

# Salient region detection using efficient wavelet-based textural feature maps

Masoumeh Rezaei Abkenar<sup>1</sup> · M. Omair Ahmad<sup>1</sup> 

Received: 13 February 2017 / Revised: 29 August 2017 / Accepted: 4 September 2017 /  
Published online: 12 September 2017  
© Springer Science+Business Media, LLC 2017

**Abstract** A salient region is part of an image that captures the highest level of attention by the human visual system. In this paper, a new salient region detection method is proposed by linearly combining the feature maps for the  $L$ ,  $a$  and  $b$  color channels. Since, the wavelet transform is capable of providing a multi-scale spatial-frequency decomposition of the image, the color feature maps are obtained using this transform. A scheme is proposed whereby the channel feature maps are linearly combined. The weights for the linear combination are determined by making use of the entropy of the channel feature maps and a Gaussian kernel, utilizing the fact that the salient objects are generally clustered and scene-centric. The salient region is further refined by making use of the proximity of the pixels to the centers of gravity in the image feature map. Extensive simulations are conducted in order to evaluate the performance of the proposed saliency detection scheme by applying it to the natural images from several datasets. Experimental results show that the proposed method provides values of precision, recall and F-measure larger than and that of the mean absolute error smaller than those provided by other existing methods. The performance of the proposed salient region detection method is also evaluated on noisy images and it is shown to be robust against noise.

**Keywords** Salient region detection · Image color feature maps · Wavelet-based feature maps · Entropy-based feature map combination

---

✉ M. Omair Ahmad  
omair@ece.concordia.ca

<sup>1</sup> Department of Electrical and Computer Engineering, Concordia University, Montreal, QC H3G 1M8, Canada

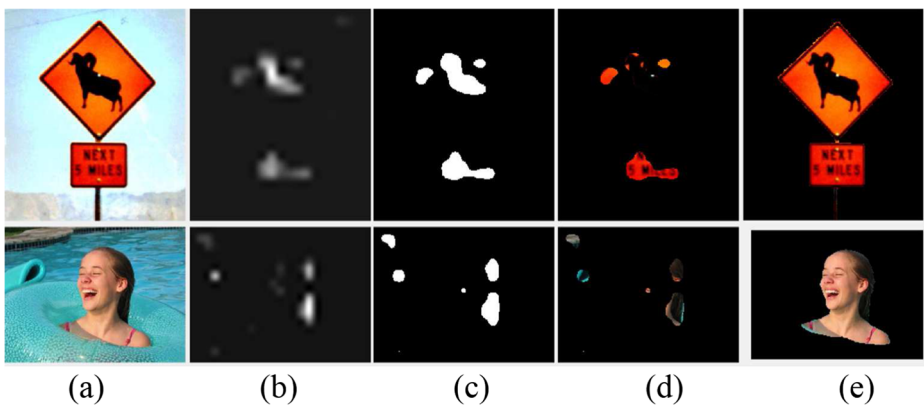
## 1 Introduction

Human visual systems (HVS) continuously receive huge amount of information. This amount of visual information is far greater than the storage and processing capability of the brain. In order to analyze this information, HVS is able to detect the most distinctive region of the scene rapidly. This region, which is called the salient region, is known to be conspicuous and has more contrast with respect to the local or global surrounding regions. It is more distinctive in terms of color, edges, boundaries, etc. Salient region detection methods try to detect these regions automatically. Figure 1 shows examples of natural images and the gray level and binary saliency maps obtained by using the Itti's method [19], a pioneering saliency detection method, the corresponding extracted salient regions, and the ground truth for the salient regions. Recently, researches have shown a great deal of interest in developing saliency detection techniques due to its variety of applications, such as object of interest image segmentation [20, 38], adaptive image or video compression [17], object-based image retrieval [25], image retargeting [5, 37], and medical imaging [35]. Saliency detection can also be used to improve the efficiency of high resolution displays [32].

There are several saliency detection techniques that have been developed in the past couple of decades, both in the spatial domain and in the frequency domain, with an objective to detect the entire salient objects [3, 14, 16, 18, 22, 24, 25, 31]. A comprehensive review of the works in saliency detection can be found in the survey papers [7–9].

In most of the saliency detection methods, some image features, such as color contrast, texture and structure, are extracted and a saliency value is assigned to each pixel based on these features. Choosing a feature that is effective for the purpose of saliency detection is a challenging task. These features or attributes should be able to provide significant information towards the local and global characteristics of the individual pixels of the image.

To detect the salient objects, several spatial domain methods have been developed [3, 24, 25, 31]. Spatial domain methods first try to extract different features from the image using its pixel values. Then, the feature maps, each using one type of feature, are computed by employing a center-surround operation [19, 24], contrast computation [25], Shannon's self-information measure [10], or graph-based random walk [15]. Finally, the various feature maps are normalized and linearly or non-linearly combined to obtain a saliency map. In [25],



**Fig. 1** **a** Sample natural images. **b** Gray level saliency maps obtained by using the Itti's method [19]. **c** The corresponding binary saliency maps. **d** Extracted salient objects. **e** Ground truth for the salient objects

the contrast between the two pixels is determined by the Gaussian distance between their *LUV* color components. Then, to extract salient regions, instead of using a fixed threshold, fuzzy theory is employed in a region-growing process. In [31], a biologically plausible method of forming the proto-objects from the saliency map is proposed based on the Itti's method [19], when proto-objects are described as volatile units of visual information. However, this method does not attempt to obtain a segmentation of salient objects themselves in the image. In [24], a salient object detection method is proposed by using multi-scale contrast, center-surround histogram, and color spatial distribution. A conditional random field is learned to effectively combine these features for salient object detection. It aims to separate the entire salient object from the background. However, it has some limitations for images with multiple salient objects. Also, the salient object is identified within a rectangular box and the actual boundary of the object cannot be detected. In [3], a salient object detection method is proposed utilizing a difference of Gaussian filters with appropriate range of spatial frequencies. The authors determine the saliency value of each pixel by the difference between the *CIELAB* color values of each pixel and the mean *CIELAB* color value of the entire image. High computational cost and choice of appropriate parameters, are the main weaknesses of most of the spatial domain saliency detection methods [36]. In addition, some of these methods cannot detect the salient objects from the background with clear boundaries [15, 19]. This may be due to an inappropriate reduction of the high frequency content of the original image. Some other spatial domain methods obtain saliency maps in which the salient region boundaries are clear, but the entire salient region is not uniformly highlighted [25], or the textured regions are highlighted regardless of their contrast with their surrounding regions [10]. From a frequency domain perspective, these limitations are the consequence of retaining an inappropriate range of frequency content from the original image [3].

In order to address the above mentioned limitations of the spatial domain methods, and because of the characteristics of the frequency domain representations of images, frequency domain methods have been developed [4, 14, 16, 18, 22, 23, 27–29, 33]. In frequency domain methods, saliency is detected by following the steps of applying the frequency transform to the input image, modulating the frequency spectrum in order to suppress the background and enhance the salient regions, and finally generating the gray level saliency map through the inverse transform. As the first saliency detection method in the frequency domain, Hou et al. [16] argued that the averaged log amplitude spectrum of the Fourier transform (FT) of the image contains redundancies. Also, statistical singularities in the amplitude spectrum are responsible for the salient objects. Thus, the spectral residual (SR), which is the difference between the log amplitude spectrum and the averaged log amplitude spectrum represents the salient regions of the image. The saliency map is determined by applying the inverse FT on the SR and the original phase. Later, Guo et al. [14] argued that the SR is not essential to obtain the saliency map, and the saliency map can be obtained by reconstructing the two-dimensional (2D) signal using only the phase spectrum. In [22], the color and intensity features of the image are combined into a quaternion, and the quaternion Fourier transform (QFT) is applied to this feature vector. To suppress the non-salient regions, the amplitude spectrum is convolved with a low-pass Gaussian kernel of different scales. At each scale, the saliency maps is obtained by reconstructing the 2D signal using the smoothed amplitude spectrum and the original phase, and the final saliency map is chosen based on a minimum entropy criterion. While the FT-based frequency domain methods have the potential to more successfully detect the entire salient object, they still have tendency to highlight the boundaries rather than entire salient regions. In the Fourier transforms of images, the global contrast is more dominant than the local contrast. Therefore, these methods result in a

saliency map in which the global irregularities, such as edges from a textured background or objects' boundaries, are more dominant than the local irregularities, such as homogeneous salient regions. Moreover, it is known that the FT cannot provide simultaneous spatial and frequency localization, and it is not useful for analyzing non-stationary signals such as most of the natural images. The short-time-Fourier transform (STFT) can be utilized to perform local frequency analysis. It segments the signal into narrow spatial intervals (i.e., narrow enough to be considered stationary) and take the FT of each segment. But fixed size of the interval is the main drawback of STFT. For low frequencies, a proper frequency resolution is needed, while for high frequencies, the spatial resolution is more important [12, 26]. In addition, trying several spatial intervals on the image increases the computation time.

To overcome the shortcoming of the Fourier-based methods, some saliency detection methods have been proposed utilizing the discrete wavelet transform (DWT) [18, 27, 29]. Basis functions of the DWT are functions with varying frequency and limited duration, which enables to do a spatial and frequency analysis at the same time. The DWT is able to provide a multi-scale analysis, in which the signal is represented and analyzed at more than one resolution. Therefore, the ability of doing a localized multi-resolution spatial and frequency analysis in DWT makes it an appropriate option for extracting oriented details of an image and detecting the salient regions of different sizes. In [29], a DWT-based salient point detector for image retrieval is proposed. In this method, the points with higher global variation based on the absolute wavelet coefficients at courser scales are selected, and are tracked along the finer scales to detect the salient points. In [27], for each color sub-band, the saliency map is obtained by an inverse DWT over a set of scale-weighted center-surround responses. Where the weights are derived from the high-pass wavelet coefficients of each level. However, the methods in [27, 29] are only able to obtain salient points rather than salient regions. In [18], DWT is applied to the image, and then at each decomposition level a feature map is generated by setting the low-pass coefficients of that level to 0, and applying the inverse DWT. Then, local contrast is obtained by a linear combination of the feature maps, and global distinctiveness is computed based on normal distribution of the feature maps. However, since the last decomposition level consists of all the detail coefficients of the previous levels, no additional information is obtained by applying the inverse transform at each decomposition level. Also, generating these feature maps increases the computational complexity of the method.

Although there are a number of methods in the frequency domain, it is still a challenging task to develop accurate methods that can obtain saliency maps with uniformly highlighted salient regions and sharp boundaries between salient and non-salient regions. As discussed earlier, the ability to do a localized multi-resolution spatial and frequency analysis in DWT, makes it a superior tool to extract image details at different scales. These details can be used to detect salient regions in an image. Moreover, since, in a color image different features can be extracted from each color channel, it is desirable to investigate an efficient approach to combine the extracted channel features.

In this paper, a saliency detection scheme is proposed by making use of the texture maps obtained from the high-pass coefficients of the DWT decompositions of the three channels of the *CIELAB* color space of images. The key points in devising a saliency detection scheme are in recognizing and extracting the features that correctly differentiate the salient regions from the non-salient ones of an image, and then utilizing them to detect the salient regions accurately. The core ideas of the the proposed saliency detection scheme are recognizing the DWT-based color channel textural details as suitable features to distinguish salient regions from non-salient ones, and devising a scheme for the weighted linear combination of the color channel features in order to detect and extract the salient objects. A

new method based on the entropy of the individual color channels within the image is proposed for determining the weights of the linear combination. The proposed scheme can be expected to yield an accurate saliency map in view of its ability in incorporating the efficient DWT-based textural details of the three color channels using the proposed weighting scheme.

The paper is organized as follows. In Section 2, the proposed scheme of saliency detection is developed by devising the various steps involved in obtaining the channel feature maps, their associated weights, and the image feature map that lead to obtaining the final saliency map. In Section 3, experiments are carried out in order to demonstrate the effectiveness of the proposed saliency detection scheme and compare its performance with that of some of the existing schemes in the literature. Also, performance of the proposed method in the presence of Gaussian and impulsive noise is studied. Finally, in Section 4, the work of this paper is summarized and the significant features of the proposed scheme highlighted.

## 2 Proposed saliency detection method

The proposed method attempts to detect salient regions of images using DWT-based textural features of the three color channels and incorporates them using an efficient weighting scheme. In this method, the input image is converted to the *CIELAB* color space, having luminance, red/green, and blue/yellow channels, denoted by  $L$ ,  $a$ , and  $b$ , respectively, since, this color space is perceptually more uniform than the *RGB* color space. Figure 2 shows the overall framework of the proposed saliency detection method. The proposed method consists of the steps of extracting feature maps corresponding to the three color channels using DWT decomposition, linearly combining the three channel feature maps based on the concept of entropy and a border avoidance criterion, modifying the feature map by making use of the centers of gravity of the image [21], and obtaining the final saliency map by bilateral filtering [30] of the modified image feature map. The following sub-sections describe these four steps in detail.

### 2.1 Wavelet-based feature maps of color channels

In the proposed technique, the DWT is utilized to extract textural details of the image at different scales,  $s \in \{1, \dots, N\}$ , where depending on the size of the input image and size of the wavelet filter used,  $N$  is the maximum level of DWT decomposition. The operation of DWT carried out individually to the  $L$ ,  $a$  and  $b$  channels of the image yields,

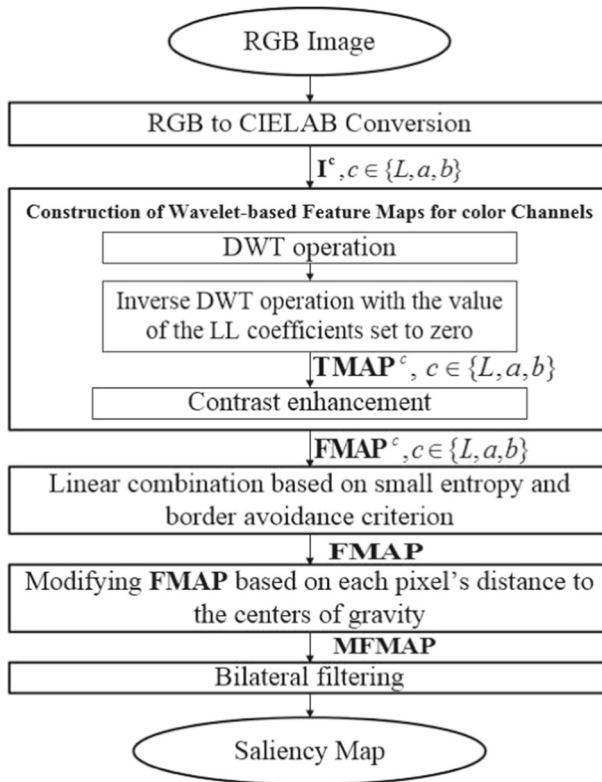
$$\left\{ A_N^c, \{H_s^c, V_s^c, D_s^c\}_{s \in \{1, \dots, N\}} \right\} = \text{DWT}(I^c), \quad (1)$$

where  $\text{DWT}(\cdot)$  denotes the discrete wavelet transform,  $I^c$ ,  $c \in \{L, a, b\}$ , represents the color channels of the input image,  $A_N^c$  is the approximation coefficients at the coarsest level,  $N$ , and  $H_s^c$ ,  $V_s^c$ ,  $D_s^c$  are the matrices of appropriate sizes representing the horizontal, vertical and diagonal sub-band coefficients of the channel  $c$  at level  $s$ , respectively.

To extract textural details from the DWT decomposition of the image, the  $LL$  coefficients at the coarsest decomposition level  $N$  are set to 0, and the 2D signal is reconstructed by applying the inverse DWT as

$$TMAP^c = \text{IDWT}(A_N^c = 0, \{H_s^c, V_s^c, D_s^c\}, s \in \{1, \dots, N\}), \quad (2)$$

where  $\text{IDWT}(\cdot)$  denotes the inverse wavelet transform, and  $TMAP^c$  is the texture map of the channel  $c$ .



**Fig. 2** Block diagram of the proposed wavelet-based saliency detection method

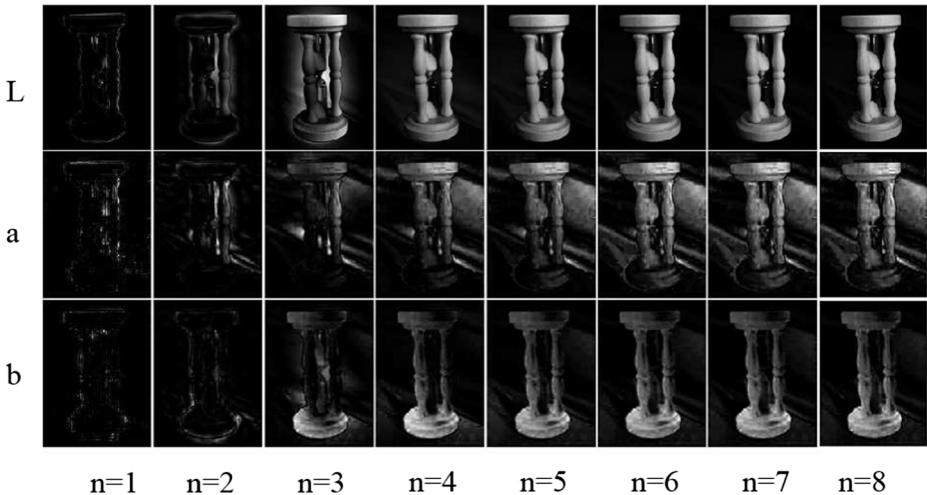
The channel feature map of a location  $(x, y)$ ,  $\text{FMAP}^c(x, y)$  for the channel  $c$ , is obtained by enhancing the high-intensity values and suppressing the low-intensity values of the channel texture map at that location,  $\text{TMAP}^c(x, y)$ , as

$$\text{FMAP}^c(x, y) = (\text{TMAP}^c(x, y))^2. \quad (3)$$

In order to investigate the effect of the number of decomposition levels used on the quality of the feature maps, we stop decomposition at different levels,  $n = 1, \dots, N$ , and construct a feature map for each  $n$  by setting the  $LL$  coefficients of the  $n$ th level to 0 and reconstructing the 2D signal by using the detail coefficients of all the levels from 1 to  $n$ . Figure 3 shows an example of the feature maps for the three channels resulting from a  $200 \times 150$  color image after each of the decomposition levels from  $n = 1$  to  $n = 8$ , the maximum decomposition level for the image of this size, by employing the *Daubechies* wavelets (Daub.7). It is seen from this figure that, as the number of decomposition levels is increased, the proposed method results in increasingly improved textural features for each of the channels. It is noted that for  $n < N$  the number of sub-bands used to construct the channel feature maps is a subset of the number of sub-bands used for constructing these channel feature maps for  $n = N$ . Thus, a linear combination of the feature maps for  $n = 1, \dots, N$ , for a given  $c$ , cannot be expected to improve the quality of the channel feature maps over that constructed by using  $n = N$  levels of decompositions. Thus, we advocate to construct the channel feature



(a)



(b)

**Fig. 3** **a** Original  $200 \times 150$  color image, **b** Channel feature maps obtained after  $n$  decomposition levels for  $L$ ,  $a$  and  $b$  channels

maps only after the last level of decomposition, i.e.,  $n = N$ , which should also result in reduced computational complexity.

### 2.2 Image feature map

After constructing channel feature maps for each of the  $L$ ,  $a$ , and  $b$  channels, an image feature map is obtained through a weighted linear combination of the channel feature maps as follows.

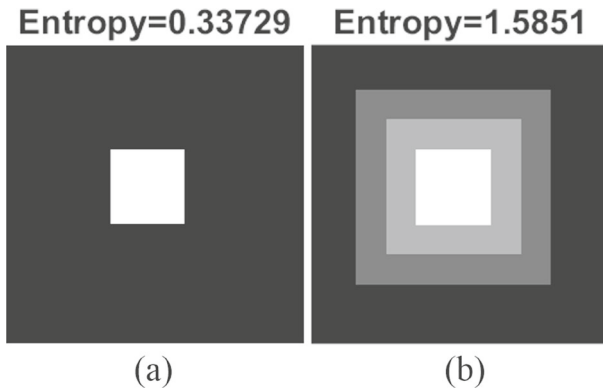
$$FMAP = \sum_{c \in \{L, a, b\}} \omega^c FMAP^c, \tag{4}$$

where  $\omega^c$  is the weight assigned to the feature map corresponding to the channel  $c$ .

In determining the values of weights,  $\omega^c$ , we focus on two aspects of a desirable feature map. A desirable feature map should have a cluster of pixels with high values of the gray levels corresponding to the salient region and the rest of pixels with low values. The other consideration is the fact that since salient objects are generally center biased, weights should be determined so as to provide less importance to the borders of the channel feature map.

In order to take into consideration the first aspect of a desirable feature map, we use the entropy of the channel feature maps. A channel feature map having only two levels of





**Fig. 4** Synthetic images with **a** two gray levels of 0 and 1, and **b** four gray levels of 0, 0.25, 0.5, 1 and their entropy values

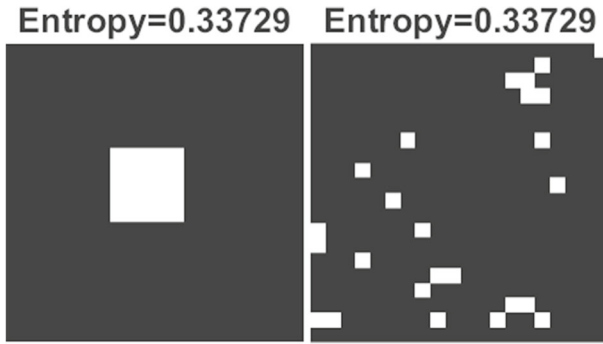
pixels would have an entropy value lower than that of a channel feature map having a larger number of gray levels of the pixels. Figure 4 shows an example of synthetic images with pixel values of 2 and 4 gray levels and their entropy values. Thus, the channel feature map with a smaller entropy should be assigned a larger weight and vice-versa. However, the problem with determining a weight based on the entropy value of the channel feature map is that it does not take into account the spatial distribution of the pixel gray levels across the map. Figure 5 shows an example of two binary images of size  $20 \times 20$  pixels. The number of pixels with a given gray level are the same in both the images. Despite the fact that distribution of these two gray levels are different in the two images, they both have the same entropy value. In order to use the entropy value for determining the weights, we would like to increase the entropy value in a situation in which the pixel gray levels of the channel feature map are more scattered. Therefore, we propose the channel feature map to undergo a low-pass filtering, a process through which the number of gray levels in the two images with different distributions will increase. Accordingly, their entropy value will increase. However, the increase in the entropy value of the image with a scattered distribution is larger than that of the image with a clustered distribution. Figure 6 shows the same two images as shown in Fig. 5 after they are filtered using a  $3 \times 3$  low-pass Gaussian filter, along with their entropy values. It is noted that after low-pass filtering, the entropy value of the image with the scattered distribution of pixels with the gray level of unity is increased much more than that of the image in which the pixels with gray level of unity are clustered together. Thus, the entropy value of the low-pass filtered channel feature map can be considered as a parameter in assigning weights to the channel feature maps. The entropy value is computed as

$$\epsilon^c = H(FMAP^c * G), \quad (5)$$

where  $H(\cdot)$  is the entropy,  $G$  is a low-pass Gaussian filter, and  $*$  represents the 2D convolution of the two associated matrices. The standard deviation of the Gaussian filter should be large enough so as to include sufficient number of neighboring pixels around each pixel in the filtering process.

In order to take into consideration the fact that in most of the natural images the salient region is located close to the center rather than on or near to the borders, in the linear combination of (4) a smaller weight is assigned to a channel feature map with a strong





**Fig. 5** Synthetic images of size  $20 \times 20$  pixels, each of two images has 25 pixels with the gray level of 1 and 375 pixels with the gray level of 0

response at the border. In this work, utilizing the border-avoidance criterion in [22], the strength of the channel feature map is computed with a greater emphasis given to a salient-like region at the center rather than at the border of the image, as

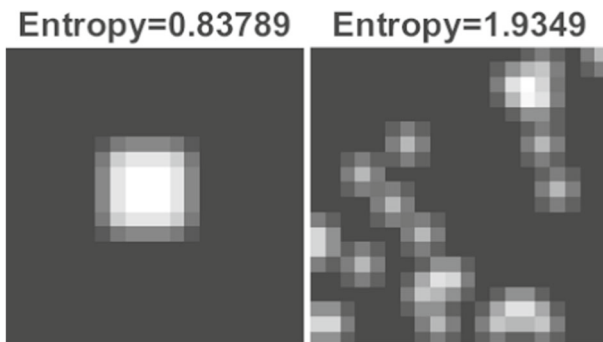
$$\beta^c = \sum_x \sum_y K(x, y) \cdot N(\text{FMAP}^c(x, y)), \tag{6}$$

where  $K(x, y)$  is the  $(x, y)$ th element of a Gaussian mask  $\mathbf{K}$ , of the same size as that of the channel feature map and its entries normalized to a maximum value of 1, and  $N(\text{FMAP}^c(x, y))$  represents the  $(x, y)$ th element of the normalized image feature map  $\text{FMAP}^c / \sum_x \sum_y \text{FMAP}^c(x, y)$ .

Thus, in order to emphasize the presence of a salient region at the center and de-emphasize any possible salient-like regions at the border of the image, we propose the weights in (4) to be chosen as

$$\omega^c = (\beta^c / \epsilon^c)^4. \tag{7}$$

Then, the image feature map is computed through the linear weighted combination of channel feature maps given by (4).



**Fig. 6** Images of Fig. 5 after low-pass filtering and their entropy values

### 2.3 Modification of the image feature map by considering centers of gravity

The image feature map obtained as above is further refined by taking centers of gravity into account. Centers of gravity are defined as one or several pixels about which the visual form of the image is organized [21]. The regions surrounding the centers of gravity attract our attention. Thus, the saliency value at the locations around the centers of gravity should be greater than those of the locations that are far away. In view of this, first the pixels whose intensity values in the feature map exceed a certain threshold are identified as centers of gravity in the same way as in [13]. Then, all the other pixels are weighted according to their Euclidean distances from the closest center of gravity in order to obtain a refined image feature map, as

$$\begin{aligned} \text{MFMAP}(x, y) \\ = \text{FMAP}(x, y) \left( 1 - \min \left\{ \|(x, y), (\acute{x}, \acute{y})\| \mid (\acute{x}, \acute{y}) \in \text{Centers of Gravity} \right\} \right), \end{aligned} \quad (8)$$

where  $\|\cdot\|$  represents the Euclidean positional distance between the pixel at location  $(x, y)$  and the center of gravity at  $(\acute{x}, \acute{y})$ .

In order to have a smooth saliency map, in most of the existing saliency detection methods, a low-pass Gaussian filter is applied to the saliency map in the last step [18, 22]. It is known that in Gaussian low-pass filtering, the pixel value of filtered image at a given location is computed as the weighted average of pixel values in a neighborhood specified by the standard deviation of the filter. The weight decreases as the distance of a pixel from the neighborhood center increases. Since the nearby pixels are likely to have the same intensity values, it is appropriate to average them together. However, the idea fails at the edges where there is an abrupt change in the intensity values of the neighboring pixels, and it results in blurred edges. Thus, low-pass filtering of the saliency map destroys the borders of the salient region.

In this work, in order to smooth the modified image feature map, *MFMAP*, while preserving the strong edges between salient and non-salient regions, a bilateral filter [30] is applied to it. Two pixels at an edge which are close to each other spatially could be very different in terms of their intensity values. The basic idea of bilateral filtering is considering both spatial closeness and intensity similarity of the pixels in assigning the weights in the filtering process. Thus, bilateral filtering can preserve high-contrast edges while removing low-contrast or gradual changes. A simple case of bilateral filtering is shift-invariant filtering, in which a Gaussian closeness filter and a Gaussian similarity filter are simultaneously used. In this work, the saliency map, *S*, is obtained by bilateral filtering of the modified image feature map as

$$S = \text{BF}(\text{MFMAP}), \quad (9)$$

where BF denotes the bilateral filtering operation. The value of a pixel at location  $p$  of the *MFMAP* after bilateral filtering is obtained as

$$\begin{aligned} \text{BF}[\text{MFMAP}_p] \\ = \frac{1}{W_p} \sum_{q \in \Omega} G_{\sigma_d}(\|p - q\|) G_{\sigma_r}(\|\text{MFMAP}_p - \text{MFMAP}_q\|) \text{MFMAP}_q, \end{aligned} \quad (10)$$

where  $W_p$  is a normalization factor given by

$$W_p = \sum_{q \in \Omega} G_{\sigma_d}(\|p - q\|) G_{\sigma_r}(|MFMAP_p - MFMAP_q|), \tag{11}$$

where  $p$  denotes a location  $(x, y)$  in  $MFMAP$ ,  $q$  denotes a location  $(\acute{x}, \acute{y}) \in \Omega$ ,  $\Omega$  being the set of possible positions in  $MFMAP$ , and  $G_{\sigma_d}$  and  $G_{\sigma_r}$  are the Gaussian closeness (domain) and similarity (range) functions with the standard deviations of  $\sigma_d$  and  $\sigma_r$ , respectively.

### 3 Experimental results

The performance of the proposed saliency detection method is evaluated on four commonly used datasets of natural images, namely, MSRA-1000 [3], complex scene saliency detection (CSSD) [34], CMU-Cornell iCoseg dataset [6] and MSRA-10K [11]. Each of these datasets contains images of different sizes and the corresponding pixel-level human-labeled ground truth. The MSRA-1000 is a widely used dataset and contains 1000 images. The CSSD dataset contains 200 images with more complex backgrounds. The CMU dataset contains 643 images of 38 related groups, where each group of images are of the same size. The MSRA-10K is a dataset of 10,000 images. To the best of our knowledge, this dataset is the largest available dataset of its kind. Therefore, it enables us to do a more extensive evaluation of the proposed and other saliency detection algorithms.

In the proposed method, the *Daubechies* wavelets (Daub.7) are used. The size of the filter results in a trade-off between the time complexity and promising quality of the saliency maps. The standard deviation of the low-pass Gaussian filter,  $G$ , in (5) is set to  $\sigma = 0.02 \times \frac{L+W}{2}$ , and the standard deviation of the Gaussian mask,  $K$ , in (6) is set to  $\acute{\sigma} = 0.25 \times \frac{L+W}{2}$ , where  $L$  and  $W$  are, respectively, the number of rows and columns in the original image. The threshold to identify centers of gravity in (8) is chosen to be 0.8. The standard deviations of the Gaussian closeness and similarity filters in the bilateral filter given by (10) are set to  $\sigma_d = \frac{\min(L,W)}{16}$  and  $\sigma_r = \frac{\max(S)-\min(S)}{10}$ , respectively.

Salient region detection can be considered as binary classification of salient and non-salient regions in the saliency map  $S$ . Since the ground truth is a binary image, the gray level saliency map  $S$  should be converted into a binary map  $S_b$  using a threshold value. In this work, two different schemes are used to convert the gray level saliency map into a binary map, fixed and adaptive. In fixed thresholding, all possible thresholds in the range of  $[0, 255]$  are applied to gray level saliency maps, and a binary map is generated corresponding to each of the threshold values and then compared to the ground truth. However, the use of fixed thresholds is not sufficient to evaluate the performance of a method, since it is image independent. Thus, an adaptive image dependent threshold is also utilized. In the adaptive thresholding, the threshold is defined to be twice the mean of the saliency values of all the pixels in the gray level saliency map as given by [3]

$$T_{adp} = \frac{2}{L \times W} \sum_x \sum_y S(x, y). \tag{12}$$

The performance of a saliency detection method is quantitatively evaluated using precision  $P$ , recall  $R$ , and F-measure  $F$ , values. Precision, also called positive predictive value, is the fraction of retrieved instances that are relevant, while recall, also known as sensitivity, is the

fraction of relevant instances that are retrieved, and F-measure is used as a combination of precision and recall values. The metrics  $P$ ,  $R$  and  $F$  are computed as

$$\begin{aligned} P &= \frac{\sum_x \sum_y S_b(x, y) \times G(x, y)}{\sum_x \sum_y S_b(x, y)}, \\ R &= \frac{\sum_x \sum_y S_b(x, y) \times G(x, y)}{\sum_x \sum_y G(x, y)}, \\ F &= \frac{(1 + \eta) \times P \times R}{(\eta \times P) + R}, \end{aligned} \quad (13)$$

where  $G(x, y)$  is the ground truth value at the location  $(x, y)$ , and  $\eta$  is a positive parameter specifying the relative importance of the precision and recall. In order to be consistent in comparing the performance of different methods, we choose the value of  $\eta$  to be 0.3.

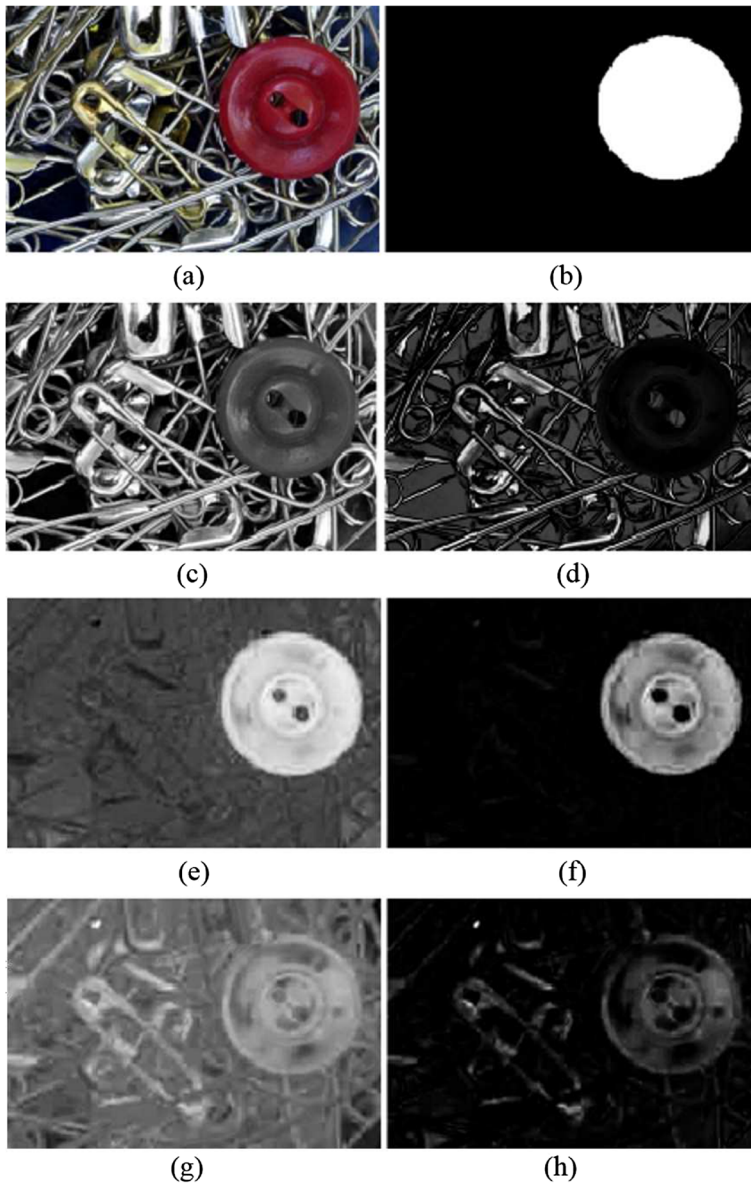
In order to make a comparison, the mean absolute error (MAE) between the binary saliency map  $S_b$  obtained by applying the adaptive threshold of (12) and the ground truth is also computed. The MAE value evaluates directly how similar is the binary saliency map to the ground truth, and it is given by

$$MAE = \frac{1}{L \times W} \sum_x \sum_y |S_b(x, y) - G(x, y)|. \quad (14)$$

### 3.1 Evaluating the impact of the main phases of the proposed method

In this section, the effect of each phase of the proposed method discussed in Section 2 is studied both subjectively and objectively. To this end, first, the wavelet-based channel feature maps (as given in (3)) for a sample image obtained by employing the proposed method. Figure 7 shows the three color channels and the channel feature maps obtained. It can be seen that for this sample image the channel feature map corresponding to the channel  $a$  can represent the salient region better than represented by the maps corresponding to the channels  $L$  and  $b$ . It is also seen that some details of the salient region has been captured by the  $b$  channel feature map, but not by the  $L$  channel feature map.

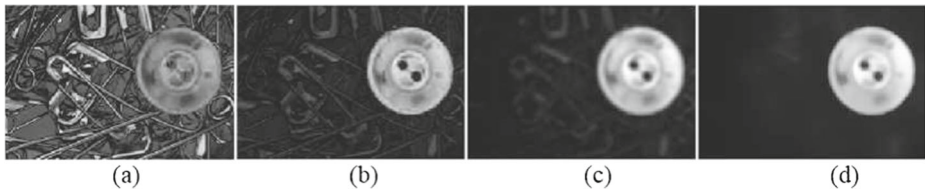
The effects of the succeeding steps that use the three channel feature maps given in Fig. 7d, f and h are depicted in Fig. 8. Figure 8a is simply the average of the three channel maps, that is, a map obtained by linearly combining the three maps with equal weights. Figure 8b shows the image feature map obtained through a weighted linear combination of the three channel feature maps. A comparison of the maps in Fig. 8a and b clearly shows a positive effect of the proposed weighting scheme, as described in Section 2.2, on the image feature map. It is seen that after applying the linear combination using the proposed weights given by (7), the non-salient regions are suppressed effectively compared to the simple averaging of the channel feature maps. Figure 8c shows the modified image feature map obtained by taking into consideration the centers of gravity in the image feature map of Fig. 8b. A comparison of the maps in Fig. 8b and c shows that the saliency values of the salient regions have increased, while those of the non-salient regions have decreased or remained unchanged. However, there are still some regions in the background which have been wrongly detected as salient. Finally, Fig. 8d shows the gray level saliency map by applying the bilateral filtering on the modified image feature map of Fig. 8c. It is seen that, after the bilateral filtering the wrongly detected regions in the background are further



**Fig. 7** **a** Original image. **b** Ground truth. **c** *L* color channel. **d** *L* channel feature map. **e** *a* color channel. **f** *a* channel feature map. **g** *b* color channel. **h** *b* channel feature map

suppressed, and that the salient region has become more uniform. It is noted that enhancement in the saliency map has been achieved while still preserving the boundary between the salient and non-salient regions.

The impact of each step is also evaluated quantitatively on the images of the MSRA-1000 dataset. The binary saliency maps are obtained by applying the adaptive threshold. Table 1 gives the precision, recall, F-measure and MAE values after successively applying



**Fig. 8** The maps obtained after the application of the proposed method. **a** Average of the channel feature maps. **b** Weighted linear combination of the channel feature maps. **c** Refined map by considering centers of gravity. **d** The final saliency map after bilateral filtering

each step of the proposed scheme. From this table it is seen that by applying the proposed linear combination to the three channel feature maps, the precision recall and F-measure values are increased by 5.8%, 1% and 4.7% respectively, and the MAE value is decreased by 1.1%. It is also seen from this table that the applications of the succeeding steps of the proposed scheme help in further improving the values of all the metrics except the recall value.

### 3.2 Results of applying the proposed scheme to images of four datasets and comparison with the existing methods

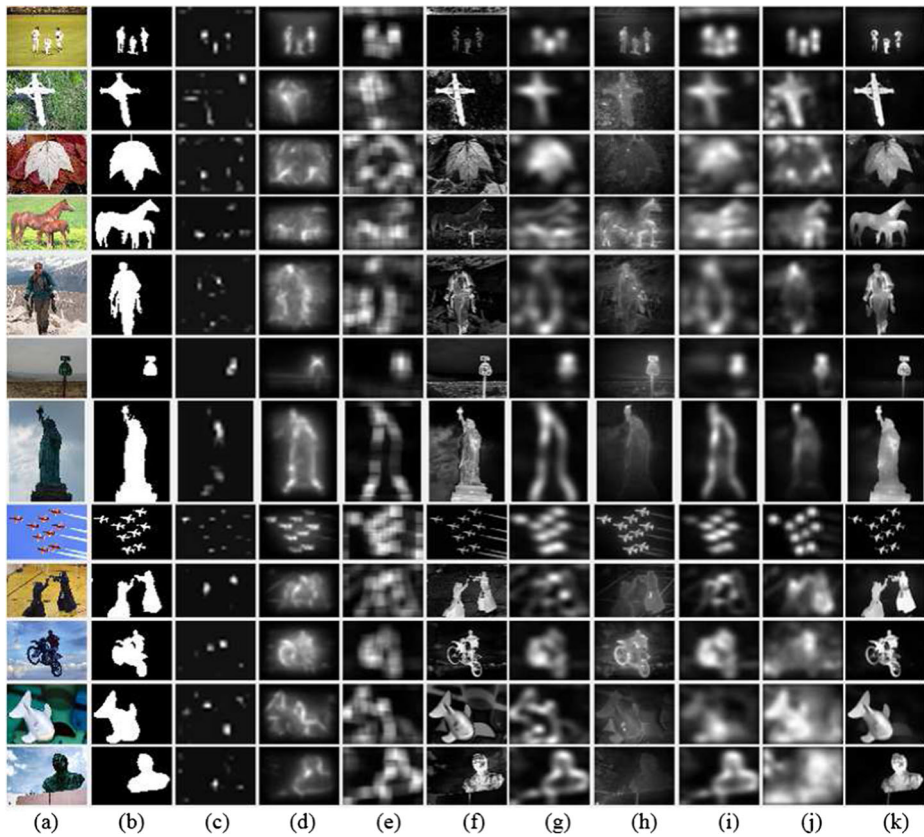
Performance of the proposed salient region detection method is evaluated both subjectively and objectively and compared to that of the six other frequency domain saliency detection methods, namely, superpixel-based wavelet method (SW) [2], weighted quaternion-based method (WQ) [1], wavelet-based method (WAV) [18], hyper-complex Fourier-base method (HFT) [22], frequency-tuned method (FTU) [3], spectral residual method (SR) [16], and also two of the most cited saliency detection methods, namely, graph-based visual saliency method (GBVS) [15] and Itti's method (IT) [19].

Figure 9 shows the saliency maps obtained by utilizing the proposed method as well as the other methods for three sample images from each of the datasets, MSRA-1000, CSSD, CMU and MSRA-10K. It is seen from this figure that the saliency maps obtained using the proposed method are more similar to the ground truth in comparison to the other methods. Also, in the saliency maps obtained, the salient regions are uniformly highlighted with a sharp boundary. It is seen from Fig. 9 that some other methods are not able to detect the entire salient object. For instance, IT [19] detects only parts of the salient object and SR [16] detects edges of the salient object. The other method, GBVS [15], highlights a large part of the image as the salient object which is not accurate. In the saliency maps obtained by FTU [3], the non-salient regions are not clean. Some other methods such as HFT [22] are not successful in detecting salient regions of large size (see sample images in rows 7,

**Table 1** Precision, Recall, F-measure, and MAE values of the main phases of the proposed method

	Precision	Recall	F-measure	MAE
Average of the Channel Feature Maps	0.8099	0.7127	0.7852	0.090763
Weighted Linear Combination	0.8684	0.7227	0.8298	0.079742
Distance to Centers of Gravity	0.8832	0.7227	0.8402	0.079043
Bilateral Filtering	0.8992	0.7147	0.8486	0.077597





**Fig. 9** Saliency maps obtained by applying the proposed and other methods on three images from MSRA-1000, CSSD, CMU and MSRA-10K datasets. **a** Original image. **b** Ground truth. **c** Itti's method (ITI) [19]. **d** Graph-based visual saliency method (GBVS) [15]. **e** Spectral residual method (SR) [16]. **f** Frequency-tuned method (FTU) [3]. **g** Hyper-complex Fourier-base method (HFT) [22]. **h** Wavelet-based method (WAV) [18]. **i** Weighted quaternion-based method (WQ) [1]. **j** Superpixel-based wavelet method (SW) [2]. **k** Proposed method

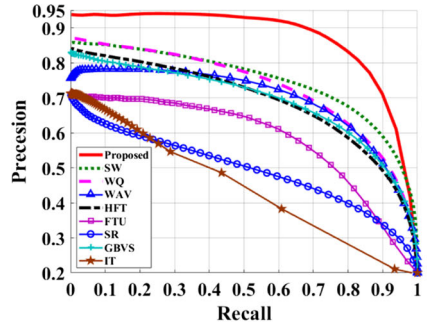
11 and 12 of Fig. 9). The saliency maps obtained by the method WAV [18] are very blurred and the salient region is not detected accurately.

Figure 10 depicts the average precision-recall curves obtained by applying the proposed and the other eight schemes with the fixed thresholds. It is seen from this figure that the proposed scheme outperforms all the other schemes in terms of the precision-recall performance. For the entire range of the fixed thresholds on each of the four datasets, the proposed method obtains the largest values of precision and recall compared to the other methods. There are two other wavelet-based methods amongst the methods considered for comparison, namely SW [2] and WAV [18], which have smaller precision and recall values compared to those of the proposed method.

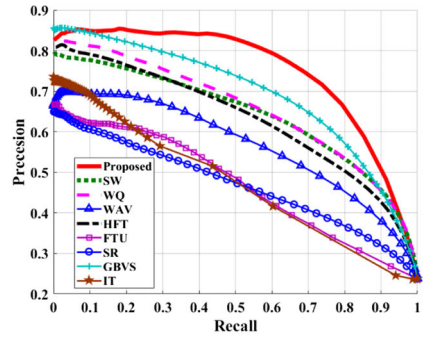
Figure 11 shows the average precision, recall and F-measure values obtained using the adaptive threshold given by (12). It is seen from this figure that the proposed method provides the largest values for the precision metric amongst all the methods regardless of the datasets on which the methods are applied. In terms of the recall metric, the proposed



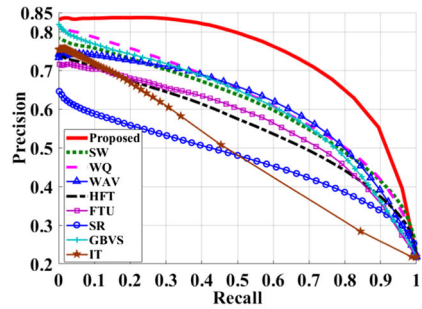
**Fig. 10** Precision-recall curves obtained by applying the proposed and other saliency detection methods on three datasets. **a** MSRA-1000 dataset. **b** CSSD dataset. **c** CMU dataset. **d** MSRA-10K dataset



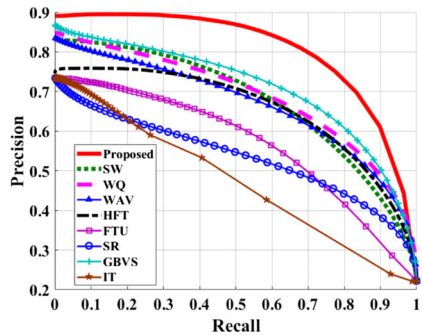
(a)



(b)



(c)



(d)

scheme is second only to GBVS scheme in the cases of the CSSD, CMU and MSRA-10K datasets. However, the GBVS scheme has provided larger values for the recall metric in comparison to that of the proposed method at the expense of smaller values for the precision metric. Since, both of the precision and recall values are considered in computing the F-measure metric, the overall performance of a saliency detection method can be evaluated using the F-measure metric. It is seen from Fig. 11 that, for all the four datasets, the proposed method provides the largest value for the F-measure metric amongst all the methods including the GBVS method.

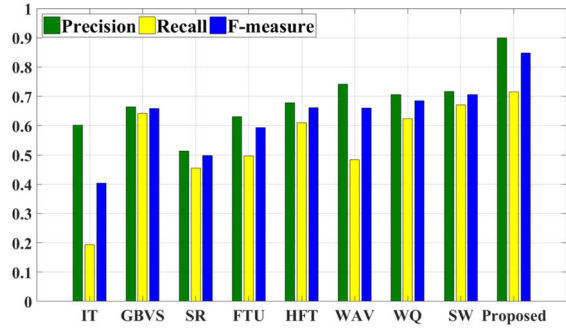
The MAE values obtained using the proposed method as well as by using the other methods are depicted in Fig. 12. It is seen from this figure that the proposed method provides the lowest value for the MAE metric amongst all the methods, irrespective of the datasets used in our experiments, indicating a strong similarity between the saliency maps obtained by applying the proposed method and the ground truth.

As seen from Fig. 9, the saliency maps obtained using the proposed method are more similar to the ground truth in comparison to those obtained using the other methods, thus indicating the superiority of the proposed method. In addition, as seen from Figs. 10, 11 and 12, the proposed saliency detection method outperforms the other existing methods in terms of precision-recall performance for the fixed thresholds, F-measure values for the adaptive threshold, and MAE values. This performance improvement can be attributed to the incorporation of the wavelet-based textural feature maps that are able to represent the saliency-related information of each color channel, and their efficient linear combination using the proposed weighting scheme.

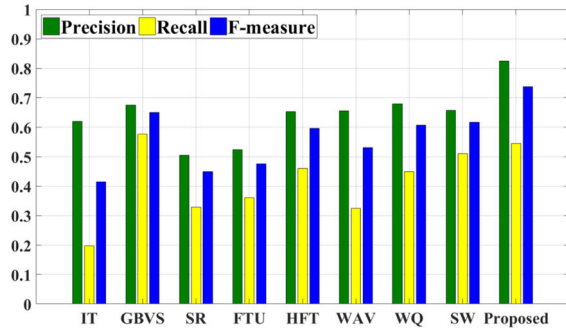
Using features that cannot suitably distinguish the salient regions from the non-salient ones, such as the spectral residual in the SR method [16], has resulted in inaccurate saliency maps (as seen from Fig. 9e rows 3, 5 and 7, for some test images), small precision and recall values (as seen from Figs. 10 and 11), and large MAE values (as seen from Fig. 12). In the FTU method [3], the use of only the color components as features, while ignoring other effective features such as textures, has led to the detection of the non-salient regions wrongly as the salient regions (Fig. 9f rows 3, 5-7, 9, and 11). In the HFT method [22], the features have been extracted using the hyper-complex Fourier transform, which is more suitable for extracting global irregularities. As a result, this method has detected only the borders of the salient region and has failed in detecting the entire salient object in images with large salient regions (see rows 7, 11, and 12 of Fig. 9). In the SW [2] and the WAV [18] methods, the wavelet-based features have been used. The wavelet transform is particularly suitable in representing the image details. However, since all the details employed in these methods are not related to the salient regions, they have failed in detecting the salient regions in some of the test images (see Fig. 9h and j, rows 3, 11, and 12). As seen from Figs. 10, 11 and 12, these two methods have yielded smaller precision and recall values and larger MAE values. In the DWT decomposition, the first decomposition levels extract edges rather than textures, while the coarsest decomposition level consists of all the saliency-related textural details of the image. In the proposed method, the extraction of a feature map only after the wavelet decomposition at the coarsest level has resulted in more accurate saliency maps, larger precision and recall values, and smaller MAE values as seen from Figs. 9k, 10, 11 and 12, respectively.

The way the extracted features are utilized has a significant impact on the saliency detection. In the HFT method [22], the maps have been generated at different levels but only one of them has been selected as the final saliency map, resulting in not necessarily an accurate detection of the salient objects (see Fig. 9g, rows 4 and 9). In the WQ method [1], using a set of pre-specified weights to linearly combine the features has led to an inaccurate detection

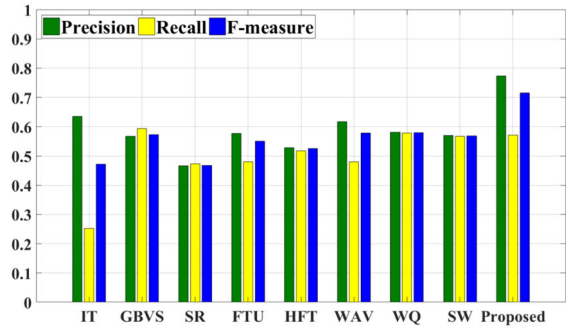
**Fig. 11** Precision, recall, and F-measure obtained by applying the proposed and other saliency detection methods on three datasets. **a** MSRA-1000 dataset. **b** CSSD dataset. **c** CMU dataset. **d** MSRA-10K dataset



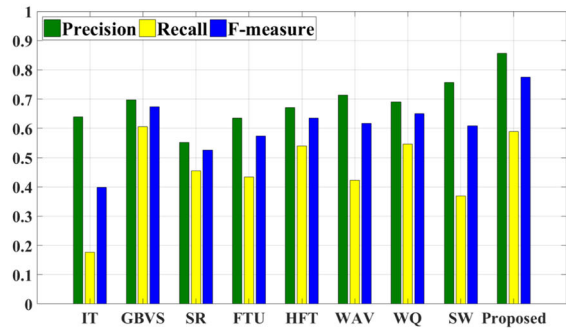
(a)



(b)

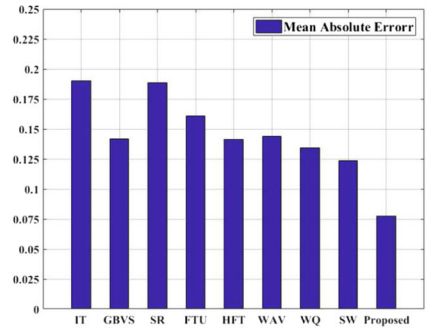


(c)

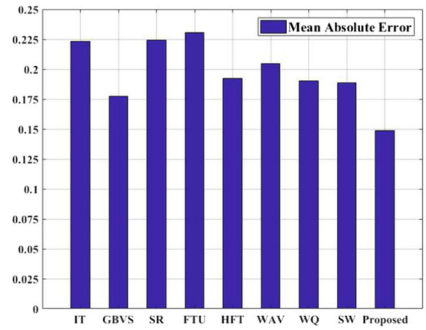


(d)

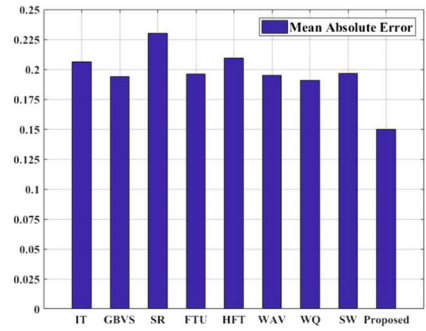
**Fig. 12** MAE obtained by applying the proposed and other saliency detection methods on three datasets. **a** MSRA-1000 dataset. **b** CSSD dataset. **c** CMU dataset. **d** MSRA-10K dataset



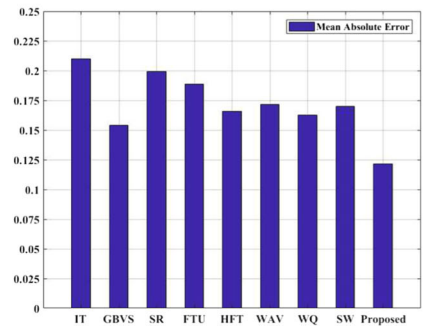
(a)



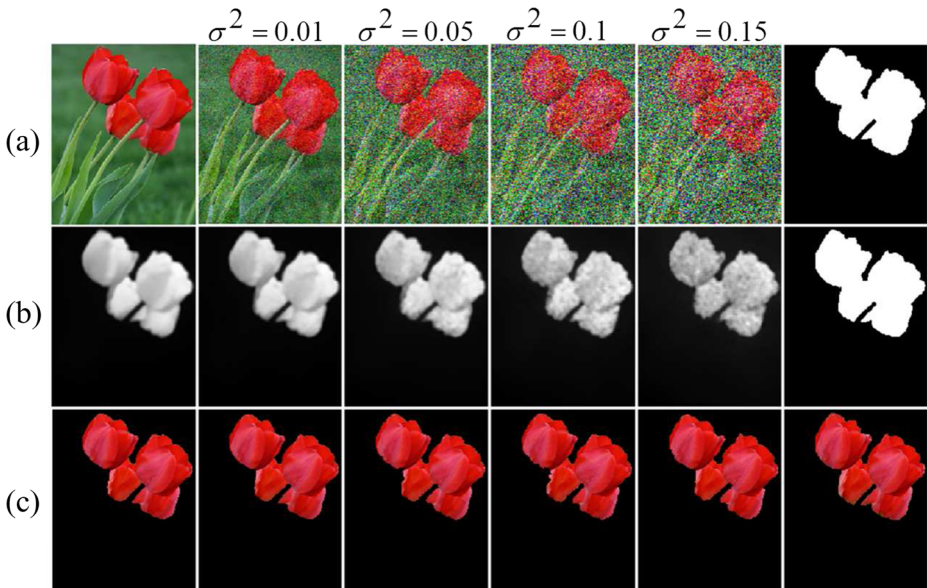
(b)



(c)



(d)

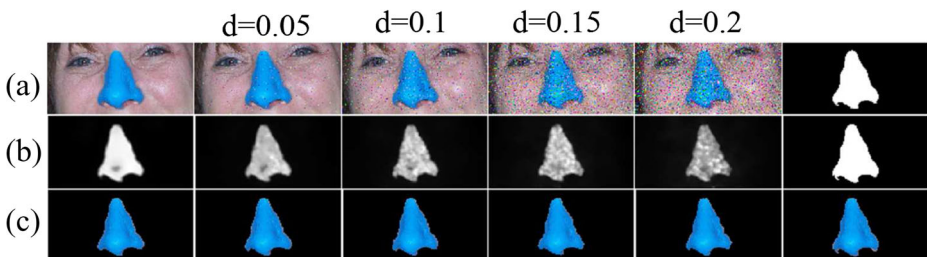


**Fig. 13** Saliency maps obtained by applying the proposed scheme on a sample image from the MSRA-1000 dataset and segmented salient objects. **a** Left to right: the original image, images with various levels of Gaussian noise and the ground truth. **b** Corresponding saliency maps. **c** Corresponding segmented salient objects

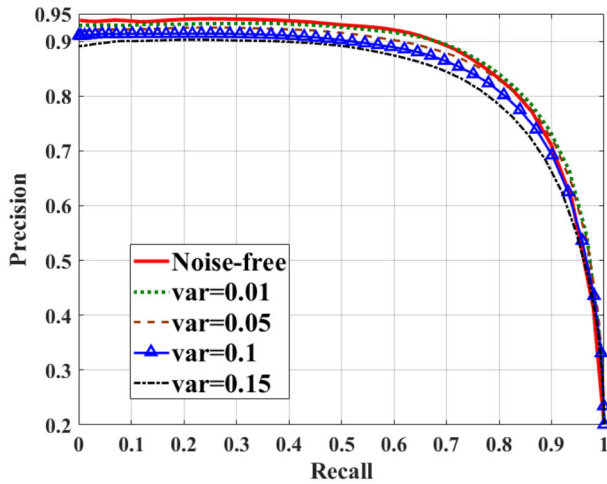
of the salient regions in some images (see Fig. 9i, rows 3-5). In the proposed method, the channel feature maps have been combined by applying a weighting scheme, in which a larger weight is assigned to a channel feature map that can represent the salient region more efficiently.

### 3.3 Evaluating the performance of the proposed method against noise

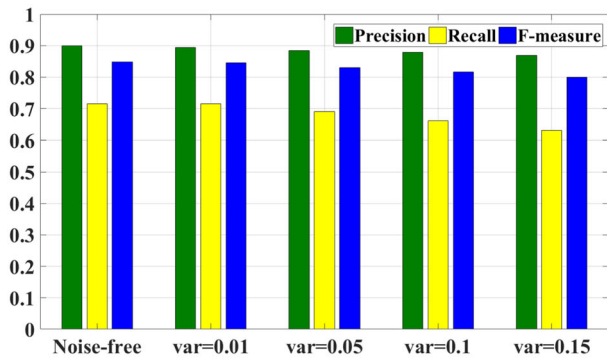
In the previous sections, the input image was assumed to be noise-free. Since, noise is inevitable in real situations, it is worth investigating the robustness of the proposed method against noise. In this section, performance of the proposed method is evaluated against



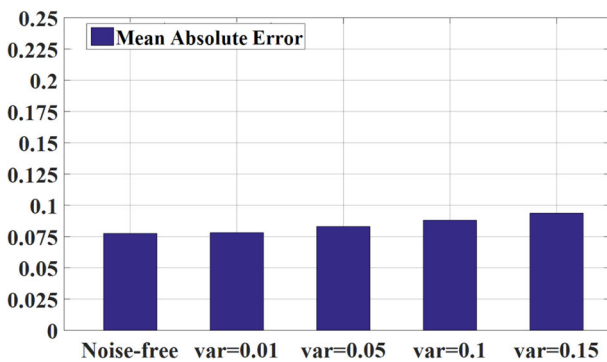
**Fig. 14** Saliency maps obtained by applying the proposed scheme on a sample image from the MSRA-1000 dataset and segmented salient objects. **a** Left to right: the original image, images with various levels of impulsive noise and the ground truth. **b** Corresponding saliency maps. **c** Corresponding segmented salient objects



(a)

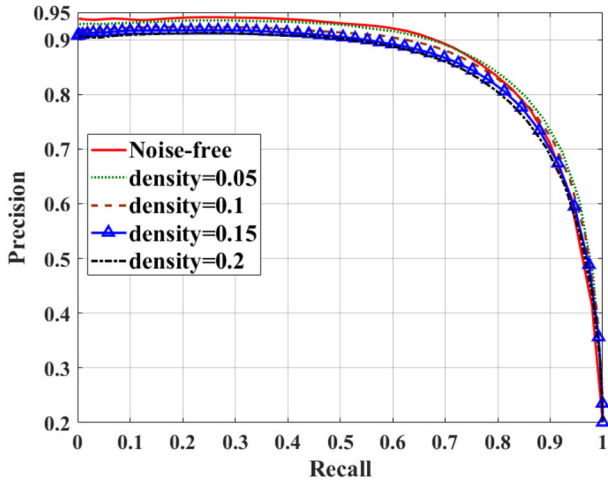


(b)

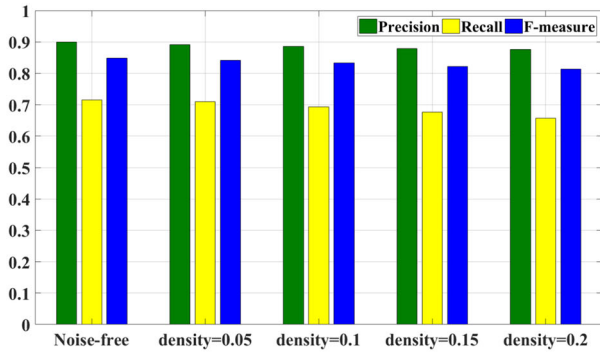


(c)

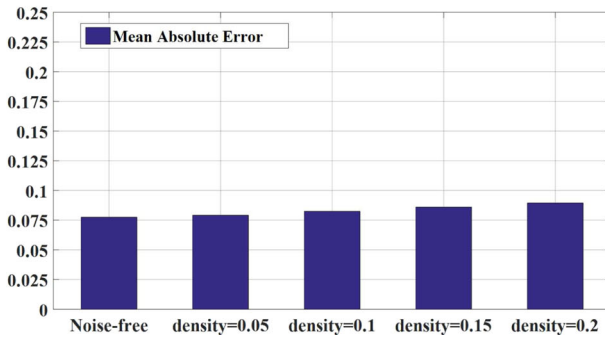
**Fig. 15** Results obtained by applying the proposed saliency detection method on images of MSRA-1000 dataset with additive white Gaussian noise with variance values of 0.01, 0.05, 0.1 and 0.15. **a** Precision-recall curves. **b** Precision, recall and F-measure values. **c** MAE values



(a)



(b)

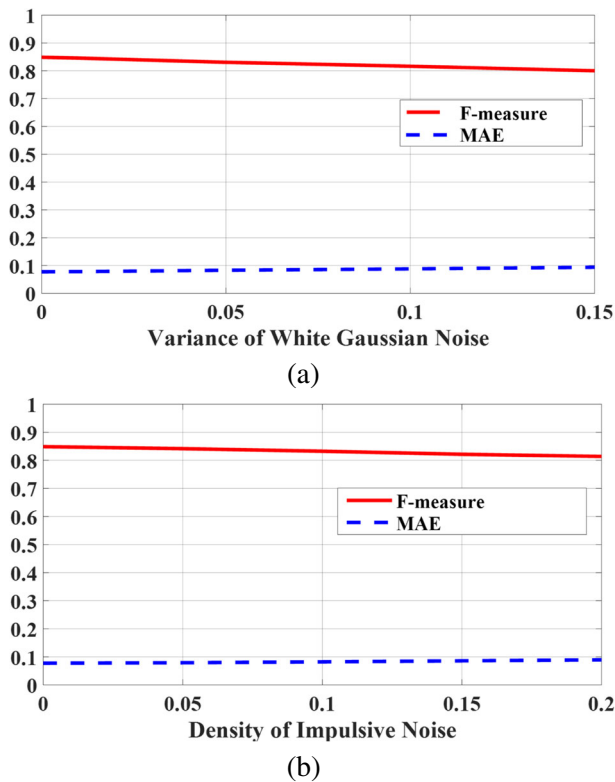


(c)

**Fig. 16** Results obtained by applying the proposed saliency detection method on images of MSRA-1000 dataset with impulsive noise with density values of 0.05, 0.1, 0.15 and 0.2. **a** Precision-recall curves. **b** Precision, recall and F-measure values. **c** MAE values

additive white Gaussian noise and impulsive salt and pepper noise of different levels. The experiments are conducted on images from the MSRA-1000 dataset.





**Fig. 17** F-measure and MAE values for different levels of **a** additive white Gaussian noise **b** impulsive noise

First, Gaussian noise with the variance values of 0.01, 0.05, 0.1 and 0.15 are added to a typical image from the dataset and results of applying the proposed scheme are depicted in Fig. 13. The images in the first row of this figure are the original image, noise corrupted images and the ground truth. Images in the second and third rows are, respectively, the corresponding saliency maps obtained by applying the proposed scheme and the segmented objects. Next, the salt and pepper noise with density values of 0.05, 0.1, 0.15 and 0.2 are applied to another sample image from the dataset. Figure 14 shows the results of applying the proposed method on the original and noisy images. It is seen from Figs. 13 and 14 that the salient object is successfully detected under various levels of noise. Thus, the proposed scheme can be regarded to be generally robust against Gaussian and impulsive noise.

Plots of the precision-recall curves, and the values of the metrics precision, recall, F-measure and MAE obtained by applying the proposed scheme to each of the images of the MSRA-1000 dataset corrupted by the various levels of the Gaussian and impulsive noise are shown in Figs. 15 and 16, respectively. It is seen from these figures that the precision, recall and F-measure values decrease only moderately and the MAE values increase slightly, as the noise level is increased.

To evaluate the robustness of the proposed method further, the F-measure and MAE values for different levels of Gaussian and impulsive noise are depicted in Fig. 17a and b, respectively. It is again seen from this figure that despite the high levels of these two types of noise, performance of the proposed method deteriorates only slightly.

## 4 Conclusion

In this paper, a saliency detection method has been proposed by using a new weighted linear combination of the wavelet-based feature maps. The textural features of the image have been extracted using the wavelet coefficients of the three color channels, and an effective feature map fusion scheme based on the concept of entropy and border avoidance criterion has been proposed. In order to take into consideration both the spatial and intensity information of the pixels, in this scheme, the entropy value of the low-pass filtered map has been utilized. The map thus obtained has been further refined based on the image centers of gravity. Finally, unlike most of the existing methods, a bilateral filter has been applied to the resulting map in order to smooth it while preserving the sharp boundaries between salient and non-salient regions.

The contributions of the work carried out in this paper are as follows: i) The channel feature maps containing the textural features, have been constructed only after the wavelet decomposition at the coarsest level. ii) A new weighting scheme to linearly combine the channel feature maps have been proposed. The idea in this scheme has been to linearly combine the maps of the individual channel features in order to construct a final feature map that can best distinguish the salient regions from the non-salient ones. An image dependent scheme for the linear combination in which the weights to the individual channel feature maps have been assigned based on their capability to distinguish the two regions of the image has been developed. The weights of the linear combination of the three feature maps have been determined using the concept of entropy and a criterion of border avoidance.

Several experiments have been carried out by applying the proposed scheme on the images from several datasets in order to evaluate its performance and to compare it to other existing methods. It has been shown that the saliency maps obtained using the proposed method is more similar to the salient regions detected by HVS. The proposed method provides the values of precision, recall and F-measure higher than those provided by the other methods. In addition, performance of the proposed method in the presence of different levels of Gaussian and impulsive noise has also been studied. Experimental results have shown that the proposed method is able to detect the salient object in images with various levels of noise. Thus, the proposed saliency detection scheme can be regarded to be noise robust.

**Acknowledgements** This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and in part by the Regroupement Strategique en Microelectronique du Quebec (ReSMiQ).

## References

1. Abkenar MR, Ahmad MO (2015) Quaternion-based salient region detection using scale space analysis. In: Proceedings signal processing and intelligent systems conference (SPIS), pp 78–82
2. Abkenar MR, Ahmad MO (2016) Superpixel-based salient region detection using the wavelet transform. In: Proceedings IEEE international symposium on circuits and systems (ISCAS), pp 2719–2722
3. Achanta R, Hemami SS, Estrada FJ, Susstrunk S (2009) Frequency tuned salient region detection. In: Proceedings IEEE conference on computer vision and pattern recognition (CVPR), pp 1597–1604
4. Arya R, Singh N, Agrawal R (2016) A novel hybrid approach for salient object detection using local and global saliency in frequency domain. *Multimedia Tools and Applications* 75(14):8267–8287
5. Avidan S, Shamir A (2007) Seam carving for content-aware image resizing. *ACM Trans Graph* 26(3):10
6. Batra D, Kowdle A, Parikh D, Luo J, Chen T (2010) iCoseg: Interactive Co-segmentation with Intelligent Scribble Guidance. In: Proceedings IEEE conference on computer vision and pattern recognition (CVPR), pp 3169–3176
7. Borji A (2015) Salient object detection: a benchmark. *IEEE Trans Image Process* 24(12):5706–5722

8. Borji A, Itti L (2013) State-of-the-art in visual attention modeling. *IEEE Trans Pattern Anal Mach Intell* 35(1):185–207
9. Borji A, Cheng M, Jiang H (2014) Salient object detection: A survey. arXiv:1411.5878
10. Bruce N, Tsotsos J (2005) Saliency based on information maximization. *Adv Neural Inf Proces Syst* 18:155–162
11. Cheng, Mitra N, Huang X, Torr P, Hu S (2015) Global contrast based salient region detection. *IEEE Trans Pattern Anal Mach Intell* 37(3):569–582
12. Fugal D (2009) Conceptual wavelets in digital signal processing: an in-depth practical approach for the non-mathematician. Space and Signals Technical Publishing, San Diego
13. Goferman S, Zelnik-Manor L, Tal A (2012) Context-aware saliency detection. *IEEE Trans Pattern Anal Mach Intell* 34(10):1915–1926
14. Guo C, Ma Q, Zhang L (2008) Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform. In: Proceedings IEEE conference on computer vision and pattern recognition (CVPR), pp 1–8
15. Harel J, Koch C, Perona P (2006) Graph-based visual saliency. *Adv Neural Inf Proces Syst* 1(2):545–552
16. Hou X, Zhang L (2007) Saliency detection: A spectral residual approach. In: Proceedings IEEE conference on computer vision and pattern recognition (CVPR), pp 1–8
17. Itti L (2004) Automatic foveation for video compression using a neurobiological model of visual attention. *IEEE Trans Image Process* 13(10):1304–1318
18. Imamoglu N, Lin W, Fang Y (2013) A saliency detection model using low-level features based on wavelet transform. *IEEE Trans Multimedia* 15(1):96–105
19. 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
20. Ko B, Nam J (2006) Object-of-interest image segmentation based on human attention and semantic region clustering. *J Opt Soc Am A* 23(10):2462–2470
21. Koffka K (1935) Principles of Gestalt psychology. Routledge, London
22. Li J, Levine M, An X, Xu X, He H (2013) Visual saliency based on scale-space analysis in the frequency domain. *IEEE Trans Pattern Anal Mach Intell* 35(4):996–1010
23. Liu S, Hu J (2016) Visual saliency based on frequency domain analysis and spatial information. *Multimedia Tools and Applications* 75(23):16699–16711
24. Liu T, Yuan Z, Sun J, Wang J, Zheng N, Tang X, Shum H (2011) Learning to detect a salient object. *IEEE Trans Pattern Anal Mach Intell* 33(2):353–367
25. Ma YF, Zhang HJ (2003) Contrast-based image attention analysis by using fuzzy growing. In: Proceedings ACM international conference on multimedia, pp 374–381
26. Merry RJ (2005) Wavelet theory and application: a literature study. Eindhoven University of Technology, Eindhoven
27. Murray N, Vanrell M, Otazu X, Parraga CA (2011) Saliency estimation using a non-parametric low-level vision model. In: Proceedings IEEE conference on computer vision and pattern recognition (CVPR), pp 433–440
28. Sun L, Tang Y, Zhang H (2016) Visual saliency detection based on multi-scale and multi-channel mean. *Multimedia Tools and Applications* 75(1):667–684
29. Tian Q, Sebe N, Lew M, Loupas E, Huang T (2001) Image retrieval using wavelet-based salient points. *J Electron Imaging* 10(4):835–849
30. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: Proceedings IEEE conference on computer vision and pattern recognition (CVPR), pp 839–846
31. Walther D, Koch C (2006) Modeling attention to salient proto-objects. *Neural Netw* 19(9):1395–1407
32. Xie X, Zaitsev Y, Velasquez-Garcia L, Teller S, Livermore C (2014) Compact, scalable, high-resolution, MEMS-enabled tactile displays. In: Proceedings solid-state sensors, actuators, and microsystems workshop, pp 127–130
33. Xu X, Mu N, Chen L, Zhang X (2016) Hierarchical salient object detection model using contrast-based saliency and color spatial distribution. *Multimedia Tools and Applications* 75(5):2667–2679
34. Yan Q, Xu L, Shi J, Jia J (2013) Hierarchical saliency detection. In: Proceedings IEEE conference on computer vision and pattern recognition (CVPR), pp 1155–1162
35. Yuan Y, Wang J, Li B, Meng MQ (2015) Saliency based ulcer detection for wireless capsule endoscopy diagnosis. *IEEE Trans Med Imaging* 34(10):2046–2057
36. Zhang L, Lin W (2013) Selective visual attention: computational models and applications. Wiley, New York
37. Zhang GX, Cheng MM, Hu SM, Martin RR (2009) A shape-preserving approach to image resizing. *Comput Graphics Forum* 28(7):1897–1906. Blackwell, New York
38. Zhou C, Liu C (2015) An efficient segmentation method using saliency object detection. *Multimedia Tools and Applications* 74(15):5623–5634



**Masoumeh Rezaei Abkenar** received the B.Sc. degree from K. N. Toosi University of Technology (KNTU), Tehran, Iran, in 2006 and the M.Sc. degree from Semnan University (SU), Semnan, Iran, in 2010, both in electrical engineering. She is currently working toward the Ph.D. degree in electrical and computer engineering at Concordia University, Montreal, QC, Canada. She has been a Research Associate in Digital Signal Processing Lab at Concordia University, since January 2014. Her research area includes image processing and computer vision with applications in salient object detection, visual fixation detection, feature extraction and image segmentation. She has served as a reviewer for several journals and major conferences.



**M. Omair Ahmad** received the B.Eng. degree from Sir George Williams University, Montreal, QC, Canada, and the Ph.D. degree from Concordia University, Montreal, QC, Canada, both in electrical engineering. From 1978 to 1979, he was a Faculty Member with the New York University College, Buffalo, NY, USA. In September 1979, he joined the Faculty of Concordia University as an Assistant Professor of computer science. He joined the Department of Electrical and Computer Engineering, Concordia University, where he was the Chair with the department from June 2002 to May 2005 and is currently a Professor. He holds the Concordia University Research Chair (Tier I) in Multimedia Signal Processing. He has published extensively in the area of signal processing and holds four patents. His current research interests include the areas of multi-dimensional filter design, speech, image and video processing, non-linear signal processing, communication DSP, artificial neural networks, and VLSI circuits for signal processing. He was a Founding Researcher at Micronet from its inception in 1990 as a Canadian Network of Centers of Excellence until its expiration in 2004. Previously, he was an Examiner of the order of Engineers of Quebec. Dr. Ahmad was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS PART I: FUNDAMENTAL THEORY AND APPLICATIONS from June 1999 to December 2001. He was the Local Arrangements Chairman of the 1984 IEEE International Symposium on Circuits and Systems. In 1988, he was a member of the Admission and Advancement Committee of the IEEE. He has served as the Program Co-Chair for the 1995 IEEE International Conference on Neural Networks and Signal Processing, the 2003 IEEE International Conference on Neural Networks and Signal Processing, and the 2004 IEEE International Midwest Symposium on Circuits and Systems. He was a General Co-Chair for the 2008 IEEE International Conference on Neural Networks and Signal Processing. He is the Chair of the Montreal Chapter IEEE Circuits and Systems Society. He is a recipient of numerous honors and awards, including the Wighton Fellowship from the Sandford Fleming Foundation, an induction to Provosts Circle of Distinction for Career Achievements, and the Award of Excellence in Doctoral Supervision from the Faculty of Engineering and Computer Science of Concordia University. Dr. Ahmad is a Fellow of the Institute of Electrical and Electronics Engineers.