Neurocomputing 315 (2018) 234–245

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A prior regularized multi-layer graph ranking model for image saliency computation

Yun Xiao[a], Bo Jiang[a], Zhengzheng Tu[a], Jixin Ma[b], Jin Tang[a,*]

[a] School of Computer Science and Technology, Anhui University, Hefei 230601, China
[b] Department of Computing and Information Systems, University of Greenwich, London SE10 9LS, UK

ARTICLE INFO

Article history:
Received 26 January 2018
Revised 22 June 2018
Accepted 29 June 2018
Available online 30 July 2018

Communicated by Junwei Han

Keywords:
Graph ranking
Boundary connectivity
Background possibility
Foreground possibility
Multiple layer

ABSTRACT

Bottom-up saliency detection has been widely studied in many applications, such as image retrieval, object recognition, image compression and so on. Saliency detection via manifold ranking (MR) can identify the most salient and important area from an image efficiently. One limitation of the MR model is that it fails to consider the prior information in its ranking process. To overcome this limitation, we propose a prior regularized multi-layer graph ranking model (RegMR), which uses the prior calculating by boundary connectivity. We employ the foreground possibility in the first stage and background possibility in the second stage based on a multi-layer graph. We compare our model with fifteen state-of-the-art methods. Experiments show that our model performs well than all other methods on four public databases on PR-curves, F-measure and so on.

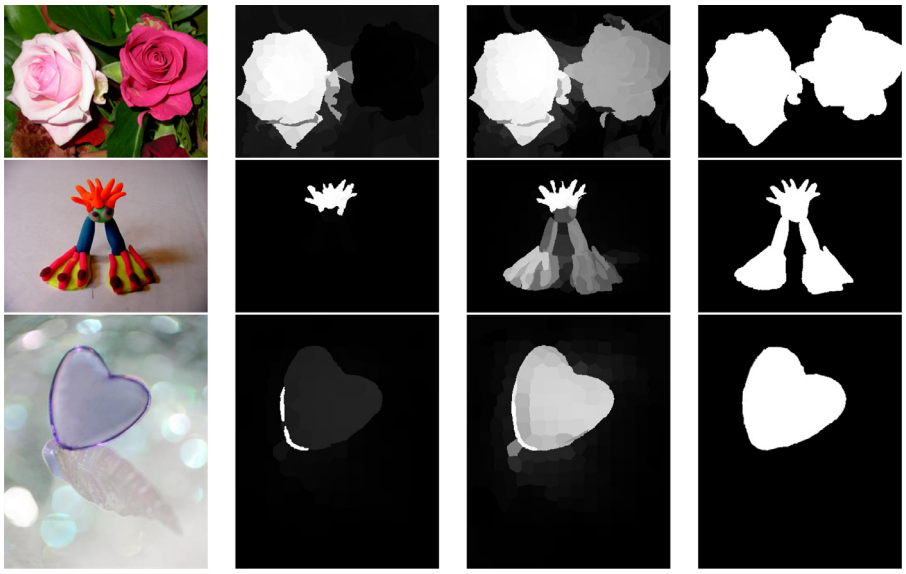
1. Introduction

Humans can identify the most salient and important area in a scene. In order to simulate this ability of human vision system in computer vision, more and more researchers pay attention to the study of the visual saliency detection. It has been a pre-processing procedure and widely used in many applications, such as image retrieval [1,2], object recognition [3,4], image compression [5,6] and so on.

Visual saliency detection can be classified into video saliency detection [7,8] and image saliency detection by using different input data. We focus on the second. Image saliency detection tends to find the salient regions, while the mission of segmentation [9,10] is to divide the digital image into multiple image subregions. Image saliency detection can be generally fell into three categories. Top-down methods[11–13] are task-driven by using the high-level knowledge. Bottom-up methods [14–16] are data-driven and rely on the assumptions of the background and foreground. Mixed models [17–19] are combined by top-down and bottom-up models. According to the technical method used, saliency detection can be divided into deep learning models [20–24] and traditional models. To avoid time-consuming training by deep learning models, we focus on traditional and bottom-up methods. Some state-of-the-art bottom-up saliency detection methods are presented in [25]. Boundary prior, contrast prior, boundary connectivity and so on are widespread to use in many models [14,26–28]. Itti et al. [14] propose a fusional model as a pioneer, which get saliency maps by fuse the color, direction and gray features of input images. With decades of development, more and more effective methods are appeared. Harel et al. [15] propose a graph-based visual saliency detection model. Gopalakrishnan et al. [29] firstly construct a completely graph and a k-regular graph by image patch and obtain the global and local properties of image. And then, they regard the saliency detection as Markov random walk on graphs to get the final results. Classically, Wei et al. [26] exploit geodesic saliency by using boundary and connectivity priors, which focus more on the background instead of the object. Yang et al. [27] calculate the saliency values by using a manifold ranking function and the relationships of all super-pixels. Zhu et al. [28] describe the boundary connectivity and present a general energy optimization framework to optimize the final results. Li et al. [30] propose a novel approach that take advantage of both region-based features and image details. Wang et al. [31] put forward a saliency detection approach by exploiting both local graph structure and background priors. Tu et al. [32] explain an minimum spanning tree representation instead of the super-pixel representation and propose an exact and iteration free solution on the tree. Xia et al. [33] try to find what is and what is not a salient object, and present a new model by ensembling linear exemplar regressors.

* Corresponding author.
E-mail address: tj@ahu.edu.cn (J. Tang).

https://doi.org/10.1016/j.neucom.2018.06.072
0925-2312/



(a) Input (b) MR (c) Ours (d) GT

Fig. 1. (a) Input image; (b) The results by using MR method; (c) The results by using our method; (d) The ground truth.

2. Related work

Manifold ranking is firstly proposed to exploit the intrinsic manifold structure of data [34], which is used in many computer vision problems including image retrieval [35], person re-identification [36], video concept detection [37], co-saliency detection [38,39], object tracking [40,41] and object co-segmentation [42]. He et al. [35] propose a a general transductive learning framework, in which they initialize a pseudo seed vector firstly and then spread its scores via manifold ranking to all the unlabeled images. Loy et al. [36] obtain results by propagating the query information along the unlabeled data manifold in their model. Tang et al. [37] raise structure sensitive manifold ranking model by taking local distribution differences into account to more accurately measure pairwise similarity by instead of using distance only. Yao et al. [38] present a novel co-saliency detection framework to solve the two sub-problems by using two-stage multi-view spectral rotation co-clustering. Han et al. [42] employ a novel two-stage co-segmentation framework to address the robustness issue by using the background prior instead of strong prior knowledge.

As a popular graph-based method, manifold ranking model plays an important role in saliency detection. But the traditional manifold ranking saliency model [27] is not considering the existing prior information. In actual conditions, the prior is useful to the saliency detection, which can lead to raise a better performance. We can get many prior information from multiple ways. For better to employ them, we propose a prior regularized graph ranking model (RegMR) to obtain the saliency maps. Our preliminary work on RegMR has been present in work [43]. Moreover, we extend our prior regularized graph ranking model (RegMR) to multi-layer case and propose a multi-layer RegMR model by using a multi-layer graph. The results are showed in Fig. 1. The main contributions of this work are enunciated as follows: Firstly, we propose a prior regularized multi-layer graph manifold ranking model to make better use of the existing prior information. Secondly, we can get the close-form solution and obtain the final results of the function. At last but not the least, we get more efficient results by many experiments.

3. Brief review of manifold ranking

For an image, Yang et al. [27] use simple linear iterative clustering (SLIC) algorithm [44] to gain n super-pixels as graph nodes in manifold ranking(MR) method. A graph $G = (V, E)$ is first constructed, where nodes V represent the super-pixels X and edges E denote the affinities $\mathbf{W}$ between pairs of super-pixels. Formally, let $X = \{x_1, x_2, \cdots, x_n\}$ be the set of super-pixels as graph nodes. Let $\mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_n)$ be the indication vector of queries. If x_i is a query super-pixel, $\mathbf{q}_i = 1$, else $\mathbf{q}_i = 0$. The aim of MR is to gain a ranking value $\mathbf{r}_i$ for each node $x_i \in X$ according to its relevance to the queries $\mathbf{q}$. Then, MR computes the optimal ranking $\mathbf{r}$ by solving

$$\min_{\mathbf{r}} \quad J_{\mathrm{MR}} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{W}_{ij}\left(\frac{\mathbf{r}_i}{\sqrt{\mathbf{d}_i}} - \frac{\mathbf{r}_j}{\sqrt{\mathbf{d}_j}}\right)^2 + \mu_0\sum_{i=1}^{n}(\mathbf{r}_i - \mathbf{q}_i)^2, \quad (1)$$

where $\mathbf{d}_i = \sum_{j=1}^{n}\mathbf{W}_{ij}$. It is known that the above MR model has a closed-from solution and the optimal solution [27] $\mathbf{r}^*$ is given by

$$\mathbf{r}^* = (\mathbf{I} - \frac{1}{1+\mu_0}\mathbf{S})^{-1}\mathbf{q},$$

where $\mathbf{S} = \mathbf{D}^{-\frac{1}{2}}\mathbf{W}\mathbf{D}^{-\frac{1}{2}}$, $\mathbf{D} = \mathrm{diag}(\mathbf{d}_1, \mathbf{d}_2, \ldots \mathbf{d}_n)$ and $\mathbf{I}$ is an identity matrix.

To get more effective result, MR model obtains another ranking function [27] by using the un-normalized Laplacian matrix as,

$$\mathbf{r}^* = (\mathbf{D} - \frac{1}{1+\mu_0}\mathbf{W})^{-1}\mathbf{q} \quad (2)$$

4. Prior regularized multi-layer graph ranking model

Model formulation. Let $X^k = \{x_1^k, x_2^k, \ldots, x_{N_k}^k\}$ be the set of graph nodes in L_k layer ($k = 1, 2, \ldots, K$), $\mathbf{W}^k$ is an affinity matrix of relationship between pairs of nodes in the L_k layer, $\mathbf{C}^{kk'}$ is an affinity matrix of relationship between pairs of nodes separately from the L_k and L_k' layers. Let $\mathbf{q}^k = (\mathbf{q}_1^k, \mathbf{q}_2^k, \ldots \mathbf{q}_{N_k}^k)$ be the indication vector of queries. if x_i^k is a query node, $\mathbf{q}_i^k = 1$, otherwise, $\mathbf{q}_i^k = 0$. Then, we can get the ranking value r_i^k of the node x_i in the L_k layer by

the multi-layer manifold ranking model as,

$$\min_{\mathbf{r}} J = \sum_{k=1}^{K} \frac{1}{2} \left[\sum_{i,j=1}^{N_k} \mathbf{W}_{ij}^k (\mathbf{r}_i^k - \mathbf{r}_j^k)^2 + \mu \sum_{i=1}^{N_k} \mathbf{d}_i^k (\mathbf{r}_i^k - \frac{\mathbf{q}_i^k}{\sqrt{\mathbf{d}_i^k}})^2 \right] + \eta \sum_{k,k'=1,k\neq k'}^{K} \sum_{i,j=1}^{N_k} \mathbf{C}_{ij}^{kk'} (\mathbf{r}_i^k - \mathbf{r}_j^{k'})^2,$$

where $\mathbf{d}_i^k = \sum_j \mathbf{W}_{ij}^k$. The first term is the smoothness constraint in the L_k layer, the second term is the fitting constraint and the third term is the smoothness between different layers L_k and $L_{k'}$.

Let $\mathbf{p}_i^k$ be the prior of the node x_i^k, the larger is $\mathbf{p}_i^k$, the less reference of the node x_i^k ranking value. Formally, by incorporating the prior regularization $\Psi(\mathbf{r}_i^k, \mathbf{p}_i^k), k = 1, \ldots, K$ in multi-layer manifold ranking model, the model can be formulated as

$$\min_{\mathbf{r}} J = \sum_{k=1}^{K} \frac{1}{2} \left[\sum_{i,j=1}^{N_k} \mathbf{W}_{ij}^k (\mathbf{r}_i^k - \mathbf{r}_j^k)^2 + \mu \sum_{i=1}^{N_k} \mathbf{d}_i^k (\mathbf{r}_i^k - \frac{\mathbf{q}_i^k}{\sqrt{\mathbf{d}_i^k}})^2 \right] + \eta \sum_{k,k'=1,k\neq k'}^{K} \sum_{i,j=1}^{N_k} \mathbf{C}_{ij}^{kk'} (\mathbf{r}_i^k - \mathbf{r}_j^{k'})^2 + \lambda \sum_{k=1}^{K} \sum_{i=1}^{N_k} \Psi(\mathbf{r}_i^k, \mathbf{p}_i^k),$$

where $\mathbf{p}_i^k$ denotes some prior information in the L_k layer. Many regularization functions can be used here. In this paper, we set $\Psi(\mathbf{r}_i^k, \mathbf{p}_i^k) = \mathbf{p}_i^k \mathbf{r}_i^{k2}$ and propose a prior regularized multi-layer graph ranking model (RegMR) as follows,

$$\min_{\mathbf{r}} J_{\text{RegMR}} = \sum_{k=1}^{K} \frac{1}{2} \left[\sum_{i,j=1}^{N_k} \mathbf{W}_{ij}^k (\mathbf{r}_i^k - \mathbf{r}_j^k)^2 + \mu \sum_{i=1}^{N_k} \mathbf{d}_i^k (\mathbf{r}_i^k - \frac{\mathbf{q}_i^k}{\sqrt{\mathbf{d}_i^k}})^2 \right] + \eta \sum_{k,k'=1,k\neq k'}^{K} \sum_{i,j=1}^{N_k} \mathbf{C}_{ij}^{kk'} (\mathbf{r}_i^k - \mathbf{r}_j^{k'})^2 + \lambda \sum_{k=1}^{K} \sum_{i=1}^{N_k} \mathbf{p}_i^k \mathbf{r}_i^{k2}, \quad (3)$$

Optimization. Our RegMR model is convex and the global optimal solution can be computed. Using vector representation, problem Eq.(3) is equivalently formulated as

$$\min_{\mathbf{r}} J_{\text{RegMR}} = \mathbf{r}^T (\mathbf{D} - \mathbf{E}) \mathbf{r} + \mu (\mathbf{r}^T \mathbf{D} \mathbf{r} - 2\mathbf{r}^T \mathbf{D}^{\frac{1}{2}} \mathbf{q} + \mathbf{q}^T \mathbf{q}) + \lambda \mathbf{r}^T \mathbf{P} \mathbf{r},$$

where $\mathbf{d}_i^k = \sum_j \mathbf{W}_{ij}^k$, $\mathbf{D}^k = diag(\mathbf{d}_1^k, \mathbf{d}_2^k \ldots, \mathbf{d}_{N_k}^k)$, $\mathbf{D} = diag(\mathbf{D}^1, \mathbf{D}^2, \ldots, \mathbf{D}^k)$, $\mathbf{P}^k = diag(\mathbf{p}_1^k, \mathbf{p}_2^k \ldots, \mathbf{p}_{N_k}^k)$, $\mathbf{P} = diag(\mathbf{P}^1, \mathbf{P}^2, \ldots, \mathbf{P}^k)$, and also

$$\mathbf{E} = \begin{pmatrix} \mathbf{W}^1 & 2\eta\mathbf{C}^{12} & \ldots & 2\eta\mathbf{C}^{1K} \\ 2\eta\mathbf{C}^{21} & \mathbf{W}^2 & \ldots & 2\eta\mathbf{C}^{2K} \\ \vdots & \vdots & \ddots & \vdots \\ 2\eta\mathbf{C}^{K1} & 2\eta\mathbf{C}^{K2} & \ldots & \mathbf{W}^K \end{pmatrix},$$

$$\mathbf{q} = \begin{pmatrix} \mathbf{q}^1 \\ \mathbf{q}^2 \\ \vdots \\ \mathbf{q}^K \end{pmatrix}, \mathbf{r} = \begin{pmatrix} \mathbf{r}^1 \\ \mathbf{r}^2 \\ \vdots \\ \mathbf{r}^K \end{pmatrix},$$

The optimal solution is computed by setting the first derivative of the above function $J_{\text{RegMR}}(\mathbf{r})$ w.r.t $\mathbf{r}$ to be zero, i.e.,

$$\frac{\partial J_{\text{RegMR}}(\mathbf{r})}{\partial \mathbf{r}} = (\mathbf{D}\mathbf{r} - \mathbf{E}\mathbf{r}) + \mu(\mathbf{D}\mathbf{r} - \mathbf{D}^{\frac{1}{2}}\mathbf{q}) + \lambda \mathbf{P}\mathbf{r} = \mathbf{0}$$

Thus, we obtain the result,

$$\mathbf{r}^* = \left(\mathbf{D} - \frac{1}{1+\mu}\mathbf{E} + \frac{\lambda}{1+\mu}\mathbf{P} \right)^{-1} \mathbf{q} \quad (4)$$

5. Saliency detection

Saliency detection model via super-pixels can get more details but ignore the overall information when the super-pixel is small. On the other hand, if the super-pixel is too large, then the small salient object can not be detected efficiently. And also, the above MR model [27] has one limitation, which fails to consider the prior (background or foreground possibility) information in its ranking process. For saliency detection tasks, the prior information has been shown importantly in saliency computation problem. Our aim in this section is to propose a prior regularized multi-layer graph ranking model to overcome the limitations. We describe the details of the saliency detection progress. At first, we construct graphs in multiple layers. Additionally, we calculate the prior vector by using the boundary connectivity. At last, we describe the saliency detection progress based on the prior regularized multi-layer graph model.

5.1. The multi-layer graph construction

We construct graphs in this section. One bottom layer L_1 is constructed by super-pixels, which can get more details of image information. We use simple linear iterative clustering (SLIC) algorithm [44] to gain N_1 super-pixels at the bottom layer, which is showed in the left of the Fig. 2. The higher $K-1$ layers $L_k, k = 2, 3, \ldots, K$ are constructed by lager region, which are clustering by super-pixels and can get more global information. In $K-1$ higher layer, we cluster the bottom super-pixels as a region. We use spectral segmentation as detailed in [45] to get $K-1$ high layers($L_k, k = 2, 3, \ldots, K$) graph nodes that is explained in the right of Fig. 2, the nodes number are $N_2, N_3, \ldots, N_K$.

The details are explained as follows.

(1) Inner the L_k layer: Each node is connected to those neighbor nodes within the layer L_k. And also the nodes are connected to the 2-hop neighbor nodes in the L_1 layer, in which all nodes on four boundaries are adjacent by constricting the graph as a close-loop graph. When $x_i, x_j \in L_k, k = 1, 2, \ldots, K$ and x_i is connected to x_j,

$$\mathbf{W}_{ij}^k = e^{-\frac{\|x_i - x_j\|}{\sigma_k^2}}, \quad (5)$$

where x_i and x_j denote the mean of the super-pixels corresponding to two nodes in the CIE LAB color space. σ_k is a parameter.

(2) Inter the L_k layer: The bottom layer nodes are clustering into higher layer graph nodes by using spectral segmentation methods, then these bottom layer graph nodes are connected to the higher layer graph node.

$$\mathbf{C}_{ij}^{kk'} = \begin{cases} \gamma & \text{if} \quad k = 1 \quad \text{or} \quad k' = 1, k \neq k' \\ 0 & \text{else} \end{cases}, \quad (6)$$

where γ is a parameter.

5.2. The prior by boundary connectivity

By reduce the influence of wrong region prior information of the lager region, we focus on the prior of super-pixels in the bottom layer, which can be calculated by any method. We introduce the boundary connectivity here to indicate the background and foreground prior value. The connotation of the boundary connectivity [28] is the proportion of the region occupied the boundary comparing with the square root of the whole area. If boundary connectivity $BndCon(x_i)$ of node x_i is low, the foreground probability $\mathbf{p}_i^{fg}$ of node x_i is high, the background probability $\mathbf{p}_i^{bg}$ is low. Then, the foreground probability and the background probability by boundary connectivity [28] are calculated as,

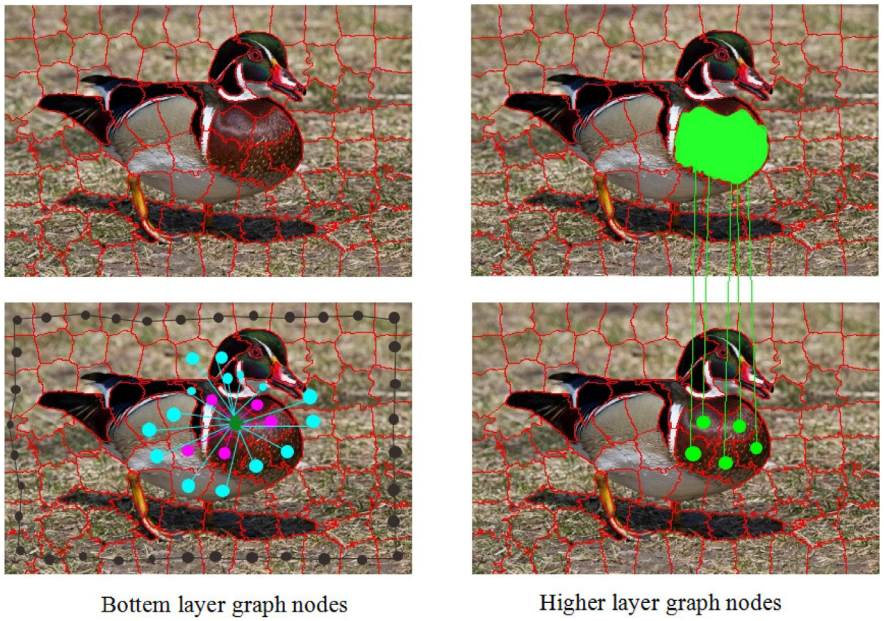


Fig. 2. Left: Supe-rpixels are gained by SLIC algorithm on the original image to get the L_1 layer. The pink and blue dots are the directly neighbors and 2-hop neighbors of the green dot. The black dots are the boundary nodes. **Right**: The $L_k(k = 2, 3, \cdots, K)$ layer nodes are gained by clustering the L_1 layer nodes such as the green region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

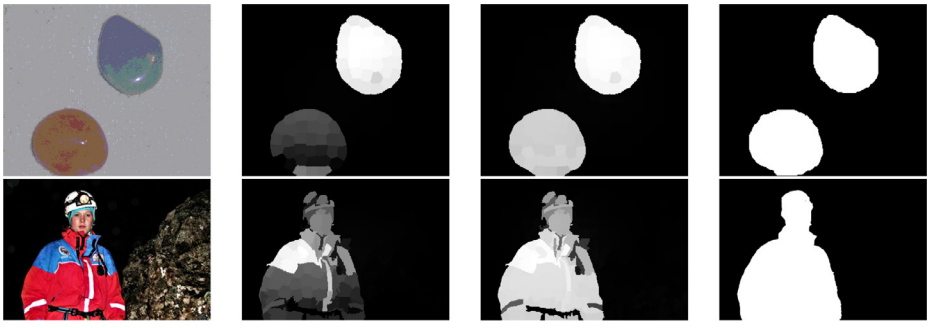


Fig. 3. (a) Input image; (b) Saliency map of the first stage results by our model; (c) Saliency map of the second stage results by our model(final results); (d) The ground truth.

$$\mathbf{p}_i^{fg} = \begin{cases} e^{-\frac{BndCon^2(x_i)}{2\sigma_p^2}} & \text{if } x_i \in X^1 \\ 0 & \text{if } x_i \in X^k, k = 2, 3, \ldots, K \end{cases} \tag{7}$$

$$\mathbf{p}_i^{bg} = \begin{cases} 1 - e^{-\frac{BndCon^2(x_i)}{2\sigma_p^2}} & \text{if } x_i \in X^1 \\ 0 & \text{if } x_i \in X^k, k = 2, 3, \ldots, K \end{cases} \tag{8}$$

where $BndCon(x_i)$ is the boundary connectivity of super-pixel x_i and σ_p is a parameter.

5.3. Saliency detection via prior regularized multi-layer graph ranking model

We describe the saliency detection progress based on the prior regularized multi-layer graph model in this section. We use two stages ranking method to get the final saliency maps. Fig. 3 give example images ranking result in two stage.

In the first stage, boundary super-pixels on the bottom layer are used as query nodes. We separately use the super-pixels in the four boundary as background queries, which is top, down, left and right. $\mathbf{q}_1^m (m = top, down, left, right)$ are the query vectors. The prior matrix $\mathbf{P}^{fg}$ can be gained by prior value $\mathbf{p}_i^{fg}$, which is the foreground probability of node x_i.

We can get the result as,

$$\mathbf{r}_1^m = \left(\mathbf{D} - \frac{1}{1+\mu}\mathbf{E} + \frac{\lambda}{1+\mu}\mathbf{P}^{fg}\right)^{-1} \mathbf{q}_1^m, m = top, down, left, right \tag{9}$$

Then, the first N_1 ranking results are selected in the first stage result and be normalized to [0,1] as $\tilde{\mathbf{r}}_1^m$.

$$\mathbf{s}_{1i} = \prod_{i=1,\ldots,N_1; m=top,down,left,right} \mathbf{s}_{1i}^m, \tag{10}$$

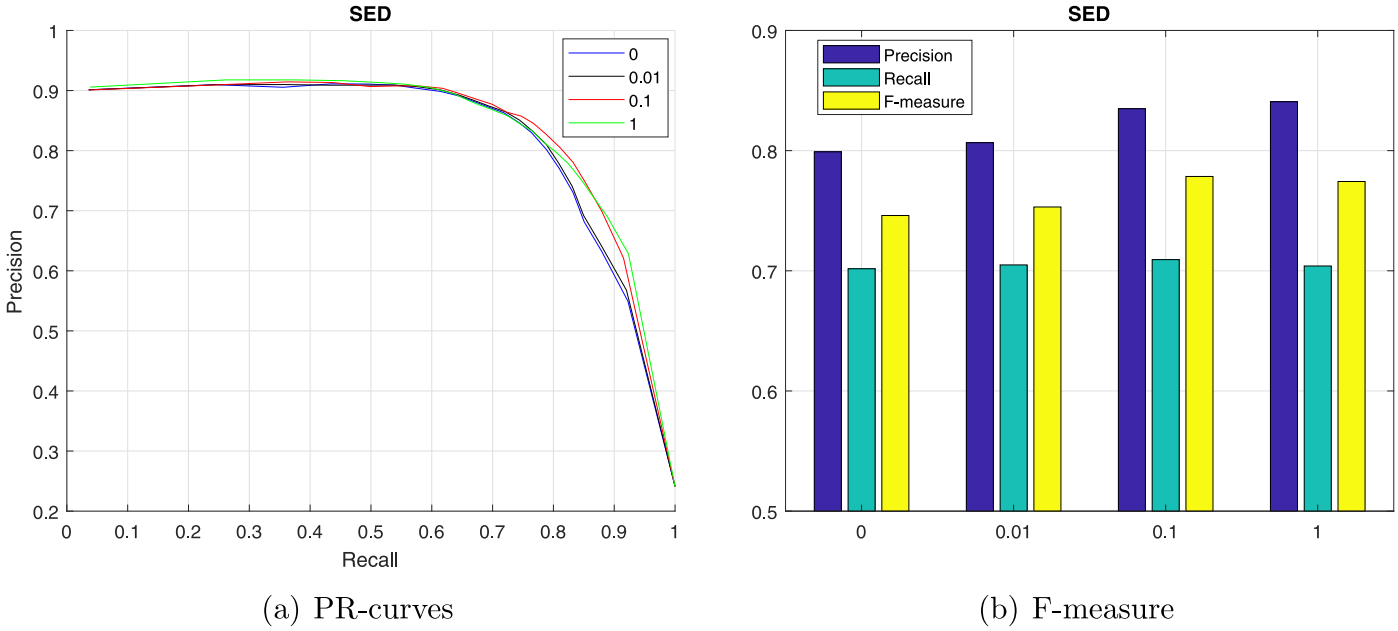
where

$$\mathbf{s}_{1i}^m = 1 - \tilde{\mathbf{r}}_{1i}^m, m = top, down, left, right. \tag{11}$$

In the second stage, we choose some salient nodes as queries, if $\mathbf{s}_{1i} \geq mean(\mathbf{s}_1)$, then $\mathbf{q}_{2i} = 1$, else $\mathbf{q}_{2i} = 0$, we can summarize the query nodes as follows,

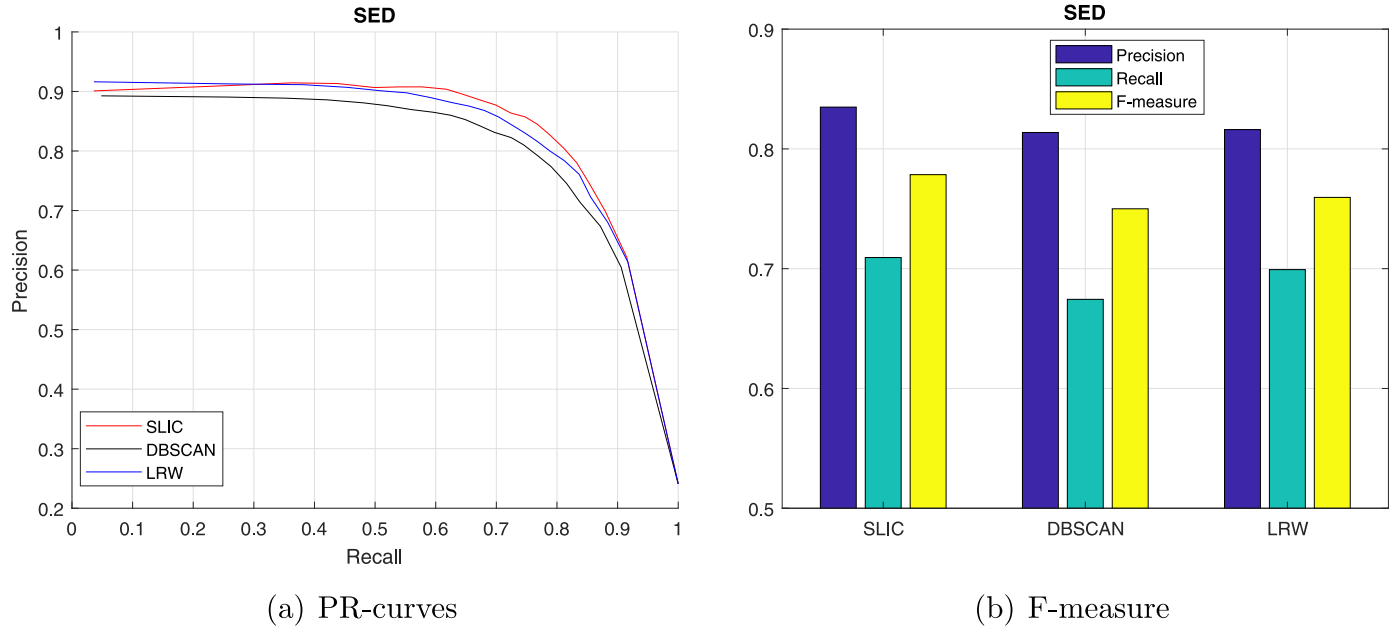
$$\mathbf{q}_{2i} = \begin{cases} 1 & \text{if } \mathbf{s}_{1i} \geq mean(\mathbf{s}_1) \text{ and } x_i \in X^1 \\ 0 & \text{if } \mathbf{s}_{1i} < mean(\mathbf{s}_1) \text{ and } x_i \in X^1 \\ 0 & \text{if } x_i \in X^k, k = 2, 3, \ldots, K \end{cases} \tag{12}$$

The prior matrix $\mathbf{P}^{bg}$ can be gained by prior value $\mathbf{p}_i^{bg}$, which is the background probability of node Then, we compute the refinement



(a) PR-curves (b) F-measure

Fig. 4. Results of our model with different value of β comparing on SED database.



(a) PR-curves (b) F-measure

Fig. 5. Comparison of results by using different superpixel segmentation method. From top to bottom, the results are obtained by SLIC, DBSCAN, and LRW method, respectively.

saliency map by,

$$\mathbf{s} = \mathbf{r}_2 = \left(\mathbf{D} - \frac{1}{1+\mu}\mathbf{E} + \frac{\lambda}{1+\mu}\mathbf{P}^{bg}\right)^{-1}\mathbf{q}_2, \tag{13}$$

where $\mathbf{P}^{bg} = \mathrm{diag}(\mathbf{p}_i^{bg})$. Then, we get the second stage results. And also, we choose the first N_1 ranking results and normalize to [0,1] as $\bar{\mathbf{s}}$, then give the super-pixel ranking value to all pixel as $\mathbf{f}$, which is the saliency results of all pixels.

The overall algorithm is summarized in Algorithm 1.

6. Experiments

For verifying the effectiveness of our method, we compare with fifteen state-of-the-art saliency object detection methods including CA [46], FT [47], SEG [48], BM [49], SWD [50], SF [51], GCHC [52], LMLC [53], HS [54], PCA [55], MR [27], MS [56], RR [30], MST [32] and MAP [57].

6.1. Datasets

We compare our proposed method on four public datasets: The first one is ECSSD [54], which contains 1000 images. The second one is ASD databse [47], which also contains 1000 images. SED [58] contains 200 images, in which 100 images have only one salient object and the other 100 images have two salient objects. PASCAL-S dataset [59] is one of the most challenging saliency datasets, which is composed of 850 natural images.

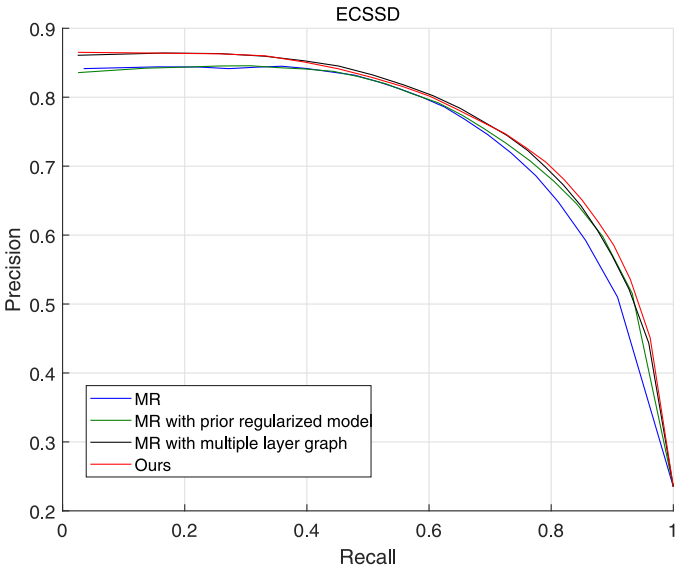


Fig. 6. Results of our model with different cues comparing on ECSSD database: The blue curve is the PR-curve of MR model; The green curve is the PR-curve of the our model on only bottom layer with N_0 super-pixel ; The black curve is the PR-curve of MR model with multi-layer graph; The red curve is our model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6.2. Parameter settings

In all experiments, Multi-layer K is set to 4, in which we set the number of super-pixels $N_1 = 300$ in the bottom layer L_1, $N_2 = 80$, $N_3 = 50$, $N_4 = 30$. Multi-layer graph construction parameters $\gamma = 0.1$, $\sigma_1^2 = 10$, $\sigma_2^2 = \sigma_3^2 = \sigma_4^2 = 5$. The parameter σ_p in prior is set to 1. Besides, there are three controlling parameters, $\alpha = \frac{1}{1+\mu}$ is setting to 0.99 as [27], $\beta = \frac{\lambda}{1+\mu}$ is setting to 0.1 by experiments in Fig. 4, and $\eta = \frac{1}{2}$ by experience.

Algorithm 1 Saliency computation by the proposed model.

Require: An image and required parameters.

1: Segment the input image into N_1 super-pixels as bottom layer nodes, construct 2-neighbor graph and get $\mathbf{W}^1$ by Eq. (5);
2: Achieve $K-1$ higher layer regions by clustering bottom layer super-pixels as graph nodes $N_2, \ldots, N_K$, construct graphs and get $\mathbf{W}^2, \ldots, \mathbf{W}^K$ by Eq. (5);
3: Compute the edge weight between higher layer and bottom layer to get $\mathbf{C}^{k_1 k_2}, k_1 \neq k_2, k_1, k_2 = 1, \ldots, K$ by Eq. (6);
4: Compute the foreground probability and the background probability with Eqs. (7) and (8);
5: Get the saliency results with Eqs. (9) and (10) in the first stage;
6: Get the saliency results with Eqs. (12) and (13) in the second stage ;
7: Use the normalized ranking score to get the saliency value $\bar{\mathbf{s}}$ of super-pixels, then set the super-pixel values to all pixels and get the final saliency map $\mathbf{f}$.

Ensure: A saliency map $\mathbf{f}$ with the same size as the input image.

6.3. Evaluation criterions

In the comparisons, we use precision-recall curves (PR-curves), F-measure and mean absolute error (MAE) to evaluate all methods with our model. PR-curves are constructed by comparing the saliency map with ground truth by thresholds in the range [0:0.05:1]. Precision is the ratio of the number of correctly salient pixels to the number of all ground truth salient pixels. Recall is the percentage of all selected salient pixels number to the number of all ground truth pixels. After the calculating of the precision and recall, then we can gain the F-measure, which is an overall performance and calculated as,

$$F = \frac{(1+\xi)Precision \cdot Recall}{\xi Precision + Recall}, \tag{14}$$

referred to [27,47], we set $\xi = 0.3$ in our experiments.

The mean absolute error (MAE) is another measurement, which is the average difference between saliency map and ground truth. To gain MAE value, we should normalize both of saliency map and ground truth in the range [0, 1]. The difference can be measured

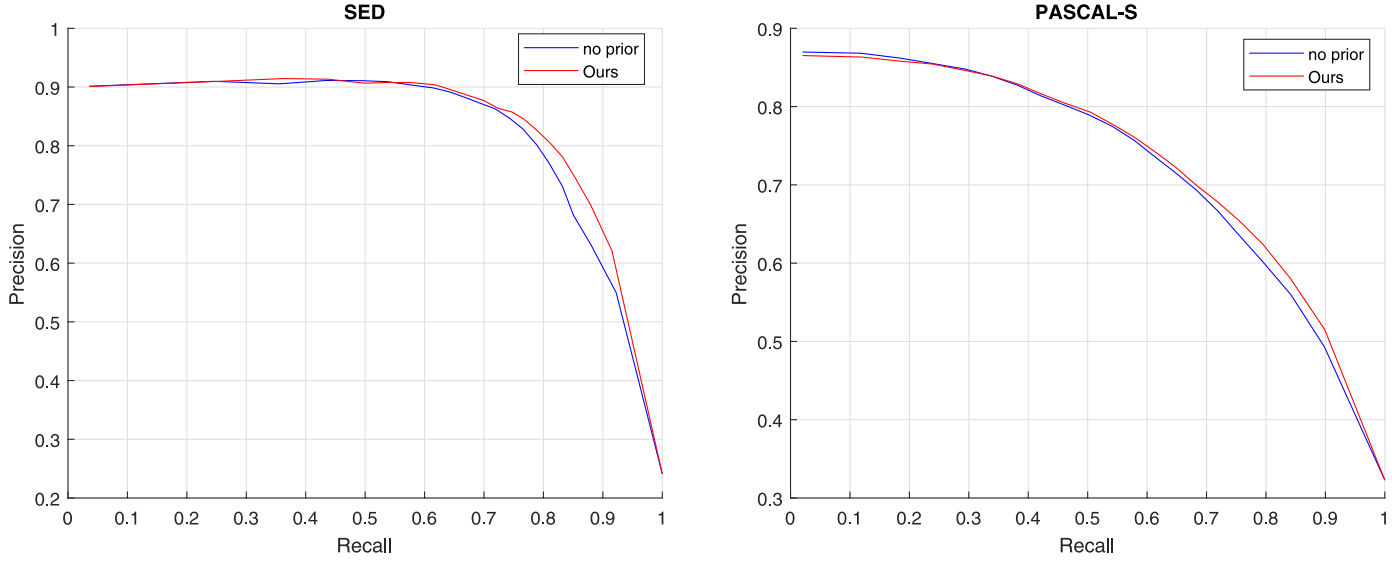
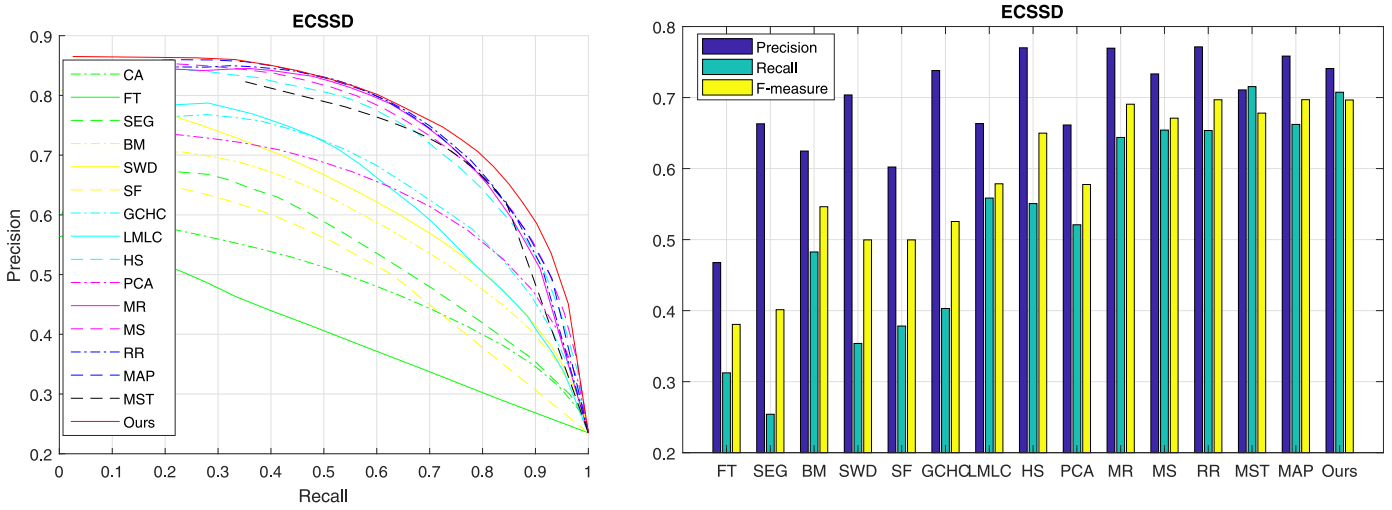
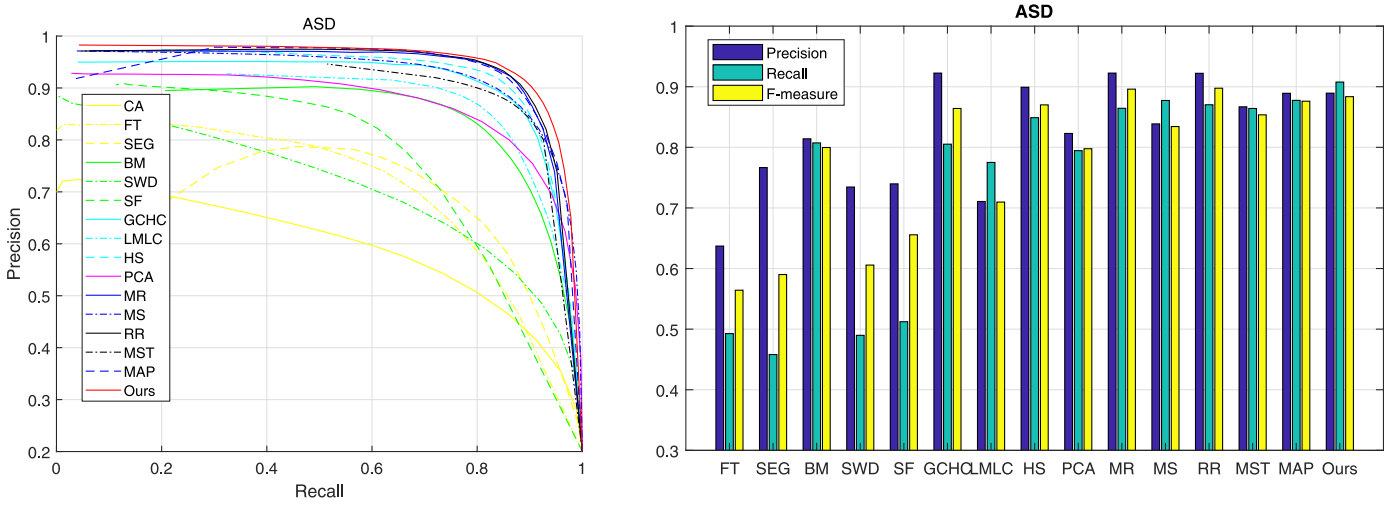


Fig. 7. Results of our model with no prior information comparing on SED and PASCAL-S databases: The blue curve is the PR-curve of MR model; The green curve is the PR-curve of the our model on only bottom layer with N_0 super-pixel ; The black curve is the PR-curve of MR model with no prior information; The red curve is our model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



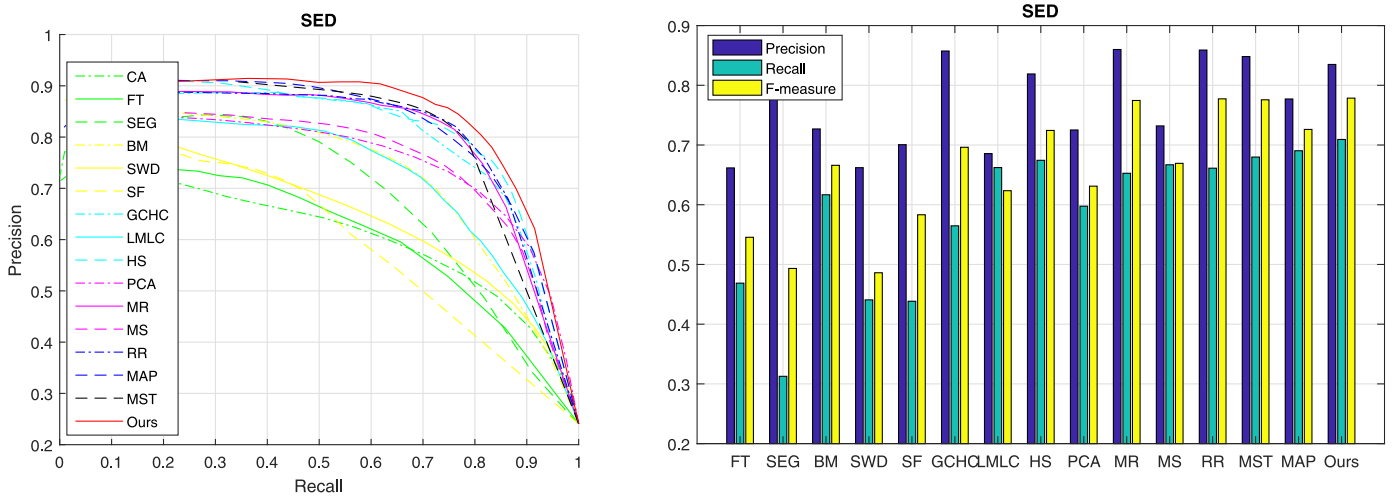
(a) PR-curves (b) F-measure

Fig. 8. Precision-recall curves and F-measures comparing with different methods on ECSSD database.



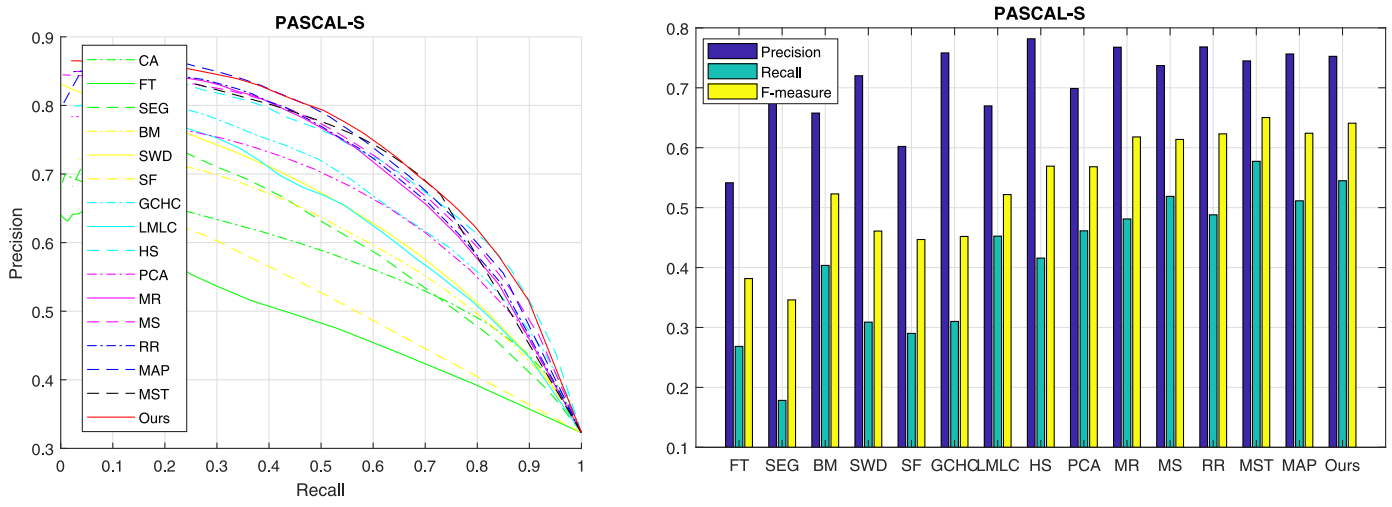
(a) PR-curves (b) F-measure

Fig. 9. Precision-recall curves and F-measures comparing with different methods on ASD database.



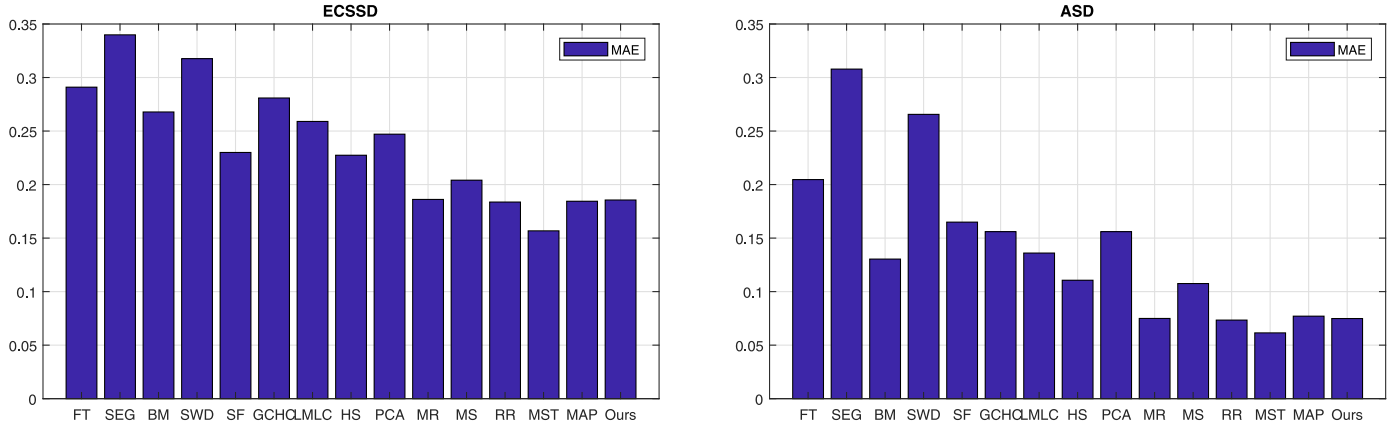
(a) PR-curves (b) F-measure

Fig. 10. Precision-recall curves and F-measures comparing with different methods on SED database.



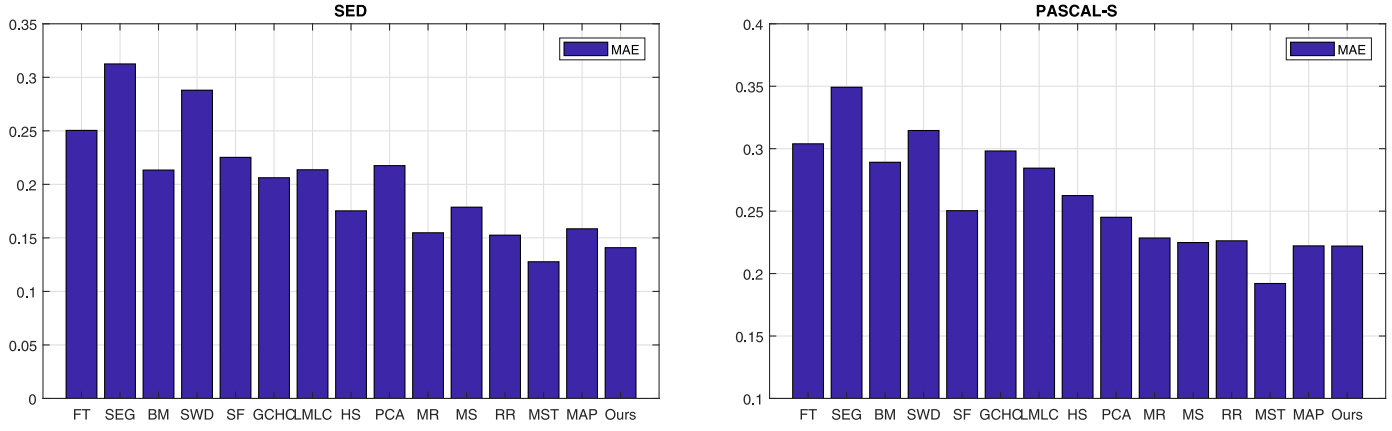
(a) PR-curves (b) F-measure

Fig. 11. Precision-recall curves and F-measures comparing with different methods on PASCAL-S database.



(a) ECSSD (b) ASD

Fig. 12. MAE values comparing with different methods on ECSSD and ASD datasets.



(a) SED (b) PASCAL

Fig. 13. MAE values comparing with different methods on SED and PASCAL-S datasets.

Fig. 14. Examples of output saliency maps results using different algorithms on the ECSSD, ASD, SED and PASCAL-S datasets.

as [32],

$$MAE = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} |\mathbf{f}(i, j) - GT(i, j)|, \tag{15}$$

where H, W denote the height and width of image, respectively. $\mathbf{f}(i, j)$ is the saliency value of pixel level, $GT(i, j)$ is the ground truth of pixel level.

6.4. Examinations of design options

The selection of the superpixel algorithm [60] has a strong effect to the saliency detection results. We do some experiments by different superpixel segmentation method including SLIC [44], DBSCAN [61] and LRW [62]. The SLIC method is as quick as the DBSCAN, but the LRW pays much time on segmentation. And also, the

Fig. 15. Failure cases. (a) Input images. (b) Ground truth. (c) Saliency maps.

Table 1
Comparison of average run time (seconds per image) on SED datasets.

Methods	SEG	BM	SWD	SF	GCHC	PCA
Time(s)	2.83	0.96	0.07	0.34	0.58	2.59
Methods	MR	MS	RR	MST	MAP	Ours
Time(s)	0.49	2.03	0.91	0.40	0.05	1.54

results of PR-curves and F-value from Fig. 5 show that SLIC method is the best for our model.

To analyze the components considered in our mode, we do experiments on ECSSD under different circumstance and give the PR-curve on Fig. 6, which shows that our method is better than other conditions. The blue curve is the PR-curve of MR model; The green curve is the PR-curve of the MR with prior regularized model; The black curve is the PR-curve of MR model with only multi-layer graph; The red curve is the prior regularized multi-layer graph ranking model. Fig. 7 shows the results of our model comparing with no prior information model on SED and PASCAL-S datasets.

6.5. Comparisons with other models

We have many experiments on ECSSD, ASD, SED and PASCAL-S datasets. Fig. 8 displays the results of our method comparing with other methods on ECSSD database. Figs. 9–Fig. 11show the comparison results of PR-curves and F-measure separately on ASD, SED and PASCAL-S datasets.

From above evaluation, our method is better than fourteen state-of-the-art methods by measuring PR-curves, which also performances well by comparing with others on F-measures except MR and RR algorithm on ASD database. The MAE value of the proposed method is slightly lower than MST method on all four databases which can be see in Figs. 12 and13.

The average runtime of some state-of-the-art models are shown in Table. 1 over all 200 images of SED using an Intel i5-5200U 2.20 GHz CPU with 8 GB RAM. CA and LMLC are obviously slower than our method. Here, we just list some methods which has similar computational cost with our method. MAP model here is the fastest (about 0.05 s per image), others are in a same order of magnitude.

Fig. 14 shows some sample saliency maps from four datasets. Our method can get more accurately saliency maps intuitively. On the whole, the proposed method is better than other fifteen methods.

For deep learning based models, such as DLS [21] and dss [20], they generally performs better than the proposed method which is unsupervised. Indeed, supervised methods usually performs better than unsupervised methods because they utilize a larger number of training data to learn a more discriminative mode for saliency model detection problem. When the training data are not available in some special case, the unsupervised method can be used.

6.6. Failure cases

In this work, we propose a prior regularized multi-layer graph ranking model by using multiple graph model and existing prior information. The proposed algorithm is highly effective for most tasks of saliency detection. However, when salient objects are very small and have similar appearances with the background and the images are filled with the cluttered background, our algorithm cannot make the entire salient object be highlighted homogeneously as shown in Fig. 15.

7. Conclusion

We propose a prior regularized multi-layer graph ranking model, which can make full use of existing prior information and overcome the influence of only super-pixel layer to a certain extent. The proposed model effectively combines the prior information with the manifold structure information and also get more global information by multi-layer graph model. We further demonstrate that the new method achieve higher results comparing with fifteen state-of-the-art methods. In the future, we will extend our method in two ways. Firstly, our model is efficiently on the images which have salient objects. But the model can not estimate that whether the image have a salient object or not in actually. We will do some research on the judgement of whether the image have a salient object before the saliency detection. Secondly, more kinds of features will be used to obtain the intrinsic relationship in the images.

Acknowledgments

This work was sponsored by the National Natural Science Foundation of China (Nos. 61472002, 61602001, 61602006), the Natural Science Foundation of Anhui Higher Education Institution of China (KJ2015A110) and by Co-Innovation Center for Information Supply & Assurance Technology (No. ADXXBZ201607), Anhui University.

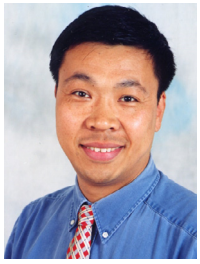
References

[1] T. Chen, M.M. Cheng, P. Tan, A. Shamir, S.M. Hu, Sketch2Photo: internet image montage, ACM Trans. Gr. (TOG) 28 (5) (2009) 1–10.
[2] M.M. Cheng, N.J. Mitra, X. Huang, S.M. Hu, Salientshape: group saliency in image collections, V. Comput. 30 (4) (2014) 443–453.
[3] Z. Ren, S. Gao, L.T. Chia, W.H. Tsang, Region-based saliency detection and its application in object recognition, IEEE Trans. Circuits Syst. Video Technol. 24 (5) (2014) 769–779.
[4] D. Tao, L. Xu, L. Jin, X. Li, Principal component 2-D long short-term memory for font recognition on single chinese characters, IEEE Trans. Cybern. 46 (3) (2016) 25–27.
[5] L. Itti, Automatic foveation for video compression using a neurobiological model of visual attention, IEEE Trans. Image Process. 13 (10) (2004) 1304–1318.
[6] Y. Fang, Z. Chen, W. Lin, C.W. Lin, Saliency detection in the compressed domain for adaptive image retargeting, IEEE Trans. Image Process. 21 (9) (2012) 3888–3901.
[7] D. Zhang, J. Han, L. Jiang, S. Ye, X. Chang, Revealing event saliency in unconstrained video collection, IEEE Trans. Image Process. 26 (4) (2017) 1746–1758.
[8] W. Wang, J. Shen, L. Shao, Video salient object detection via fully convolutional networks, IEEE Trans. Image Process. 27 (1) (2018) 38–49.
[9] X. Dong, J. Shen, L. Shao, L. Van Gool, Sub-Markov random walk for image segmentation, IEEE Trans. Image Process. 25 (2) (2016) 516–527.
[10] J. Shen, J. Peng, X. Dong, L. Shao, F. Porikli, Higher order energies for image segmentation, IEEE Trans. Image Process. 26 (10) (2017) 4911–4922.
[11] J. Wang, H. Jiang, Z. Yuan, M.M. Cheng, X. Hu, N. Zheng, Salient object detection: a discriminative regional feature integration approach, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2083–2090.
[12] S. Lu, V. Mahadevan, N. Vasconcelos, Learning optimal seeds for diffusion-based salient object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2790–2797.
[13] J. Kim, D. Han, Y.W. Tai, J. Kim, Salient region detection via high-dimensional color transform and local spatial support., IEEE Trans. Image Process. 25 (1) (2016) 9–23.
[14] L. Itti, C. Koch, E. Niebur, A model of saliency-based visual attention for rapid scene analysis, IEEE Trans. Pattern Anal. Mach. Intell. 20 (11) (1998) 1254–1259.
[15] J. Harel, C. Koch, P. Perona, Graph-based visual saliency, Adv. Neural Inf. Process. Syst. 1 (5) (2006) 545–552.
[16] B. Jiang, L. Zhang, H. Lu, C. Yang, M.H. Yang, Saliency detection via absorbing Markov chain, in: Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 1665–1672.
[17] L. Itti, C. Koch, Feature combination strategies for saliency-based visual attention systems, J. Electron. Image. 10 (1) (2001) 161–169.
[18] V. Navalpakkam, L. Itti, An integrated model of top-down and bottom-up attention for optimizing detection speed, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2006, pp. 2049–2056.
[19] A. Borji, D.N. Sihite, L. Itti, Probabilistic learning of task-specific visual attention, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2012, pp. 470–477.
[20] Q. Hou, M.-M. Cheng, X. Hu, A. Borji, Z. Tu, P. Torr, Deeply supervised salient object detection with short connections, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, pp. 5300–5309.
[21] P. Hu, B. Shuai, J. Liu, G. Wang, Deep level sets for salient object detection, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, pp. 2300–2309.
[22] D. Zhang, J. Han, Y. Zhang, Supervision by fusion: towards unsupervised learning of deep salient object detector, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4048–4056.
[23] W. Wang, J. Shen, Deep visual attention prediction[J]., IEEE Trans. Image Process. 27 (5) (2018) 2368–2378.
[24] J. Han, D. Zhang, G. Cheng, N. Liu, D. Xu, Advanced deep-learning techniques for salient and category-specific object detection: a survey, IEEE Signal Process. Mag. 35 (1) (2018) 84–100.
[25] A. Borji, M.M. Cheng, H. Jiang, J. Li, Salient object detection: a benchmark, IEEE Trans. Image Process. 24 (12) (2015) 5706–5722.
[26] Y. Wei, F. Wen, W. Zhu, J. Sun, Geodesic saliency using background priors, in: Proceedings of the European Conference on Computer Vision, 2012, pp. 29–42.
[27] C. Yang, L. Zhang, H. Lu, X. Ruan, M.H. Yang, Saliency detection via graph-based manifold ranking, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3166–3173.
[28] W. Zhu, S. Liang, Y. Wei, J. Sun, Saliency optimization from robust background detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2814–2821.
[29] V. Gopalakrishnan, Y. Hu, D. Rajan, Random walks on graphs for salient object detection in images, IEEE Transactions on Image Processing 19 (12) (2010) 3232–3242.
[30] C. Li, Y. Yuan, W. Cai, Y. Xia, D.D. Feng, Robust saliency detection via regularized random walks ranking, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2710–2717.
[31] Q. Wang, W. Zheng, R. Piramuthu, Grab: visual saliency via novel graph model and background priors, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 535–543.
[32] W.C. Tu, S. He, Q. Yang, S.Y. Chien, Real-time salient object detection with a minimum spanning tree, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2334–2342.
[33] C. Xia, J. Li, X. Chen, A. Zheng, Y. Zhang, What is and what is not a salient object? learning salient object detector by ensembling linear exemplar regressors, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4399–4407.
[34] D. Zhou, J. Weston, A. Gretton, O. Bousquet, B. Schölkopf, Ranking on data manifolds, in: Proceedings of the Advances in Neural Information Processing Systems, 2004, pp. 169–176.
[35] J. He, M. Li, H.-J. Zhang, H. Tong, C. Zhang, Generalized manifold-ranking-based image retrieval, IEEE Trans. Image Process. 15 (10) (2006) 3170–3177.
[36] C.C. Loy, C. Liu, S. Gong, Person re-identification by manifold ranking, in: Proceedings of the Twentieth IEEE International Conference on Image Processing (ICIP), IEEE, 2013, pp. 3567–3571.
[37] J. Tang, X.-S. Hua, G.-J. Qi, M. Wang, T. Mei, X. Wu, Structure-sensitive manifold ranking for video concept detection, in: Proceedings of the ACM Multimedia, 2007, pp. 852–861.
[38] X. Yao, J. Han, D. Zhang, F. Nie, Revisiting co-saliency detection: a novel approach based on two-stage multi-view spectral rotation co-clustering, IEEE Trans. Image Process. 26 (7) (2017) 3196–3209.
[39] D. Zhang, D. Meng, J. Han, Co-saliency detection via a self-paced multiple-instance learning framework, IEEE Trans. Pattern Anal. Mach. Intell. 39 (5) (2017) 865–878.
[40] C. Li, H. Cheng, S. Hu, X. Liu, J. Tang, L. Lin, Learning collaborative sparse representation for grayscale-thermal tracking, IEEE Trans. Image Process. 25 (12) (2016) 5743–5756.
[41] C. Li, X. Liang, Y. Lu, N. Zhao, J. Tang, RGB-T object tracking: benchmark and baseline, arXiv:1805.08982(2018).
[42] J. Han, R. Quan, D. Zhang, F. Nie, Robust object co-segmentation using background prior, IEEE Trans. Image Process. 27 (4) (2018) 1639–1651.
[43] Y. Xiao, B. Jiang, Z. Tu, J. Tang, Visual saliency detection via prior regularized manifold ranking, in: Proceedings of the CCF Chinese Conference on Computer Vision, 2017, pp. 711–722.
[44] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Ssstrunk, Slic superpixels, Technical Report 149300, Epfl(2010).
[45] L. Zhang, C. Yang, H. Lu, R. Xiang, M.H. Yang, Ranking saliency, IEEE Trans. Pattern Anal. Mach. Intell. 39 (9) (2017) 1892–1904.
[46] S. Goferman, L. Zelnik-Manor, A. Tal, Context-aware saliency detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2010, pp. 2376–2383.
[47] R. Achanta, S. Hemami, F. Estrada, S. Susstrunk, Frequency-tuned salient region detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 1597–1604.
[48] E. Rahtu, J. Kannala, M. Salo, J. Heikkil?, Segmenting salient objects from images and videos, in: Proceedings of the Computer Vision-ECCV, 2010, pp. 366–379.
[49] Y. Xie, H. Lu, Visual saliency detection based on Bayesian model, in: Proceedings of the IEEE International Conference on Image Processing, 2011, pp. 645–648.
[50] L. Duan, C. Wu, J. Miao, L. Qing, Y. Fu, Visual saliency detection by spatially weighted dissimilarity, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 473–480.
[51] F. Perazzi, P. Krahenbuhl, Y. Pritch, A. Hornung, Saliency filters: contrast based filtering for salient region detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 733–740.
[52] C. Yang, L. Zhang, H. Lu, Graph-regularized saliency detection with convex-hull-based center prior, IEEE Signal Process. Lett. 20 (7) (2013) 637–640.
[53] Y. Xie, H. Lu, M.H. Yang, Bayesian saliency via low and mid level cues, IEEE Trans. Image Process. 22 (5) (2013) 1689–1698.
[54] Q. Yan, L. Xu, J. Shi, J. Jia, Hierarchical saliency detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 1155–1162.
[55] R. Margolin, A. Tal, L. Zelnik-Manor, What makes a patch distinct? in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 1139–1146.
[56] N. Tong, H. Lu, L. Zhang, X. Ruan, Saliency detection with multi-scale superpixels, IEEE Signal Process. Lett. 21 (9) (2014) 1035–1039.
[57] J. Sun, H. Lu, X. Liu, Saliency region detection based on markov absorption probabilities, IEEE Trans. Image Process. 24 (5) (2015) 1639–1649.
[58] S. Alpert, M. Galun, R. Basri, A. Brandt, Image segmentation by probabilistic bottom-up aggregation and cue integration, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1–8.
[59] Y. Li, X. Hou, C. Koch, J.M. Rehg, A.L. Yuille, The secrets of salient object segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2014, pp. 280–287.
[60] X. Dong, J. Shen, L. Shao, Hierarchical superpixel-to-pixel dense matching, IEEE Trans. Circuits Syst. Video Technol. (2016) 1–9.
[61] J. Shen, X. Hao, Z. Liang, Y. Liu, W. Wang, L. Shao, Real-time superpixel segmentation by DBSCAN clustering algorithm, IEEE Trans. Image Process. 25 (12) (2016) 5933–5942.
[62] J. Shen, Y. Du, W. Wang, X. Li, Lazy random walks for superpixel segmentation, IEEE Trans. Image Process. 23 (4) (2014) 1451–1462.

Yun Xiao received the B.Eng. degree in computer science in 2011 from Anhui University, Hefei, China, in 2011. She is currently a Lecturer and a Ph.D. student in computer science at Anhui University. Her current research interests include computer vision and saliency detection.

Bo Jiang received the Ph.D. degree in computer science from Anhui University, Hefei, China, in 2015. He is currently an Associate Professor at computer science at Anhui University. His current research interests include image and graph matching, image feature extraction, and graph-based representation and learning.

Zhengzheng Tu received the Ph.D. degree in computer science from Anhui University, Hefei, China, in 2009. She is currently an Associate Professor in computer science at Anhui University. Her current research interests include computer vision, pattern recognition theory and application and digital image processing and video intelligent analysis.



Jixin Ma obtained his BSc and MSc of Mathematics in 1982 and 1988, respectively, and PhD of Computer Sciences in 1994. He is a Reader of Computer Science in the Department of Computing and Information Systems, and the Director of the Centre for Computer and Computational Science, as well as the Leader of Artificial Intelligence Research Group, at University of Greenwich, United Kingdom. His main research areas include Artificial Intelligence, Software Engineering and Information Systems, with special interests in Temporal Logic, Temporal Databases, Reasoning about Action and Change, Case-Based Reasoning, Pattern Recognition and Graph Matching.

Jin Tang received the B.Eng. degree in automation in 1999, and the Ph.D. degree in computer science in 2007 from Anhui University, Hefei, China. Since 2012, he has been a Professor in the School of Computer Science and Technology at Anhui University. His research interests include image processing, pattern recognition, machine learning and computer vision.