LETTER

Co-saliency Detection Linearly Combining Single-View Saliency and Foreground Correspondence

Huiyun JING†, Xin HE††(a), Qi HAN†††, Nonmembers, and Xiamu NIU†††, Member

SUMMARY The research of detecting co-saliency over multiple images is just beginning. The existing methods multiply the saliency on single image by the correspondence over multiple images to estimate co-saliency. They have difficulty in highlighting the co-salient object that is not salient on single image. It is caused by two problems. (1) The correspondence computation lacks precision. (2) The co-saliency multiplication formulation does not fully consider the effect of correspondence for co-saliency. In this paper, we propose a novel co-saliency detection scheme linearly combining foreground correspondence and single-view saliency. The progressive graph matching based foreground correspondence method is proposed to improve the precision of correspondence computation. Then the foreground correspondence is linearly combined with single-view saliency to compute co-saliency. According to the linear combination formulation, high correspondence could bring about high co-saliency, even when single-view saliency is low. Experiments show that our method outperforms previous state-of-the-art co-saliency methods.

key words: saliency detection, co-saliency, foreground-correspondence

1. Introduction

It is common that similar objects of the same category repeatedly appear on multiple online images. Detecting the same (similar) salient objects from multiple images has become one of the most important and challenging machine vision problems. To solve this problem, co-saliency detection on multiple images is proposed. It discovers the common saliency from multiple images [1], which is different from the traditional visual attention methods [2] detecting the saliency on single image. The co-saliency can be useful priori knowledge for image segmentation [3] to improve the accuracy of segmentation.

Currently, most of the existing co-saliency detection methods [4]–[6] are formulated on detecting common saliency between a pair of images. However, they are hard to be generalized to the case of more than two images. A few researchers [1], [7] devised novel co-saliency detection methods on multiple images, MBCS (multiplication based co-saliency method) and CBCS (cluster-based co-saliency method).

These methods regard the object, being simultaneously salient on each single image and frequently repeated on multiple images, as the co-salient object. According to this understanding, they employ the multiplication formulation of single-view saliency and correspondence to compute co-saliency. In MBCS method, correspondence is computed by matching the SIFT features of points sampled from multiple images. At the same time, CBCS method compute correspondence by estimating the distribution of the cluster that each pixel belongs to among the multiple images.

The above mentioned methods [1], [7] prefer to detect the object that is salient on each single image and frequently repeated over multiple images. However, in practice, co-salient object may not be salient on single image [5]. These co-salient objects are difficult to detect by the existing methods. For example, in Fig. 1 the Venus statue is not salient on the first input image highlighted by a red rectangle, whereas it is the co-salient object over the four input images. The current co-saliency methods [1], [7] are unable to label the Venus statue as salient on the co-saliency map corresponding to the first input image (Fig. 1 c and d).

There are mainly two causes of the above issues. The first one is that the correspondence computation lacks precision, especially when the co-salient object greatly changes its appearance or the background areas are very similar over the multiple images. Another important reason is the multiplication formulation of co-saliency underestimates the intrinsic influence of correspondence for co-saliency detection. It is well known that high correspondence could significantly boost the co-saliency of one object on multiple images [8]. However, under the multiplication formulation,
when the single-view saliency of co-salient object is very low, its co-saliency would not be high, even though the correspondence is high.

To address the above mentioned issues, we propose a novel co-saliency detection method linearly combining foreground correspondence and single-view saliency. The proposed method is introduced in detail in Sect. 2. And experimental results and conclusion are presented in Sects. 3 and 4.

2. The Proposed Co-saliency Detection Mode

2.1 Overview of the Proposed Model

The co-salient object repeatedly appears on multiple images and always lies in the image foreground. It implies co-salient object always belongs to the correspondent image foreground regions between multiple images. In another word, the correspondent image foreground regions are very likely to be the co-salient object. Highlighting the correspondent foreground regions benefits detecting co-salient object. Thus the proposed method firstly locates the correspondent foreground regions between multiple images by utilizing the progressive graph matching approach, which is robust to the in-class variations of co-salient object.

The co-salient object may not be salient on single image. High correspondence turns non-salient object on single image to the co-salient one on multiple images. To improve the importance of correspondence, we compute co-saliency by linearly combining single-view saliency and foreground correspondence, rather than multiplying. The block diagram of the proposed method is shown in Fig. 2.

2.2 Locating Correspondent Foreground Regions and Computing Single-View Saliency

For more accurately locating the correspondent image foreground, we firstly roughly estimate the foreground areas for each image. Then we utilize progressive graph matching between the multiple estimated image foreground areas to locate the correspondent foreground regions.

1) Estimating image foreground: Image foreground areas are the complementary regions of image background. Compared with the kaleidoscopic image foreground, the sub-regions of image background are similar to each other. Thus the image foreground can be obtained by estimating the complementary regions of image background. Based on the basic rule of photographic composition, photographers usually would not crop salient objects along the view frame. Consequently, it is reasonable that declaring the image boundary to be mostly background. We thus use the image boundary as the initial possible image background, and utilize GrabCut to obtain more image background areas. Then the complementary areas of the image background segmented by GrabCut can be regarded as the estimated image foreground.

2) Locating correspondent foreground regions by utilizing progressive graph matching: Since the progressive graph matching approach is robust to the in-class variations of co-salient object, we utilize progressive graph matching between the multiple estimated image foreground areas to locate the correspondent foreground regions.

3) Background contrast based single-view saliency computation: Here, our previously proposed background contrast based salient region detection method (BCSS) [11] is utilized, due to its better performance. The BCSS method measures the saliency of each region by computing its contrast to the estimated image background.

2.3 Co-saliency Computation

After locating the correspondent foreground regions (CFR)
and computing the single-view saliency, we have obtained the map spotlighting CFR (MCFR) and the single-view saliency map (SSM). However, the two kind maps have difficulty in highlighting the co-salient object. For example, in some cases the co-salient object is not highlighted on some single-view saliency map (Fig. 3 (a)), in other cases the located correspondent foreground regions are only one part of the co-salient object (Fig. 3 (b)). For synthetically considering the influence of correspondence and single-view saliency cues on co-saliency computation, the two kind maps MCFR and SSM needs to be fused to generate the co-saliency map. And since the co-salient object may be not salient on some single-view saliency map, we thus adopt the linear combination mode rather than the multiplication mode to generate co-saliency map:

\[
CS = MCFR + SSM
\]

3. Experimental Results

In order to completely verify the proposed method, we collect six image sets from MSRC database [12] as the testing dataset. Similar as the dataset used in previous works [1], [7], each image set comprises the images containing co-salient object of the same category in MSRC database. The image numbers of the six testing sets (airplane, car, cat, cow, face and sheep) vary from 22 to 30, and there are 166 images in the six testing image sets. Note that some testing image sets have similar backgrounds, such as airplane and cow, and some testing image sets have co-salient object greatly varying the appearance and size, such as cat and car. The similar background areas or the substantially variant co-salient object on the multiple images usually confuse the correspondence computation and causes unsatisfactory co-saliency maps. Thus the two kind image sets are the typical challenge cases for co-saliency detection methods. Furthermore, similar as previous works, precision-recall curve is used as as performance metric.

(1) Evaluation of co-saliency in the whole dataset

As mentioned above, currently, only a few approaches [1], [7] are formulated to detect co-saliency from multiple images (more than two images). We compare our method with them (MBCS [7] and CBCS [1]). Figure 4 shows the resulting precision-recall curves. To further evaluate the improvement of these co-saliency methods comparing with their employed single-view saliency approaches, we also plot the single-view saliency methods CA [13] and CSSS [1] respectively utilized by MBCS [7] and CBCS [1].

From the precision-recall curves in Fig. 4, we have the following two observations. First, our co-saliency method performs better than the current existing co-saliency method MBCS and CBCS. Second, the performance of co-saliency methods MBCS and CBCS is very close to their employed single-view saliency approaches. On the contrary, the performance of our co-saliency method is promoted higher than the utilized single-view saliency method BCSS and the correspondence computation method generating MCFR. Since MCFR contains more pixels with the saliency value 255, the minimum recall values of MCFR are higher than those of the other methods.

(2) Evaluating co-saliency in the image set having similar background areas

As stated above, the image set having similar background areas is the challenge case. Like the images of airplane image set, the image background sky and meadow hardly changes. We compare the performance of our method with MBCS [7] and CBCS [1] on the airplane image set. The resulting precision-recall curves and visual comparison of co-saliency maps (Fig. 5 and Fig. 6) indicate two points. First, our co-saliency method outperforms the other two methods on the image set with similar image background. Second, our method performs better than its single-view saliency method BCSS and correspondence computation method generating MCFR.

(3) Evaluating co-saliency in the image set with substantially variant co-salient object

We further compare our method with MBCS [7] and CBCS [1] on the image set with substantially variant co-salient object, such as cat image set. In this image set, the co-salient object cat changes its color and size. Such intra-class variations challenge the correspondence algorithms employed by co-saliency methods. The performance comparison of these co-saliency methods can be seen in
4. Conclusion

In this paper, a novel co-saliency detection method for multiple images is proposed. The method provides an appropriate mode of linearly combining single-view saliency and foreground correspondence to compute co-saliency. The proposed mode is beneficial to highlight the estimated corresponding co-salient parts which are darkened in the single-view saliency map. Background-contrast based saliency method is employed to obtain the single-view saliency. Progressive graph matching method is utilized to robustly compute the correspondence between the estimated foreground areas on the multiple images. The experimental results demonstrate that our method outperforms previous state-of-the-art co-saliency methods.

Acknowledgments

This work is supported by the National Key Technology Research and Development Program (2012BAH46B02).

References

[1] H. Fu, X. Cao, and Z. Tu, “Cluster-based co-saliency detection,” IEEE Trans. Image Process., vol.22, pp.3766–3778, 2013.
[2] A. Kimura, R. Yonetani, and T. Hirayama, “Computational models of human visual attention and their implementations: A survey,” IEICE Trans. Inf. Syst., vol.E96-D, no.3, pp.562–578, March 2013.
[3] N.R. Pal and K.P. Sankar, “A review on image segmentation techniques,” Pattern Recognit., vol.26, no.9, pp.1277–1294, 1993.
[4] H.T. Chen, “Preattentive co-saliency detection,” 2010 17th IEEE International Conference on Image Processing (ICIP), pp.1117–1120, 2010.
[5] D.E. Jacobs, D.B. Goldman, and E. Shechtman, “Cosaliency: Where people look when comparing images,” ACM Symposium on User Interface Software and Technology, pp.219–228, 2010.
[6] H. Li and K.N. Ngan, “A co-saliency model of image pairs,” IEEE Trans. Image Process., vol.20, no.12, pp.3365–3375, 2011.
[7] K.Y. Chang, T.L. Liu, and S.H. Lai, “From co-saliency to cosegmentation: An efficient and fully unsupervised energy minimization model,” 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.2129–2136, 2011.
[8] G. Kim, E.P. Xing, L. Fei-Fei, and T. Kanade, “Distributed cosegmentation via submodular optimization on anisotropic diffusion,” 2011 IEEE International Conference on Computer Vision (ICCV), pp.169–176, 2011.
[9] J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions,” Image and Vision Computing, vol.22, no.10, pp.761–767, 2004.
[10] M. Cho and K.M. Lee, “Progressive graph matching: Making a move of graphs via probabilistic voting,” 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.398–405, 2012.
[11] J. Huiyun, H. Qi, H. Xin, and N. Xiamu, “Background contrast based salient region detection,” Neurocomputing, vol.124, pp.57–62, 2014.
[12] J. Shotton, J. Winn, C. Rother, and A. Criminisi, “Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context,” Int. J. Comput. Vis., vol.81, no.1, pp.2–23, 2009.
[13] S. Goferman, L. Zelnik-Manor, and A. Tal, “Context-aware saliency detection,” IEEE Trans. Pattern Anal. Mach. Intell., vol.34, no.10, pp.1915–1926, 2012.