

Saliency Detection Based on Directional Patches Extraction and Principal Local Color Contrast

Muwei Jian^{1, 2*}, Wenyin Zhang², Hui Yu³, Chaoran Cui¹, Xiushan Nie¹, Huaxiang Zhang⁴, Yilong Yin^{5*}

¹School of Computer Science and Technology,

Shandong University of Finance and Economics, Jinan, China.

²School of Information Science and Engineering, Linyi University, Linyi, China.

³School of Creative Technologies, University of Portsmouth, Portsmouth, UK

⁴School of Information Science and Engineering, Shandong Normal University, Jinan, China.

⁵School of Software Engineering, Shandong University, Jinan 250101, China

* (Correspondence author) E-mail: jianmuweihk@163.com

Abstract

Saliency detection has become an active topic in both computer vision and multimedia fields. In this paper, we propose a novel computational model for saliency detection by integrating the holistic center-directional map with the principal local color contrast (PLCC) map. In the proposed framework, perceptual directional patches are firstly detected based on discrete wavelet frame transform (DWFT) and sparsity criterion, then the center of the spatial distribution of the extracted directional patches are utilized to locate the salient object in an image. Meanwhile, we proposed an efficient local color contrast method, called principal local color contrast (PLCC), to compute the color contrast between the salient object and the image background, which is sufficient to highlight and separate salient objects from complex background while dramatically reduce the computational cost. Finally, by incorporating the complementary visual cues of the global center-directional map with the PLCC map, a final compounded saliency map can be generated. Extensive experiments performed on three publicly available image databases, verify that the proposed scheme is able to achieve satisfactory results compared to other state-of-the-art saliency-detection algorithms.

Keywords: saliency detection, wavelet frame transform, principal local color contrast, directional patches

1. Introduction

Saliency detection focuses on separating the most salient objects/regions from background in images or videos [1], which is different from the conventional image-segmentation algorithms. Essentially, saliency detection tries to reveal and model the visual attention mechanism of Human Visual System (HVS) for perceiving salient objects or regions in a neurobiological way [2, 3]. Currently, a plenty of saliency detection models have been designed, which have been extended to various computer-vision and multimedia applications. Among them, typical applications include object detection [4, 5, 8], traffic

congestion analysis [7, 8], facial analysis [9, 10], underwater vision [36, 37], health prediction [51, 52, 53], visual prediction [56 - 61], activity recognition [62, 63], image retrieval [64 - 68] and so on.

Over the past few decades, although the research of saliency detection has attracted wide attentions of scientists from different disciplines such as neurobiology, physiology of vision, cognitive psychology, and computer vision, reliably modeling the mechanism of visual selective attention still remains a challenge and an open issue, especially for the images under complex scenes and without prior knowledge in advance.

Compared to these previous works, in this paper, we mainly focus on simulating how the human visual system (HVS) utilizes perception-directional visual stimuli to locate the most salient object in an image from the perspective of neurobiological perception. In our proposed framework, we first concentrate on designing a perceptual directional patch detector, based on discrete wavelet frame transform (DWFT) and sparsity criterion, to mimic how cortical cells that in the HVS utilize visual directional cues to detect salient objects in a computational way. Furthermore, the proposed perceptual directional patch detector can be applied universally, which is independent of image contents and no need of prior knowledge of image scenes. After locating the salient object, we also proposed an efficient local color contrast method, which is named principal local color contrast (PLCC) to calculate the color contrast between the salient object and the background rapidly. In contrast to traditional local color contrast methods, which require computing local contrast similarity over the entire image pixel by pixel leading to huge computational consumption. Our devised PLCC can not only soften and formulate principal local color contrast as a binary classification computation to avoid the NP-hard problem, but also has the ability to differentiate salient objects from complex background. Experiments are carried out on three publicly available image databases to evaluate the proposed model. And the comparison results have shown that the proposed method is effective and reliable for saliency detection.

The remainder of the paper is organized as follows. In Section 2, we will describe the related work about saliency detection. In Section 3, we will introduce the proposed saliency-detection model in detail. Experimental results are presented in Section 4. The paper closes with a conclusion and discussion in Section 5.

2. Related work

Generally, saliency detection models can be categorized into two distinctive types [11, 12]: either "the bottom-up" or "the top-down" case. The bottom-up visual attention modeling framework, which is based on low-level visual cues in an image including intensity, color, contrast, and orientation features, with the understanding that visual attention is in a biologically plausible way with unconsciously driven [2, 3, 4]. In contrast, the top-down computational mechanism mainly relies on the high-level comprehension of the image contents and requires prior knowledge of image scenes. Therefore, the top-down based saliency detection models are usually task-dependent or application-oriented, which demand large-scale training data and time-consuming human labeling images in a supervised learning manner. Since the training process is depended heavily on the training data, thus the generalization abilities of the top-down models normally lack of expansibility.

At present, most research works of saliency detection are focused on the bottom-up computational model [11, 12]. A pioneering exploration for bottom-up saliency detection was proposed by Itti *et al.* in 1998 [1]. In their computational model, an input image was decomposed into three separately feature maps of color, intensity, and orientation. Then, a center-surround contrast method was implemented to produce multi-scale image features, which were fused into a compounded saliency map to estimate each pixel's saliency value ultimately. In [13], a local contrast based image attention framework using a fuzzy growing algorithm is designed to produce saliency maps for salient object detection. Based on luminance information and spatiotemporal cues, Zhai and Shah [14] employed global contrast to implement visual attention detection in video sequences. In [15], Harel *et al.* proposed a Graph-Based Visual Saliency (GBVS) method for saliency detection, which was executed via normalizing and fusing the feature maps to highlight the salient regions in an image. Hou and Zhang proposed a spectral-residual method based on the log-spectrum analyzes of an image to generate a saliency map for saliency detection [16]. With the benefit of multi-scale contrast, center-surround histogram and color spatial distribution, a salient object detection method was designed to separate the conspicuous objects from the image background [17]. In [18], by formulating the bottom-up saliency detection as a center-surround process, a discriminant saliency model was devised to predict eye fixations in an image. Based on color contrast scheme by calculating each pixel's color difference between the average image color, a frequency-tuned approach was proposed for salient region detection in [19]. Rahtu *et al.* [20] presented a conditional random field (CRF) based segmentation model to separate salient objects from images and videos, which was based on local feature contrast in illumination, color, and motion clues. In [21], a non-parametric low-level saliency detection method was proposed. In their method, a spatial pooling mechanism and a scale-weighting function were employed to achieve saliency estimation. Goferman *et al.* introduced a context-aware saliency detection model based on local contrast scheme in [22], which applied local low-level features, visual organization rules, and high-level features to highlight salient objects. Based on selective-contrast strategy, Wang *et al.* [23] proposed a supervised Markov-random-field framework based on differential threshold for saliency detection. In [24], a saliency detection model based on multiple-instance learning by combining low-, mid-, and high-level features in the saliency estimation progress was effectively devised. By analyzing the nonlinear integration of different visual features and their intrinsic correlations, a saliency detection method by nonlinearly incorporating first-order statistics of features and region covariances was proposed in [25]. Yang *et al.* [26] proposed Graph-regularized saliency detection model, which integrated contrast, smoothness priors and convex-hull-based center prior to accomplish salient object detection. In [27], a salient object detection method based on multi-scale analysis on superpixels was introduced. In their method, visual cues from local contrast, integrity and center bias were optimized by a guided filter to produce a final saliency map. Zhu *et al.* [28] proposed a saliency optimization framework by fusing low level features via background detection with a principled optimization algorithm to generate saliency maps. In [29], Cheng *et al.* designed a histogram-based contrast (HC) framework, which was then incorporated into spatial information to obtain

region-based contrast (RC) saliency maps for detecting salient regions. Qin *et al.* [30] introduced a novel cellular automata based propagation mechanism for saliency detection, which incorporated color and space contrast in a Bayesian framework to produce a robust saliency map. In [31], superpixel-wise model based on homology-similarity was proposed, which took advantage of two saliency principles of spatial compactness and color contrast. Kong *et al.* [32] presented a framework to improve the performance of existing saliency detection models via an Extended Random Walk (ERW) selection mechanism, which could separate salient foreground from image background more effectively. With the aim for localizing salient objects, a relevance feedback algorithm was employed to estimate weights for integrating texture-insensitive and region-based saliency measures [33]. Jian *et al.* [34] proposed a saliency detection method by combining quaternionic distance based weber local descriptor and low-level priors. In [35], Li *et al.* design a novel learning-to-rank method for salient object detection, which elaborated saliency detection as a cost-sensitive label ranking problem by ranking saliency values in a descending order. Recently, underwater saliency detection becomes a hot issue. A large-scale underwater image database was constructed in [36]. Later, Jian *et al.* [37] designed an underwater saliency detection model by integrating Quaternionic Distance Based Weber Descriptor (QDWD) with pattern distinctness and Local Contrast, which incorporated quaternion number system and principal components analysis simultaneously.

In contrast to the bottom-up saliency detection mechanisms driven by low-level visual features, the top-down saliency model is dependent on the understanding of the scene context. Since the top-down strategy normally needs supervised learning and lack of generalization abilities, little work has been introduced. Oliva *et al.* [38] proposed a top-down saliency detection framework based on global scene configuration for object detection. In [39], a top-down computational mechanism with prior knowledge and region dividing strategy was devised for image segmentation. Cholakkal *et al.* [40] designed a top-down saliency detection approach via locality-constrained contextual sparse coding. In their approach, three locality constraints were combined to label an object category. In [41], a two-stage deep method was proposed to learn the intra-class association between the exemplars and query objects for top-down saliency detection. Recently, based on joint conditional random field and dictionary learning, Yang *et al.* [42] proposed a top-down saliency framework with a max-margin algorithm to train the detection model.

In this paper, we consider the more common category of the bottom-up computational saliency detection model. Then, based on perception-directional visual stimuli and sparsity criterion, we propose a novel saliency detection model based on directional patches extraction and principal local color contrast. The merits and advantages of the proposed saliency detection model are listed as follows:

- We design a perceptual directional patch detector based on DWFT and sparsity criterion, to mimic how cortical cells that utilize visual directional cues in HVS to detect salient objects.
- Compared with conventional local color contrast requiring calculating local contrast similarity over the entire image, the proposed principal local color contrast (PLCC) can not only distinguish salient objects from complex natural scenes, but also significantly reduce computational consumption.

- Finally, to further highlight the boundaries of the salient objects and adequately suppress the background noises, a compounded saliency map is incorporated by integrating the holistic center-directional cue with the PLCC feature simultaneously to achieve reliable saliency detection.

3. The proposed saliency-detection framework

In this section, we present a novel framework, which integrates the perceptual directional patches and principal local color contrast cue for saliency detection. Fig. 1 illustrates the designed framework for saliency detection based on directional patches extraction and principal local color contrast (PLCC). We will first introduce the proposed procedure for detecting and extracting perception-based directional patches via DWFT according to sparsity criterion, and then propose the principal local color contrast method. All different sources of visual cues are combined into a single saliency map for saliency detection.

3.1 Perception-based directional patch extraction

For human visual system (HVS), the perception of directionality/orientation cue has arisen with the birth of human beings. Consequently, directionality/orientation information in an image is a very important and subconscious stimulus for HVS to perceive and understand the image [5, 48, 49, 54]. Hence, directionality/orientation cue has been applied to many computer-vision related tasks, such as image fusion [43] and content-based image retrieval [44, 55].

In the following subsection, we will describe and design a DWFT based algorithm to extract the directional patches in an image. The proposed perception-based directional patch extraction scheme, which is independent of image contents and prior knowledge of image contents, is biologically plausible and aims to mimic how HVS perceives the directionally visual stimuli in the saliency detection process.

The discrete wavelet frame transform (DWFT), which is an improved variant of discrete wavelet transform (DWT), can present and reflect image directional variations at different scales in a multi-resolution analysis (MRA) manner [45, 46]. Compared to the DWT, DWFT has an obvious advantage and it is thus an overcomplete wavelet representation by omitting the down-sampling manipulations in DWT, as is shown in Fig. 1 (a). According to the definition of wavelet transform, a wavelet is an oscillating and attenuated function and its integrals equal to zero $\int_{-\infty}^{\infty} \psi(x) dx = 0$, a one-dimensional wavelet transform can be written as follows:

$$\varphi_{j,k}(x) = 2^{-j/2} \varphi(2^{-j}x - k); \text{ and} \quad (1)$$

$$\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k), \quad (2)$$

where: $\varphi(x)$ and $\psi(x)$ are the scaling function and wavelet function, respectively; $\{\varphi_{j,k}(x)\}$ and $\{\psi_{j,k}(x)\}$ are two orthogonal function basis sets.

The computation process of the DWT of a 2D signal involves recursive filtering and sub-sampling, as shown in Fig. 2 (a). Then the two-dimensional DWT of the image $f(x, y)$ is defined [45, 46]:

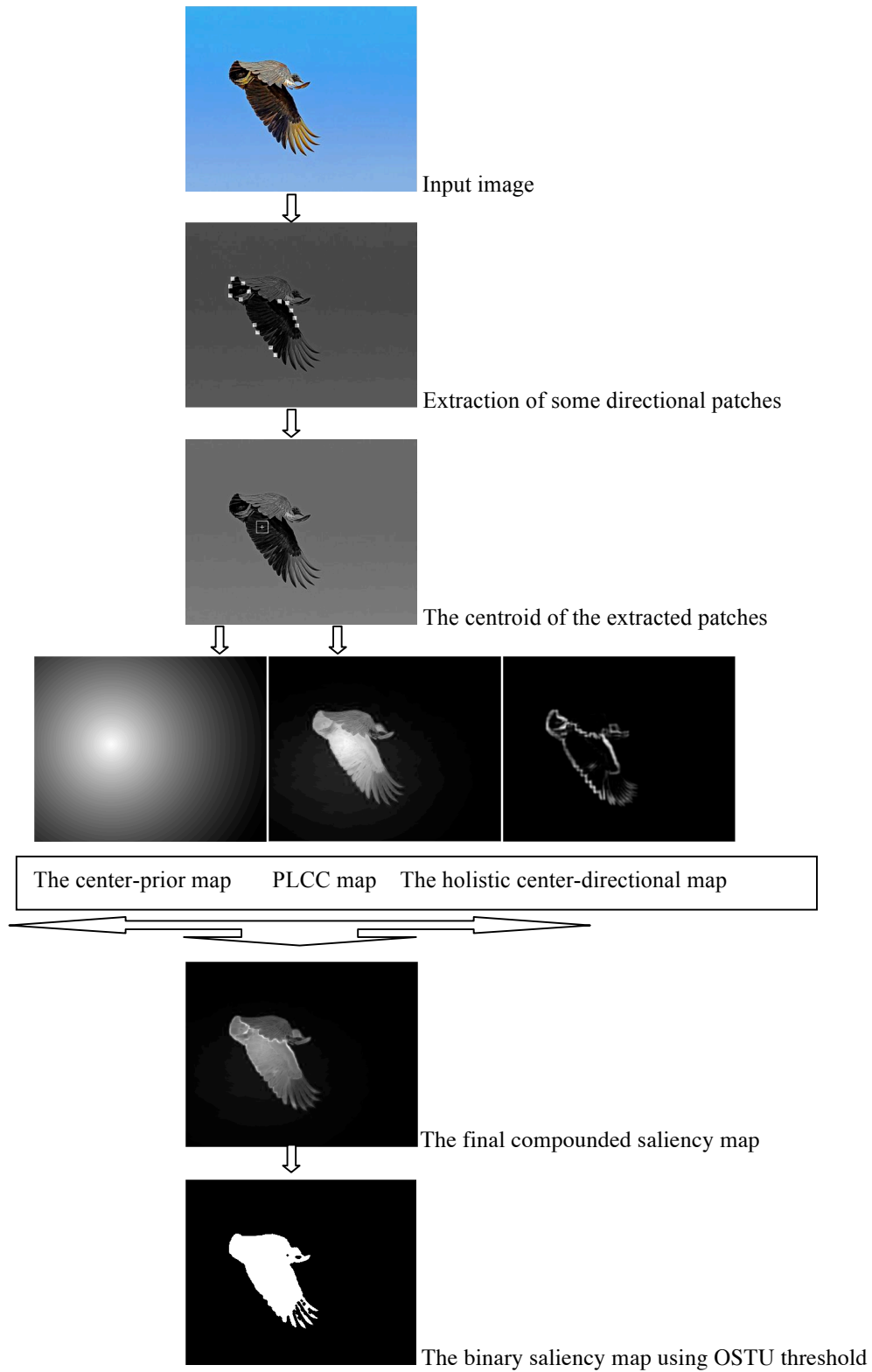


Fig. 1. Designed framework for saliency detection based on directional patches extraction and PLCC.

$$W_{\varphi}(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f \varphi_{j_0, m, n}; \quad (3)$$

$$W_{\psi^H}(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f \psi_{j_0, m, n}^H; \quad (4)$$

$$W_{\psi^V}(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f \psi_{j_0, m, n}^V; \text{ and} \quad (5)$$

$$W_{\psi^D}(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f \psi_{j_0, m, n}^D; \quad (6)$$

where $M \times N$ denotes the size of the input $f(x, y)$; j_0 is an arbitrary initial scale; $\varphi(x)$ is also the scaling function, and $\psi^i(x)$ (with $i = \{H, V, D\}$) are the wavelet functions; H denotes the horizontal high-frequency information, V the vertical high-frequency information, and D the diagonal high-frequency information, respectively. A two-level decomposition of DWT is as shown in Fig. 2 (a). We denote the detail sub-images (sub-band) as LH^{DWT} (contains the high frequency horizontal information), HL^{DWT} (contains the high frequency vertical information), and HH^{DWT} sub-band (contains the high frequency diagonal information), accordingly. The decomposition of DWT also produces one approximation sub-image, denoted by LL^{DWT} , containing the low frequency information of the input image. The DWT can decompose the LL^{DWT} band recursively. Therefore, two-level level wavelet decomposition yields 6 detail image. We use LH_1^{DWT} , HL_1^{DWT} , HH_1^{DWT} , LH_2^{DWT} , HL_2^{DWT} , HH_2^{DWT} , and an additional approximation image LL_2^{DWT} to denote all the sub-band images (See Fig. 2 (a)).

However, an obvious drawback using DWT is that it is not translation-invariant for image analysis. To handle and overcome this disadvantage, an over-complete and redundant wavelet representation, called the DWFT, with a dilated fact $\uparrow 2^j$ at decomposition level j is proposed by Unser in [47]. Similar to DWT, the DWFT with J decomposition levels yields a total of $3 \times J$ detail sub-bands and one approximation sub-band, while all of them are of the same size as the input image. The detail sub-bands of DWFT contain LH^{DWFT} (reflects the high frequency horizontal information), HL^{DWFT} (reflects the high frequency vertical information), and HH^{DWFT} sub-band (reflects the high frequency diagonal information), respectively. Fig. 2(b) shows a two-level decomposition of DWFT, which produces 6 detail images - LH_1^{DWFT} , HL_1^{DWFT} , HH_1^{DWFT} , LH_2^{DWFT} , HL_2^{DWFT} , HH_2^{DWFT} , and one low-pass filtered approximation image LL_2^{DWFT} of the input image. The main differences between the DWT and DWFT are illustrated in Fig. 2.

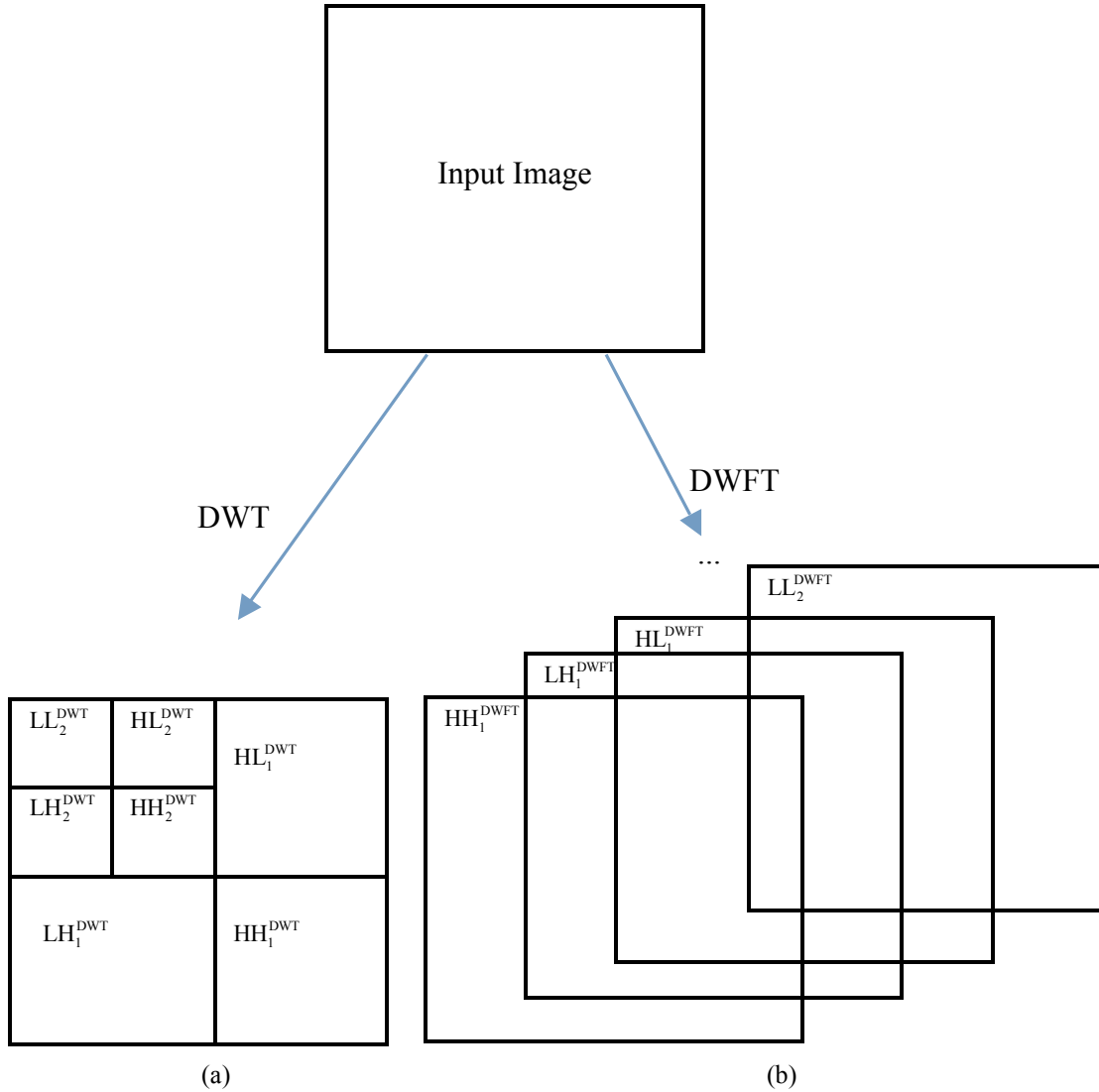


Fig. 2. Two-level decomposition of DWT and DWFT. (a) The decomposition progress of DWT; (b) The decomposition progress of DWFT.

Since the $\{ LH^{DWFT}, HL^{DWFT}, HH^{DWFT} \}$ sub-images of DWFT reflect three different direction high-frequency information, we propose a procedure to compute the vertical, horizontal, and diagonal directionality maps based on these three $\{ LH^{DWFT}, HL^{DWFT}, HH^{DWFT} \}$ sub-images, accordingly. Then, these directionality maps are finally normalized to fuse into a composite directionality/orientation map for locating the perceptual-meaningful directional patches. Concerning the horizontal directionality/ orientation map estimation, which can be calculated from the HL^{DWFT} sub-image (covering the vertical high-frequency information) of the DWFT. Suppose $LH^{DWFT}(x,y)$, $HL^{DWFT}(x,y)$ and $HH^{DWFT}(x,y)$ denote the sub-images' coefficients of the DWFT, where x and y denote the row coordinate and column coordinate of the sub-

images, accordingly. With regard to each row of the $HL^{DWFT}(x, y)$ sub-image, the horizontal directionality map is defined as follows:

$$Dir_H(x, y) = \frac{1}{2W} |HL^{DWFT}(x, y)| \left\{ \sum_{r=-W}^W |HL^{DWFT}(x+r, y)| \right\}, \quad (7)$$

where W represents the width of adjacent coefficients to be extracted, and in practice it can be set to p (where p is the wavelet regularity of an orthogonal wavelet with a compact support).

With respect to the calculation of the vertical directionality, the vertical directionality map is expressed along each column of the $LH^{DWFT}(x, y)$ sub-image in the following form:

$$Dir_V(x, y) = \frac{1}{2W} |LH^{DWFT}(x, y)| \left\{ \sum_{r=-W}^W |LH^{DWFT}(x, y+r)| \right\}. \quad (8)$$

Compared with the horizontal and vertical directionality maps, the calculation of diagonal directionality map is taken two perpendicularly diagonal directions, corresponding to $\pi/4$ and $3\pi/4$, into consideration together. Based on the $HH^{DWFT}(x, y)$ sub-image, the computation process of diagonal directionality map is defined as the average of the two perpendicularly diagonal directions of $\pi/4$ and $3\pi/4$ as follows:

$$Dir_D(x, y) = \frac{1}{2} (Dir_{D_1}(x, y) + Dir_{D_2}(x, y)), \quad (9)$$

where Dir_{D_1} and Dir_{D_2} denotes the diagonal directionalities of $\pi/4$ and $3\pi/4$, respectively.

Since all the vertical, horizontal and diagonal directionality maps are with the same size of the original image, these different directionality maps can be straightforwardly normalized and then merged together to construct a composite orientation/directionality map, as is shown in Fig. 4 (b):

$$Dir(x, y) = \frac{1}{3} (Dir_H(x, y) + Dir_V(x, y) + Dir_D(x, y)). \quad (10)$$

Moreover, in order to keep the sparsity of the extracted patches, while prevent the detected patches from being gathered together too concentrated, we utilize the following ‘sparsity criterion’ to ensure the detected patches scattered around the salient object. Once a directional patch was detected, the pixel values of its neighborhood patches are set as zeros. The diagram of sparsity criterion is defined as in Fig. 3.

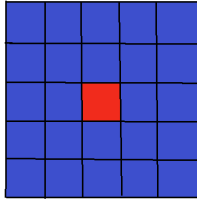


Fig. 3. Scattering criterion for the directional patches extraction. Red patch denotes the detected directional patch, and blue patches indicate its neighborhood with 2 patch-blocks.

Then, these patches with first K (e.g. $K = 15$) largest values can be extracted from the composite orientation/directionality map. Fig. 4 (c) shows some detected perception-based directional patches with the maximum values extracted from the composite directionality map for detecting and locating the salient

object. As illustrated in Fig. 4 (b), the proposed method can detect the directional patches in the image efficiently, which can be seen as a typical local feature to locate the salient object in an image for saliency detection.

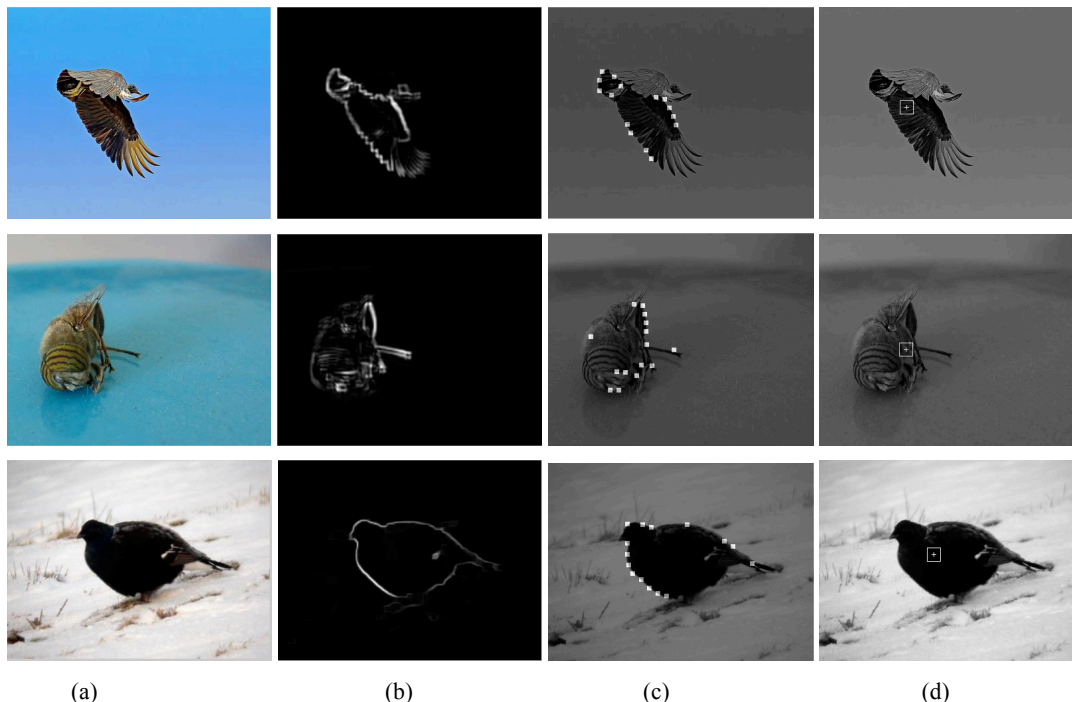


Fig. 4. Extraction of directional patches based on DWFT and sparsity criterion: (a) input images, (b) composite orientation/directionality maps, (c) extraction of some directional patches, (d) the centroids using the spatial distribution of the extracted patches

3.2 Determining the center using the spatial distribution of the extracted patches

Assume K directional patches have been extracted in an image, and each directional patch utilizes one center-point to represent the spatial distribution of the directional patch, accordingly. Then, we can determine the center based on the spatial distribution of the extracted patches easily. Let $Cen=(x', y')$ be the centroid of the detected directional patches (the white cross as is shown in Fig. 4 (d)), a square region Sq with the length Z as its sides can be defined to represent the local characteristics of the salient object, as is illustrated in Fig. 4 (d).

3.3 Principal Local Color Contrast (PLCC)

In most of the existing saliency-detection modeling approaches based on the bottom-up computational strategy, local color contrast is widely used as a low-level visual feature in images for salient object detection, which has achieved good results in highlighting the foreground objects. Compared to previous methods, in this paper, we propose an efficient and effective local color contrast method called principal local color contrast (PLCC) for saliency detection. The main idea is that we utilize the square region Sq inside the salient object to estimate the principal local color contrast between the salient object and the

image background. Instead of calculating local contrast or blindly comparing similarity over each pixel in the entire image one by one resulting in vast computational cost, the designed PLCC method is not only sufficient to highlight semantic objects from complicated scenarios, but also can significantly reduce the computing complexity during distinguishing salient objects from the image background.

With the aid of the square region Sq determined in the previous step, for an input image we first set plc as the principal local color vector by taking the mean value of the pixels inner the square region Sq in the image color space, and Cen is the spatial position center of the square region Sq . Let $pos(x, y)$ denote the spatial position with pixel coordinates (x, y) in the input image, and $pc(x, y)$ is used to represent its color vector of the pixel in an input image. Then, the computational scheme of the principal local color contrast (PLCC) can be defined as follows:

$$PLCC(x, y) = \frac{1}{\|pc(x, y) - plc\| \exp\left(\frac{-\|pos(x, y) - Cen\|}{2\sigma^2}\right)}, \quad (11)$$

where σ^2 controls the strength of spatial weighting.

Fig. 5 shows some typical principal local color contrast (PLCC) maps of the input images. Among PLCC computational procedure, Fig. 5 (b) illustrates the center-prior map $Cenpri(x, y)$, generated by the second part of the equation (11) using the corresponding centroid of (as is shown in Fig. 4 (d)) spatial distribution of the extracted directional patches; while Fig. 5 (c) shows the final PLCC maps of the original images.

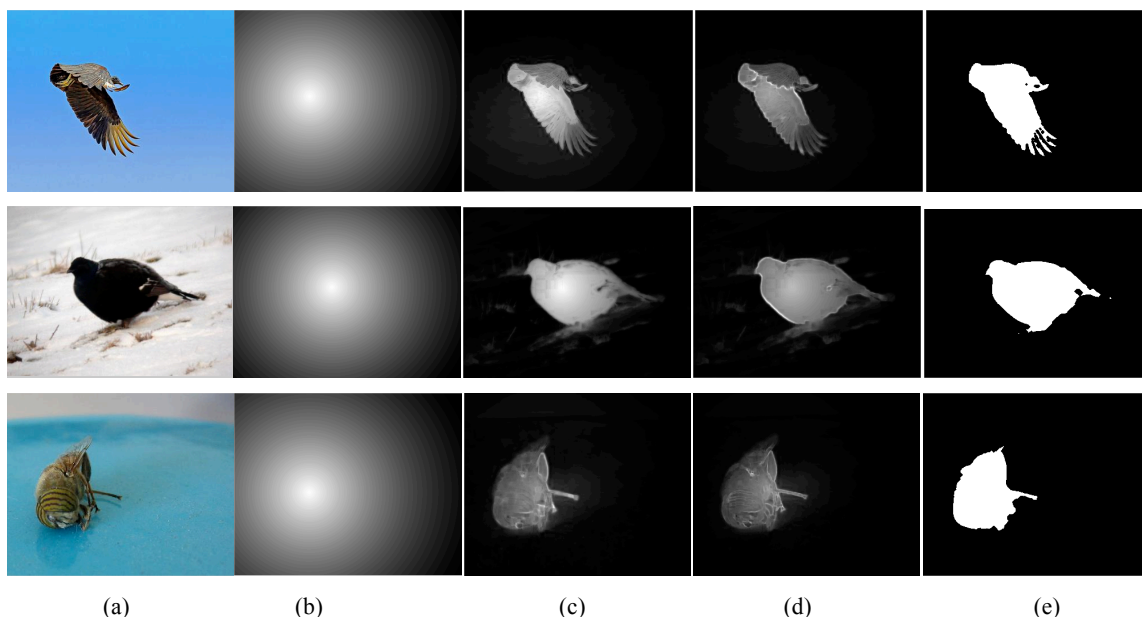


Fig. 5. Computational progress of the principal local color contrast (PLCC): (a) input images, (b) the center-prior map $Cenpri(x, y)$ using the spatial distribution of the extracted directional patches, (c) PLCC maps, (d) the final compounded saliency maps by integrating the holistic center-directional map with the PLCC map, (e) the binary saliency map using classical OSTU threshold method.

3.4 The final compounded saliency map

The aim of the final stage is to incorporate these two saliency cues as an overall saliency map, by combining the merits of both sides, complement each other and make the saliency map more robust [50]. Besides, with an attempt to further highlight the boundaries of the salient objects and adequately suppress the image background, a compounded saliency map is fused by integrating the holistic directional map $Dir(x, y)$ with the center-prior map $Cenpri(x, y)$ and the PLCC map $PLCC(x, y)$ simultaneously, to generate the final compounded saliency map as follows:

$$Salmap(x, y) = Dir(x, y) \cdot Cenpri(x, y) + PLCC(x, y) , \quad (12)$$

where \cdot is the pixel-by-pixel multiplication operator. The purpose of the multiplication processing is that we utilize the center-prior map $Cenpri(x, y)$ to preserve the edges of salient objects while simultaneous stifle and removal of the background noises in the border of images. Then after integration of the complementary visual cues, the final compounded saliency map $Salmap(x, y)$ enhances the contrast as well as retaining the edges of salient objects in order to facilitate saliency-detection tasks. Fig. 5 (d) displays the final compounded saliency maps produced by the proposed method. And Fig. 5 (e) shows the binary saliency map using the classic OSTU threshold method. It can be noted that the proposed method can preserve the boundaries of the salient objects completely while can highlight the whole salient objects uniformly and smoothly.

4. Experimental results

In order to evaluate the performance of the proposed model, there publicly available and widely used datasets, including MSRA10K dataset [17, 28, 29], ECSSD database [50], and DUT-OMRON image database [26, 27, 30] were tested and evaluated for various saliency-detection models. In the experiments, the number of directional patches with the first largest values K is set at 15, which can be employed to estimate the center using their spatial distribution. A square region Sq with the length $Z = 5$ is adapted, and σ^2 is set at 8 for controlling the strength of spatial weighting. These parameters are empirically set and experiments show that using the above settings can produce a satisfactory and robust overall performance.

For quantitatively comparing the state-of-the-art saliency-detection approaches, the average precision, recall, and F -measure are utilized. Precision measures the proportion of the detected salient-object rectangle, inside the ground-truth rectangle, over the detected salient-object rectangle, while recall is defined as the ratio of the detected salient-object rectangle, inside the ground-truth rectangle, over the ground-truth rectangle:

$$\text{Precision} = \frac{\text{the detected salient object rectangle inside the ground truth rectangle}}{\text{the detected salient object rectangle}} , \text{ and} \quad (12)$$

$$\text{Recall} = \frac{\text{the detected salient object rectangle inside the ground truth rectangle}}{\text{the ground truth rectangle}} . \quad (13)$$

The F -measure is an overall evaluation to weight precision more than recall, and is defined as follows:

$$F_{\beta} = \frac{(1 + \beta) \times \text{Precision} \times \text{Recall}}{\beta \times \text{Precision} + \text{Recall}}, \quad (14)$$

where β is a real positive value and is set at $\beta = 0.3$ [4, 14, 15, 16].

In our extensive experiments, we also compare the proposed algorithm with other state-of-the-art saliency-detection frameworks, containing nine representative bottom-up methods with different computational strategies, namely the Itti [1], Graph-Based (GB) [15], Context-aware (CA) [22]; Non-Parametric (NP) [21], Nonlinearly Covariance (NC) [25], Graph-regularized (GR) [26], Multi-Scale Superpixel (MSS) [27]; Robust Background Detection (RBD) [28] and Background-Single Cellular Automata (BSCA) [30]. The following sub-sections will give the detail of the experimental results performed on three individual databases.

4.1 The MSRA10K database

We firstly evaluated the proposed algorithm on the popular MSRA10K database, which contains 10,000 images with their corresponding pixel-wise ground truth for saliency detection assessment.

Fig. 6 shows some typical saliency detection results of different methods performed on the MSRA10K database. It can be noted that the proposed algorithm generates saliency maps more uniformly with contrast and little background pixels than other state-of-the-art methods. In addition, the merit of our proposed method is that the boundaries of the salient objects are well-preserved. Through visual comparing, the proposed strategy has a superior ability to highlight the salient objects and retain the edges of objects efficiently, as illustrated in Fig. 6.

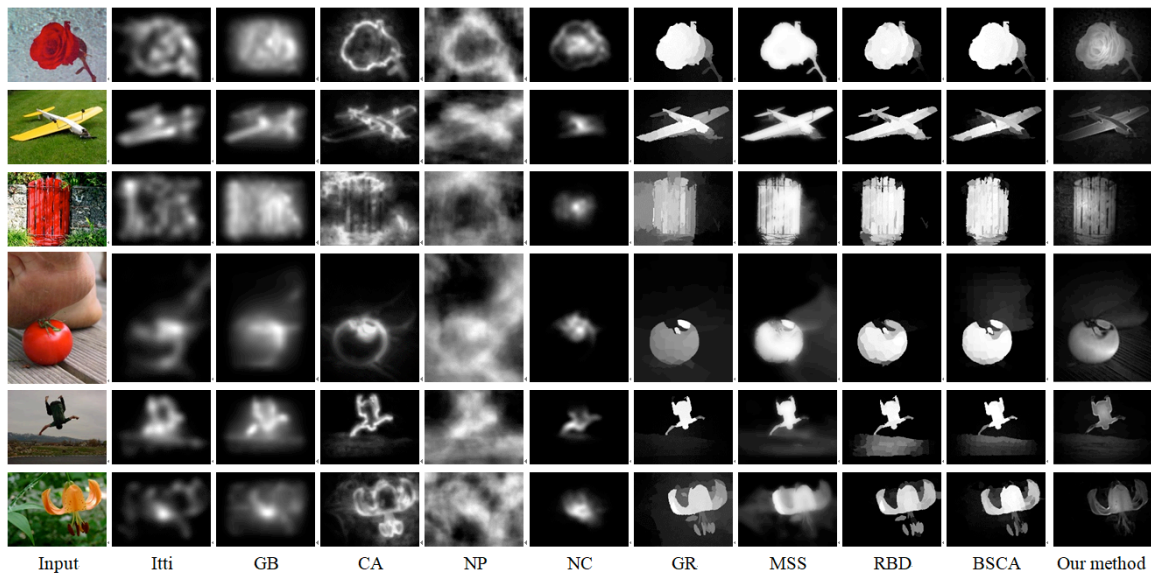


Fig. 6. Some saliency-detection results of different state-of-the-art saliency-detection models on the MSRA10K dataset.

The quantitative comparisons (with precision, recall and the F -measure values) between the proposed method and other nine state-of-the-art saliency detection algorithms are firstly assessed on the MSRA10K database. Fig. 7 shows the precision, recall and the F -measure criteria of all of the ten performed models. From the comparisons, we can find that most of the methods produces saliency-detection results with precision higher than 70%, namely CA [22], NC [25], MSS [27]; RBD [28], BSCA [30] and our proposed algorithm. For the overall criterion F -measure, all of the saliency-detection models, except the early method in Itti [1], are higher than 50%, and our proposed model achieves better performance over other nine existing state-of-the-art saliency-detection methods in terms of F -measure. Experiment results on the MSRA10K dataset show that our proposed scheme is efficient and precise for general images.

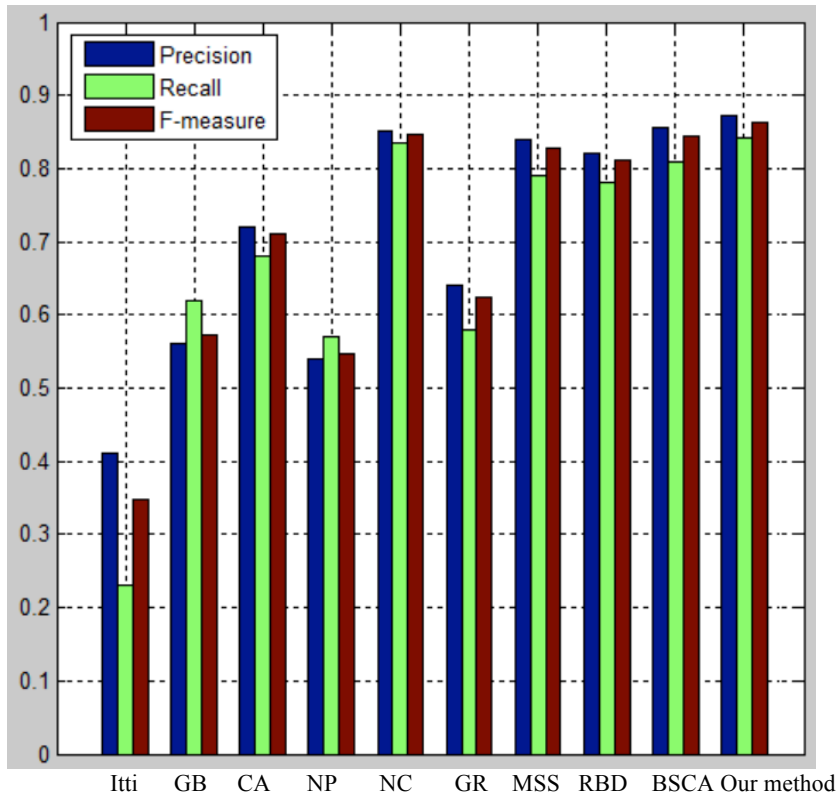


Fig. 7. Performance of the proposed model compared with other state-of-the-art methods on the MSRA10K dataset, in terms of average precision, recall, and F -measure.

4.2 The ECSSD database

We have also tested our proposed method against other state-of-the-art saliency detection models on the widely used and more challenging ECSSD database, which includes many semantically meaningful but structurally complex natural-scene images, for performance comparisons and evaluations. The ECSSD database contains 1,000 images in total. These images in the ECSSD database are collected from the internet and five helpers are asked to label the homologous ground truth masks.

Fig. 8 shows some representative results of saliency maps generated by different saliency detection algorithms including our proposed method. From Fig. 8, we can note that the proposed method can accurately detect almost entire salient objects and produce better object outlines and edges than other methods. The main reason can be attributed to the contribution of holistic center-directional maps based on DWFT, which is designed to efficiently represent the salient object’s global shape cues in an input image.

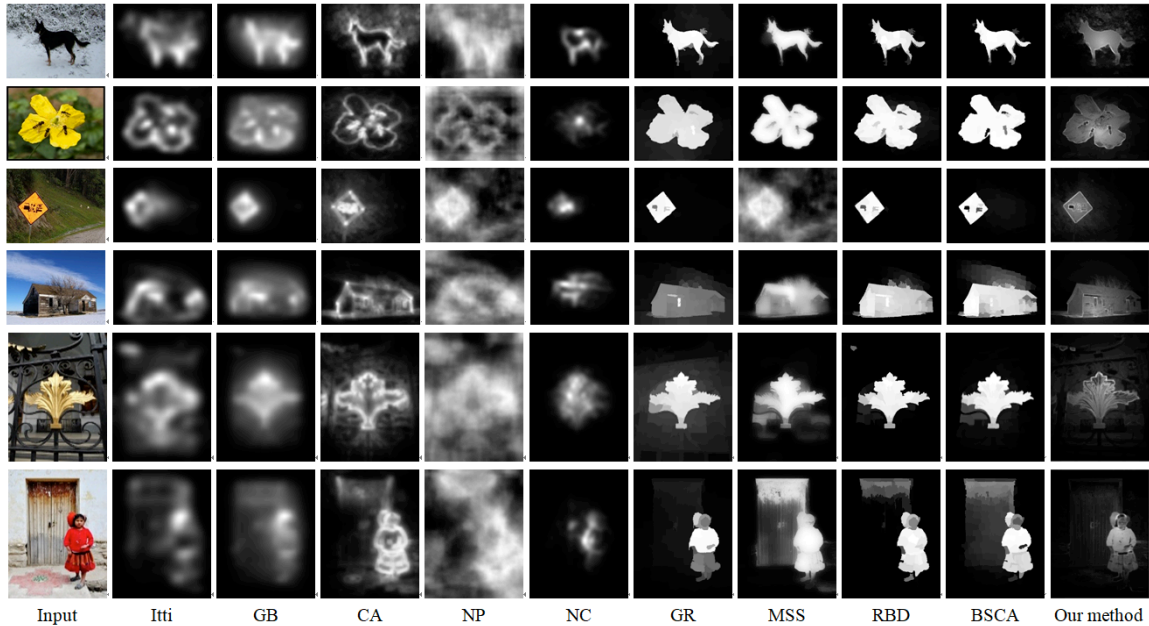


Fig. 8. Some representative results of saliency maps generated by different saliency detection algorithms preformed on the ESSCN dataset.

We also report the performance of the proposed method and other nine state-of-the-art saliency detection methods. Quantitatively comparisons are performed and assessed according to the measurements of precision, recall and the F -measure . Fig. 9 shows the comparisons of different methods under the three different evaluation criterions. From the comparisons as shown in Fig. 9, it can be seen that our proposed method outperforms the other ten different state-of-the-art methods on the ECSSD database, in terms of the overall F -measure values. These experimental results, carried on the ECSSD dataset, also verify that our proposed model is able to produce reliable and promising saliency-detection outcomes.

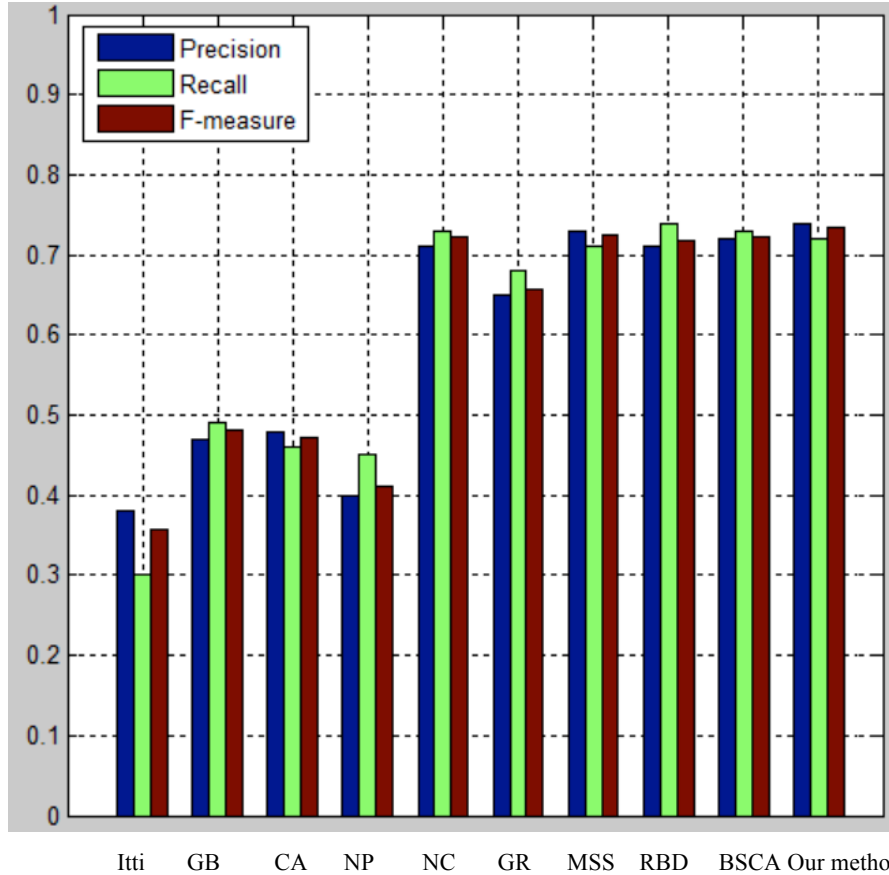


Fig. 9. Comparisons of different saliency-detection methods in terms of average precision, recall, and F -measure on the ECSSD dataset.

4.3 The DUT-OMRON database

Moreover, most current state-of-the-art models are sensitive to images with strong textures and complicated backgrounds. Those species of images are tough to dispose and are a still challenging issue for saliency detection. Thus, we further evaluate our proposed algorithm on the publicly available DUT-OMRON database, which has 5,168 challenging images selected from more than 140,000 images. The images in DUT-OMRON database are difficult but carefully collected, since these images in this dataset contain complicated background; besides, most of the images is with low color-contrast between the salient objects and the image background. For instance, there are some images with animal protective coloration that enable them to blend with their surroundings and backgrounds.

Fig. 10 illustrates some typical saliency detection results of various algorithms performed on the DUT-OMRON dataset. As shown in Fig. 10, we can observe that our proposed model can highlight the salient objects more smoothly while contain less background regions than other state-of-the-art algorithms. Compared with other existing methods, the saliency maps generated by our model are more precise and it is also capable of suppressing the background regions effectively.

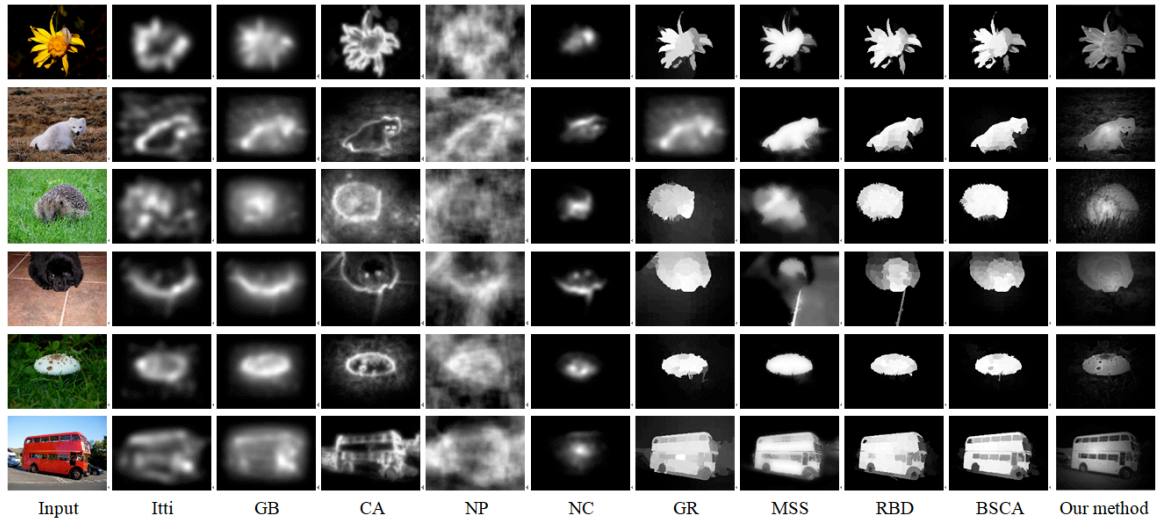


Fig. 10 Some typical saliency detection results of various algorithms performed on the DUT-OMRON dataset.

For objective evaluation, Fig. 11 shows quantitative comparison results on DUT-OMRON database according to precision, recall and F -measure. From the F -measure scores in Fig. 11, it can be observed that the proposed model gives the more accurate saliency-detection results compared with other existing algorithms. Extensive experiments, implemented on the DUT-OMRON dataset with images of rough textures, low color-contrast and complex background, also prove that our designed framework is capable of generating satisfactory saliency-detection results.

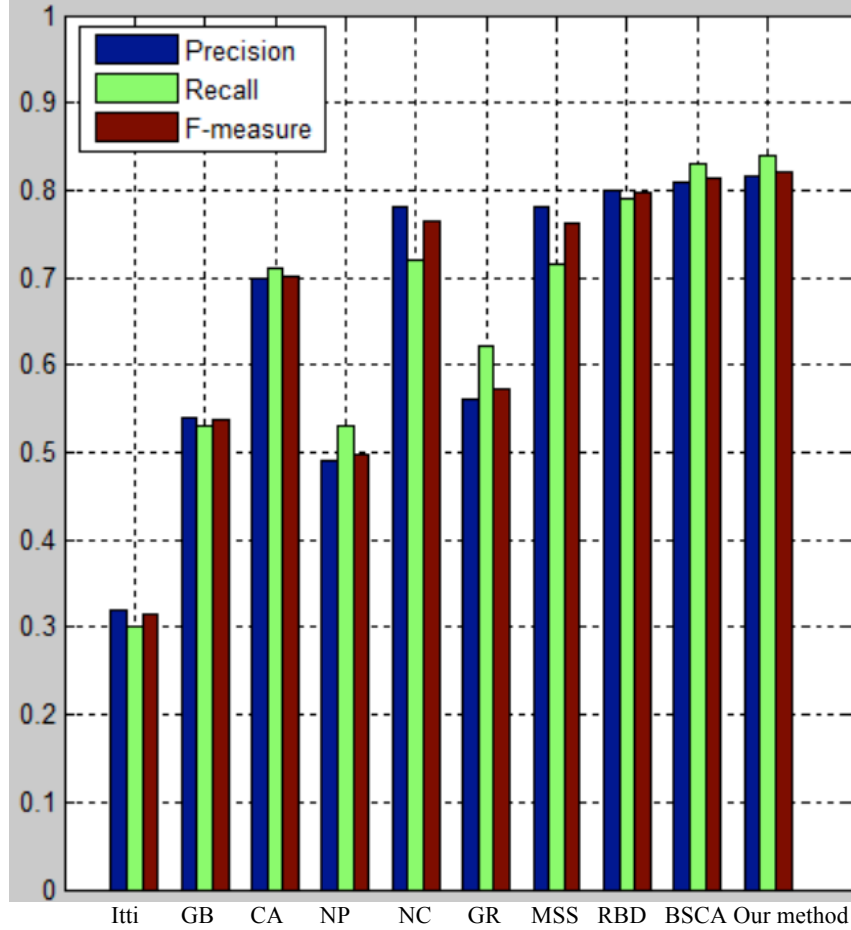


Fig. 11. Performance of the proposed model compared with other state-of-the-art methods on the DUT-OMRON database, in terms of average precision, recall, and F -measure .

5. Conclusion and discussion

In this paper, a novel bottom-up method is proposed for efficient and reliable saliency detection. The designed framework effectively takes both the holistic center-directional cue and the principal local color contrast (PLCC) cue into consideration simultaneously. Experiments show that the proposed framework can highlight and distinguish salient objects from complicated background while significantly diminishes the complexity of the computational consumption. Furthermore, extensive experiments on three widely used databases verify that our proposed method produces satisfactory saliency detection results compared with other state-of-the-art methods.

Acknowledgments

This work was supported by National Natural Science Foundation of China (NSFC) (61601427, 61602229, 61771230, 61573219); Natural Science Foundation of Shandong Province (ZR2016FM40); Shandong Provincial Key Research and Development Program of China (2017CXGC0701, 2017CXGC1504); Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions.

References

- [1] 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):1254-1259, 1998.
- [2] C. Koch and S. Ullman, "Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry," *Human Neurobiology*, vol. 4, pp. 219–227, 1985.
- [3] B. A. Olshausen, C.H. Anderson, and D.C. Van Essen, "A Neurobiological Model of Visual Attention and Invariant Pattern Recognition Based on Dynamic Routing of Information," *J. Neuroscience*, vol. 13, no. 11, pp. 4,700–4,719, Nov. 1993
- [4] M. Jian, K. M. Lam, J. Dong, et al., "Visual-patch-attention-aware Saliency Detection," *IEEE Transactions on Cybernetics*, vol. 45, no. 8, pp. 1575-1586, 2015.
- [5] J. Zhang, Y.Zhao, S. Chen, Object-level saliency: Fusing objectness estimation and saliency detection into a uniform framework, *Journal of Visual Communication and Image Representation*, Vol. 53, 2018, pp. 102-112.
- [6] X. Yan, Y. Wang, Q. Song, K. Dai, Salient object detection via boosting object-level distinctiveness and saliency refinement, *Journal of Visual Communication and Image Representation*, Vol. 48, 2017, pp. 224-237.
- [7] Q. Wang, J. Wan, Y. Yuan, Locality constraint distance metric learning for traffic congestion detection, *Pattern Recognition*, vol. 75, pp. 272-281, 2018.
- [8] Q. Wang, J. Wan, Y. Yuan, "Deep metric learning for crowdedness regression", *IEEE Trans. Circuits System and Video Technology*, DOI: 10.1109/TCSVT.2017.2703920.
- [9] M. Jian, K.M. Lam, J. Dong, "Facial-Feature Detection and Localization Based on a Hierarchical Scheme", *Information Sciences*, Vol. 262, pp. 1-14, 2014.
- [10] M. Jian, K.M. Lam, "Face-Image Retrieval Based on Singular Values and Potential-Field Representation", *Signal Processing*, Vol. 100, pp. 9-15, 2014.
- [11] A. Toet, "Computational versus psychophysical bottom-up image saliency: A comparative evaluation study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 11, pp. 2131–2146, Nov. 2011.
- [12] A. Borji, D. N. Sihite, and L. Itti, "Quantitative analysis of human-model agreement in visual saliency modeling: A comparative study," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 55–69, Jan. 2013.
- [13] Y.F. Ma, H.J. Zhang: Contrast-based image attention analysis by using fuzzy growing. *ACM International Conference on Multimedia*, November 2003, pp. 374–381, 2003.
- [14] Y. Zhai and M. Shah. Visual attention detection in video sequences using spatiotemporal cues. *ACM International Conference on Multimedia*, pages 815–824, 2006.
- [15] J. Harel, C. Koch, P. Perona, "Graph-based visual saliency," *Advances in Neural Information Processing Systems*, pp. 545-552, 2006.
- [16] X. Hou & L. Zhang. Saliency Detection: A spectral residual approach. *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1-8, 2007.
- [17] T. Liu, J. Sun, N. Zheng, X. Tang, and H. Shum. Learning to detect a salient object. *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1-8, 2007.
- [18] D. Gao, V. Mahadevan, N. Vasconcelos, "The discriminant center-surround hypothesis for bottom-up saliency," *Advances in Neural Information Processing Systems*, pp. 497-504, 2007.
- [19] R. Achanta, S. Hemami, F. Estrada, & S. S. Sussstrunk. Frequency-tuned salient region detection. *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1597-1604, 2009.

- [20] E. Rahtu, J. Kannala, M. Salo, and J. Heikkilä, "Segmenting salient objects from images and videos," European Conference on Computer Vision, pp. 366–379, 2010.
- [21] N. Murray, M. Vanrell, X. Otazu, et al., "Saliency estimation using a non-parametric low-level vision model," IEEE Conference on Computer Vision and Pattern Recognition, pp. 433-440, 2011.
- [22] S. Goferman, L. Zelnik-Manor, A. Tal., Context-aware saliency detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012, 34(10): 1915-1926.
- [23] Q. Wang, Y. Yuan, P. Yan, X. Li, "Visual saliency by selective contrast", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 7, pp. 1150-1155, Jul. 2013.
- [24] Q. Wang, Y. Yuan, P. Yan, X. Li, "Saliency detection by multiple-instance learning", *IEEE Trans. Cybern.*, vol. 43, no. 2, pp. 660-672, Apr. 2013.
- [25] E. Erdem, A. Erdem, "Visual saliency estimation by nonlinearly integrating features using region covariances," *Journal of vision*, vol. 13, no. 4, pp. 11-11, 2013.
- [26] C. Yang, L. Zhang, H. Lu, "Graph-regularized saliency detection with convex-hull-based center prior," *IEEE Signal Proces. Let.*, vol. 20, no. 7, pp. 637-640, 2013.
- [27] N. Tong, H. Lu, L. Zhang, et al., "Saliency detection with multi-scale superpixels," *IEEE Signal Processing Letters*, vol. 21, no. 9, pp. 1035-1039, 2014.
- [28] W. Zhu, S. Liang, Y. Wei, and J. Sun, "Saliency optimization from robust background detection," *IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2014, pp. 2814–2821.
- [29] M. M. Cheng, N. J. Mitra, X. Huang, et al., "Global contrast based salient region detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 569–582, 2015.
- [30] Y. Qin, H. Lu, Y. Xu, et al., Saliency detection via cellular automata, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.110-119, 2015.
- [31] Z. Chen, H. Wang, L. Zhang, et al., "Visual saliency detection based on homology similarity and an experimental evaluation," *Journal of Visual Communication and Image Representation*, vol. 40, pp. 251-264, 2016.
- [32] Y. Kong, L. Wang, X. Liu, H. Lu, X. Ruan, Pattern Mining Saliency, *European Conference on Computer Vision*, vol. 9910, pp. 583-598, 2016.
- [33] A. Hati, S. Chaudhuri, R. Velmurugan, "An Image Texture Insensitive Method for Saliency Detection," *Journal of Visual Communication and Image Representation*, vol. 43, pp. 212–226, 2017.
- [34] M. Jian, Q. Qi, J. Dong, et al., "Saliency Detection Using Quaternionic Distance Based Weber Local Descriptor and Level Priors," *Multimedia Tools and Applications*, pp. 1-18, 2017.
- [35] Z. Li, C. Lang, S. Feng, T. Wang, Saliency ranker: A new salient object detection method, *Journal of Visual Communication and Image Representation*, Vol. 50, pp.16-26, 2018.
- [36] M. Jian, Q. Qi, J. Dong, Y. Yin, W. Zhang, K. M. Lam, "The OUC-Vision Large-Scale Underwater Image Database". 2017 *IEEE International Conference on Multimedia and Expo (ICME)*, Hong Kong, 2017, pp. 1297-1302.
- [37] M. Jian, Q. Qi, J. Dong, et al., Integrating QDWD with Pattern Distinctness and Local Contrast for Underwater Saliency Detection, *Journal of Visual Communication and Image Representation*, Vol. 53, pp. 31–41, 2018.
- [38] A. Oliva, A. Torralba, M. S. Castelhano, et al, "Top-down control of visual attention in object detection," *IEEE ICIP*, vol. 1, 2003.
- [39] Y.T. Wu, F.Y. Shih, J. Shi, Y.T. Wu, A Top-Down Region dividing approach for image segmentation, *Pattern Recognition*, 41(6), 1948–1960, 2008.

- [40] H. Cholakkal, D. Rajan, J. Johnson, "Top-down saliency with locality-constrained contextual sparse coding," BMVC, 2015.
- [41] S. He, R. W. H. Lau, Q. Yang, "Exemplar-Driven Top-Down Saliency Detection via Deep Association," International Journal of Computer Vision, vol. 115, no. 3, pp. 330-344, 2015.
- [42] J. Yang, M. H. Yang, Top-down visual saliency via joint crf and dictionary learning, IEEE Trans. Pattern Anal. Mach. Intell., 2017, 39, (3), pp. 576-588.
- [43] S. Li, J. T. Kwok, & Y. Wang, Using the discrete wavelet frame transform to merge landsat tm and spot panchromatic images. Information Fusion, 3(1), pp. 17-23, 2002.
- [44] M. Jian, J. Dong, J. Ma. "Image retrieval using wavelet-based salient regions", Imaging Science Journal, The. Vol. 59, No. 4, pp. 219-231, 2011.
- [45] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell., Vol. 11, pp. 674-693, July 1989.
- [46] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," IEEE Trans. on Information Theory, Vol.36, pp.961-1005, 1990.
- [47] M. Unser, Texture classification and segmentation using wavelet frames, IEEE Transactions on Image Processing, 4 (11), pp. 1549–1560, 1995.
- [48] Mario Villena-González, Vladimir López, Eugenio Rodríguez, Orienting attention to visual or verbal/auditory imagery differentially impairs the processing of visual stimuli, NeuroImage, vol. 132, pp. 71–78, 2016.
- [49] F .Daneshvarfard, N. Maarefi, H. A. Moghaddam, F. Wallois, A survey on stimuli for visual cortical function assessment in infants, Brain & Development, vol. 40, pp. 2–15, 2018.
- [50] T. Xi, W. Zhao, H. Wang, W. Lin, Salient object detection with spatiotemporal background priors for video, IEEE Trans. Image Process., Vol. 26, No. 7, pp. 3425-3436, 2017.
- [51] L. Nie, L. Zhang, Y. Yang, M. Wang, R. Hong, T. S. Chua. Beyond Doctors: Future Health Prediction from Multimedia and Multimodal Observations. ACM International Conference on Multimedia, pp. 591-600, 2015.
- [52] L. Nie, M. Wang, L. Zhang, S. Yan, Bo Zhang, T. S. Chua. Disease Inference from Health-Related Questions via Sparse Deep Learning. IEEE Transactions on Knowledge & Data Engineering, 27 (8), pp.2107-2119, 2015.
- [53] L. Nie, Y. Zhao, M. Akbari, J. Shen, T. S. Chua. Bridging the Vocabulary Gap between Health Seekers and Healthcare Knowledge. IEEE Transactions on Knowledge & Data Engineering, 27(2), pp. 396-409, 2014.
- [54] L. Nie, Y. Zhao, X. Wang, J. Shen, T. S. Chua. Learning to Recommend Descriptive Tags for Questions in Social Forums. ACM Transactions on Information Systems, 32 (1), pp. pp.1-23, 2014.
- [55] M. Jian, Y. Yin, J. Dong, K. M. Lam, Content-based image retrieval via a hierarchical-local-feature extraction scheme, Multimedia Tools and Applications, 2018. DOI : 10.1007/s11042-018-6122-2
- [56] P. Jing, Y. Su, L. Nie, H. Gu, J. Liu, M. Wang, A framework of joint low-rank and sparse regression for image memorability prediction. IEEE Trans. on Circuits and Systems for Video Technology, 2018, DOI: 10.1109/TCSVT.2018.2832095.
- [57] P. Jing, Y. Su, L. Nie, X. Bai, J. Liu, M. Wang, Low-rank multi-view embedding learning for micro-video popularity prediction. IEEE Transactions on Knowledge and Data Engineering, 30 (8), pp. 1519 – 1532, 2018.
- [58] P. Jing, Y. Su, L. Nie, H. Gu, Predicting Image Memorability Through Adaptive Transfer Learning from External Sources, IEEE Transactions on Multimedia, 19 (5), pp. 1050-1062, 2017.
- [59] J. Zhang, X. Li, P. Jing, J. Liu, Y. Su, Low-rank regularized heterogeneous tensor decomposition for subspace clustering. IEEE Signal Processing Letters, 25 (3), pp. 333-337, 2018.

- [60] A. Liu, Y. Shi, P. Jing, J. Liu, Y. Su, Structured Low-rank Inverse-covariance Estimation for Visual Sentiment Distribution Prediction, *Signal Processing*, Vol. 152, pp. 206-216, 2018.
- [61] J. Chen, X. Song, L. Nie, X. Wang, H. Zhang, T. S. Chua, Micro Tells Macro: Predicting the Popularity of Micro-Videos via a Transductive Model, *ACM International Conference on Multimedia*, pp. 898-907, 2016.
- [62] Y. Liu, L. Nie, L. Liu, D. S. Rosenblum, From action to activity: Sensor-based activity recognition, *Neurocomputing*, 181(12) pp.108-115, 2016.
- [63] Y. Liu, L. Nie, L. Han, L. Zhang, D. S. Rosenblum, Action2Activity: recognizing complex activities from sensor data, *International Conference on Artificial Intelligence (IJCAI'15)*, pp. 1617-1623, 2015.
- [64] L. Zhu, Z. Huang, Z. Li, L. Xie, H. Shen, Exploring auxiliary context: discrete semantic transfer hashing for scalable image retrieval, *IEEE Transactions on Neural Networks and Learning Systems*, 2018. DOI: 10.1109/TNNLS.2018.2797248.
- [65] J. Li, K. Lu, Z. Huang, L. Zhu, H. Shen, Transfer Independently Together: A Generalized Framework for Domain Adaptation, *IEEE Transactions on Cybernetics*, 2018.
- [66] L. Zhu, Z. Huang, X. Chang, J. Song, H. Shen, Exploring Consistent Preferences: Discrete Hashing with Pair-Exemplar for Scalable Landmark Search, *ACM International Conference on Multimedia*, pp. 726-734, 2017.
- [67] L. Zhu, Z. Huang, X. Liu, X. He, J. Sun, X. Zhou, Discrete multimodal hashing with canonical views for robust mobile landmark search, *IEEE Transactions on Multimedia*, 19 (9), pp. 2066-2079, 2017.
- [68] L. Xie, J. Shen, J. Han, L. Zhu, L. Shao, Dynamic multi-view hashing for online image retrieval, *International Joint Conference on Artificial Intelligence (IJCAI'17)*, pp.3133-3139, 2017.



Muwei Jian received the PhD degree from the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, in October 2014. He was a Lecturer with the Department of Computer Science and Technology, Ocean University of China, from 2015 to 2017. Currently, Dr. Jian is a Professor and Ph.D Supervisor at the School of Computer Science and Technology, Shandong University of Finance and Economics.

His current research interests include human face recognition, image and video processing, machine learning and computer vision. Prof. Jian was actively involved in professional activities. He has been a member of the Program Committee and Special Session Chair of several international conferences, such as *SNPD 2007*, *ICIS 2008*, *APSIPA 2015*, *EEECs 2016*, *ICTAI2016*, *ICGIP 2016* and *ICTAI 2017*. Dr. Jian has also served as a reviewer for several international SCI-indexed journals, including *IEEE Trans.*, *Pattern Recognition*, *Information Sciences*, *Computers in Industry*, *Machine Vision and Applications*, *Machine Learning and Cybernetics*, *The Imaging Science Journal*, and *Multimedia Tools and Applications*.

Prof. Jian holds 3 granted national patents and has published over 40 papers in refereed international leading journals/conferences such as *IEEE Trans. on Cybernetics*, *IEEE Trans. on Circuits and Systems for Video Technology*, *Pattern Recognition*, *Information Sciences*, *Signal Processing*, *ISCAS*, *ICME* and *ICIP*.



Wenyin Zhang is currently a professor at the School of Information Science and Engineering, Linyi University, Linyi, China. His research interests include machine learning, bioinformatics, and image processing.



Hui Yu is currently a professor in the School of Creative Technologies. Dr. Yu is the leader of Visual Computing Group at the University of Portsmouth. He previously held an appointment with the University of Glasgow. He has won prizes for his study and research including Excellent Undergraduate Prize (provincial level), the Best PhD Thesis Prize, EPSRC DHPA Awards (PhD) and Vice Chancellor Travel Prize. Dr. Yu is an *Associate Editor of IEEE Transactions on Human-Machine Systems*. He is a member of the Peer Review College of the Engineering and Physical Sciences Research Council (EPSRC), UK.



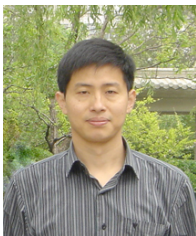
Chaoran Cui received the B.S. degree in Software Engineering in 2010 and the Ph.D. degree in Computer Science and Technology in 2015, both from Shandong University, Jinan, China. During 2015-2016, he was a research fellow at Singapore Management University. He is now a Professor in School of Computer Science and Technology at Shandong University of Finance and Economics. His research interests include information retrieval, analysis and understanding on multimedia information, and computer vision.



Xiushan Nie received the Ph.D. degree from Shandong University, Shandong, China, in 2011. He is currently a Professor with Shandong University of Finance and Economics, Shandong, China. From 2013 to 2014, he was a Visiting Scholar with the University of Missouri-Columbia, Columbia, MO, USA. His research interests include multimedia retrieval and indexing, multimedia security, and computer vision.



Huaxiang Zhang is currently a professor with the School of Information Science and Engineering & the Institute of Data Science and Technology, Shandong Normal University, China. He received his Ph.D. from Shanghai Jiaotong University in 2004, and worked as an associated professor with the Department of Computer Science, Shandong Normal University from 2004 to 2005. He has authored over 180 journal and conference papers and holds 13 granted national patents. His current research interests include machine learning, pattern recognition, evolutionary computation, cross-media retrieval, web information processing, big data analysis, etc.



Yilong Yin received the Ph.D. degree from Jilin University, Changchun, China, in 2000. From 2000 to 2002, he was a Post-Doctoral Fellow with the Department of Electronics Science and Engineering, Nanjing University, Nanjing, China.

He is currently the Director of the data Mining, Machine Learning, and their Applications Group and a Professor of the School of Software Engineering, Shandong University, Jinan, China. His research interests include machine learning, data mining, and computational medicine.