Saliency detection based on global color and space contrast

Yajie Zeng[1, a], Heng Li[1,b] and Xinyu Chai[1,c]

[1]School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

[a]zengyajie@sjtu.edu.cn, [b]liheng@sjtu.edu.cn, [c]xychai@sjtu.edu.cn

Keywords: Saliency detection, global contrast, space contrast

Abstract. Saliency detection acts as an important role in many computer vision tasks. We propose a novel bottom-up, object-level saliency detection method based on global color and space contrast. The proposed method not only considers the global color contrast of the image but also takes the space information into account, leading to an advanced multi-feature combined saliency detection model. Through qualitative and quantitative evaluation, our method consistently outperformed 4 global color (or intensity)–based saliency detection methods, resulting in higher precision.

Introduction

In various kinds of visual tasks, through the visual attention mechanism people with normal sight identify the most prominent object, which is called salient object, and concentrate on the salient object region. However, the detection of the salient region in an image by computational modeling still remains a challenge in computer vision. The saliency maps are made use of in numerous applications, including object recognition, image segmentation, image retrieval, content-aware image resizing, and prosthetic vision [1-5].

Visual attention is regarded as two stages according to the attention theories [6]: bottom-up, fast and data driven and top-down, slower and goal driven. Saliency widely focuses on bottom-up saliency detection. At present, there are approximately two categories of existing bottom-up saliency detection methods: local and global approaches. The local method is very time-consuming, which computes visual features in parallel, such as color, intensity, motion, and orientation [7-9]. By comparison, global methods, less time-consuming, detect the saliency map under a full resolution based on the object level and can uniformly highlight entire object. In 2006 Zhai [10] proposed LC algorithm to calculate the saliency value of each pixel according to intensity. Then Hou [11] proposed SR method based on frequency domain analysis, generating saliency map consistent with human perception. And Achanta [12] proposed FT method, defining the saliency by the difference between the color of a pixel and the average color of the image after smoothing. Afterwards histogram based contrast (HC) algorithm was proposed by Cheng [13], which is based on global color contrast. Even if HC can achieve real-time processing, this global method ignores the space information that is quite important to saliency detection, especially when the color of the object is similar to that of background.

Here, we develop a novel object-level saliency detection method based on global color and space contrast (GCSC). Firstly, we calculate a color saliency map using the histogram based contrast method. In parallel, a space saliency map is generated with the space saliency detection (SS) based on the watershed segmentation. The weighted saliency map integrates both the color and space information of the image. Through qualitative and quantitative comparison with LC, SR, FT and HC methods, our method has better performance overall.

Methods

For the deficiency of those saliency detection approaches purely considering about color contrast of the image, we propose a novel global color and space contrast method. In details, we combine the color saliency map generated from histogram based contrast and the space saliency map calculated by space saliency detection. The entire algorithm flowchart is shown in Fig.1.
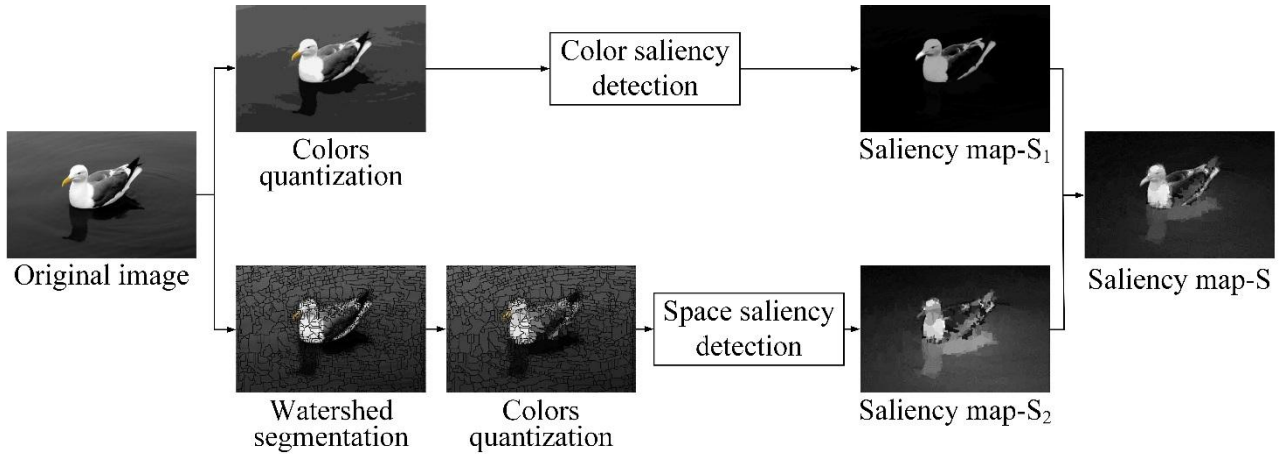
Fig.1: Schematic diagram of the proposed saliency detection method GCSC. The original image is firstly calculated by histogram based contrast method to obtain the saliency map-S_1. Meanwhile in the space saliency detection, after watershed segmentation and colors quantization, resulting in the saliency map-S_2. And the saliency map-S is generated by the weighted average of the two maps above.

Histogram based Contrast Saliency Detection. In the HC method, saliency values for image pixels are defined using color statistics of the input image. That is, the saliency of a pixel is defined using its color contrast to all other pixels in the image, i.e., the saliency value of a pixel I_k in image I is:

$$S(I_k) = \sum_{\forall I_i \in I} D(I_k, I_i). \tag{1}$$

Where $D(I_k, I_i)$ is the color distance metric in the L*a*b space for perceptual accuracy. Because pixels with the same color have the same saliency under this definition, the equation (1) can be transformed to:

$$S(I_k) = S(c_l) = \sum_{j=1}^{n} f_j D(c_l, c_j). \tag{2}$$

Where n is the number of distinct pixel colors, and f_j is the probability of pixel color in image I.

In order to speed up the computational efficiency, Cheng used the full color space to reduce the number of colors, using the simple histogram based quantization instead of optimizing for an image specific color palette. In order to reduce noisy saliency results caused by quantization randomness, a smoothing procedure was used to refine the saliency value for each color. The saliency value of each color was replaced by the weighted average of the saliency values of similar colors.

Space Saliency Detection. The first step of SS is watershed segmentation. The watershed segments the image according to the gradient and extracts the edges information of objects. Even if when the foreground has the similar color to the background, the edge information still can distinguish them. Here, the edge information is not used directly but the block information generated from the edge boundary is needed. In order to improve computation efficiency and saliency effect, the image colors need to be quantized. After the watershed segmentation, the color of each block is averaged:

$$\begin{cases} R_i = \sum_{n=1}^{k} R(n)/k \\ G_i = \sum_{n=1}^{k} R(n)/k \\ B_i = \sum_{n=1}^{k} R(n)/k \end{cases} \tag{3}$$

Where R_i, G_i, and B_i are the block colors, and k is the number of the pixels in the block.

In SS detection model, the saliency value is still based on the color information, which is obtained by computing the Euclidean distance between each block and the other blocks:

$$S(I) = \sum_{i-1}^{n} D(I, i). \tag{4}$$

Because of the feature of watershed algorithm, the pixels of edges are assigned as zero, not included in the saliency computation. Therefore, we make smoothing on the whole image to get a better saliency result.

Results

In order to evaluate the proposed method, we analyze the results from qualitative and quantitative aspects. The qualitative comparison of our method with SR, LC, FT and HC is shown in Fig.2. Apparently, our saliency maps enhance saliency level and are closer to the ground truth (GT) than the others.
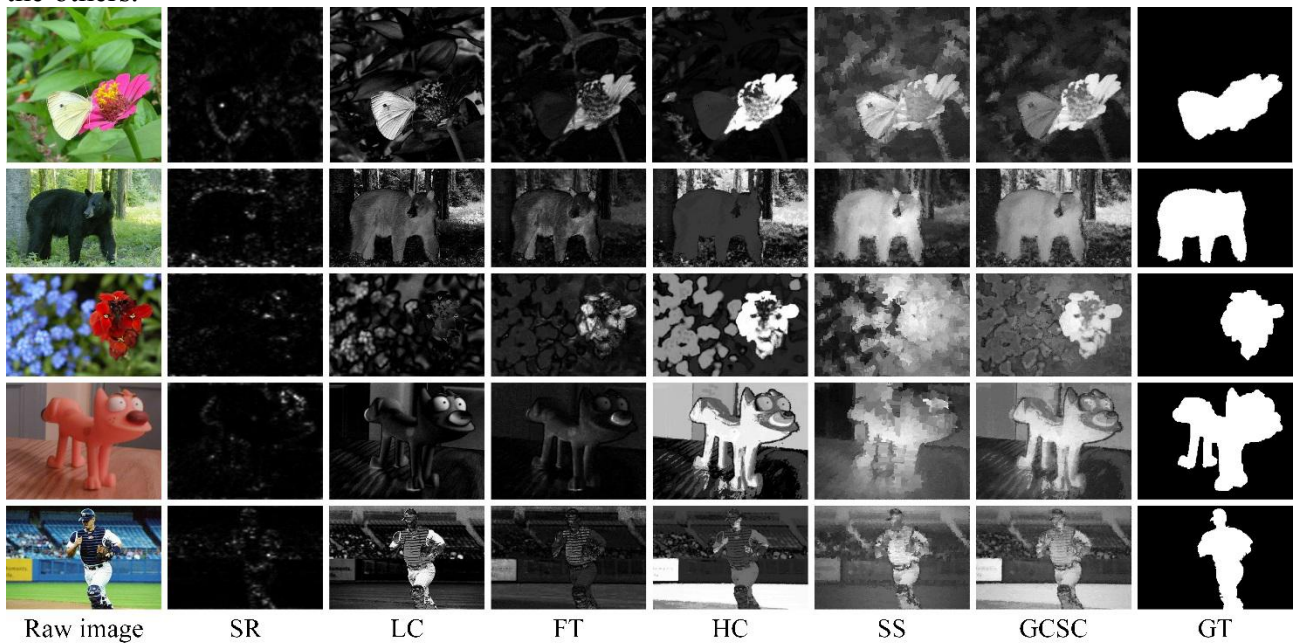

Fig.2: Some results of saliency maps using different saliency detection methods

We compared the precision, recall and F-Measure statistics with other methods on the MSRA1000 database. Fig.3 shows the resulting precision and recall curves. The result shows that our method GCSC had higher precision and F-Measure.
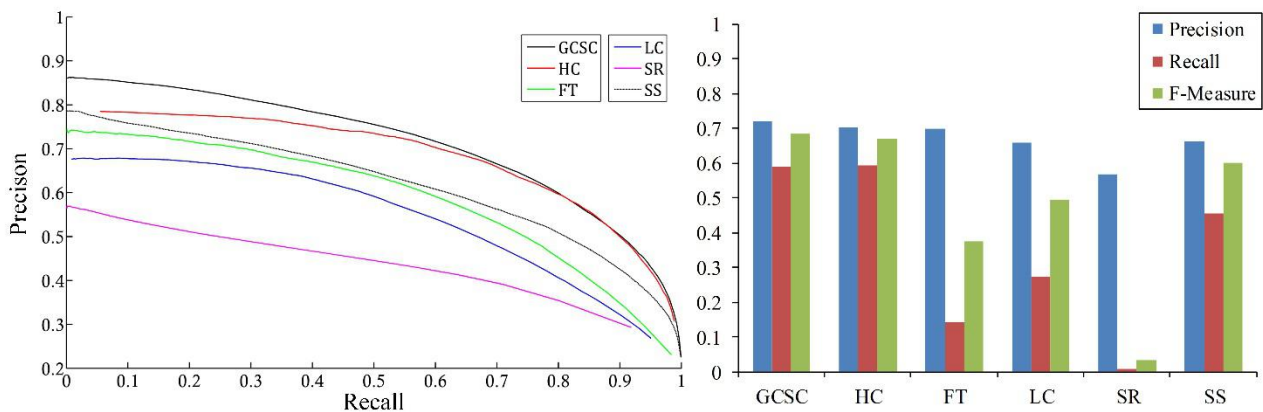

Fig.3: Precision and recall curves based on the MSRA1000 database

Conclusions

In this study, we proposed a novel object-level saliency detection method based on global color and space contrast, which combines the color saliency map generated from histogram based contrast and the space saliency map calculated by space saliency detection. We evaluated our method on the MSRA1000 database and compared it with other 4 global color (or intensity)–based saliency detection methods. The results indicate that the proposed method enhances saliency level and is superior in precision.

Acknowledgements

References

[1] J. Han, K. Ngan, M. Li and H. Zhang: IEEE TCSV, 16(1) (2006), p. 141–145

[2] B. Ko and J. Nam: J Opt Soc Am, 23(10) (2006), p. 2462

[3] U. Rutishauser, D. Walther, C. Koch and P. Perona: CVPR (2004), p. 37–44

[4] M. Ding and R.-F. Tong: The Viusal Computer, 26 (2010), p. 721–729

[5] Wang, J., Lu, Y. Y. and Gu, L. J.: Inform Sciences, Sep 1 (2014), p. 512-524.

[6] C. Koch and S. Ullman: Human Neurbiology, vol. 4 (1985), p. 219–227

[7] A. Borji, D.N. Sihite and L. Itti: Computer Vision–ECCV (2012), Springer (2012), p. 414-429

[8] J. Harel, C. Koch and P. Perona: Advances in neural information processing systems (2006), p. 545-552

[9] L. Itti, C. Koch and E. Niebur: IEEE Transactions on pattern analysis and machine intelligence, 20 (1998), p. 1254-1259

[10] Zhai Y and Shah M: ACM International Conference on Multimedia, Santa Barbara, Ca, Usa, October (2006), p.478-482

[11] Hou X and Zhang L: IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society (2007), p. 1-8

[12] R. Achanta, S. Hemami, F. Estrada and S. S¨usstrunk: CVPR (2009), p. 1597–1604

[13] Cheng, M. M., Zhang, G. X., Mitra and N. J.: IEEE Conference on Computer Vision And Pattern Recognition (Cvpr) (2011), p. 409-416