AutoSoft®

The Application of Sparse Reconstruction Algorithm for Improving Background Dictionary in Visual Saliency Detection

Lei Feng[1,2], Haibin Li[1], Yakun Gao[1], Yakun Zhang[1]

[1]Yanshan University, School of Electrical Engineering, Qinhuangdao, China;
[2] Xingtai Polytechnic College, China
Full Mailing Address: No. 438, west section of Hebei Street, Qinhuangdao City, Hebei Province, China

ABSTRACT

In the paper, we apply the sparse reconstruction algorithm of improved background dictionary to saliency detection. Firstly, after super-pixel segmentation, two bottom features are extracted: the color information of LAB and the texture features of the image by Gabor filter. Secondly, the convex hull theory is used to remove object region in boundary region, and K-means clustering algorithm is used to continue to simplify the background dictionary. Finally, the saliency map is obtained by calculating the reconstruction error. Compared with the mainstream algorithms, the accuracy and efficiency of this algorithm are better than those of other algorithms.

KEY WORDS: Saliency detection; sparse reconstruction; image features; K-means clustering algorithm.

1 INTRODUCTION

THE visual saliency detection is to enable computers to detect visual attention areas as quickly and accurately as the human visual system. If the computer can accurately detect the visual saliency area, it can effectively allocate the computing resources, focus on the visual saliency area, and greatly improve the efficiency of the computing resources utilization. It originates from the human visual attention mechanism and is dedicated to finding out the most interesting regions of the human eye. The human can easily identify areas of significance in an image and notice the important parts of the image. If a person can focus on an area in an image in a very short time, that area is the salient area of the image. Each image contains one or more salient objects, and saliency detection is to imitate the visual attention mechanism to obtain important information. In the process of the image target detection, image features are usually extracted, and then matched with the corresponding template to determine the target to be detected. If the visual saliency region can be judged beforehand, it can be sampled directly in the visual saliency region when extracting features. In this way, the interference of the non-target regions on the feature extraction is reduced, and the accuracy of the important image processing tasks such as target

detection, recognition and classification are improved. If the function of the visual saliency detection is added to the device, the versatility will be increased, and it will have broad application prospects in the field of intelligent devices.

Koch (1985) put forward the concept of the salient graph and thought that the binary image was used to represent the saliency map. This model has a far-reaching impact on subsequent research algorithms. The Itti model (1998) was based on the human visual cognitive system to simulate human perception characteristics of color, brightness and direction information. At the same time, the model used the difference between the underlying information of these images and the surrounding background area to measure the saliency. The graph-based saliency detection method (2006) used Itti's model framework and basic principles. The Markov random field model was used to simulate the complex attention transfer process in perceiving image information. This method had a strong correlation with the visual attention mechanism and was a very representative saliency detection method. The method proposed by Hou et al. (2007) calculated the saliency maps from the frequency domain perspective according to the idea of the residual spectrum. First, the statistical characteristics of the Fourier spectrum of the natural background was used, and then the saliency map was

calculated by the inverse Fourier transform of the residual spectrum. This method was too simple to detect small targets, because salient objects tend to be sparse in natural scenes and sparse representation has been used in the saliency detection. A context-based model (2014) was the successful application of the iterative multi-scale cumulative process in the field of the visual saliency detection. Huo et al. (2016) used global cues and local objective cues to calculate the saliency of the whole image and candidate objects respectively, and the multi-cue fusion method highlights the saliency of the image and suppresses the interference.

There are two innovations in this paper. First, the texture feature is a basic feature that the human visual system can feel directly. The human visual mechanism is like the method based on the multi-channel filtering, so the Gabor filter is used to extract texture features. The appearance value of the Gabor amplitude feature is transformed into the strength value by the principal component analysis, then it is combined with the LAB color space and the position space to form an eigenvector. Second, the background dictionary is constructed from a single image by using the background priori, but the foreground image is often included in the boundary super-pixel. In this paper, after removing the foreground area, the remaining boundary super-pixels are used as the background dictionary. The redundancy in the background dictionary can be further reduced, and the super-pixels in the template can be more discriminant, so that the foreground of the image is prominent.
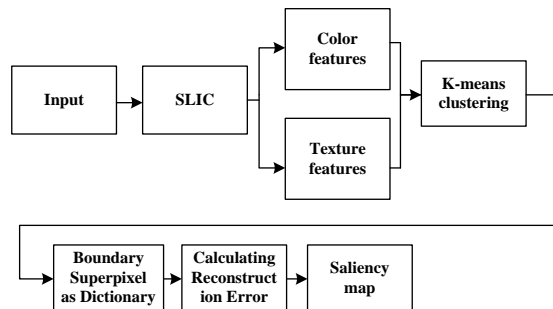

Figure. 1. The Flow Chart of the Algorithm.

2   COLOR FEATURE INFORMATION

THE RGB color space is used in most display devices but the RGB color space has some shortcomings in the color description: (1) The color is linearly combined by R, G and B, so it is difficult to express different colors with precise values. (2) In the RGB color space, the linear combination of the R, G and B results in the high correlation among components. If the brightness of the natural image changes, the three components will change accordingly. (3) The uniformity of the RGB color space is poor, and the sensitivity of the human eye to

the three color components is different. It leads to a large deviation between the color and the human vision.

The LAB color space is formed by the second encoding of the human eye to color. Any value in the Lab color space represents a color, and the value is independent of the device you use. The color space of the LAB forms signals of three new channels: Black white (B-W), red green (R-G) and yellow blue (Y-R). Among them, the B-W channel is named the L channel with a value range of [0-255], and the range of the R-G and Y-R channel is [– 127~128]. Because the use of the RGB and the CIEXYZ color space is complex and not intuitive, the CIElab color space contains almost all the colors that human beings can feel. The XYZ links the RGB and the Lab as intermediate quantities:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.240479 & -1.537150 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2}$$

$$\begin{cases} L^* = 116 f(Y/Y_n) - 16 \\ a^* = 500[f(X/X_n) - f(Y/Y_n)] \\ b^* = 200[f(Y/Y_n) - f(Z/Z_n)] \end{cases} \tag{3}$$

$$f(t) = \begin{cases} t^{1/3} & if \ t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29} & otherwise \end{cases} \tag{4}$$

where, L^*, a^*, b^*, X, Y, Z are the values of their respective color spaces. X_n, Y_n, Z_n generally are 95.047, 100.0, 108.883.

Since in the Lab color space, the color information is contained in a and b components, the color space conversion better enables the color blocks in the image to be clustered in the following k-means clustering analysis.

3   THE TEXTURE FEATURE EXTRACTION

3.1   The Gabor Filter

THE Gabor wavelet transform simulates the human visual system. It magnifies the image feature (local texture information). It measures the direction and frequency characteristics in the spatial and frequency domain. In the image processing, the Gabor wavelet transform cannot only carry out multi-scale and multi-directional analysis to capture the local structure information of the image but also has good robustness to illumination change and image deformation. Moreover, the Gabor wavelet has low sensitivity to

light changes and can adapt better to the phenomenon of uneven light in the image above the characteristics proving that the Gabor wavelet can be widely used in the visual system, making the two-dimensional Gabor filter a tool for the multi-scale and multi-direction representation of images. It realizes the optimal time of the signal, and the spatial position, frequency, direction, phase and the phase bandwidth of the two-dimensional signal can be obtained from the image processed by the two-dimensional Gabor filter. The texture features are filtered as approximate periodic signals. The Gabor filter can extract high quality local information and frequency domain information of targets. Its formulaic expression is as follows:

$$g(x, y, \lambda_1, \theta, \sigma_1, \gamma) = \frac{1}{2\pi\sigma_{x1}\sigma_{y1}} \exp\left(-\frac{x^2 + \gamma^2 y^2}{2\sigma_1^2}\right) \exp\left(i\frac{2\pi x_\theta}{\lambda_1}\right) \tag{5}$$

where, λ_1 represents the wavelength of the Gabor function, and θ represents the filtering Angle of the filter. σ_1 is a constant, usually set to half the size of the filter, and γ denotes the aspect ratio and determines the ellipticity of the Gabor function shape.

3.2   The Gabor Filter Parameter Selection

The two-dimensional Gabor filter belongs to the band pass filter. It can describe the local information of the spatial location, scale selectivity and direction selectivity with high quality. The two-dimensional Gabor filters generate a two-dimensional Gabor wavelet, which can extract texture features of images through scaling and rotation. The selection of the filter parameters is the focus of the feature extraction. For this reason, we use the Gabor filter Banks with multiple center frequencies and directions to describe the image. It can be seen from the above that the effect of two-dimensional Gabor filter is determined by the parameters of image direction and scale. In this paper, the Gabor filter is applied to the response graph of the direction $\theta = \{0, \pi/4, \pi/2, 3\pi/4\}$ at five different scales. Figure 1 shows the frequency real part, imaginary part, amplitude and phase information of Gabor kernel function (2008).
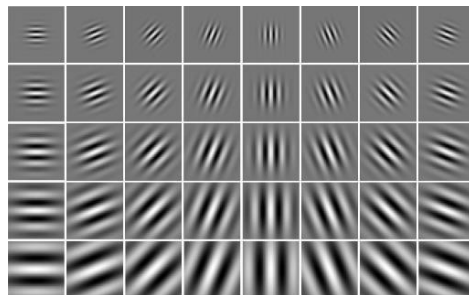

Figure. 2. The Real Part of the Gabor Filter the Kernel Frequency Real Part, Imaginary Part and Feature Images.
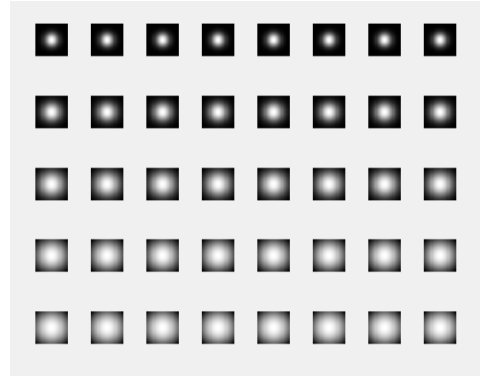

Figure. 3. The Amplitude of The Gabor Filter

The Figure 2 and 3 are the real part and the amplitude of the Gabor feature images. The two-dimensional Gabor filter has unique properties of spatial locality, spatial frequency and directional selectivity. Each Gabor kernel function can extract significant features of edges and local textures of texture images at different frequencies and directions. According to theoretical analysis, the more filtering directions or scale levels are selected, the better the final effect will be. However, the number of filters will increase, and the real-time performance will be reduced.

In order to show more intuitively that the Gabor filter can effectively extract the texture features of the image in different directions, this paper selects the images with the obvious texture directions from the Brodatz texture library. In Figure 2, the figure is the extraction effects of the texture features in their respective directions.
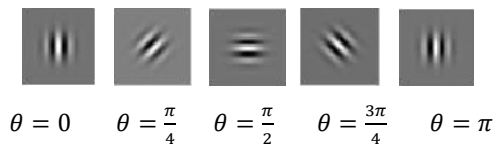

$$\theta = 0 \quad \theta = \frac{\pi}{4} \quad \theta = \frac{\pi}{2} \quad \theta = \frac{3\pi}{4} \quad \theta = \pi$$

Figure. 4. The Texture Feature Information Extraction in Different Directions

3.3   From the Gabor Amplitude Image to the Gabor Characteristics

(1) Each Gabor amplitude image contains some local variations. The simple Gabor low-pass filtering can be used to compensate for these changes to the smooth Gabor amplitude information.

(2) The function set of the spatial information includes not only the data points represented by the pixels but also the features represented by the plane. Each filter in the Gabor filter bank has a separate function and the two additional functions of the spatial information. Finally, each pixel in the input image has 24 Gabor features and 2 spatial features.

(3) The normalization of the features contained in each pixel are obtained. In order to visualize the feature and simplify the calculation, the multi-

dimensional array representing each pixel feature is transformed into a numerical value by the principal component analysis.

4   THE SPARSE RECONSTRUCTION ERROR

4.1   The Mathematical Form of the Sparse Reconstruction

THE standard sparse coding model is used to simulate the human nerve cells. Each element in the image is represented as a linear combination of a set of dictionary basis functions. Let $x \in R^m$ represent the feature descriptor of a sample and U represent the dictionary of an image. As a dictionary, it is generally required that the number of bases should be at least greater than the dimension of the elements in the sample. Under the given conditions, the coding of the sample υ is as follows:

$$\min \|x - Uv\|^2 + \lambda \|v\|_0 \qquad (6)$$

$$\text{s.t.} \ \|u_i\|^2 \leq 1$$

where λ is the equilibrium factor of the two equilibrium terms. $U = [u_1, u_2, \cdots, u_n] \in R^{n \times k}$, u_i is the base of a dictionary. $\| * \|$ represents the two norms, and $\|*\|_0$ denotes the zero norm.

The first item of the formula is to describe the reconstruction error of the original image block and the reconstructed original image block. If its value is small, it shows that the image block reconstructed by the basis function and the coding coefficient can approximate the original image block. This item describes the quality of the coding results. The second item is the sparse item, which represents the sparse degree of coding. The value of λ determines the sparsity of coding υ.

Suppose an image is divided into n super-pixels $\{x_1, x_2, \cdots, x_n\}$, then they are divided into S segments $\{S_1, S_2, \cdots, S_n\}$. The super-pixel of the image boundary is coded as a dictionary, and its sparse coding u is calculated by the formula. The reconstruction error of this super-pixel is as follows:

$$e_i = \|x_i - U_k v_i\|_2 \qquad (7)$$

The algorithm uses the sparse reconstruction error to solve the salient target detection problem. The basic idea of the algorithm is that the reconstruction error of each image region is proportional to the saliency value of the image region when the background template is chosen as the base vector in the reconstruction process. For equation (6), a series of sparse coefficients are solved to minimize the corresponding reconstruction error of equation (7). The reconstruction error shows the significant score of the target area, to assist the detection of the significant objects and achieve the suppression of the foreground noise. In the process, every pixel region is generated by the SILC segmentation, and the n-dimensional feature representation x is extracted. Then the image boundary area is selected as the background template B, and a sparse coding framework is established based on this. For the i-th region, the background template B is used as the base vector for the sparse representation according to formula (6), and then the sparse reconstruction errors of each region are calculated according to formula (7). For this region, if the reconstruction error calculated by sparse representation of each background template is small, it is like the background and is not salient. For the same set of background templates, the reconstruction errors calculated from the object and background regions are generally quite different. For the relatively complex background in the image, the sparse representation can express the difference. In this paper, the sparse reconstruction error is used to measure whether a super-pixel region is a saliency region and quantify its saliency value.

The larger the reconstruction error of a super-pixel, the more it tends to be part of the salient target. Its segmentation tends to be more salient. If some foreground segmentation is grouped into the background (salient objects appear in the edge of the image), the salient regions will be mis-detected as the background.

4.2   The Selection of the Dictionary

In this section, the color-enhanced Harris corner operator is used to detect the corners of the potential target areas in the image. These salient points can confirm the spatial location of the salient targets. These saliency points are encapsulated in the convex hulls to estimate the saliency target regions. Since the saliency objects appear at the edge of an image, the accuracy of saliency detection will be reduced if all the edge regions of the image are used as the background dictionaries. After modifying the part of the boundary area, they are used as a background dictionary. It guarantees more stable and reliable background information. The redundancy in the background dictionary can be further reduced after the boundary super pixels are clustered.

In this paper, the super-pixels contain many pixel units, and the average CIELAB color features. The Gabor texture features, and location information of these pixel units constitute the characteristics of the super-pixels, $sf_i = \{l, a, b, G, x, y\}$. The K super-pixel is recorded as $SK = \{sf_1, \dots sf_i, \dots, sf_k\}$. The back ground template $BG = \{Bg_1, \dots Bg_i, \dots, Bg_l\} \in R^{D*L}$ is constructed by selecting the super-Pixel at the image boundary from the SK. Although the background template is still used as a dictionary for sparse representation, the background template is pre-processed before the sparse reconstruction. The image boundary can be divided into four directions; upper and lower, left and right. Because the super-pixel features of the image on each boundary may have great similarity, if not processed, the uniqueness of the

super-pixel in the background template is not obvious. This will lead to redundancy in the sparse dictionaries, so this paper uses the background template pre-processing to reduce the redundancy in dictionaries. The background templates are classified into four categories according to the four directions. The similar super-pixels of each boundary are merged by the k-means algorithm, then the merged super-pixel features are represented by the feature mean of the merged super-pixels. The new background template is recorded as $BG' = \{Bg'_1, ... Bg'_i, ..., Bg'_M\} \in R^{D*L}$. Bg'_i represents the eigenvector of the second feature vector in the sparse dictionary. The basic idea of the sparse representation is to represent as much data as possible with as few dictionary features as possible. After pre-treatment, the super-pixels in the background template are more discriminant. The reduction of the super-pixel feature in the template means the reduction of the sparse dictionary feature, and it effectively solves the redundancy problem of the dictionary and improves the efficiency of the sparse reconstruction error calculation.

4.3 The K-means Algorithm

Because of its unsupervised and fast operation speed, the K-means algorithm is widely used. The data is clustered into different clusters in the algorithm according to the distance difference between the data. There is an observation set $(x_1, x_2, ..., x_n)$, and each observation is a multi-dimensional vector. The observation values are clustered in the K sets, so formula (8) calculates the closest class to each sample.

$$\arg\min_S \sum_{i=1}^{k} \sum_{x_j \in s_i} \|x_j - \mu_i\|^2 \tag{8}$$

where, $S = (s_1, s_2, ..., s_k)$, and μ_i is the average value in s_i.

The iterative process of the K-means algorithm is as follows:

1. First, $(\mu_1^{(1)}, \mu_2^{(1)}, ..., \mu_k^{(1)})$ is the initial set of the k-means and assigning each observed value to the class. Calculate the class it should belong to, as shown in the formula 9:

$$s_i^{(t)} = \left\{ x_p : \|x_p - \mu_i^{(t)}\|^2 \le \|x_p - \mu_j^{(t)}\|^2 \ \forall j, 1 \le j \le k \right\} \tag{9}$$

2. The centroid of each class j can be recalculated by

$$\mu_i^{(t+1)} = \frac{1}{|s_i^{(t)}|} \sum_{x_j \in s_i^{(t)}} x_j \tag{10}$$

formula

where, $\left|s_i^{(t)}\right|$ is the value of $s_i^{(t)}$.

The initial central point moves to the position of the average value of the data, and after several iterations, when the central point does not change any more, the clustering center is obtained.

The K - means algorithm
Input: number of clustering K, and data samples N
Output: K clusters satisfying the minimum standard deviation
1. Randomly select the K objects from n samples as the initial clustering center.
2. While (if the clustering changes).
3. Use formula (10) to calculate each cluster center.
4. Use formula (9) to calculate the clustering of each sample and its clustering center and re-classify the samples according to the minimum distance.
5. End.

The k-means algorithm satisfies formula (9), which obtains the minimum value and ends the algorithm. Because of the association between the image features, in the process of the feature clustering，the saliency details are retained, and the saliency information is extracted.

5 EXPERIMENT

IN this paper, the super-pixel is used as the basic processing unit, and the SLIC method (2010) is used to simplify the image structure information and preserve the edges of the salient map objects. The image block is represented by the average color feature, texture feature value and position coordinate of all the pixels contained in the super-pixel. The features of the super pixel are composed of color features, texture features and location, $\mathbf{x} = \{L, a, b, G, x, y\}^T$. The image features are represented by the N super-pixel features $\mathbf{X} = [x_1, x_2, \cdots, x_N] \in R^{D \times N}$, where D denotes the feature dimension.

5.1 The Algorithm Comparison and Analysis

In this paper, the MSRA-1000 data set is used for the experiment. The MSRA dataset is provided by the Asian Research Institute. Its subset is one of the benchmark datasets widely used by many saliency detection algorithms. The database was first compiled by Liu et al (2012). Most of the image resolutions in this database is 400*300, and the saliency region is labelled with a rectangular box, Cheng et al. (2014) simplified the MSRA data set, and 10000 images are randomly selected by the precise manual labelling. It referred to as the MSRA-10K. The images in the data set (Figure 5) has the characteristics of the complex background and low contrast of the target area. Many representative methods were chosen to test on the data set.
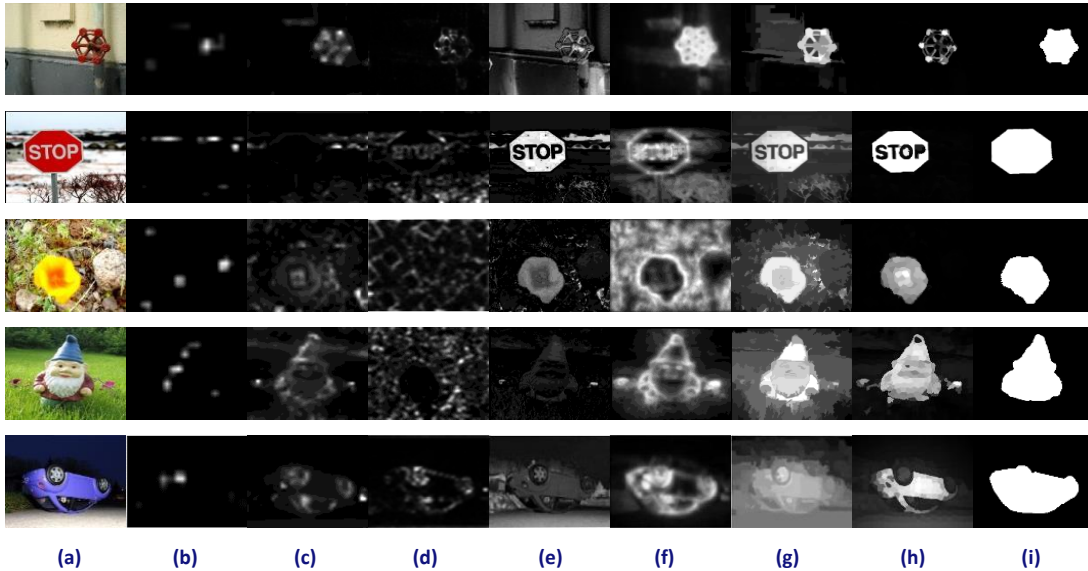
(a) The MSRA-10K original Map

(b) The MSRA-10K Manual Labelling Map

Figure 5. The MSRA-10K Original and Manual Labelling Maps.

The experimental results of Figure 6 show several classical algorithms, (IT, GB, SR, FT, CA, RC) for comparison. By comparing with the manual annotated binary images, we can visually see whether the algorithm improves the accuracy of the saliency detection. The validity of the algorithm can be quantitatively verified by the recall rate and the F measure. Figure 4 intuitively shows the results of the algorithm from a qualitative point of view. Section 5.2 will select several classical indicators to compare the results from a quantitative point of view.

(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)

Figure. 6. The Saliency Map Comparison of Seven Models.((a)Input image; (b) IT;(c) GB;(d) SR; (e) FT;(f) CA; (g) RC;(h) Ours;(i) Ground truth)

5.2 The Performance Comparison and Analysis

In Figure 4, the sample image is used for the binarization and the acquired binary image is used as the quasi-basis of the saliency region. Precision (P) denotes the proportion of the salient regions detected by the algorithm, recall rate (R) denotes the matching degree between the salient regions of the algorithm and the salient regions of the truth maps, and the F-measure is a comprehensive index. The aaccuracy represents the proportion of the salient pixels in the whole region detected by the saliency detection algorithm. The recall rate represents the degree to which the number of salient pixels detected by the saliency detection algorithm agrees with the salient region in the ground truth (GT) of the manual labelling. The F-measure is a harmonic average obtained by combining the accuracy and recall, which reflects the comprehensive quality of the saliency detection. The formula is as follows:

$$P = \frac{\sum_{(x,y)} S(x,y)B(x,y)}{\sum_{(x,y)} S(x,y)} \tag{11}$$

$$R = \frac{\sum_{(x,y)} S(x,y)B(x,y)}{\sum_{(x,y)} B(x,y)} \tag{12}$$

$$F = \frac{(1+\beta^2) \times P \times R}{\beta^2 \times P + R} \tag{13}$$

where, $S(x,y)$ is the binary saliency graph of the algorithm in this paper, $B(x,y)$ is the manually calibrated binary graph, and β is the index of sample 1, and sample 2 by taking 0.2. The value of the F-measure is determined by the uniformity of the background texture in the experimental image.

The saliency map is the set to a threshold value of $T_f \in [0, 255]$ for segmentation, and the salient map after the two values can be obtained. In order to fully compare the effect of the saliency detection of different saliency detection methods, the threshold

value of the threshold T_f is changed between $[0, 255]$, and the dynamic threshold segmentation of the salient map is achieved to draw the PR curve. Figures 6 and 7 show the obtained accuracy and recall curves. Most of the time, high accuracy and high recall rate needs to use the F-measure as the overall performance evaluation mechanism.

In this paper, the average value of three indicators of the experimental sample graph for statistical analysis is calculated, and their mean values are 85.92%, 76.83%, 0.8526.

Three performance indicators of the Binary Images of the binary images under their respective thresholds were obtained by experiments on 1000 images in the MSRA test database. Figure 6 to 8 show three performance indicators of each algorithm.

Considering the above three objective evaluation criteria, the proposed algorithm is superior to other algorithms. As seen from Figure 5, under the same recall rate, the accuracy of this algorithm is higher than other algorithms. The false positive rate is also a measure index. From this index, the algorithm is superior to other methods, as shown in Figure 6. In Figure 7, the accuracy of the RC method is not much different from that of the proposed algorithm, but the recall rate is low. Because the saliency maps of some objects are enhanced by the RC method, compared with the RC method, the proposed algorithm can also guarantee a higher recall rate while guaranteeing a higher accuracy. The proposed algorithm makes the detection results of the saliency areas improve significantly. It is more consistent with the human visual perception and can highlight the saliency areas.

The robustness of the algorithm is also an important index. The box graph demonstrates the robustness of the algorithm and it is introduced as one of the evaluation criteria in this paper. By finding out the outliers, it directly reflects the symmetry, discreteness and anomaly of the data, and facilitates the comparison of multiple groups of data.
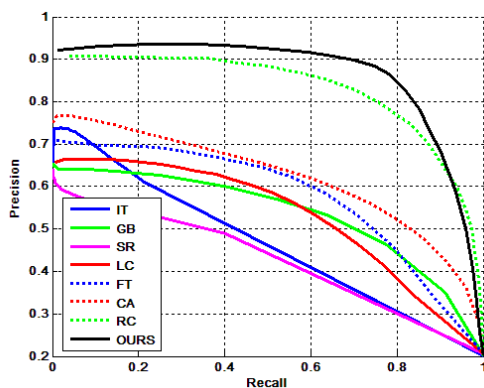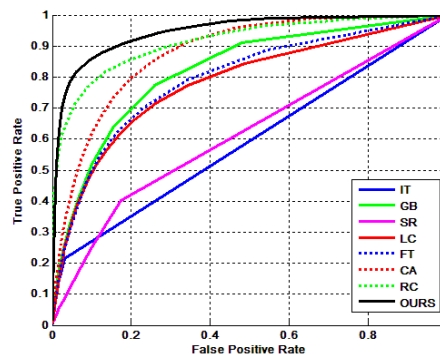

Figure 8. **The Receiver Operating Characteristic Curves of the Saliency Detection Algorithms.**
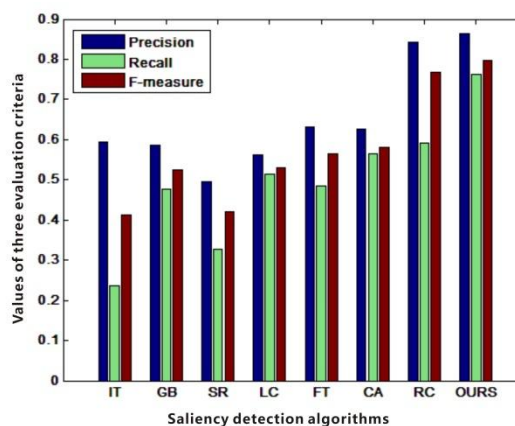

Figure 9. **The F-measure Bars of the Saliency Detection Algorithms.**

6   CONCLUSION

THIS paper introduces a sparse reconstruction algorithm of the improved background dictionary. Considering the correlation of the information inside the image, the boundary super-pixels are clustered according to the feature vectors of the super-pixels. The new boundary super-pixel is used as the base template to calculate the sparse reconstruction error. The algorithm further improves the saliency detection index and reduces the interference of the background area on the feature extraction. According to the experimental comparison and data analysis, the saliency detection method based on the clustering analysis is better than the saliency detection method listed in this paper. In the next stage of the research, we plan that our approach is combined with the top-down technique to improve the efficiency and accuracy in the saliency detection.

7   DISCLOSURE STATEMENT

NO potential conflict of interest was reported by the authors.

Figure 7. **The Precision-Recall Curves of the Saliency Detection algorithms.**

8   REFERENCES

Arunkumar, R., and Nagaraj Balakrishnan. "Medical image classification for disease diagnosis by DBN methods." Pak J Biotechnol 15, no. 1 (2018): 107-110.

Bangalore, S.M., Jens-Rainer, O., Vinod, V.V., & Akio, Y. (2001). Color and Texture Descriptors. IEEE Trans on Circuits and Systems for Video Technology, 11(6), 703 -715.

Bartkevicius, S., Fiodorova, O., Knys, A., Derviniene, A., Dervinis, G., Raudonis, V., Lipnickas, A., Baranauskas, V., Sarkauskas, K., Balasevicius, L. (2017) Mobile Robots Navigation Modeling in Known 2D Environment Based on Petri Nets, Intelligent Automation & Soft Computing, DOI: 10.1080/10798587.2016.1264695

Sbaiz, L., Vandewalle, P., & Vetterli, M. (2008). Groebner basis methods for multichannel sampling with unknown offsets. Applied and Computational Harmonic Analysis, 25(3), 277-294.

Cheng, M.M., Mitra, N.J., Huang, X., Torr, P.H., & Hu, S.M. (2015). Global Contrast Based Salient Region Detection. In: Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(3), 569-582.

Costantini, R., Sbaiz, L., & Susstrunk, S. (2008). Higher order svd analysis for dynamic texture synthesis. IEEE Transactions on Image Processing, 17(1), 42-52.

Goferman S., Zelnik-Manor, L., & Tal, A. (2012) Context-aware saliency detection. In: Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(10), 1915-1926.

Gopalakrishnan, V., Hu, Y.Q., & Rajan, D. (2010) Random walks on graphs for salient object detection in images. IEEE Transactions on Image Processing, 19(12), 3232-3242.

Harel, J., Koch, C., & Perona, P. (2007). Graph-Based Visual Saliency. Advances in Neural Information Processing Systems, 19, 545-552.

Alleysson, D., Meylan, L., Süsstrunk, Sabine. (2006). HDR CFA Image Rendering. In: Proceedings of the European Signal Processing Conference.

He, K., Sun, J., & Tang, X. (2013). Guided Image Filtering. IEEE Transactions on Software Engineering, 35(6), 1397-1409.

Hou, Q.B., Cheng, M. M., Hu, X. W., Ali, B., Tu, Z.W., & Philip, T. (2018). Deeply supervised salient object detection with short connections. In: Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.1-14.

Huibin, W., Shengnan, Z., & Xin, W.（2013）.An Approach for Target Detection and Extraction Based on Biological Vision. Intelligent Automation & Soft Computing, 17(7), 909-921

Huo, L., Jiao, L., Wang, S., Yang, S. (2016). Object-level saliency detection with color attributes. Pattern Recognit, 49, 162－173.

Itti, L., Koch, C., Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. In: Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11), 1254-1259.

Joost, V.D.W., Rein, V.D.B. (2001). Local Mode Filtering. In: Proceedings of the 2001 IEEE Computer Society Conference, 422(2), 428-433.

Khelifi, F, Jiang, J.M. (2011). K-NN regression to improve statistical feature extraction for texture retrieval. IEEE Transactions on Image Processing, 20(1), 293-298.

Koch, C., Unman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. Hum Neurobiol, 4(4), 219-227.

Ma, Z.Y., He, K.M., & Wei, Y.C. (2014). Constant Time Weighted Median Filtering for Stereo Matching and Beyond. In: Proceedings of the IEEE International Conference on Computer Vision, pp 49-56.

Philippe T, Daniel S, & Michael U. (2012). Bi-Exponential Edge-Preserving Smoother. IEEE Transactions on Image Processing, 21(9), 3924−3936.

Schillaci, G., Bodiroza, S., & Hafiier, V. (2013). Evaluating the effect of saliency detection and attention manipulation in human-robot interaction. International Journal of Social Robotics, 5(1), 139-152.

Tamburrino, D., Dicarlo, J.M., Rodricks, B.G., Alleysson, D., Meylan, L, & Süsstrunk, S. (2008). Digital Camera Workflow for High Dynamic Range Images Using a Model of Retinal Processing. In: Proc. IS&T/SPIE Electronic Imaging 2008: Digital Photography IV, pp. 1-12.

Wang Y., Peng G., & Zhou M. (2017). Saliency detection by hierarchically integrating compactness, contrast and boundary connectivity. Multimedia Tools and Applications, 77(10), 11883–11901.

Zhang B.Y., Jan P.A. (2008). Adaptive Bilateral Filter for Sharpness Enhancement and Noise Removal. IEEE Transactions on Image Processing, 17(5), 664-678

9   NOTES ON CONTRIBUTORS
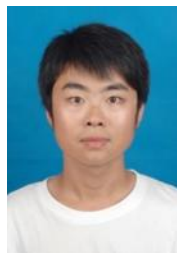
Lei Feng received B.E. and M.E. degrees from the Department of Electrical Engineering, Yanshan University, Qinhuangdao, China, in 2008 and 2011, respectively. He has been teaching at Xingtai Polytechnic College since 2011. Since 2016, he has studied for a doctorate degree at the Electrical Engineering College of Yanshan University. His main research fields include image and video processing, image segmentation, visual saliency detection.

INTELLIGENTAUTOMATIONANDSOFTCOMPUTING 839

Haibin Li received B.S. degree in Automation from the YanShan University, Qinhuangdao, China, in 1999. In the same year, he earned his Master's degree in Control Theory and Control Engineering and obtained M.S. degree in 2002. In 2002, he continued to pursue doctoral at the Department of Electrical Circuits and Systems. During this period, he held many provincial and ministerial level topics, published several high-level papers, and received a doctorate degree in 2006. He is currently a Professor with the College of Electrical Engineering, Yanshan University, Qinhuangdao. His main research areas include image processing, stereo vision, 3D reconstruction and deep learning.

Yakun Gao has studied for master's degree and doctorate degree in control science and engineering at the school of electrical engineering, Yanshan University since 2013 During the period, he participated in many provincial and ministerial projects and more than 10 enterprise projects. His main research fields include image and video processing, stereo vision, 3D reconstruction.

Yakun Zhang has studied for a doctorate in the school of electrical engineering since 2017. During this period, she participated in many provincial and ministerial projects and more than 10 enterprise projects. Her main research interests include computer vision, pattern recognition and stereovision.