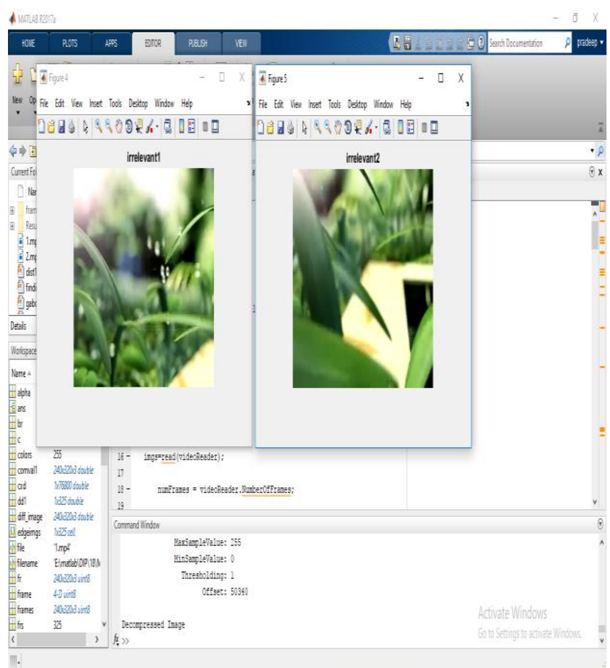
HAAR WAVELET TRANSFORMATION:

In this section we shall introduce the basic notions connected with the Haar transform.

First, we need to define the type of signals that we shall be analyzing with the Haar transform. Throughout this book we shall be working extensively with discrete signals. A discrete signal is a function of time with values occurring at discrete instants. Generally we shall express a discrete signal in the form f = (f1, f2,...,fN), where N is a positive even integer which we shall refer to as the length of f. The values of f are the N real numbers f1, f2,...,fN . These values are typically measured values of an analog signal g, measured at the time values $t = t1, t2,...,tN$. That is, the values of f are $f1 = g(t1), f2 = g(t2), ..., fN = g (tN)$. For simplicity, we shall assume that the increment of time that separates each pair of successive time values is always the same. We shall use the phrase equally spaced sample values, or just sample values, when the discrete signal has its values defined in this way. An important example of sample values is the set of data values stored in a computer audio file, such as a .wav file. Another example is the sound intensity values recorded on a compact disc. A non-audio example, where the analog signal g is not a sound signal, is a digitized electrocardiogram. Like all wavelet transforms, the Haar transform decomposes a discrete signal into two subsignals of half its length. One subsignal is a running average or trend; the other subsignal is a running difference or fluctuation. Let's begin by examining the trend subsignal. The first trend subsignal, $a1 = (a1, a2,...,aN/2)$, for the signal f is computed by taking a running average in the following way. Its first value, a1, is computed by taking the average of the first pair of values of f: $(f1 + f2)/2$, and then multiplying it by $\sqrt{2}$. That is, $a1 = (f1 +f2)/ \sqrt{2}$. Similarly, its next value a2 is computed by taking the average of the next pair of values of f: $(f3 + f4)/2$, and then multiplying it by $\sqrt{2}$. That is, $a2 = (f3 + f4)/ \sqrt{2}$. Continuing in this way, all of the values of a1 are produced by taking averages of successive pairs of values of f, and then multiplying these averages by $\sqrt{2}$. A precise formula for the values of a1 is $am = f2m-1 + f2m \sqrt{2}$, (1.2) for m = 1, 2, 3, . . . , N/2. For example, suppose f is defined by eight values, say f = (4, 6, 10, 12, 8, 6, 5, 5); then its first trend subsignal is $a1 = (5\sqrt{2}, 11\sqrt{2}, 7 \sqrt{2}, 5 \sqrt{2})$. This result can be obtained using Formula (1.2). Or it can be calculated as indicated in the following diagram: f: 4 6 10 12 8 6 5 5 511 7 5 ↓ ↓ ↓↓ a1: $5\sqrt{2}$ $11\sqrt{2}$ $7\sqrt{2}$ $5\sqrt{2}$. You might ask: Why perform the extra step of multiplying by $\sqrt{2}$? Why not just take averages? These questions will be answered in the next section, when we show that multiplication by $\sqrt{2}$ is needed in order to ensure that the Haar transform preserves the energy of a signal. The other subsignal is called the first fluctuation. The first fluctuation of the signal f, which is denoted by $d1 = (d1, d2,...,dN/2)$, is computed by taking a running difference in the following way. Its first value, d1, is calculated by taking half the difference of the first pair of values of f: $(f1-f2)/2$, and multiplying it by $\sqrt{2}$. That is, $d1 = (f1-f2)/ \sqrt{2}$. Likewise, its next value d2 is calculated by taking half the difference of the next pair of values of f: $(f3 - f4)/2$, and multiplying it by $\sqrt{2}$. In other words, $d2 = (f3 - f4)/ \sqrt{2}$. Continuing in this way, all of the values of d1 are produced according to the following formula: $dm = f2m-1 - f2m \sqrt{2}$, (1.3) for m = 1, 2, 3, . . . , N/2. For example, for the signal f = (4, 6, 10, 12, 8, 6, 5, 5) considered above, its first fluctuation d1 is $(- \sqrt{2}, - \sqrt{2}, \sqrt{2}, 0)$. This result can be obtained using Formula (1.3), or it can be calculated as indicated in the following diagram: f: 4 6 10 12 8 6 5 5 −1 −1 1 0 ↓ ↓ ↓↓ d1: $-\sqrt{2}$ $-\sqrt{2}$ $\sqrt{2}$ 0.

RESULT:





CONCLUSION:

A new video cosegmentation approach, which is based on common proposal clusters searching and adaptive recognition for irrelevant frames, is introduced in this letter. The common targets' mining strategy greatly improves the segmentation of each video, and the adaptive criterion condition adds flexibility for irrelative frames involved cases. The proposed framework is still extensible for multiclass problems. A straightforward extension could be conducting multiple selections for different common targets. HAAR compression ration performance and existing technique is compared an analyzed the best method for digital image compression

REFERENCES:

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. TPAMI 34(11), 2274–2282 (2012) 2. Alexe, B., Deselaers, T., Ferrari, V.: What is an object? In: CVPR. pp. 73–80 (2010) 3. Avidan, S.: SpatialBoost: Adding spatial reasoning to adaboost. In: ECCV. pp. 386–396 (2006) 4. Bai, X., Wang, J., Simons, D., Sapiro, G.: Video SnapCut: robust video object cutout using localized classifiers. In: ACM Trans. on Graphics. vol. 28, p. 70 (2009) 5. Batra, D., Kowdle, A., Parikh, D., Luo, J., Chen, T.: iCoseg: Interactive co-segmentation with intelligent scribble guidance. In: CVPR. pp. 3169–3176 (2010) 6. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient ND image segmentation. IJCV 70(2), 109–131 (2006) 7. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. TPAMI 26(9), 1124–1137 (2004) 8. Brox, T., Malik, J.: Object segmentation by long term analysis of point trajectories. In: ECCV. pp. 282–295 (2010) 9. Chen, D.J., Chen, H.T., Chang, L.W.: Video object cosegmentation. In: ACM Multimedia. pp. 805–808 (2012) 10. Chiu, W.C., Fritz, M.: Multi-class video co-segmentation with a generative multi-video model. In: CVPR. pp. 321–328 (2013) 11. Dai, J., Wu, Y.N., Zhou, J., Zhu, S.C.: Cosegmentation and cosketch by unsupervised learning. In: ICCV (2013) 12. Grundmann, M., Kwatra, V., Han, M., Essa, I.: Efficient hierarchical graph-based video segmentation. In: CVPR. pp. 2141–2148 (2010) 13. Guo, J., Li, Z., Cheong, L.F., Zhou, S.Z.: Video co-segmentation for meaningful action extraction. In: ICCV (2013) 14. Harel, J., Koch, C., Perona, P., et al.: Graph-based visual saliency. NIPS pp. 545–552 (2006) 15. Joulin, A., Bach, F., Ponce, J.: Discriminative clustering for image co-segmentation. In: CVPR. pp. 1943–1950 (2010) 16. Lee, Y.J., Kim, J., Grauman, K.: Key-segments for video object segmentation. In: ICCV. pp. 1995–2002 (2011) 17. Li, F., Kim, T., Humayun, A., Tsai, D., Rehg, J.M.: Video segmentation by tracking many figure-ground segments. In: ICCV (2013) 18. Liu, D., Chen, T.: A topic-motion model for unsupervised video object discovery. In: CVPR. pp. 1–8 (2007) 19. Liu, D., Hua, G., Chen, T.: A hierarchical visual model for video object summarization. TPAMI 32(12), 2178–2190 (2010) 20. Ma, T., Latecki, L.J.: Maximum weight cliques with mutex constraints for video object segmentation. In: CVPR. pp. 670–677 (2012)
21. Ochs, P., Brox, T.: Object segmentation in video: a hierarchical variational approach for turning point trajectories into dense regions. In: ICCV. pp. 1583–1590 (2011) 22. Ochs, P., Brox, T.: Higher order motion models and spectral clustering. In: CVPR. pp. 614– 621 (2012) 23. Papazoglou, A., Ferrari, V.: Fast object segmentation in unconstrained video. In: ICCV (2013) 24. Prest, A., Leistner, C., Civera, J., Schmid, C., Ferrari, V.: Learning object class detectors from weakly annotated video. In: CVPR. pp. 3282–3289 (2012) 25. Rubinstein, M., Joulin, A., Kopf, J., Liu, C.: Unsupervised joint object discovery and segmentation in internet images. In: CVPR. pp. 1939–1946 (2013)