IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 24, 2018, accepted May 6, 2018, date of publication May 14, 2018, date of current version June 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2835826

A Saliency Map Fusion Method Based on Weighted DS Evidence Theory

BING-CAI CHEN[1,2], (Member, IEEE), XIN TAO[1], MAN-ROU YANG[1], CHAO YU[1], WEI-MIN PAN[2], AND VICTOR C. M. LEUNG[3], (Fellow, IEEE)

[1]School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China
[2]School of Computer Science and Technology, Xinjiang Normal University, Urumqi 830054, China
[3]Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Corresponding author: Xin Tao (taoxin510@mail.dlut.edu.cn) and Chao Yu (cy496@dlut.edu.cn)

This work was supported by the Natural Science Foundation of China under Grants 61771089, 61671169, and 61502072.

ABSTRACT In this paper, we propose a weighted Dempster–Shafer (DS) evidence theory-based fusion algorithm to take advantages of state-of-the-art salient object detection methods. First, we define the mass function value for each saliency detection method to be fused at the pixel level, based on which we further calculate the similarity coefficient and similarity matrix. The credibility of each saliency detection method will be computed by considering to what degree it is supported by other saliency detection methods. Second, using the credibility of each image saliency detection method as the weight, we compute the weighted mass function value of each method and get a saliency map. Third, we use the synthetic rules of DS evidence theory to fuse the weighted mass function values and get the other saliency map. The final saliency map will be obtained by fusing the aforementioned two saliency maps. Extensive experiments on three publicly available benchmark datasets demonstrate the superiority of the proposed weighted DS evidence theory-based fusion model against each individual saliency detection algorithm in terms of three evaluation metrics of precision-recall rate, F-measure, and average absolute error. The saliency map after fusion utilizing weighted DS evidence theory is closer to the ground-truth map.

INDEX TERMS Salient object detection, DS evidence theory, fusion algorithm, mass function, pixel level.

I. INTRODUCTION

Salient object detection, which aims to detect the most important part of the input image, has become more and more popular in the field of image processing [1]. It functions as an important pre-processing step in many computer vision tasks, such as content-aware video resizing [2], visual tracking [3], [4], image segmentation [5]–[7], image/video retrieval [8], [9], video summarization [10], [11], image categorization [12], to name a few.

Existing saliency detection methods can be generally divided into two categories: top-down methods and bottom-up methods. Top-down methods [13]–[20] are task-driven, *i.e.,* they are usually designed for a specific task. Specifically, they learn various characteristics of the target in a supervised manner and apply the learned information to detect the object. One main drawback is that they can not generalize well from the specific task that is originally trained for to another. In contrast, bottom-up methods [21]–[25] are data-driven. These methods usually rely on low-level visual cues such as colors, spatial distances, texture features and so forth. Besides, they usually follow strong assumptions, for example, the boundary priority [26], [27], which assume that regions along the image boundary are highly possible to be the background. Another typical assumption is center-surround contrast principle [23], [28], [29]. Bottom-up methods are widely adopted as they are easy and fast to compute saliency regions for subsequent image processing. However, as they rely on low-level features and strong assumptions, it is very likely to misclassify the salient object in noisy backgrounds.

We note each existing state-of-the-art salient object detection algorithm has its own advantages and disadvantages, which makes them unreliable in many situations. Therefore, it is important to develop a more robust method, in order to obtain more accurate saliency maps. In this paper, we exploit the superiority in the fusion of the uncertain events of Dempster-Shafer(DS) Evidence theory [30] and propose a fusion based saliency detection algorithm to take advantages of existing methods. Experiments demonstrate the effectiveness of the proposed method.

II. RELATED WORK

Recently, more and more saliency detection fusion algorithms have been applied to fuse different saliency maps in order to get a saliency map which is closer to the ground-truth map. And the fusion effect has been gradually optimized. For example, Li *et al.* [31] first obtain two initial saliency maps using dense sparse reconstruction method and then utilize the Bayesian algorithm to fuse the two initial saliency maps to get a more accurate saliency map. Mai *et al.* [32] use a Conditional Random Field framework to fuse different methods and achieve good results. Qin *et al.* [26] directly take advantages of the results obtained by other saliency detection algorithms and treat the results obtained by each algorithm as a layer cellular automaton. They use a multi-layer cellular automata to fuse a variety of saliency detection algorithms which achieves good results. Inspired by above ideas and considering extensive applications of DS Evidence theory in the multi-source information fusion, we try to adopt DS Evidence theory to fuse the results of several saliency detection algorithms to get better results.

DS Evidence theory is originally based on the works of Dempster which models the uncertainty by a probability interval rather than a single probability value. The results show its superiority in the fusion of uncertainty. DS Evidence theory has been widely applied in the field of medical diagnosis, target detection, military command and so on. For example, Liu [33] applies DS Evidence theory to vehicle target detection, in which they use DS Evidence theory to integrate three characteristics of vehicles to detect whether the corresponding area contains vehicles. However, DS Evidence Theory has not been applied in the field of saliency detection till today. We compare the results of multiple saliency detection algorithms to multiple features and introduce weighted DS Evidence theory into the fusion of several saliency detection algorithms.

In the classic DS Evidence theory, the fusion is based on that all evidences have the same weight. However, for some decision in reality, the credibility of each evidence is different in most cases. So Murphy presented a modified model [30] which still retain the synthetic rules of the classic DS Evidence theory. As in the field of saliency detection, different saliency detection algorithms perform differently at different scenarios, and the credibility of the results obtained by various saliency detection algorithms differs at different pixel, so that the weighted DS Evidence theory is appropriate for with the scenarios of saliency detection.

III. SALIENCY MAP FUSION BASED ON WEIGHTED DS EVIDENCE THEORY

In this section, we will apply DS Evidence theory to image saliency detection. We fuse the saliency maps obtained by multiple saliency detection algorithms at the level of pixel(in the following sections the operation is at the pixel level if there is no special instruction) and produce a saliency map which is closer to the ground-truth map.

The overall flow of the algorithm is as follows:

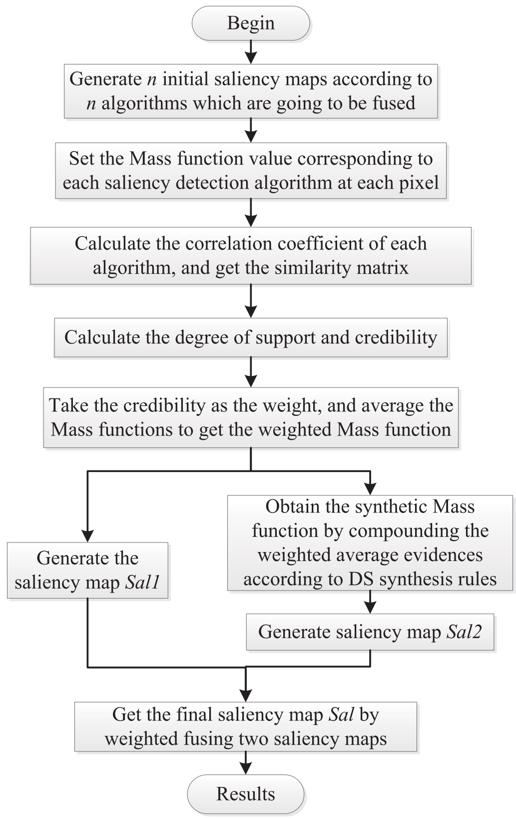


FIGURE 1. The flowchart of the algorithm.

A. INITIAL MASS FUNCTION

In this paper, we first define the corresponding recognition framework of DS Evidence theory. For each pixel in the image, we define the environment as $\Theta = \{FG, BG\}$ which is a collection of mutually exclusive and exhaustive elements. It means that a pixel either belongs to foreground or background. Where FG means the pixel belongs to the foreground and BG means the pixel belongs to the background. The recognition framework contains 2^2 subsets, which define the power sets. These subsets represent four cases that the pixel is neither foreground nor background, the pixel is the foreground, the pixel is the background, and the case that whether the pixel is the foreground or the background is uncertain. In this paper, the power set 2^Θ is defined as

$$2^\Theta = \{\phi, \{FG\}, \{BG\}, \{FG \cup BG\}\} \tag{1}$$

For the mass function, it satisfies:

$$\begin{cases} m(\phi) = 0 \\ \sum\limits_{X \in 2^\Theta} m(X) = 1 \end{cases} \tag{2}$$

Where X represents a subset of the recognition framework and $m(X)$ represents a Mass function of event X which indicates the degree of trust for X in the evidence.

The Mass function at each pixel is defined as

$$m_i(FG) = p_i \tag{3}$$

$$m_i(BG) = 1 - p_i \tag{4}$$

$$m_i(\phi) = m_i(FG \cup BG) = 0 \tag{5}$$

Where m_i represents the Mass function corresponding to the i-th saliency detection algorithm. FG indicates that the pixel to be fused is the foreground and BG indicates that the pixel is determined to be the background. ϕ indicates that the corresponding pixel is neither foreground nor background and $FG \cup BG$ represents that whether the pixel is foreground or background is uncertain. p_i represents the saliency value of the i-th saliency detection method at the corresponding pixel.

B. GENERATION OF WEIGHTED MASS FUNCTION AND INITIAL SALIENCY MAP

In the synthesis rules of former DS Evidence theory, the calculation is usually performed based on the same weight of each evidence. However in the process of saliency detection, the saliency values of a certain pixel calculated by different saliency detection algorithms tend to be different. And sometimes the difference is very obvious. It can be said that for a certain pixel, the reliability of the results obtained by different saliency detection algorithms is disparate. So it is more reasonable and effective to perform weighted synthesis on evidences. We refer to the method proposed by Murphy for the generation of a weighted Mass function.

1) SIMILARITY COEFFICIENT AND SIMILARITY MATRIX

At a certain pixel, we assume that the Mass functions of the saliency detection algorithms i and j correspond to m_i and m_j respectively. And we represent these two saliency detection algorithms as evidences E_i and E_j. The similarity coefficient d_{ij} between evidence E_i and E_j can be expressed as:

$$d_{ij} = \frac{\sum\limits_{A_x \cap B_y \neq \phi} m_i(A_x) m_j(B_y)}{\sqrt{(\sum m_i^2(A_x))(\sum m_j^2(B_y))}} \tag{6}$$

Where the similarity coefficient d_{ij} ($d_{ij} \in [0, 1]$) is used to describe the similarity between the results of the saliency detection algorithms E_i and E_j. The larger the value of d_{ij} is, the more similar the results of the two algorithms are. When $d_{ij} = 1$, it indicates that the results of the saliency detection algorithms E_j and E_j are completely consistent. When $d_{ij} = 0$, it means that the saliency detection algorithms E_i and E_j have the opposite conclusions. Here A_x and B_y are the focal elements (The concept is stated in the appendix), and the values of which can be BG or FG. When they take the value of BG, it indicates that the saliency detection algorithm determines the pixel as the background, while the value of FG indicates that the saliency detection algorithm determines the pixel as the foreground.

Assuming that the number of saliency detection algorithms to be fused is n, we can give the similarity matrix between saliency detection algorithms by Eq.7. The similarity matrix S is defined as

$$S = \begin{bmatrix} 1 & d_{12} & \cdots & d_{1n} \\ d_{21} & 1 & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & 1 \end{bmatrix} \tag{7}$$

2) SUPPORT DEGREE AND CREDIBILITY

We add each row of the similarity matrix S to obtain the support degree to the saliency detection algorithm E_i that supported by other saliency object detection algorithms. It is defined as

$$Sup(m_i) = \sum_{j=1}^{n} d_{ij} \quad (i, j = 1, 2, \cdots, n) \tag{8}$$

Where $Sup(m_i)$ indicates the support degree of the saliency detection algorithm E_i which is supported by other saliency detection algorithms. If the results obtained by a saliency detection algorithm are similar to those obtained by other saliency detection algorithms, we consider the degree of their mutual support is higher, otherwise it is believed that the degree of their mutual support is lower.

The credibility of the evidences can be obtained by normalizing the support degree of the evidence, and the credibility can be defined as

$$Crd(m_i) = \frac{Sup(m_i)}{\sum\limits_{i=1}^{n} Sup(m_i)} \tag{9}$$

The credibility $Crd(m_i)$ reflects the degree of which the results obtained by saliency detection algorithm E_i can be trusted. Where $\sum\limits_{i=1}^{n} Crd(m_i) = 1$. The higher the support degree of a saliency detection algorithm which is supported by other saliency detection algorithms, the more credible the result of the saliency detection algorithm is. Conversely, the less reliable the result is.

3) WEIGHTED MASS FUNCTION VALUE

Taking the credibility of each saliency detection algorithm as the weight, we use the method proposed by Murphy to weight each saliency detection algorithm to get the weighted Mass function. The weighted Mass function $m_{ave}(FG)$ is defined as

$$m_{ave}(FG) = \sum_{i=1}^{n} Crd(m_i) \cdot m_i(FG) \tag{10}$$

Where m_i is the basic probability assignment (The concept is stated in the appendix) of the i-th saliency detection algorithm. And FG means that the i-th saliency detection algorithm judge the pixel belongs to the foreground. Here we take the weighted Mass function value as the saliency value of the saliency map to get a preliminary composite saliency map Sal_1. This saliency map can effectively preserve the foreground area. It is defined as

$$Sal_1 = m_{ave}(FG) \tag{11}$$

C. SYNTHESIS RULES OF WEIGHTED MASS FUNCTION

The synthesis rule of Mass function in DS evidence theory is as follows

$$m(X) = m_1 \oplus m_2 \oplus \cdots \oplus m_n(X) = \frac{\sum_{x_1 \cap x_2 \cdots \cap x_n = X} m_1(x_1)m_2(x_2)\cdots m_n(x_n)}{k} \tag{12}$$

$$k = 1 - \sum_{x_1 \cap x_2 \cdots \cap x_n = \phi} m_1(x_1)m_2(x_2)\cdots m_n(x_n) \tag{13}$$

In this paper, we use the DS synthesis rules to fuse the weighted Mass functions obtained by n saliency detection algorithms and synthesize the weighted Mass function for $n-1$ times to get the synthetic Mass function as follows

$$m(FG) = \frac{(m_{ave}(FG))^n}{k} \tag{14}$$

$$k = (m_{ave}(FG))^n + (1 - m_{ave}(FG))^n \tag{15}$$

D. GENERATION OF THE FINAL SALIENCY MAP

In this paper, we use the following formula to calculate the confidence function Bel and likelihood function Pls after fusing saliency detection algorithms. Concepts of Bel and Pls are stated in the appendix.

$$Bel(FG) = \sum_{X \subseteq FG} m(X) = m(FG) \tag{16}$$

$$Pls(FG) = 1 - \sum_{X \subseteq BG} m(X) = m(FG) \tag{17}$$

Where $X \in 2^{\Theta}$. From Equation (16) and Equation (17), we can get that at a certain pixel, the confidence function value is the same as the likelihood function value after fusing different algorithms using the DS evidence theory in this paper. In this paper, we define the other saliency map Sal_2 which highlights the foreground area as

$$Sal_2 = Bel(FG) \tag{18}$$

We fuse two saliency maps obtained from Equation (11) and Equation (18), and then we can get the final saliency map Sal. It is defined as

$$Sal = \mu_1 \times Sal_1 + \mu_2 \times Sal_2 \tag{19}$$

where μ_1 and μ_2 are weight values of synthesis. In this paper, we empirically set $\mu_1 = 0.35$ and $\mu_2 = 0.65$ to highlights the foreground area.

The main process of our proposed fusion algorithm is summarized in Algorithm 1.

Algorithm 1 Saliency Map Fusion Method Based on DS Evidence Theory

Input: Several saliency maps of different image salient object detection algorithms

Output: a fusion saliency map with the same size as input images.

1: Define the recognition framework of DS evidence theory with Eq.1.
2: Define the initial Mass function at each pixel with Eq.3, Eq.4 and Eq.5.
3: Calculate similarity coefficient between different algorithms with Eq.6.
4: Acquire the similar matrix S of saliency maps with Eq.7.
5: Acquire the degree being supported and the credibility of different algorithms with Eq.8 and Eq.9.
6: Acquire weighted Mass function value $m_{ave}(FG)$ and a saliency map Sal_1 with Eq.10 and Eq.11.
7: Use the DS synthesis rules to fuse the weighted Mass function $m_{ave}(FG)$ and get the synthetic Mass function $m(FG)$ with Eq.14 and Eq.15.
8: Calculate the confidence function Bel and likelihood function Pls and get saliency map Sal_2 with Eq.16, Eq.17, and Eq.18.
9: Fuse two saliency maps Sal_1 and Sal_2 to get the final saliency map Sal with Eq.19.

IV. EXPERIMENT

In this paper, we extensively present evaluation and analysis of the proposed DS Evidence Theory saliency model against the state-of-the-art methods on three standard salient object databases with the labeled ground truth. Considering the cases that natural images generally fall into, we introduce the dataset ECSSD [34] which cantains 1000 semantically meaningful but structurally complex images with pixel-level saliency labeling and the dataset DUT-OMRON [35] which are used to compare models on a large scale. We also introduce the dataset MSRA10K [36] which is a descendant of the MSRA dataset [37]. It contains 10000 annotated images that covers all the 1000 images in the popular ASD dataset [38].

In this paper, We firstly fuse several state-of-the-art methods including BSCA [26], DSR [31], HS [35], wCO [39], MR [35] and RCRR [40] to get a saliency map of fusion. After that we evaluate the saliency map of fusion with these state-of-the-art methods. In the following experiments, we find that on all the tested databases, our model can yield comparable results when compared with the best baselines.

A. EVALUATION OF SALIENCY MAP MODELS

1) EXPERIMENTAL PARAMETERS AND EVALUATION CRITERIA

Following most existing works, we use the precision-recall(PR) curve, the F-measure and the mean absolute error(MAE) to evaluate all algorithms. For a saliency map, we can convert it to a binary mask M and compute Precision and Recall by comparing M with ground-truth GT. The specific implementation is to quantify the saliency map to [0, 255], and then set a threshold for every five values. In a saliency map, if the saliency value of a pixel is larger than the threshold, it indicates that the pixel belongs to the foreground, otherwise the pixel belongs to the background. On each threshold, a pair of precision/recall scores are computed, and are finally combined to form a precision-recall(PR) curve

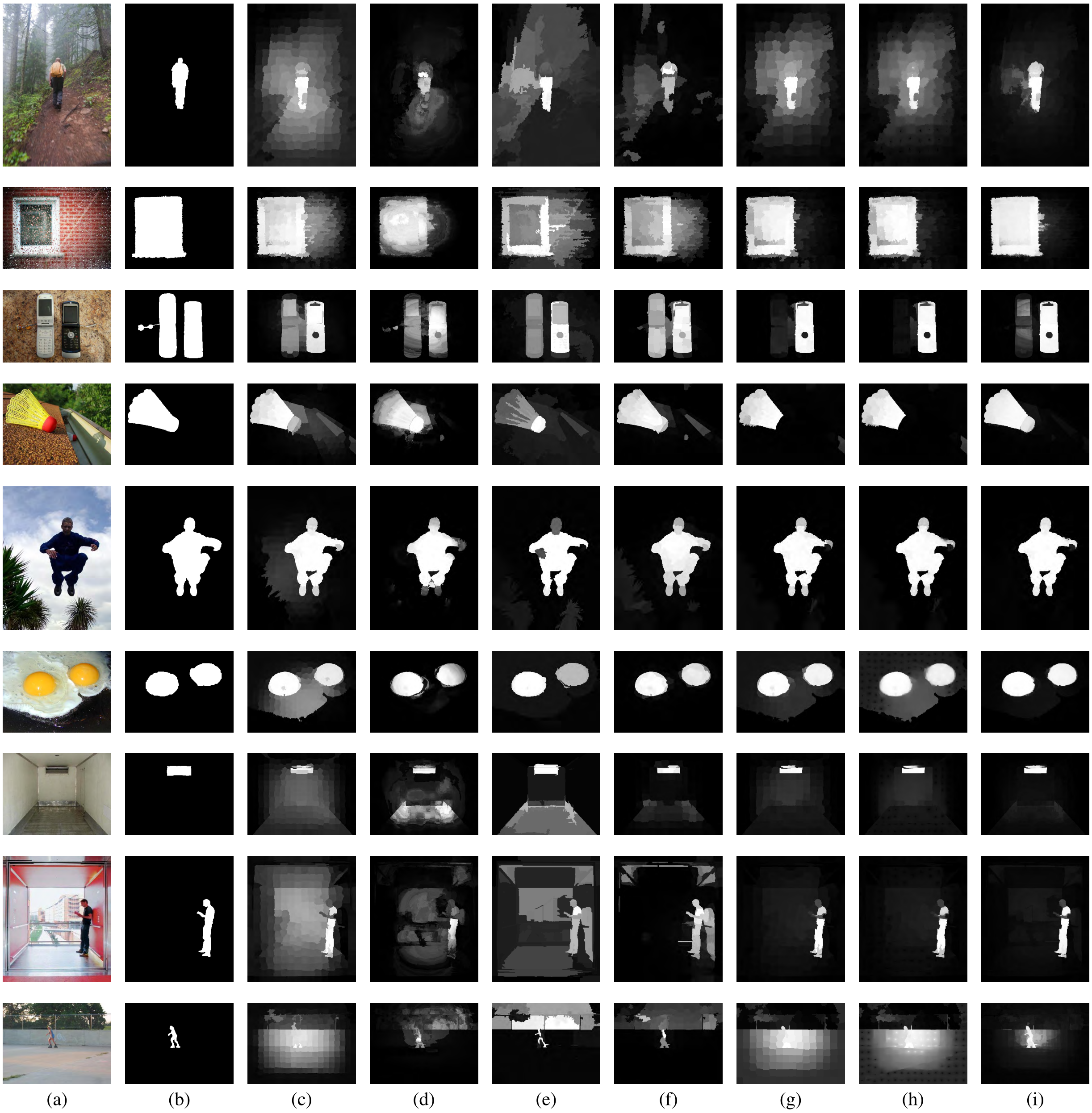


FIGURE 2. Visual comparison among different methods on DUT-OMRON(5168), MSRA10K(10000) and ECSSD(1000) databases. (a) Input. (b) GT. (c) BSCA. (d) DSR. (e) HS. (f) wCO. (g) MR. (h) RCRR. (i) Ours.

to describe the model performance at different situations. Precision and recall can be calculated by

$$precision = \frac{|M \cap GT|}{|M|}, \ recall = \frac{|M \cap GT|}{|GT|} \tag{20}$$

Usually, neither precision nor recall can comprehensively evaluate the quality of a saliency map. In this paper, we use F-measure as the overall performance evaluation, according to [28], the specific calculation method is:

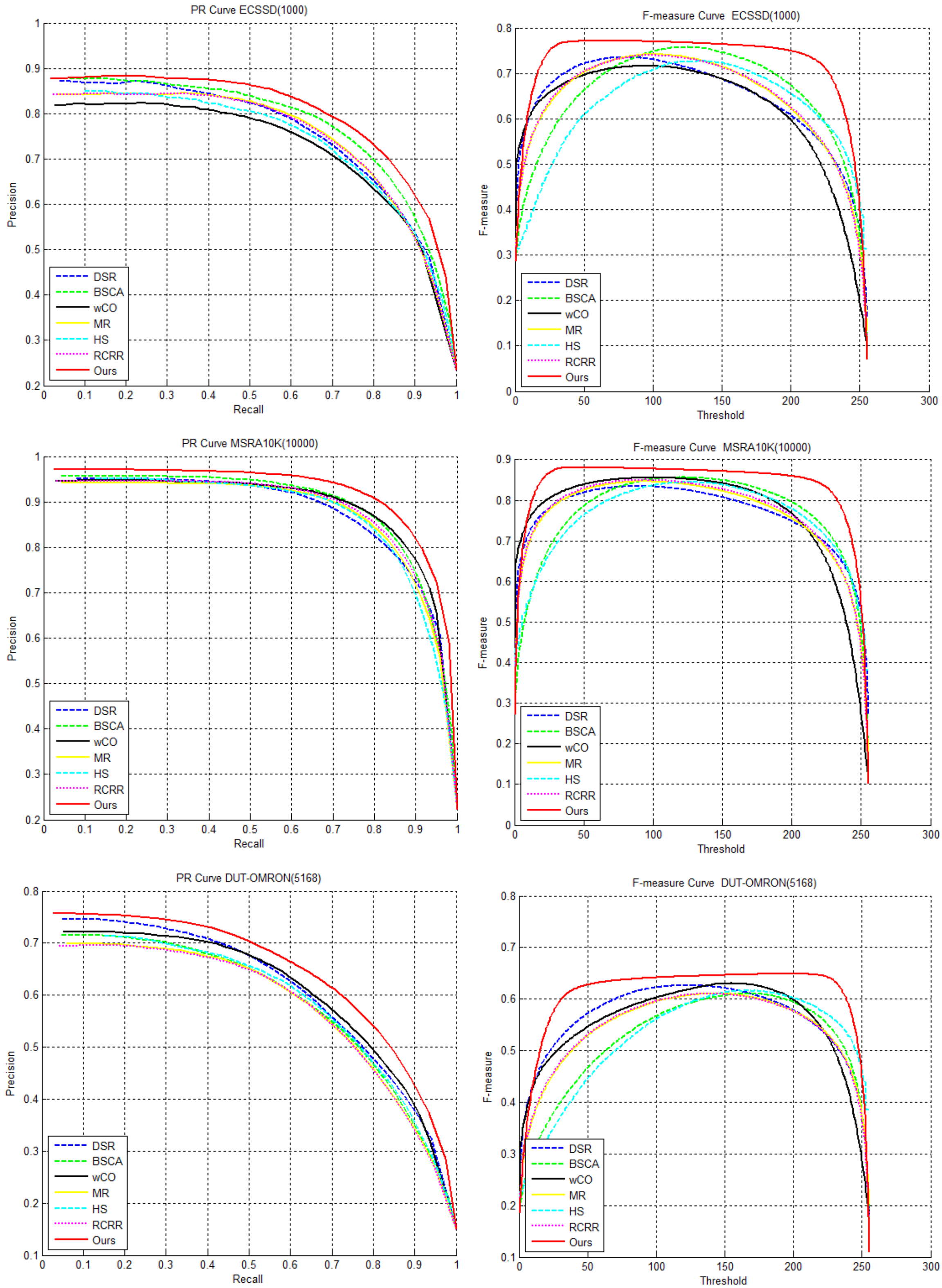
$$F_\beta = \frac{(1+\beta^2)precision \cdot recall}{\beta^2 \cdot precision + recall} \tag{21}$$

If β^2 is set to 1, then the result of the F-measure is the harmonic mean of both the precision and the recall. If $\beta^2 > 1$, then the recall has a greater impact on the results of the F-measure. If $\beta^2 < 1$, then the precision has a greater impact on the results of the F-measure. As suggested by many salient object detection works(e.g., [21], [38], [41], [42]), we set β^2 as 0.3 to raise more importance to the precision value.

We also introduce the mean absolute error (MAE) which indicates the average difference between the saliency map and the ground-truth at pixel level. The specific calculation method is

$$MAE = \frac{1}{H}\sum_{h=1}^{H}|S(h) - GT(h)| \tag{22}$$

Where S is saliency map, GT is the ground-truth map, H is the total number of pixels, and h denotes each pixel.



(a) (b)

FIGURE 3. Precision-Recall curve and F-measure curve of different algorithms. From top to bottom: ECSSD, MSRA10K, DUT-OMRON. (a) Precision-Recall curve. (b) F-measure curve.

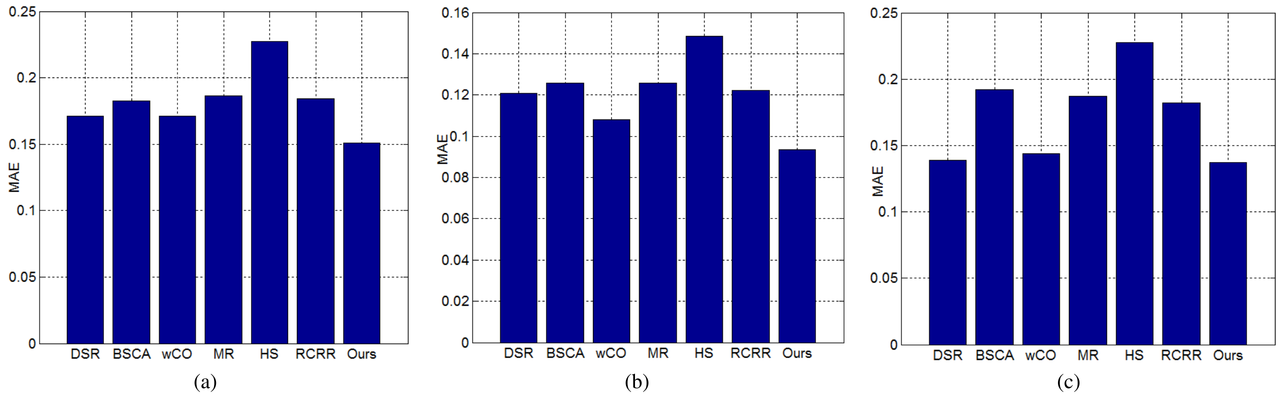


FIGURE 4. Mean absolute error of different algorithms.(a) ECSSD. (b) MSRA10K. (c) DUT-OMRON.

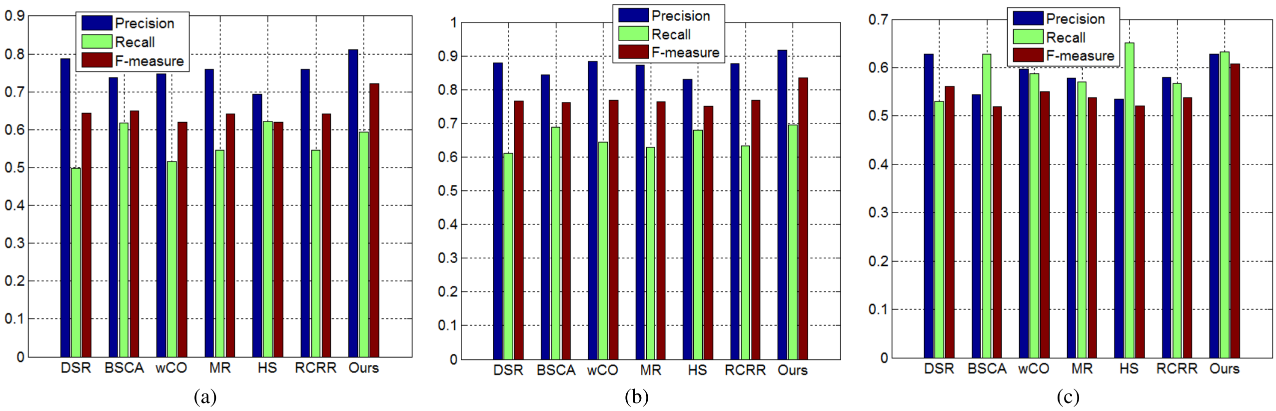


FIGURE 5. Mean precision, recall and F-measure values of different algorithms. (a) ECSSD. (b) MSRA10K. (c) DUT-OMRON.

This measure indicates how similar a saliency map is to the ground-truth map.

2) COMPARISON WITH EXISTING ALGORITHMS

In order to illustrate the universality of the algorithm to the saliency detection, we makes a comparison between our method and other six algorithms which are fused in three published databases. These six methods are BSCA, DSR, HS, wCO, MR and RCRR. Fig.2 shows some samples for qualitative comparison to several classic approaches, from which we can see that our saliency maps of fusion achieves the best performance on these images. The test images of the top three rows are selected from ECSSD(1000) database. The center three rows are from MSRA10K(10000) database. And images of the last three rows belong to the DUT-OMRON(5168) database. It can be seen from the figure that after being fused by our DS Evidence Theory, saliency maps will be closer to ground-truth maps among all the candidates. Our fusion method can perform very well, no matter there are multiple objects in the image or the scrambled backgrounds and heterogeneous foregrounds.

For quantitative evaluation, we plot the precision-recall curves and F-measure curves in Fig.3, mean absolute error bar in Fig.4 and mean precision, recall and F-measure values bar in Fig.5 to compare our method with six state-of-the-art approaches. As observed from Fig.3, our model gets the highest precision value in almost the whole recall interval [0, 1].

TABLE 1. Elapsed time of the proposed method on ECSSD, MSRA10K and DUT-OMRON databases.

	ECSSD	MSRA10K	DUT-OMRON
Total time(s)	978.076	10524.813	4724.046
Number of images(i)	1000	10000	5168
Average time(s/i)	0.978	1.052	0.914

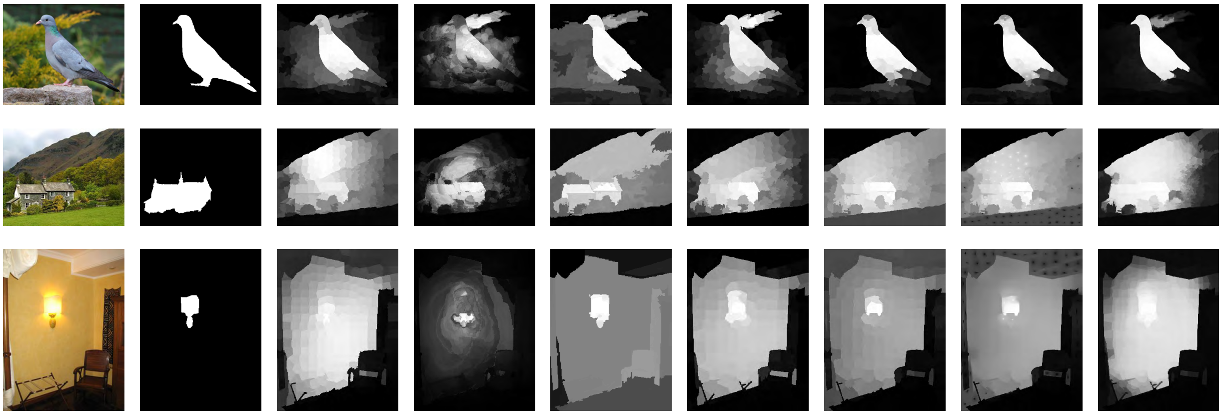
And the F-measure curves of our fusion algorithm in Fig.3 are fixed at high values which are insensitive to almost every selective thresholds. In Fig.4, we can see that our approach has the lowest mean absolute error compared to six state-of-the-art approaches on different datasets. And our fusion algorithm also performs well on mean precision, recall and F-measure as observed from Fig.5. The result clearly demonstrate that our method is effective in fusing salient object detection methods.

3) RUNTIME ANALYSIS

The experiments are run on the single thread of an Intel Core i5-2450M CPU of 2.50GHz with 4 GB RAM and the code are written in MATLAB. The time-consuming of fusing BSCA, DSR, HS, wCO, MR and RCRR six algorithms using the proposed method is shown in Table 1.

B. FAILURE CASES

Our model performs favorably against algorithms to be fused with higher precision, recall and F-measure. And we also have the lowest mean absolute error. However, as our DS



(a) (b) (c) (d) (e) (f) (g) (h) (i)

FIGURE 6. **Failure cases of saliency map models.**

Evidence Theory fuses other state-of-the-art methods and then get a saliency map. Our method depends on methods being fused. The proposed method does not work well if more than half of methods to be fused get wrong answers at one pixel. On the other hand, the other saliency map models do not perform well in such cases as shown in Fig.6.

V. CONCLUSION

In this paper, we study the fusion of multiple saliency detection algorithms in the field of image salient object detection. For the first time, we introduce the DS Evidence theory into the image salient object detection at the pixel level. We fuse the results of multiple salient object detection algorithms at the pixel level. We conduct a large number of experiments on three public datasets DUT-OMRON, ECSSD and MSRA10K. We compare saliency maps of the fusion with maps of several algorithms that are fused. The experimental results show that our fusion algorithm is better than algorithms that are fused in the precision, recall, F-measure and mean absolute error indicators. All of these fully prove the effectiveness and robustness of our fusion algorithm. In the future work, we will consider using DS Evidence theory to fuse multiple saliency detection algorithms at the feature level for better fusion effect.

APPENDIX

To further understand how we use DS evidence theory in the saliency detection research area. We would like to give some basic concepts of this paper including the recognition framework, the mass function, the trust function, the likelihood function and the law of synthesis of DS evidence theory.

Definition 1 (Recognition Framework): In DS evidence theory, we use Θ to represent the environment, which is a collection of mutually exclusive and exhaustive elements. We use $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ to represent the set of topics of interest. In general, a set of size N consists of 2^N subsets, which define the power set, denoted as 2^Θ. All subsets of the recognition framework are exhaustive and disjoint.

Definition 2 (Mass Function): In DS evidence theory, the Mass function is defined as the basic probability assignment (BPA). Each Mass function can be formally represented as a function of the mapping of each element in the power set to the interval [0, 1]. It is often referred to as the basic probability assignment function ($BPAF$). It is defined as $m : 2^\Theta \rightarrow [0, 1]$. Under normal circumstances, $BPAF$ need to meet the following two conditions:

(1) The Mass function of an empty set is defined as 0. It is $m(\phi) = 0$.

(2) The sum of the Mass functions for all subsets of the power set 2^Θ in Θ is 1. It is defined as $\sum_{X\in 2^\Theta} m(X) = 1$.

Any subset of power set 2^Θ is called the focal element if its Mass function value is greater than 0.

Definition 3 (Trust Function $Bel(X)$ and Likelihood Function $Pls(X)$): For any proposition set X, the trust function $Bel(X)$ and the likelihood function $Pls(X)$ are defined as

$$Bel(X) = \sum_{Y\subseteq X} m(Y) \tag{23}$$

$$Pls(X) = 1 - Bel(\overline{X}) = 1 - \sum_{Y\subseteq \overline{X}} m(Y) \tag{24}$$

where $Bel(X)$ is the total belief between a proposition set and all its subsets. It is the smallest trust based on the evidence. The likelihood (Pls) of a set X is defined as the degree to which X is not opposed. It is the maximum trust based on the evidence. $[Bel(X), Pls(X)]$ represents the evidence interval of X.

Definition 4 (The Synthetic Rule of DS Evidence Theory): Assuming that Bel_1 and Bel_2 are two trust functions on the same discriminant frame Θ, m_1 and m_2 are their corresponding basic probability distribution functions respectively, then the synthesis law of DS evidence theory can be expressed as

$$m(Z) = m_1(Z) \oplus m_2(Z) = \frac{\sum_{X_i\cap Y_j=Z} m_1(X_i)m_2(Y_j)}{k} \tag{25}$$

$$k = 1 - \sum_{X_i \cap Y_j = \phi} m_1(x_i) m_2(Y_j) \tag{26}$$

where $X_i (i = 1, 2, \ldots, m)$ and $Y_j (j = 1, 2, \ldots, n)$ are focal elements and $Z \subseteq \Theta$.

REFERENCES

[1] A. Borji, M.-M. Cheng, H. Jiang, and J. Li, "Salient object detection: A benchmark," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5706–5722, Dec. 2015.

[2] Y. Niu, F. Liu, X. Li, and M. Gleicher, "Warp propagation for video resizing," in *Proc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2010, vol. 238, no. 6, pp. 537–544.

[3] T. Cane and J. Ferryman, "Saliency-based detection for maritime object tracking," in *Proc. Comput. Vis. Pattern Recognit. Workshops*, 2016, pp. 1257–1264.

[4] V. Mahadevan and N. Vasconcelos, "Saliency-based discriminant tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 1007–1013.

[5] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut': Interactive foreground extraction using iterated graph cuts," in *Proc. ACM SIGGRAPH*, 2004, pp. 309–314.

[6] M. Donoser, M. Urschler, M. Hirzer, and H. Bischof, "Saliency driven total variation segmentation," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep. 2010, pp. 817–824.

[7] Z. Liu, R. Shi, L. Shen, Y. Xue, K. N. Ngan, and Z. Zhang, "Unsupervised salient object segmentation based on kernel density estimation and two-phase graph cut," *IEEE Trans. Multimedia*, vol. 14, no. 4, pp. 1275–1289, Aug. 2012.

[8] Z. Xu, C. Hu, and L. Mei, "Video structured description technology based intelligence analysis of surveillance videos for public security applications," *Multimedia Tools Appl.*, vol. 75, no. 19, pp. 12155–12172, 2016.

[9] Z. Xu, L. Mei, C. Hu, and Y. Liu, "The big data analytics and applications of the surveillance system using video structured description technology," *Cluster Comput.*, vol. 19, no. 3, pp. 1283–1292, 2016.

[10] Q.-G. Ji, Z.-D. Fang, Z.-H. Xie, and Z.-M. Lu, "Video abstraction based on the visual attention model and online clustering," *Signal Process., Image Commun.*, vol. 28, no. 3, pp. 241–253, 2012.

[11] Y. J. Lee, J. Ghosh, and K. Grauman, "Discovering important people and objects for egocentric video summarization," in *Proc. Comput. Vis. Pattern Recognit.*, 2012, pp. 1346–1353.

[12] C. Siagian and L. Itti, "Rapid biologically-inspired scene classification using features shared with visual attention," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 12–300, Feb. 2007.

[13] B. Alexe, T. Deselaers, and V. Ferrari, "What is an object?" in *Proc. Comput. Vis. Pattern Recognit.*, 2010, vol. 23, no. 3, pp. 73–80.

[14] G. Li and Y. Yu, "Visual saliency detection based on multiscale deep CNN features," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5012–5024, Nov. 2016.

[15] H. Lu, X. Zhang, J. Qi, N. Tong, X. Ruan, and M.-H. Yang, "Co-bootstrapping saliency," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 414–425, Jan. 2017.

[16] K. Fu, I. Y.-H. Gu, and J. Yang, "Saliency detection by fully learning a continuous conditional random field," *IEEE Trans. Multimedia*, vol. 19, no. 7, pp. 1531–1544, Jul. 2017.

[17] H. Du, Z. Liu, H. Song, L. Mei, and Z. Xu, "Improving RGBD saliency detection using progressive region classification and saliency fusion," *IEEE Access*, vol. 4, pp. 8987–8994, Dec. 2016.

[18] Q. Hou, M.-M. Cheng, X. Hu, A. Borji, Z. Tu, and P. H. S. Torr, "Deeply supervised salient object detection with short connections," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 5300–5309.

[19] P. Zhang, D. Wang, H. Lu, H. Wang, and B. Yin, "Learning uncertain convolutional features for accurate saliency detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 212–221.

[20] F. Huang, J. Qi, H. Lu, L. Zhang, and X. Ruan, "Salient object detection via multiple instance learning," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1911–1922, Apr. 2017.

[21] Y. Xie, H. Lu, and M.-H. Yang, "Bayesian saliency via low and mid level cues," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1689–1698, May 2013.

[22] N. Tong, H. Lu, Y. Zhang, and X. Ruan, "Salient object detection via global and local cues," *Pattern Recognit.*, vol. 48, no. 10, pp. 3258–3267, 2015.

[23] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2007, pp. 1–8.

[24] F. Han, A. Han, and J. Hao, "Saliency detection method using hypergraphs on adaptive multiscales," *IEEE Access*, Jan. 2018, doi: 10.1109/ACCESS.2018.2797880.

[25] K.-Y. Chang, T.-L. Liu, H.-T. Chen, and S.-H. Lai, "Fusing generic objectness and visual saliency for salient object detection," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 914–921.

[26] Y. Qin, H. Lu, Y. Xu, and H. Wang, "Saliency detection via cellular automata," in *Proc. Comput. Vis. Pattern Recognit.*, 2015, pp. 110–119.

[27] A. Borji, D. N. Sihite, and L. Itti, *Salient Object Detection: A Benchmark*. Berlin, Germany: Springer, 2012.

[28] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[29] 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, Oct. 2012.

[30] F. Yang, *Combination Method of Conflictive Evidences in D-S Evidence Theory*. Beijing, China: National Defense Industry Press, 2010.

[31] X. Li, H. Lu, L. Zhang, X. Ruan, and M.-H. Yang, "Saliency detection via dense and sparse reconstruction," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2976–2983.

[32] L. Mai, Y. Niu, and F. Liu, "Saliency aggregation: A data-driven approach," in *Proc. Comput. Vis. Pattern Recognit.*, 2013, pp. 1131–1138.

[33] P. Liu, "Research on vehicle detection and tracking in vehicle active safety," Ph.D. dissertation, College Comput. Sci. Technol., Jilin Univ., Changchun, China, 2015.

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

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

[36] M. Cheng. *Thur 15000*. [Online]. Available: http://mmcheng.net/gsal/

[37] T. Liu, J. Sun, N.-N. Zheng, X. Tang, and H.-Y. Shum, "Learning to detect a salient object," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2007, pp. 1–8.

[38] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 1597–1604.

[39] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," in *Proc. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2814–2821.

[40] Y. Yuan, C. Li, J. Kim, W. Cai, and D. D. Feng, "Reversion correction and regularized random walk ranking for saliency detection," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1311–1322, Mar. 2018.

[41] M.-M. Cheng, G.-X. Zhang, N. J. Mitra, X. Huang, and S.-M. Hu, "Global contrast based salient region detection," in *Proc. Comput. Vis. Pattern Recognit.*, 2011, pp. 409–416.

[42] F. Perazzi, P. Krähenbühl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *Proc. Comput. Vis. Pattern Recognit.*, 2012, pp. 733–740.

BING-CAI CHEN received the M.S. degree and the Ph.D. degree in information and communication engineering from the Harbin Institute of Technology, Harbin, China, in 2003 and 2007, respectively. He joined The University of British Columbia, Canada, as a Visiting Scholar, in 2015.

He is currently an Associate Professor with the School of Computer Science and Technology, Dalian University of Technology, Dalian, China, and also with the School of Computer Science and Technology, Xinjiang Normal University, Urumqi, China. His current research interests include computer vision, wireless ad hoc network, and network and information security.

Dr. Chen received the National Science Foundation Career Award of China in 2009. He is serving as a Reviewer for project proposals to the National Science Foundation of China, Ministry of Education of China. He is also serving as a Reviewer for some refereed journals, including the IEEE/ACM Transactions on Networking, the *Journal of Electronics*, and the *Journal of Communication*.

XIN TAO received the B.E. degree in computer science and technology from Shihezi University, Shihezi, China, in 2015. He is currently pursuing the master's degree in computer science and technology with the Dalian University of Technology, Dalian, China. His current research interest is saliency detection.

MAN-ROU YANG is currently pursuing the master's degree in computer science and technology with the Dalian University of Technology, Dalian, China. Her current research interests include wireless and ad hoc network.

CHAO YU received the M.Sc. degree in computer science and technology from Huazhong Normal University, Wuhan, in 2010, and the Ph.D. degree in computer science from the University of Wollongong, Wollongong, NSW, Australia, in 2013.

He is currently an Associate Professor with the School of Computer Science and Technology, Dalian University of Technology, Dalian, China. His current research interests include multiagent systems and learning, with their wide applications in modeling and solving various real-world problems.

WEI-MIN PAN is currently a Professor with the School of Computer Science and Technology, Xinjiang Normal University, Urumqi, China. His current research interests include wireless ad hoc network.

VICTOR C. M. LEUNG (S'75–M'89–SM'97–F'03) received the B.A.Sc. degree (Hons.) in electrical engineering from The University of British Columbia (UBC) in 1977 and the Ph.D. degree in electrical engineering from the Graduate School, UBC, in 1981. He received the Natural Sciences and Engineering Research Council Postgraduate Scholarship from the Graduate School, UBC. He was the Head of the Graduating Class, Faculty of Applied Science, where he received the APEBC Gold Medal.

From 1981 to 1987, he was a Senior Member of Technical Staff and a Satellite System Specialist with MPR Teltech Ltd., Canada. In 1988, he joined the Department of Electronics, The Chinese University of Hong Kong, as a Lecturer. He returned to UBC as a Faculty Member in 1989, where he is currently a Professor and the TELUS Mobility Research Chair in advanced telecommunications engineering with the Department of Electrical and Computer Engineering. He has co-authored over 900 technical papers in international journals and conference proceedings and 31 book chapters and co-edited 11 book titles. Several of his papers had been selected for best paper awards. His research interests are in the areas of wireless networks and mobile systems.

Dr. Leung is a registered Professional Engineer with the Province of British Columbia, Canada. He is a fellow of the Royal Society of Canada, The Engineering Institute of Canada, and The Canadian Academy of Engineering. He was a Distinguished Lecturer of the IEEE Communications Society. He is a member of the editorial boards of the IEEE Wireless Communications Letters, the IEEE Journal on Selected Areas in Communications Series on Green Communications and Networking, the IEEE Access, the *Computer Communications*, and several other journals, and previously served on the editorial boards for the IEEE Journal on Selected Areas in Communications C Wireless Communications Series, the IEEE Transactions on Wireless Communications, the IEEE Transactions on Vehicular Technology, the IEEE Transactions on Computers, and the *Journal of Communications and Networks*. He was a recipient of the IEEE Vancouver Section Centennial Award and the 2012 UBC Killam Research Prize. He has guest edited many journal special issues and provided leadership to the organizing committees and technical program committees of numerous conferences and workshops.