Saliency Detection Based on Non-uniform Quantification for RGB Channels and Weights for Lab Channels

Yahui Yuan[1], Aili Han[1(✉)], and Feilin Han[2]

[1] Department of Computer Science, Shandong University, Weihai 264209, China
hanal@sdu.edu.cn
[2] College of Computer Science and Technology, Zhejiang University, Hangzhou 310007, China

Abstract. We propose a non-uniform quantification method for RGB channels based on the different sensitivities of human eyes to the red, green and blue primary colors, which quantifies the R, G, B channels to different ranges, and then design a method of computing the weights for Lab channels and a method of computing the weighted color distance based on the different contributions of L,a,b channels to pixel saliency. Based on this, we present a saliency detection method using the non-uniform quantification method and the weighted color distance. First, we do the non-uniform quantification on an RGB image and convert the result into Lab space. And then, we compute the weights w_L, w_a, w_b for L,a,b channels by means of histogram, and compute the weighted color distance of each pixel I_k to all other pixels using the channel weights. Finally, the saliency of each pixel is computed using the weighted color distances. The proposed non-uniform quantification method, the weight computing method and the weighted color distance can be used in the early processing step for various applications based on color features. Experimental results show that our methods can improve the quality and efficiency for saliency detection to some extent.

Keywords: Saliency detection · Visual computing · Non-uniform quantification · Channel weight

1 Introduction

Human beings can understand complex scenes in real time because human visual process can choose a subset of perceptible information to reduce the complexity of scene analysis. The visual computing system should also have such function. How to select a valid subset of the given image is the key to improve the efficiency of visual computing system.

In order to select a valid subset of the given image, the existing saliency detection methods usually quantify the pixel values of RGB channels to the same range, which is called the uniform quantification method in this paper. In 1998, Itti et al. [1] proposed a visual attention system of combining multi-scale image features into a single saliency map, which reduces the input image to 640×480 pixels. In 2006, Harel et al. [2]

H. Zha et al. (Eds.): CCCV 2015, Part I, CCIS 546, pp. 258–266, 2015.
DOI: 10.1007/978-3-662-48558-3_26

proposed a graph-based visual saliency model (GBVS), in which they first reduce the input image to 600×400 pixels, and then use the linear filters to get one or more feature maps that are down-sampled to 25×37 pixels. In 2007, Hou et al. [3] proposed a spectral residual approach for saliency detection, in which they first adjust the input image to 64×64 pixels by means of low pass filter and down-sample method, and then do the saliency detection on the 64×64 images. In 2011, Cheng et al. [4] proposed a global contrast based salient detection method, which includes two contrast methods, i.e., a histogram-based contrast (HC) method and a region contrast (RC) method. The HC method first uniformly quantifies each channel of the input image to 12 values, and then computes the contrast between a pixel and all the other pixels.

The above uniform quantification methods can reduce the number of image colors, which result in less computations, but they do not consider the sensitivity differences of human eyes to different colors. We propose a non-uniform quantification method based on the different sensitivities of human eyes to the red, green and blue primary colors. Different from the previous quantification methods, our method quantifies the R, G, B channels to different ranges, which will help for the performance improvement of saliency detection method.

2 Non-uniform Quantification for RGB Channels

For an RGB image, the values in R,G,B channels are in the range of 0 to 255, corresponding to $256^3 \approx 1.68*10^7$ kinds of colors. According to the visual attention theory [11,12], only when the difference between pixel values reaches a certain quantity, human eyes can perceive it. Therefore, the pixel values in R,G,B channels in an image can be separately quantified according to the different sensitivities of human eyes to different colors so as to reduce the amount of computations. The previous methods quantify the pixel values in R,G,B channels to the same range, which can reduce the total number of colors, but do not consider the sensitivity differences of human eyes to different colors. In this paper, we propose a non-uniform quantification method, which quantifies the values in R,G,B channels to different ranges based on the different sensitivities of human eyes to the red, green and blue primary colors.

2.1 Visual Sensitivity Tests on RGB Channels

In order to understand the sensitivities of human eyes to different colors, we did the sensitivity tests on a three-primary-colors image, in which we separately changed the pixel values in RGB channels step by step.

Experimentally, we observed that when the pixel value in R channel decreases to about 40, it is hard to distinguish the red region for human eyes; when the pixel value in G channel decreases to about 30, it is hard to distinguish the green region for human eyes; when the pixel value in B channel decreases to about 65, it is hard to distinguish the blue region for human eyes. The experimental results are shown in Fig.1. The experimental results show that the visual sensitivities of human eyes to the red, green and blue colors are different, and the sensitivity order is: green > red > blue.

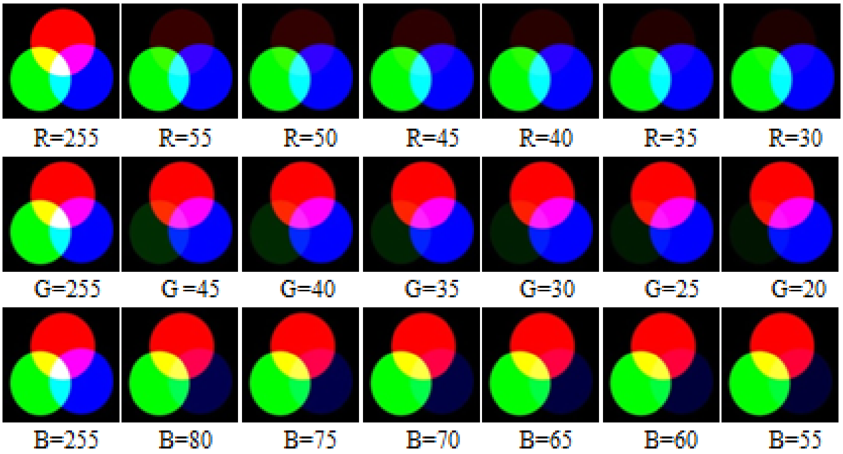

Fig. 1. The stepwise results of separately decreasing the pixel values in R,G,B channels in a three-primary-colors image.

From the visual sensitivity tests, we can know that the red in the black background is not salient when its pixel value is less than 40; the green is not salient when its pixel value is less than 30; the blue is not salient when its pixel value is less than 65.

2.2 Theory Basis of the Visual Sensitivity Order

The visible spectrum is shown in Fig.2(a), which are from [10]. The sensitivity of human eyes to a single color can be denoted by a luminosity function [10]. The luminosity functions under the light and dark environments are shown in Fig.2(b). From the luminosity function under the dark environment, we can know that human eyes are most sensitive to

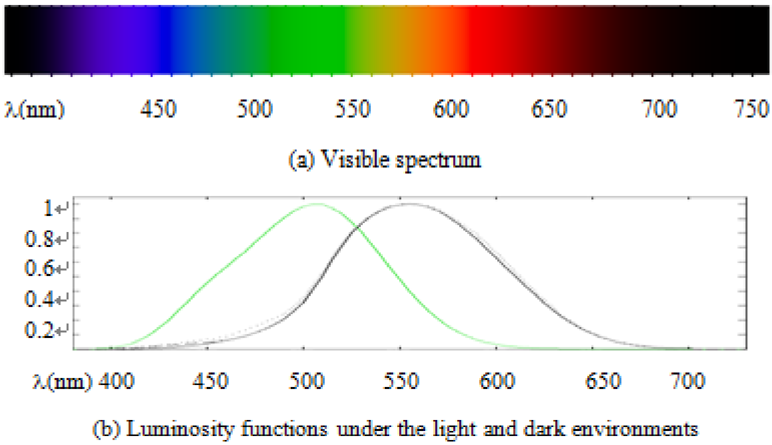

Fig. 2. Visible spectrum and luminosity functions [10].

the yellow-green color with 555nm wavelength, followed by green, red and blue. The closer the distance to both ends of the visible spectrum, the less sensitive the human eyes are. The experimental results in our visual sensitivity tests agree with this theory.

2.3 Non-uniform Quantification for RGB Channels

We design an algorithm of non-uniform quantification according to the experimental results in our visual sensitivity tests. Let k_1, k_2, k_3 denote the ranges of pixels values in RGB channels after non-uniform quantification, respectively. The total color number after non-uniform quantification is $k_1 * k_2 * k_3$, which is far less than that in the original RGB image. Therefore, the non-uniform quantification algorithm can reduce the amount of computations and improve the computational efficiency.

Algorithm 1. Non-uniform quantification
Input: an RGB image I;
Output: an RGB image I' after non-uniform quantification
(1) Quantify the pixel value in R channel of image I in step length of 40. The range of the pixel value in R channel after non-uniform quantification is $k_1 = \lceil 256/40 \rceil$;
(2) Quantify the pixel value in G channel in step length of 30. The range of the pixel value in G channel after non-uniform quantification is $k_2 = \lceil 256/30 \rceil$;
(3) Quantify the pixel value in B channel in step length of 65. The range of the pixel value in B channel after non-uniform quantification is $k_3 = \lceil 256/65 \rceil$.

We applied the non-uniform quantification algorithm to a number of RGB images, three of them are shown in Fig.3. From the visual effects before and after the non-uniform quantification, we observed that after the non-uniform quantification, the difficulty of distinguishing the object and the background in an image does not increase. The non-uniform quantification algorithm can reduce the color number and so improve the computational efficiency.

Fig. 3. The visual effects before and after non-uniform quantification. (a) The original RGB images; (b) The images after non-uniform quantification.

From the non-uniform quantification algorithm, we know that $k_1 = \lceil 256/40 \rceil = 7$, $k_2 = \lceil 256/30 \rceil = 9$, $k_3 = \lceil 256/65 \rceil = 4$. Thus, the total color number after the non-uniform quantification is $k_1 * k_2 * k_3 = 252$. The total color number decreases from about $1.68*10^7$ to 252, which results in the significant improvement in the computational efficiency.

3 Computing the Weights for Lab Channels

The histograms of RGB channels in the images shown in Fig.4(a) are shown in Fig.4(b)~(d), and the histograms of Lab channels in the images shown in Fig.5(a) are shown in Fig.5(b)~(d). We observed that the differences between gray distributions in Lab color space are more obvious than that in RGB space. Therefore, in the following saliency detection algorithm, we compute the weights for Lab channels.

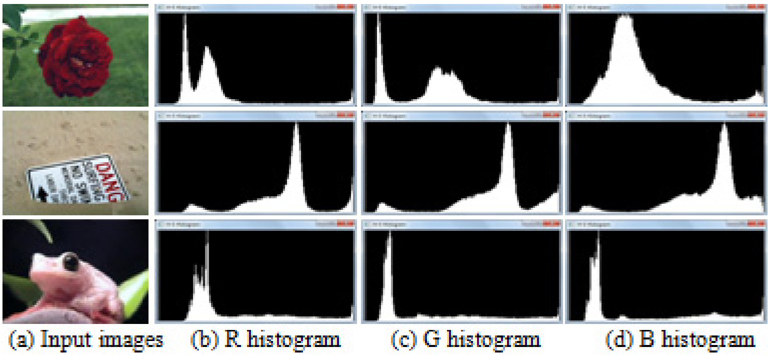

Fig. 4. The input images and their histograms of R, G, B channels.

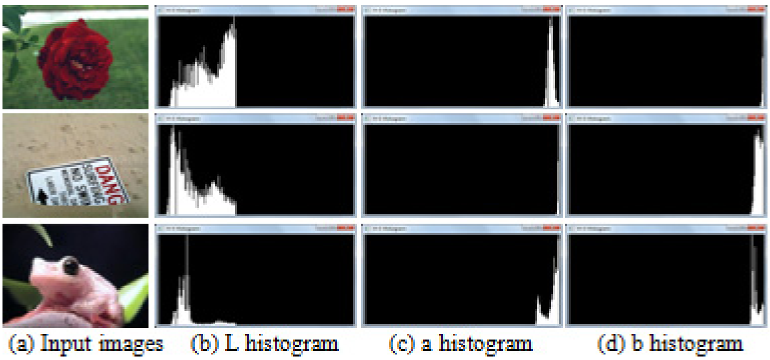

Fig. 5. The input images and their histograms of L, a, b channels.

Let miH_L, miH_a, miH_b denote the minimum coordinates of the bins with height greater than 5 in the L, a, b histograms, and maH_L, maH_a, maH_b denote the maximum coordinates in the L, a, b histograms, respectively. We use the following formula to compute the weights w_L, w_a, w_b for the L, a, b channels:

$$w_L = \frac{maH_L - miH_L}{\sum_{x \in L,a,b}(maH_x - miH_x)},$$

$$w_a = \frac{maH_a - miH_a}{\sum_{x \in L,a,b}(maH_x - miH_x)}, \tag{1}$$

$$w_b = \frac{maH_b - miH_b}{\sum_{x \in L,a,b}(maH_b - miH_b)}.$$

4 Saliency Detection Algorithm Based on Non-uniform Quantification and Channel Weights

The saliency of a pixel is defined as the color contrast of the pixel to all other pixels in the image in the HC method [4]. Our saliency definition is similar with that in the HC method, but we use the weighted color distance. By testing the visual sensitivity of human eyes to different colors, we know that different channels have different contributions to the saliency of pixels. Base on this observation, we give a weighted color distance $D_w(I_k,I_u)$ between pixels I_k and I_u in Lab color space:

$$D_w(I_k, I_u) = \frac{1}{w_L} \cdot \| v_{L_k} - v_{L_u} \| + \frac{1}{w_a} \cdot \| v_{a_k} - v_{a_u} \| + \frac{1}{w_b} \cdot \| v_{b_k} - v_{b_u} \| \tag{2}$$

For each pixel I_k, we compute its saliency $S(I_k)$ using the weighted distance $D_w(I_k,I_u)$:

$$S(I_k) = \sum_{\forall I_u \in I} D_w(I_k, I_u) \tag{3}$$

The saliency detection algorithm based on the proposed non-uniform quantification method for RGB channels and the weight computing method for Lab channels is given as follows.

Algorithm 2. Saliency detection algorithm

Input: an RGB image I;
Output: the saliency map of the input RGB image I

(1) Non-uniformly quantify each channel in the RGB image by Algorithm 1. The resulted image I' contains less colors.
(2) Convert the resulted image I' into Lab color space, and generate the histograms of L, a, b channels by means of histogram function.
(3) Compute the weights w_L, w_a, w_b for Lab channels by means of formula (1).
(4) Compute the weighted color distances by means of formula (2).
(5) Compute the saliency value $S(I_k)$ of each pixel I_k by means of formula (3) to get the saliency map of the given image.

Take the rose image shown in Fig.6(a) as an example. The image is first non-uniformly quantified into an image that contains less colors, as shown in Fig.6(b), which is then converted into Lab space to generate the histograms of L,a,b channels. From the histograms, we can get $minH_L$=14, $maxH_L$=99, $minH_a$=101, $maxH_a$=119, $minH_b$=117, $maxH_b$=119. The weights w_L, w_a, w_b are computed by the following formula:

$$\begin{aligned} w_L &= (99-14)/[(99-14)+(119-101)+(119-117)]=83/(83+18+2) = 0.8058\ , \\ w_a &= (119-101)/[(99-14)+(119-101)+(119-117)]=1/(83+18+2) = 0.1748\ , \\ w_b &= (119-117)/[(99-14)+(119-101)+(119-117)]=2/(83+18+2) = 0.0194\ . \end{aligned} \tag{4}$$

The final saliency map is shown in Fig.6(c). The experimental results of using different methods to the rose image are shown in the first line of Fig.7.

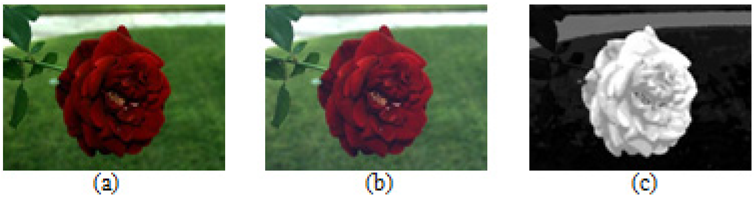

Fig. 6. Saliency detection based on non-uniform quantification and channel weights. (a) The original image; (b) The image after non-uniform quantification; (c) The saliency map.

5 Experimental Results and Analysis

We test our method on the MSRA-1000 [7] image dataset. In order to illustrate the effectiveness of our method, we compare the saliency maps obtained by our method with those obtained by the state-of-the-art saliency methods FT [7], HC [4], LC [8], RC [4] and SR [3], some of the experimental results are shown in Fig.7.

The experimental results show that, for the most images in the MSRA-1000 image dataset, the saliency maps obtained by our method are better than those by the LC and SR methods, and the overall results are not lower than those by the FT, HC and RC methods. For some images, the experimental results obtained by our method are more accurate. Take the rose image shown in Fig.6(a) as an example. The saliency maps obtained by FT, HC and RC methods include the rose region as well as the white horizontal region on the top side of the given image, and the white horizontal region is even more clearly than the rose region by using the FT and HC methods. The saliency map obtained by our method highlights the rose region better than the other methods, as shown in Fig.7.

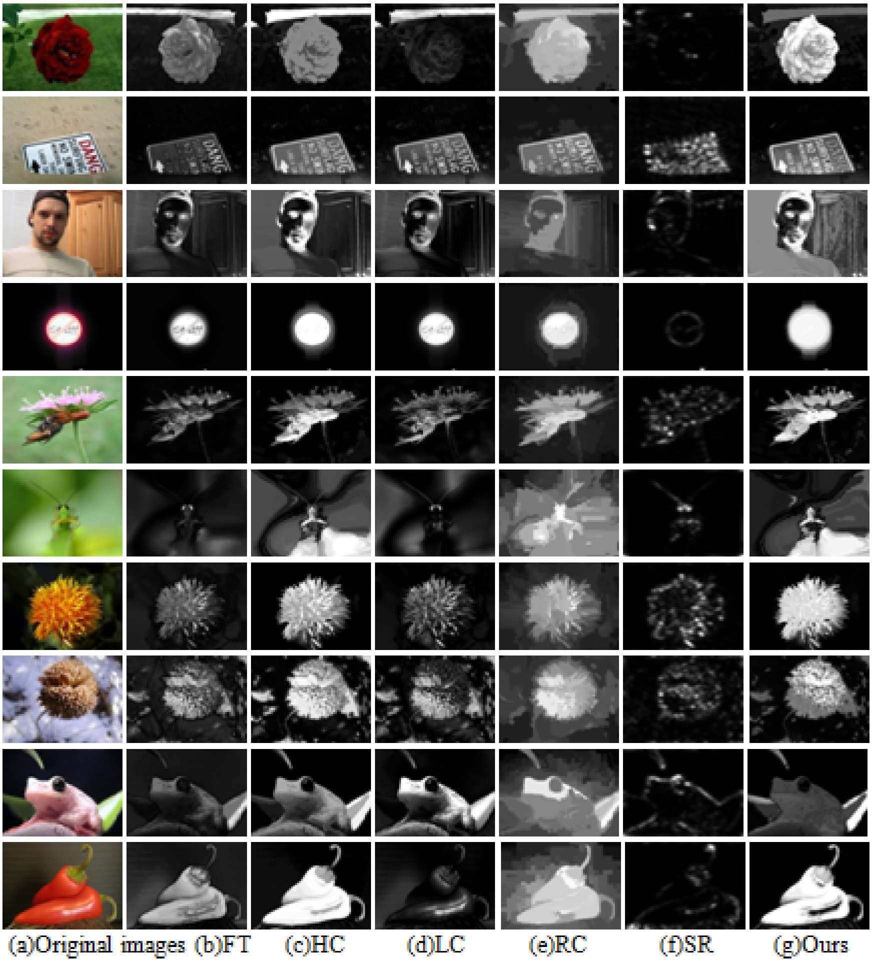

Fig. 7. Saliency maps by using our method and FT, HC, LC, RC, SR methods.

6 Conclusion

For the big data such as images or video, the usual method of decreasing the amount of computations is to down-sample the image. Another effective way is to decrease reasonably the range of pixel values in each channel. The existing methods do this usually by reducing the size of the image, while our method decreases the range of pixel values in each channel. Intuitively, the existing methods usually reduce the length and width of the image, which results in the loss of many image details, and our method reduce the range of pixel values in each channel, which does not lost the details of image but reduce the intensity of image details.

References

1. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence **20**(11), 1254–1259 (1998)
2. Harel, J., Koch, C., Perona, P.: Graph-based visual saliency. Advances in Neural Information Processing Systems **19**, 545–552 (2006)
3. Hou, X., Zhang, L.: Saliency detection: a spectral residual approach. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007), pp. 1-8 (2007)
4. Mingming, C., Guoxin, Z., Mitra, N.J., Xiaolei, H., Shimin, H.: Global contrast based salient region detection. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011), vol. 42(7), pp. 409-416 (2011)
5. Achanta, R., Estrada, F.J., Wils, P., Süsstrunk, S.: Salient region detection and segmentation. In: Gasteratos, A., Vincze, M., Tsotsos, J.K. (eds.) ICVS 2008. LNCS, vol. 5008, pp. 66–75. Springer, Heidelberg (2008)
6. Peng, J., Haibin, L., Jingyi, Y., Jingliang, P.: Salient region detection by UFO: uniqueness, focusness and objectness. In: 2013 IEEE International Conference on Computer Vision (ICCV 2013), pp. 1976-1983 (2013)
7. Achanta, R., Hemami, S., Estrada, F., Süsstrunk, S.: Frequency-tuned salient region detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009), pp. 1597-1604 (2009)
8. Yun, Z.: Visual attention detection in video sequences using spatiotemporal cues. ACM MM (2006)
9. Lu, S., Mahadevan, V., Vasconcelos, N.: Learning optimal seeds for diffusion-based salient object detection. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2014), pp. 2790-2797 (2014)
10. Wikipedia explanation of luminosity function. https://en.wikipedia.org/wiki/Luminosity_function
11. Desimone, R., Duncan, J.: Neural mechanisms of selective visual attention. Annual review of neuroscience **18**(1), 193–222 (1995)
12. Zeki, S.: Inner vision: an exploration of art and the brain.Oxford University Press (1999)