Proceedings of the 37th Chinese Control Conference
July 25-27, 2018, Wuhan, China

Vehicle Recognition Method Based on Color Invariant SIFT Features

Yiming Xu, Juan Zhang, Chengcheng Liu, Juping Gu, Liang Hua

School of Electrical Engineering, Nantong University, Nantong 226019
E-mail: gu.jp@ntu.edu.cn

Abstract: Both aiming at the problems of the color information lack for SIFT features and the interference of color on vehicle identification, a vehicle identification method based on color invariant SIFT features is proposed. Firstly, the color edge regions of the images are calculated by using the RGB information, and the SIFT color invariant feature description vectors are generated in conjunction with the gray information. While ensuring the integrity of the vehicle color features, the method can effectively suppress the false matching caused by the redundant feature points. Then the nearest neighbor matching method is used to find the matching pairs, and the vehicle identification is finished by the discrimination degree weighting strategy. In the vehicle matching experiment with the same type and different colors, the matching number of the algorithm is less affected by the change of the color, and the stability of image matching is enhanced. The vehicle recognition experiment based on 25 kinds of vehicle model database shows that the method can eliminate the interference of the color features on the vehicle identification, and the total recognition accuracy of the method is improved to 88.2%-92.2%.

Key Words: Color Invariant Features, Scale-invariant Feature Transform, Vehicle Recognition, Color Edge Detection, Discrimination Degree Weighting

1 Introduction

In recent years, vehicle information detection and recognition technology based on computer vision has been developed rapidly, where vehicle type recognition technology plays a powerful auxiliary role in traffic management and safety monitoring. At present, most surveillance systems use color video, and rich vehicle colors also provide a large amount of available information for vehicle type identification. However, when the traditional SIFT algorithm processes color images, color images are often converted into grayscale images and then the features are extracted and matched. Discarding color information and using only the grayscale information of the image may cause mismatches and interfere with the subsequent vehicle type identification process. Therefore, based on the SIFT algorithm to remove the interference of the color on the vehicle type identification is the focus of this article.

In 2004, LOWE proposed Scale Invariant Feature Transform(SIFT), which is invariant to rotation, scale scaling and brightness variation. It is also robust to change of perspective, affine transformation and noise, so it is widely used in image feature matching[1]. In 2008, BAY et al proposed Speeded Up Robust Features(SURF). Although it inherits the advantages of the SIFT algorithm, it often gets fewer matches. When there are fewer image feature points, the matching real-time error is larger[2]. XU Yanlu et al[3] proposed a SIFT image matching algorithm combining color invariance and shape context. The color information was integrated on the basis of the traditional SIFT algorithm. The shape context histogram based on each contour point was replaced by the shape context histogram based on the center of gravity point, and the two cascaded to form a new eigenvectors, which effectively improved the matching accuracy. In the literature[4], Wang Rui et al considered the problem of ignoring the color information in the local features of the image, and proposed a scale-invariant feature transform method that integrated the global color information. The global shape information and color invariant information were added to the original SIFT frame by establishing the concentric coordinate system. What's more, the Euclidean distance was used as the matching cost function to reduce the false matching rate and enhance the matching stability and robustness. The algorithms proposed in [3-4] could effectively improve the matching accuracy of images in the presence of a large number of similar shaped regions, but none of them discussed the interference of color features on the image matching.

This paper presents a vehicle identification method based on color invariant SIFT features for the vehicle frontal images of the same size. In order to solve the interference problem of color on vehicle image matching, the color edge detection based on RGB space is firstly integrated into the vehicle frontal image to make the extracted vehicle features more complete. The color invariant SIFT feature descriptor is generated by combining the gray information to avoid the interference of different colors to SIFT feature matching and improve the stability and robustness of image matching. Finally, the discrimination degree weighting strategy is constructed according to the matching pairs to realize the vehicle type recognition.

2 SIFT Feature Extraction and Matching

The detection of the extreme points of SIFT operator is based on the difference of Gaussian (DOG) scale space. The implementation is to convolute the image with Gaussian kernel function with different scale factors and then subtract the Gaussian image[5-6]. Image I in the DOG scale space of dimension σ is defined as:

$$\begin{aligned} D(x,y,\sigma) &= (G(x,y,k\sigma)-G(x,y,\sigma))*I(x,y) \\ &= L(x,y,k\sigma)-L(x,y,\sigma) \end{aligned} \quad (1)$$

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

*This work is supported by National Natural Science Foundation (NNSF) of China under Grant 61673226, Jiangsu Provincial College of Natural Science Research Project under Grant 17KJB460009 and Nantong City Applied Basic Research Project under Grant GY2016018.

Where $G(x, y, \sigma)$ is the convolution kernel of Gaussian function under the scale σ, $L(x, y, \sigma)$ represents Gaussian scale space of image I under the scale σ. k is a constant, generally taken as $\sqrt{2}$.

In order to find the key points in the scale space, each sample points needs to be compared with the points in its neighborhood under the same scale 3×3 and the points in the 3×3 neighborhood corresponding to the upper and lower adjacent scales(a total of 26 points). Only when the pixel is the maximum or minimum of the 26 points, it may be regarded as the key point. Figure 1 shows the local extreme detection in DOG scale space, the pixels marked as × in the middle are compared with the 18 points on the adjacent scales at the same time when they are compared with the 8 adjacent points on the same scale, ensuring that extreme points are detected within the DOG scale space[7].

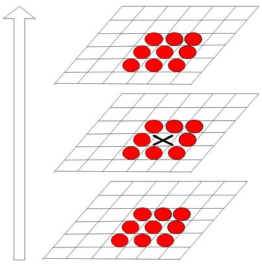


Fig. 1: Extreme points detection in DOG space

The stability of candidate extreme points are detected, that is, the feature points are precisely located, the low contrast points and the unstable edge response points are filtered out. Then the extreme points detected can be described as SIFT feature points.

In order to make the feature descriptor have the direction rotation invariance, the main direction of the feature points must be determined. The main direction can be calculated according to the gradient magnitude and the direction distribution of the pixels around the key points[8],which is defined as:

$$f(x,y)=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2} \quad (3)$$

$$\theta(x,y)=\tan^{-1}((L(x,y+1)-L(x,y-1)) / (L(x+1,y)-L(x-1,y))) \quad (4)$$

Where, $f(x, y)$ is the magnitude of the gradient and $m(x, y)$ is the gradient direction.

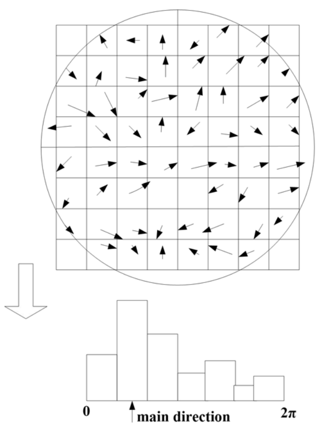


Fig. 2: Main direction generation

A neighborhood window centered on the feature points is sampled. The histogram is used to calculate the accumulated value of each gradient direction in the neighborhood pixels. The range of the gradient histogram is 0-360 degrees, and one direction for every 10 degrees. The direction located in the histogram peak value is the main direction of the feature point. As shown in Figure 2, the main direction is determined by the histogram within the region.

In the feature point descriptor generation stage, the 16×16 region around the feature point is usually selected and evenly divided into eight 4×4 sub-regions. Statistical histograms of 8 gradient directions are calculated in each small region block and sorted according to positions, thus forming a $4\times4\times8$ feature vector descriptor.

SIFT performs feature points matching by calculating the Euclidean distance of two feature point description vectors. The common matching method is the Nearest Neighbor Ratio (NNR) matching, which finds out the nearest point and the second nearest point of Euclidean distance, and compares the distance ratio and the nearest neighbor ratio threshold to judge whether the nearest point is correctly matched[9]. When the threshold is set to a small value, it is possible to ensure that the result does not contain false matching points, but may miss the correct matching points. When the threshold is set to a larger value, all the correct matching points can be guaranteed, but at the same time, the wrong matching points will be introduced. When the number of false matching points is too large, it will seriously affect the recognition results.

3 Color Invariant SIFT Feature Extraction and Matching

3.1 Color Edge Detection

Image edge is the most basic feature of the image, it can not only transmit a large amount of information, but also depict the basic contours of the object[10]. The first step of the image processing and analysis is often the edge detection, but most of the research on images is based on grayscale images, such as the common edge detection operator Sobel operator, Roberts operator and so on. In fact, the edge detection on the gray image can easily lose the color information of the image, and the extracted edge is not detailed and complete. If the above operators are applied directly to the color image, some of the main edge information will not be detected. So it is necessary to construct a new color image edge detection algorithm by studying the edge detection of gray image, which can fully preserve the color features of the image and enhance the effects of edge detection.

In general, color images are stored and expressed in the form of RGB. Edge detection using RGB color space avoids losing the original image information during non-linear conversion because of without having to switch to other color spaces, and at the same time reduces the computational complexity and time consuming. In this paper, the image gradients are calculated and the edges of color image are enhanced on the basis of RGB space. The specific steps are as follows:

(1)Any point is picked in the image;

(2)The color of the point is divided into R, G, B three components;

(3)The gradient values of the R, G and B components in the x direction are respectively calculated by the formulas (5), (6) and (7):

$$H\_rx = |RGB(i-1, j+1, 1) + 2*RGB(i, j+1, 1) + RGB(i+1, j+1, 1) - RGB(i-1, j-1, 1) - 2*RGB(i.j-1, 1) - RGB(i+1, j-1, 1)| / 255 \quad (5)$$

$$H\_gx = |RGB(i-1, j+1, 2) + 2*RGB(i, j+1, 2) + RGB(i+1, j+1, 2) - RGB(i-1, j-1, 2) - 2*RGB(i.j-1, 2) - RGB(i+1, j-1, 2)| / 255 \quad (6)$$

$$H\_bx = |RGB(i-1, j+1, 3) + 2*RGB(i, j+1, 3) + RGB(i+1, j+1, 3) - RGB(i-1, j-1, 3) - 2*RGB(i.j-1, 3) - RGB(i+1, j-1, 3)| / 255 \quad (7)$$

(4)The gradient values of the R, G and B components in the y direction are respectively calculated by the formulas (8), (9) and (10):

$$H\_ry = |RGB(i+1, j-1, 1) + 2*RGB(i+1, j, 1) + RGB(i+1, j+1, 1) - RGB(i-1, j-1, 1) - 2*RGB(i-1.j, 1) - RGB(i-1, j+1, 1)| / 255 \quad (8)$$

$$H\_gy = |RGB(i+1, j-1, 2) + 2*RGB(i+1, j, 2) + RGB(i+1, j+1, 2) - RGB(i-1, j-1, 2) - 2*RGB(i-1.j, 2) - RGB(i-1, j+1, 2)| / 255 \quad (9)$$

$$H\_by = |RGB(i+1, j-1, 3) + 2*RGB(i+1, j, 3) + RGB(i+1, j+1, 3) - RGB(i-1, j-1, 3) - 2*RGB(i-1.j, 3) - RGB(i-1, j+1, 3)| / 255 \quad (10)$$

(5)The gradient values H_x in the x direction and the gradient values H_y in the y direction are calculated, and the larger value is taken as the color value of the center point:

$$H_x = RGB(H\_rx, H\_gx, H\_bx) \quad (11)$$

$$H_y = RGB(H\_ry, H\_gy, H\_by) \quad (12)$$

(6)The above five steps are performed on each pixel in the image to obtain a color edge image;

(7)Grayscale threshold segmentation is performed for the color edge image, if it is greater than the threshold, the pixel is set to1, otherwise set to 0.

3.2 CI-SIFT Feature Extraction and Matching

The steps of color invariant SIFT (CI-SIFT) feature extraction and matching are basically the same as the SIFT algorithm, including the extreme detection of scale space, key point screening, main direction determination, feature point description and nearest neighbor ration matching. The difference is that in the scale space detection process, the color invariant value $H(x, y)$ detected by the color edge detection is used instead of $I(x, y)$ in the original SIFT algorithm. The formula is defined as:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * H(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (13)$$

Therefore, the CI-SIFT algorithm does not improve the original structure of the SIFT descriptor, but uses the mean of the color components in each sub-region to form the eigenvector. Since the feature point description stage is to calculate the gradient values of the pixels in the x and y directions within the neighborhood of the feature points, the robustness of the generated descriptor is greater if the change of the pixel gradient around the feature point is larger. The feature points in the edge region relatively change greatly in gradient, so the generated color invariant descriptors are more robust, which can restrain the mismatches caused by some feature points with less information. In addition, the color difference caused by human factors such as vehicle body color change as well as environmental factors such as lighting will reduce the matching pairs between the reference image and the image to be matched, but the stable color invariant features generated in the process of feature point description just weaken the effect to the vehicle type recognition, which can effectively restrain the interference of the color information on vehicle identification while improving the matching stability of the algorithm.

4 Vehicle Identification Based on CI-SIFT Features

4.1 Vehicle Body Color Recognition Based on HSV Space

The standard color template is quantified according to the change relationship between the H (hue), S (saturation) and V (value) in the standard HSV color space[11]. First of all, part of the front is intercepted from the car hood according to the proportion as the body color identification area, and then the image is converted from RGB space into HSV space. Color histograms are used to statistic the number of color appears, and the color that appears most often is the color to be recognized. The color histogram shows the proportion of different colors in the entire image, regardless of the spatial location of each color[12]. The color histogram statistics instead of the average color component obtaining, to a certain extent, reduces the impact of light and improve the accuracy of color recognition. Finally, the vehicle body color is judged according to the parameters of the quantification template. Common vehicle colors include red, blue, green, yellow, black, silver and white. The red H component has a parameter range of 0°-15°or 245°-255°, 20°-65°for yellow, 66°-130°for green and 200°-250°for blue while the H components of the three colors of black, silver and white all fall within the range of 0°-180°.

4.2 Vehicle Identification Base on CI-SIFT Feature Matching and Discrimination Degree Weighting

First of all, M samples are selected as the matching samples from the frontal images of each type of vehicle, and the sample library is set up. If the number of detected vehicles is N, the sample library contains N×M images.

Then the front of the vehicle to be recognized is taken as the input image. The color edge region of the image is obtained by using the color edge detection algorithm based on RGB, and the feature points are detected in the region. What's more, the CI-SIFT feature descriptors are generated according to the gray information and matched with the sample images to get N×M matching results.

For a certain car $C \in [1, 2, ..., N]$, the number of matching points between the input images and sample images is respectively $M_{C1}, ..., M_{CM}$. If the input image model belongs to the sample model, the variance of matching points in the sample database should be less than that of non-corresponding cases, so variance δ_C can be used to reflect the stability of the model similarity data set.

$$M_{CW} = \overline{M_C} / (\sum_{i=1,...,M} (M_i - \overline{M_C})^2 / M) \quad (14)$$

Where, M_{CW} is the matching result after weighting the input images and sample images, $\overline{M_C}$ is the average value of the matching points, and the discrimination degree of vehicle matching can be enhanced by weighting δ_C .

$M_{CW}, C \in [1,2,...,N]$ is sorted, if the ratio between the maximum value and the second largest value is less than T_{dis} (T_{dis} is the discrimination degree threshold, and the empirical value is 0.75), the sample type corresponding to the maximum value is taken as the recognition result of the input image, otherwise, the discrimination degree of the two vehicles is insufficient, then the sample type corresponding to the second largest value is taken as an alternative model.

5 Experimental Results

5.1 Experimental Conditions

The development environment of this algorithm is MATLAB R2015b. The detection and matching of SIFT operator uses correlation function in Computer Vision System Toolbox. Algorithm test environment is PC with the Intel i7-2670 and 8G memory. The operating system is Windows7.

5.2 Body Color Recognition in HSV Space

A certain range of the front hood was selected in accordance with the proportion as the color area to be identified. Figure 3 shows the red Cruze body color extraction area. The color histogram of this feature area in the HSV color space was calculated after RGB conversion. The statistical result of the H component was shown in Figure 4. The color that appeared most times was selected as the color of the feature area to be recognized, and the color was determined by the parameter interval of the color quantization template. In order to prove the accuracy of color recognition, 300 vehicle samples under normal lighting conditions were selected for experiment. The body color included red, yellow, green, blue, white, silver and black. The correct number of samples was 257, so the accuracy rate of identification was 85.67%.

Fig. 3: Body color feature extraction area

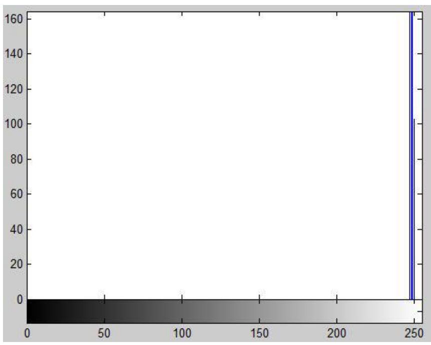
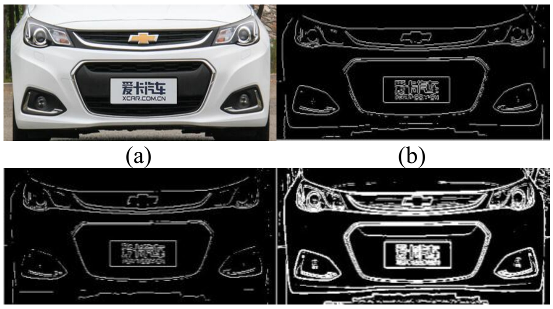


Fig. 4: Feature area color histogram

5.3 Feature Detection in Color Edge Area

In order to verify the effectiveness of the introduction of color edge detection, the color edge detection algorithm based on RGB space was compared with the traditional Sobel operator and Roberts operator. Taking the white Cruze as an example, the edge detection results were shown in Figure 5. Figure 5(b) showed the detection result of the Sobel operator on the gray image. Figure 5(c) showed the detection result of the Roberts operator on the gray image. Figure (d) was the result obtained by the method proposed in this paper. Compared with the traditional algorithm, the color edge detection algorithm based on RGB color space could better preserve the color information and edge information of the original image, and the edges were clear and complete.



(a) (b)

(c) (d)

Fig. 5: Comparison of edge detection effects(a)original front image(b)sobel operator(c)roberts operator(d)detection method of this paper

The SIFT algorithm was used to extract the local SIFT feature points of the front edge image. Figure 6(a) were feature points extracted by the standard SIFT algorithm. Figure 6(b) were the color invariant feature points extracted after the introduction of color edge detection combined with gray information. It could be seen from the comparison of the two cases that the feature points detected in Figure (b) were fewer than (a), which could effectively eliminate the redundant feature points of the standard SIFT algorithm and reduce the influence of irrelevant features on the recognition effects, so as to ensure the accuracy of subsequent processing.

(a) (b)

Fig. 6: Comparison of feature points detection(a)traditional edge detecting feature points(b)color edge detecting feature points

5.4 CI-SIFT Feature Matching and Vehicle Identification

After the local SIFT features were extracted, the NNR was used to match the front color edge image, where the nearest neighbor threshold was 0.75. Considering that the color change was likely to cause the edge direction to jump and appear mismatch points, the traditional SIFT algorithm and the method proposed in this paper were respectively used to conduct NNR matching on the vehicle image to further explore the impacts of body color on vehicle identification. The white BMW 3 series was taken as a reference image, and the BMW 3 series of different colors were selected as images to be matched, including white, blue, black, silver and red. White, blue, black, silver and red were the common BMW 3 series body color.

The matching results were shown in Figure 7-16, where Figure 7 and Figure 8 showed the matching results of the white BMW 3 series and the white BMW 3 series, Figure 9 and Figure 10 showed the matching results of the white BMW 3 series and the blue BMW 3 series, Figure 11 and Figure 12 showed the matching results of the white BMW 3 series and the black BMW 3 series, Figure 13 and Figure 14 showed the matching results of the white BMW 3 series and the silver BMW 3 series, Figure 15 and Figure 16 showed the matching results of the white BMW 3 series and the red BMW 3 series.

Fig. 7: Matching results of SIFT algorithm for the white BMW and the white BMW

Fig. 8: Matching results of the algorithm in this paper for the white BMW and the white BMW

Fig. 9: Matching results of SIFT algorithm for the white BMW and the blue BMW

Fig. 10: Matching results of the algorithm in this paper for the white BMW and the blue BMW

Fig. 11: Matching results of SIFT algorithm for the white BMW and the black BMW

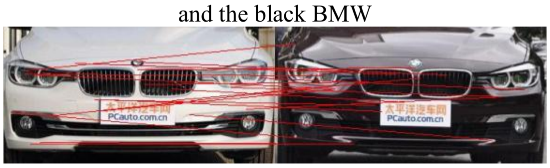


Fig. 12: Matching results of the algorithm in this paper for the white BMW and the black BMW

Fig. 13: Matching results of SIFT algorithm for the white BMW and the silver BMW

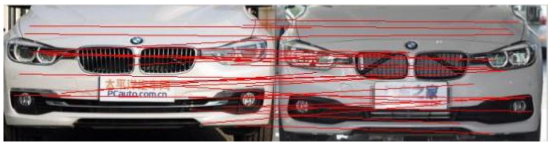


Fig. 14: Matching results of the algorithm in this paper for the white BMW and the silver BMW

Fig. 15: Matching results of SIFT algorithm for the white BMW and the red BMW

Fig. 16: Matching results of the algorithm in this paper for the white BMW and the red BMW

Related parameters in the above matching process were shown in Table 1, and the line chart of the matching number was shown in Figure 17. For the same type of vehicle, due to the change of the body color, the matching number of the standard SIFT algorithm was significantly affected by the color change, but the reduction magnitude of the matching pairs obtained by the algorithm in this paper was far lower than the standard SIFT algorithm.It could be seen that the introduction of the color edge detection and the preserve of the body color information could reduce the impact of color on vehicle matching.

Table 1: Related Parameters in Matching Process

Reference image	Image to be matched	Standard SIFT algorithm	Algorithm in this paper
White BMW 3 series	White BMW 3 series	78	46
	Blue BMW 3 series	42	37
	Black BMW 3 series	31	43
	Silver BMW 3 series	9	36
	Red BMW 3 series	3	41

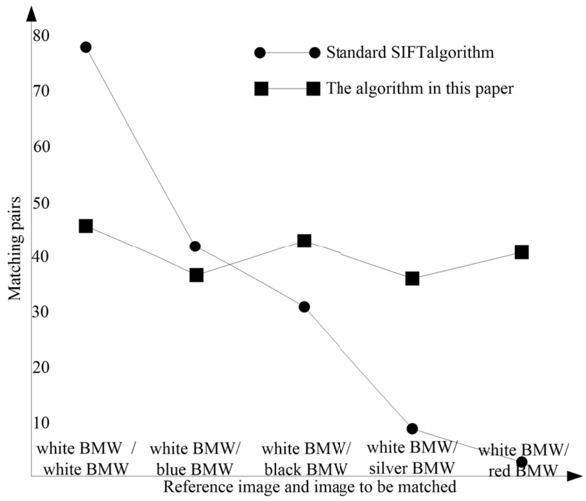


Fig. 17: Line chart of matching pairs between the reference image and the image to be matched

25 kinds of models were selected from the vehicle frontal image database, per type had 20 images, 2,3,4,5 images were selected as matching sample base from them. The color invariant SIFT features of the front region images were extracted to build the feature library file, and the other images of these 25 models were identified. The recognition results were shown in Table 2, in which the total recognition rate was the recognition result including the candidate vehicle model.

Table 2: Recognition Results for 25 Kinds of Vehicle Models

Number of samples per type /Total number of samples	Number of input images	Recognition rate	Average time-consuming /s	Total recognition rate (Including candidate vehicle model)
2/50	450	84.2%	4.365	88.2%
3/75	425	86.1%	6.742	88.9%
4/100	400	89.3%	7.868	91.8%
5/125	375	90.9%	8.113	92.2%

6 Conclusion

In this paper, a vehicle recognition method with robustness and suppression of color information interference is proposed, which is characterized by: ①Color edge detection of the image is achieved based on RGB space to make the vehicle features extracted more complete; ②The gray information is combined to generate CI-SIFT features on the basis of the color edge detection, which reduces the impact of different color features on the vehicle identification; ③The reasonable recognition strategy is given by enhancing the discrimination degree of vehicle matching to ensure the reliability of the recognition. The vehicle identification method in this paper is based on the number of the correct matching pairs, so it is necessary to improve the matching method to improve the accuracy of vehicle identification.

References

[1] D. G. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60(2): 91-110, 2004.

[2] H. Bay, T. Tuytelaars, and L. V. Gool, SURF: speeded up robust features, *Computer Vision and Image Understanding*, 110(3): 346–359, 2008.

[3] Y. L. Xu, Y. Ma, S. B. Li, and X. F. Zhang, Algorithm of image matching based on color SIFT and shape context, *Computer Science*, 41(11A): 144–146, 2014.

[4] R. Wang, Z. D. Zhu, SIFT matching with color invariant characteristics and global context, *Optics and Precision Engineering*, 23(1): 295–301, 2015.

[5] D. Z. Cheng, Y. J. Li, and R. X. Yu, Image matching method based on improved SIFT algorithm, *Computer Simulation*, 28(7): 285-289, 2011.

[6] H. X. Fu, L. L. Han, Y. C. Wang, and P. Li, Electronic image stabilization based on SIFT corner detection, in *Proceedings of the 36th Chinese Control Conference*, 2017: 11029–11033.

[7] J. Xin, B. J. Ran, and X. M. Ma, Robust visual sliding mode servoing using SIFT features, in *Proceedings of the 35th Chinese Control Conference*, 2016: 4723–4729.

[8] J. J. Liu, X. H. He, and W. L. Chen, Image matching algorithm by wavelet transform combined with SIFT characteristics, *Computer Simulation*, 28(1): 257-260, 2011.

[9] C. Huang, Recognition of car based on local feature, *Shanghai: Shanghai Jiao Tong University*, 2010.

[10] W. Ju, J. L. Liu, and S. Z. Jin, An improved clustering based on edge detection method, in *Proceedings of the 35th Chinese Control Conference*, 2016: 4027–4030.

[11] Z. Y. Hu, Y. D. Cao, and Y. Li, Car color recognition algorithm based on HSV color space, *Journal of Liaoning University of Technology(Natural Science Edition)*, 37(1): 10-12, 2017.

[12] Z. Y. Hu, Design and implementation of vehicle information recognition system, *Liaoning: Liaoning University of Technology*, 2016.