

# An Improved Saliency Detection for Different Light Conditions

Yongfeng Ren<sup>1,2</sup>, Jingbo Zhou<sup>2</sup>, Zhijian Wang<sup>1</sup> and Yunyang Yan<sup>2</sup>

<sup>1</sup>College of Computer and Information, Hohai University,  
Nanjing, 211100, P. R. China  
[e-mail: renyongfeng@hyit.edu.cn]

<sup>2</sup>Faculty of Computer Engineering, Huaiyin Institute of Technology  
Huai'an, 223003, P. R. China

\*Corresponding author: Yongfeng Ren

*Received May 27, 2014; revised January 6, 2015; accepted February 27, 2015;  
published March 31, 2015*

---

## Abstract

In this paper, we propose a novel saliency detection framework based on illumination invariant features to improve the accuracy of the saliency detection under the different light conditions. The proposed algorithm is divided into three steps. First, we extract the illuminant invariant features to reduce the effect of the illumination based on the local sensitive histograms. Second, a preliminary saliency map is obtained in the CIE Lab color space. Last, we use the region growing method to fuse the illuminant invariant features and the preliminary saliency map into a new framework. In addition, we integrate the information of spatial distinctness since the saliency objects are usually compact. The experiments on the benchmark dataset show that the proposed saliency detection framework outperforms the state-of-the-art algorithms in terms of different illuminants in the images.

---

**Keywords:** Saliency detection, Local sensitive histograms, Illumination invariant features, Different light conditions

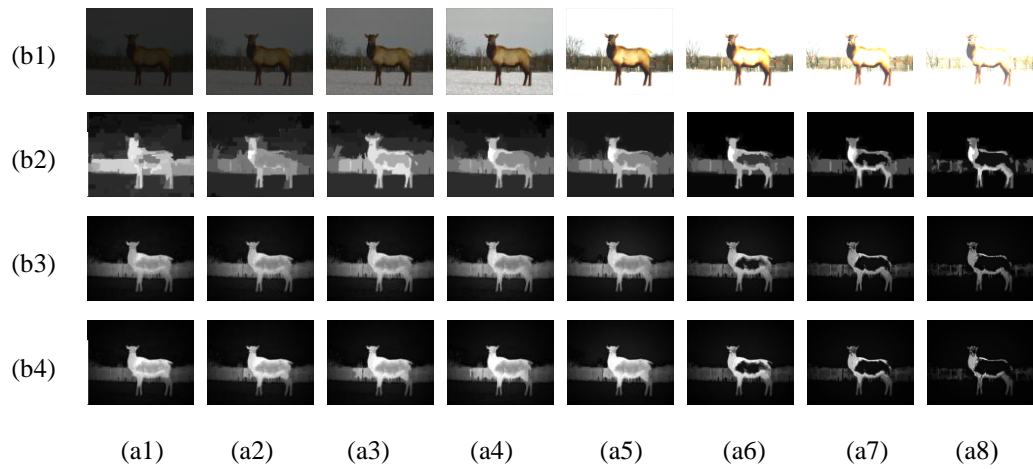
## 1. Introduction

Saliency detection, the task to detect objects attracted by the human visual system in an image/video, has potential applications in computer vision and multimedia tasks, such as object recognition [1], image resizing [2], image retrieval [3], automatic video-to-comics conversion [24], automatic multimedia tagging [25], video accessibility enhancement [26], visual attention [27], and content-aware image editing [4].

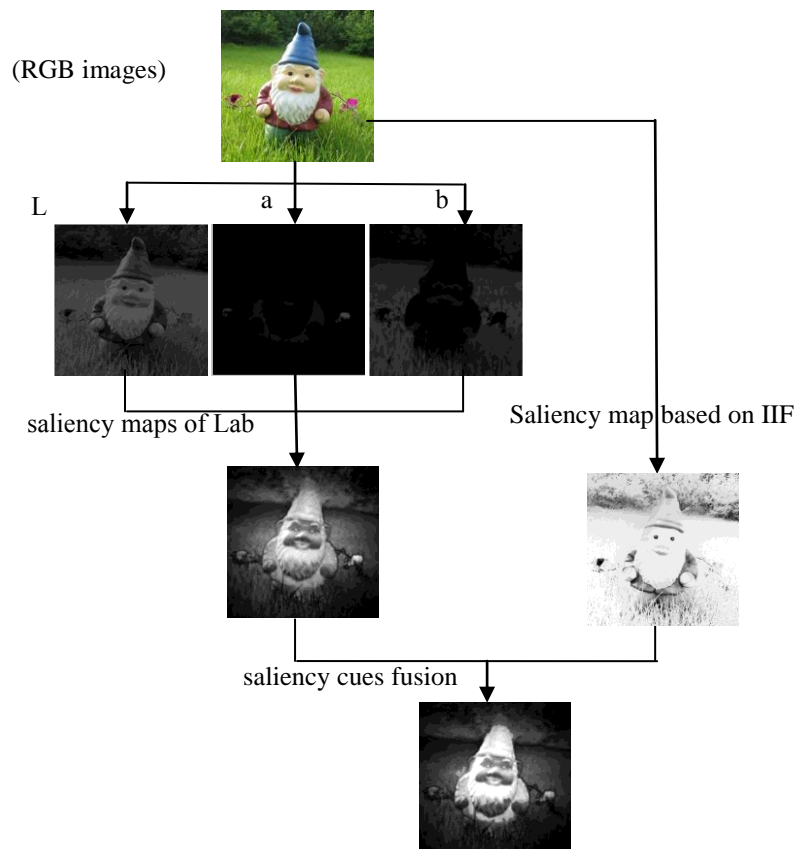
Saliency models have been developed for top-down based and bottom-up based approaches. The former is related to the recognition processing influenced by the prior knowledge such as tasks to be performed, the feature distribution of the target, the context of the visual scene and so on [5-7, 28]. The latter, which is data-driven and task-independent, is approached mainly by three steps, namely, feature selection, saliency calculation, and map normalization [9-17]. First, low-level features, such as color, intensity, orientation and motion, are selected as the basic elements for supporting saliency detection. Second, the saliency value for each pixel in an input image is computed according to a predefined model. In the end, saliency maps obtained from different sources are integrated and normalized to get a final result. In this paper, we focus on the bottom-up salient object detection tasks.

There have been several studies related to saliency detection in recent years. Itti et al. [8] defined a salient region by counting the differences in the areas surrounding the central region across multi-scales images. Ma and Zhang [9] used a fuzzy growth model to generate saliency maps. Harel et al. [10] combined the saliency maps of Itti et al. [8] with other feature maps to highlight the distinctive regions of an image. Hou and Zhang [11] constructed a saliency map by extracting the spectral residual of an image in the spectral domain. Achanta et al. [12] determined the salient regions in images using low-level luminance and color features. Recently, Goferman et al. [4] considered both local and global features to highlight the salient objects enhanced by means of visual organization. Moreover, Achanta et al. [13] proposed a frequency-tuned method that defines the pixel saliency based on a frequency-domain analysis and the differences in color from the average image color. Zhai and Shah [14] define the pixel-level saliency by constructing spatial and temporal attention models. Cheng et al. [15] used histogram- and region-based contrast to compute saliency maps with high precision and recall rates.

With the development of computer and algorithm, the effects of saliency detection have been improved tremendously. However, the illumination in the image in practical application is not always as ideal as the ones in the datasets in this research field. Many of the existing saliency detection methods do not attach great importance to the illumination problem. However, detecting salient objects in the same images which with different illumination is always a challenge task. For example, we simulate the different illumination conditions and run the codes of the early works, i.e., RC [15] and SDSP [16], on MSRA 1000 database [13]. Fig. 1. shows the results. Fig. 1 (a4) is the normal illuminant in the aforementioned database. We reduce the light gradually and demonstrate them as Fig. 1 (a1) ~ (a3). Similarly, Fig. 1 (a5) ~ (a8) is showed as bright image. It can be seen that both RC and SDSP, which represented as Fig. 1 (b2) and Fig. 1 (b3) respectively, are influenced by the case of illumination in images. Since the preferable light condition is very difficult to guarantee in application, it is very necessary to improve the accuracy of saliency detection under the worse illumination. In this paper, we propose a novel saliency detection framework based on illumination invariant features to improve the accuracy of the saliency detection under the different light conditions.



**Fig. 1.** Examples of the saliency maps in different illumination conditions.



**Fig. 2.** Diagram of our proposed model

In this paper, our work includes two parts mainly. First, we add the information of different illumination conditions to a publicly available database MSRA1000 provided by Achanta et al. [13] which includes 1000 images. As we known, this database is the largest and the best of the image datasets to research the saliency detection, and has ground truth in the form of accurate human-marked labels for salient regions. Each image in the dataset is changed to different

light condition by reducing or increasing the light gradually. Second, we propose a new algorithm to detect the salient object based on illumination invariant features [18]. The proposed algorithm extracts the illumination invariant features from the given image and uses the region growing method to detect the salient regions. Since the illuminant invariant features are independent of the light and retain the saliency information of the original image, we can exploit this characteristic to improve the precision of saliency detection under different illuminations.

The contributions of this paper are summarized as follows:

(1) We propose a new framework for the issue of reducing the impact of illumination in saliency detection, since it is not discussed before as we know. In our proposed framework, we exploit illumination invariant features (IIF) which extracts the invariant features efficiently and ignores the illumination conditions of the image as shown in Fig. 2;

(2) The proposed algorithm fuses some priors, such as color distinctness and spatial distinctness into a framework to improve the accuracy of the saliency detection. According to the mechanism of the human visual system that warm colors, such as red and yellow, are more pronounced to our eyes than cold ones, such as green and blue, we propose a simple method to model it by analyzing the color space of the image.

The remainder of the paper is organized as follows: we state the foundations in section 2. In Section 3, we demonstrate framework of our saliency detection method in detail. Then, we demonstrate our experimental results based on three public image datasets and compare the results with other state-of-the-art saliency detection methods in Section 4. The final section concludes the paper by summarizing our findings.

## 2. Illumination invariant features

First of all, we introduce the illumination invariant features (IIF) [18] which can help us to find the invariant features in the images ignoring the illumination conditions.

Let  $I_p$  and  $I'_p$  denote the intensity values of pixel  $p$  before and after an affine illumination change. There is a relationship between them:

$$I'_p = A_p(I_p) = a_{1,p}I_p + a_{2,p} \quad (1)$$

where  $a_{1,p}$  and  $a_{2,p}$  are two parameters of the affine transform  $A_p$  at pixel  $p$ . Let  $H_p^s$  denote the histogram computed from a window  $S_p$  centered at the pixel  $p$ , and  $b_p$  denote the bin corresponding to the intensity value  $I_p$ . According to the definition of the histogram, the number of pixels in  $S_p$  whose intensity value resides in  $[b_p - r_p, b_p + r_p]$  is:

$$S_p = \sum_{b=b_p-r_p}^{b_p+r_p} H_p^s(b) \quad (2)$$

where parameter  $r_p$  controls the interval of integration at pixel  $p$ . We denote  $r_p = \kappa |I_p - \bar{I}_p|$  where  $\kappa = 0.1$  is a constant,  $\bar{I}_p = \frac{1}{|S_p|} \sum_{q \in S_p} I_p$  is the mean intensity value of window  $S_p$  and  $|S_p|$  is the number of pixels in  $S_p$ . With an additional assumption that the affine illumination changes are locally smooth so that the affine

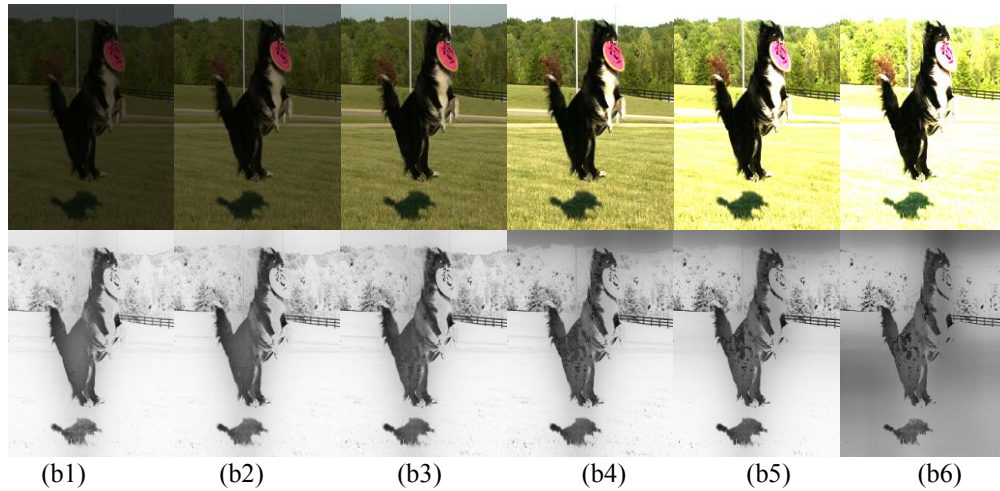
transform is the same for all pixels inside window  $S_p$ . The interval of integration at pixel  $p$  can be expressed as follows:

$$\begin{aligned}
 r'_p &= \kappa \mid I'_p - \overline{I'_p} \mid \\
 &= \kappa \mid a_{1,p} + a_{2,p} - \frac{1}{|S_p|} \sum_{q \in S_p} (a_{1,p} I_q + a_{2,p}) \mid \\
 &= a_{1,p} \kappa \mid I_p - \overline{I_p} \mid \\
 &= a_{1,p} r_p
 \end{aligned} \tag{3}$$

The integrated value  $r'_p$  obtained under a different illumination condition corresponds to the number of pixels with intensity value resides in

$$\begin{aligned}
 &[b'_p - r'_p, b'_p + r'_p] \\
 &= [a_{1,p} b_p + a_{2,p} - a_{2,p} r_p, a_{1,p} b_p + a_{2,p} + a_{2,p} r_p] \\
 &= [a_{1,p} (b_p - r_p) + a_{2,p}, a_{1,p} (b_p + r_p) + a_{2,p}] \\
 &= [A_p (b_p - r_p), A_p (b_p + r_p)]
 \end{aligned} \tag{4}$$

Ignoring the quantization error  $S_p$  is equal to  $S'_p$ . Thus  $S_p$  is independent of affine illumination changes and can be used as a matching invariant under different illumination conditions.



**Fig. 3.** Examples of extracting the IIF. The images on the top are shown under different illumination conditions and their transitional image of their illumination invariant features are shown below them.

In practice, it is inaccurate to define an exact local window inside which the affine illumination transform remains unchanged. Hence, we replace histogram  $H_p^S$  in equation (2) with the locality sensitive histogram  $H_p^E$ , which adaptively is taken into account the contribution from all image pixels [18]. In addition, we use a “soft” interval to reduce the

quantization error. When  $\overline{I_p} = \sum_{b=1}^B H_p^E(b)$ . In experiments,  $a_{2,p}$  is relatively small.

Thus  $I_p$  can be replaced by  $\kappa I_p$ .  $S_p$  becomes:

$$S_p = \sum_{b=1}^B \exp\left(-\frac{(b - b_p)^2}{2 \max(\kappa, I_p)^2}\right) H_p^E(b) \quad (5)$$

The examples of extracting the illumination invariant features are shown in Fig. 3 (b1) ~ (b5). Since we ignore  $a_{2,p}$  when  $S_p$  is computed, the illumination invariant features will have many errors when the light is changed greatly as shown in Fig. 3 (b6). However, the light condition in Fig. 3 (b6) rarely happens in practice. Therefore, the method can be used to improve saliency detection for different light conditions.

### 3. Proposed saliency framework

Based on the discussion in section 2, we propose the novel algorithm to improve saliency detection for different light conditions. In the proposed algorithm, we first compute the salience maps by some priors, such as color distinctness and spatial distinctness [8]. Then, we compute the saliency map based on illumination invariant features by the method which is introduced in section 2. Lastly, we obtain the saliency maps by fusing the information generated by the first and the second step.

In this proposed saliency framework, the color distinctness is considered first. Some studies [19] conclude from daily experiences that warm colors, such as red and yellow, are more pronounced to the human visual system than cold ones, such as green and blue. In this paper, we propose a simple yet effective method to model this prior.

As stated before, we convert RGB color space to CIE Lab color space. Lab color space is an opponent color system, in which a-channel represents green-red information while b-channel represents blue-yellow information. If a pixel has a smaller (greater) a value, it would seem greenish (reddish). With the same manner, if a pixel has a smaller (greater) b value, it would seem bluish (yellowish). Hence, if a pixel has a higher a or b value, it would seem “warmer”; otherwise, it would seem “colder”.

Based on the aforementioned analysis, we devise a metric to evaluate the “color saliency” for a given pixel. At first, we perform linear mappings  $f_a(x) \mapsto f_{an}(x) \in [0,1]$  and  $f_b(x) \mapsto f_{bn}(x) \in [0,1]$  by

$$f_{an}(x) = \frac{f_a(x) - \min a}{\max a - \min a} \quad (6)$$

$$f_{bn}(x) = \frac{f_b(x) - \min b}{\max b - \min b} \quad (7)$$

where  $\min a$  ( $\max a$ ) is the minimum (maximum) value of  $\{f_a(x) : x \in \Omega\}$  and  $\min b$  ( $\max b$ ) is the minimum (maximum) value of  $\{f_b(x) : x \in \Omega\}$ . Thus, each pixel  $x$  can be mapped to one point in the color plane  $(f_{an}, f_{bn}) \in [0,1] \times [0,1]$ . Intuitively, in this color plane, the point  $(f_{an}, f_{bn})$  is the “coldest” point and thus it is the “least salient” one. Therefore, we define the color saliency of a point  $x$  in a straightforward manner as



$$S_c(x) = 1 - \exp\left(-\frac{f_{an}^2(x) + f_{bn}^2(x)}{\sigma_c^2}\right) \quad (8)$$

where  $\sigma_c$  is a parameter.

After computer the color distinctness, a general area of the saliency map in the current image can be got. But the result of this step is less accurate for lack of the information of spatial distinctness. In some studies, a conclusion that objects near the image center is more attractive to human than others [20] have been concluded. That implies spatial near the center of the image will be more likely to be “salient” than the ones far away from the center. Therefore we generate a prior map using a Gaussian distribution based on the distances of the pixels to the image center in this paper, such as

$$S_I(x) = \exp(-d(x, c) / \sigma_I^2) \quad (9)$$

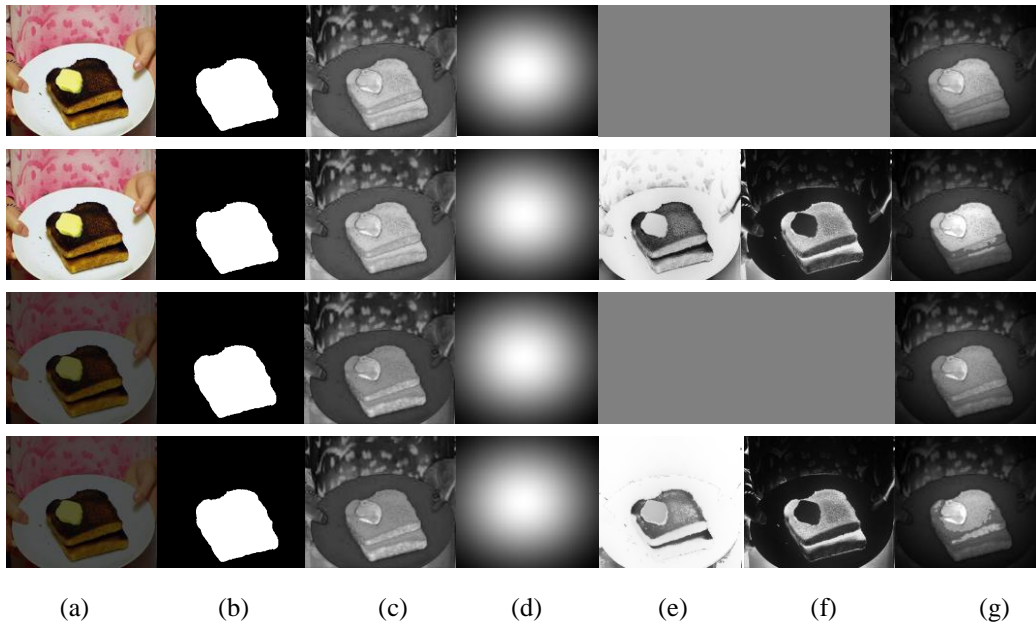
where  $\sigma_I$  is a parameter,  $d(x, c)$  is the distance between point  $x$  and  $c$ .

According to the above methods, three results will be obtained.  $S_p$  is the illumination invariant features,  $S_c$  is the color saliency, and  $S_I$  is the spatial saliency. To combine the saliency maps generated from different cues, we exploit the Bayesian method similar to [21]. Let  $p(x_f \mid S_j)$  be saliency map such ones as side cues, the fusion map would be:

$$p(x_f = 1 \mid S_p, S_c, S_I) \propto \frac{1}{Z} \prod p(x_f \mid S_j), j = P, C, I \quad (10)$$

where  $Z$  is chosen in a way that the final map is a probability density function (pdf).

At last the region growing algorithm [23] is used to find the salient image regions based on the pixels of the outline in the transitional image of the illumination invariant features.



**Fig. 4.** Examples of saliency maps based on illumination invariant features.

An example of saliency maps based on illumination invariant features is shown in Fig. 4. The first and second rows are the images with normal illumination. The third and fourth are the images with 0.8 times of the illumination compared with the first and second rows. Fig. 4 (a) are the images in database of the MSRA 1000, and Fig. 4 (b) are their binary ground truth. We compute the color and spatial distinctness which is shown in Fig. 4(c) and Fig. 4 (d) respectively. The illumination invariant features mentioned in section 2 is shown in Fig. 4 (e) and accordingly saliency map is shown in Fig. 4 (f). Last, we fuse the color distinctness, spatial distinctness and the saliency map based on IIF into a final saliency map which is shown in Fig. 4 (g). From the comparison, we can see that the saliency map which added the information of illumination invariant features is more accurate than that without this information.

## 4. Experiments and analysis

In order to verify the proposed method, we have evaluated the results of our approach on the publicly available database provided by Achanta et al. [13]. As we know, the database is the largest and best of the image datasets to research the saliency detection, and has ground truth in the form of accurate human-marked labels for salient regions. The experiments include two aspects. First, we add the information of illumination to the image database and compare our method with other seven state-of-the-arts saliency detection methods, including FT [13], RC [15], LC [14], HC [15], SDSP [16], SR [11], GBVS [10] on this database. Second, we also evaluate our algorithm and state-of-art saliency detection methods under normal illumination for the purpose of fair comparison. To evaluate the results of saliency maps generated by the methods aforementioned, we use the precision-recall curve and the F-measure similar to [13, 15].

### 4.1 Datasets

Since there is no image database available for us to test the effect of illumination on saliency detection, we first build an image database on the MSRA 1000. For the color space of RGB having not a direct correlation with the information of the light, we transfer the RGB color space to CIE Lab color space. Lab color space is an opponent color system, in which L-channel represents light information. We reduce or increase the value of L-channel, and then we translate the CIE Lab color space to RGB color space.

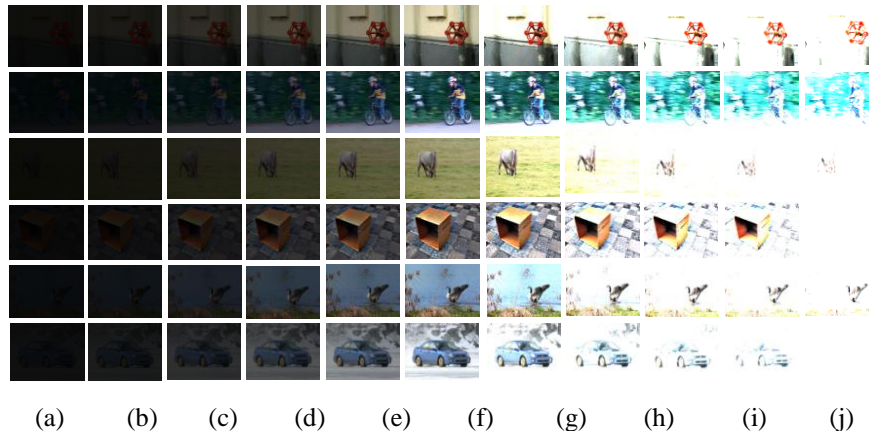


Fig. 5. Examples of images under different illuminant conditions.



Those operations are the simulation to the change of the light from every image into 20 images under different illumination. The Fig. 5 shows the 11 images of the 20 images under different illumination based on an example. The images in the Fig. 5 can be observed by human well. The other 9 images which is under the illuminant conditions as 0 ~ 0.4 or 1.6 ~ 2.0 times as normal ones are too dark or too bright to be observed by human. In Fig. 5, (f) is the images under the normal illuminant condition; (a) ~ (e) is under the illuminant conditions as 0.5~ 0.9 time as normal illumination; and (g) ~ (k) is under the illuminant conditions as 1.1~ 1.5 time as normal illumination. From those images, we can find that the angle of the light remains the same when the light is changing. However, the effect of this method is close to the actual situation to the greatest extent. By modifying the illumination in each image, we can expand the image dataset what is the basement to the experiments in this paper.

## 4.2 Evaluating measure

The results generated by the proposed model are evaluated in two different ways. For the first evaluation, a fixed threshold within [0, 255] is used to construct a binary foreground mask from the saliency map. Then, the binary mask is compared with the ground truth mask to obtain a precision-recall (PR) pair. To compare the proposed model with others, we always use the precision value and recall value to measure different algorithms, for the precision value is the ratio of the correctly detected region over the whole detected region, and recall is calculated as the ratio of correctly detected salient region to the ground-truth salient region. We vary the threshold over its entire range to obtain the PR curve for one image. The average precision-recall curve is obtained by averaging the results from all testing images. For the second evaluation, we follow [13] to segment a saliency map by an adaptive threshold, i.e.,

$$\tau = \frac{2}{H \cdot W} \sum_{x=1}^H \sum_{y=1}^W S(x, y) \quad (11)$$

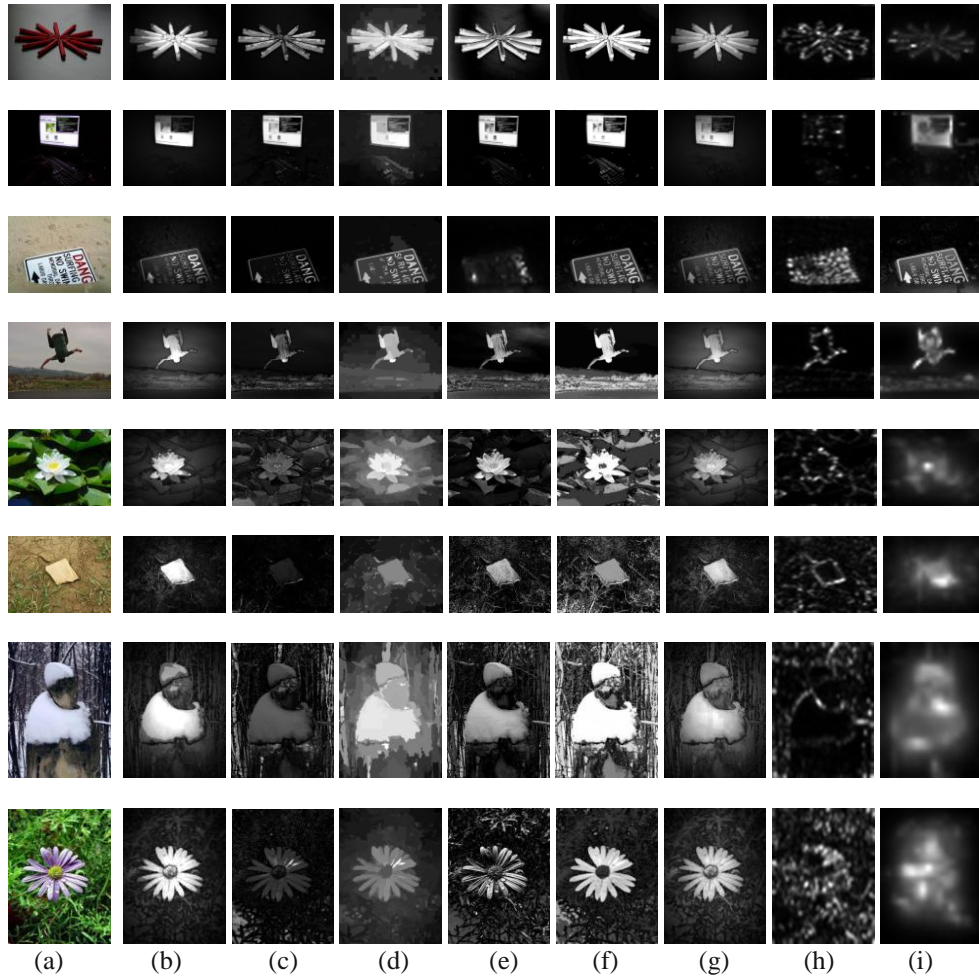
where  $S(x, y)$  is the saliency value in position  $(x, y)$ ,  $H$  and  $W$  represent the high and width respectively. If the saliency value of a pixel is larger than threshold, the pixel is considered as foreground. In many applications, high precision and high recall are both required. We thus estimate the F-Measure [15] as:

$$F_{\beta} = \frac{(1 + \beta^2) \cdot Precision \cdot Recall}{\beta^2 \cdot Precision + Recall} \quad (12)$$

If the saliency of a pixel is larger than the threshold, the pixel is considered as foreground. In many applications, high precision and high recall are both required. We thus estimate the F-Measure as equation (12) where we set  $\beta^2 = 0.3$  to emphasize the precision.

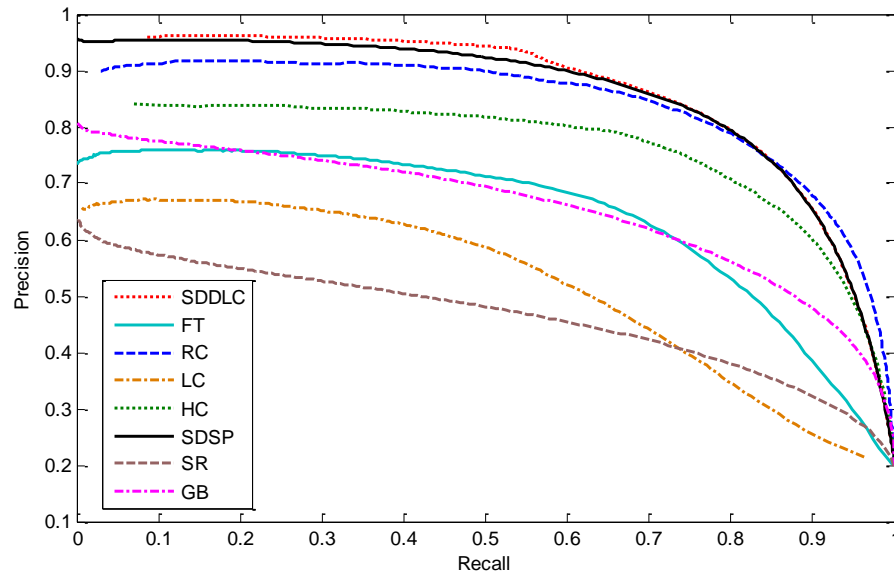
## 4.3 The comparison under normal illuminant condition

In this subsection, we show the experimental results when comparing the proposed method with state-of-the-art saliency detection methods on MSRA 1000.



**Fig. 6.** Examples of saliency maps by different algorithms under normal illuminant condition. (a) input images; (b) SDDL; (c) FT[13]; (d) RC[15]; (e) LC[14]; (f) HC[15]; (g) SDSP[16]; (h) SR[11]; (i) GBVS[10]

Visual comparison of saliency maps obtained by our method SDDL and other algorithms are shown as **Fig. 6**, **Fig. 7** and **Table 1**. In **Fig. 6**, the results show that the saliency maps generated by SDDL, RC [15], HC [15] and SDSP [16], are always better than FT [13], LC [14], SR [11] and GBVS [10]. In **Fig. 7**, the precision and the recall curve of our method SDDL is higher than those of the other methods. It is indicated that the saliency maps computed by our method are smoother and containing more pixels with the saliency threshold 255. At last, we compute the F-measure of different algorithms according to equation (10). The results are shown in **Table 1**. It can be seen that the F-measure of SDDL, RC, HC and SDSP are all higher than 0.7. In addition, the results of FT, LC, SR and GBVS are all less than 0.6.

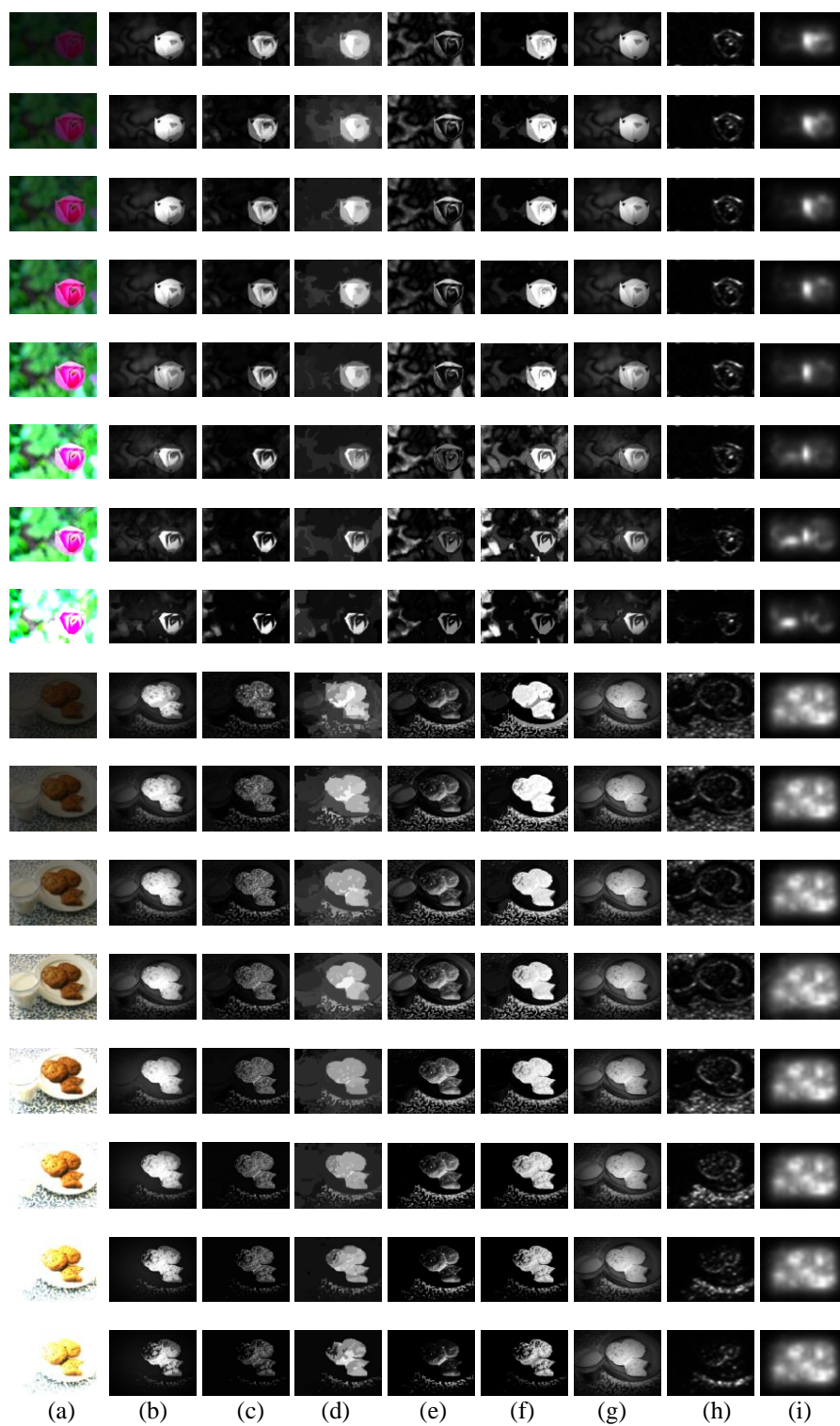


**Fig. 7.** Precision and recall for all algorithms under normal illuminant condition.

**Table 1.** F-measure for each algorithm

Algorithm	SDDL	FT[13]	RC[15]	LC[14]	HC [15]	SDSP [16]	SR [11]	GBVS[10]
F-measure	0.7804	0.3722	0.7631	0.4655	0.7273	0.7516	0.4137	0.5230

The models of FT, LC, SR and GBVS generally detect the foreground from input images. However, they are easy to influence the saliency maps of them by the background that the salient area contains not only salient object but also clutter background. The model of HC gets saliency maps by the contrast between the colors. The information of colors is more than that of the gray image. The results generated by HC showed in [Fig. 6 \(f\)](#) and [Fig. 7](#) are preferable. However, it is influenced by the fine objects since the spatial information is out of consideration. The model of RC showed in [Fig. 6 \(d\)](#) and [Fig. 7](#) is added the information of the size and location of salient area based on the model of HC. Thus its results of saliency map are better than HC. However, similar to FT, it is also influenced by the clutter background. The reason might be that RC compute color contrast based on color histogram to measure the difference between two regions. It is failed when the regions located in the salient object and in background having the same color histogram. From the fifth to eighth rows in [Fig. 6 \(d\)](#), the salient regions detected by RC contain not only the salient object, but also the clutter background.

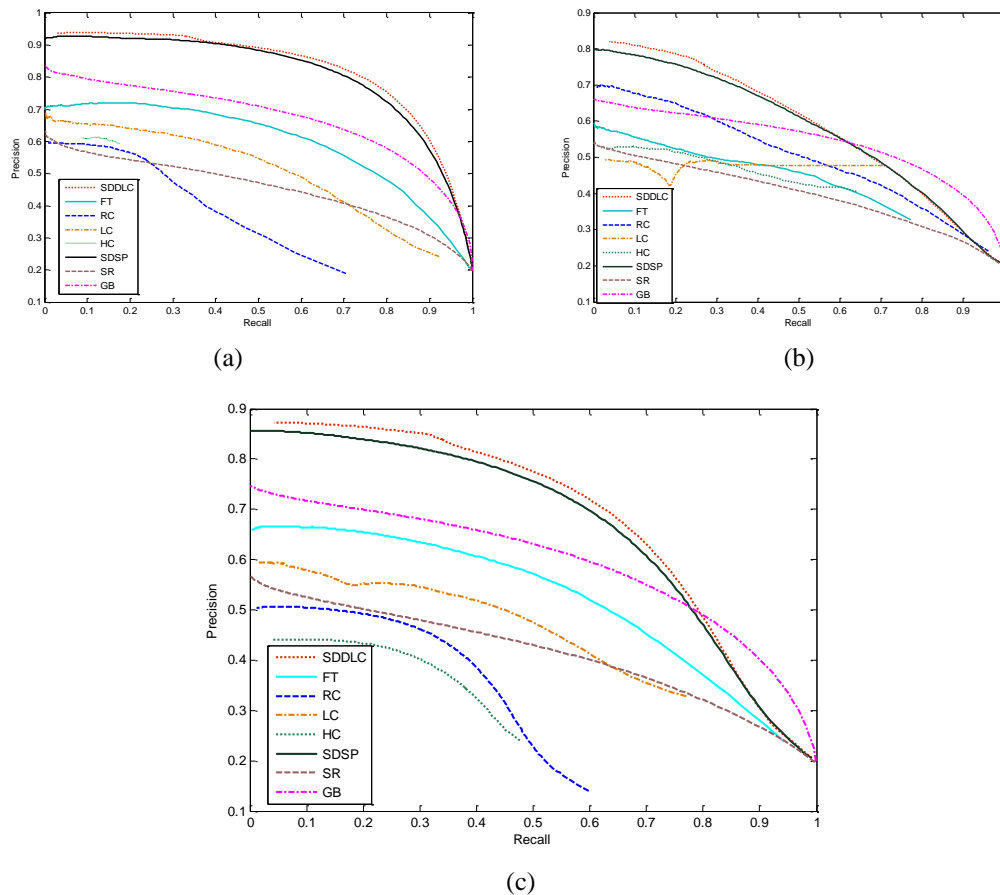


**Fig. 8.** Saliency maps by different methods under different illuminant conditions. (a) input images ; (b) SDDLc; (c) FT [13]; (d) RC [15]; (e) LC [14]; (f) HC [15]; (g) SDSP [16]; (h) SR [11]; (i) GBVS[10]

The model of SDSP [16] showed in Fig. 6 (g) and Fig. 7 generates saliency maps by the contrast of L-channel, a-channel and b-channel in the CIE Lab color space which contains the information of the light and the colors. Therefore, the saliency map of SDSP is better than that of RC when the salient regions have the same color histogram as the background. Our method SDDLCL mixes the information of illumination invariant features to reduce the effect of the illumination and emphasize the salient object in each image. The results of the proposed method are showed in Fig. 6 (b) and Fig. 7. We can find that the proposed algorithm outperforms other state-of-the-art methods under normal illumination.

#### 4.4 The comparison under different illuminant conditions

In this section, we compare the results of the proposed algorithm with the other algorithms under different illuminant conditions. The results of different algorithms under different illuminant conditions are shown in Fig. 8. In addition, we use the precision and recall to measure the results generated by the algorithms mentioned before. The comparison is shown in Fig. 9. At last, we compute the F-measure of the results of different algorithms which is shown in Table 2.



**Fig. 9.** Precision and recall by different methods. (a) 0.5 times of normal illuminant conditions; (b) 1.5 times of normal illuminant conditions; (c) is the average precision and average recall by different methods under different illuminant conditions.

It can be seen that the results generated by our method SDDLDC is more robust than that by other methods in Fig. 8. For FT, SR and LC, they are affected by not only light changes but also the noises of the background. GBVS, which seems robust to the variety of the illumination, can not detect the object accurately. RC and SDSP, which is comparable to the proposed algorithm, are heavily influenced by the mess of the background. Fig. 9 (a) and (b) are the precision versus recall curves for all algorithms under 0.5 and 1.5 of normal illuminant conditions. When the illuminant conditions become higher than 1.5 or less than 0.5, some algorithms would fail to detect the salient areas. Hence, the extreme illuminant conditions can test the saliency map of all algorithms completely. As shown in Fig. 9 (a), the precision and the recall curve of our method SDDLDC is higher than those of other methods. For the other models, the precision and the recall of the method HC is worse than other methods when the light is 0.5 times of normal illustration or low. In Fig. 9 (b), our method SDDLDC is higher than those of the other methods. SR is poor than those of the other methods when detect the salient objects in bright images. In addition, LC becomes the most unstable when the light is 1.5 times of normal illumination. Fig.9 (c) is the result that compares average precision versus average recall curves for all algorithms under 0.1~2.0 times of normal illuminant conditions. It can also be seen that HC is poor in both precision and recall when comparing with those of the other methods. The precision and the recall curve of our method SDDLDC is higher than those of the other methods since the saliency maps computed by our method are smoother and contain more pixels with the saliency value 255.

**Table 2.** F-measure by different methods for different illumination

Illuminant	SDDLDC	FT[18]	RC[19]	LC[20]	HC[19]	SDSP[1]	SR[33]	GBVS[21]
0.1	0.4541	0.1852	0	0.0538	0	0.2368	0.3824	0.4006
0.2	0.4792	0.1959	0	0.1142	0	0.3186	0.3930	0.4553
0.3	0.5180	0.1986	0.0045	0.1750	0.0024	0.3699	0.4014	0.5074
0.4	0.5548	0.2013	0.0384	0.2310	0.0343	0.4440	0.4036	0.5281
0.5	0.6221	0.2249	0.2368	0.3103	0.2251	0.5139	0.4087	0.5371
0.6	0.6790	0.2468	0.6750	0.3710	0.5732	0.6768	0.4094	0.5426
0.7	0.7668	0.2827	0.7623	0.4064	0.6889	0.7186	0.4133	0.5394
0.8	0.7735	0.3145	0.7711	0.4271	0.7069	0.7504	0.4156	0.5276
0.9	0.7717	0.3348	0.7621	0.4276	0.7273	0.7601	0.4156	0.5243
1.0	0.7804	0.3722	0.7655	0.4655	0.7273	0.7631	0.4137	0.5227
1.1	0.7614	0.3346	0.7006	0.4880	0.6350	0.6991	0.4156	0.5061
1.2	0.6869	0.2679	0.6019	0.3987	0.5099	0.5987	0.3872	0.4621
1.3	0.6023	0.2256	0.5115	0.3370	0.4305	0.5052	0.3550	0.4115
1.4	0.5249	0.2013	0.4441	0.2950	0.3629	0.434	0.3306	0.3739
1.5	0.4679	0.1986	0.3939	0.2604	0.3147	0.3736	0.3134	0.3481
1.6	0.4236	0.1959	0.2121	0.2310	0.1835	0.3621	0.2992	0.3251
1.7	0.3842	0.1852	0.0384	0.1750	0.0343	0.3199	0.2840	0.3045
1.8	0.3459	0.1818	0.0045	0.1650	0.0024	0.2843	0.2741	0.2860
1.9	0.3219	0.1503	0	0.1142	0	0.255	0.2649	0.2729
2.0	0.2980	0.0745	0	0.0538	0	0.2278	0.2561	0.2592
Average F-measure	0.5608	0.2194	0.3461	0.2723	0.3079	0.4688	0.3618	0.4317

We compute the F-measure of different algorithms under 0.1~2.0 times of normal illuminant conditions in Table 2 and Fig. 10. As shown in Table 2, all algorithms except FT, LC and SR can detect the salient regions accurately. With the light becomes higher or lower, the F-measure of each algorithm decreases gradually. As shown in the Fig. 10, the results of



the proposed method outperform other methods when the image with dark illumination. With the image becomes bright, the precision and recall decay faster. For other methods, the results of RC and HC decline fastest since both RC and HC are dependent on the color histogram. As we know, color histogram is affected by the light changing severely [22]. When we select the illumination as 0.1~0.2 and 1.9~2.0 times of normal illumination, the results of RC and HC are already close to zero. However, the results of our method SDDL and GBVS are robust to the light changing when compared with other models.

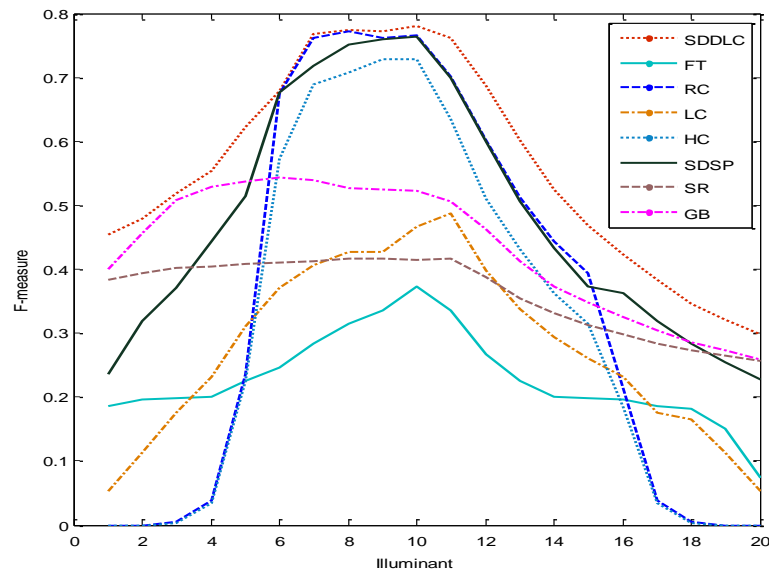


Fig. 10. F-measure by different methods under different illuminant conditions

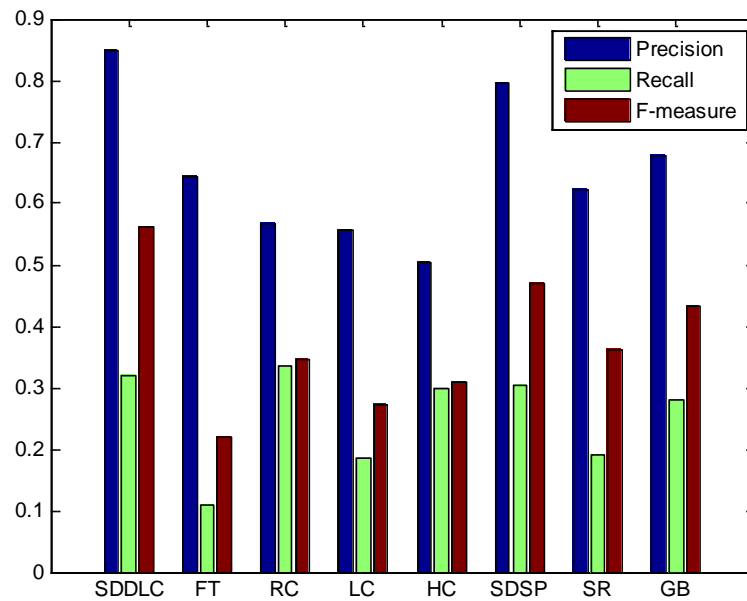


Fig. 11. Average precision, recall and F-measure by different methods under different illuminant conditions

In **Fig. 11**, we show the average precision, the average recall and the average F-measure by different methods with the illumination under 0.1~2.0 times of normal illuminant conditions. Generally speaking, the precision indicates the performance of the saliency detection algorithms compared with ground-truth saliency map. To compare the proposed model with others, we always see the precision value for different algorithms, for the precision value is the ratio of the correctly detected region over the whole detected region. From **Fig. 11**, we find out that the performance of our algorithm is close to that of SDSP, but be better than others.

## 5. Conclusion

We propose a novel saliency detection framework based on illumination invariant features to improve saliency detection for different light condition. In our framework, an input image is extracted from the illuminant invariant features that can help us to reduce the effect of the illumination to set up a transitional image based on the local sensitive histograms. Then a preliminary saliency map of the image is obtained in the CIE Lab color space. We use the method of the region growing to fuse the illuminant invariant features and the preliminary saliency map with the information of spatial distinctness lastly. We evaluate our method on an image dataset based on a publicly available datasets and compare our scheme with the state-of-art models under the normal and revised illustration conditions respectively. The resulting saliency maps are much less sensitive to background texture under the normal illustration condition.

Our future work will focus on high-level knowledge, which could be beneficial from handling more challenging cases and other kinds of saliency cues or priors to be embedded into our framework.

## References

- [1] U. Rutishauser, D. Walther, C. Koch, and P. Perona, "Is bottom-up attention useful for object recognition?" *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp.37-44, 2004. [Article \(CrossRef Link\)](#)
- [2] Chen, L. Q., Xie, X., Fan, X., "A visual attention model for adapting images on small displays," *Multimedia systems*, vol. 9, no. 4, pp.353-364, 2003. [Article \(CrossRef Link\)](#)
- [3] E. Louprias, "Wavelet-based salient points for image retrieval," in *Proc. of Image Processing, Proceedings, International Conference on*, IEEE, vol. 9, pp. 518-521, 2000. [Article \(CrossRef Link\)](#)
- [4] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.34, no. 10, pp. 1915-1926, 2012. [Article \(CrossRef Link\)](#)
- [5] L. Itti, "Models of Bottom-Up and Top-Down Visual Attention," *PhD thesis, California Institute of Technology*, Pasadena, 2000.
- [6] C. Kanan, M. Tong, L. Zhang and G. Cottrell, "SUN: Top-down saliency using natural statistics," *Visual Cognition*, vol.17, no. 6, pp. 979-1003, 2009. [Article \(CrossRef Link\)](#)
- [7] Z. Lu, W. Lin, X. Yang, E. Ong and S. Yao, "Modeling visual attention's modulatory aftereffects on visual sensitivity and quality evaluation," *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1928-1942, 2005. [Article \(CrossRef Link\)](#)
- [8] Jun Yu, Yong Rui, Yuan Yan Tang, Dacheng Tao, "High-Order Distance-Based Multiview Stochastic Learning in Image Classification," *IEEE Transactions on Cybernetics*, vol.44, no.12, pp.2431-2442, 2014. [Article \(CrossRef Link\)](#)
- [9] Ma Y-F, Zhang H, "Contrast-based image attention analysis by using fuzzy growing,"

- MULTIMEDIA '03 Proceedings of the eleventh ACM international conference on Multimedia*, pp. 374–381, 2003. [Article \(CrossRef Link\)](#)
- [10] J. Harel, C. Koch, P. Perona, “Graph-based visual saliency,” *Advances in neural information processing systems*, pp.545–552, 2007.
- [11] X. Hou and L. Zhang, “Saliency detection: a spectral residual approach,” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-8, 2007. [Article \(CrossRef Link\)](#)
- [12] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk, “Salient region detection and segmentation,” in *Proc. of ICVS*, Santorini, Greece, vol. 5008, pp. 66-75, 2008. [Article \(CrossRef Link\)](#)
- [13] R. Achanta, Hemami S, Estrada F, Süssstrunk S, “Frequency-tuned salient region detection,” *IEEE conference on computer vision and pattern recognition (CVPR)*, Miami Beach, pp. 1597–1604, 2009. [Article \(CrossRef Link\)](#)
- [14] Y. Zhai and M. Shah, “Visual attention detection in video sequences using spatiotemporal cues,” in *Proc. of MULTIMEDIA '06 Proceedings of the 14th annual ACM international conference on Multimedia*, pp. 815-824, 2006. [Article \(CrossRef Link\)](#)
- [15] M. Cheng et al, “Global contrast based salient region detection,” *IEEE International Conference on Computer Vision (CVPR)*, pp. 409–416, 2011. [Article \(CrossRef Link\)](#)
- [16] L. Zhang, Z. Gu , H. Li, “SDSP: A novel saliency detection method by combining simple priors,” in *Proc. of Image Processing (ICIP), 2013 20th IEEE International Conference on*, pp. 171-175, 2013. [Article \(CrossRef Link\)](#)
- [17] J. Zhou, Z. Jin, “A New Framework for Multiscale Saliency Detection Based on Image Patches,” *Neural Processing Letters*, vol. 38, no.3, pp. 361-374, 2013. [Article \(CrossRef Link\)](#)
- [18] S. He et al. “Visual Tracking via Locality Sensitive Histograms,” in *Proc. of IEEE International Conference on Computer Vision (CVPR)*, pp. 409–416, 2013. [Article \(CrossRef Link\)](#)
- [19] X. Shen, Y. Wu, “A unified approach to salient object detection via low rank matrix recovery,” *IEEE International Conference on Computer Vision (CVPR)*, pp. 853-860, 2012. [Article \(CrossRef Link\)](#)
- [20] T. Judd, K. Ehinger, F. Durand, A. Torralba, “Learning to predict where humans look,” *Computer Vision, 2009 IEEE 12th International Conference on (ICCV)*, pp. 2106-2113, 2009. [Article \(CrossRef Link\)](#)
- [21] Borji, A., Sihite, D. N., & Itti, L, “Salient object detection: A bench-mark,” in *Proc. of 12th European Conference on Computer Vision (ECCV)*, Springer Berlin Heidelberg, pp. 414-429, 2012. [Article \(CrossRef Link\)](#)
- [22] Mark S. Drew, et al, “Illumination-invariant color object recognition via compressed Chromaticity histograms of color-channel-normalized images,” in *Proc. of Computer Vision, 1998. Sixth International Conference on (ICCV)*, pp. 1-8, 1998. [Article \(CrossRef Link\)](#)
- [23] Sen Y, Qian Y, Avolio A, et al, “Image Segmentation Methods for Intracranial Aneurysm Haemodynamic Research,” *Journal of Biomechanics*, vol. 47, no. 5, pp. 1014-1022, 2014. [Article \(CrossRef Link\)](#)
- [24] Wang, M., Hong, R., Yuan, X.-T., Yan, S, Chua,”Movie2Comics: towards a lively video content presentation,” *IEEE Transactions on Multimedia*, IEEE, vol.14, pp.858–870, 2012. [Article \(CrossRef Link\)](#)
- [25] Wang M, Ni B, Hua X-S, Chua, “Assistive tagging: A survey of multimedia tagging with humancomputer joint exploration,” *ACM Comput Surv (CSUR)*, vol. 44, no 4, p. 25, 2012. [Article \(CrossRef Link\)](#)
- [26] Hong, R., Wang, M., Xu, M., Yan, S, Chua, “Dynamic captioning: video accessibility enhancement for hearing impairment,” in *Proc. of ACM multimedia*, pp. 421–430, 2010. [Article \(CrossRef Link\)](#)
- [27] Yanxiang Chen, Tam V. Nguyen, Mohan Kankanhalli, et al, “Audio Matters in Visual Attention,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol.24, no.11, pp. 1992–2003, 2014. [Article \(CrossRef Link\)](#)
- [28] Hong RC, Wang M, Yuan XT, Xu MD, Jiang JG, Yan SC, Chua TS, “Video accessibility enhancement for hearing-impaired users,” *ACM Trans Multimed Comput, Commun, Appl* 7: Article 24, 2011. [Article \(CrossRef Link\)](#)



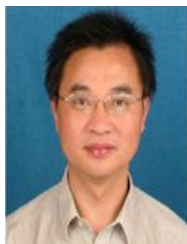
**Yongfeng Ren** is a doctoral student majoring in College of Computer and Information at Hohai University from September 2012 to present. He is a lecturer at Huaiyin Institute of Technology. His research interests include pattern recognition, image processing, etc.



**Jingbo Zhou** received his PhD degree in control science and engineering from Nanjing University of Science and Technology (NUST) in 2013. He is a lecturer at Huaiyin Institute of Technology. His research interests include pattern recognition, image processing, etc.



**Zhijian Wang** received his BS degree in Software Engineering from Nanjing University in 1982. He received his MS degree in Software Engineering from Nanjing University in 1986 and his PhD degree from Nanjing University, on the subject of Software Engineering, in 1989. His current research interests include pattern recognition, formal method, web service, software engineering, software testing, software component and water resources information study.



**Yunyang Yan** received his BS degree in computer application from Nanjing University of Aeronautics and Astronautics in 1988. He received his MS degree in computer application from Southeast University in 2002 and his PhD degree from NUST, on the subject of pattern recognition and intelligence systems, in 2008. His current research interests include pattern recognition and computer vision.