Real-Time Visual Saliency Detection Using Gaussian Distribution

Haoqian Wang, Chunlong Zhang, Xingzheng Wang
Graduate School at Shenzhen, Tsinghua University
Shenzhen 518055, China
Email: wanghaoqian@tsinghua.edu.cn

***Abstract*—Image visual saliency detection without prior knowledge of image details is fundamental for many computer vision tasks including object recognition, image retrieval, and image segmentation. In order to achieve more accurate and quick detection, this paper proposed a novel global contrast method to generate full resolution saliency maps using Gaussian distribution model. Compared with existing methods, this developed algorithm could be implemented in real-time with a higher accuracy. After a reasonable estimation of the parameters in our method, comparison experiments were conducted with five typical algorithms, experimental results demonstrate our approach is faster than the current real time approaches and accurate in maintaining high quality.**

***Index Terms*—Visual Saliency Detection; Gaussian Distribution; Real-Time**

I. Introduction

Salient region detection is to make an object, a region, or some pixels stand out of the image. People can easily and exactly focus on the important regions in images and ignore the parts they are not interested in. Computationally detecting the visual salient regions is essential for many computer vision tasks, and the saliency map is useful in many applications, *i.e.*, adaptive region-of-interest based image compression [1], object-of-interest image segmentation [2], object recognition [3], and 2D-to-3D conversion [4].

Detecting salient regions from a single image often utilizes clues like gradient, edges, color, intensity[5]. Computational efficiency and accuracy of salience map are the two most important evaluation standard for salience detection algorithms. In this research, we focus on the approaches which utilize the contrast of regions to their neighbors. Most of these kinds of saliency models could be classified into two groups [6]: local and global contrast processes.

Itti et al.[7] built their model motivated by Koch and Ullman [8], and estimated the saliency value by Difference of Gaussians (*DoG*) method. However, this model only produces low resolution saliency maps. Ma and Zhang [9] proposed a saliency map method based on local contrast analysis and used a fuzzy growing method to extract the interesting regions. Liu et al. [10] use multi-scale contrast, center-surround histogram, and color spatial distribution to generate saliency maps. Goferman et al. [11] propose a context-aware saliency method, based on the principles of visual attention, local low-level clues, global considerations, and high-level factors to enhance salient regions. Recently, Katramados and Breckon [6] propose a real time method using Division of Gaussians (*DIVoG*). These above methods use local contrast to generate saliency maps, but their output results is not so accurate which often have high saliency values near edges.

Relative to local saliency detecting method, global contrast based methods estimate saliency value using the contrast of whole image pixels. Zhai and Shah [12] propose a method using both spatial and temporal saliency maps. But they only use luminance factor in order to improve operation efficiency. Achanta et al. [5] propose a frequency tuned method using *DoG* filter. Cheng et al. [13] propose a global contrast based method and use the results to create segmentation masks. They segment the input image into regions and then compute the region color contrast and spatial weights. These methods have been proven to generate full resolution saliency maps, but can not give good performance both in computational efficiency and accurate saliency regions.

In order to achieve high quality full resolution saliency maps in real-time applications, we propose in this paper a novel visual saliency detection model by using gaussian distribution. Most existing saliency detecting approaches use multi-scale factors to gain a reasonable measurement, but it is computationally complex and time-consuming. As we know, gaussian mixture model (*GMM*) [14] is almost the most statistically mature methods for clustering (though they are also used intensively for density estimation) which could be employed to model the colors of an object to perform tasks such as real-time color-based tracking and segmentation, hence, we try to combine the *GMM* model with the salience detection task to improve the accuracy and speed. The *GMM* model use optimization method which is called Expectation Maximization (*EM*) to estimate the parameters. It consists of several single gaussian functions as follows:

$$p(x) = \sum_{i=1}^{n} \alpha_i g(x; \mu_i, \delta^2) \tag{1}$$

Three single gaussian functions were utilized to estimate the salient value for a pixel, and each gaussian function corresponds to a channel of a pixel. After parameter generation, the full resolution map could be derived in real-time.

The rest of the paper is organized as follows: Section II

This work was supported by the National Science Foundation of China under Grant U0935001, 61170195.

978-1-4799-4013-4/14 $31.00 © 2014 IEEE
DOI 10.1109/ICMB.2014.41

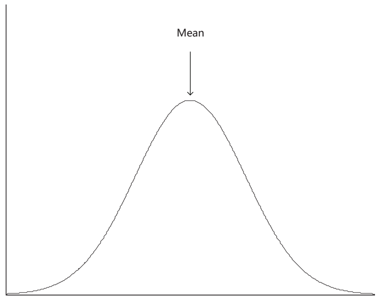


Fig. 1. Graph of gaussian function.

(a) (b) (c) (d)

Fig. 2. Saliency map. From left to right: input image, saliency map of Itti et al.[7], saliency map of Achanta et al.[5], our saliency map. Both our method and [5] are real time approaches.

elaborately describes our gaussian distribution based saliency detection algorithm. In section III, experimental results of our algorithm and comparisons with several other approaches are introduced and finally section IV concludes our paper.

II. The Proposed Algorithm

The details of our algorithm for saliency detection are presented in this section. Compared with most previous methods, we firstly utilize gaussian probability model to compute the global contrast ratio.

$$p(x) = a \exp^{-\frac{(x-\mu)^2}{2\delta^2}} \tag{2}$$

Based on the gaussian model, we utilize three gaussian functions to estimate the saliency values for pixels, and then we propose a simple method to estimate the parameters of each gaussian model. We also utilize the idea of histogram statistic to improve the efficiency of our algorithm.

A. Contrast Calculation

Our human vision system is sensitive to contrast in images, based on the biological vision survey. Most of the factors obey to gaussian distribution, so we utilize gaussian distribution to calculate the contrast between a pixel and the global color statistics of the input image. In detail, the contrast of a pixel is defined using its color to gaussian mean value and variance value, we will estimate mean value and variance value in next step. The contrast value of a pixel *X* in image *I* is defined as,

$$C(X) = \sum_{i=1}^{3} \alpha_i a \exp^{-\frac{(x_i-\mu_i)^2}{2\delta_i^2}} \tag{3}$$

where x_i is the Lab color space value [9] corresponds to a channel, a is the gaussian parameter, in order to control each channel contrast value from 0 to 1, we set $a = 1$, where α_i is the ratio of each gaussian function value.

B. Parameter Estimation

Based on section 2.1, we should estimate the parameters α_i, μ_i, δ_i, respectively. For a input image *I*, we classify the pixels to salient regions and background, and we know x_i is the pixel that needs to calculate the contrast value, so it is easily to deduce that α_i, μ_i should tend to represent the global background. Salient regions generally have higher luminance values, and the pixels near edges are transitional regions. Based on these two criterions, we utilize the pixels that have low luminance values and out of edges to compute α_i and μ_i. Denoting α_i satisfies $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Based on the method of [12], we know if there is only luminance values to be used, it also could get a rough result. However, their approach ignores the difference of color information. In our work, we use the full color space instead of one factor only, through our experiment, we set α_1 a higher value than $\alpha_2 = \alpha_3$.

C. Saliency Map Generation

In this section, we transform the contrast value to saliency value and then generate the saliency map, we also utilize a method to improve the efficiency. Based on our contrast model and the graph of gaussian function Fig.1,

we know that the mean value, in other words, the symmetry axis represents the background, so if a pixel tend to be salient, it will be have a low contrast value, far away from the symmetry axis. In section 2.1 and 2.2, we have controlled the contrast value from 0 to 1, so the saliency value of a pixel *X* in image *I* is defined as,

$$S(X) = 1 - C(X) \tag{4}$$

Normalizing the saliency value to 0-255, we can generate a saliency map. It is easy to find that the each channel of pixels with the same color value has the same contrast value under our definition, so we need not compute every pixel once. Building a lookup table to each channel of the pixels is an effective way. Our method only needs to compute 256×3 times if three gaussian function lookup tables were established. The equation representation in Equation 3 only takes $O(N)$ time.

D. Implementation notes

We implemented our algorithm in CIE *LAB* color space for it was showed could achieve a more accurate result, and a lookup table was generated for each color change hence to improve the computational efficiency. Furthermore, pixels with low-level luminance values and out of edges were utilized to estimate gaussian parameters, and all saliency maps were normalized to 0-255 so as for displaying purpose.

III. EXPERIMENTS AND COMPARISONS

We tested our method on the publicly available database provided by Achanta et al. [5]. The database has 1000 images, and it also has ground truth which is accurate human-marked labels for salient regions. In this paper, we compared our method with several other classical and popular saliency detection approaches. We choose these approaches on the basis of number of citations (IT [7], LC [12]), date of publication (RC [13], CA [11]), computational efficiency (FT [5], LC [12]). Detection accuracy was firstly compared between our proposed method and other algorithms, then, Since the main contribution of this paper is the real-time performance of the approach, we compare the computational efficiency of our approach with other approaches.

A. Accuracy of Salience Map

Fig.2 shows the detection result of our method with two classical methods on two simple content images. It shows that our method could detect more accurate and robust salient content than other two methods. Besides, since our method could be implemented in real-time, it greatly outperforms other two method.

Fig.3 shows more visual results, because of the limitation of space, we present the saliency map of 4 methods and our method. Obviously, our method get a better performance than the previous approaches. The Precision-Recall curve for our method and four other methods shows in Fig.4. Our saliency maps are obtained by the method as described in Section 2 with the parameters $\alpha_1 = 0.4$, $\alpha_2 = \alpha_3 = 0.3$, and the luminance value less than 150.

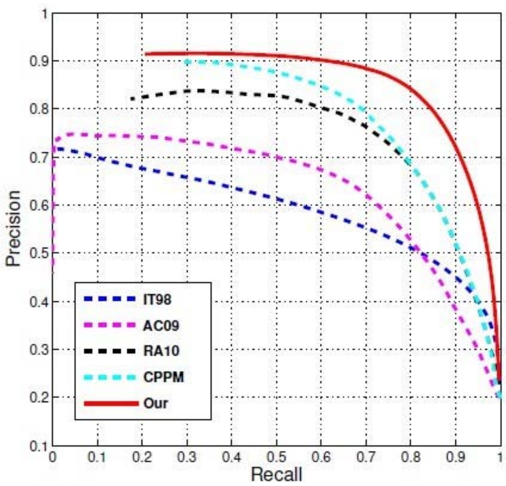


Fig. 4. Precision-Recall curve for our method and four other methods.

TABLE I
COMPARISON OF AVERAGE EXECUTION TIME OF 1000 IMAGES.

Method	Time(s)	Code
IT	0.611	C
RC	0.253	C++
CA	3.1	C
LC	0.018	C++
FT	0.016	C++
GC	0.014	C++

B. Running-time Comparison

To evaluate the execution time of our approach and other approaches, a Dual Core 2.6 GHz processer was used with 4GB RAM. Table 1. shows a comparison in time efficiency between our approach and other approaches. It is the average time taken to compute saliency map for images in the database by Achanta [5]. Most of the images provided by the database have resolution 400 $\times$ 300.

Compared with the two real time method Achanta et al. [5] and Zhai and Shah [12], our method gets a fraction of time superiority, and our approach demonstrates higher quality saliency maps than Achanta et al. [5], even compared with the time-consuming approach RC[13], our saliency maps perform good because of the maps obtained by our method are smoother and contain more pixels with the saliency value 255. There is a linear relationship between the image size and execution time, because we implement our method through building lookup tables.

IV. CONCLUSIONS

In this paper, we propose a novel computational global contrast method for visual saliency detection using Normal distribution, namely Gaussian function based Contrast. The presented method is real-time and generates full resolution saliency maps with high quality. We evaluated our method on the public database provided by Achanta et al. and compared our approach with five other methods. Experiments show the proposed method is better in terms of precision than existing real time methods, while still easy to be implemented.

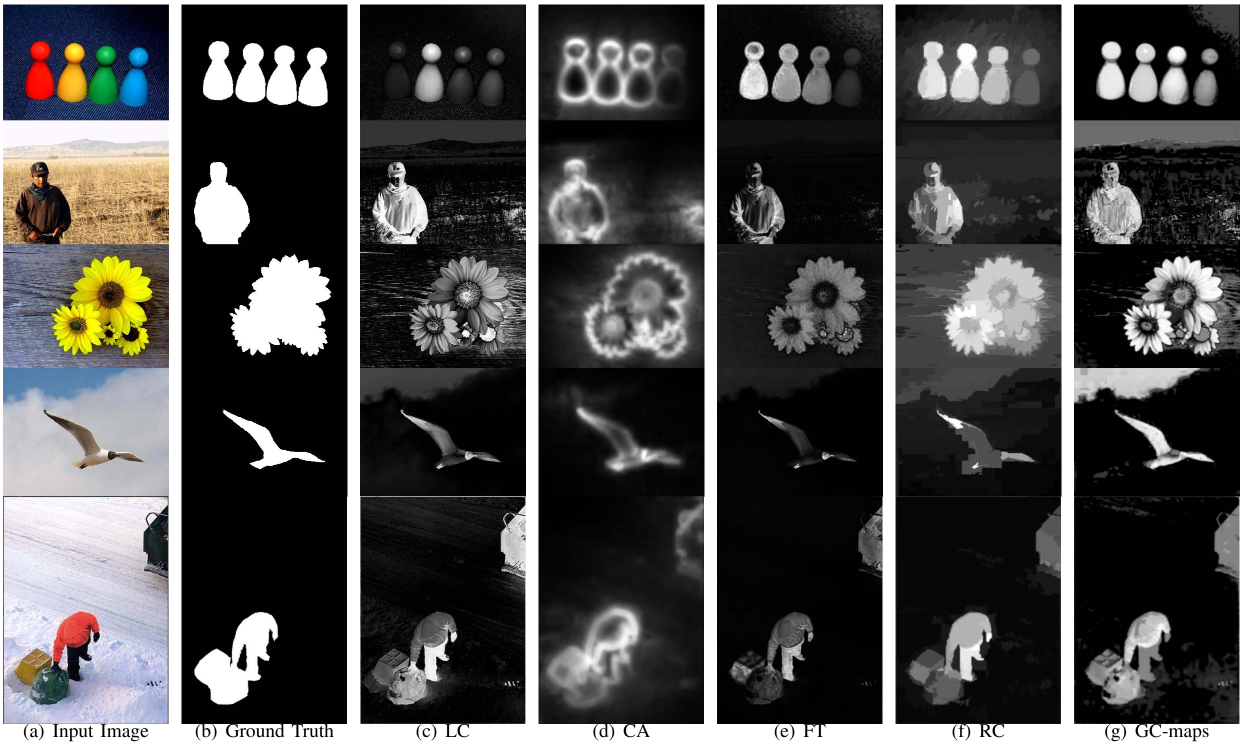
Future work could be done to focus more on the efficient algorithms that incorporate global contrast in real time saliency computation while implementing the work well saliency detection method. Also, it is essential for us to find some accurate methods to estimate the priori knowledge. Finally it would be great to investigate the application of this technology in the 2D to 3D conversion area.

REFERENCES

[1] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The jpeg2000 still image coding system: an overview," *Consumer Electronics, IEEE Transactions on*, vol. 46, no. 4, pp. 1103–1127, 2000.

[2] J. Han, K. Ngan, M. Li, and H. Zhang, "Unsupervised extraction of visual attention objects in color images," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 16, no. 1, pp. 141–145, 2006.

[3] U. Rutishauser, D. Walther, C. Koch, and P. Perona, "Is bottom-up attention useful for object recognition?" in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2. IEEE, 2004, pp. II–37.



(a) Input Image (b) Ground Truth (c) LC (d) CA (e) FT (f) RC (g) GC-maps

Fig. 3. Comparison of 4 previous approaches with our algorithm. (a) input images. (b) ground truth images. (c) saliency maps produced using Zhai and Shah [12], (d) Goferman et al.[11], (e) Achanta et al. [5], (f) Cheng et al.[13] region contrast method, (g) our method GC. The uniformly highlighted salient regions is generated by our method.

[4] J. Kim, A. Baik, Y. Jung, and D. Park, “2d-to-3d conversion by using visual attention analysis,” in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2010, pp. 752 412–752 412.

[5] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, “Frequency-tuned salient region detection,” in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 1597–1604.

[6] I. Katramados and T. Breckon, “Real-time visual saliency by division of gaussians,” in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 1701–1704.

[7] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 11, pp. 1254–1259, 1998.

[8] C. Koch and S. Ullman, “Shifts in selective visual attention: towards the underlying neural circuitry.” *Hum Neurobiol*, vol. 4, no. 4, pp. 219–27, 1985.

[9] Y. Ma and H. Zhang, “Contrast-based image attention analysis by using fuzzy growing,” in *International Multimedia Conference: Proceedings of the eleventh ACM international conference on Multimedia*, vol. 2, no. 8, 2003, pp. 374–381.

[10] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum, “Learning to detect a salient object,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 2, pp. 353–367, 2011.

[11] S. Goferman, L. Zelnik-Manor, and A. Tal, “Context-aware saliency detection,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 10, pp. 1915–1926, 2012.

[12] Y. Zhai and M. Shah, “Visual attention detection in video sequences using spatiotemporal cues,” in *Proceedings of the 14th annual ACM international conference on Multimedia*. ACM, 2006, pp. 815–824.

[13] M. Cheng, G. Zhang, N. Mitra, X. Huang, and S. Hu, “Global contrast based salient region detection,” in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 409–416.

[14] H. Permuter, J. Francos, and I. Jermyn, “A study of gaussian mixture models of color and texture features for image classification and segmentation,” *Pattern Recognition*, vol. 39, no. 4, pp. 695–706, 2006.