IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 20, 2018, accepted April 24, 2018, date of publication May 9, 2018, date of current version June 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2834545

Hypergraph Optimization for Salient Region Detection Based on Foreground and Background Queries

JINXIA ZHANG[1], SHIXIONG FANG[1], KRISTA A. EHINGER[2], HAIKUN WEI[1], WANKOU YANG[1], KANJIAN ZHANG[1], AND JINGYU YANG[3]

[1]Key Laboratory of Measurement and Control of CSE, Ministry of Education, School of Automation, Southeast University, Nanjing 210096, China
[2]Centre for Vision Research, York University, Toronto, ON M3J 1P3, Canada
[3]School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

Corresponding author: Haikun Wei (hkwei@seu.edu.cn)

This work was supported in part by the National Natural Science Fund of China under Grant 61233011, Grant 61374006, Grant 61473086, Grant 61463008, Grant 61703100, Grant 61773118, and Grant 61773117, in part by the Major Program of National Natural Science Foundation of China under Grant 11190015, in part by the Natural Science Foundation of Jiangsu under Grant BK20131300 and Grant BK20170692, in part by the Innovation Fund of Key Laboratory of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, Nanjing University of Science and Technology, under Grant JYB201601, in part by the Innovation Fund of Key Laboratory of Measurement and Control of Complex Systems of Engineering, Southeast University, under Grant MCCSE2017B01, and in part by the Fundamental Research Funds for the Central Universities under Grant 2242016k30009.

ABSTRACT Graph-based methods have been widely adopted to detect salient objects in images. However, there are two limitations of these methods. First, only one kind of query is employed for saliency propagation on the graph. Second, these methods only represent pairwise relations between vertices and thus give an incomplete representation of the relationships between image regions. In this paper, we propose a foreground- and background-queries-based hypergraph optimization framework for salient region detection. In this framework, both foreground queries and background queries are explicitly exploited to uniformly highlight the salient foreground and suppress the non-salient background. Furthermore, to include both the pairwise and the higher-order relations among two or more vertices, a probabilistic hypergraph is constructed based on local spatial correlation, global spatial correlation, and color correlation to represent the relations among image regions from different views. Extensive experimental results demonstrate the effectiveness of the proposed framework.

INDEX TERMS Hypergraph, queries, optimization framework, salient region detection.

I. INTRODUCTION

Visual saliency detection has attracted a great deal of attention and achieved major progress during the past few years due to its application to computer vision tasks, including image compression [1], image quality assessment [2], image segmentation [3], object recognition [4], and content based image retrieval [5]–[7].

The literature of visual saliency detection is vast [8]–[18]. It can be roughly categorized into unsupervised methods and supervised methods. While supervised methods are able to automatically combine various features and in general achieve better performance than unsupervised approaches, a large amount of sample data with pixel-wise ground truth annotations are needed and it is very expensive to carry out the training process. A fast saliency detection technique can be an important preprocessing step for many computer vision tasks.

In recent years, unsupervised graph-based saliency detection methods have gained much popularity. These methods model each input image as a graph and propagate saliency information via weighted edges of the graph connecting image regions. The graph-based saliency propagation approaches achieve competitive performances. At the same time, they do not require labor-intensive annotated samples and expensive training processes. However, there are two major limitations of these methods. Firstly, only one kind of query is explicitly employed in one optimization step. Since salient region detection aims to separate the salient object foreground from the non-salient background, both the foreground queries and the background queries are critical

for the final saliency detection result. Secondly, representing the relations among image regions in a pairwise simple graph gives an incomplete representation of the image. It is important to consider not only the pairwise relation between two vertices but also local grouping information among two or more vertices [19]. These higher-order relations can be represented by a hypergraph, a generalization of a simple graph [20]. In a hypergraph, hyperedges are used to describe complex relations among any number of vertices.

In this work, we introduce a novel foreground- and background-queries-based hypergraph optimization framework to compute the saliency values (the likelihood belonging to the salient foreground) of different image regions for salient region detection. In the optimization framework we represent similarity relationships among image regions in multiple feature spaces and explicitly employ both the foreground queries and the background queries to uniformly highlight all of the salient regions and suppress the non-salient background. These two kinds of queries are acquired based on a hypergraph-based ranking algorithm using the boundary prior of visual saliency and the local grouping information among multiple vertices.

Our second contribution is constructing a probabilistic hypergraph for an input image. In the proposed probabilistic hypergraph, vertices are defined as homogeneous regions in the input image and different correlations defined based on the distances between regions (in image space or feature space) are used to model local grouping information among various image regions and compute the probability that an vertex belongs to a hyperedge. The probabilistic hypergraph is constructed according to three different kinds of correlations: local spatial correlation, global spatial correlation, and color correlation, representing the relations among multiple vertices from different views through multiple types of hyperedges. According to local spatial correlation, each vertex in the hypergraph serves as the centroid for a local spatial hyperedge which connects the centroid vertex to its local spatial neighbors which share a boundary with the centroid vertex in the input image. According to global spatial correlation, each vertex serves as the centroid of a global spatial hyperedge which connects the centroid vertex to the global border vertices which are located on the four borders of the input image. Based on color correlation, each vertex serves as the centroid of a color hyperedge which connects the centroid vertex to its neighbors in the color space. In our work, the probability that a vertex belongs to a hyperedge is defined to be the similarity between this vertex and the corresponding centroid vertex of the hyperedge.

As we will demonstrate in experiments, the proposed hypergraph optimization framework outperforms state-of-the-art methods on various databases. The remainder of this paper is organized as follows: related works are introduced in Section II; the foreground- and background-queries-based hypergraph optimization framework is described in Section III; Section IV describes how we construct a probabilistic hypergraph for an input image; Section V elaborates the method to acquire the foreground and background queries; Section VI illustrates our testing procedure and experimental comparison. Finally, the conclusion of the paper is made in Section VII.

II. RELATED WORK

Our work is related to graph-based ranking methods as in [12] and [16], hypergraph-based approaches as in [21], boundary prior modeling as in [22], and queries acquisition as in [12].

A. SUPERVISED VS. UNSUPERVISED

Unsupervised saliency detection methods [12], [16], [21], [23] aim at separating the salient foreground objects from the non-salient background based on visual cues from the input image only. Various visual cues have been proven to be useful for salient region detection, including contrast [10], background prior [24], objectness [25], etc. Without the need for training, unsupervised saliency detection methods can be conveniently integrated into various computer vision applications. By comparison, supervised methods [25]–[28] acquire knowledge from the pixel-wise ground truth annotations. Recent progress in deep learning methods show promising saliency detection results using supervised methods on public benchmark databases. However, it is very expensive to set up the learning framework, collect the hand-marked images and operate the training procedure.

B. GRAPH-BASED

In recent years, graph-based saliency detection methods have increased in popularity [12]–[16], [18]. These methods do not need manually-annotated image samples and overcome a major flaw of classical pixel-level saliency detection methods: namely, these methods tend to highlight object boundaries instead of object interiors. Graph-based methods [12]–[16], [18] employ a pairwise simple graph to represent the relationships between two image regions. In a simple graph, image regions are defined as vertices and two related vertices are connected by an edge, reflecting the similarity between regions. Yang *et al.* [12] propagate the saliency information first with background queries and then with foreground queries via a graph-based ranking algorithm. In this algorithm, a smoothness term is used to include the relations between pairs of vertices and a fitting term is used to constrain the result to be close to the queries. Gong *et al.* [14] propagate saliency information via the graph first to easily-classified image regions and then to more ambiguous regions. Li *et al.* [22] improve the graph-based methods by generating pixel-wise saliency detection results based on a regularized random walk ranking algorithm. Zhang *et al.* [18] propose a salient object detection method which integrates simple graphs and incorporates a popular cognitive property about visual saliency, namely visual rarity, into the graph-based optimization framework. In the above graph-based methods, one kind of queries (foreground queries or background queries) is explicitly exploited in one step.

C. HYPERGRAPH-BASED

Hypergraphs have been shown to be effective in computer vision applications [19], [20], [29]. Based on hypergraph, a ranking algorithm was proposed [20] for classification and further applied to image retrieval [19]. In the hypergraph-based ranking algorithm, a smoothness term includes the relations among multiple vertices and a fitting term constrains the final result to be close to the queries. Li *et al.* [21] has successfully adopted the hypergraph in the saliency detection field. In this work, binary hypergraphs at multiple scales are constructed for an input image using a clustering algorithm and describe the binary relations between vertices and hyperedges, i.e. whether a vertex is contained in a hyperedge. Li et al. then propose a hypergraph modeling approach to capture the contextual information of image regions and the final saliency map is essentially the mean of saliency detection results from multiple scales.

D. CENTER VS. BOUNDARY PRIOR

Human observers freely viewing natural scenes show a bias to attend to the center of the image [30]–[32]. Multiple works [32]–[34] have attempted to employ the central bias for saliency detection. This center prior helps when the salient objects are near the center of the input image. More recent saliency detection methods prefer to use a boundary prior, which assumes that pixels on the image boundaries are more likely to be part of the background. This boundary prior is more robust and effective than the center prior because the foreground object rarely intersects all four image borders. In the work of Yang *et al.* [12], the image borders are chosen as the background queries to perform saliency propagation. Jiang *et al.* [35] choose the virtual boundary nodes as the absorbing nodes and formulate saliency detection in an image via an absorbing Markov chain.

III. FOREGROUND- AND BACKGROUND-QUERIES-BASED HYPERGRAPH OPTIMIZATION FRAMEWORK

Let V denote a finite set of vertices in a hypergraph. Let E denote a set of hyperedges, which is a family of subsets of V such that $\bigcup_{e\in E} = V$. A hypergraph can be represented by a $|V| \times |E|$ incidence matrix $H(v_i, e_j)$ where $v_i \in V$ is a vertex and $e_j \in E$ is a hyperedge.

$$H(v_i, e_j) = \begin{cases} 1 & \text{if } v_i \in e_j \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

If a vertex v_i is contained in a hyperedge e_j (i.e. $v_i \in e_j$), then the corresponding incident value $H(v_i, e_j)$ equals to 1; otherwise, the incident value equals to 0. By this definition, the incidence matrix H indicates different hyperedges in a hypergraph and each column of the incidence matrix contains the composition information of a hyperedge.

The weights of different hyperedges can be recorded in a diagonal matrix W, where each hyperedge e has a positive weight $W(e)$. Based on the incidence matrix H and the hyperedge weight matrix W, the degree of each vertex $D_v(v_i)$ which is recorded in a diagonal matrix D_v can be defined as $D_v(v_i) = \sum_{e_j\in E} W(e_j)H(v_i, e_j)$, and the degree of each hyperedge $D_e(e_j)$ which is recorded in a diagonal matrix D_e can be defined as $D_e(e_j) = \sum_{v_i\in V} H(v_i, e_j)$.

A ranking algorithm based on a hypergraph was proposed [20] for classification. In this algorithm, a smoothness term represents the relations among multiple vertices and a fitting term constrains the final classification result to be close to the queries. This hypergraph-based ranking algorithm is appropriate for the case where there is only one kind of specific queries, e.g. image retrieval [19]. However, salient region detection aims to separate the salient object foreground from the non-salient background. Thus, both the foreground queries and the background queries are vital for salient region detection in images.

In this paper, we propose a novel hypergraph-based optimization framework which considers both the foreground queries and the background queries to compute the saliency values (the likelihood belonging to the salient foreground) of different vertices for salient region detection. One goal of the proposed optimization framework is to group similar regions (parts of the same object or surface) and assign them to the same saliency class, which is accomplished by a smoothness term. Besides, the proposed optimization framework aims to make use of heuristics about the likelihood of different regions being foreground or background. This is accomplished by two terms which separately represent the foreground and background constraints. By considering the smoothness constraint, the foreground fitting constraint and the background fitting constraint, an optimal saliency vector $S : V \Rightarrow \mathbb{R}$, which assigns a saliency value $S(v_i)$ to each vertex v_i can be computed.

$$\begin{aligned} S^* &= \arg\min_S \left(\frac{1}{2}\Omega + \lambda_f \Psi + \lambda_b \Phi\right), \ \textit{where} \\ \Omega &= \sum_{e_k\in E} \sum_{v_i, v_j \in e_k} \frac{W(e_k)}{D_e(e_k)} H(v_i, e_k) H(v_j, e_k)(S(v_i) - S(v_j))^2 \\ \Psi &= \sum_{i=1}^{n} Q_f(v_i) \left\| S(v_i) - L_f \right\|^2 \\ \Phi &= \sum_{i=1}^{n} Q_b(v_i) \left\| S(v_i) - L_b \right\|^2 \end{aligned} \tag{2}$$

In the term Ω of the above optimization function, $S(v_i)$ and $S(v_j)$ are the saliency values of the vertices v_i and v_j respectively. $H(v_i, e_k)$ and $H(v_j, e_k)$ indicate whether v_i and v_j belong to the hyperedge e_k. $W(e_k)/D_e(e_k)$ can be interpreted as the normalized weight of the hyperedge e_k. Thus, Ω is a smoothness term which represents the smoothness constraint as it is defined in a hypergraph: if two vertices v_i and v_j are frequently members of the same hyperedge and these hyperedges have high weights, then v_i and v_j should have similar saliency values. This reflects the fact that nearby, similar-looking image regions are very likely to be parts of the same object or surface, and should therefore have the same saliency values.

In the term Ψ, L_f is the label of the salient foreground, which indicates the class of the salient foreground and is set to 1 in our paper. $Q_f(v_i)$ indicates whether the vertex v_i is a salient foreground query: if v_i is a salient foreground query, $Q_f(v_i)$ can be assigned a value of the probability to be a foreground query; otherwise, $Q_f(v_i)$ equals to 0. Thus, Ψ can be understood as a foreground fitting term, which encourages a vertex v_i which is a foreground query to take a saliency value close to the foreground label.

Similarly, the term Φ can be interpreted as a background fitting term. L_b indicates the label of the non-salient background, which indicates the class of the non-salient background and is set to -1 in our paper. $Q_b(v_i)$ indicates whether the vertex v_i is a background query: if v_i is a background query, $Q_b(v_i)$ can be set to a value of the probability to be a background query; otherwise, $Q_b(v_i)$ equals to 0. Thus term Φ encourages the saliency value of a vertex v_i to be close to the label of the background if this vertex is a background query.

In this work, we obtain the foreground and background queries based on the boundary prior and similarity relationships among image regions in multiple feature spaces. It should be biased to mark regions dissimilar to the image borders as salient foreground queries and regions similar to the image borders as background queries. The implementation details can be seen in Section V.

In the above optimization function of Formula 2, λ_f and λ_b are the weighting parameters, which specify the relative balance of three terms: Ω, Ψ and Φ. By setting the derivative of the optimization function to be zero, the saliency values of different vertices can be calculated. The normalized saliency value of each vertex indicates the probability that the vertex belongs to the salient foreground.

$$S^* = (D_v - HWD_e^{-1}H^T + \lambda_f Q_f + \lambda_b Q_b)^{-1}(\lambda_f Q_f - \lambda_b Q_b) \tag{3}$$

From the result, we can see that there are two critical factors for our foreground- and background-queries-based hypergraph optimization framework in determining the final saliency values: the constructed hypergraph represented by the incidence matrix H and the hyperedge weight matrix W for the smoothness constraint term Ω; and the foreground queries Q_f and the background queries Q_b for the fitting constraint terms Ψ and Φ. In Section IV, we explain how we construct the hypergraph based on multiple features for salient region detection. And in Section V, we describe how we compute the foreground queries Q_f and the background queries Q_b based on the boundary prior and our constructed hypergraph.

IV. PROBABILISTIC HYPERGRAPH CONSTRUCTION

In the work of Li *et al.* [21], a binary hypergraph which models the binary relations between the vertices and the hyperedges is used to detect salient objects in images. However, the binary relations between the vertices and the hyperedges which treat all the vertices in a hyperedge equally is unable to represent information about similarity relationships within the hyperedge. In our work, we construct a probabilistic hypergraph which captures not only the binary relations but also the probability that a vertex belongs to a hyperedge to better capture the grouping information in an input image. If a vertex v_i is contained in a hyperedge e_j, the corresponding incident value $H(v_i, e_j)$ is equal to $p(v_i|e_j)$, i.e. the probability that v_i belongs to e_j; otherwise, the incident value is 0.

$$H(v_i, e_j) = \begin{cases} p(v_i|e_j) & \text{if } v_i \in e_j \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

We first employ the Simple Linear Iterative Clustering (SLIC) algorithm [36] to over-segment an input image I into n homogeneous superpixels, as in many previous graph-based salient object detection methods [12], [14], [22], [37]. We define these homogeneous superpixels of an image as the vertices of our probabilistic hypergraph. In this work, the number of the superpixels in each image is set to 300. For each superpixel, we compute its mean in the CIELab color space and the image spatial coordinate space separately, producing a color feature and a spatial feature for each superpixel.

A. HYPEREDGES

In addition to the vertices, we need to identify the hyperedges and compute the probabilities that the vertices belong to these hyperedges to create a probabilistic hypergraph. The above information can be recorded in an incidence matrix H. In this paper, given a finite set of vertices V, we choose different hyperedges according to three kinds of correlations: local spatial correlation, global spatial correlation, and color correlation. For each vertex v_i, three hyperedges are chosen: a local spatial hyperedge, which contains v_i itself as a centroid vertex and its immediate spatial neighbors (vertices which share a boundary with the centroid vertex in the input image); a global spatial hyperedge, which contains v_i as a centroid vertex and the global boundary vertices on the four borders of the input image; and a color hyperedge which contains v_i as a centroid vertex and its neighbors in the CIELab color space. We further compute the probability that a contained vertex belongs to a hyperedge as the similarity between this vertex and the centroid vertex of the hyperedge.

Let R_1, R_2 and R_3 denote the above three rules based on three feature correlations. We define a $n * n$ adjacency matrix A: If two vertices v_i and v_j are spatially adjacent in the input image, then $A(v_i, v_j) = A(v_j, v_i) = 1$, otherwise $A(v_i, v_j) = A(v_j, v_i) = 0$. In addition, let a $n * n$ similarity matrix SIM encode the similarities between each pair of vertices and $\mathcal{B} \in V$ denote a list of boundary vertices on the four sides of the image. Then the hyperedge set $E = \{E_1, E_2, E_3\}$ and the corresponding incidence matrix $H = [H_1, H_2, H_3]$ can be

produced as follows:

$$\begin{aligned} R_1 &: \forall v \in V,\ E_1(v) = \{v' \in V | v' = v \text{ or } A(v, v') = 1\} \\ & H_1(v', E_1(v)) = SIM(v, v') \\ R_2 &: \forall v \in V,\ E_2(v) = \{v' \in V | v' = v \text{ or } v' \in \mathcal{B}\} \\ & H_2(v', E_2(v)) = SIM(v, v') \\ R_3 &: \forall v \in V, E_3(v) = \{v' \in V | v' = v \text{ or } D_c(v, v') < \varepsilon\} \\ & H_3(v', E_3(v)) = SIM(v, v') \end{aligned} \tag{5}$$

The first rule enables the hypergraph to make use of the proximity of spatial neighbors. The second rule effectively utilizes the global spatial correlation. The third rule helps to group image regions with similar appearances into the same class, which is particularly useful in images with multiple salient objects or a single object with complex patterns. For the third rule, we use Euclidean distance to compute the color distance between two vertices $D_c(v, v')$ and the parameter ε is set to 0.15 in our work. Our method to compute the similarity between two vertices is given in Section IV-C.

B. HYPEREDGE WEIGHTS

In addition to the hyperedges described by the incidence matrix H, the hyperedge weights in the diagonal matrix W are also important for salient region detection. For efficiency and simplification, we compute the hyperedge weights based on the constructed incidence matrix H and the similarity matrix SIM. The weight of each hyperedge $w(e_j)$ is defined as follows.

$$w(e_j) = \sum_{v_i \in e_j} H(v_i, e_j) SIM(v_i, v_{e_j}) \tag{6}$$

In the above formula, v_{e_j} denotes the centroid vertex of the hyperedge e_j. If all vertices contained in a hyperedge e_j have a high probability to be part of the hyperedge and have a close similarity to the centroid vertex, this hyperedge has high inner group similarity and thus should be assigned a high weight; otherwise, this hyperedge should be assigned a low weight.

C. SIMILARITY MATRIX

The $n * n$ matrix SIM used in Formula 5 and Formula 6 indicates the similarity between each pair of the vertices. In this work, we compute this matrix based on both the color distance and the spatial distance between pairs of vertices.

$$SIM(v_i, v_j) = exp(-\frac{\rho D_c(v_i, v_j) + (1 - \rho) D_s(v_i, v_j)}{2\sigma}) \tag{7}$$

The parameter ρ balances the weights of the color distance and the spatial distance. A larger ρ will put more emphasis on the color distance. The scale parameter σ controls the strength of the combined distance. If σ has a small value, only vertices with close color features and spatial positions would make a contribution. If σ has a larger value, then vertices with larger distances would also have a strong influence on each other. The parameters ρ and σ are set as 0.35 and 0.05 respectively.

$D_c(v_i, v_j)$ represents the color difference between vertices v_i and v_j, while $D_s(v_i, v_j)$ represents the spatial distance between those vertices. We use the Euclidean distance to compute the color distance $D_c(v_i, v_j)$ and employ the sine spatial distance proposed in our previous work [38] to compute the spatial distance $D_s(v_i, v_j)$, as shown in Formula 8.

$$D_s(v_i, v_j) = \sqrt{(\sin(\pi \cdot |x_i - x_j|))^2 + (\sin(\pi \cdot |y_i - y_j|))^2} \tag{8}$$

In the above formula, x_i and y_i represent the horizontal and vertical image coordinates of a vertex v_i, which have been normalized to be in the range [0, 1]. Experimental comparisons in previous work [38] have shown that the sine spatial distance helps to uniformly suppress the non-salient background.

V. FOREGROUND AND BACKGROUND QUERIES ACQUISITION

In this section, we describe how we obtain the foreground and background queries using the boundary prior and our constructed hypergraph. According to the boundary prior, the vertices on the four borders of the image are very likely to be part of the background. We use each of the four image sides as the initial background queries separately to get four background maps based on the following hypergraph-based ranking algorithm.

$$\begin{aligned} B^* &= \arg\min_B (\frac{1}{2}\Omega + \alpha\Lambda),\ \ where \\ \Omega &= \sum_{e_k \in E} \sum_{v_i, v_j \in e_k} \frac{W(e_k)}{D_e(e_k)} H(v_i, e_k) H(v_j, e_k) (B(v_i) - B(v_j))^2 \\ \Lambda &= \sum_{i=1}^{n} \|B(v_i) - O(v_i)\|^2 \end{aligned} \tag{9}$$

In the above formula, $O(v_i)$ indicates whether a vertex v_i is an initial background query which is on one of the four image borders. If a vertex v_i is on the image border, $O(v_i) = 1$ indicating that this vertex is a background vertex; otherwise, $O(v_i) = 0$. Each of the four background maps is computed based on the smoothness constraint defined by our constructed hypergraph and one of the image borders. By setting the derivative of the above formula to be zero, the background map can be computed as:

$$B^* = \alpha(D_v - HWD_e^{-1}H^T + \alpha I)^{-1} O \tag{10}$$

According to Formula 10, the vertices which have the smoothness constraint with the image borders would have a value close to 1, and otherwise would have a value close to 0. Using the four image boundary sides (O_{top}, O_{down}, O_{left} and O_{right}) separately, we can create four background maps (B_{top}, B_{down}, B_{left} and B_{right}). We then combine them to get an initial saliency map based on the following formula.

$$T = (1 - B_{top}). * (1 - B_{down}). * (1 - B_{left}). * (1 - B_{right}) \tag{11}$$

In this formula, we use node multiplication to combine different background maps. The final foreground queries and

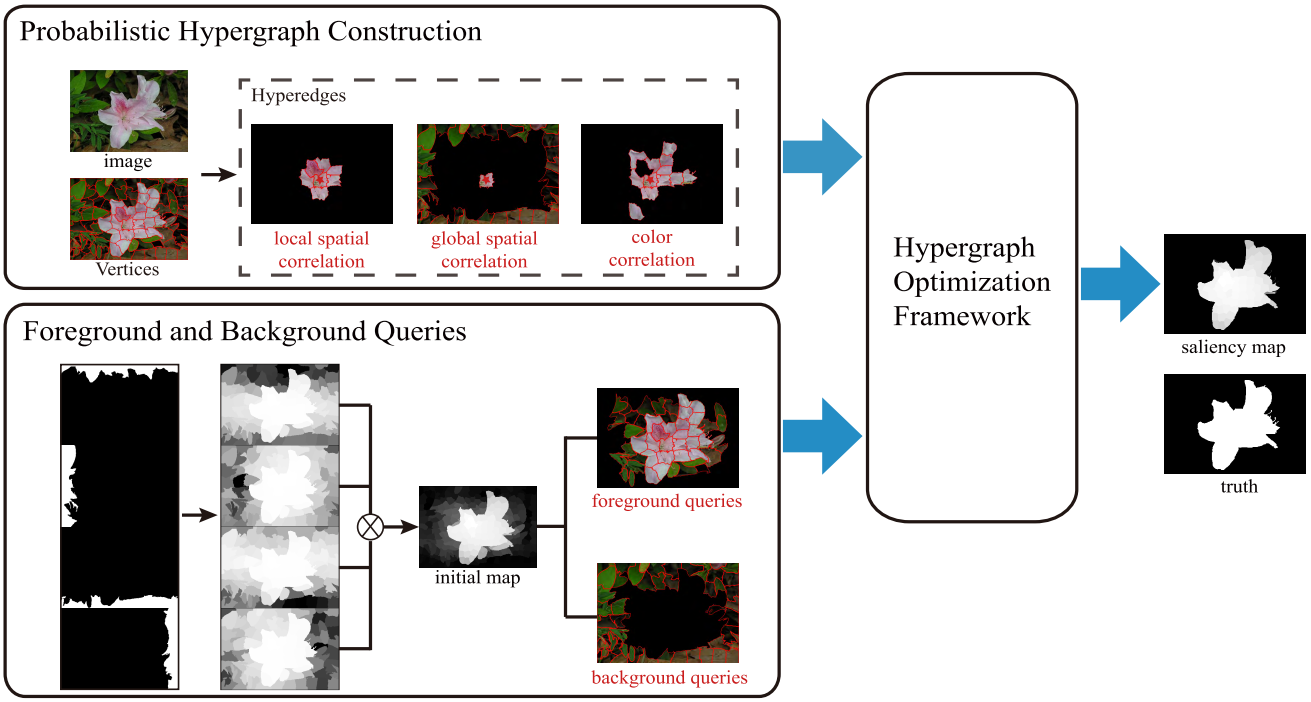

FIGURE 1. Diagram of our foreground- and background-queries-based hypergraph optimization method for salient region detection. The original image is first over-segmented into superpixels, which are the vertices of our probabilistic hypergraph. Each vertex forms hyperedges based on local spatial correlation, global spatial correlation and color correlation. In the diagram, the centroid vertex of a hyperedge is marked with a small red star. Each of the four image borders is set as the initial background queries separately to get four background maps. An initial saliency map is obtained by combining these background maps and then further thresholded to get the final foreground queries and background queries. Both the probabilistic hypergraph and the queries are incorporated into the hypergraph optimization framework to detect salient objects.

background queries are obtained from this initial result T using different thresholds th_f and th_b:

$$Q_f(v_i) = \begin{cases} T(v_i) & \text{if } T(v_i) \geq th_f \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

$$Q_b(v_i) = \begin{cases} 1 - T(v_i) & \text{if } T(v_i) \leq th_b \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

If a vertex v_i has a value $T(v_i) \geq th_f$, then this vertex is a foreground query and the probability to be a foreground query is set as $T(v_i)$. If a vertex v_i has a value $T(v_i) \leq th_b$, this vertex is a background query and the corresponding probability is set as $1 - T(v_i)$. The complete procedure of our proposed foreground- and background-queries-based hypergraph optimization framework for salient region detection is summarized in Algorithm 1 and Figure 1.

VI. EXPERIMENTAL COMPARISON

A. EXPERIMENTAL SETUP

1) DATABASES

We evaluate the performance of our framework on five public databases: MSRA10K [10], ECSSD [39], SOD [40], iCoseg [41] and PASCAL-S [42]. The MSRA10K database contains 10,000 images with a variety of image contents. Most of the images have a single salient object. The ECSSD (Extended Complex Scene Saliency Database) contains 1,000 natural images with complex foreground and background patterns. The iCoseg database contains 643 images and many of them contain multiple foreground objects. The SOD database contains 300 challenging images, originally from the Berkeley segmentation database [43]. In the SOD database, many images contain multiple foreground objects and complex patterns. The PASCAL-S database contains 850 natural images with object masks labeled by 12 subjects. We threshold the masks at 0.5 to obtain binary ground truth as suggested by [42].

2) EVALUATION CRITERIA

We evaluate performance using precision-recall (PR) curves [44], max F-measure [45] and S-measure [46]. The precision-recall curve of a saliency map is computed by converting a continuous saliency map to a binary map using each threshold in the range [0:1:255] and comparing this binary map against the ground truth. The precision-recall curve of a database is obtained by averaging the precision-recall values over saliency maps of all images in the database. As a measure of overall performance, we also report the max F-measure values on different databases. The F-measure is defined as $F_\beta = \frac{(1+\beta^2)\cdot precision \cdot recall}{\beta^2 \cdot precision + recall}$, where β^2 is set to 0.3. We report the maximum F-measure calculated from all precision-recall pairs as suggested in [45]. Since human visual system is highly sensitive to the structures in image

Algorithm 1 Foreground- and Background-Queries-Based Hypergraph Optimization Framework for Salient Region Detection

Input: An image and required parameters.

1. Construct a probabilistic hypergraph based on local spatial correlation, global spatial correlation and color correlation: Choose hyperedges and compute the probability that a vertex belongs to a hyperedge according to Formula 5 to get the hyperedge set E and an incident matrix H; Compute the hyperedge weights according to Formula 6 to get a hyperedge weight matrix W. The similarity matrix SIM which is used for the incident matrix H and the hyperedge weight matrix W can be computed according to Formula 7 and Formula 8.
2. Set each of the four image borders as the initial background queries separately to get four background maps based on Formula 10. Combine them to get the initial saliency map based on Formula 11. The final foreground queries and the final background queries are then obtained from this initial saliency detection result using two thresholds separately based on Formula 12 and Formula 13.
3. Compute the final saliency result by Formula 3 according to the constructed probabilistic hypergraph, the final foreground queries, and the final background queries.

Output: A saliency map.

scenes, in this paper we also report the S-measure metric which puts more emphasis on the region-aware and object-aware structural similarity between a saliency map and the ground truth. The implementation detail of S-measure can be found in Fan *et al.*'s work [46].

3) PARAMETER SETUP

We choose parameter empirically. The weights λ_f and λ_b in Formula 3 are set to 0.05 and 0.1, respectively. The weight α in Formula 10 is set to 0.2. The thresholds th_f and th_b in Formula 12 and Formula 13 are empirically chosen, $th_f = 0.2$ and $th_b = 0.5$, for all the experiments. Our method is implemented using Matlab on a computer with Intel Core i7-4790 3.6 GHz CPU and 8GB RAM.

B. QUANTITATIVE COMPARISON

We compare our method (HO2) against twenty-one recent methods, including GB [8], FT [9], MSS [47], CB [11], HC [10], RC [10], GS [24], SF [48], AM [35], GR [12], HM [21], HS [39], BD [13], BSCA [49], CL [14], GP [15], RRWR [22], PM [16], MST [17], GF [18] as well as a preliminary version of our method HO1 [50]. CL, GP, PM and RRWR are the latest graph-based methods. HM is the previous hypergraph-based method. For fair comparison, we use the implementations provided by the authors.

1) MSRA10K

Figure 2 (a) compares our method with state-of-the-art methods using precision-recall curves on the MSRA10K database where most images contain one salient object. Recent graph-based methods, e.g. PM, RRWR, GP, BD, GR and AM show competitive performance. However, the proposed method outperforms these methods which demonstrates the effectiveness of our approach. HM is a hypergraph-based model, but the algorithms proposed by HM and our method are quite different. For one thing, HM constructs a binary hypergraph based on a clustering algorithm while our method constructs a probabilistic hypergraph based on three correlations: local spatial correlation, global spatial correlation and color correlation. In addition, HM detects the salient region using a hypergraph modeling algorithm which is essentially the average of saliency detection maps from different image scales, while in this work we propose a novel foreground- and background-queries-based hypergraph optimization framework. The quantitative comparison with PR curves demonstrates that our method outperforms HM by a large margin on MSRA10K database. The max F-measure values of different methods on MSRA10K database is shown in Table 1. As we can see, our method has the largest max F-measure value, which demonstrates that our method gives the best overall performance when considering both precision and recall values. The S-measure values of different methods in Table 2 further demonstrate that our method can better extract the region-aware and object-aware structures of the salient objects on the MSRA10K database.

2) ECSSD

We further evaluate our method on the ECSSD database which contains various images with complex foreground and background patterns, shown in Figure 2 (b), Table 1, and Table 2. The performances of different methods as measured by PR curves, max F-measure and S-measure are slightly lower in comparison to the MSRA10K database. This indicates that it is harder to detect salient objects in complex images. However, the proposed method outperforms other state-of-the-art methods, demonstrating that the proposed hypergraph-based optimization framework and the usage of both the foreground queries and the background queries helps to detect salient foreground objects in complex natural images.

3) ICOSEG

The iCoseg database contains a number of images with multiple salient objects. The performance comparison in Figure 2 (c), Table 1, and Table 2 shows that the proposed method outperforms other methods. It demonstrates that the proposed foreground-queries- and background-queries-based optimization framework and the construction of the probabilistic hypergraph can help to detect multiple salient objects in images.

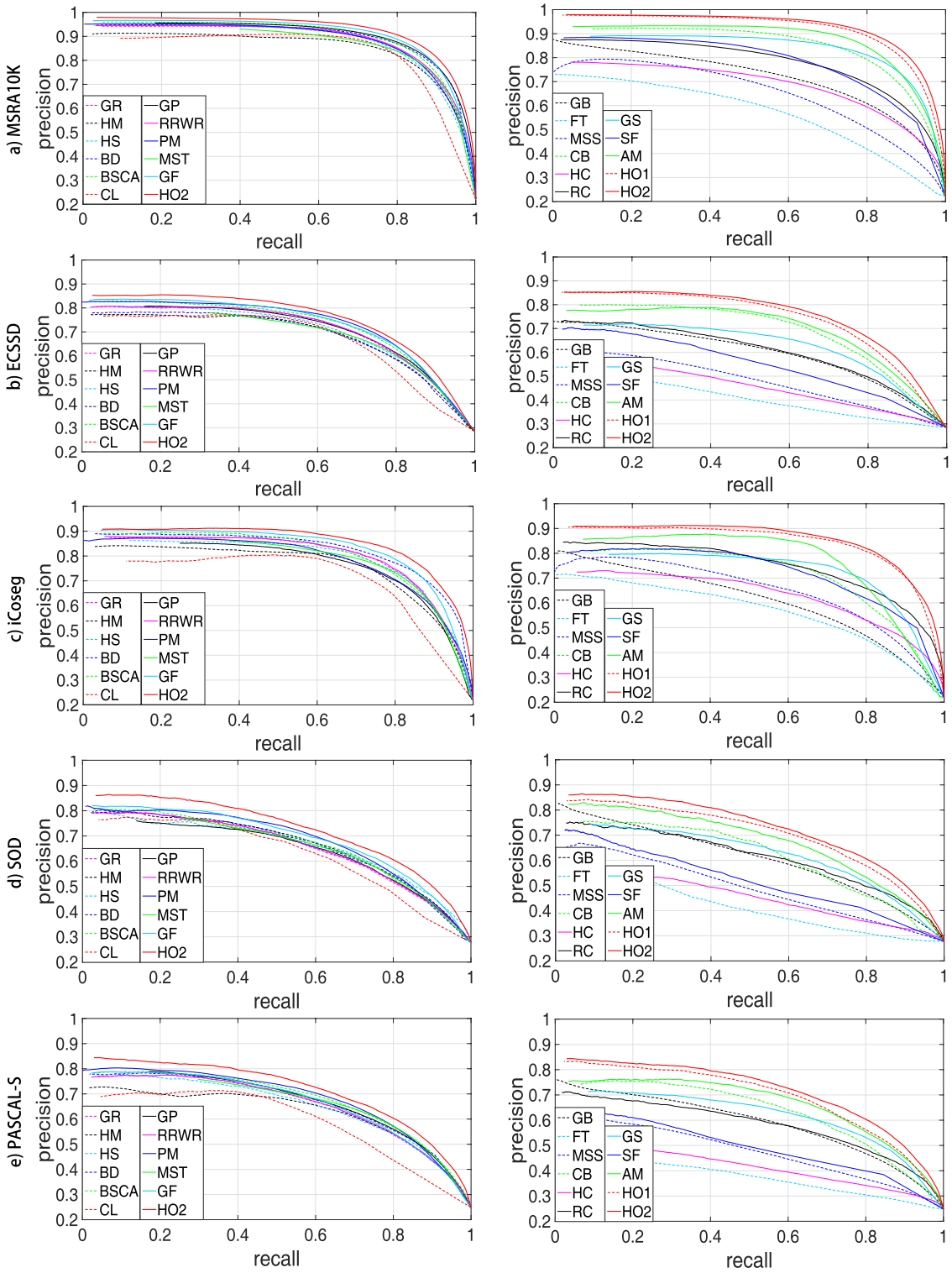

FIGURE 2. The quantitative comparison of different methods on MSRA10K, ECSSD, SOD, iCoseg and PASCAL-S databases based on precision-recall curves. Our method gives the best performance on all databases.

4) SOD

The SOD database is a difficult saliency detection database, which contains a number of natural images with multiple salient objects and complex patterns. Figure 2 (d) shows that our method can better detect salient objects in challenging images with higher precision and recall values. Furthermore, our method improves on the maximum F-measure achieved by the state-of-the-art methods, shown in Table 1. Table 2

TABLE 1. Performance Comparison Based on Max F-Measure. The First, Second, and Third Best Results on Each Database Are Labeled in Red, Green, and Blue, Respectively.

$F-measure$	GB	FT	MSS	CB	HC	RC	GS
MSRA10K	0.648	0.536	0.609	0.775	0.645	0.696	0.792
ECSSD	0.545	0.384	0.454	0.631	0.432	0.537	0.600
iCoseg	0.538	0.532	0.595	0.653	0.588	0.656	0.686
SOD	0.538	0.375	0.446	0.537	0.423	0.535	0.572
PASCAL-S	0.525	0.366	0.448	0.568	0.403	0.527	0.578
$F-measure$	SF	AM	GR	HM	HS	BD	BSCA
MSRA10K	0.713	0.809	0.809	0.797	0.812	0.831	0.834
ECSSD	0.505	0.638	0.650	0.646	0.646	0.639	0.679
iCoseg	0.650	0.720	0.729	0.699	0.733	0.746	0.737
SOD	0.465	0.590	0.562	0.593	0.584	0.595	0.601
PASCAL-S	0.455	0.599	0.586	0.593	0.578	0.602	0.615
$F-measure$	CL	GP	RRWR	PM	MST	GF	HO2
MSRA10K	0.786	0.827	0.811	0.802	0.813	0.835	0.855
ECSSD	0.625	0.664	0.652	0.680	0.648	0.667	0.694
iCoseg	0.689	0.694	0.730	0.712	0.722	0.753	0.786
SOD	0.539	0.588	0.564	0.600	0.594	0.598	0.634
PASCAL-S	0.528	0.607	0.586	0.610	0.613	0.584	0.626

TABLE 2. Performance Comparison Based on S-Measure. The First, Second, and Third Best Results on Each Database Are Labeled in Red, Green, and Blue, Respectively.

$S-measure$	GB	FT	MSS	CB	HC	RC	GS
MSRA10K	0.671	0.532	0.531	0.759	0.663	0.679	0.775
ECSSD	0.572	0.427	0.445	0.621	0.479	0.551	0.611
iCoseg	0.606	0.558	0.554	0.676	0.641	0.670	0.711
SOD	0.575	0.421	0.441	0.564	0.468	0.557	0.597
PASCAL-S	0.584	0.426	0.463	0.611	0.472	0.566	0.622
$S-measure$	SF	AM	GR	HM	HS	BD	BSCA
MSRA10K	0.591	0.785	0.786	0.755	0.787	0.808	0.813
ECSSD	0.451	0.639	0.643	0.608	0.635	0.637	0.666
iCoseg	0.581	0.698	0.712	0.691	0.735	0.757	0.732
SOD	0.420	0.606	0.586	0.567	0.599	0.589	0.622
PASCAL-S	0.439	0.632	0.625	0.595	0.623	0.641	0.652
$S-measure$	CL	GP	RRWR	PM	MST	GF	HO2
MSRA10K	0.767	0.819	0.789	0.798	0.810	0.821	0.832
ECSSD	0.628	0.658	0.645	0.667	0.648	0.661	0.677
iCoseg	0.699	0.716	0.712	0.727	0.736	0.764	0.776
SOD	0.563	0.620	0.588	0.614	0.609	0.618	0.648
PASCAL-S	0.579	0.651	0.626	0.644	0.651	0.636	0.664

demonstrates that the proposed method outperforms other methods by better extracting the structures of the salient objects.

5) PASCAL-S

We also compare different methods on the PASCAL-S database, which contains various natural images with the ground truth labeled by 12 subjects. The quantitative comparison in Figure 2 (e), Table 1, and Table 2 demonstrates that our method outperforms other salient object detection methods on the PASCAL-S database in terms of higher precision-recall values, max F-measure and S-measure.

The proposed method HO2 also consistently outperforms the preliminary version of this method HO1 [50] on the above five databases as shown by the precision-recall curves, max F-measure and S-measure in Figure 2 and Table 3. This is because the proposed method HO2 improves on the preliminary version's methods for acquiring the foreground and background queries and constructing the probabilistic hypergraph [50].

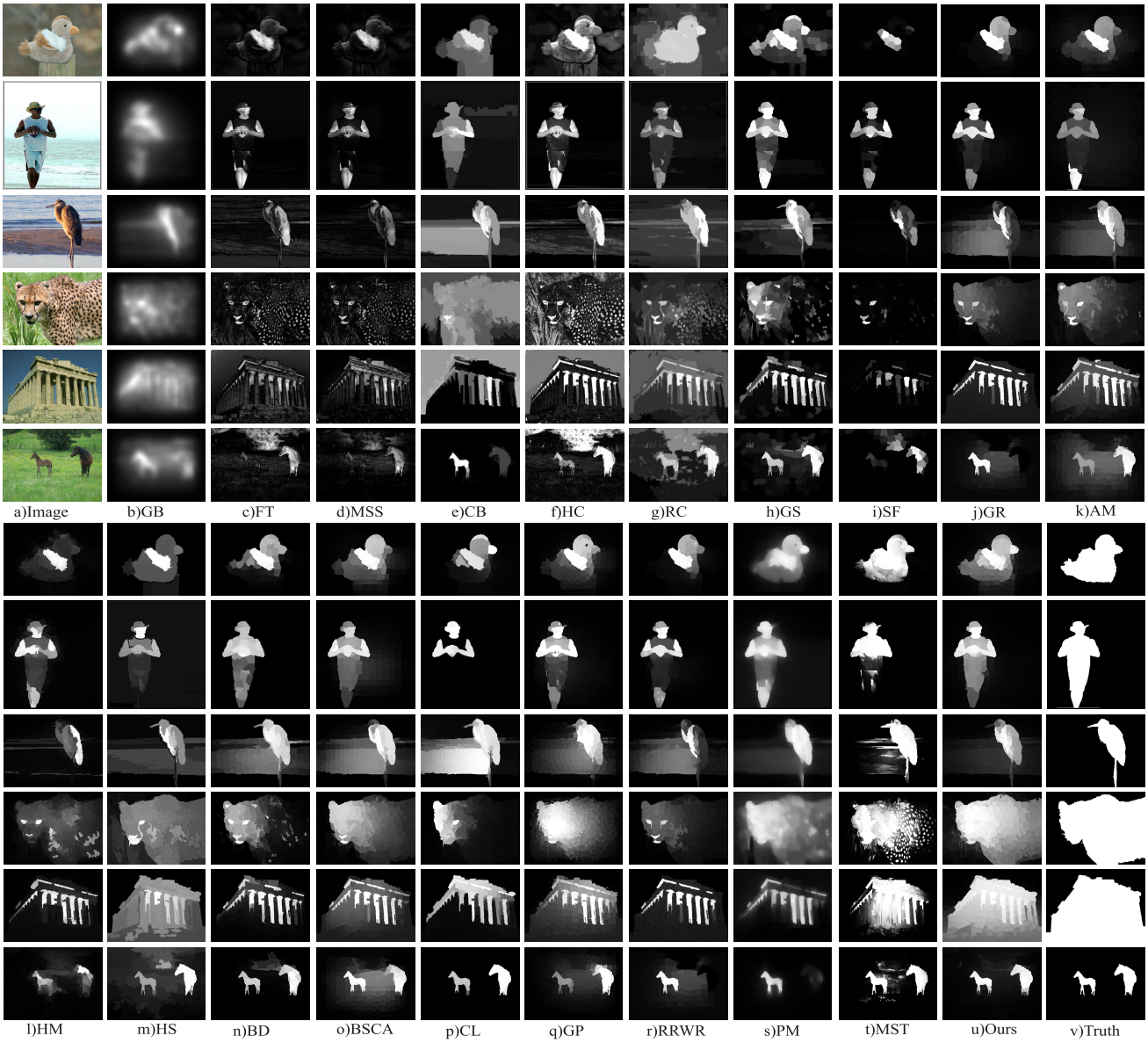

FIGURE 3. Visual comparison of saliency maps generated from different salient object detection methods. The first column is the original image and the last column is the ground truth. The second-to-last column is our saliency detection result and the remaining columns are results of other evaluated methods.

TABLE 3. Performance Comparison of the Proposed Method With the Preliminary Version Based on Max F-Measure and S-Measure.

Metric	Method	MSRA10K	ECSSD	iCoseg	SOD	PASCAL-S
F-measure	HO1	0.845	0.684	0.770	0.611	0.606
	HO2	0.855	0.694	0.786	0.634	0.626
S-measure	HO1	0.806	0.663	0.763	0.619	0.636
	HO2	0.832	0.677	0.776	0.648	0.664

C. VISUAL COMPARISON

We present a few saliency maps generated by our method as well as other representative methods, including graph-based methods(PM [16], RRWR [22], GP [15], CL [14], BD [13], GR [12], AM [35] and GB [8]) and a hypergraph-based method (HM [21]), for visual comparison in Figure 3.

Representative images have been chosen to highlight the differences between different methods. The first and second images have low contrast between the foreground object and the background. Most other methods only detect part of the foreground while our method can uniformly highlight the entire salient object. The third image has a

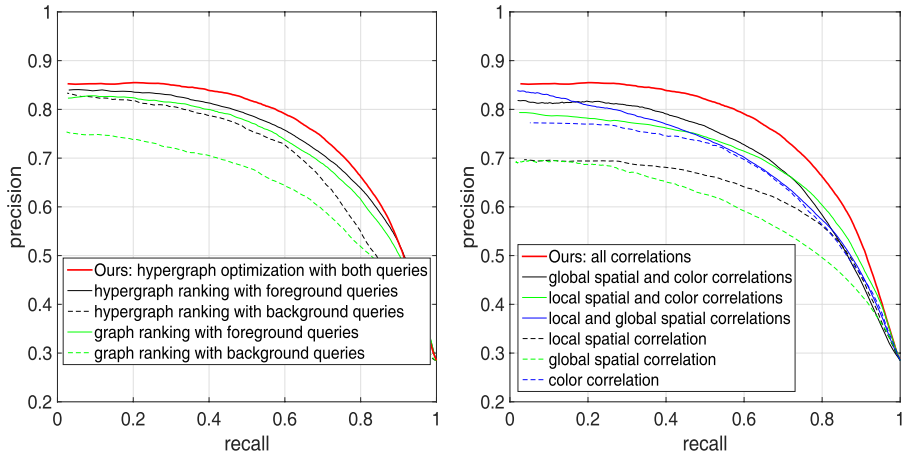

FIGURE 4. The PR curves on ECSSD database with different design options.

complex background. Many other methods wrongly highlight a portion of the background, but our method can suppress the background because we introduce the background queries in our hypergraph optimization framework in addition to the foreground queries, which can help to inhibit the image background. The forth and fifth images show salient objects which intersect the image boundaries. Most methods cannot highlight the foreground regions near the image borders. Our method can highlight the whole foreground object due to the constructed probabilistic hypergraph based on local spatial correlation, global spatial correlation and color correlation, which can better represent the grouping information in the image. When there are multiple disconnected salient objects in an image (last image), our method can uniformly highlight all of the salient objects and largely suppress the background while some methods detect only part of the salient objects or wrongly highlight the background. The major reason why our method works better is that we construct a probabilistic hypergraph based on three different correlations in the proposed method, which can help to detect all the salient regions even when they are not contiguous. In short, the proposed foreground- and background-queries-based probabilistic hypergraph optimization framework generates more accurate saliency maps, which highlight more of the salient objects with less noise in a variety of challenging cases.

D. EXAMINATION OF DESIGN OPTIONS

We examine different design options on the ECSSD database, shown in Figure 4. We first examine the major innovations of our proposed method in Figure 4 (a). The green full curve and the black full curve show the performances of the graph-based ranking framework with only foreground queries and the hypergraph-based ranking framework with only foreground queries. The green dashed curve and the black dashed curve represent the performances of the graph-based ranking framework with only background queries and the hypergraph-based ranking framework with only background queries. Both the comparisons between the graph-based ranking framework and the hypergraph-based ranking framework with only foreground queries and with only background queries demonstrate that the hypergraph is helpful for detecting salient regions in images. The red curve shows the performance of the final output using the proposed hypergraph optimization framework with both the foreground queries and the background queries. The complete method outperforms the methods with only foreground queries and with only background queries, which demonstrates that including both the foreground queries and the background queries can help to uniformly highlight the salient foreground and suppress the background. Based on the above observations, both the hypergraph and the inclusion of two kinds of queries in our optimization framework contribute to the overall performance.

We further compare the performances of this method using different hypergraphs based on various subsets of features (local spatial correlation, global spatial correlation, and color correlation) in Figure 4 (b). Hypergraphs which use only one of these three features have the lowest performance. Hypergraph which use two of the three features perform somewhat better, while the complete method based on all three kinds of correlations gives the best performance. This demonstrates that local spatial correlation, global spatial correlation and color correlation are all helpful for salient region detection.

VII. CONCLUSION

In this work, we present a novel foreground- and background-queries-based hypergraph optimization framework for salient region detection. Unlike the traditional graph-based method, we construct a probabilistic hypergraph based on local spatial correlation, global spatial correlation, and color correlation to include both the pairwise and the higher-order relations among two or more vertices. Furthermore, both foreground queries and background queries are explicitly exploited in the optimization framework to uniformly highlight the salient foreground and suppress the non-salient background. An extensive experimental evaluation demonstrates the effectiveness of our method on five representative saliency databases.

APPENDIX A
CALCULATE THE SOLUTION OF FORMULA 2

The optimization function $\mathcal{F}(S)$ of Formula 2 is:

$$\begin{aligned}\mathcal{F}(S) = {} & S^T(D_v - HWD_e^{-1}H^T)S \\ & + \lambda_f Q_f(S - L_f)^T(S - L_f) \\ & + \lambda_b Q_b(S - L_b)^T(S - L_b)\end{aligned} \tag{14}$$

The derivative of the above optimization function is:

$$\begin{aligned}\frac{\partial \mathcal{F}(S)}{\partial S} = {} & 2(D_v - HWD_e^{-1}H^T)S \\ & + 2\lambda_f Q_f(S - L_f) \\ & + 2\lambda_b Q_b(S - L_b)\end{aligned} \tag{15}$$

Set the derivative to be zero and the optimal S can be computed:

$$S^* = (D_v - HWD_e^{-1}H^T + \lambda_f Q_f + \lambda_b Q_b)^{-1}(\lambda_f Q_f - \lambda_b Q_b) \tag{16}$$

APPENDIX B
CALCULATE THE SOLUTION OF FORMULA 9

The optimization function $\mathcal{F}(B)$ of Formula 9 is:

$$\begin{aligned}\mathcal{F}(B) = {} & B^T(D_v - HWD_e^{-1}H^T)B \\ & + \alpha(B - O)^T(B - O)\end{aligned} \tag{17}$$

The derivative of the above optimization function is:

$$\frac{\partial \mathcal{F}(B)}{\partial B} = 2(D_v - HWD_e^{-1}H^T)B + 2\alpha(B - O) \tag{18}$$

Set the derivative to be zero and the optimal B can be computed:

$$B^* = \alpha(D_v - HWD_e^{-1}H^T + \alpha I)^{-1}O \tag{19}$$

REFERENCES

[1] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 185–198, Jan. 2010.
[2] A. Li, X. She, and Q. Sun, "Color image quality assessment combining saliency and FSIM," in *Proc. Int. Conf. Digit. Image Process.*, 2013, pp. 88780I-1–88780I-5.
[3] Q. Li, Y. Zhou, and J. Yang, "Saliency based image segmentation," in *Proc. Int. Conf. Multimedia Technol.*, 2011, pp. 5068–5071.
[4] F. Moosmann, D. Larlus, and F. Jurie, "Learning saliency maps for object categorization," in *Proc. Eur. Conf. Comput. Vis. Workshops*, 2006, pp. 1–14.
[5] G.-H. Liu, J.-Y. Yang, and Z. Li, "Content-based image retrieval using computational visual attention model," *Pattern Rcognit.*, vol. 48, no. 8, pp. 2554–2566, 2015.
[6] M. De Marsico, R. Distasi, M. Nappi, and D. Riccio, "Fractal based image indexing and retrieval," in *Intelligent Computing Based on Chaos*. Berlin, Germany: Springer, 2009, ch. 4, pp. 73–92.
[7] A. F. Abate, M. Nappi, G. Tortora, and M. Tucci, "IME: An image management environment with content-based access," *Image Vis. Comput.*, vol. 17, no. 13, pp. 967–980, 1999.
[8] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Proc. Adv. Neural Inf. Process. Syst.*, 2006, pp. 545–552.
[9] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1597–1604.
[10] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M. Hu, "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 569–582, Mar. 2015.
[11] H. Jiang, J. Wang, Z. Yuan, T. Liu, N. Zheng, and S. Li, "Automatic salient object segmentation based on context and shape prior," in *Proc. Brit. Mach. Vis. Conf.*, vol. 6. 2011, pp. 1–7.
[12] 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.
[13] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2814–2821.
[14] C. Gong *et al.*, "Saliency propagation from simple to difficult," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2531–2539.
[15] P. Jiang, N. Vasconcelos, and J. Peng, "Generic promotion of diffusion-based salient object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 2015, pp. 217–225.
[16] Y. Kong, L. Wang, X. Liu, H. Lu, and X. Ruan, "Pattern mining saliency," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 583–598.
[17] W.-C. Tu, S. He, Q. Yang, and S.-Y. Chien, "Real-time salient object detection with a minimum spanning tree," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 2334–2342.
[18] J. Zhang, K. A. Ehinger, H. Wei, K. Zhang, and J. Yang, "A novel graph-based optimization framework for salient object detection," *Pattern Recognit.*, vol. 64, pp. 39–50, Apr. 2017.
[19] Y. Huang, Q. Liu, S. Zhang, and D. N. Metaxas, "Image retrieval via probabilistic hypergraph ranking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3376–3383.
[20] D. Zhou, J. Huang, and B. Schölkopf, "Learning with hypergraphs: Clustering, classification, and embedding," in *Proc. Adv. Neural Inf. Process. Syst.*, 2006, pp. 1601–1608.
[21] X. Li, Y. Li, C. Shen, A. Dick, and A. van den Hengel, "Contextual hypergraph modeling for salient object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 3328–3335.
[22] C. Li, Y. Yuan, W. Cai, Y. Xia, and D. D. Feng, "Robust saliency detection via regularized random walks ranking," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 2710–2717.
[23] Q. Wang, W. Zheng, and R. Piramuthu, "Grab: Visual saliency via novel graph model and background priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 535–543.
[24] Y. Wei, F. Wen, W. Zhu, and J. Sun, "Geodesic saliency using background priors," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 29–42.
[25] P. Jiang, H. Ling, J. Yu, and J. Peng, "Salient region detection by UFO: Uniqueness, focusness and objectness," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1976–1983.
[26] H. Jiang, J. Wang, Z. Yuan, Y. Wu, N. Zheng, and S. Li, "Salient object detection: A discriminative regional feature integration approach," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 2083–2090.
[27] G. Li and Y. Yu, "Visual saliency based on multiscale deep features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 5455–5463.
[28] N. Liu and J. Han, "DHSNet: Deep hierarchical saliency network for salient object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 678–686.
[29] S. Kim, S. Nowozin, P. Kohli, and C. D. Yoo, "Higher-order correlation clustering for image segmentation," in *Proc. Adv. Neural Inf. Process. Syst.*, 2011, pp. 1530–1538.
[30] B. W. Tatler, "The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions," *J. Vis.*, vol. 7, no. 14, pp. 1–17, 2007.
[31] M. Bindemann, "Scene and screen center bias early eye movements in scene viewing," *Vis. Res.*, vol. 50, no. 23, pp. 2577–2587, 2010.
[32] T. Judd, K. Ehinger, F. Durand, and A. Torralba, "Learning to predict where humans look," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep. 2009, pp. 2106–2113.
[33] Q. Zhao and C. Koch, "Learning a saliency map using fixated locations in natural scenes," *J. Vis.*, vol. 11, no. 3, pp. 1–9, 2011.
[34] D. Parkhurst, K. Law, and E. Niebur, "Modeling the role of salience in the allocation of overt visual attention," *Vis. Res.*, vol. 42, no. 1, pp. 107–123, 2002.
[35] B. Jiang, L. Zhang, H. Lu, C. Yang, and M.-H. Yang, "Saliency detection via absorbing markov chain," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1665–1672.
[36] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[37] S. Lu, V. Mahadevan, and N. Vasconcelos, "Learning optimal seeds for diffusion-based salient object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2790–2797.
[38] J. Zhang, K. A. Ehinger, J. Ding, and J. Yang, "A prior-based graph for salient object detection," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 1175–1178.
[39] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 1155–1162.
[40] V. Movahedi and J. H. Elder, "Design and perceptual validation of performance measures for salient object segmentation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2010, pp. 49–56.
[41] D. Batra, A. Kowdle, D. Parikh, J. Luo, and T. Chen, "iCoseg: Interactive co-segmentation with intelligent scribble guidance," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3169–3176.
[42] Y. Li, X. Hou, C. Koch, J. M. Behg, and A. L. Yuille, "The secrets of salient object segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 280–287.
[43] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. 8th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2. Jul. 2001, pp. 416–423.
[44] D. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *J. Mach. Learn. Technol.*, vol. 2, pp. 37–83, Dec. 2011.
[45] G. Li, Y. Xie, L. Lin, and Y. Yu, "Instance-level salient object segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 2386–2395.
[46] D.-P. Fan, M.-M. Cheng, Y. Liu, T. Li, and A. Borji, "Structure-measure: A new way to evaluate foreground maps," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 1–10.
[47] R. Achanta and S. Susstrunk, "Saliency detection using maximum symmetric surround," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 2653–2656.
[48] F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung, "Saliency filters: Contrast based filtering for salient region detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 733–740.
[49] Y. Qin, H. Lu, Y. Xu, and H. Wang, "Saliency detection via cellular automata," in *Proc. IEEE Conf. CVPR*, Jun. 2015, pp. 110–119.
[50] J. Zhang, S. Fang, K. A. Ehinger, W. Guo, W. Yang, and H. Wei, "Probabilistic hypergraph optimization for salient object detection," in *Proc. Int. Conf. Intell. Sci. Big Data Eng.*, 2017, pp. 368–378.

JINXIA ZHANG received the B.S. and Ph.D. degrees from the Department of Computer Science and Engineering, Nanjing University of Science and Technology, China, in 2009 and 2015, respectively. She was a Visiting Scholar with the Visual Attention Lab, Brigham and Women's Hospital, and also with the Harvard Medical School from 2012 to 2014. She is currently a Lecturer with the School of Automation, Southeast University. Her research interests include visual attention, visual saliency detection, computer vision, and machine learning.

SHIXIONG FANG received the B.S. degree from the School of Electrical and Electronic Engineering, Hubei University of Technology, China, in 2009, and the M.S. and Ph.D. degrees in control theory and control engineering from Southeast University, China, in 2002 and 2009, respectively. He was a Visiting Scholar with the European Organization for Nuclear Research from 2004 to 2006. He is currently a Lecturer with the School of Automation, Southeast University. His research interests include computer vision and cyber physical systems.

KRISTA A. EHINGER received the B.S. degree in engineering and applied science from the California Institute of Technology in 2003, the B.Sc. degree (Hons.) in psychology from The University of Edinburgh in 2007, and the Ph.D. degree from the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, in 2013. She was a Post-Doctoral Fellow with the Visual Attention Lab, Brigham and Women's Hospital, and also with the Harvard Medical School from 2013 to 2016. She is currently a Post-Doctoral Fellow with the Centre for Vision Research, York University. Her research interests are scene recognition, spatial representation, and visual search.

HAIKUN WEI received the B.S. degree from the Department of Automation, North China University of Technology, China, in 1994, and the M.S. and Ph.D. degrees from the Research Institute of Automation, Southeast University, China, in 1997 and 2000, respectively. He was a Visiting Scholar with the RIKEN Brain Science Institute, Japan, from 2005 to 2007. He is currently a Professor with the School of Automation, Southeast University. His research interest is real and artificial in neural networks and industry automation.

WANKOU YANG received the B.S., M.S., and Ph.D. degrees from the School of Computer Science and Technology, Nanjing University of Science and Technology, China, in 2002, 2004, and 2009, respectively. From 2009 to 2011, he was a Post-Doctoral Fellow with the School of Automation, Southeast University. From 2011 to 2016, he was an Assistant Professor with the School of Automation, Southeast University. He is currently an Associate Professor with the School of Automation, Southeast University, China. His research interests include pattern recognition and computer vision.

KANJIAN ZHANG received the B.S. degree in mathematics from Nankai University, China, in 1994, and the M.S. and Ph.D. degrees in control theory and control engineering from Southeast University, China, in 1997 and 2000, respectively. He is currently a Professor with the School of Automation, Southeast University. His research is in nonlinear control theory and its applications, with particular interest in robust output feedback design and optimization control.

JINGYU YANG received the B.S. degree in computer science from the Nanjing University of Science and Technology (NUST), Nanjing, China. From 1982 to 1984, he was a Visiting Scientist with the Coordinated Science Laboratory, University of Illinois at Urbana–Champaign. From 1993 to 1994, he was a Visiting Professor with the Department of Computer Science, University of Missouri. In 1998, he was a Visiting Professor with Concordia University, Canada. He is currently a Professor with the Department of Computer Science and Engineering, NUST. He has authored over 300 scientific papers in computer vision, pattern recognition, and artificial intelligence. His current research interests are in the areas of pattern recognition, robot vision, image processing, data fusion, and artificial intelligence. He was a recipient of over 20 provincial and national awards.