

# The Method of Vehicle Type Identification Based on Improved Corner Feature

ZHANG Tong<sup>1</sup> and WEN Bochong<sup>1</sup>

<sup>1</sup>Network information security department of GuangDong police officer college, Guangzhou, 510230, China

Email:susanzhangtong@126.com

**Abstract** This paper present a method of vehicle type identification using improved corner points as key points for feature description. The corner points can reflect the information of different type. Therefore, we adopt the features of corner points to recognize which type the vehicle belongs to. The method of Harris corner detection was improved, and the ratio of eigenvalues of M matrix in different directions was added to the corner response. The corner points were detected by improved corner response and results of test showed that the count of the detected corner points was increased. Considering rotation invariance and scale invariance, feature description is carried out around detected corner points. The vector of feature is determined by the Harr wavelet response of the pixels around the corner point, but the direction is only one of the eight values, and the length of the vector is 16. The matching rate of feature vectors between different types is used to determine which type the vehicle belongs to. The results of experiment confirmed the validity of the method.

## 1. Introduction

In China the automobile is increasingly becoming the most important means of transportation in people's daily life. While bringing convenience to people, cars have also become the preferred way for criminals to commit crimes and escape, which brings great difficulties to the investigation and inspection work of public security. In many cases the lucky escape of criminals often resulted in the descending detection speed of criminal cases because criminals are not captured in time. How to lock the target suspect vehicle in the shortest time is an urgent problem to be solved by the public security system. The most prominent feature of the target suspect vehicle is the license plate number. Once the license plate number of the escaping vehicle is obtained, the suspect vehicle can be quickly tracked down using a strategic pass system that covers the city. However, in real life, due to the weather, light, time and other reasons, many informers failed to record the number of escape vehicle license plate. Searching a suspect car among the city's millions of vehicles is like looking for a needle in sea. Although many informer have not been able to record the number of the escape vehicle license plate, but generally they know the suspect vehicle's colour, brand, additionally considering when and where the case happen, combined with the vehicle type recognition technology, police officers will be able to lock the suspect vehicle in a very small range, which brings great convenience to detect the case. The technology of vehicle type recognition aims at finding out which brand the vehicle belongs to. For example, the car is detected to the brand of Audi, BMW or Volkswagen etc. based on vehicle's face features[1].

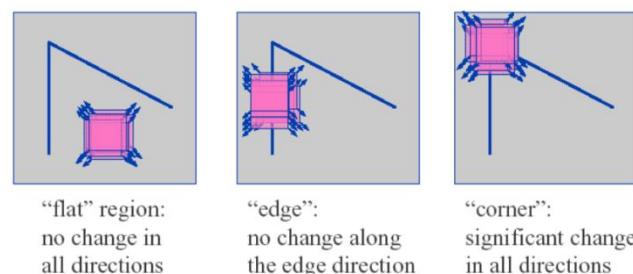


Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

The front structure of vehicle which belongs to different brand is different, and the corner points of image can reflect the information of different structure. Therefore, we adopt the features of corner points to recognize which type the vehicle belongs to. Corner point detection was first proposed by Moravec in 1977 that used the operator of gray variance to extract point features, which mainly selects the points with the maximum and minimum gray variance as feature points in four directions[2]. Although Moravec's corner detection algorithm has the advantages of fast calculation speed, it also has some problems, such as using only four directions for autocorrelation, the response is anisotropic, and because a square window is used, the response also contains noise etc. To solve these problems, Harris and Stephen have made major changes to Moravec corner detection algorithm and proposed Harris edge and corner detection algorithm[3]. Harris operator is an effective feature point extraction operator, and its advantages mainly include: (1) the points of extracted features are uniform and reasonable. Harris operator calculates its response value for each point in the image, and then selects the best point in the neighborhood. (2) Simple calculation. Harris operator only uses the first difference of gray image and filtering, and the whole process is highly automated. (3) Stability. It is a stable feature extraction operator even in the presence of image rotation, gray scale change, and noise effect and viewpoint transformation. (4) Feature points can be extracted quantitatively. Since the last step of Harris operator is to sort all local extremum points, a certain number of optimal points can be extracted as needed. This technology must have scale invariance and rotation invariance. When people recognize an object, regardless of whether the object is far or near, they can correctly identify it, which is called scale invariance. When an object rotates, we still recognize it correctly. This is called rotation invariance. The feature description of the pixels around the corner point can solved this problem.

## 2. The Theory Of Harris Corner Point

People usually observe the gray value of point through a small window to recognize the corner point. We move the window to any direction. If the change of gray value in the point is rather large, and then we think the point to be a corner point(see Figure.1).



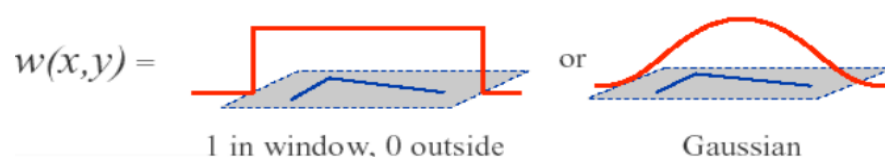
**Figure 1.** Determination of the corner according to the window response

For the transformation of intensity after moving the point  $[x, y]$  to  $[u, v]$ , the following formula can be obtained:

$$E(u, v) = \sum_{x, y} w(x, y) [I(x + u, y + v) - I(x, y)]^2 \quad (1)$$

$I(x+u, y+v)$  is the intensity after translation, and  $I(x, y)$  is the pixel of the original image. For the values in square brackets, if it's a region of constant intensity, then it's close to zero, whereas if the intensity changes dramatically it's going to be very large, so we expect  $E(u, v)$  to be very large.

Where  $w(x, y)$  is window function, it may be weight function or Gaussian functions(see in Figure.2)



**Figure 2.** window function

$I(x+u, y+v)$  is expanded as Taylor series

$$I(x+u, y+v) = I(x, y) + (I_x u + I_y v)I(x, y) \quad (2)$$

as

$$E(u, v) = \sum_{x,y} w(x, y) [I(x, y) + (I_x u + I_y v)I(x, y) - I(x, y)]^2 \quad (3)$$

then

$$E(u, v) = \sum_{x,y} w(x, y) I^2(x, y) [I_x u + I_y v]^2 \quad (4)$$

The following can be obtained

$$E(u, v) = \sum_{x,y} w(x, y) I^2(x, y) (I_x^2 u^2 + 2I_x I_y uv + I_y^2 v^2) \quad (5)$$

The matrix form is

$$I_x^2 u^2 + 2I_x I_y uv + I_y^2 v^2 = [u \ v] \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (6)$$

Write equation (5) as

$$E(u, v) = [u \ v] \sum_{x,y} w(x, y) I^2(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \quad (7)$$

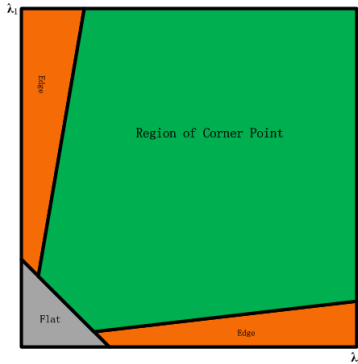
then

$$M = G(\sigma) * \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (8)$$

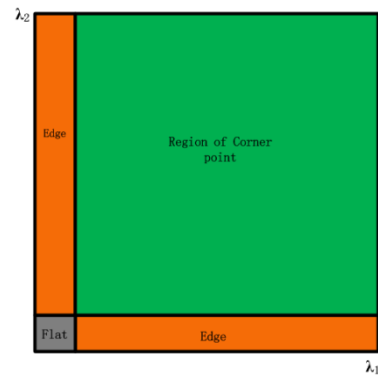
Where  $I_x$  is the first derivative of the horizontal  $\frac{\partial I}{\partial x}$ ,  $I_y$  is the first derivative of the vertical  $\frac{\partial I}{\partial y}$ , and  $G(\sigma)$  is the standard difference gaussian filter.

According to the value of  $\lambda_1, \lambda_2$  which were the eigenvalue of the matrix  $M$  the points can be classified into three categories, showed in Figure.3:

1.  $\lambda_1$  and  $\lambda_2$  are small and close, so  $E$  is close to constant in the all directions, and this region is “flat”.
2.  $\lambda_1 \gg \lambda_2$ , or  $\lambda_2 \gg \lambda_1$ ,  $E$  is going to be big in some direction, and this region is “edge”.
3.  $\lambda_1$  and  $\lambda_2$  are both big, so  $E$  is very large in that direction, and this region is “corner”.



**Figure 3.** Determination of the region by the eigenvalue of the matrix



**Figure 4.** Determination of the region by Shi-Tomasi response

Define  $R$  as Harris corner response of each pixel:

$$R = \det(M) - k(\text{trace}(M))^2 \quad (9)$$

Where  $k$  is a constant, usually 0.04.

$$\det(M) = \lambda_1 \lambda_2 \quad (10)$$

$$\text{trace}(M) = \lambda_1 + \lambda_2 \quad (11)$$

Where  $\lambda_1, \lambda_2$  is the eigenvalue of the matrix M.

### 3. Improvement of Corner Detection

Since the pixel is corner point or not depends on the eigenvalue  $\lambda_1, \lambda_2$  of the m-matrix, and the response value defined by Harris is complicated to calculate and relatively unclear in meaning, Shi-Tomasi [4] changed the response value to the following equation

$$R = \min(\lambda_1, \lambda_2) \quad (12)$$

In other words, The region determination is replaced by figure 4 in figure 3

While  $\min(\lambda_1, \lambda_2)$  is small, She-Tomasi corner point response had the no influence to the results of determination, that is the flat region and edge region. For the corner region, while both of  $\lambda_1, \lambda_2$  is big, difference is very small. Only when one of  $\lambda_1, \lambda_2$  is small and the other is big, the result of determination is wrong. Considering this case we joined the ratio of  $\lambda_1, \lambda_2$  to the response of corner, R is defined as

$$R = \min(\lambda_1, \lambda_2) + k \frac{\min(\lambda_1, \lambda_2)}{\max(\lambda_1, \lambda_2)} \quad (13)$$

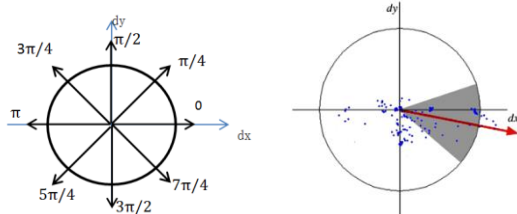
Where K is a weighting coefficient considering orders of magnitude of  $\min(\lambda_1, \lambda_2)$ .

The effect of the improvement on determination of the region is:

- 1) For the flat region,  $\min(\lambda_1, \lambda_2)$  is small,  $\max(\lambda_1, \lambda_2)$  is small too, then  $\min(\lambda_1, \lambda_2) / \max(\lambda_1, \lambda_2)$  is bigger, the improvement had an effect, and the influence can be reduced by increasing the threshold.
- 2) For the edge region,  $\min(\lambda_1, \lambda_2)$  is small,  $\max(\lambda_1, \lambda_2)$  is big, and then  $\min(\lambda_1, \lambda_2) / \max(\lambda_1, \lambda_2)$  is small, the improvement had no effect.
- 3) For the corner region,  $\min(\lambda_1, \lambda_2)$  is big,  $\min(\lambda_1, \lambda_2) / \max(\lambda_1, \lambda_2)$  is bigger, the improvement increased accuracy of determination.

### 4. Searching the Orientation of the Corner Point

To satisfy the demand of rotary invariant we need to determine a direction for the corner point[5]. Only one of the following eight values  $[0 \ \pi/4 \ \pi/2 \ 3\pi/4 \ \pi \ 5\pi/4 \ 3\pi/2 \ 7\pi/4]$  can be taken for the corner direction to simplify the calculation(see in Figure.5), therefore, we calculate the Haar wavelet response in x and y directions in the circular neighborhood of 6 pixels radius of corner point. We can also use the integral graph to speed up the filtering. Figure 6 shows the filter used. It only takes four steps at any scale to get the x or y response. A circular neighborhood of radius 6 pixels around the corner point was divided into eight areas and every area denoted one direction. The Haar wavelet response of pixel was calculated out in x and y direction. The horizontal and vertical responses are summed up. The two summed responses then yield a new vector. The longest vector determined the orientation of the corner point. It must be one of  $[0 \ \pi/4 \ \pi/2 \ 3\pi/4 \ \pi \ 5\pi/4 \ 3\pi/2 \ 7\pi/4]$ .



**Figure 5.** values of directions and a sliding window



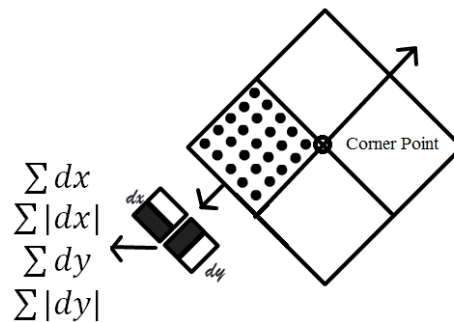
**Figure 6.** Haar wavelet filter

The direction is estimated by calculating the response sum in a sliding window of  $\pi/3$ , as shown in Figure.5. The horizontal and vertical responses within the window are added and then the horizontal sum and vertical sum formed a vector of orientation, and the longest vector in all the Windows was

defined as the direction of the point of interest. The size of the sliding window needs to be carefully selected, a small window can only find a dominant gradient, and nevertheless a large window can only output a meaningless maximum. Neither result is the right direction.

### 5. Description and Matching of the Corner Features

Considering of scale invariant we cannot adopt the position of the corner points to describe the feature. But the difference between pixels around the corner point can reflect the feature. A square region was built around the corner point and its edge is 10 pixels and its direction is along with the orientation of the corner point, and then this square region is regularly split into smaller  $2 \times 2$  squares, which are sub-regions (shown in Figure.7). For each sub-region of  $5 \times 5$  pixels,  $dx$  and  $dy$  of every pixel can be calculated out, where  $dx$  denotes the Haar wavelet response in horizontal direction and  $dy$  denotes the Haar wavelet response in vertical direction. "Horizontal" is defined as along the orientation of the corner point and "vertical" is defined as vertical to that. Then all the wavelet responses of  $dx$  and  $dy$  are summed up over each sub-region. Considering of information of the summit of the intensity changes, the sum of the absolute values of the responses,  $|dx|$  and  $|dy|$  were add into the vector. Thus each sub-region has a describing vector consisted of four parts  $v = (dx, dy, |dx|, |dy|)$ . Therefore a describing vector for all  $2 \times 2$  sub-regions was generated, and the length of it is 16. The Euclidean distance of the vector is calculated to determine the matching of feature.



**Figure 7.** Four parts of the feature vector

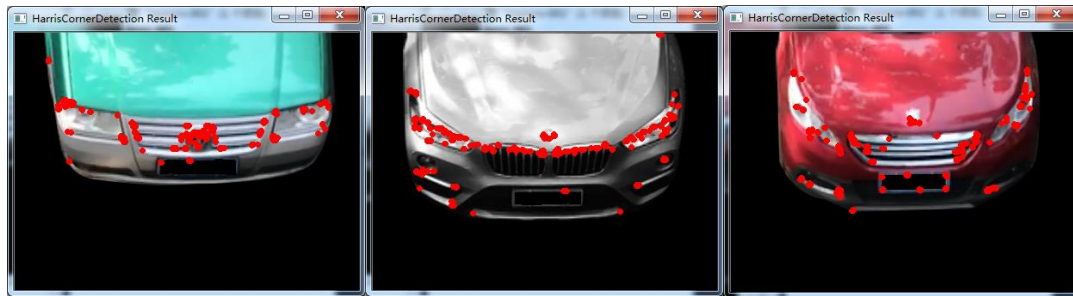
### 6. Results of experiment

In our experiment we process video and image by Visual Studio 2015 supported with OpenCV (Open Computer Vision library). Some popular vehicle types were selected out, and the car images in good quality were choose as standard type. Feature vectors of corn point were figure out and stored in database. The  $400 \times 300$  image of cars' face were segmented real time from the traffic video. Corner points are extracted from the segmented car face images by Harris method and improved method respectively (shown in Figure.8). The Table 1 show that the method of the improved corner point extraction can find more and better corner points, and the method is more stable.

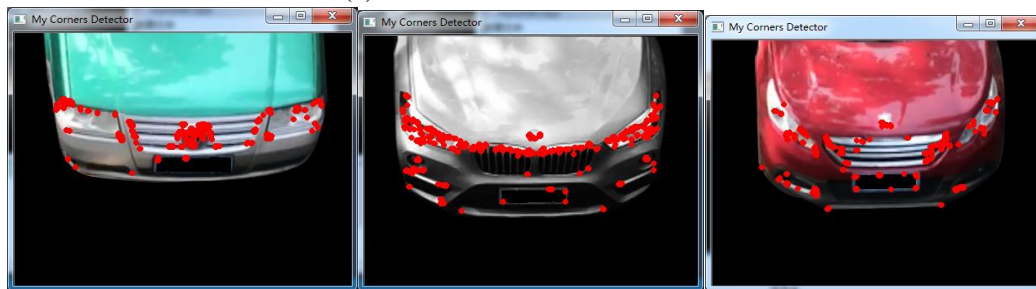
When  $k=0.0021$  in Equation 13, the number of corner point detected by improved method is increased by 8% than that of Harris corner. One factor is that the precision of response is improved and corner point amount is much, and another factor is that  $k * \min(\lambda_1, \lambda_2) / \max(\lambda_1, \lambda_2)$  increases the amount of corner point response, much higher than with the threshold point. Combining the two factors, the number of corner points is still within the range of processing, which is suitable for vehicle type identification.

After corner points of the car image were stored in structure of list, the direction of the feature vector was calculated out, and then feature vector was built along this direction. The feature vector of the image to be detected is matched with the features of standard vehicle type in the database, and the matching ratio determines whether it is the same car model. The matching process of four groups of vehicles is shown in the Figure.9,10,11,12. The table 2 shows the identification process of four groups

of vehicles. In the experiment, a large number of vehicles video were collected on the same road section for identification, and the identification results are shown in table 3.



(a)Result of Harris corner detection

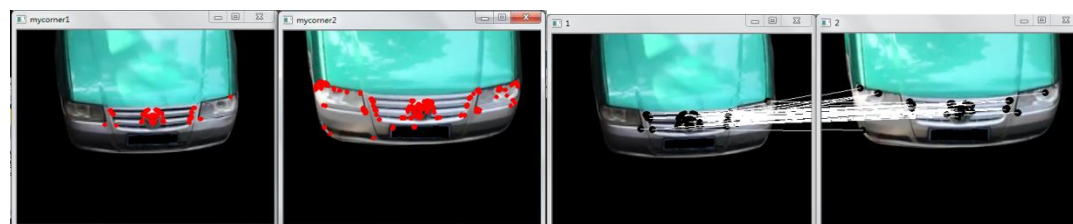


(b)Result of Improved corner detection

**Figure 8.** Comparison of corner point

**Table 1.** The number of corner point in Figure.8

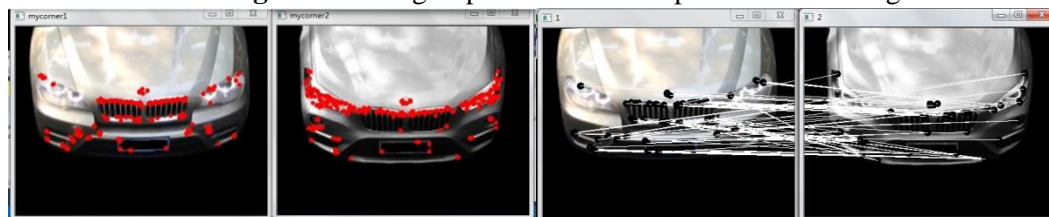
	Green Volkswagen	Siler BMW	Red Peugeot
Harris method	304	332	302
Our improved method	358	495	334



(a)Corner points of the image to be tested(left) and the stadard image(right)

(b)the matching of the image to be tested(left) and the stadard image(right)

**Figure 9.** No.1 group result of corner point and matching

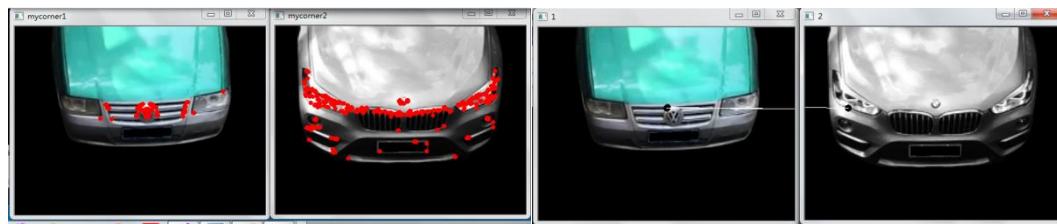


(a)Corner points of the image to be tested(left) and the stadard image(right)

(b)the matching of the image to be tested(left) and the stadard image(right)

**Figure 10.** No.2 group result of corner point and matching

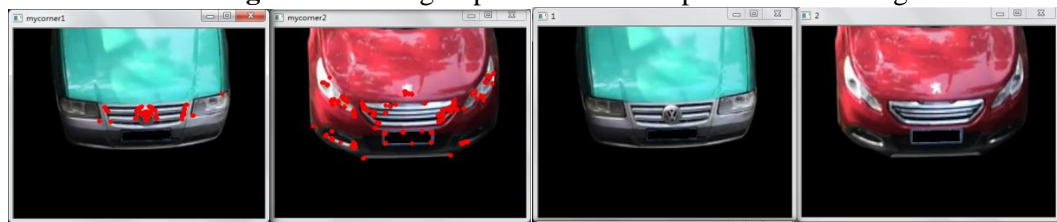




(a) Corner points of the image to be tested (left) and the standard image (right)

(b) the matching of the image to be tested (left) and the standard image (right)

**Figure 11.** No.3 group result of corner point and matching



(a) Corner points of the image to be tested (left) and the standard image (right)

(b) the matching of the image to be tested (left) and the standard image (right)

**Figure 12.** No.4 group result of corner point and matching

**Table 2.** vehicle type identification of 4 Groups in above figures

	The number of features detected	The number of matching	The ratio of matching	Is the same type or not
NO.1 group	93	93	100%	Yes
NO.2 group	223	117	52%	Yes
NO.3 group	93	2	2%	No
NO.4 group	223	0	0%	No

**Table 3.** identification results of different models

vehicle type	count	identification
Volkswagen	10	10
BMW	10	9
Audi	10	10
Buick	10	10
Peugeot	10	9

It can be seen from the table 3 that the accuracy rate of vehicle type recognition is more than 95%. Compared with other methods, such as SURF features and SIFT features, the algorithm is simple, fast and has a high accuracy rate.

## 7. Conclusion

In this paper we improved Harris corner detection method, the ratio of the eigenvalues of different directions of M matrix to join the corner point response. The corner point information was used as the key point of car face image to build feature vector. The feature vector is determined by the harr wavelet response of the pixels around the corner points. The description of the eigenvector is a 16-dimensional vector. Through a large number of experiments, the effectiveness of the recognition method is verified.

**Acknowledgements**

The authors gratefully acknowledge the support from Key research platforms and projects of universities in Guangdong province (2018KTSCX157) and Guangdong Science and Technology Plan Project(2018KJYZ009).

**References**

- [1] Zhang Tong, Mao Li. The Method of Cutting Image of Vehicle Face based on Haar feature and improved cascade classifier[C]. International Conference on Computer Graphics and Digital Image Processing, 2019, 7.
- [2] H. Moravec. Towards Automatic Visual Obstacle Avoidance[C], Proc. IJCAI, 1977, 8, p 584-590.
- [3] Chris Harris, Mike Stephens. A Combined Corner and Edge Detector[C]. The Fourth Alvey Vision Conf., 1988, p147-151.
- [4] Shi and C. Tomasi. Good Features to Track.[C] Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1994, 6, p593-600.
- [5] Zhang Tong, Peng Han, Yan Donsheng. The method of Vehicle type recognition based on improved SURF feature[C]. International Conference on Computer Graphics and Digital Image Processing, 2017, 7.