

Saliency detection using quaternionic distance based weber local descriptor and level priors

Muwei Jian^{1,2} · Qiang Qi² · Junyu Dong² · Xin Sun² ·
Yujuan Sun³ · Kin-Man Lam⁴

Received: 19 November 2016 / Revised: 12 May 2017 / Accepted: 11 July 2017 /
Published online: 22 July 2017
© Springer Science+Business Media, LLC 2017

Abstract In this paper, a novel and efficient framework by exploiting Quaternionic Distance Based Weber Local Descriptor (QDWLD) and object cues is proposed for image saliency detection. In contrast to the existing saliency detection models, the advantage of the proposed approach is that it can combine quaternion number system and object cues simultaneously, which is independent of image contents and scenes. Firstly, QDWLD, which was initially designed for detecting outliers in color images, is used to represent the directional cues in an image. Meanwhile, two low-level priors, namely the Convex-Hull-Based center and color contrast cue of the image, are utilized and fused as an object-level cue. Finally, by combining QDWLD with object cues, a reliable saliency map of the image can be computed and estimated. Experimental results, based on two widely used and openly available database, show that the proposed method is able to produce reliable and promising salient maps/estimations, compared to other state-of-the-art saliency-detection models.

Keywords Weber descriptor · Saliency detection · Object cues · Directional cues

✉ Muwei Jian
jianmuweihk@163.com

¹ School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China

² Department of Computer Science and Technology, Ocean University of China, Qingdao 266100, China

³ College of Information and Electrical Engineering, Ludong University, 186 Shixue Road, Yantai 264025, China

⁴ Centre for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong

1 Introduction

Visual saliency detection is the process of identifying the most conspicuous locations of objects in images and has become a hot and active research area in recent years. Serving as a pre-processing step, it can efficiently focus on the interesting image regions or objects related to the current task and broadly applied to various computer vision tasks, such as image classification/retrieval, image retargeting, video coding [35, 36], vision tracking, video processing [37], visual motion estimation [38], etc.

Human can accurately identify the salient regions or objects in image scenes. Simulating such an ability of the human vision attention system in computer vision is critical for many real-world multimedia applications. The key issue to simulate human visual system (HVS) is the distinctness of images, which is standing out from their neighbors and may cause human visual stimuli [14]. Recently, many research papers have been published on simulating such intelligent behavior in human visual system (HVS). However, effectively and accurately modeling the process of salient-object detection is a tough task and still remains a challenge right now, especially when no prior knowledge is available, or without any context for images.

In this paper, a novel saliency-detection framework based on a bottom-up mechanism, integrating Quaternionic Distance Based Weber Local Descriptor (QDWLD), center prior and color cues, is proposed for saliency detection. This proposed model includes quaternion number system and object cues, which aims to simulate how humans process the visual stimuli in the detection saliency process. In the proposed method, we first computed the QDWLD, which was initially designed for detecting outliers in color images, to represent the directional cues in an image. Secondly, we incorporated the Convex-Hull-Based center method to measure the location of the salient objects and applied color contrast to estimate the color saliency. Finally, these three different maps are combined to represent the most important stimuli for saliency detection. In order to evaluate the performance of our proposed method, we carried out extensive experiments on two publicly available and widely used datasets, and the comparison results with other state-of-the-art saliency-detection algorithms show that the proposed approach is effective and efficient for saliency detection.

A preliminary conference version appeared in [15]. In contrast to the preliminary work in [15], we adopt a more precise center cue estimation algorithm, called Convex-Hull-Based center estimation, to compute the spatial position of the salient object more accurately than the preliminary work [15]. In addition, we have analyzed the proposed model and the individual contributions of the Quaternionic Distance Based Weber Local Descriptor (QDWLD), center prior, and color prior, which show that the new proposed saliency model can generate more accurate results than the preliminary model. At last, we also performed extensive experiments on another widely used database, comparing with the other eight state-of-the-art saliency-detection methods, to show our method is efficient and effective.

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

2 Related work

Generally, there are two approaches to modeling the saliency-detection mechanism: the top-down and the bottom-up algorithms. Currently, most of the existing visual detection modeling methods [1,

5, 7, 10, 12, 13, 22, 24, 41] are based on the bottom-up computational scheme, which uses low-level visual features in images, such as intensity, color, contrast, and orientations, to achieve saliency detection. A pioneer work of saliency detection was proposed by Itti et al. [13]. In the model, a final saliency map, to represent the saliency of each pixel in an image, is produced by using three individual feature maps: the color, intensity, and orientation maps. In [24], an alternative local contrast analysis approach was proposed for calculating saliency maps based on a fuzzy growth algorithm for saliency-attention detection. Koch and Perona [10] proposed a model, namely Graph-Based Visual Saliency (GBVS), which aims to fuse different feature maps so as to strengthen those remarkable regions in an image. In [41], a global contrast method, using the luminance channel, was proposed to estimate the salient regions in an image. Liu et al. [22] designed a saliency detection task as an image segmentation problem by segmenting salient objects from input images. In this method, the multi-scale contrast, the center-surround histogram, and the color spatial distribution are simultaneously considered for detecting a salient object. Hou and Zhang proposed a spectral-residual algorithm by analyzing the log-spectrum of an image to estimate the corresponding saliency map [12]. In [1], a frequency-tuned model for saliency detection was proposed by calculating the color difference from the average color in an image. Goferman et al. [7] proposed a method based on local contrast and visual organization rules to detect salient objects. Chen et al. [5] proposed a histogram-based contrast (HC) and a region-based contrast (RC) method for saliency detection. In [14], a visual-patch, attention-aware saliency-detection model was proposed. Later, Zhong et al. proposed a generalized nonlocal mean framework with object-level cues for saliency detection in [43]. A spatio-temporal saliency detection method based on the amplitude or phase spectrum of the Fourier transform of an image in frequency domain was proposed in [8]. In [23], a Gaussian-based saliency detection approach was proposed for salient object segmentation. In [25], an unsupervised model was proposed to estimate the pixel-wise saliency in an image. In [45], a saliency detection method based on linear neighbourhood propagation was proposed. Li et al. designed a salient object detection model by combing meanshift filtering with colour information in an image [20]. Recently, Song et al. [29] proposed a RGBD co-saliency method via bagging-based clustering to detect the salient objects in an image. Zhang and Sclaroff designed a boolean-map based method for saliency detection in [42]. In [31], a novel context features auto-encoding algorithm based on regression tree was proposed to handle numerical features and select effective features. In [30], a multistage saliency detection framework, which is based on multilayer cellular automata (MCA), was proposed to detect the saliency of images. A novel saliency-detection model based on probabilistic object boundaries was proposed in [19]. Li et al. [21] presented an effective deep neural network framework embedded with low-level features (LCNN) for salient object detection. In [44], Zhong et al. proposed a novel video saliency detection model based on perceptual orientation to detect the attended regions in video sequences.

In contrast with the bottom-up visual-attention model, which is driven by low-level visual features, the top-down method is based on the high-level cues of images and usually requires prior knowledge and context-aware understanding [27]. In general, the top-down model is task dependent or application oriented, therefore it is difficult to refine the attention targets in an image. Furthermore, the top-down mechanism usually needs supervised learning and is lacking in extendibility and scalability, so not much work has been proposed, based on the top-down model. In [24], a top-down saliency method by using the global scene configuration was proposed for saliency detection. The works in [11], a novel saliency-detection model was proposed to learn the intra-class association between the exemplars and query objects. Rahman et al. [28] proposed a top-down contextual weighting (TDCoW) saliency model, which incorporates high-level knowledge of the gist context of images, to appropriate weights to the features. In [39], a novel top-down

saliency model, which learns a Conditional Random Field (CRF) and a visual dictionary was proposed for saliency detection. The shortcoming of top-down methods is lack of expandability and scalability, and also takes long time to train the corresponding parameters in the models.

According to the attention mechanism of human visual systems (HVS), directional information of an image is an important cue for saliency detection. Relative to the previously preliminary work [15], this paper introduces a more precise center cue to produce better saliency estimation. In this paper, an efficient saliency-detection method is designed by utilizing both visual-directional stimuli and two low-level-object cues. The proposed model, which is independent of image contents and scenes, is biologically plausible, and the directional cues can be utilized as neuronal visual stimuli for HVS to detect salient objects. Thus, the main contributions of our proposed saliency detection model are listed as follows:

- We utilized the Quaternionic Distance Based Weber Local Descriptor (QDWLD), which was initially designed for detecting outliers in color image, to represent the directional cues in an image for saliency detection. To the best of our knowledge, it is the first time to incorporate the QDWLD into saliency detection issue.
- We applied a simple but powerful algorithm to estimate the location of the salient objects, which is more precise than the traditional center prior estimation method.
- We further employed the color contrast to define the color saliency of an image, which plays an important role in related saliency detection tasks.

3 The proposed method

In this section, we present a novel framework, which incorporates the Quaternionic Distance Based Weber Local Descriptor (QDWLD) with object cues, to simulate saliency detection. QDWLD will first be described, followed by the object cues based on the center prior and color prior. All these different types of information are fused to form a saliency map.

3.1 Quaternionic distance based weber local descriptor (QDWLD)

Neurobiological studies have indicated that directional information of an image, such as edges and outliers, plays an important role in HVS for saliency detection [14]. In our proposed framework, QDWLD, which was initially proposed for detecting the outliers and edges in an image [18], is utilized to represent the directional cues for the HVS to detect saliency. QDWLD has also been applied to texture classification and face recognition [18]. In this section, we will briefly introduce QDWLD, which is utilized to represent the directional cues in our method.

According to quaternion algebra, a quaternion \dot{q} is made up of one real part and three imaginary parts as follows:

$$\dot{q} = a + ib + jc + kd, \quad (1)$$

where $a, b, c, d \in \mathcal{R}$; i, j, k are complex operators; a is the real part; and $\{ib, jc, kd\}$ are the imaginary parts. In the polar form with $S(\dot{q}) = a$ and $V(\dot{q}) = ib + jc + kd$, the quaternion \dot{q} can be rewritten as:

$$\dot{q} = |\dot{q}|e^{i\theta} = |\dot{q}|(\cos\theta + usin\theta), \quad (2)$$

where $\dot{u} = V(\dot{q})/|V(\dot{q})|$ and $\theta = \tan^{-1}(|V(\dot{q})|/|S(\dot{q})|)$ are the eigenaxis and phase of \dot{q} , respectively. In [18], $V(\dot{q})$ is used to encode a RGB image as follows:

$$\dot{Q}(x, y) = iR(x, y) + jG(x, y) + kB(x, y), \tag{3}$$

where $\dot{Q}(x, y)$ is the Quaternionic Representation (QR) of the pixel at location (x, y) , and $R(x, y)$, $G(x, y)$, and $B(x, y)$ are the red, green, and blue components in a color image, respectively.

By considering two pixels in a color image, $\dot{q}_1 = r_1i + g_1j + b_1k$ and $\dot{q}_2 = r_2i + g_2j + b_2k$, different types of distances can be defined to measure the distance of quaternions. Let $D_t(\dot{q}_1, \dot{q}_2)$ represent the t^{th} type of QR of quaternions \dot{q}_1 and \dot{q}_2 . Since the modulus of a quaternion \dot{q} is nonnegative, the quaternionic distance (QD), denoted as $D_1(\dot{q}_1, \dot{q}_2)$, can be defined as follows:

$$D_1(\dot{q}_1, \dot{q}_2) = |\dot{q}_1 - \dot{q}_2| \tag{4}$$

Let $\dot{\rho} = e^{\dot{\alpha}u}$. By defining the intensity and the chromaticity components of \dot{q}_1 as $(\dot{\rho}\dot{q}_1\dot{\rho}^* + \dot{\rho}^*\dot{q}_1\dot{\rho})/2$ and $(\dot{\rho}\dot{q}_1\dot{\rho}^* - \dot{\rho}^*\dot{q}_1\dot{\rho})/2$, respectively, Quaternionic distances $D_2(\dot{q}_1, \dot{q}_2)$ and $D_3(\dot{q}_1, \dot{q}_2)$ can be expressed as follows [3]:

$$D_2(\dot{q}_1, \dot{q}_2) = \frac{1}{2} |(\dot{\rho}\dot{q}_1\dot{\rho}^* + \dot{\rho}^*\dot{q}_1\dot{\rho}) - (\dot{\rho}\dot{q}_2\dot{\rho}^* + \dot{\rho}^*\dot{q}_2\dot{\rho})|, \tag{5}$$

$$D_3(\dot{q}_1, \dot{q}_2) = \frac{1}{2} |(\dot{\rho}\dot{q}_1\dot{\rho}^* - \dot{\rho}^*\dot{q}_1\dot{\rho}) - (\dot{\rho}\dot{q}_2\dot{\rho}^* - \dot{\rho}^*\dot{q}_2\dot{\rho})| \tag{6}$$

In consideration of the differences between the intensity and chromaticity components simultaneously, the quaternionic distance $D_4(\dot{q}_1, \dot{q}_2)$ is defined as follows [6]:

$$D_4(\dot{q}_1, \dot{q}_2) = D_2(\dot{q}_1, \dot{q}_2) + D_3(\dot{q}_1, \dot{q}_2). \tag{7}$$

Let $\dot{u} = (i + j + k) / \sqrt{3}$, which represents an axis in RGB space with values $R = G = B$. In [16], by rotating one quaternion towards the gray line \dot{u} , a quaternion \dot{q}_3 should be close to the gray line. If \dot{q}_1 and \dot{q}_2 are closed to each other, $\dot{q}_3 = \dot{q}_2 = u\dot{q}_1u^* = r_3i + g_3j + b_3k$. Then, the quaternionic distance $D_5(\dot{q}_1, \dot{q}_2)$ can be expressed as follows:

$$D_5(\dot{q}_1, \dot{q}_2) = |(r_3 - \mu)i + (g_3 - \mu)j + (b_3 - \mu)k|, \tag{8}$$

where $\mu = r_3 + g_3 + b_3/3$.

By adding the luminance component into the quaternionic distance $D_5(\dot{q}_1, \dot{q}_2)$, a new quaternionic distance $D_6(\dot{q}_1, \dot{q}_2)$ is proposed in [17], as follows:

$$D_6(\dot{q}_1, \dot{q}_2) = \lambda D_5(\dot{q}_1, \dot{q}_2) + (1 - \lambda) |L(\dot{q}_1, \dot{q}_2)| \tag{9}$$

where $L(\dot{q}_1, \dot{q}_2) = k_1(r_2 - r_1) + k_2(g_2 - g_1) + k_3(b_2 - b_1)$ is the luminance difference, λ is the weight to balance the difference between the chromaticity and luminance component, and k_1 , k_2 and k_3 are the weights of the different color channels to luminance.

Extensive analysis and intensive discussions of the above quaternionic distances are not provided because it is beyond the scope of this paper. It should be noted that Lan et al. [18] gave a proof that the quaternionic distances $D_3(\dot{q}_1, \dot{q}_2)$ and $D_5(\dot{q}_1, \dot{q}_2)$, which were defined from different viewpoints, have an equal relationship, i.e. $D_3(\dot{q}_1, \dot{q}_2) = D_5(\dot{q}_1, \dot{q}_2)$.

In order to ensure the derived features are robust and discriminative, we take the Weber's law (WL) into consideration. The essence of Weber's law, which assumed that the ratio between the smallest perceptual change in a stimulus (ΔI_{\min}) and the background level of the stimulus I is a constant, is proposed by the German physiologist Ernst Weber. The Weber's law can be defined as follows:

$$\frac{\Delta I_{\min}}{I} = u \quad (10)$$

where u is called Weber fraction and is a constant.

Inspiring by the Weber's law, the above-described quaternionic distances, which can be defined as the increments between two quaternions, can also be used to measure the similarity between different quaternions from different perspective viewpoints. That is to say, the quaternionic increment between quaternions can be computed by using quaternionic distances (QDs). By integrating the quaternionic distances (QDs) and Weber's law, the Quaternionic Distance Based Weber Local Descriptor (QDWLD) can be derived for image feature extraction. Assume that \dot{q}_c denotes the center quaternion in a local patch, and \dot{q}_l ($l \in L$, where L is the index set) represents the residual quaternions in the patch. Let $\xi_2^t(\dot{q}_c)$ denote the differential feature of QDWLD defined by D_t . Then, the total quaternionic increment in a local patch can be written as $\sum_{l=0}^{l=7} D_t(\dot{q}_c, \dot{q}_l)$. With the aid of $|\dot{q}_c|$ as the quaternionic intensity, $\xi_2^t(\dot{q}_c)$ can be represented as follows:

$$\xi_2^t(\dot{q}_c) = \arctan\left(\frac{\sum_{l \in L} D_t(\dot{q}_c, \dot{q}_l)}{|\dot{q}_c|}\right). \quad (11)$$

The nonlinear mapping, $\arctan(\cdot)$, aims to enhance $\xi_2^t(\dot{q}_c)$ to become more robust [9, 18]. Actually, Eq. (10) is the Euclidean distance between the center quaternion \dot{q}_c and the original point in the color space. With the differential features $\xi_2^t(\dot{q}_c)$ of QDWLD produced by using differential D_t ($t = 1, 2, 3, 4, 6$), the increments in all directions can be contained.

Figure 1a shows an example of the QDWLD features for an input color image. Figure 1b-f are the differential feature maps $\xi_2^t(\dot{q}_c)$ of QDWLD obtained by using D_t ($t = 1, 2, 3, 4, 6$). The image generated by the 5th QD is not given since it is equivalent to the 3rd QD. As shown in Fig. 1b-f, these differential features can be utilized to reflect the directional information of an image. Then, we normalize these directional maps and add them to form an integrated holistic directional map Dir . It should be noted that we consider the five different quaternionic distances (QDs) have the equal weights and the final integrated directional map is the mean of these different feature maps of QDWLD obtained by using different D_t ($t = 1, 2, 3, 4, 6$). Figure 1g shows the fusion of the different directionality maps to form an integrated directional map, which can be utilized for saliency detection.

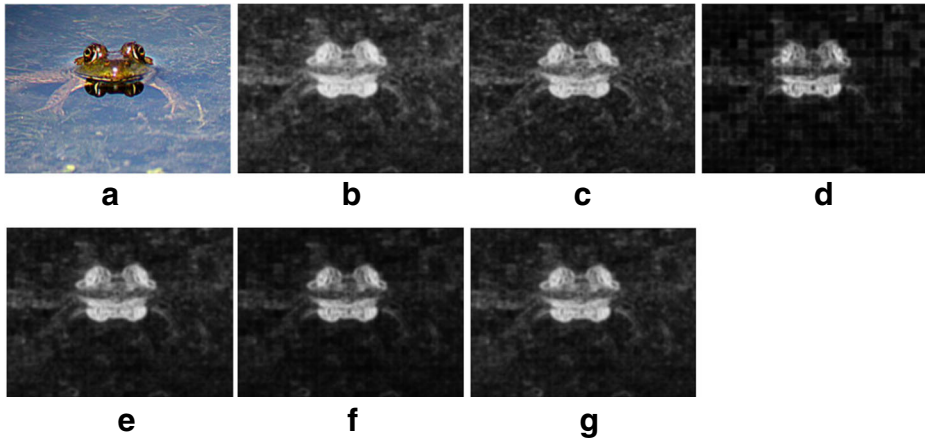


Fig. 1 An example of the QDWLD features. **a** An input color image. **(b–f)** The differential feature maps of QDWLD produced by different QDs. **g** The corresponding integrated directional map

3.2 Object cues – center prior

Spatial position is also a useful cue for saliency detection. Some pristine center priors proposed in [2, 7, 22, 24] are generally based on a Gaussian distribution from the center of an image. These typical center cues can strengthen the saliency-detection power in human-captured photographs. However, in most circumstances, a salient object may be far away from the image center, and it will then be affected or disabled. Considered the real situation for spatial position, an improved method by choosing the N top most bounding boxes detected in an image [4, 43], which can estimate the object spatial position slight more precisely than the traditional center prior, was proposed in [43] and also employed in [15]. However, extensive experimental results show that although the improved method [15, 43] can produce object spatial position more efficiently than the traditional center prior, it still far from estimating the accurate object spatial position, as shown in Fig. 2b.

To alleviate this problem, we introduce and employ the Convex-Hull-Based center prior instead of the previous one [15, 43]. The convex hull is proposed in [32, 33, 40] in order to improve the inference of Bayesian saliency model, which achieves a robust and promising performance. The main advantage is that it uses the convex hull of interest points to estimate the center of the salient object rather than directly use the image center which makes the saliency result more robust to the location of objects. In this paper, we first compute a convex hull which encloses interesting points to orientate the location of salient region. Then, we regard the centroid of the convex hull as the center to achieve the convex-hull-based center prior.

The saliency of a pixel can be defined as follows:

$$Cen_i = \exp\left(-\frac{\|x_i - x_0\|}{2\psi_x^2} - \frac{\|y_i - y_0\|}{2\psi_y^2}\right), \quad (12)$$

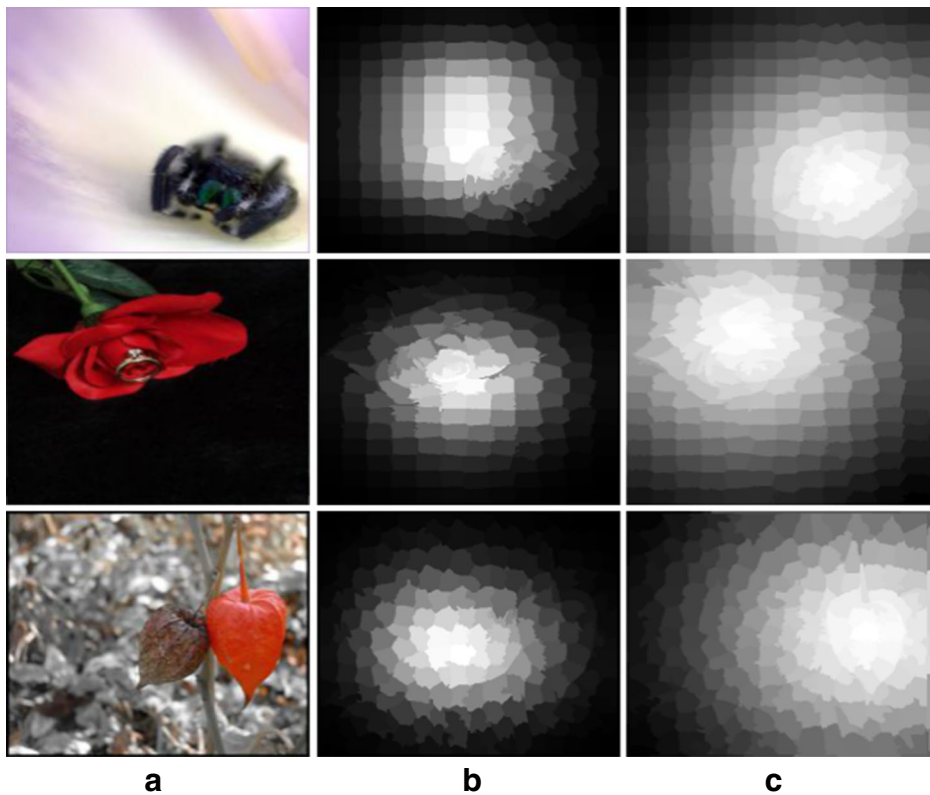


Fig. 2 Comparison of center priorestimation using different methods. **a** An input color image. **b** The method in [15, 43]. **c** The Convex-Hull-Based center prior

where (x_0, y_0) is the center of the pixel i , x_i and y_i are the mean horizontal and vertical positions of the pixel i respectively, ψ_x and ψ_y denote the horizontal and vertical variances. We use a centered anisotropic Gaussian distribution to model the center prior. So, we set $\psi_x = \psi_y$ with pixel coordinates normalized to $[0, 1]$ in our implementation. The convex-hull-based center prior map is more reasonable and robust owing to the convex hull provides a rough location of the salient object. With this reliable center prior, the object's spatial position can be estimated more accurately than those traditional center priors, as shown in Fig. 2c.

3.3 Object cues – color prior

Color is another important cue for saliency detection [5, 14, 43]. For each patch v_i , its feature vector $c = c_i, 1 \leq i \leq n$, is produced, based on its mean color in a color space [43]. A global color contrast with its spatial contrast can be expressed as follows:

$$Cor_i = \sum_{j=1}^n \phi(v_j) \varphi(v_i, v_j) \left| |c_i - c_j| \right| i = 1, 2, \dots, n, \quad (13)$$

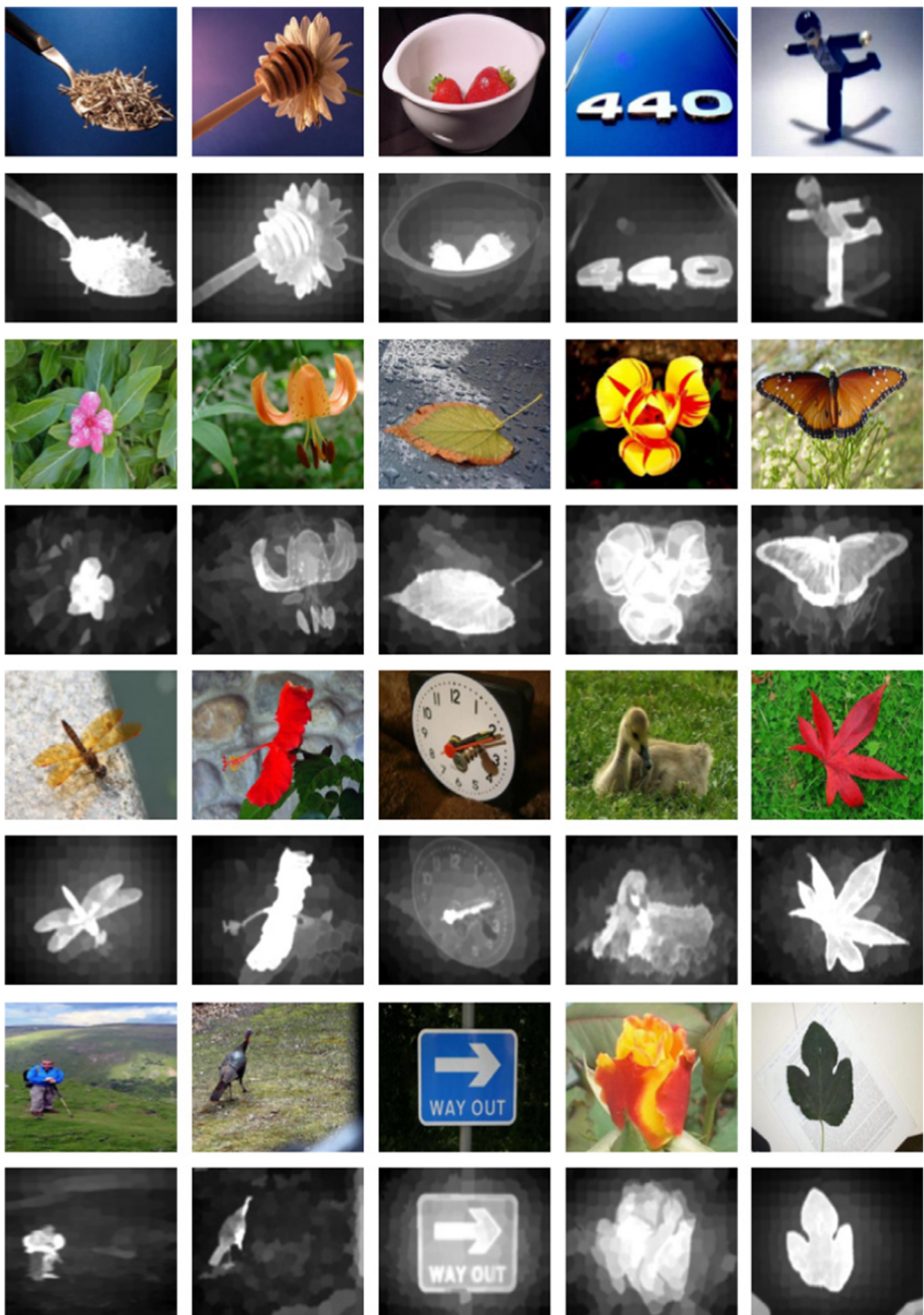


Fig. 3 Some distinctive saliency-detection results using our method based on the MSRA database

where $\phi(v_j)$ is the total number of pixels in v_j . The idea is that the contribution of those regions with more pixels should be larger than those with fewer pixels. The local contrast parameter $\varphi(v_i, v_j)$ is set to $\exp(-d(v_i, v_j)/\delta^2)$, and $d(v_i, v_j)$ is the Euclidean distance between the patch pair (v_i, v_j) .

3.4 Final saliency fusion

In the final stage of our algorithm, the integrated holistic directional map Dir is fused with the object cues of the center prior map, Cen , and the color prior map, Cor , to produce the final saliency map as follows:

$$Sal = Dir \cdot * Cen \cdot * Cor, \quad (14)$$

where $\cdot *$ is the pixel-by-pixel multiplication operator.

4 Experimental results

In this section, we present experimental results based on two publicly available benchmark datasets: MSRA [26] and ECSSD [34], and compare our method with other eight state-of-the-art methods: including the Itti [13], fuzzy growing (FG) [24], graph-based (GB) [10], multi-scale contrast MC [22], spectral residual (SR) [12], linear neighbourhood propagation (LNP) [45], meanshift filtering (MF) [20], multi-scale contrast and colour information (MCI) [46]. In order to quantitatively compare the state-of-

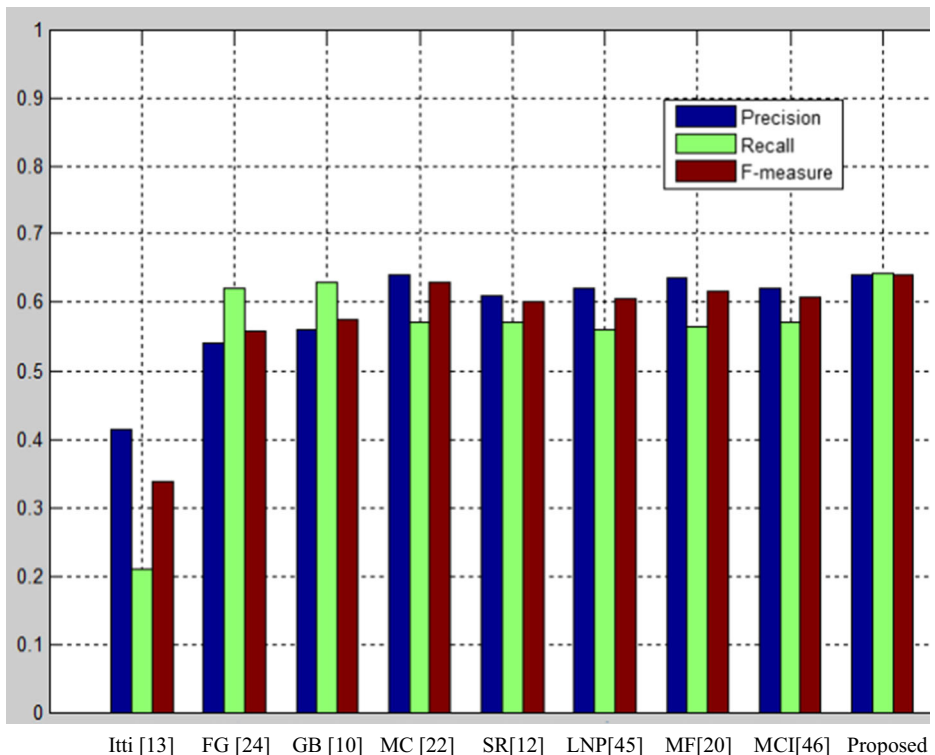


Fig. 4 Comparison of different saliency-detection methods in terms of average precision, recall, and F -measure on the MSRA dataset. Itti [13] FG [24] GB [10] MC [22] SR[12] LNP[45] MF[20] MCI[46] Proposed

the-art saliency-detection methods, the average precision, recall, and F -measure are utilized to measure the quality of the saliency maps. The adaptive threshold is twice the average value of the whole saliency map to get the accurate results. Each image is

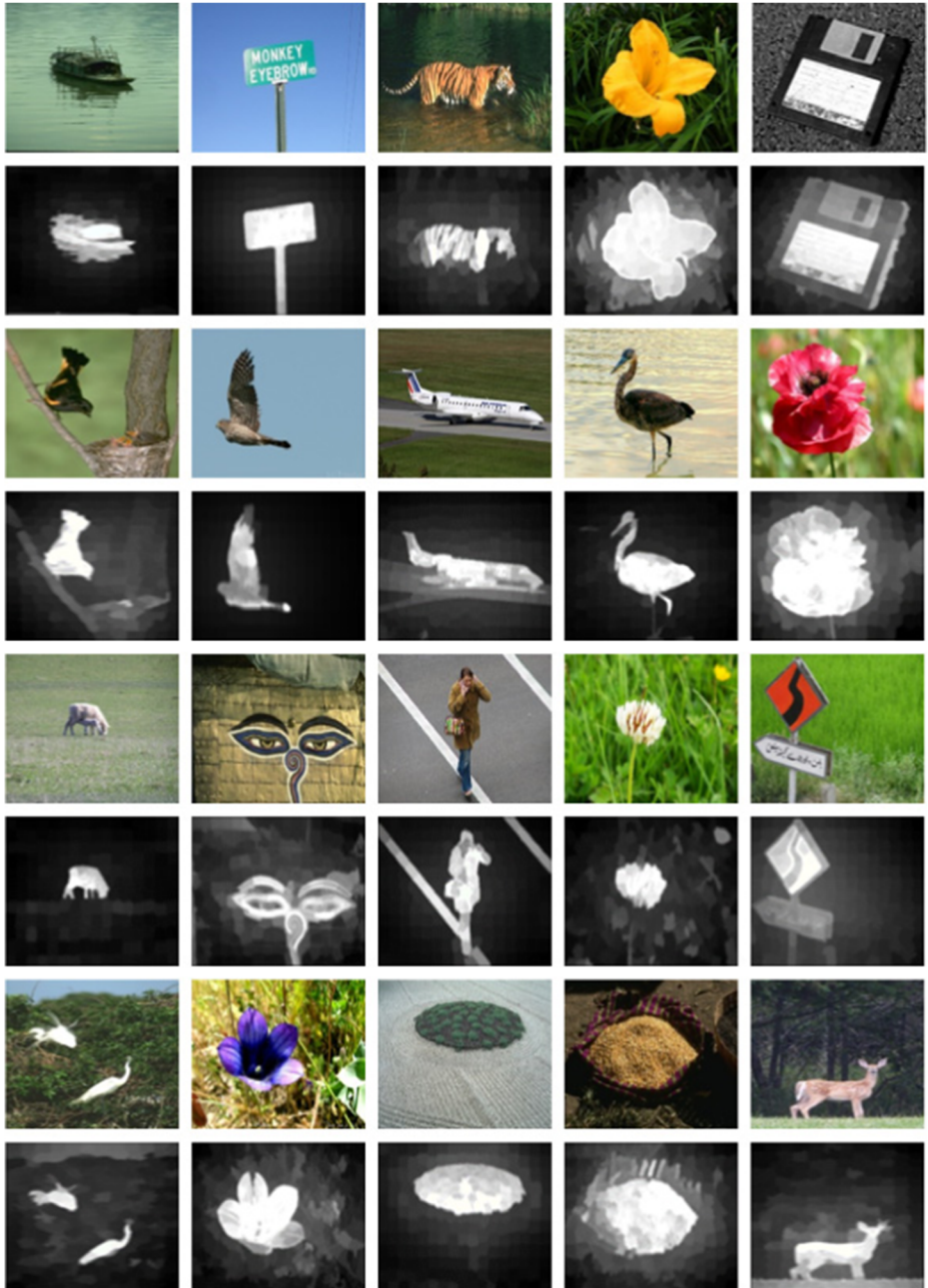


Fig. 5 Samples of the saliency detection results based on ECSSD database

segmented with superpixels and masked out when the mean saliency values are lower than the adaptive threshold. The F -measure is defined as follows:

$$F_{\beta} = \frac{(1 + \beta) \times \text{Precision} \times \text{Recall}}{\beta \times \text{Precision} + \text{Recall}} \quad (15)$$

where β is a real positive value and is set at $\beta = 0.3$ according to [5, 34].

4.1 MSRA database

We first evaluate our proposed algorithm on the MSRA database. It contains 5000 images with pixel-wise ground truth, which is widely used for saliency detection and most of the images include only one salient object with high contrast to the background.

Figure 3 shows some saliency detection results based on MSRA database. It can be seen that the proposed scheme is able to produce reliable and promising saliency-detection results. We also compared the performance of the proposed method with other eight state-of-the-art saliency detection methods. Figure 4 shows the precision, recall and the F -measure values of all the different methods. From the comparison, we can see that most have results higher than 55%, namely FG [24], GB [10], MC [22], SR [12], LNP [45], MF [20], MCI [46] and our proposed algorithm. For the overall F -measure, all the methods, except Itti [13], are higher than 50%, and our proposed model achieves the highest performance according to the overall among the nine different state-of-the-art saliency-detection models. Our proposed method

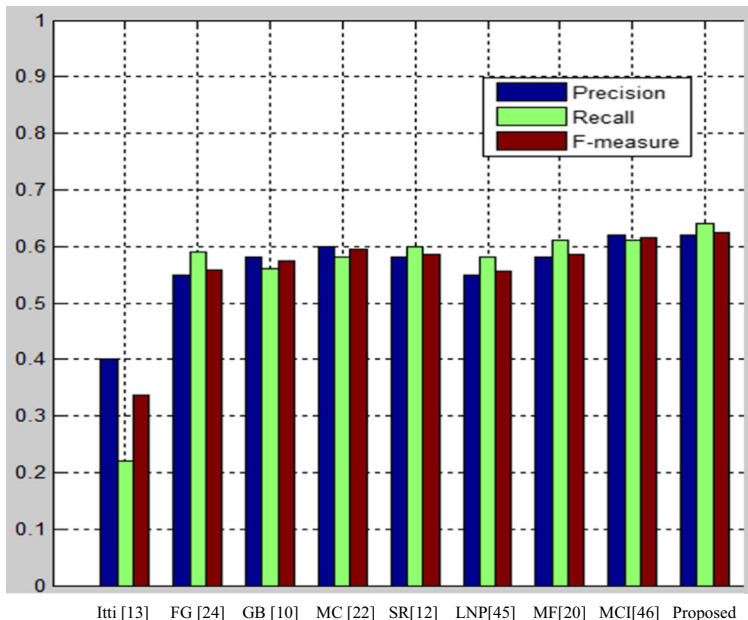


Fig. 6 Performance of the proposed model compared with state-of-the-art methods on ECSSD dataset in terms of average precision, recall, and F -measure

outperforms the other eight methods in terms of detection accuracy. The extensive experiment results show that our proposed model is efficient and accurate.

4.2 ECSSD database

We also test the proposed model on the openly available ECSSD database [34], which includes many semantically meaningful but structurally complex images for performance evaluation. There are 1000 images in this database. The images are acquired from the internet and 5 helpers were asked to produce the ground truth masks.

Figure 5 shows some results of saliency maps generated by our method based on the ECSSD database. The results show that our final saliency maps can accurately detect almost entire salient objects and preserve the salient object's contours clearly. Similar to the experiments on the MSRA database, we also used the precision, recall and the F -measure to evaluate the performance of our proposed method. Figure 6 shows the comparisons of different methods according to different evaluation criterions (precision, recall and the F -measure). As can be seen from Fig. 6, our proposed method outperforms the other eight methods in terms of detection accuracy and the proposed method achieves the best overall saliency-detection performance (with precision = 62.0%, recall = 63.0%), and the F -measure is 62.2%. The experiment results show that the proposed model is efficient and effective.

4.3 Model analysis and limitation

In this subsection, we provide the model analysis to analyze the performance of each individual component in the proposed scheme. Figure 7a shows the performance of each step in our proposed algorithm, we can see that the components of Quaternionic Distance Based Weber Local Descriptor (QDWLD), center prior, and color prior are all contributed to our final saliency maps and are complementary for each other. In Fig. 7b, we also present and analyze the results of the proposed method according to different quaternionic distances (QDs). From the comparison, we can see that the individual quaternionic distance has its own contributory effect on the final saliency detection results. In order to achieve a robust and better

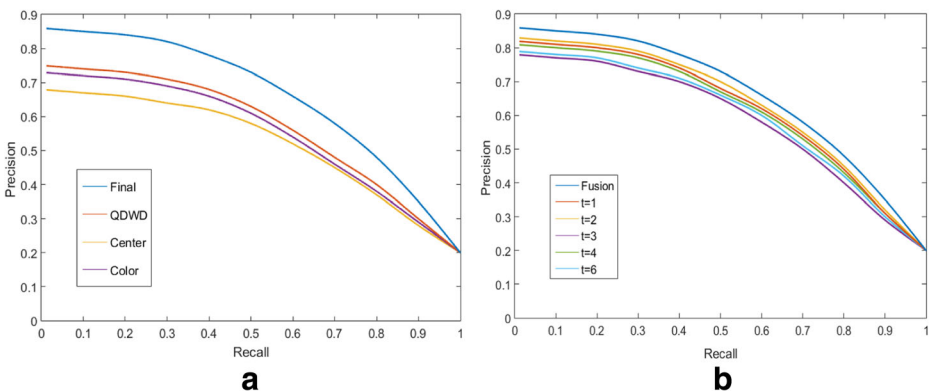


Fig. 7 Model analysis. **a** shows performance of each component in the proposed method, **(b)** shows the results of our method in different quaternionic distance. Both of them are evaluated on the MSRA dataset

performance, the fusion of QDWLD from different quaternionic distances (QDs) is used to represent the final directional cues in an image for saliency detection.

5 Conclusion and discussion

In this paper, we have proposed a new bottom-up method for efficient and accurate saliency detection. In the proposed approach, the integrated holistic directional map generated by QDWLD and the object cues are utilized to estimate the final saliency map. We have performed our method on two publicly available datasets, and experimental results show that our algorithm is effective and efficient.

Acknowledgements We would like to thank Dr. Rushi Lan in the Faculty of Science and Technology, University of Macau for providing the QDWLD Matlab code.

This work was supported by National Natural Science Foundation of China (NSFC) (61601427, 61602229); Natural Science Foundation of Shandong Province (ZR2015FQ011); China Postdoctoral Science Foundation funded project (2016 M590659); Postdoctoral Science Foundation of Shandong Province (201603045); Qingdao Postdoctoral Science Foundation funded project (861605040008) and Applied Basic Research Project of Qingdao (16-5-1-4-jch); The Fundamental Research Funds for the Central Universities (201511008, 30020084851); & Technology Cooperation Program of China (ISTCP) (2014DFA10410).

References

1. Achanta R, Hemami S, Estrada F, et al (2009) Frequency-tuned salient region detection. *IEEE Computer vision and pattern recognition*, pp 1597–1604
2. Achanta R, Shaji A, Smith K et al (2012) SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans Pattern Anal Mach Intell* 34(11):2274–2282
3. Cai C, Mitra SK (2000) A normalized color difference edge detector based on quaternion representation. *IEEE Int Conf Image Proces (ICIP)* 2:816–819
4. Cheng MM, Zhang Z, Lin WY, et al (2014) BING: binarized normed gradients for objectness estimation at 300 fps. *IEEE conference on computer vision and pattern recognition*, pp 3286–3293
5. Cheng MM, Mitra NJ, Huang X et al (2015) Global contrast based salient region detection. *IEEE Trans Pattern Anal Mach Intell* 37(3):569–582
6. Geng X, Hu X, Xiao J (2012) Quaternion switching filter for impulse noise reduction in color image. *Signal Process* 92(1):150–162
7. Goferman S, Zelnik-Manor L, Tal A (2012) Context-aware saliency detection. *IEEE Trans Pattern Anal Mach Intell* 34(10):1915–1926
8. Guo C, Ma Q, Zhang L (2008) Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform. *IEEE conference on computer vision and pattern recognition (CVPR)*, pp 1–8
9. Guo L, Dai M, Zhu M (2014) Quaternion moment and its invariants for color object classification. *Inf Sci* 273:132–143
10. Harel J, Koch C, Perona P (2006) Graph-based visual saliency. *Advances in neural information processing systems*, pp 545–552
11. He S, Lau RWH, Yang Q (2015) Exemplar-driven top-down saliency detection via deep association. *IJCV* 115(3):330–344
12. Hou X, Zhang L (2007) Saliency detection: a spectral residual approach. *IEEE Conference on Computer Vision and Pattern Recognition*, pp1–8
13. Itti L, Koch C, Niebur E (1998) A model of saliency based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Mach Intell* 20(11):1254–1259
14. Jian M, Lam KM, Dong J et al (2015) Visual-patch-attention-aware saliency detection. *IEEE Trans. Cybern* 45(8):1575–1586
15. Jian M, Qi Q, Sun Y, et al (2016) Saliency detection using quaternionic distance based weber descriptor and object cues. *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC'2016)*, Korean

16. Jin L, Li D (2007) An efficient color-impulse detector and its application to color images. *IEEE Signal Process Lett* 14(6):397–400
17. Jin L, Liu H, Xu X et al (2013) Quaternion-based impulse noise removal from color video sequences. *IEEE Trans Circuits Syst Video Technol* 23(5):741–755
18. Lan R, Zhou Y, Tang Y (2015) Quaternionic weber local descriptor of color images. *IEEE Trans Circuits Syst Video Technol*. doi:10.1109/TCSVT.2492839
19. Lei H, Xie H, Zou W et al (2017) Hierarchical saliency detection via probabilistic object boundaries. *Int J Pattern Recognit Artif Intell* 31(06):1755010
20. Li J, Chen H, Li G et al (2015) Salient object detection based on meanshift filtering and fusion of colour information. *IET Image Process* 9(11):977–985
21. Li H, Chen J, Lu H et al (2017) CNN for saliency detection with low-level feature integration. *Neurocomputing* 226:212–220
22. Liu T, Sun J, Zheng N et al (2011) Learning to detect a salient object. *IEEE Trans Pattern Anal Mach Intell* 33(2):353–367
23. Liu Z, Xue Y, Yan H et al (2011) Efficient saliency detection based on gaussian models. *IET Image Process* 5(2):122–131
24. Ma YF, Zhang, HJ (2003) Contrast-based image attention analysis by using fuzzy growing. *Proceedings of the Eleventh ACM International Conference on Multimedia*, pp 374–381
25. Manipoonchelvi P, Muneeswaran K (2014) Region-based saliency detection. *IET Image Process* 8(9):519–527
26. MSRA, http://research.microsoft.com/en-us/um/people/jiansun/SalientObject/salient_object.htm
27. Oliva A, Torralba A, Castelhano MS et al (2003) Top-down control of visual attention in object detection. *IEEE ICIP* 1
28. Rahman I, Hollitt C, Zhang M (2016) Contextual-based top-down saliency feature weighting for target detection. *Mach Vis Appl* 27(6):893–914
29. Song H, Liu Z, Xie Y et al (2016) RGBD co-saliency detection via bagging-based clustering. *IEEE Signal Process Lett* 23(12):1707–1711
30. Wang A, Wang M (2017) RGB-D salient object detection via minimum barrier distance transform and saliency fusion. *IEEE Signal Process Lett*
31. Wu W, Zhao J, Zhang C, et al (2017) Improving performance of tensor-based context-aware recommenders using bias tensor factorization with context feature auto-encoding. *Knowledge-Based Systems*
32. Xie Y, Lu H (2011) Visual saliency detection based on Bayesian model. *IEEE International Conference on Image Processing*, pp 645–648
33. Xie Y, Lu H, Yang MH (2013) Bayesian saliency via low and mid-level cues. *IEEE Trans Image Process* 22(5):1689–1698
34. Yan Q, Xu L, Shi J, Jia J (2013) Hierarchical saliency detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
35. Yan C, Zhang Y, Xu J, Dai F, Li L, Dai Q, Wu F (2014) A highly parallel framework for HEVC coding unit partitioning tree decision on many-core processors. *IEEE Signal Process Lett* 21(5):573–576
36. Yan C, Zhang Y, Dai F, Wang X, Li L, Dai Q (2014) Parallel deblocking filter for HEVC on many-core processor. *Electron Lett* 50(5):367–368
37. Yan C, Zhang Y, Dai F, Zhang J, Li L, Dai Q (2014) Efficient parallel HEVC intra-prediction on many-core processor. *Electron Lett* 50(11):805–806
38. Yan C, Zhang Y, Xu J, Dai F, Zhang J, Dai Q, Wu F (2014) Efficient parallel framework for HEVC motion estimation on many-core processors. *IEEE Trans Circuits Syst Video Technol* 24(12):2077–2089
39. Yang J, Yang MH (2017) Top-down visual saliency via joint crf and dictionary learning. *IEEE Trans Pattern Anal Mach Intell* 39(3):576–588
40. Yang C, Zhang L, Lu H (2013) Graph-regularized saliency detection with convex-hull-based center prior. *IEEE Signal Process Lett* 20(7):637–640
41. Zhai Y, Shah M (2006) Visual attention detection in video sequences using spatiotemporal cues. *ACM international conference on multimedia*, pp 815–824
42. Zhang J, Sclaroff S (2016) Exploiting Surroundedness for saliency detection: a Boolean map approach. *IEEE Trans Pattern Anal Mach Intell* 38(5):889–902
43. Zhong G, Liu R, Cao J et al (2016) A generalized nonlocal mean framework with object-level cues for saliency detection. *Vis Comput* 32(5):611–623
44. Zhong SH, Liu Y, Ng TY, Liu Y (2016) Perception-oriented video saliency detection via spatio-temporal attention analysis. *Neurocomputing* 207:178–188
45. Zhou J, Gao S, Yan Y et al (2014) Saliency detection framework via linear neighbourhood propagation. *IET Image Process* 8(12):804–814
46. Zhou W, Song T, Li L et al (2014) Multi-scale contrast-based saliency enhancement for salient object detection. *IET Comput Vis* 8(3):207–215



Muwei Jian received the PhD degree from the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, in October 2014. He joined the Department of Computer Science and Technology, Ocean University of China in January 2015. His current research interests include image processing, wavelet analysis, 3-D multimedia analysis, and human face hallucination/recognition. Dr. Jian has served as a reviewer for several international SCI-indexed journals, including *Pattern Recognition*, *Computers in Industry*, *Machine Vision and Applications*, *Machine Learning and Cybernetics*, *The Imaging Science Journal*, and *Multimedia Tools and Applications*. He holds 3 granted national patents and has published over 30 papers in refereed international leading journals/conferences such as *IEEE Trans. on Cybernetics*, *IEEE Trans. on Circuits and Systems for Video Technology*, *Information Sciences*, *Pattern Recognition*, *Signal Processing*, *ISCAS*, *ICME* and *ICIP*.

Qiang Qi is pursuing his Master's degree at the Department of Computer Science and Technology, Ocean University of China. His research interests include image processing, pattern recognition, and underwater vision.



Junyu Dong received his B.Sc. and M.Sc. in Applied Mathematics from the Ocean University of China (formerly called Ocean University of Qingdao) in 1993 and 1999, respectively. He won the Overseas Research Scholarship and James Watt Scholarship for his PhD study in 2000 and was awarded a Ph.D. degree in Image Processing in 2003 from the School of Mathematical and Computer Sciences, Heriot-Watt University, UK. Dr. Junyu Dong joined Ocean University of China in 2004. From 2004 to 2010, Dr. Junyu Dong was an associate professor at the Department of Computer Science and Technology. He became a Professor in 2010 and is currently the Head of the Department of Computer Science and Technology. Prof. Dong was actively involved in professional activities. He has been a member of the program committee of several international conferences,

including the 4th International Workshop on Texture Analysis and Synthesis (associated with ICCV2005), the 2006 British Machine Vision Conference (BMVC 2006) and the 3rd International Conference on Appearance (Predicting Perceptions 2012). Currently, Prof. Dong is the Chairman of Qingdao Young Computer Science and Engineering Forum (YOCSEF Qingdao). He is a member of ACM and IEEE. Prof. Dong's research interest includes texture perception and analysis, 3D reconstruction, video analysis and underwater image processing.

Xin Sun is currently an associate professor at the Department of Computer Science and Technology, Ocean University of China. His research interests include machine learning, bioinformatics, and image processing.



Yujuan Sun is a lecturer in the College of Information and Electrical Engineering at Ludong University of China. She has published more than 20 journal and international conference papers (indexed by EI and ISTP). Her research interests include 3D shape reconstruction from images, pattern recognition, and machine learning.



Kin-Man Lam received the Associateship in Electronic Engineering with distinction from The Hong Kong Polytechnic University (formerly called Hong Kong Polytechnic) in 1986, the M.Sc. degree in communication engineering from the Department of Electrical Engineering, Imperial College of Science, Technology and Medicine, London, U.K., in 1987, and the Ph.D. degree from the Department of Electrical Engineering, University of Sydney, Sydney, Australia, in August 1996. He was a Lecturer with the Department of Electronics

Engineering, The Hong Kong Polytechnic University, from 1990 to 1993. He joined the Department of Electronics and Information Engineering, The Hong Kong Polytechnic University, as an Assistant Professor, in 1996, where he became an Associate Professor in 1999. He was involved in professional activities. He is currently a Professor with The Hong Kong Polytechnic University. His research interests include human face recognition, image and video processing, and computer vision. Prof. Lam has been a member of the Organizing Committee or Program Committee of many international conferences. He was a Secretary of the IEEE International Conference on Acoustics, Speech, and Signal Processing in 2003, the Technical Chair of the International Symposium on Intelligent Multimedia, Video and Speech Processing in 2004, the Technical Co-Chair of the International Symposium on Intelligent Signal Processing and Communication Systems in 2005, a Secretary of the International Conference on Image Processing in 2010, the Technical Co-Chair of the Pacific-Rim Conference on Multimedia in 2010, and the General Co-Chair of the IEEE International Conference on Signal Processing, Communications and Computing in Hong Kong in 2012. He was the Chairman of the IEEE Hong Kong Chapter of Signal Processing from 2006 to 2008. He has served as an Associate Editor of ***IEEE TRANSACTIONS ON IMAGE PROCESSING*** and the Director of Student Services of the IEEE Signal Processing Society. He is the Vice President of Member Relations and Development of the Asia-Pacific Signal and Information Processing Association (APSIPA) and the Director of Membership Services of the IEEE Signal Processing Society. He is an *Area Editor of IEEE Signal Processing Magazine* and an *Associate Editor of Digital Signal Processing, APSIPA Transactions on Signal and Information Processing, and EURASIP International Journal on Image and Video Processing*. He is also an Editor of *HKIE Transactions*.