Brief Analysis on Typical Image Saliency Detection Methods

Wenwen Pan, Xiaofei Sun, Xia Wang, Wei Zhang
College of Information Science and Engineering Zaozhuang University,Zaozhuang,277160,China

Keywords: saliency detection; computer vision; spatial domain; frequency domain； image segmentation.

Abstract. Image saliency detection is becoming more and more important in the field of computer vision. The purpose is to detect salient objects in complex background, at the same time, reduce time complexity to get salient objects with high-resolution and clear edge. In this paper some brief analysis on the state-of-the-art studies relating to image saliency detection has been implemented at first. Next, two kinds of the relating methods has been separated, the former is based on space domain and the latter is based on frequency domain. And then some typical methods respective in space domain and frequency domain are introduced and tested in image databases used for image saliency detection, the experimental results of these method are analyzed with running time and saliency map and the results of image segmentation using fixed threshold and adaptive threshold are compared. Finally, the conclusion and prospects of studies on image saliency detection are proposed.

Introduction

The most important part of the image often focuses on some key small region, namely the so-called salient region. Image saliency detection task is to find out which areas in the image are more likely to become the focus of human visual attention. Image saliency detection can be used for the content transmission, image compression, image segmentation, object recognition, image zoom, image restoration, image editing etc.

This paper studies on image saliency algorithms. First, different state-of-the-art methods were classified and illustrated, and then through the experiment of a typical image database, the experimental results and operation time with different methods were compared. Finally the image saliency detection is summarized and prospected.

The main methods of saliency detection

There is a variety of classification about saliency detection methods. This paper classifies saliency detection methods according to the processing space.

Saliency detection algorithm based on spatial domain

Itti model [2] is a selective visual attention model based on biological visual attention mechanism. Itti model uses intensity feature I=(r+g+b)/3, broadly-tuned color channels R=r-(g+b)/2, G=g-(r+b)/2, B=b-(r+g)/2, Y=(r+g)/2-|r-g|/2-b. Local orientation information is obtained from I using oriented Gabor pyramids O(s, q), where s Θ =[0..8] represents the scale and θ={0,45,90,135} is the preferred orientation.

The image size is an important factor affecting the human visual attention mechanism [45]. The model divides images into 8 scales. Centred-surround is implemented in the model as the difference between fine and coarse scales. The intensity map I is as follows,

$$I(c,s) = \left|I(c)\Theta I(s)\right| \quad (1)$$

The color maps are created is as follows,

$$RG(c,s) = \left|\left(R(c)-G(c)\right)\Theta\left(G(s)-R(s)\right)\right|$$
$$BY(c,s) = \left|\left(B(c)-Y(c)\right)\Theta\left(Y(s)-B(s)\right)\right| \quad (2)$$

Orientation feature map is as follows,

$$O(c,s,\theta) = \left|O(c,\theta)\Theta O(c,\theta)\right| \quad (3)$$

Finally, all feature maps are normalized and combined. The most important contribution of Itti model is the use of feature integration theory and inhibition of return mechanism, and it makes the visual selective attention mechanism used for the engineering application of quantitative calculation [8]. The model intensifies the noise interference when combining the different feature maps.

Goferman proposed image saliency detection algorithm based on context CA [1], which can detect salient region rather than salient object. Divide the input image into blocks, and then calculate the difference between any two blocks,

$$d\left(p_i, p_j\right) = \frac{d_{color}\left(p_i, p_j\right)}{1 + c\cdot d_{position}\left(p_i, p_j\right)} \tag{4}$$

Each pixel block's local and global saliency is calculated,

$$S_i^r = 1 - \exp\left\{-\frac{1}{K}\sum_{k=1}^{K} d\left(p_i^r, p_k^r\right)\right\} \tag{5}$$

CA applies to the case when surrounding environment and the salient object are equally important.

GBVS is based on graph theory. The image is divided into blocks, and the weights between pixel blocks are determined on the basis of visual features such as color, brightness and direction. It assumes the feature map of the input image has been got: M:[n]2→R. Difference between the two point's features is as follows,

$$d\left((i,j)\,\|\,(p,q)\right) \overset{\Delta}{=} \left|\log\frac{M(i,j)}{M(p,q)}\right| \tag{6}$$

Each node of salient map *M* is connected to each other and the weight between the node (*i*, *j*) and node (*p*, *q*) is defined as follows,

$$\omega\left((i,j),(p,q)\right) \overset{\Delta}{=} d\left((i,j)\,\|\,(p,q)\right)\cdot F(i-p, j-q) \tag{7}$$

$$F(a,b) \overset{\Delta}{=} \exp\left(-\frac{a^2+b^2}{2\sigma}\right) \tag{8}$$

The GBVS algorithm has strong robustness, and it gets more obvious salient region. Moreover, AC, HC and RC methods also get good image visual saliency. AC method calculates the difference between a region R_1 and its neighbour R_2, and saliency of different scale adds up to the final saliency. HC method calculates the color histogram and then the saliency map comes into being. RC method improves HC with spatial weight.

Saliency detection method based on the frequency domain

Saliency detection method based on the frequency domain tries to find the relationship between spectrum and saliency features.SR [5] is a typical saliency detection method based on the frequency domain. This method considers there are salient information and redundant information in images. The distribution trends of the log spectrum about different images are similar. The different parts of the log spectrums associate with the salient regions. The 2-D discrete Fourier transform is applied to the input image I(x), and then the image is transformed from the spatial spectrum to the frequency spectrum. Next, the image amplitude is calculated. The spectrum equalization is as follows,

$$A(f) = \Re\left(\Im\left[I(x)\right]\right) \tag{9}$$

$$P(f) = \aleph\left(\Im\left[I(x)\right]\right) \tag{10}$$

$P(f)$ is the phase spectrum, and $\Im(I)$ is the Fourier transform of I. A(f) is spectrum of the Fourier transform about I(x).

$$L(f) = \log(A(f)) \tag{11}$$

$$R(f) = L(f) - h_n(f) * L(f) \tag{12}$$

The log spectrum subtracts the mean filter of itself, and the saliency map is as follows,

$$S(x) = \left|F^{-1}\left[\exp\left\{R(f) + P(f)\right\}\right]\right|^2 \tag{13}$$

SR removes some simple background and obtains relatively satisfactory result. It is more effective than methods based on spatial domain.SR is based on the principle of the constant scale of natural images.

FT [6] is a new saliency detection algorithm of image based on frequency analy-sis. A continuous band of the image is filtered by multiple band pass filters, and all the output are combined into the final result saliency map. FT gets full resolution saliency map. Gauss filter is as follows,

$$DoG(x,y)=\frac{1}{2\pi}\left[\frac{1}{\sigma_1^2}e^{-(x^2+y^2)/2\sigma_1^2}-\frac{1}{\sigma_2^2}e^{-(x^2+y^2)/2\sigma_2^2}\right]=G(x,y,\sigma_1)-G(x,y,\sigma_2) \quad (14)$$

A plurality of filter after superposition is as follows,

$$\sum_{n=0}^{N-1}G\left(x,y,\rho^{n+1}\sigma\right)-G\left(x,y,\rho^{n}\sigma\right)=G\left(x,y,\sigma\rho^{N}\right)-G\left(x,y,\sigma\right) \quad (15)$$

Saliency value is defined as $S(x,y)=\left\|I_\mu-I_{\omega_{hc}}(x,y)\right\|$, this algorithm can get saliency map with clear boundary and homogeneously salient region, and its resolution is same as image, its calculation speed is fast. At the same time, it also emphasizes the biggest salient object, can get the overall contour of salient object, but differences of brightness in salient regions and non salient regions is not obvious.

Achanta et al in 2010 made the improvement to the FT method and put forward MSS method [10]. The method according to the distance between pixel and image edge changes the centred-surround bandwidth, thus with the average of maximum possible symmetry area around to replace the mean of feature vector calculation of whole image in FT algorithm. The saliency value of symmetrical surrounding area around pixel (x, y) in image whose width is w, height is h image is expressed as,

$$S_{ss}(x,y)=\left\|I_\mu(x,y)-I_f(x,y)\right\| \quad (16)$$

The MSS method reduces the saliency of background, but if salient objects in the image being not complete, it will be easy to be considered as background.In addition, Guo proposed PQFT [9], only use the phase spectrum of input image discarding its amplitude spectrum. After Fourier transform to obtain the saliency map similar to the SR. At the same time, the PFT is extended and achieve better performance, but its saliency map only has high saliency degree on the edge being difficult to describe the salient object completely.

Experimental results and analysis

This paper introduces the Itti[2], CA[1], GBVS[3], SR[5], FT[6] and MSS[10] for experiments, calculate 200 images of ImgSal database[11], table 2 shows the average time required of each method, Fig. 1 is the saliency detection results. All of the code used is the Matlab source code, and the running software environment is MATLAB R2010a, the system environment is the Windows XP SP3, memory is 2GB, CPU frequency is 2.1GHZ.

From the results in Table 1, the running time based on frequency domain algorithms are less than that based on spatial domain algorithms.

Table 1 The running time for computing the saliency maps of 200 images

Method	Itti	CA	GBVS	SR	FT	MSS
Time(second)	3.286	114.769	3.202	0.141	0.964	2.391

From experimental results in Fig. 1 we can see that the IT algorithm and GBVS algorithm considering the contrast of local features, some region has a high salient degree relatively but the outline of region is not clear. The FT and MSS considering the overall feature contrast retain the intact outline of salient region while saliency degree of regions is low, not highlighting the most salient position. SR based on overall frequency analysis because not reserving high-frequency information enough lead to the boundary of saliency map being not too clear.

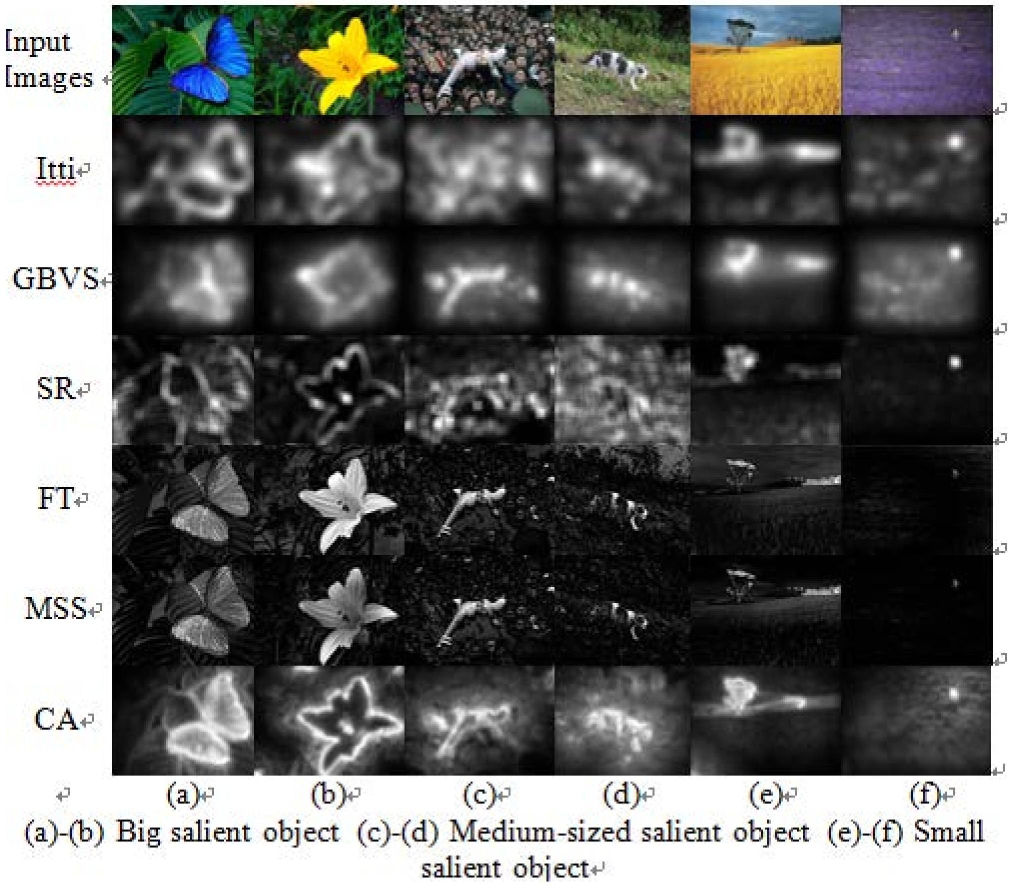

Fig.1 The detection results of various methods

The fixed threshold segmentation

Extract salient object from the saliency map by using the method of global threshold segmentation. The threshold gradually increases from 0 to 255, which is used for segmentation of saliency map. The ROC curves of six methods are computed as shown in Fig. 2.

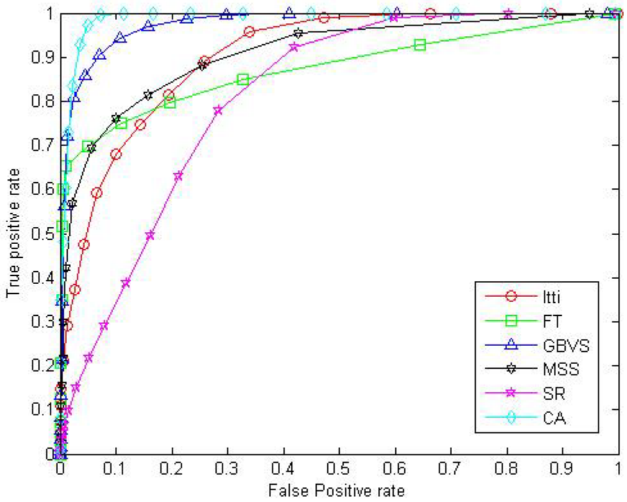

Fig.2 ROC curves

ROC curves show that the CA and the GBVS algorithms have less false positive rate and larger true positive rate, while the SR algorithm's effect is the worst.

Adaptive threshold segmentation

Set the adaptive threshold of the saliency map as：

$$T_{\alpha} = \frac{2}{W \times H} \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} S(x, y) \quad (17)$$

Where W and H are the image's width and height, $S(x, y)$ is the saliency value of the pixel (x, y).When the gray of pixels in the saliency map is larger than the threshold we set to 1, otherwise set to 0, to get a two value image. First, compute the precision P, recall R and F-measurement of each method:

$$P = precision = \frac{sum(S, A)}{sumA} \tag{18}$$

$$R = recall = \frac{sum(S, A)}{sumS} \tag{19}$$

$$F = F - measure = \frac{2 \times P \times R}{P + R} \tag{20}$$

The $sum(S, A)$ is represented as the sum of multiplication of corresponding pixels in saliency map S and artificial segmentation map A. $sum(S)$ and $sum(A)$ are respectively the sum of all pixels gray value in saliency map S and segmentation map A. The larger the value of P, R, F are, the better the saliency map is. Comparison of adaptive threshold segmentation results as shown in Fig. 3.

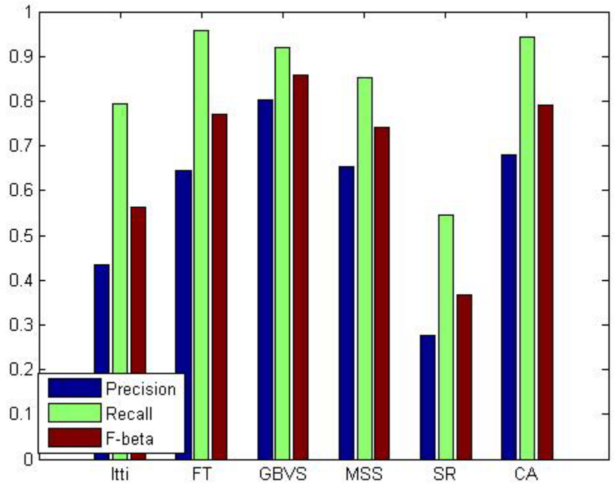

Fig.3 The *P*, *R*, *F* of adaptive threshold segmentation

From the comparison, the results of GBVS and CA are best, FT and MSS take second place, Itti and SR are the worst.

Conclusions and Prospect

This paper focuses on the typical saliency detection methods, through the analysis on the experimental results, advantages and disadvantages of various methods are discussed.

With the in-depth study, image saliency detection for further research includes:

(1) Reduce the time cost. The existing algorithms with high complexity are not suitable for real-time processing.

(2) Pay attention to the development of a top-down algorithm. This paper mainly introduces several important algorithms of bottom-up visual attention mechanism, now the top-down visual attention mechanism research also has some achievements. In the future to pay more attention to the development of top-down visual attention mechanism, and use it to recognize object and make judgment will have important research significance and research space.

(3) Increase the application scope of detection algorithm. For example, the current saliency detection algorithm has achieved good results in small size image processing, but the effect in processing the large size image containing large scene is not satisfying.

(4)The extension of saliency detection methods, Algorithms can be extended to the methods based on signal processing, based on computer vision, based on method of data mining or machine learning.

Acknowledgements

Our work is supported by a project of Shandong province higher educational science and technology program (grant No. J12LN53) and a project supported by youth foundation of Zaozhuang university (grant No. 2011QN43).

References

[1] GOFERMAN S, ZELNIK-MANOR L, TAL A. Context-aware saliency detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10): 1915-1926.

[2] ITTI L, KOCH C, NIEBUR E. A Model of Saliency-based visual attention for rapid scene analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.

[3] HAREL J, KOCH C, PERONA P. Graph-based visual saliency [C]// Proceedings of the 2006 Conference Advances in Neural Information Processing Systems. Vancouver: MIT Press,2006: 545-552.

[4] ACHANTA R, ESTRADA F, WILS P, et al. Salient region detection and segmentation [C]// Proceedings of the International Conference on Computer Vision Systems. Springer, 2008:66-75.

[5] Hou Xiaodi ,Zhang Liqing . Saliency Detection: A spectral residual approach [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Minnesota: IEEE, 2007: 1-8.

[6] ACHANTA R, HEMAMI S, ESTRADA F, et al. Frequency-tuned salient region detection [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Miami:IEEE, 2009: 1597-1604.

[7] TREISMAN A M, GELADE G. A feature-integration theory of attention [J]. Cognitive Psychology, 1980, 12 (1): 97-136.

[8] Qu Yanyun, Zheng Nanning, Li Cuihua et al, Salient building detection based on support vector machine, computer research and development 2007, 44 (1) : 141-147.

[9] Guo Chenlei , Ma Qi , Zhang Liming . Spatio-temporal saliency detection using phase spectrum of quaternion Fourier transform [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Anchorage: IEEE, 2008: 1-8.

[10] CHANTA R, SÜSSTRUNK S. Saliency detection using maximum symmetric surround [C]// Proceedings of the IEEE International Conference on Image Processing. Hong Kong: IEEE, 2010: 2653-2656.

[11] AN X J, LRTINE L, HE H G, et al. ImgSal: A benchmark for saliency detection V1.0 [EB/OL]. (2001) [2013-12-21]. http://www.cim.mcgill.ca/~lijian/database.htm.