SALIENCY BASED IMAGE CO-SEGMENTATION

Shaik. Aawez Ahmed [1], N. Subbarao[2]

[1]PG Scholar, Dept. of ECE - DS&CE , QUBA College of Engineering & Technology, Andhra Pradesh, India
[2]Associate Professor, Dept. of ECE, QUBA College of Engineering & Technology, Andhra Pradesh, India

---***---

***Abstract** - Most existing high-performance Co-Segmentation algorithms are usually complex due to the way of co-labeling a set of images as well as the common need of fine-tuning few parameters for effective Co-Segmentation. In this method, instead of following the conventional way of co-labeling multiple images, we propose to first exploit inter-image information through co-saliency, and then perform single-image segmentation on each individual image. To make the system robust and to avoid heavy dependence on one single saliency extraction method, we propose to apply multiple existing saliency extraction methods on each image to obtain diverse salient maps. Our major contribution lies in the proposed method that fuses the obtained diverse saliency maps by exploiting the inter-image information, which we call saliency co fusion. Experiments on five benchmark datasets with eight saliency extraction methods show that our saliency co fusion based approach achieves competitive performance even without parameter fine-tuning when compared with the state-of-the-art methods.*

***Key Words:** Co-Segmentation algorithms, diverse saliency maps, Inter-image information, saliency extraction method, and saliency co fusion.*

1. INTRODUCTION

An image may be defined as a two-dimensional function f(x, y), where x & y are spatial coordinates, & the amplitude off at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y & the amplitude values of f are all finite discrete quantities, we call the image a digital image. The field of DIP refers to processing digital image by means of digital computer. Digital image is composed of a finite number of elements, each of which has a particular location & value. Digital image processing, as already defined is used successfully in a broad range of areas of exceptional social & economic value.

1.1 Fundamental steps of Image processing:

In Image processing there are some fundamental basic steps which help in easy analysis and enhancing the quality of an image.

A. Image Acquisition

It is the process of creation of digitally encoded representation of visual characteristics of an object such as physical scene and interior structure of an object.

B. Image Enhancement

Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. Such as, changing brightness & contrast etc.

C. Image restoration

Image restoration is an area that also deals with improving the appearance of an image. Image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

D. Color image processing

It is the area that gains that has crucial role due to rapid increment in the use of Digital Images over internet. It includes color, modeling and processing in digital domains.

E. Wavelets and multi resolution processing

Wavelets are the foundation for representing images in various degrees of resolution. Images subdivision successively into smaller regions for data compression and for pyramidal representation.

F. Image Compression

The compression technique is mostly widen process for reducing the storage that is required to save image or to transmit bandwidth.

G. Morphological processing

It deals with the tools for extracting image components that are useful in description and in representation of an image.

H. Segmentation

Segmentation [1] is a process in which an image is made into several pixels by partitioning. In this the pixels having same characteristics comparing to its beside pixels then such similar pixels will form as a Super pixel.

I. Representation and description

Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. Description deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.

J. Object recognition

Recognition is the process that assigns a label, such as, "vehicle" to an object based on its descriptors.

K. Knowledge base

It is the process in which detailing the region of an image where information of the region is known to be calculated. It is quite complex because an interrelated list of all major possible defects in materials inspection problem or an image database containing high-resolution satellite images of a region in connection with change-detection applications.

First, most of the state-of-the-art Co-Segmentation algorithms [1]–[3] require fine-tuning of quite a few parameters and the co-labeling of multiple images simultaneously, which are very complex and time-consuming, especially for large diverse datasets. Second, as seen in the existing works, co-segmenting images might not perform better than single image segmentation for some datasets. This might be due to the additional energy term commonly used to enforce inter-image consistency, which often results into unsmooth segmentations in individual images. In this method, we focus on binary image Co-Segmentation, i.e. extracting a common foreground from a given image set. Instead of following the conventional way of co-labeling multiple images, we aim to exploit inter-image information through co-saliency, and then perform single-image segmentation on each individual image. Moreover, to make the system robust and avoid heavy dependence on one single saliency extraction method for generating co-saliency, we propose to apply multiple saliency extraction methods on each image.. We call the proposed method saliency co fusion, whose objectives include: Boosting the saliency of common foreground regions and Suppressing the saliency of background regions.

The process flow of the proposed saliency confusion based on the image Co-Segmentation. The key component lies in the developed saliency co fusion process, which is performed at the super pixel level. Particularly, we define each saliency map region (produced by one saliency detection method) of one super pixel as an element and give a weight for each element. We formulate the weight selection as an energy minimization problem, where we incorporate saliency recommendations from similar elements, foreground/background priors through similar element voting, and neighbor smoothness constraints. Finally, the fused saliency for a super pixel is just a weighted summation of all the saliency maps of the super pixel. Experimental results show that our saliency co fusion based Co-Segmentation achieves competitive performance even without fine-tuning the parameters, i.e., at default setting, compared with the state-of-the-art Co-Segmentation algorithms.

2. LITERATURE SURVEY

Our method is closely related to Co-Segmentation and co-saliency research method. Many Co-Segmentation algorithms have been proposed in the literature. Early approaches [1]–[3] focused on segmenting a pair of images containing one common object. It was later extended to deal with multiple images containing one common object with more effective or more efficient models enforcing inter-image consistency [4], [5], [22]–[25]. However, there are also some algorithms being designed for segmenting multiple common foregrounds from a given image set [11]–[14], where the best performers make use of supervised information.

On the contrary, our method is purely an unsupervised approach. A few interactive Co-Segmentation approaches [8]–[10] have also been proposed, where users can give scribbles for one or a small number of the images. Recently, [16] applied dense SIFT matching to discover common objects, and co-segment them out from noisy image dataset, where some images do not contain common objects. They tried to enforce inter-image consistency strongly by developing matching based prior, so as to exclude noise image from participating in the Co-Segmentation process. In [19], Co-Segmentation was combined with co-sketch for effective Co-Segmentation by sharing shape templates. In [26], Co-Segmentation problem was addressed by establishing consistent functional maps across images in a reduced functional space, which requires training. Another interesting work [20], which reports state-of the- art performance, employed region-level matching.

Also, it determined a good co-segment by checking whether it can be well composed from other co-segments. Most of these methods focused on pixel level co-labelling whereas we focus on saliency co fusion. Just like we use multiple saliency maps, there are also some methods that use multiple segmentation proposals to perform semantic segmentation.

2.1 Co-Saliency:

Co-saliency typically refers to the common saliency existing in a set of images containing similar objects. The term "co-saliency" was first coined in [28], in the sense of what is unique in a set of similar images, and the concept was later linked to extract common saliency, which is very useful for many practical applications [29], [30]. For example, co-saliency object priors have been efficiently used for Co-Segmentation in [31]. Recently, a cluster based co-saliency method using various cues was proposed in [32], which learns the global correspondence and obtains cluster saliency quite well.

2.2 Implementation Details:

For optimization, since G is positive semi-definite and the constraints are linear; the objective function defined in (1) is essentially a quadratic programming problem, which is solved by the interior-point convex algorithm provided in Matlab. Once the fused saliency map is available, different single image segmentation algorithms can be applied for segmentation. In this research, we adopt two segmentation methods as two variations. One is the classical Otsu's method, which is an optimal threshold based method. The other one is GrabCut algorithm with some modification. Specifically, by noticing the final fused saliency map containing certain boundary information, following we modify the GrabCut energy equation and add another localization potential to ensure that segmentation is guided not only by color, but also by the location prescribed by the object prior contained in the fused saliency map. The foreground (*FG*) and the background (*BG*) seed locations.

3. PROPOSED METHOD

In the proposed method, the process of Co-segmentation is used, in which a number of saliency maps are generated by using edge detection method. In this the Image is made to several number of pixels using segmentation process. Each pixel compares its properties with its neighbor pixel, if its properties are same those two pixels form as a super pixel. The process of co-segmentation and co-fusing the different saliency maps obtained from source input is shown below.

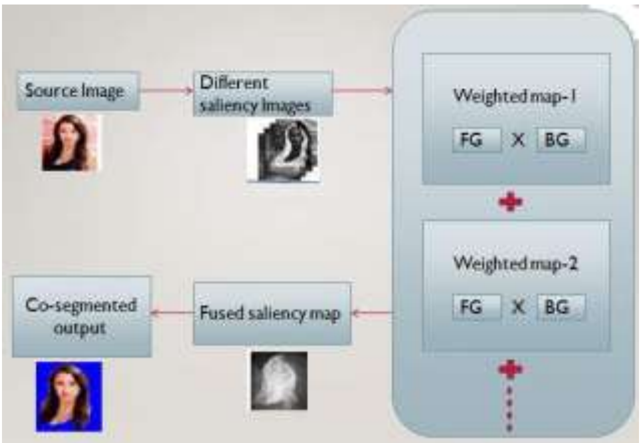

Fig -1: Block diagram of proposed method

Initially, a source image is taken as an input and by using edge detection process different saliency maps are generated by using characteristics of super pixels. The saliency maps of each unit consist of a foreground and background in it. In this process each unit multiplies with its foreground and background. Similarly every unit multiplies its foreground with its background. Finally in this stage all units adds their weights called as pixel level addition to obtain a fused saliency map. Later from this fused saliency maps co-segmented output is obtained.

In this process each unit multiplies with its foreground and background. Similarly every unit multiplies its foreground with its background. Finally in this stage all units adds their weights called as pixel level addition to obtain a fused saliency map. Later from this fused saliency maps co-segmented output is obtained. Moreover, to make the system robust and avoid heavy dependence on one single saliency extraction method for generating co-saliency, we propose to apply multiple saliency extraction methods on each image. Eventually, an enhanced saliency map is generated for each image by fusing its various saliency maps via weighted summation at super pixel level, where the weights are optimized by exploiting inter-image information.

We call the proposed method saliency co fusion, whose objectives include:

i. Boosting the saliency of common foreground regions.
ii. Suppressing the saliency of background regions.

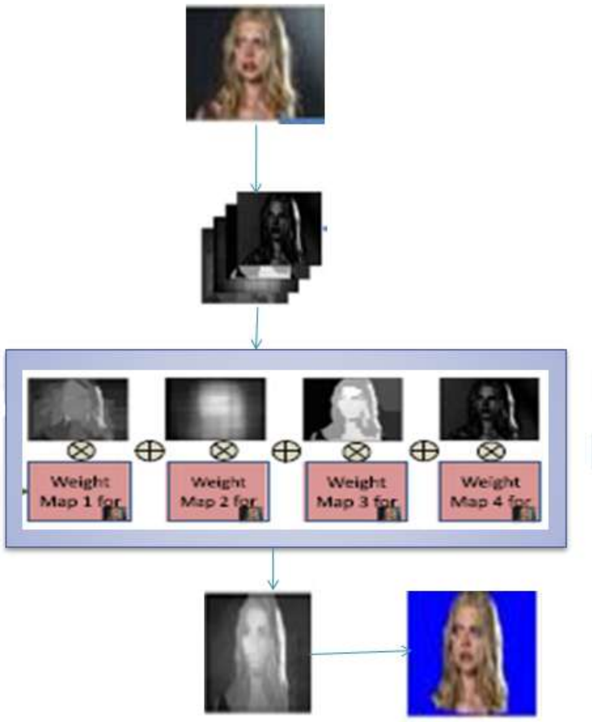

Fig -2: Flow chart representation of co-segmentation process

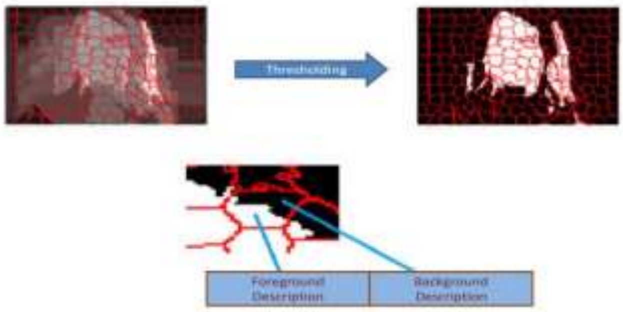

Fig -3: Separation of foreground and background

The above figure illustrates the process flow of the proposed saliency confusion based on the image Co-Segmentation. The key component lies in the developed saliency co fusion process, which is performed at the super pixel level. Particularly, we define each saliency map region (produced by one saliency detection method) of one super pixel as an element and give a weight for each element. Finally, the fused saliency for a super pixel is just a weighted summation of all the saliency maps of the super pixel. Experimental results show that our saliency co fusion based Co-Segmentation achieves competitive performance even without fine-tuning the parameters, i.e., at default setting, compared with the state-of-the-art Co-Segmentation algorithms.

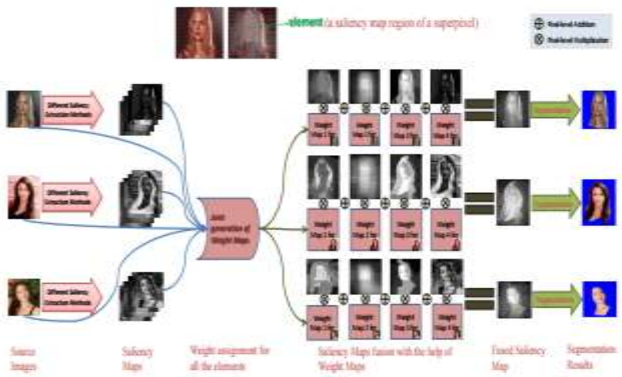

Fig -4: Flow chart representation of multiple images

In this method, we focus on binary image Co-Segmentation, i.e. extracting a common foreground from a given image set. Instead of following the conventional way of co-labeling multiple images, we aim to exploit inter-image information through co-saliency, and then perform single-image segmentation on each individual image.

3.1 Enhanced Adaptive Hybrid Median Filter algorithm:
Enhanced Adaptive Hybrid Median filtering is similar to an averaging filter, in that each output pixel is set to an average of the pixel values in the neighborhood of the corresponding input pixel. However, with median filtering, the value of an output pixel is determined by the median of the neighborhood pixels, rather than the mean. The median is much less sensitive than the mean to extreme values (called outliers). Median filtering is therefore better able to remove these outliers without reducing the sharpness of the image.

3.2 Saliency Feature Extraction using HOG (Histogram of Gradient) algorithm:
We compute Histograms of Oriented Gradients (HOG) to describe the distribution of the selected edge points. HOG is based on normalized local histograms of image gradient orientations in a dense grid. The HOG descriptor is constructed around each of the edge points. The neighborhood of such an edge point is called a cell. Each cell provides a local 1-D histogram of quantized gradient directions using all cell pixels. To construct the feature vector, the histograms of all cells within a spatially larger region are combined and contrast-normalized.

3.3 Modified Fuzzy C means algorithm with expectation maximization clustering algorithm:
First, we calculate the image saliency by using the color and space information of both local and global in single scale. Then by applying the multi-scale fusion, we can effectively inhibit outstanding but not salient region in each single scale, and different scale can also reflect salient region of the images from different aspects. Many image display devices allow only a limited number of colors to be simultaneously displayed. Usually, this set of available colors, called a color palette, may be selected by a user from a wide variety of available colors. Such device restrictions make it particularly difficult to display natural color images since these images usually contain a wide range of colors which must then be quantized by a palette with limited size. This color quantization problem is considered in two parts: the selection of an optimal color palette and the optimal mapping of each pixel of the image to a color from the palette. The entire work is divided into two stages. First enhancement of colors separation of satellite image using de-correlation stretching is carried out and then the regions are grouped into a set of five classes using Fuzzy c-means clustering algorithm. Using this two step process, it is possible to reduce the computational cost avoiding feature calculation for every pixel in the image. Although the color is not frequently used for image segmentation, it gives a high discriminative power of regions present in the image.

3.4 Advantages of Proposed system:

- It is the fastest algorithm when compared to the k means algorithm and modified Fuzzy C Means algorithm.
- The proposed anisotropic diffusion filter will completely eliminates the speckle noise from the image.
- The proposed algorithm is applicable for RGB color space images.
- We were able to successfully apply a segmentation method based on Expectation Maximization clustering.
- The use of saliency promises benefits to multimedia applications.
- It is clear that limited accuracy is one of the possible reasons for this. Another reason could be that in general saliency allows us to detect salient regions of the image rather than objects. To fill this gap we study to what extend the integration of segmentation into saliency detection allows the estimation of saliency of objects.

3.5 Video object segmentation:
The role of image processing technology has become more and more important for our life. This technology has been applied to various fields and a lot of applications such as detecting defects for manufacturing process and recognizing human face for identification system have been implemented. The requirement for image processing has become complex, various methods have been developed.

Recently, we often deal with multi-dimensional image data such as 3D volume data and video data. In the medical field, we are required to process 3D volume data, for instance CT, MRI and PET processing. In ITS it is necessary to process video image sequence to detect moving objects. 3D image processing generally requires more complex process and contains more parameters than 2D ones.

Therefore, the knowledge and experiences about image processing are necessary in this field. A lot of multi-dimensional image processing algorithms which can apply complex problem have been developed. However most of these methods need adjusting many parameters and processing procedures manually by trial and error every other problem.

4. RESULTS AND DISCUSSIONS

4.1 Input Image:

Here, an image is taken as an input as shown below. The source image contains of both the foreground and background. So by using Image co-segmentation and co-fusion process the foreground should be boosted up and background is suppressed.

Fig -5: Input image

4.2 Output Image:

Here, in the output the foreground is boosted up and background is suppressed which is the desired output.

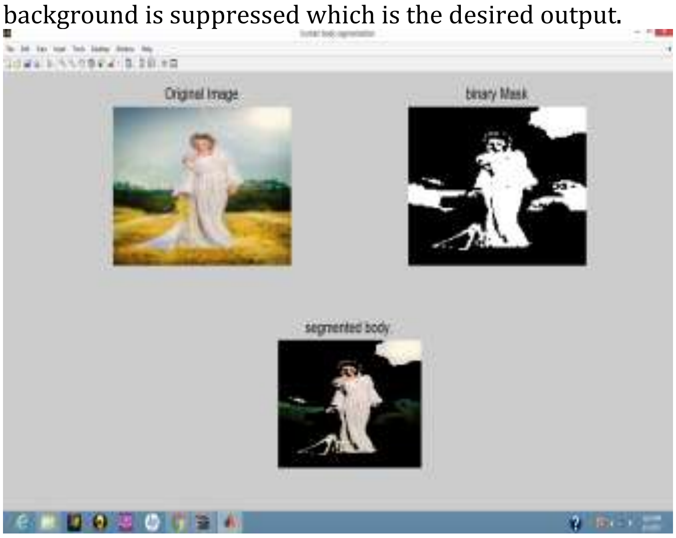

Fig -6: Output image

5. CONCLUSION

We have proposed a novel saliency co fusion approach for the purpose of image Co-Segmentation which uses the association of similar images to fuse multiple saliency maps of an image in order to boost up common foreground saliency and suppress background saliency. Experimental results on five benchmark datasets show that our method while co-fusing eight different saliency maps, achieves very competitive performance, compared to the state-of-the-art methods of image Co-Segmentation. This process of co-segmenting image depending on weights separation and finally co-fusing them by taking consideration of foreground and background is an efficient method and produces a fine output. This method have various advantages particularly in the images having clutter background, faded images and complex images and it have wider applications in the field of satellite communication, Image retrieval and Medical applications.

REFERENCES

[1] [1] C. Rother, T. Minka, A. Blake, and V. Kolmogorov, "Co-Segmentation of image pairs by histogram matching-incorporating a global constraint into MRF," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2006, vol. 1, pp. 993–1000.

[2] [2] D. S. Hochbaum and V. Singh, "An efficient algorithm for Co-Segmentation," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep.-Oct. 2009, pp. 269–276.

[3] [3] L.Mukherjee,V. Singh, andC.R.Dyer, "Half-integrality based algorithms for Co-Segmentation of images," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2009, pp. 2028–2035.

[4] [4] A. Joulin, F. Bach, and J. Ponce, "Discriminative clustering for image Co-Segmentation," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2010, pp. 1943–1950.

[5] [5] G. Kim, E. P. Xing, L. Fei-Fei, and T. Kanade, "Distributed Co-Segmentation via submodular optimization on anisotropic diffusion," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 169–176.

[6] [6] H. Li, F. Meng, andK.N.Ngan, "Co-salient object detection frommultiple images," *IEEE Trans. Multimedia*, vol. 15, no. 8, pp. 1896–1909, Dec. 2013.

[7] [7] F. Meng, H. Li, G. Liu, and K. N. Ngan, "Object Co-Segmentation based on shortest path algorithm and saliency model," *IEEE Trans. Multimedia*, vol. 14, no. 5, pp. 1429–1441, Oct. 2012.

[8] [8] D. Batra, A. Kowdle, D. Parikh, J. Luo, and T. Chen, "iCoseg: Interactive Co-Segmentation with intelligent scribble guidance," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2010, pp. 3169–3176.

[9] [9] M. D. Collins, J. Xu, L. Grady, and V. Singh, "Random walks based multi-image segmentation: Quasiconvexity results and GPU-based solutions," in *Proc., IEEE Comput. Vis. Pattern Recog.*, Jun. 2012, pp. 1656–1663.

[10] [10] D. Batra, D. Parikh, A.Kowdle, T. Chen, and J. Luo, "Seed image selection in interactive Co-Segmentation," in *Proc. IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 2393–2396.

[11] [11] F. Meng, B. Luo, and C. Huang, "Object Co-Segmentation based on directed graph clustering," in

Proc. IEEE Visual Commun. Image Process., Nov. 2013, pp. 1–5.

[12] [12] Z. Liu, J. Zhu, J. Bu, and C. Chen, "Object Co-Segmentation by nonrigid mapping," *Neurocomputing*, vol. 135, pp. 107–116, 2014.

[13] [13] T. Ma and L. J. Latecki, "Graph transduction learning with connectivity constraints with application to multiple foreground Co-Segmentation," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2013,

[14] pp. 1955–1962.

[15] [14] G. Kim and E. P. Xing, "Onmultiple foreground Co-Segmentation," in *Proc.IEEE Comput. Vis. Pattern Recog.*, Jun. 2012, pp. 837–844.

[16] [15] F. Meng, J. Cai, and H. Li, "On multiple image group Co-Segmentation," in *Proc. Asian Conf. Comput. Vis.*, 2014, pp. 258–272. [17] M. Guillaumin, D. Kttel, and V. Ferrari, "Imagenet auto-annotation with segmentation propagation," *Int. J. Comput. Vis.*, vol. 110, no. 3, pp. 328–348, 2014.

[17] [18] K. R. Jerripothula, J. Cai, and J. Yuan, "Group saliency propagation for large scale and quick image Co-Segmentation," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 4639–4643.

[18] [19] J. Dai,Y. N.Wu, J. Zhou, and S.-C. Zhu, "Co-Segmentation and cosketch by unsupervised learning," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1305–1312.

[19] [20] A. Faktor andM. Irani, "Co-Segmentation by composition," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1297–1304.

[20] [21] S. Vicente, C. Rother, and V. Kolmogorov, "Object Co-Segmentation," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2011, pp. 2217– 2224.

[21] [22] S. Vicente, V. Kolmogorov, and C. Rother, "Co-Segmentation revisited: Models and optimization," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 465–479.

[22] [23] J. C. Rubio, J. Serrat, A. L´opez, and N. Paragios, "Unsupervised Co-Segmentation through region matching," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2012, pp. 749–756.

[23] [24] L. Mukherjee, V. Singh, and J. Peng, "Scale invariant Co-Segmentation for

[24] image groups," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2011, pp. 1881–1888.

[25] [25] J. Yuan *et al.*, "Discovering thematic objects in image collections and videos," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2207–2219, Apr. 2012.

[26] [26] F.Wang, Q. Huang, and L. J. Guibas, "Image Co-Segmentation via consistent functional maps," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 849–856.

[27] [27] A. Ion, J. Carreira, and C. Sminchisescu, "Image segmentation by figureground composition into maximal cliques," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2110–2117.

[28] [28] D. E. Jacobs, D. B. Goldman, and E. Shechtman, "Cosaliency: Where people look when comparing images," in *Proc. ACM Symp. User Interface Softw. Technol.*, 2010, pp. 219–228.

[29] [29] H.-T. Chen, "Preattentive co-saliency detection," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 1117–1120.

[30] [30] H. Li and K. N. Ngan, "A co-saliency model of image pairs," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3365–3375, Dec. 2011.

[31] [31] K.-Y. Chang, T.-L. Liu, and S.-H. Lai, "From co-saliency to Co-Segmentation: An efficient and fully unsupervised energy minimization model," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2011, pp. 2129–2136.

[32] [32] H. Fu, X. Cao, and Z. Tu, "Cluster-based co-saliency detection," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3766–3778, Oct. 2013.

[33] [33] K. R. Jerripothula, J. Cai, F. Meng, and J. Yuan, "Automatic image Co-Segmentation using geometric mean saliency," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 3282–3286.

[34] [34] R. Achanta *et al.*, "SLIC super pixels compared to state-of-the-art super pixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[35] [35] Y. Li, X. Hou, C. Koch, J. M. Rehg, and A. L. Yuille, "The secrets of salient object segmentation," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2014, pp. 280–287.

[36] [36] K. Tang, A. Joulin, L.-J. Li, and L. Fei-Fei, "Co-localization in realworld images," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2014, pp. 1464–1471.

[37] [37] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut: Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309-314, Aug. 2004.

[38] [38] M.-M. Cheng, V. A. Prisacariu, S. Zheng, P. H. Torr, and C. Rother, "DenseCut: Densely connected CRFS for realtime grabcut," *Comput. Graph. Forum*, vol. 34, no. 7, pp. 193-201, 2015.

[39] [39] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2013, pp. 1155–1162.

[40] [40] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *Proc. IEEE Comput. Vis. Pattern Recog.*, Jun. 2013, pp. 3166–3173.