ICIC Express Letters
Part B: Applications
Volume **7**, Number **3**, March **2016**

ICIC International ©2016 ISSN 2185-2766
pp. **555–561**

A METHOD OF DETECTING SALIENT REGIONS EMPLOYING GLOBAL AND LOCAL SALIENCY

Ibuki Kuwata[1], Joo Kooi Tan[2], Hyoungseop Kim[2] and Seiji Ishikawa[2]

[1]Graduate School of Engineering
[2]Faculty of Engineering
Kyushu Institute of Technology
Sensuicho 1-1, Tobata, Kitakyushu, Fukuoka 804-8550, Japan
{ kuwata; etheltan }@ss10.cntl.kyutech.ac.jp

Received August 2015; accepted October 2015

Abstract. *This paper proposes a method of detecting prominent regions from an image using a global and local saliency measure. Detection of prominent regions from an image is important, because it is expected to improve the precision and the processing speed of object recognition through image processing. The proposed method employs two spatial redundancies, global saliency and local saliency. Global saliency is calculated by comparing the value of an interested pixel with the mode of the values of all pixels. On the other hand, local saliency is calculated using the relation among local pixels. In calculating the local saliency, the proposed method employs human visual characteristics, i.e., complementary color harmony to detect salient regions. Saliency map is made by integrating the global saliency and the local saliency considering some weights. Experimental results show the effectiveness of the proposed method.*

Keywords: Saliency map, Global saliency, Local saliency, Complementary color harmony, Object recognition

1. **Introduction.** In recent years, the study on the detection of salient regions on an image has been actively conducted, such as the study of Itti et al. [1]. Salient regions are foreground regions that are the subjects of image processing in object recognition. Detection of salient regions on an image enables to reduce the processing time of complex and time-consuming object recognition process.

In Itti et al.'s study [1], a saliency map, the map showing salient regions on an image, is created using physiology model that is based on the feature integration theory [2]. Salient regions are estimated by analyzing the log-spectrum of an image in Hou and Zhang [3]. On the other hand, Achanta and Hemami [4] propose a method of creating high resolution saliency maps by retaining wider range of frequency than other existing techniques. In their studies, the accuracy of saliency maps has become higher. However, a detector that has the same ability as a human has not yet been implemented.

In this study, we propose a method of detecting salient regions from an image using global saliency and local saliency. In particular, the local saliency is calculated by a human visual characteristic, i.e., complementary color harmony. The objective of this study is the creation of a detector that has the same ability as a human by simulating human visual characteristics.

2. **Proposed Method.** In this paper, we propose a method of detecting prominent (or salient) regions from an image using global saliency and local saliency. They are explained precisely in the following.

2.1. **Global saliency.** Global saliency is a feature value showing the degree of saliency of an interested pixel among entire pixels. The use of global saliency enables to estimate the background area. In this study, global saliency is calculated in the procedure shown in Figure 1.

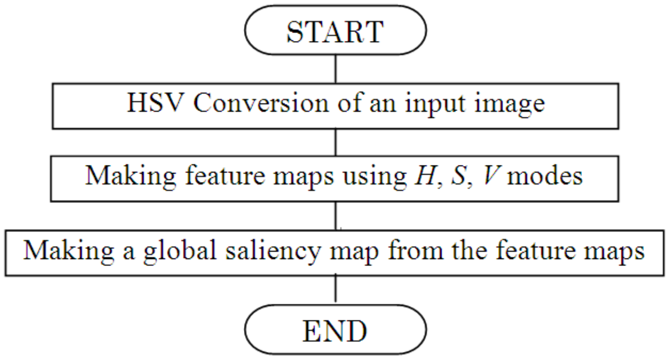

FIGURE 1. Procedure for the creation of a global saliency map

2.1.1. *HSV conversion.* In the first place, the color of an input image is converted from the RGB coordinates to the HSV coordinates. HSV coordinates have hue, saturation, and lightness. The HSV coordinates are known to be close to the human sense compared to the RGB coordinates. The formula for the conversion is given by the following,

$$\begin{aligned} H &= \begin{cases} 60 \times \dfrac{G-B}{MAX-MIN}, & \text{if } MAX = R \\ 60 \times \dfrac{B-R}{MAX-MIN} + 120, & \text{if } MAX = G \\ 60 \times \dfrac{R-G}{MAX-MIN} + 240, & \text{if } MAX = B \end{cases} \\ H &= H + 360, \text{ if } H < 0 \\ S &= \frac{MAX-MIN}{MAX} \\ V &= MAX \end{aligned} \tag{1}$$

Here H is a hue value, S is a saturation value and V is a lightness value: R, G, B are respectively the RGB values of an input image, MAX is the maximum value of the RGB values of an input image, and MIN is the minimum value of the RGB values of an input image.

2.1.2. *Feature maps.* In the second place, we find the mode (the value having the largest frequency) of HSV values of the input image by raster scan of the input image.

We then create three feature maps to represent the saliency of each feature, which is the absolute difference defined by

$$\begin{aligned} f_H(x,y) &= |m_H - i_H(x,y)| \\ f_S(x,y) &= |m_S - i_S(x,y)| \\ f_V(x,y) &= |m_V - i_V(x,y)| \end{aligned} \tag{2}$$

Here $f_k(x,y)$ is the pixel value of a feature map, m_k is the mode of HSV values, and $i_k(x,y)$ is the pixel value of an input image with respect to k ($k = H, S, V$).

2.1.3. *A global saliency map.* We define a global saliency map by the weighted sum of the three feature maps given by Equation (2). The accuracies of the three feature maps vary depending on the input image. Therefore, three weights, w_H, w_S, w_V are considered to define a weighted global saliency map as follows,

$$G(x, y) = w_H f_H(x, y) + w_S f_S(x, y) + w_V f_V(x, y) \tag{3}$$

Here $G(x, y)$ is the pixel value of a global saliency map and w_k $(k = H, S, V)$ is a normalized value of the mode $(w_H + w_S + w_V = 1)$. The larger the value of $G(x, y)$ is, the more salient pixel (x, y) is.

2.2. **Local saliency.** Local saliency is a feature showing the salient degree of an interested pixel among its neighbor pixels. The use of local saliency may enable to emphasize local prominent regions. The idea of the proposed method is to employ a certain function similar to a human perceptual sense. The defined local saliency is calculated by the procedure shown in Figure 2.

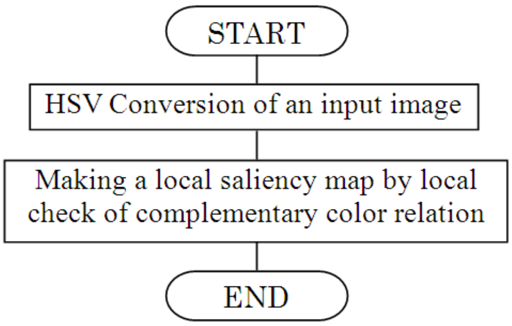

FIGURE 2. Procedure for the creation of a local saliency map

2.2.1. *Complementary color harmony.* In this paper, we calculate the local saliency using complementary color harmony [5] which is employed in human perception. Complementary color harmony is caused by combining complementary colors, the color to be diametrically opposite in the hue ring, and it has an effect to enhance the other color.

2.2.2. *HSV conversion.* The color of an input image is converted from the RGB coordinates to the HSV coordinates as in global saliency definition. In order to express the complementary color harmony, the hue value H is employed.

2.2.3. *Creation of a local saliency map.* A local saliency map is produced using Equation (4).

$$\begin{cases} L_1(x, y, u, v) = |h(x, y) - k(u, v)| & \cdots |h(x, y) - k(u, v)| < 180 \\ L_2(x, y, u, v) = 360 - |h(x, y) - k(u, v)| & \cdots |h(x, y) - k(u, v)| \geq 180 \end{cases}$$
$$L(x, y) = \sum_{(u,v) \in R(x,y)} \{L_1(x, y, u, v) + L_2(x, y, u, v)\} \tag{4}$$

Here $L(x, y)$ is the pixel value of a local saliency map, $R(x, y)$ is a local compared range on an image, $h(x, y)$ is the hue value of an interested pixel, and $k(u, v)$ $((u, v) \in R(x, y))$ is the hue value of the compared pixel. By the case classification as in Equation (4), $L(x, y)$ is maximized when the relationship of the interested pixel and the compared pixel is complementary relationship. The larger the value of $L(x, y)$ is, the more salient the pixel is locally.

2.3. Saliency map. Finally, the global saliency map and the local saliency map are integrated into a single value $S(x, y)$ using Equation (5),

$$S(x, y) = G(x, y) \times L(x, y) \tag{5}$$

This definition means that a pixel having a larger global saliency value is enhanced more by multiplication, if it has also a larger local saliency value.

In this way, an overall saliency value $S(x, y)$ at pixel (x, y) is defined and it provides an overall saliency map.

3. Experimental Results.

3.1. Experimental results. We use the dataset employed in [6] and [7] to evaluate the proposed method. The dataset has been used in many other papers as a benchmark to evaluate proposed saliency algorithms. Some of the images in the dataset are shown in Figure 3. The number of compared pixels (u, v) in Equation (4) is set to 100, which is 10 by 10 pixels centered at (x, y) with the interval of every 3 pixels.

Some experimental results are shown in Figure 4.

FIGURE 3. Some images in the dataset

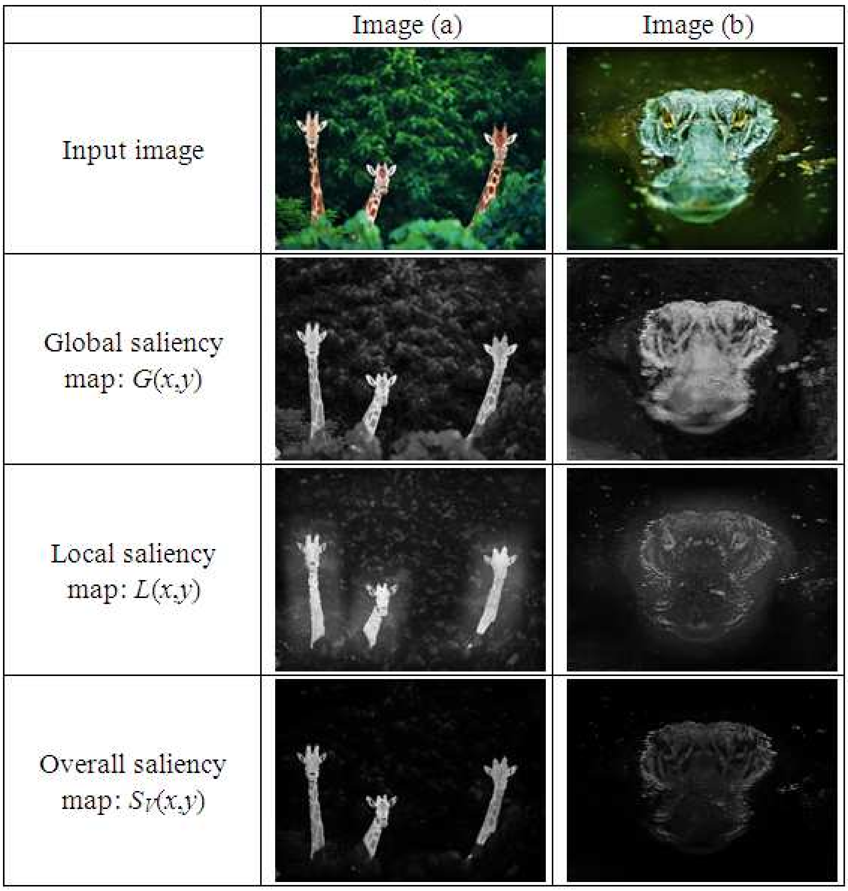

FIGURE 4. Experimental results

3.2. **Evaluation.** We evaluate the accuracy of the obtained experimental results. First, we provide a threshold to the saliency map. A segmentation image as shown in Figure 5 is obtained by making the pixel whose value is higher than the threshold a pixel in a prominent region and the pixel whose value is lower than the threshold a pixel in a non-prominent region. We compare the segmentation image to the ground truth image. Two evaluation indices *precision* and *recall* are calculated using the following,

$$\begin{aligned} precision &= \frac{TP}{TP+FP} \times 100 \\ recall &= \frac{TP}{TP+FN} \times 100 \end{aligned} \tag{6}$$

Here TP is true positive, FP is false positive and FN is false negative.

The precision vs. recall curve is then calculated by changing the threshold from 0 to 255. The precision vs. recall curve of the proposed method is shown in Figure 6(a) and the curves of other approaches are given in Figure 6(b). Processing time of the proposed method and other approaches are provided in Table 1.

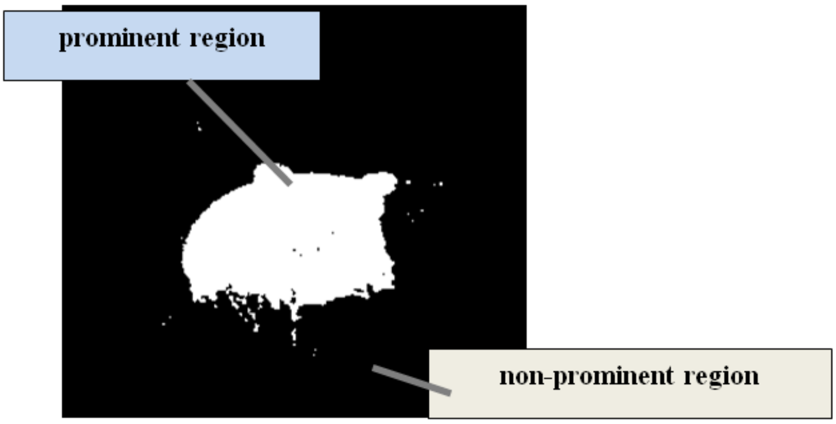

FIGURE 5. A segmentation image showing a prominent region

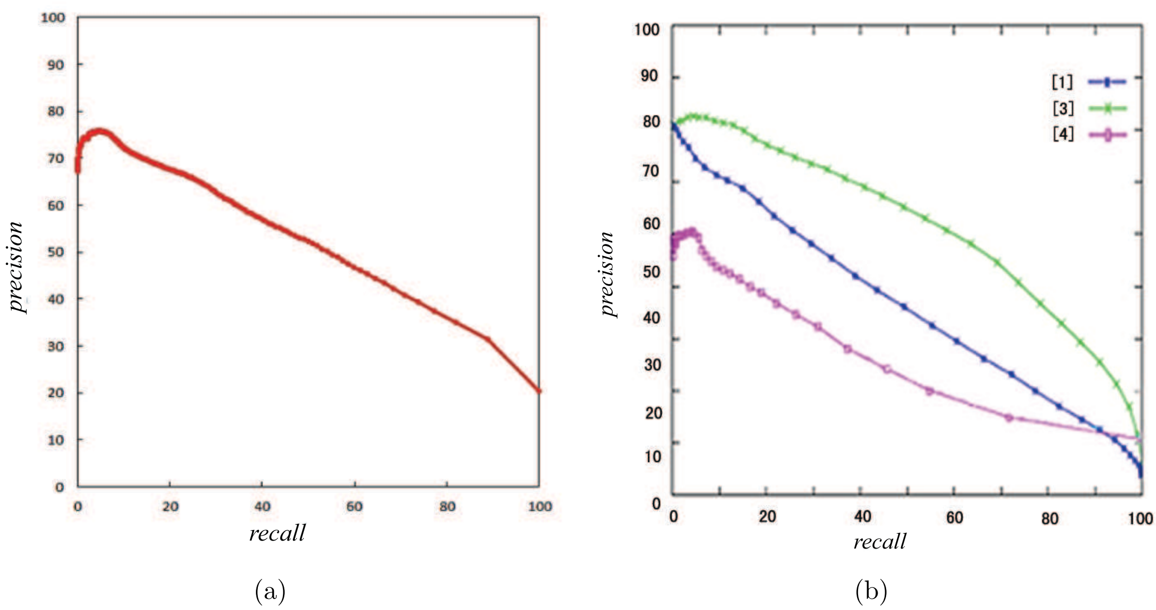

FIGURE 6. Precision vs. recall curves: (a) the proposed method, (b) other approaches [6]

Table 1. Processing time

Method	Processing time [msec]
Proposed method	243
Itti et al. [1]	994
Hou and Zhang [3]	65
Achanta et al. [4]	98

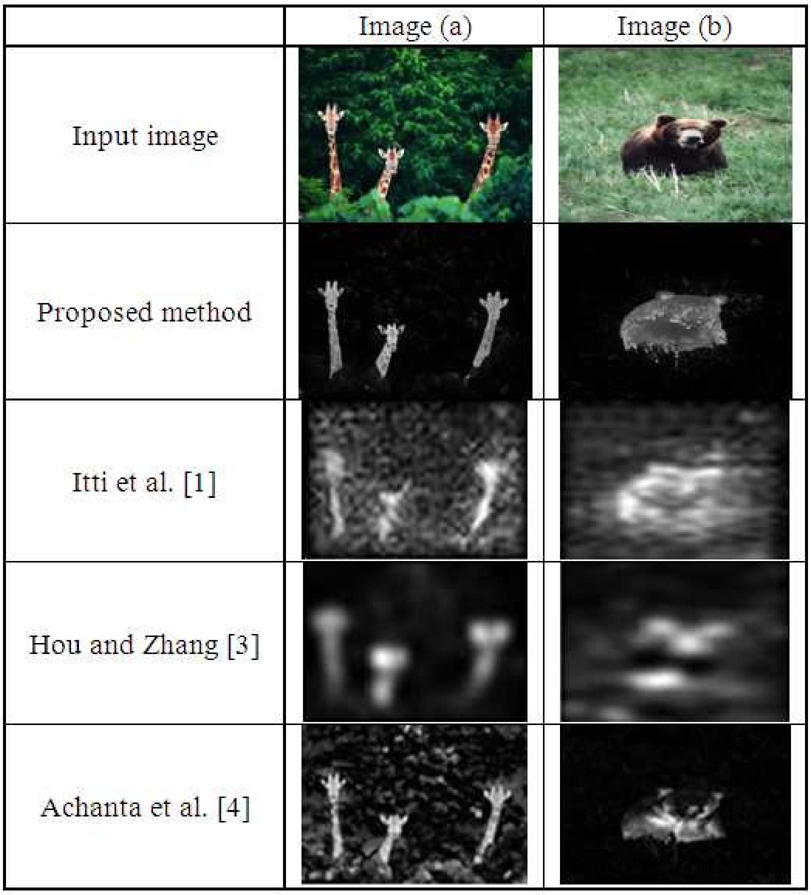

Figure 7. Results of the proposed method and other approaches

4. **Discussion.** Experimental results are satisfactory. As shown in Figure 7, it can be seen that the proposed method is able to remove the background area in a more effective way compared to other approaches. In particular, the proposed method can detect only a foreground area compared to other approaches in both images. This may come from the fact that both global saliency and local saliency are considered in the proposed method. The overall saliency measure defined by Equation (5) is the multiplication of the global saliency measure and the local saliency measure. Therefore, a region judged finally as a salient region is a prominent region in an image in the both global and local sense.

Furthermore, the area having complementary color is emphasized as shown in the circle in Figure 8. This indicates that the proposed method can produce a saliency map closer to a human sense by the employment of the local saliency considering human visual characteristics.

5. **Conclusion.** In this paper, we propose a method of detecting salient regions from an image using the saliency measure considering the global saliency and the local saliency. In particular, the local saliency is calculated taking account of one of the human visual

(a)

(b)

Figure 8. Saliency map emphasizing complementary color: (a) input image, (b) local saliency map

characteristics, i.e., complementary color harmony. Future work includes to improve the accuracy of detected regions and processing time.

Acknowledgment. This work was supported by JSPS KAKENHI Grant No. 25350477, which is greatly acknowledged.

REFERENCES

[1] L. Itti, C. Koch and E. Niebur, A model of saliency-based visual attention for rapid scene analysis, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.20, no.11, pp.1254-1259, 1998.
[2] C. Koch and S. Ullman, Shifts in selective visual attention: Towards the underlying neural circuitry, *Human Neurobiology*, vol.4, no.4, pp.219-227, 1985.
[3] X. Hou and L. Zhang, Saliency detection: A spectral residual approach, *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1-8, 2007.
[4] R. Achanta, S. Hemami, F. Estrada and S. Susstrunk, Frequency-tuned salient region detection, *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.
[5] K. Uchikawa, *Vision I: Structure of a Visual System and Primary Functions*, Asakura Shoten, pp.98-159, 2008 (in Japanese).
[6] A. Boudissa, *A Study on Saliency and Human Detection from an Image*, Ph.D. Thesis, Kyushu Institute of Technology, 2014.
[7] A. Jain, A. Wong and P. Fieguth, Saliency detection via statistical non-redundancy, *IEEE International Conference on Image Processing*, pp.1073-1076, 2012.