Co-saliency Detection Based on Superpixel Matching and Cellular Automata

Zhaofeng Zhang, Zemin Wu*, Qingzhu Jiang, Lin Du, Lei Hu
College of Communication Engineering, PLA University of Science and Technology
Nanjing, Jiangsu 210007, P.R. China
[e-mail: zhangzhaofeng5@163.com]
*Corresponding author: Zemin Wu

Received September 3, 2016; revised January 3, 2017; accepted February 25, 2017; published May 31, 2017

Abstract

Co-saliency detection is a task of detecting same or similar objects in multi-scene, and has been an important preprocessing step for multi-scene image processing. However existing methods lack efficiency to match similar areas from different images. In addition, they are confined to single image detection without a unified framework to calculate co-saliency. In this paper, we propose a novel model called Superpixel Matching- Cellular Automata (SMCA). We use Hausdorff distance adjacent superpixel sets instead of single superpixel since the feature matching accuracy of single superpixel is poor. We further introduce Cellular Automata to exploit the intrinsic relevance of similar regions through interactions with neighbors in multi-scene. Extensive evaluations show that the SMCA model achieves leading performance compared to state-of-the-art methods on both efficiency and accuracy.

Keywords: co-saliency detection, multi-scene, superpixel, Hausdorff distance, Cellular Automata

https://doi.org/10.3837/tiis.2017.05.015 **ISSN : 1976-7277**

1. Introduction

Recently, with the development of multi-scene application, co-saliency detection[1],[2] aimed at locating the most important and similar objects in multi-scene has been more and more necessary and has been widely used in collaborative segmentation, multi-scene re-target recognition, image and video summaries, etc. The saliency detection methods of single image have developed rapidly and some of them are directly used in co-saliency detection. However, they are often ineffective when the object is integrated in multi-scene.

At present, there are only a few co-saliency detection methods. For example, Li et al.[3] employed a multi-scene structure model based on single image detection for co-saliency; however, their method was only applicable to two images but not to multiple images. Fu et al.[4] constructed co-saliency image through color contrast fusion, spatial distribution and consistency of images after extracting consistency information from multiple images by clustering colors. This method used simple color clustering at pixel level and lacked the structural information of the target, which can cause wrong detection of targets and background areas with the same colors. Liu et al.[5] calculated overall similarity of segmented areas at image level from color histogram. In this method, mismatching is prone to happen since the calculation merely based on the similarity of the characteristics of the segments. Cao et al.[6] obtained first saliency map of single image using multiple algorithms, and then co-saliency map using low-rank decomposition to fuse multiple saliency maps. Accuracy is highly affected by the algorithms of single image saliency detection in this method. In reference 114, in order to extract the characteristics of both target and background, Zhang et al.[1] comprehensively trained on similar and different images by extracting the characteristics in depth. The performance was yet to be improved because they only employed average color of CIELab in the matching of regional similarities between images. Later on, they applied polyinstantiation study and autoregulation study to find coincident targets; nevertheless, both study methods are hard to extend due to problems like high complexity and necessity of precise calibration of targeted training set. In summary, two problems exist for the abovementioned methods: lack of efficient methods for matching similar areas of images, and lack of a general framework applicable for co-saliency calculation.

In this paper, combining repeatability and saliency, we firstly propose a matching algorithm based on neighboring superpixel sets of Hausdorff distance to calculate the similarity between images and to depict the occurrence of area repetition. Secondly, inspired by Ref. [7], we propose a 2-Layer cellular automata model to calculate the saliency spread of infra-images and inter-images, in order to ensure complete saliency of targeted area. The result shows that the proposed superpixel matching algorithm is effective and that the multi-scene saliency spread algorithm demonstrates robustness, whose detection performance is superior to mainstream algorithm at present.

2. Superpixel matching

For convenience of calculations and intrinsic structural information, the image was firstly segmented into a set of superpixels by simple linear iterative clustering (SLIC) algorithm[8]. The core of detecting the common salient object is the superpixel matching in different images.

We define $\boldsymbol{I}=\{I_i\}_{i=1}^{n}$ as an image set consisting of n images. Each image I_i is segmented into a superpixel set $\boldsymbol{R}_i=\{r_{i,s}\}_{s=1}^{N_i}$, where N_i is the number of superpixels. In this paper, superpixel matching means, for any superpixel $r_{i,s}$ in image I_i, finding a set of superpixels with high similarity in another image I_j. Note that not all superpixel can be matched and one superpixel can have several matching superpixels in other images.

2.1 Single superpixel

Our aim is to find the matching superpixel of $r_{i,s}$. Briefly, we can calculate the similarity of all superpixels in I_j and set the similarity threshold δ to choose eligible ones as matching superpixels. This matching method is single superpixel matching. The feature similarity between any two of superpixels can be calculated according to Eq. (1),

$$w_{i,j}^{s,t}=\exp\left(-\frac{\left\|\boldsymbol{c}_{i,s}-\boldsymbol{c}_{j,t}\right\|^2}{2\sigma_c^2}\right) \tag{1}$$

where $\boldsymbol{c}_{i,s}$ is the average color value of $r_{i,s}$ in the CIELab color space. σ_c is the intensity for weight control. It is determined to be 10 as in reference[9] because it is not sensitive in the range [5,15].

As **Fig. 1** shows, the red regions of left image is input superpixel and the similarity threshold is set as $\delta=0.9$. The red regions in the upper row of the right images are matching results by single superpixel. Obviously, there are many mismatching results, which include regions on the object and background regions with similar color. The major cause for this error is exactly the above matching method. All superpixels are calculated without the spatial structure of other superpixels.

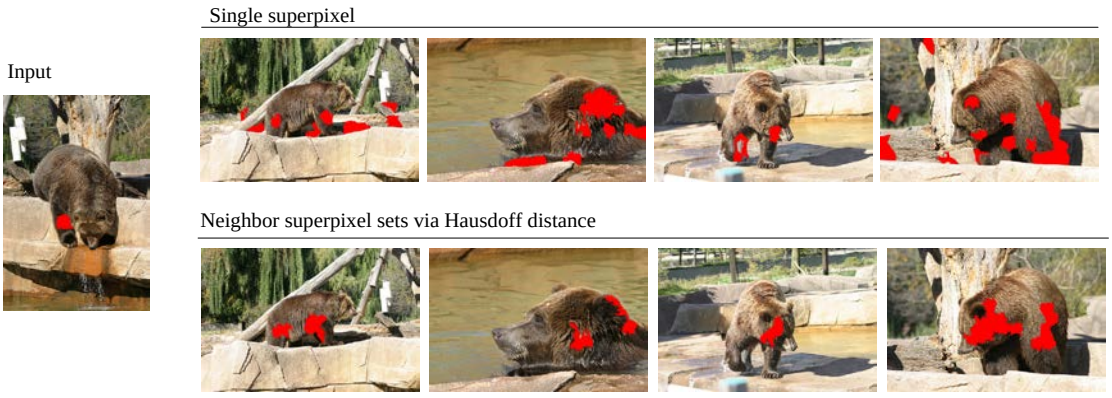


Fig. 1. Superpixel matching results of a test image (left) by single superpixel (upper row) and adjacent superpixel sets via Hausdorff distance (lower row)

2.2 Adjacent superpixel sets via Hausdorff distance

The description of a superpixel contains not only the inherent feature of its own internal pixels, but also the spatial structure of its adjacent superpixels. In detail, if the feature of a superpixel is similar to adjacent superpixels, the structure feature of this superpixel is smooth. On the

other hand, if the feature is entirely different, this superpixel is isolate. In addition, if the feature cuts both ways, this superpixel may be on the edges of the object.

Fig. 2 is an example of finding superpixel $r_{j,t}$ that matches $r_{i,s}$. Considering $N_{i,s}$, a neighbor of $r_{i,s}$, we can transform the solution to the calculation of Hausdorff distance between $N_{i,s}$ and $N_{j,t}$.

$$H(N_{i,s}, N_{j,t}) = \max(h(N_{i,s}, N_{j,t}), h(N_{j,t}, N_{i,s})) \tag{2}$$

where $H(N_{i,s}, N_{j,t})$ is Hausdorff distance between $N_{i,s}$ and $N_{j,t}$, and $h(N_{i,s}, N_{j,t})$ is the orientation-distance from $N_{i,s}$ to $N_{j,t}$. In this paper, Modified Hausdorff Distance (MHD)[10] is used.

$$h(N_{i,s}, N_{j,t}) = \frac{1}{|N_{i,s}|} \sum_{a \in N_{i,s}} d_{N_{j,t}}(a) \ , \ \ d_{N_{j,t}}(a) = \min_{b \in N_{j,t}} \left\| \boldsymbol{c}_a - \boldsymbol{c}_b \right\| \tag{3}$$

where $\boldsymbol{c}_a$ is the average color value of superpixel a in CIELab color space, and $|N_{i,s}|$ is the number of $N_{i,s}$. Now we can define the similarity between $N_{i,s}$ and $N_{j,t}$ as:

$$v_{i,j}^{s,t} = \exp\left(-\frac{H(N_{i,s}, N_{j,t})^2}{2\sigma_c^2} \right) \tag{4}$$

Take pixels whose $v_{i,j}^{s,t}$ is bigger than δ to constitute matching superpixel sets of $r_{i,s}$ as did in single superpixel matching. As the lower row of **Fig. 1** shows, compared with single superpixel matching, the adjacent superpixel sets matching can effectively eliminate some mismatching caused by isolate superpixel so that the matched superpixels are basically inside the object.

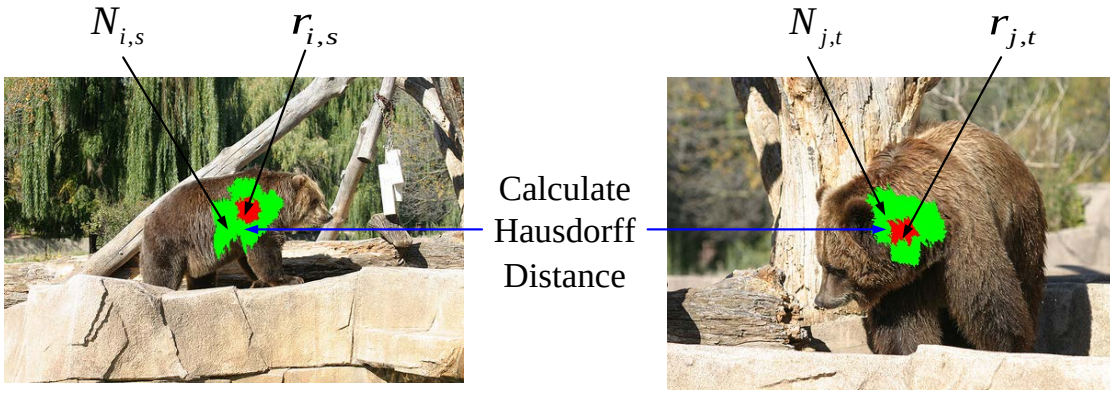


Fig. 2. Matching result by adjacent superpixel sets

3. Co-Saliency Detection via 2-Layer Cellular Automata

In Ref. [7], Cellular Automata method was proposed to calculate the saliency of a single image. The core concept of this method is that the saliency of one superpixel is affected by itself and the adjacent superpixels. All of the superpixels will converge after several times of spread. However, for co-saliency detection, as shown in **Fig. 3**, the saliency of one superpixel (red

spots) is not only affected by the adjacent superpixels (blue and yellow spots) but also affected by the matched superpixels in other images (green spots).

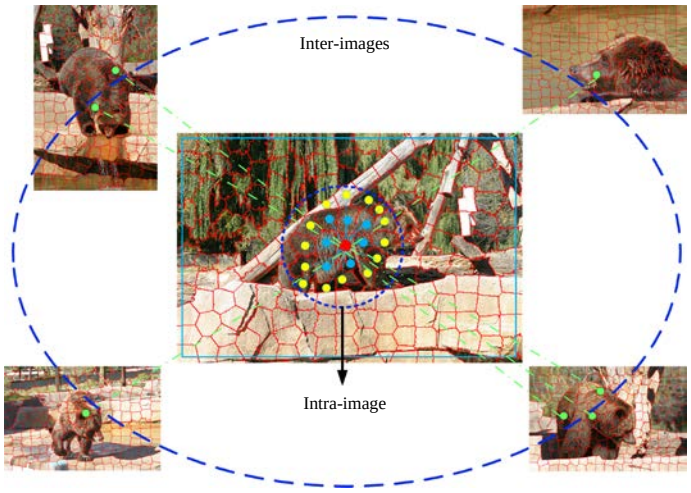


Fig. 3. Co-saliency detection model

According to this theory, we propose 2-layer Cellular Automata via intra image and inter images spread:

$$\boldsymbol{S}_i^{m+1} = (1-\lambda_1-\lambda_2)\boldsymbol{S}_i^m + \lambda_1 \boldsymbol{F}_i^{intra}\boldsymbol{S}_i^m + \lambda_2 \sum_{j=1, j\neq i}^{n} \boldsymbol{F}_{i,j}^{inter}\boldsymbol{S}_j^m \tag{5}$$

where $\boldsymbol{S}_i^m$ is the saliency of all superpixels in I_i after m times of status updates, $\boldsymbol{S}_i^0$ is the initial saliency, $\boldsymbol{F}_i^{intra}$ is the influence matrix of superpixels in I_i, $\boldsymbol{F}_{i,j}^{inter}$ is the influence matrix from I_j to I_i and λ_1 and λ_2 are impact factors. The state machine is stable after m times of status updates and the final $\boldsymbol{S}_i^m$ is the saliency of I_i. Note that, the initial saliency $\boldsymbol{S}_i^0$ can be obtained from single image saliency detection[11]-[25]. Here, MR method[16] was adopted to get $\boldsymbol{S}_i^0$. On the one hand, the saliency was obtained by image spread under superpixel level; on the other hand, image structural information used can be applied again to the calculation of intra-image influence.

3.1 Intra-image influence matrix

The similarity of intra-image superpixels is calculated by color similarity in CIELab color space. We define the initial intra-image influence matrix, $\hat{\boldsymbol{F}}_i^{intra} = [f_i^{s,t}]_{N_i \times N_i}$, to describe the influence among superpixels,

$$f_i^{s,t} = \begin{cases} \exp\left(-\dfrac{\left\|\boldsymbol{c}_{i,s}-\boldsymbol{c}_{i,t}\right\|^2}{2\sigma_c^2}\right) & t \in \mathcal{N}_{i,s} \\ 0 & t = s \text{ or others} \end{cases} \tag{6}$$

where $\mathcal{N}_{i,s}$ is the 2-layer adjacent region. As can be seen from **Fig. 3**, different from $N_{i,s}$,

$\mathcal{N}_{i,s}$ includes not only directly adjacent superpixels, shown as the blue spots, but also the indirectly adjacent superpixels, shown as the yellow spots, and the boundary superpixels, shown as the blue rectangular frame. Then the diagonal matrix is $\boldsymbol{D}_i^{intra} = diag\{d_1, d_2, ..., d_{N_i}\}$, where $d_i = \sum_t f_i^{s,t}$. Normalizing the initial intra-image influence matrix $\hat{\boldsymbol{F}}_i^{intra}$ leads to:

$$\boldsymbol{F}_i^{intra} = \left[\boldsymbol{D}_i^{intra}\right]^{-1} \cdot \hat{\boldsymbol{F}}_i^{intra} \tag{7}$$

3.2 Inter-images influence matrix

To focus on the influence of other images in the same image set, we firstly use the method introduced in section 3.2 to obtain $v_{i,j}^{s,t}$, then define inter-images influence matrix $\hat{\boldsymbol{F}}_{i,j}^{inter} = [\, f_{i,j}^{s,t}\,]_{N_i \times N_j}$ to describe the influence by any two of superpixels in different images,

$$f_{i,j}^{s,t} = \begin{cases} v_{i,j}^{s,t} & v_{i,j}^{s,t} > \delta, \\ 0 & others \end{cases} \tag{8}$$

where δ is saliency matching threshold. The parameter is set to be 0.9 according to our experience.

Similarly, diagonal matrix $\boldsymbol{D}_i^{intra} = diag\{d_1, d_2, ..., d_{N_i}\}$ is obtained, where $d_i = \sum_t f_{i,j}^{s,t}$. Normalize the initial matrix $\hat{\boldsymbol{F}}_i^{intra}$ considering the influence of the other $n-1$ images, and the final inter-images influence matrix is indicated as:

$$\boldsymbol{F}_i^{inter} = \frac{1}{n-1} \cdot \left[\boldsymbol{D}_i^{inter}\right]^{-1} \cdot \hat{\boldsymbol{F}}_i^{inter} \tag{9}$$

3.3 The termination condition for iteration

After confirming $\boldsymbol{F}_i^{intra}$, $\boldsymbol{F}_{i,j}^{inter}$ and the initial saliency map, $\{\boldsymbol{S}_i^0\}_{i=1}^n$, saliency of superpixels can be spread via iteration updates as Eq. 5. According to the experiment in Ref. [7], after 20 times of cellular automata iteration, the saliency of the superpixels stabilizes, that is, the system reaches stable status. However, it was found in our experiment that the system can sometimes terminate in advance. The difference of each image's saliency between two sequent updates can be used to determine the termination condition.

$$\Delta = \max_{i \in [1,n]} \left(\left\| \boldsymbol{S}_i^{m+1} - \boldsymbol{S}_i^{m} \right\|_2 \right). \tag{10}$$

When $\Delta < \varepsilon$ (where ε is the set minimum threshold), it is considered that the saliency does not change significantly anymore. Therefore, the termination condition for iteration is given as: when the saliency of each image does not change significantly anymore or when number of iterations reaches maximum M=20 (determined as in Ref. [7]).

3.4 Algorithm procedure and analysis

It can be seen from Eq. (5), Eq. (7) and Eq. (9) that a superpixel is primarily influenced by the saliency of adjacent infra-superpixels. If the superpixel has matching superpixels in most other parts of the image, its saliency is more influenced by other images; but if it has only a few or no

matching superpixels, its saliency becomes smaller and smaller in the iteration of Eq. (5). Thus, through iteration of 2-layer cellular automata, common superpixels in all the images have a higher saliency, while superpixels only existing in a few images have a lower saliency. Furthermore, common background superpixels can be effectively restrained due to the restriction of initial saliency (approaching 0). Co-saliency detection can be completed through this spread model.

Table 1 shows the overall procedure of our SMCA model, where lines 1-6 are to construct superpixel matching and image structure impact factor matrix and lines 7-12 are to spread saliency infra- and inter-images.

Table 1. The procedure of SMCA algorithm

Alg. SMCA
Input: image set $\boldsymbol{I} = \{I_t\}_{i=1}^{n}$
Output: co-saliency map $\{\boldsymbol{S}_i\}_{i=1}^{n}$
1. Segment each image I_i into superpixel sets $\boldsymbol{R}_i = \{r_{i,s}\}_{s=1}^{N_i}$ by SLIC algorithm.
2. Calculate the infra impact factor matrix $\boldsymbol{F}_i^{intra}$ of each image.
3. Calculate initial saliency map $\{\boldsymbol{S}_i^0\}_{i=1}^{n}$ using MR algorithm.
4. Construct adjacent superpixel sets $N_{i,s}$ for each superpixel $r_{i,s}$.
5. Calculate matching similarity $v_{i,j}^{s,t}$ of any two superpixels according to Eq. (2) and Eq. (4).
6. Construct inter-images impact factor matrix $\boldsymbol{F}_{i,j}^{inter}$.
7. while $\Delta > \varepsilon$ and $m < M$ do
8. for i=1 : n do
9. update saliency map of each image $\boldsymbol{S}_i^{m+1} = (1-\lambda_1-\lambda_2)\boldsymbol{S}_i^m + \lambda_1 \boldsymbol{F}_i^{intra}\boldsymbol{S}_i^m + \lambda_2 \sum_{j=1, j\neq i}^{n} \boldsymbol{F}_{i,j}^{inter}\boldsymbol{S}_j^m$;
10. end
11. calculate change of saliency $\Delta = \max_{i\in[1,n]}(\left\|\boldsymbol{S}_i^{m+1} - \boldsymbol{S}_i^m\right\|_2)$;
12. end

4. Experimental Evaluation

4.1 Parameters and Evaluation Metrics

Testing dataset: So far, there are only two common co-saliency datasets, ImgPair[3] and iCoseg[26], with the whole true value label at pixel level. iCoseg is the most commonly applied dataset in co-saliency measurement and co-segmentation algorithm. It contains 38 images sets of more than 5 images, with 634 images in total. ImgPair is also in widespread use. Every image set of ImgPair has two images and there are 210 images altogether. So all the algorithm measurements were based on these two datasets.

Algorithm comparison: Here comparisons were made among typical co-saliency detection algorithms from recent years: CB-C[4] (multi-scene algorithm based on color cluster), HS[5]

and SA[6] and typical single image saliency detection algorithms: CS-S[4] (single image algorithm based on color cluster), MR[16], RBD[22], HDCT[21] and RR[25].

Evaluation Measures: As is described in Ref. [27], three commonly applied performance metrics are adopted for our quantitative evaluation as follows:

Precision-recall (PR). To evaluate the final saliency map S, we can convert it to a binary mask M with a threshold sliding form 0 to 255 and compute Precision and Recall by comparing M with ground-truth G:

$$Precision = \frac{|M \cap G|}{M} \tag{11}$$

$$Recall = \frac{|M \cap G|}{G} \tag{12}$$

F-measure. In general, neither precision nor recall can comprehensively evaluate a saliency map. In other words, we both need the high precision and recall. To this end, the F-measure is put forward as the overall performance metrics:

$$F_{measure} = \frac{(1+\beta^2) \cdot Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \tag{13}$$

where $\beta^2 = 0.3$ as a non-negative weight which was proved in Ref. [28] to insure the precision.

Receiver operating characteristics (ROC) curve. Besides PR and F-measure, we can also report the false positive rate (FPR) and true positive rate (TPR) when binarizing the saliency map with a set of fixed thresholds:

$$TPR = \frac{|M \cap G|}{G} \tag{14}$$

$$FPR = \frac{|M \cap G|}{|M \cap G| + |\overline{M} \cap \overline{G}|} \tag{15}$$

Where $\overline{M}$ and $\overline{G}$ respectively denote the opposite of the binary mask M and ground-truth G. The ROC curve is the plot of TPR versus FPR by testing all possible thresholds.

Arear under ROC curve (AUC). While ROC represents the performance of a model as a two-dimensional curve, by calculating as the area under the ROC curve, the AUC distills this information into a single scalar.

Mean Absolute Error (MAE). The above evaluation measures do not consider the true negative saliency assignments, for example the pixels correctly marked as non-salient. For a comprehensive comparison, it is necessary to evaluate MAE which calculates the average absolute error between the saliency map S and the ground truth G of the image ($W \times H$ resolution) in pixel

$$MAE = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} |S(i,j) - G(i,j)| \tag{16}$$

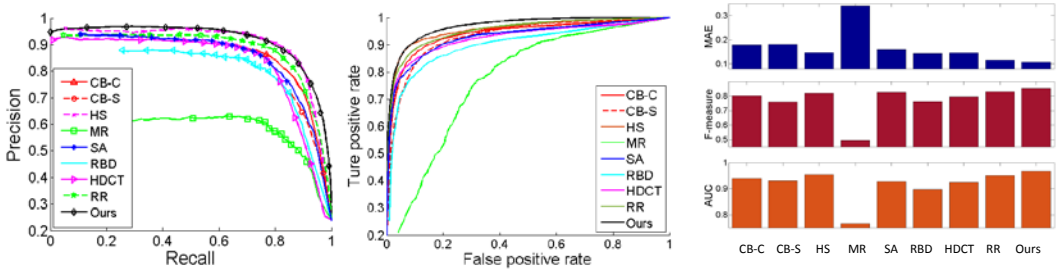
MAE score directly indicates how similar a saliency map S is to the ground truth G.

Parameter setting: SLIC superpixel segmentation number $N_i = 200$. Compactness weight value was set as 20. λ_1 and λ_2 in Eq. (5) were experiencedly chosen to be 0.5 and 0.3, respectively. Maximum iteration number $M = 20$. Threshold value for iteration termination $\varepsilon = 10^{-3}$. Threshold value for matching similarity $\delta = 0.9$.

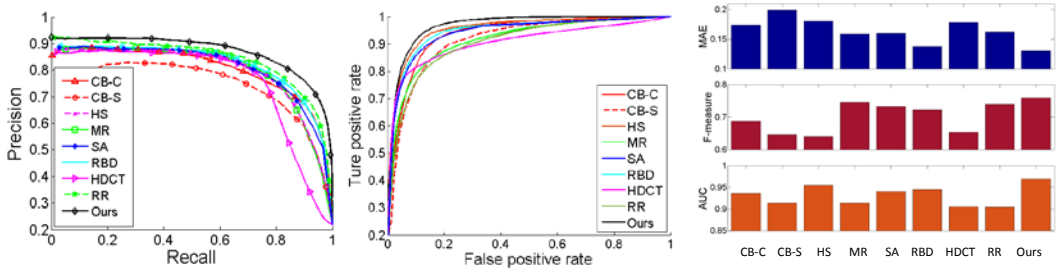
4.2 Method Comparisons

Firstly, several state-of-the-art methods are compared with our SMCA model. **Fig. 4(a)** and **Fig. 4(b)** respectively show the comparisons of PR cure, ROC cure, MAE, F-measure and AUC between the SMCA and other models on ImgPair and iCoseg. It can be seen from the figure that our SMCA model shows significant improvement in both recall rate and accuracy. The relatively high F-measure value indicates that the overall performance of the proposed SMCA model is superior to other algorithms and the relatively low MAE value indicates that our result is closer to true value.

Fig. 5(a) and **Fig. 5(b)** respectively show some examples of detection effectiveness based on ImgPair and iCoseg dataset. It is demonstrated that most co-saliency detection algorithm, such as HS, SA, and the proposed SMCA model have a better performance. Particularly, some common target areas can be highlighted. It is evident from row 2 that result of CB-C algorithm is the same color pixel value from different images as it only used color cluster at pixel level. It is especially sensitive to color difference cause by different illumination intensity and detection is not complete without considering structure properties of pixels. As for HS algorithm, mismatching shows up a lot because simple color matching at regional level is applied without regard to adjacent structure characteristics, resulting in mismatching of targets and background. SA algorithm generally performs better, but since the fusion of algorithm detection is the core, its performance highly depends on the effectiveness of algorithm detections. Take the second set of images in **Fig. 5(b)** as an example. Athletes in white are common saliency targets. Athletes in red are saliency targets in a few images but not co-saliency targets of the whole image set. Since a few images contain red pixel values, none of the three algorithms can effectively eliminate athletes in red which are not the co-saliency target. The proposed SMCA model utilizes adjacent superpixel matching algorithm, bettering detection effectiveness despite illumination difference, structure variation, background complexity, etc.

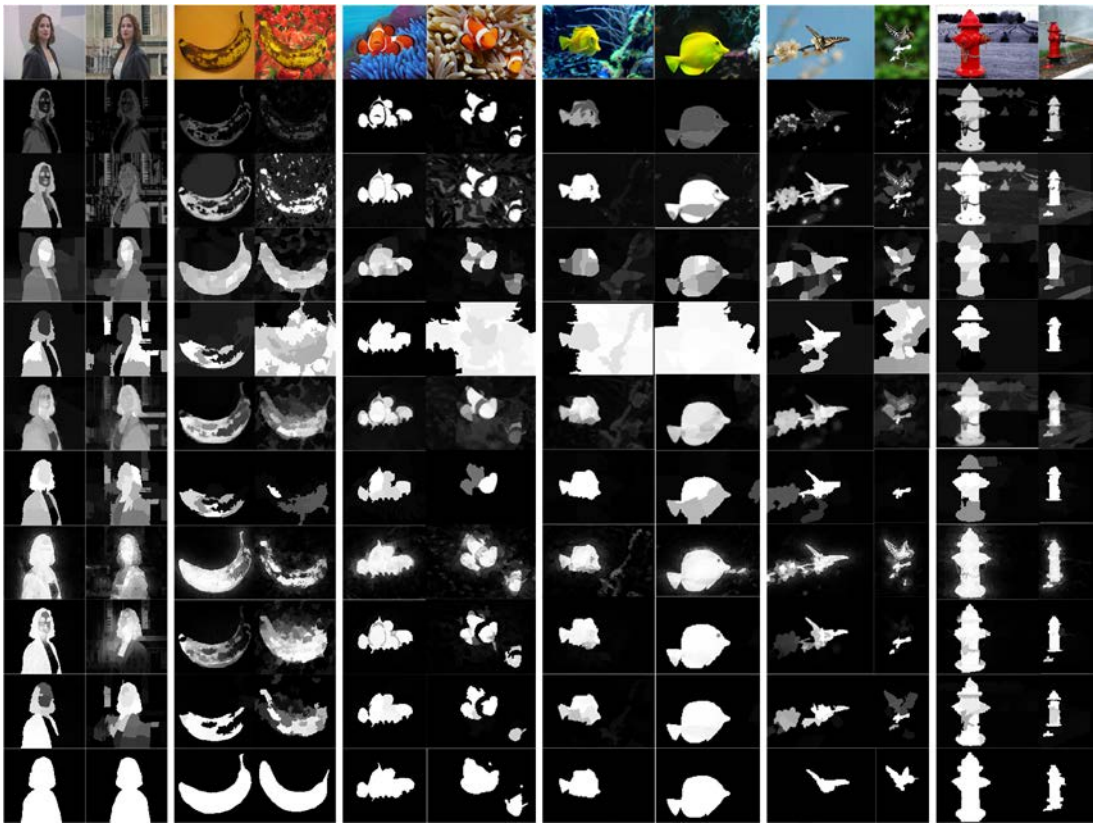


(a) Comparison results on the ImgPair dataset in terms of PR cure, ROC cure, MAE, F-measure and AUC

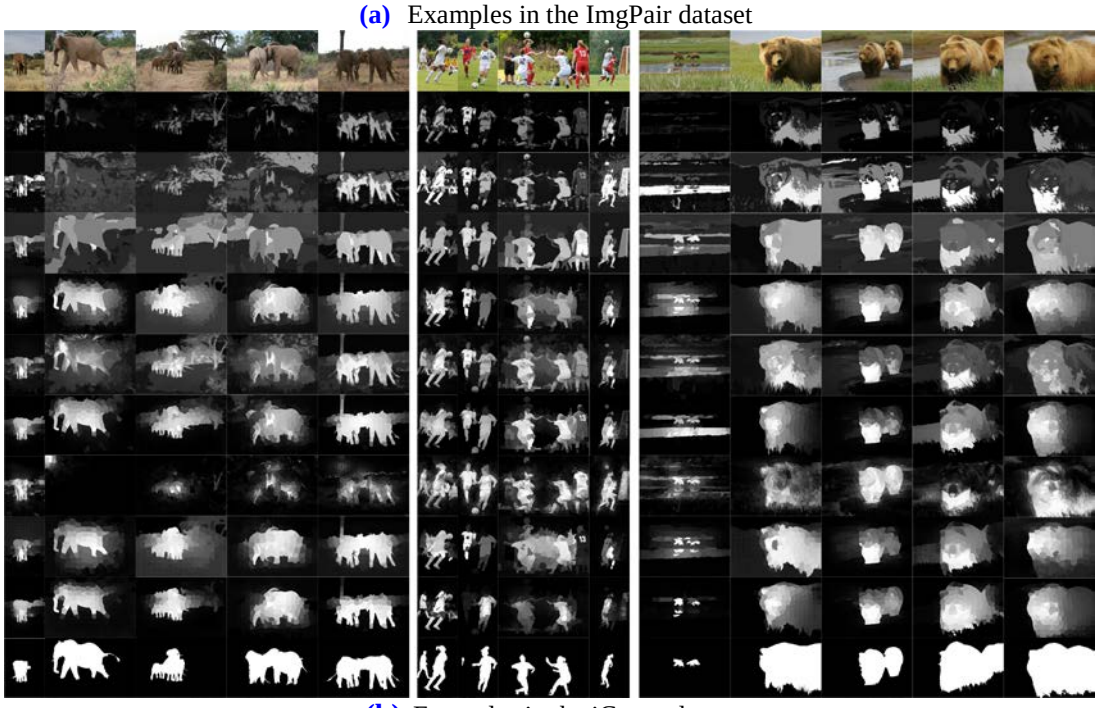


(b) Comparison results on the iCoseg dataset in terms of PR cure, ROC cure, MAE, F-measure and AUC

Fig. 4. Quantitative comparisons of the proposed method and other state-of-the-art methods



(a) Examples in the ImgPair dataset



(b) Examples in the iCoseg dataset

Fig. 5. Visual comparisons of different methods (from top to bottom: original image, CB-C, CB-S, HS, MR, SA, RBD, HDCT, RR, our proposed SMCA and true value)

4.3 Run time analysis

Table 2 shows the comparison of average run time between the proposed SMCA model and other co-saliency detection methods. The measurement environment was Intel Xeon E5-2620 2.00GHZ, RAM 16GB. All the algorithms directly operated the source code published by the author under Matlab 2012b platform. It is worth noting that the SA algorithm is a fusion algorithm. So its overall run time depends on all the other algorithms. The table only gives the time for fusion part but not the overall run time, thus relatively low. It can be seen from the table that the proposed SMCA algorithm has a similar complexity with CB-C algorithm, with run time much lower than HS algorithm. Further more, other models such as CB-C and RBD focus on single image saliency detection.

Table 2. Comparison of average run time between different co-saliency algorithms

Method	Run time(s)
CB-C	5.5
CB-S	2.1
HS	24.9
MR	1.7
SA	1.9
RBD	0.25
HDCT	4.9
RR	1.8
Ours	6.3

5. Conclusion

Based on Hausdorff distance, a novel algorithm using adjacent superpixel sets for superpixel matching is proposed. The result shows that matching effectiveness is superior to the traditional single superpixel method. Further on, impact factor matrix are constructed for infra- and inter-images, and 2-layer cellular automata spread model is constructed for multi-scene co-saliency detection. Experiment result indicates that our proposed SMCA model has significant advantages in both robustness and time complexity. In addition, different from other co-saliency detection model, the proposed 2-layer cellular automata is a uniform model combining similarity of infra-images and inter-images, with excellent extendibility.

As future work, we will consider to add more characteristics such as color histogram and textual features for more accurate superpixel matching. Moreover, only a few datasets can be used to test the co-saliency detection model. Therefore, we may focus on making a new dataset.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (NSFC) (Grant No.: 61501509).

References

[1] D. Zhang, J. Han, C. Li and J. Wang, "Co-saliency detection via looking deep and wide," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2994-3002, June 7-12, 2015. Article (CrossRef Link).

[2] D. Zhang, D. Meng, C. Li, L. Jiang, Q. Zhao and J. Han, "A Self-Paced Multiple-Instance Learning Framework for Co-Saliency Detection," in *Proc. of IEEE International Conf. on Computer Vision (ICCV)*, pp. 594-602, Dec. 13-16, 2015. Article (CrossRef Link).

[3] H. Li and K. N. Ngan, "A Co-Saliency Model of Image Pairs," *IEEE Transactions on Image Processing*, vol 20, no.12, pp.3365-3375, 2011. Article (CrossRef Link).

[4] H. Fu, X. Cao and Z. Tu, "Cluster-Based Co-Saliency Detection," *IEEE Transactions on Image Processing*, vol 22, no.10, pp.3766-3778, 2013. Article (CrossRef Link).

[5] Z. Liu, W. Zou, L. Li and L. Shen, "Co-Saliency Detection Based on Hierarchical Segmentation," *Signal Processing Letters*, vol 21, no.1, pp.88-92, 2014. Article (CrossRef Link).

[6] X. Cao, Z. Tao, B. Zhang and W Feng, "Self-adaptively Weighted Co-saliency Detection via Rank Constraint," *IEEE Transactions on Image Processing*, vol 23, no.9, pp.4175-4186, 2014. Article (CrossRef Link).

[7] Y. Qin, H. Lu, Y. Xu and H. Wang, "Saliency detection via Cellular Automata," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 110-119, June 7-12, 2015. Article (CrossRef Link).

[8] R. Achanta, A. Shaji, K. Smith, A. Lucchi , P. Fua and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 34, no.11, pp.2274-2282, 2012. Article (CrossRef Link).

[9] J. Kim, D. Han, Y. Tai and J. Kim, "Salient Region Detection via High-Dimensional Color Transform and Local Spatial Support," *IEEE Transactions on Image Processing*, vol 25, no.1, pp.9-23, 2016. Article (CrossRef Link).

[10] M. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," in *Proc. of the 12th IAPR International Conf. on Pattern Recognition*, pp. 566-568, October 9-13, 1994. Article (CrossRef Link).

[11] M. Cheng, G. Zhang, N. J. Mitra, X. Huang and S. Hu, "Global contrast based salient region detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 37, no.3, pp. 569 -582, 2016. Article (CrossRef Link).

[12] K. Chang, T. Liu, H. Chen and S. Lai, "Fusing generic objectness and visual saliency for salient object detection," in *Proc. of IEEE International Conf. on Computer Vision (ICCV)*, pp. 914-921, November 6-13, 2011. Article (CrossRef Link).

[13] F. Perazzi, P. Krähenbühl, Y. Pritch and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *Proc. of IEEE Conf. on Computer Vision and* Pattern Recognition (CVPR), pp. 733-740, June 16-21, 2012. Article (CrossRef Link).

[14] Y. Wei, F. Wen, W. Zhu and J. Sun, "Geodesic saliency using background priors," *Computer Vision–ECCV*, pp. 29-42, October 7-13, 2012.Article (CrossRef Link)

[15] Q. Yan, L. Xu, J. Shi and J. Jia, "Hierarchical saliency detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 1155-1162, June 23-28, 2013. Article (CrossRef Link).

[16] C. Yang, L. Zhang, H. Lu, X. Ruan and M. Yang, "Saliency detection via graph-based manifold ranking," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 3166-3173, June 23-28, 2013. Article (CrossRef Link).

[17] R. Margolin, A. Tal and L. Zelnik-Manor, "What makes a patch distinct?," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 1139-1146, June 23-28, 2013. Article (CrossRef Link).

[18] H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng and S. Li, "Salient object detection: A discriminative regional feature integration approach," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2083-2090, June 23-28, 2013. Article (CrossRef Link).

[19] X. Li, H. Lu, L. Zhang, X. Ruan and M. Yang, "Saliency detection via dense and sparse reconstruction," in *Proc. of IEEE International Conf. on Computer Vision (ICCV)*, pp. 2976-2983, December 1-8, 2013. Article (CrossRef Link).

[20] M. Cheng, J. Warrell, W. Lin, S. Zheng, V. Vineet and N. Crook, "Efficient salient region detection with soft image abstraction," in *Proc. of IEEE International Conf. on Computer Vision (ICCV)*, pp. 1529-1536, IEEE (2013). Article (CrossRef Link).

[21] J. Kim, D. Han, Y. Tai and J. Kim, "Salient region detection via high-dimensional color transform," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 883-890, June 24-27, 2014. Article (CrossRef Link).

[22] W. Zhu, S. Liang, Y. Wei and J. Sun, "Saliency optimization from robust background detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2814-2821, June 24-27, 2014. Article (CrossRef Link).

[23] R. Liu, J. Cao, Z. Lin and S. Shan, "Adaptive partial differential equation learning for visual saliency detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 3866-3873, June 24-27, 2014. Article (CrossRef Link).

[24] S. Lu, V. Mahadevan and N. Vasconcelos, "Learning optimal seeds for diffusion-based salient object detection," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2790-2797, June 24-27, 2014. Article (CrossRef Link).

[25] C. Li, Y. Yuan, W. Cai, Y. Xia, D. Feng, "Robust saliency detection via regularized random walks ranking," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2710-2717, June 7-12, 2015. Article (CrossRef Link).

[26] D. Batra, A. Kowdle, D. Parikh, J. Luo and T. Chen, "Interactively Co-segmentating Topically Related Images with Intelligent Scribble Guidance," *International Journal of Computer Vision*, vol 93, no.3, pp.273-292, 2011. Article (CrossRef Link).

[27] A. Borji, M. Cheng, H Jiang, and J. Li, "Salient Object Detection: A Survey," *Eprint Arxiv*, vol 16, no. 7, pp.3118-3213, 2014.Article (CrossRef Link).

[28] R. Achanta, S. Hemami, F. Estrada and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1597-1604, 20-25 June 2009. Article (CrossRef Link).

Zhaofeng Zhang is currently pursuing a master degree at PLA University of Science and Technology. His current research interests include image processing and video information processing.

Zemin Wu is a professor at PLA University of Science and Technology. He received the Ph.D. from PLA University of Science and Technology in 2002. His research interests include data integration and image processing.

Qingzhu Jiang received the M.E. degree from PLA University of Science and Technology in 2016. His current research interests include image processing and video quality evaluation.

Lin Du is currently pursuing a master degree at PLA University of Science and Technology. His current research interests include video transmission guarantee and video quality evaluation.

Lei Hu is a lecturer at PLA University of Science and Technology. He received the Ph.D. from PLA University of Science and Technology in 2013. His research interests include compressive sensing and video information processing.