International Journal of Engineering and Applied Sciences (IJEAS)
ISSN: 2394-3661, Volume-5, Issue-3, March 2018

SURF feature extraction algorithm based on visual saliency improvement

Zongyuan Zhu, Guicang Zhang, Hongjie Li

Abstract— **Feature extraction is an important link in image retrieval and image matching.Aiming at the problem of the traditional feature extraction method, which is too simple to extract valid dimension and feature points, a SURF weight algorithm combining visual significance and improvement is proposed to extract the key points:. SURF algorithm of image is utilized to extract the key point, and then through improved significant find significant area detection method, key points can be divided into two parts, the significant region and significant area, external point by weighting algorithm to judge the importance of structure information, thus retaining structure information important points.The experimental results show that the feature points extracted by this method are more comprehensive and improve the accuracy of image matching.**

***Index Terms*— visual significance; Speeded-Up Robust Features(SURF); feature extraction; image retrieval; weight;**

I. INTRODUCTION

Image feature extraction is a highly interdisciplinary subject.Almost as long as computer vision, it plays a vital role in image processing and computer field[1].The results of feature extraction also directly affect the effect of image processing [2].The study of feature extraction has two main purposes: first, the most important description of the difference between the target and the target. The second is to reduce the dimension of the target data under certain circumstances.With the development and popularization of computer and network, image feature extraction has been widely used in various fields of life.For example, SURF algorithm is used for target tracking[3], face tracking recognition[4], hand vein recognition [5], image registration and panoramic splicing [6].Image features can be roughly divided into three categories. The first type of image feature is to analyze the single image and extract the dominant features of the image, such as the shape, texture , color and other information;The second category is extracted from the overall analysis of the commonality and the opposite of the image data;The third is an image semantic feature that can be perceived directly or indirectly by humans.These three features are not conflicting, and in many cases, they can even be used together.

SURF(Speeded-Up Robust Features) is a new algorithm proposed byHrebert Bay[7] in 2008.Since this operator describes the local texture features of key points in different directions and scales of the grayscale image, it maintains invariance to rotation, scaling, and brightness changes.it also has good stability for affine transformation and noise, and it is more than other methods in the aspects of uniqueness and robustness, and it greatly improves the computational efficiency.Therefore, literature [8] applies SURF to the matching of PBC images and uses the points of interest to extract PBC images for image matching.The operator is robust to image rotation and image repetitive modes, reducing the complexity of PBC production.Bipblab Banergee[9] used SURF for automatic classification of image targets, and the combination of SURF operator and SVM kernel effectively improved the accuracy of image classification.[10] applied SURF to trademark image retrieval, literature [11] and [12] respectively used SURF for face recognition and iris recognition.

On this basis, in order to improve the feature extraction effect of color image,For the image into a gray image due to SURF operator ignore the image darker areas in the process, resulting in a decline in feature detection quantity, with the increase of image scale space, the characteristic quantity of detection has not increased effectively，but the feature detection time is increased.Some scholars proposed the C-SURF feature detection method [14] that introduced color invariants [13]. Shi Yasun [15] applied color invariants [13] to the SURF feature point extraction algorithm in his research.Because the color invariant can retain the color information of the image, the detection effect of the C-SURF feature point is obviously enhanced, and the algorithm speed is obviously accelerated.In the literature [16], Yang fan proposed a histogram equalization process, and reconstructed the r-surf algorithm of scale space.Improves the number of image feature detections and matches while maintaining a high match rate and inherits the good characteristics of the SURF algorithm. In particular, it has a strong robustness to changes in light and viewing angles, providing a new and effective solution to image feature extraction.

This paper presents a new feature extraction method that combines visual saliency analysis with improved SURF algorithm.Firstly, SURF algorithm is used to extract image features, and then the improved visual significance algorithm is used to analyze the image,and extracts the area of significance,The key points in the area of significance are important.In order not to miss the key points and make the extracted points not too numerous,The point outside the key region is judged by the weight algorithm to determine the importance of its structural information, thereby eliminating the non-essential points and using this method to improve the accuracy of retrieval and recognition.

II. ALGORITHM PRINCIPLE

2.1 Visual significant

Human beings have a strong sense of visual perception[17], which can quickly lock the objects of interest in images in the preprocessing stage, namely the elements that attract visual attention in the scene.The researchers did a lot of work to make the computer simulate the human visual system. Neurobiology and psychology divide the mechanism of

Zongyuan Zhu, Guicang Zhang, Hongjie Li, College of Mathematics and Statistics, Northwest Normal University, Lanzhou 730070, China

visually significant into two aspects[18]: selective attention from the bottom up and selective attention from the top down. Neurobiology and psychology divide the mechanism of visually significant into two aspects [18]: selective attention from the bottom up and selective attention from the top down.The bottom-up mechanism is a fast, environment-independent process; the top-down mechanism is a slow process that depends on visual processing tasks.This article focuses on the bottom-up visual saliency detection (FT visual detection algorithm).

2.2 Improved SURF feature extraction algorithm

SURF algorithm mainly consists of the following steps[19]: (1) construct Hessian matrix; (2) scale space generation; (3) use non-maximal suppression to initially determine feature points and then accurately locate feature points; (4) select main direction of feature points; (5) construct feature point descriptors .

In order to extract feature points effectively, this paper combines the FT saliency analysis method and weight algorithm to give an improved SURF feature extraction method. Proceed as follows:

(1) construct Hessian matrix;

(2) Scale space generation;

(3) Using the non-maximal suppression to initially determine the feature points and then accurately locate the feature points;

(4) Using the improved FT algorithm to find all salient regions in the image;

(5) Calculate the proportional weights of feature points outside the significant region;

(6) Extract the SURF descriptor of the selected key point.

In this paper, we mainly study the fourth stage of the original SURF algorithm. For the detected feature points, we use the improved FT detection method and feature point weight algorithm to eliminate feature points that are not important for structural information, so as to extract more representative key points.

2.2.1 Hessian matrix

Hessian matrix is the core of SURF algorithm. The local maximum of its determinant can determine the position and scale of feature points[20]. The SURF algorithm obtains stable points by using the Hessian matrix to find the extreme points, and uses the maximum value of the matrix determinant to mark the position of the blob-like structure.

2.2.2 SURF Scale Space and Feature Extraction

The scale space of the image is a representation of an image at different resolutions. It can be realized by using the convolution of the Gaussian kernel $\frac{\partial^2}{\partial x^2} g(t)$. The scale of the image is generally expressed by Gaussian standard deviation[21].

In the field of computational vision, The scale space needs to repeatedly convolute the functions of the input image with the kernel of the Gaussian function and repeatedly perform sub-sampling on it, which is symbolically represented as an image pyramid.SURF algorithm is used to process the box filter with different sizes.allowing the scale-space multi-layer image to be processed at the same time,only change filter size don't need to carry on the second sampling,thus improve the algorithm performance.

The SURF feature extraction uses interpolation techniques to find Space and size positions on the sub-pixel precision, which can be obtained by the ternary quadratic equation proposed by Brown and Lowe[22].Assume Hessian's determinant function is noted a $H(x,y,s)$, and $x=(x,y,s)^{\mathrm{T}}$,

According to the Taylor expansion you can get:

$$H(x)=H+\frac{\partial H^{\mathrm{T}}}{\partial x}x+\frac{1}{2}x^{\mathrm{T}}\frac{\partial^2 H}{\partial x^2}x \quad (1)$$

Get the extreme value $\hat{x}=(x,y,s)$ of the interpolation region by the derivative of $H(x,y,\sigma)$. When $H(x,y,\sigma)=0$, you can get:

$$\hat{x}=-\frac{\partial^2 H^{-1}}{\partial x^2}\frac{\partial H}{\partial x} \quad (2)$$

The derivative of the function can be approximated by the difference between adjacent pixels. If $\hat{x}$ is greater than 0.5 in the three directions of x , y and σ , then the position of the feature point needs to be adjusted and the interpolation algorithm used again until $\hat{x}$ is less than 0.5 in all directions, or the number of preset interpolation algorithm uses overflows.

1) 2.2.3 FT visual significance test

Achanta[23] proposed Frequency-tuned Salient Region Detection ,abbreviated as FT algorithm. This article uses the improved FT algorithm for significance testing.The FT algorithm uses a Gaussian differential filter method, which is implemented by selecting a certain frequency range through classification.First, Gaussian filtering is performed on image $\boldsymbol{T}$ of size $M\times N$ using equation (3) to obtain a smooth image and eliminate noise, and a filtered image T_N is obtained.

$$T_N=T*G \quad (3)$$

Where G represents a Gaussian filter.Then the image is converted to Lab space to obtain the brightness feature L and the color features a and b of the image. Then calculate the spatial distance between each feature and the average value of the image:

$$S_L(i,j)=\left\|L(i,j)-\frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}L(i,j)\right\| \quad (4)$$

$$S_a(i,j)=\left\|a(i,j)-\frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}a(i,j)\right\| \quad (5)$$

$$S_b(i,j)=\left\|b(i,j)-\frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}b(i,j)\right\| \quad (6)$$

Finally, feature fusion is used to obtain the significance value of each pixel:

$$S(i,j)=S_L(i,j)+S_a(i,j)+S_b(i,j) \quad (7)$$

The value obtained by the above equation is normalized to a real number between [0 1].

2.2.3.1 Improved FT visual significance test

Since the original FT algorithm uses a Gaussian filter to filter the input image, the Gaussian filter simply performs denoising adjustment on the image and then outputs the filtered image. This article uses image-guided filters [24] for the filtering part of the image.Compared to Gaussian filtering, guided filters can be used as local estimates based on the mean and variance of the pixels in the image neighborhood. so as to adaptively adjust the output image weight according to the information of the image. Equation (3) then becomes:

$$T_N = T * W_{guide} \quad (8)$$

Due to the space for traditional images，The range and speed of change of the three characteristics L, a, b are different[25].It is likely that the result of the spatial distance between each feature and the mean is not the same order of magnitude, resulting in the effect of the smaller eigenvalues not being able to highlight their effects. For this reason, this paper improves the feature fusion link (formula (7)). First normalize the spatial distances obtained by equations (4), (5) and (6):

$$S_L^P(i,j) = \frac{S_L(i,j) - S_{L\min}}{S_{L\max} - S_{L\min}} \quad (9)$$

$$S_a^P(i,j) = \frac{S_a(i,j) - S_{a\min}}{S_{a\max} - S_{a\min}} \quad (10)$$

$$S_b^P(i,j) = \frac{S_b(i,j) - S_{b\min}}{S_{b\max} - S_{b\min}} \quad (11)$$

After the new eigenvalues are obtained, the ratio of the three features corresponding to each pixel is different. Traditional direct addition will affect the proportion of each feature. So use the following formula for feature fusion:

$$S^P(i,j) = \frac{(S_l^P(i,j))^2 + (S_a^P(i,j))^2 + (S_a^P(i,j))^2}{S_l^P(i,j) + S_a^P(i,j) + S_b^P(i,j)} \quad (12)$$

The improved FT visual saliency detection algorithm and FT visual saliency detection results are shown in Figure 1.

(a)Infrared image (b}FT test result (c)Improved FT test result

Figure 1 Comparison of FT detection algorithms

2) ***2.2.4 Calculate the weight of feature points***

3) ***2.2.4.1 Local feature model of feature points***

When judging key points outside the salient region, the information around the feature points plays an important role. If only one feature point is considered, it cannot be judged whether it is a key point, so the arrangement of the surrounding feature points has a significant influence on the importance of this feature point. This article considers two information, the distance between feature points, and the density of surrounding feature points. Therefore, considering the two together, establishing the relationship to get the importance of the feature point, that is, the weight of feature point P_{weight}， Then, based on the weights of feature points, feature points with higher weights are retained, and feature points with lower weights are removed. This article sets the weight threshold $P_{weight}^{'} = 0.3$。

Assuming that $i \in V$ in image I , i is a feature point, and V is an image outside the saliency area.Centered on a feature point, r is a circle of radius O .(A represents 1/5 of the distance from the center of the salient region to the farthest point from the center in the salient region)

The circle O is a circle centered on a feature point and r is a circle (r represents $\frac{1}{5}$ the distance from the center of the significant area to the farthest point from the center in the significant area) : $\{center\_x, centre\_y, r\}$, $\{center\_x, center\_y\}$ is the center of the circle， the coordinates of the feature points i 。The weight P_{weight} of the feature point i is obtained by the following formula:

$$P_{weight} = PR_{dis} \times \mu \quad (13)$$

$$PR_{dis} = 1 - \frac{\sqrt{(x - Center\_x)^2 + (y - Center\_y)^2}}{r} \quad (14)$$

PR_{dis} shows that the smaller the distance from the center point, namely the feature point i , is, the larger PR_{dis} is, and (x, y) is the position of the neighboring feature point of i .

If the neighbor feature point is in the circle, then PR_{dis} is less than or equal to 1; if it is on the edge or outside, then PR_{dis} is equal to 0. μ denotes the density of the characteristic point i .

There is

$$\mu = \frac{k_i}{\varphi_i} \quad (15)$$

k_i represents the degree of the feature point i , that is, the number of feature points in the circle O ， φ_i denotes the total number of all feature points in a circular area having $\frac{3}{2}r$ as a radius centering on the feature point i .

Here

$$k_i = \sum_{x \in V} a_{ix} \quad (16)$$

among them

$$a_{ix} = \begin{cases} 1 & \text{distance between i and x} < r \\ 0 & \text{distance between i and x} \geq r \end{cases}$$

Since the degree of the feature point reflects the number of feature points around the selected feature point and cannot reflect the density of the neighbor feature points, it is still difficult to accurately determine the importance degree of the feature point. Therefore, the degree of feature points and the degree of neighboring feature points must be comprehensively considered. Its expression φ_i ：

$$\varphi_i = k_i + \sum_{x \in \Omega_i} k_x \quad (17)$$

k_x represents the number of feature points in the circle o with the feature point x as the center, $\frac{r}{2}$ as the radius, and Ω_i represents the set of feature points in the circle O.

In equation (15), the larger k_i is, that is, the more neighboring feature points are, and the closer φ_i is to k_i, that is, the denser the neighboring feature points are, the larger μ is, and the more important feature point i is. Parameter μ makes use of the degree and density of feature points, which fully reflects the local structure of the node.

Equation (13) is the local structure information model to be established. The weight of the feature point P_{weight} finally obtained reflects the importance of the feature point i from the aspect of structure information. Using the obtained weight value P_{weight} to compare with a given threshold, feature points greater than or equal to the threshold are retained.

In this way, it can not only effectively detect the important feature points, but also avoid missing important feature points, so as to improve the efficiency of image retrieval.

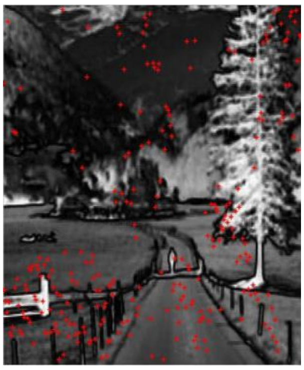
2.2.5 Extract SURF descriptors

Through the FT significance algorithm and the SURF feature point weight algorithm, the insignificant feature points are eliminated and the key points are preserved, so that the feature points to be obtained can better reflect the image information.

III. Experimental Results and Analysis

In order to illustrate the effectiveness of the method, the experiment compares the detection algorithm of SURF feature points based on visual saliency with the results of standard SURF algorithm and visual saliency algorithm.Figure 2 is a feature point extracted by the standard SURF algorithm;Figure 3 is the test result of visual saliency detection;Figure 4 is a feature detection result based on a visual saliency map;Figure 5 is the final test result of this article's extraction algorithm.Compared with Figure 2, the feature points extracted in Figure 5 not only retain the useful key points extracted by the SURF algorithm, but also can represent and identify the feature information of the images more completely, and the redundancy is relatively small and has good scale Invariability.Especially in the part of the road surface, the effect is particularly obvious.This paper illustrates that the algorithm is more effective than the standard SURF extraction algorithm.

Figure 2 SURF feature point detection Figure 3Schematic diagram of visual saliency detection



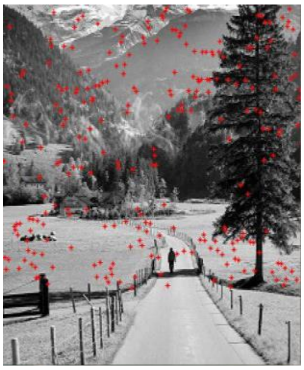


Figure4Detection of feature points in visual saliency maps Figure 5 This article algorithm

IV. Conclusion

This paper proposes an image feature extraction method combining visual saliency with SURF algorithm.The simulation experiments of feature extraction using standard SURF algorithm, visual saliency algorithm, and SURF algorithm based on visual saliency improvement were performed.After analyzing, we found that the algorithm of this article is superior to other algorithms.In the process of extracting the key points of the image, it is found that the algorithm can not only preserve the visual saliency information, but also extract the visible information outside the visually significant area.The detection results of pavement feature points in the figure show that the algorithm has obvious advantages in the identification and retrieval applications.

References

[1] Wang zhirui,yan cailiang.Overview of image feature extraction methods [J].Journal of Jishou University.(Natural Science News).2011,32(05):43-47.

[2] David G.Lowe.Distinctive Image Features from Scale-Invariant Keypoints[J].International Journal of Computer Vision,2004,V60(2):91-110.

[3] Cai Jia, Huang Panfeng.Research on real-time tracking of feature points based on improved SURF and P-KLT algorithm[J].Aviation Journal.2013,5(25):5(34):1204-1214.

[4] Shi Lei, Xie Xiaofang, Qiao Yongjun.Surface tracking based on SURF algorithm [J]. Computer Simulation,2010,27(12):227-231.

[5] Li Xiuyan, Liu Tiegen, Deng Shichao, et al. Fast hand vein recognition based on SURF operator[J]. Journal of instrumentation,2011,32(4):831-836.

[6] Luo J,Oubong G.SURF applied in panorama image stitching[C]//2010 2nd International Conference on Image Processing Theroy Tools and Applications.[S.I]:IEEE,2010:495-499.

[7] Herbert Bay,Andreas Ess,Timme Tuyteplaars,et al.Speeded-Up Robust Features[J].Computer Vision and Image Understanding,2008,110(3):346-459.

[8] Dong zhijie,Ye feng,Li di,et al.PCB matching based on SURF [J]. Circuit World, 2012, 38(3): 153-162.

[9] Biplab Banerjee,Tanusree Bhattacharjee,Nirmalya Chowdhury.Image Object Classification Using Scale Invariant Features Transform Descriptor with Support Vector Machine Classifier with Histogram Intersection Kernel [J].Communication in Computer and Information Science,2010,101(2):443-448.

[10] Shijie jia,Nan Xiao,Zeng Jie.Trademark Image Retrieval Algorithm Based on SIFT Feature [J].Lecture Notes in Electrical Engineering,2012,113(3):201-207.

[11] Minkook Cho,Hyeyoung Park. A Robust Key points Matching Strategy for SIFT :An Application to Face Recognition[J]. Lecture Notes in Computer Science,2001:716-723.

[12] Xiaomin Liu and Peihua Li.An Iris Recognition Approach with SIFT Descriptors[J].Lecture Notes in Computer Science,2012:427-434.

[13] Geusebroek J M,Boomgaard R.Color Invariance[J].IEEE Trancesactions on Pattern Analysis and Machine Intelligence,2001,23(12):1338-1350.

[14] Shi Yasun, Liu Xiaoyun, Chen Fen. Color Image Registration Based on SURF[J].Infrared Technology,2010,32(7):415-419.

[15] Shi yasun.Image registration research for improveed SURF key-points[D].Xidian:Xidian University,2008[16] Yang Fan, Deng Zhengsheng.Image Feature Extraction Based on Histogram Equalization and SURF Reconstruction[J].Computer Engineering and Applications.,2013:188-199.

[17] Yang Zuqiao,Chen Yuepeng,Zhang Qing.Research on Significant Local Feature Extraction Algorithm for Comprehensive Visual Attention Model[J], Computer Science.,2013,40(8):289-292.

[18] Koch C.Ullman S.Shifts in selective visial attention: towards the underlying neural circuitry[J].Hum Neurobiol,1985,4(4):219-227.

[19] Wu Yiquan, Tao Feixiang, Cao Zhaoqing.Image Registration Algorithm Based on Double-valued Complex Wavelet Transform and SURF[J].System Engineering and Electronics,2014,36(05):997-1003.

[20] Shi Lei, Xie Xiaofang, Qiao Yongjun.Face feature detection based on SURF algorithm and OPenCV[J].Computer and Digital Engineering,2010,38(2):124-126.

[21] Gao Jian, Huang Xinhan, Peng Gang. et al. A simplified SIFT image feature point extraction algorithm[J].Computer Applications Research,2008,25(7):2213-2215.

[22] Brown M,Lowe D.Invariant features from interest point groups [C]//British Machine Vision Conference.Cardiff,Wales,2002:656-665.

[23] R.Achanta,S Hemami,F.Estrada.Frequency-tuned salient region detection.In CVPR, pages 1597-1604,2009.1,2,4,5,6,7.

[24] HE K, SUN J, TANG X. Guided image filtering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2013, 35(6):1397-1409.

[25] Shen Zhiwei. Prepress digital Lab image processing[J]. Printing Technology.2014,(17):42-45