Fusion with CS-LBP and HOG for Vehicle Make and Model Recognition

Zhu Shanwei[1, a,*], Li Yuhui[1], Feng Jin[1] and Zhang Lingfeng[1]

[1] College of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China

[a]18288267375@163.com

*corresponding author

Keywords: vehicle recognition; local binary pattern; HOG feature; blocking CS-LBP.

Abstract: Local binary pattern (LBP) has widely used in face recognition with extracting texture feature, for the high statistical histogram dimensions of LBP and unable to effectively express the edge and direction information of image, so a method which is called fusion with blocking CS-LBP and HOG features is proposed, and applied vehicle recognition. At first, the vehicle image is extracted texture feature with blocking CS-LBP operator, which is calculated the texture histogram for each sub-block, and then the HOG feature of the original image is extracted, as well as the HOG feature which is based on the CS-LBP, finally, the blocking CS-LBP feature is fused with these two different HOG features. The experiments are implemented on the vehicle image databases, the results show that the proposed method can be obtained a higher recognition with K neighbor classification.

1. Introduction

Intelligent transportation system (ITS) has the great effect on alleviating traffic pressure and improving traffic management efficiency. Vehicle recognition is an important branch of intelligent transportation system, which is a hot topic for domestic and foreign scholars. The methods of vehicle recognition which are researched include geomagnetic induction coil method, infrared detection method, dynamic piezoelectric detection method and video image detection method, etc.

Predecessors have made a lot of research in vehicle recognition with the video image detection method, in the paper [1, 2, 3] the SIFT feature is used to classify the vehicle which is dimensioned by PCA, pooling, etc. In the paper [4] the Harris feature points extracted from the outline of the vehicle shape are used to identify the models. These methods for the vehicle recognition have the good effect, but only classifying the cars, trucks and buses, etc., the specific models can not be classified.

Face recognition is through the facial and facial features of the different distribution to distinguish between different people. For the vehicle, the different brand of vehicle has different logo, the same brand vehicle contains some series vehicle which have the different grid and headlights. These logo, grid and headlights constitute the face characteristics of vehicles which provide evidence to the identification of vehicles. In the paper [5, 6, 7, 8], the vehicle's face image is attempted to identify the specific vehicle models, for example, paper [7] used the shape histogram features of the vehicle's face image and the Bayesian transcendental model to identify the specific vehicle model, paper [8] used the HOG feature which is extracted from the face image of the vehicle to identify 80 vehicle models of 15 brands. The above methods verify the feasibility of identifying the specific vehicle models based on the vehicle's face images, but the recognition rate needs to be improved.

The vehicles in the real road are affected by environmental factors, such as light, and the resulting vehicle's face images contain noise, which is troubling the classification of the vehicle models. LBP is a non-statistical texture description operator, it is widely used in the field of face recognition [9, 10, 11] because of not sensitive to monotonous light changes and high computational efficiency. The method of traditional LBP is comparing the current pixel value with the pixel value in its neighborhood. This pattern will produce 256 different binary codes, which will cause the histogram to be too high to reduce the computational speed. Therefore, the paper [12] proposed the Center

Symmetric LBP to reduce the dimension and enhance the anti-noise ability. LBP is a non-directional one-dimensional operator, using LBP to extract the characteristics of the image will lose the direction of the image information, and the HOG feature is one of the best features that describe the edge and direction information of an image. In this paper, the combination of global feature and local feature will be applied to the recognition of the vehicle images with CS-LBP and HOG feature fusion.

2. CS-LBP and HOG feature

2.1 CS-LBP feature

The traditional LBP operator's window size is defined as a 3×3 rectangular window, the grayscale value of a pixel in a gray scale image is compared with the gray value of its neighborhood pixel. If the gray value of the neighborhood pixel is larger than the center pixel, it is set to one, otherwise it is zero. The coding process is shown in Fig 1.

Binary Code:
10101000

89	76	90
65	80	53
59	70	85

Threshold

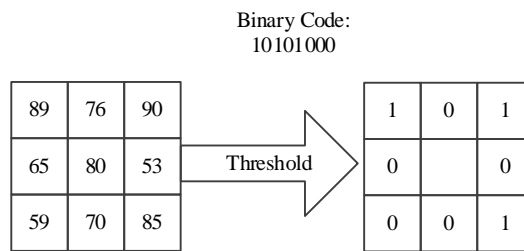| 1	0	1
0		0
0	0	1

Fig.1 LBP coding process diagram

So an 8-bit binary number can be got, each bit is multiplied by the corresponding weigh, then the cumulative sum will be LBP value, the formula is

$$LBP_{1,8}=\sum_{i=0}^{7}s(g_i - g_c)2^i \qquad (1)$$

Among the formula, g_c is the value of the center pixel, g_i is the value of the neighborhood pixel, s is the weigh, the value of s accords to the formula

$$s(x)=\begin{cases}0, & x<0 \\ 1, & x\geq0\end{cases} \qquad (2)$$

Unlike the traditional LBP, the CS-LBP need not to compare the values of the central pixels with the neighborhood pixels in turn, but compares the values of two pixels about the center symmetry, the coding process is shown in Fig 2.
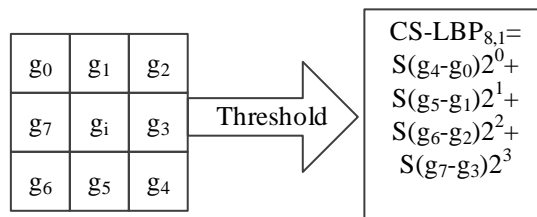
g_0	g_1	g_2
g_7	g_i	g_3
g_6	g_5	g_4

Threshold

$CS\text{-}LBP_{8,1}=$
$S(g_4\text{-}g_0)2^0+$
$S(g_5\text{-}g_1)2^1+$
$S(g_6\text{-}g_2)2^2+$
$S(g_7\text{-}g_3)2^3$

Fig.2 LBP coding process diagram

The value of s accords to the formula

$$s(x)=\begin{cases}1, & x\geq T \\ 0, & others\end{cases} \qquad (3)$$

In the formula, the value of the threshold T is 0 or 0.01 [13], usually it is 0. Compared with LBP, the dimensions of the statistical histogram features of CS-LBP reduces from 2^p to $2^{p/2}$,so the computational complexity is reduced. Because of the encoding method, CS-LBP has a stronger

anti-noise capability than LBP. A vehicle image is processed by the CS-LBP operator and the grayscale value is enlarged by 15 times, the extracted histogram feature is shown in Fig 3.
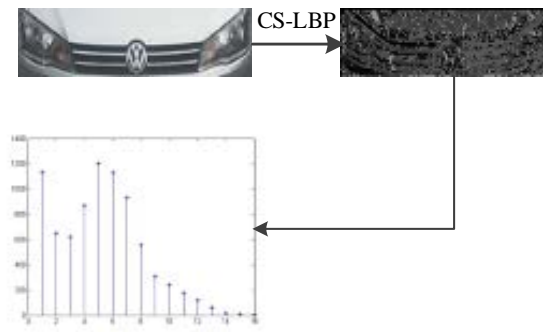

Fig.3 CS-LBP feature extraction histogram

2.2 HOG feature

Both HOG and SIFT are features that describe the shape and edge information of the image, there are similar steps in operation. Unlike the SIFT feature, the HOG feature has no scale and rotation invariance, but quantization in the direction space and position can slightly suppress the effect of image rotation and translation. The effect of color and light on the image can be neglected to a certain extent so that the data dimensions required for describing the image feature is reduced, the computational complexity of the HOG feature is small.

At first the HOG feature was applied from the beginning to pedestrian detection, which was later used for face recognition[14]. The HOG feature is described by the normalization of the gradient histogram of the cells of the image. The extraction process of the HOG feature is as follows:

Step1. The color image turns into a grayscale image, and the grayscale image is normalized with Gamma correction.

Step2. The image is divided into several cells, each unit consists of 3×3 pixels, the blocks are slid by one unit as the step size.

Step3. The horizontal and vertical gradients of each pixel point (x, y) are calculated in the image.

$$G_x(x,y)=H(x+1,y)-H(x-1,y) \tag{4}$$

$$G_y(x,y)=H(x,y+1)-H(x,y-1) \tag{5}$$

The gradient direction and the gradient amplitude at the pixel (x, y) can be obtained.

$$\alpha(x,y)=\arctan\left(\frac{G_y(x,y)}{G_X(x,y)}\right) \tag{6}$$

$$f(x,y)=\sqrt{(G_x(x,y))^2+(G_y(x,y))^2} \tag{7}$$

The gradient direction of each unit is evenly divided into 9 bins, and the histogram features of each unit are obtained by counting the gradient direction of all the pixels in the unit, so the histogram features of each block are obtained.

Step4. The HOG features of the entire image are obtained by cascading of the histogram features of each blocks sequentially.

3. Feature fusion

The main part of the vehicle image is the logo, lights and grid, the identification depends on the details, outline and edge information. The CS-LBP operator is robust compared with global feature in illumination and so on. However, the CS-LBP operator expresses the texture information of the image, and the extracted feature is not complete, direction and edge information are not effectively expressed.

The HOG feature describes the change of the local shape of the image, the partition processing makes the relationship have a good representation between the local pixels of the image which can describe the edge information of the image well.

Fusing with CS-LBP texture feature and HOG edge feature can improve the recognition rate, this paper combines two kinds of features to get relatively complete information.

3.1 Fusion with the block CS-LBP and original HOG feature

If the CS-LBP histogram is calculated with the entire face image, a 3×3 rectangular window is taken as an example, the statistical histogram features are only 16 dimensions. The amount of data is too small to describe the information of the face image, and the data of each dimension is too large to cause the calculation error to be large, the recognition rate is too low, so the vehicle image is divided. The vehicle image extracted by CS-LBP operator is divided into M×N sub-blocks, the histogram features of each sub-block image are statistically counted. Then, the histogram features of each sub-block image are combined to form the whole image's CS-LBP features, the specific operation is shown in Figure 4.
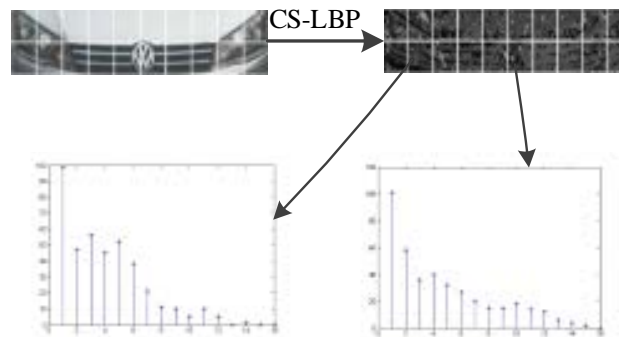

Fig.4 Statistical histogram of the block CS-LBP

After the CS-LBP feature of the whole image is obtained, it is fused with the HOG edge feature extracted by the original image to form the final feature of the image.

Algorithm1. Fusion with the block CS-LBP and original HOG feature

Step1. The CS-LBP feature extraction is performed on the gray scale image.

Step2. The feature extraction image is divided into sub-blocks with the same size and non-overlapping.

Step3. The histogram features of each sub-block image are counted, and the histogram features of all sub-block images are cascaded to form the CS-LBP eigenvector of the whole image, that is the CS-LBP feature.

Step4. The original HOG feature of the whole vehicle image is obtained by extracting the HOG feature of the original gray image.

Step5. The CS-LBP feature is fused with the original HOG feature to obtain the fusion feature, expressed as CS-LBPHOG(1,0).

Algorithm1 process is shown in Figure 5, the vehicle image feature extracted effect is shown in Figure 6.
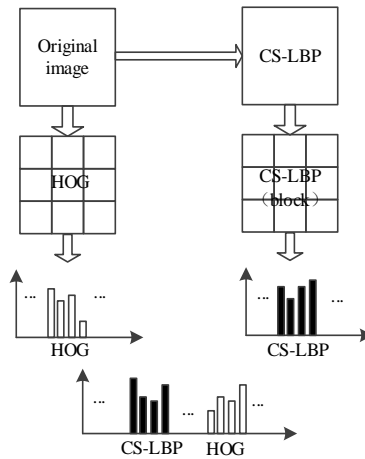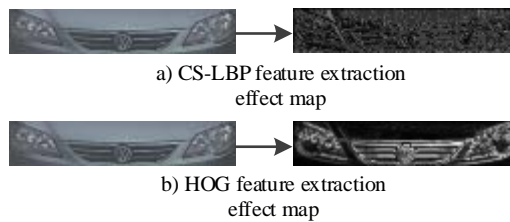
Fig.5 Algorithm 1 flow chart

a) CS-LBP feature extraction
effect map

b) HOG feature extraction
effect map

Fig.6 Algorithm 1 image extraction feature map

3.1 Fusion with the block CS-LBP and original HOG feature

The basic idea is to extract the CS-LBP feature to the image. Based on the CS-LBP feature image, the HOG feature is extracted and the original HOG feature is added to fuse the three features to form the final feature of the image.

Algorithm2. Fused block CS-LBP, original HOG with the HOG feature based on CS-LBP

Step1. It is the same as Step1 ~ Step4 of Algorithm1.

Step2. The feature of CS-LBPHOG is obtained by extracting the HOG feature from the CS-LBP image.

Step3. The CS-LBP feature, HOG feature and CS-LBPHOG feature are sequentially cascaded to obtain the fusion feature, expressed as CS-LBPHOG(1,1).

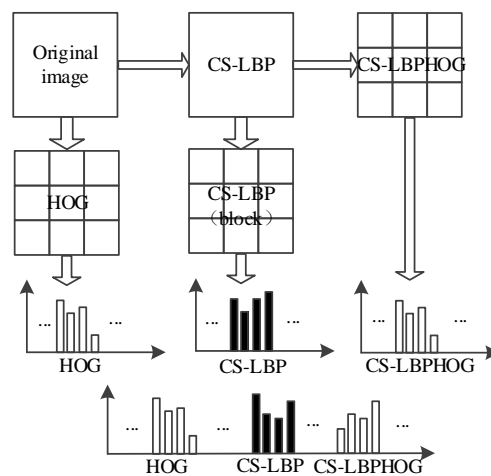Algorithm2 process is shown in Figure 7, the vehicle image feature extracted effect is shown in Figure 8.
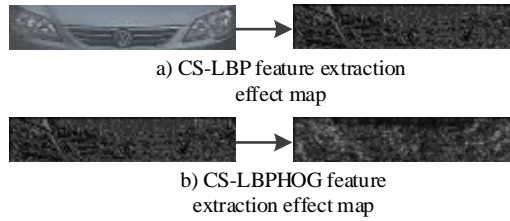

Fig.7 Algorithm 2 flow chart

a) CS-LBP feature extraction
effect map

b) CS-LBPHOG feature
extraction effect map

Fig.8 Algorithm 2 image extraction feature map

4. Experiment and analysis

The experimental environment is the Intel i5-4210M CPU, 2.60GHz, 8G memory and Win10 operating system, the programming platform is the MatlabR2014a software. The experimental data is the vehicle images from the expressway, including three brands of vehicles, a total of 240 images.

The vehicle images include the different illumination, angles and 20% range of scale changes, and they are preprocessed gray scale image with gray level of 256, and unified as 200×40 pixels. 80% of the entire vehicle image database is randomly selected as the training samples, the rest of the vehicle images as the test samples. Part of the vehicle images are shown in Figure 9.

Fig.9 Part of the vehicle sample images

In this paper, CS-LBP, HOG, CS-LBPHOG(1, 0) and CS-LBPHOG(1,1) are used to compare the recognition performance with the traditional PCA. The difference between the two images is determined by the similarity measure of the vehicle images' histogram features, and the commonly used Euclidean distance is chosen as the similarity measure. The European distance is calculated as follows:

$$d(S,M)=\sqrt{\sum_{i=1}^{n}(S_i-M_i)^2} \tag{8}$$

Where S is the histogram eigenvector of the test samples, M is the histogram eigenvector of the training samples, and n is the dimension of the eigenvector. Using the similarity measure method, the optimal classification classes of the vehicle images are searched by the K neighbor classifier (KNN).

4.1 Block size selection

Recognition rate of the vehicle images is compared in different block, such as 2×2, 2×4, 4×4 and 4×8, the recognition rate results are shown in Fig10.
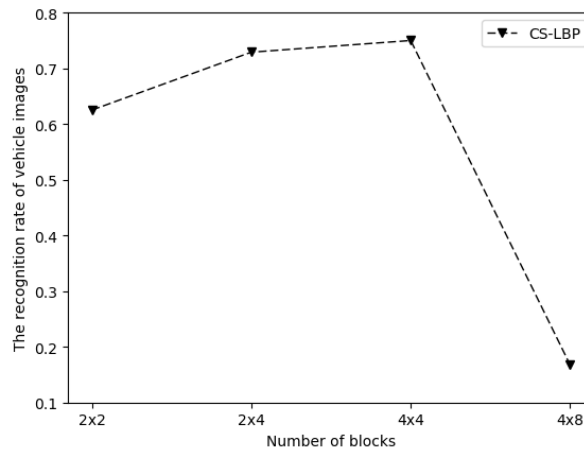
Fig.10 The recognition rate of different block sizes

It can be seen from the figure that when the block size is 4×4, the recognition rate of the vehicle image is the highest, so the number of image blocks is selected 4×4.

4.2 K value selection of classifier

In the same sample database, the vehicle image is selected with the optimal 4×4 block number. The recognition rate of different features is verified by the K nearest neighbor classifier with different K values, the experimental results are shown in Fig11.
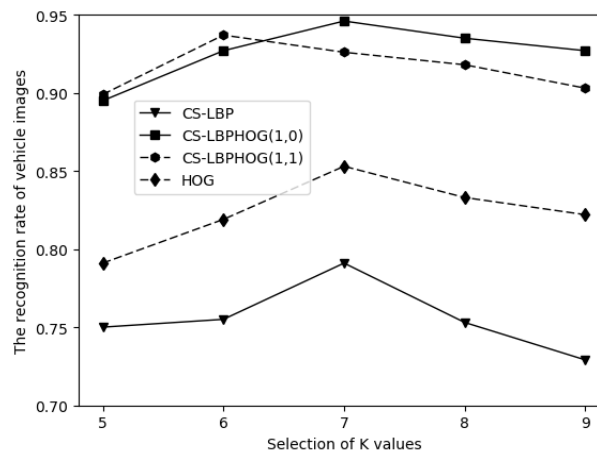

Fig.11 The recognition rate of different K values

As shown in Figure 11, the recognition rate of the vehicle images increases gradually with the K value increasing for the CS-LBP, HOG, CS-LBPHOG(1,0) and CS-LBPHOG(1,1), the recognition rate is gradually reduced when the K value continues to increase. In order to get the recognition rate of the vehicle images under four different features, the optimal K value of K nearest neighbor classifier is 7.

4.3 Comparison of experimental results

Based on the optimal number of blocks and the optimal parameter K, the dimensions, recognition rate and running time of CS-LBP, HOG, CS-LBPHOG (1, 0), CS-LBPHOG (1,1) are obtained. With the experimental results of PCA algorithm, the results are shown in Table 1.

Table 1 Experimental results of different algorithms

Algorithm	Dimension	Accuracy (%)	Time(ms)
PCA	174	65.1	19
CS-LBP	256	79.1	35
HOG	3456	85.3	167
CS-LBPHOG(1,0)	3712	94.6	174
CS-LBPHOG(1,1)	7168	92.6	305

As shown in Table 1, it can be seen that the time of recognition increases with the increase of the dimension. In contrast, the recognition rates of two fusion methods are higher than the traditional PCA, CS-LBP and HOG feature extraction methods. Among them, CS-LBP and original HOG feature fusion method has the highest recognition rate, up to 94.6%. The feasibility of the application of CS-LBP and original HOG feature fusion method in vehicle images' recognition is described.

5. Conclusions

The CS-LBP feature mainly expresses the texture information of the vehicle's face image, the HOG feature expresses the edge and contour information. The fusion method of CS-LBP and original HOG feature which is proposed in this paper can obtain a relatively complete image representation and the higher recognition rate. Compared with the predecessors' article, the recognition rate of vehicle's face images is slightly lower than the recognition rate of face images in the ORL face database, Yale face database and GT face database. The result is that vehicle images from the highway are affected by uneven illumination, multi-attitude, etc., reducing the influence of these factors can improve the recognition rate of vehicle's face images.

6. Acknowledgments

References

[1] Hua Liqin, Xu Wei, Wang Tuo. Vehicle Recognition Using Improved SIFT and Multi-View Model[J]. Journal of Xi'an Jiaotong University, 2013, 47(4):92-99.

[2] Zhang Peng, Chen Xiangjun, Ruan Yaduan. A Vehicle Classification Technique Based on Sparse Coding[J]. Journal of Xi'an Jiaotong University, 2015, 49(12):137-143.

[3] Huang Yi, Chen Xiangjun, Ruan Yaduan. The whitening-sparse coding vehicle classification algorithm for low resolution video[J]. Journal of Nanjing University, 2015,51(2):257-263.

[4] Zhang Jun, Sun Wenben, Yang Zhengling. Vehicle Recognition System Based on Mixing Characteristics of Contour Moment and Harris Corner[J]. Computer Applications and Software, 2016, 33(2): 142-149.

[5] Chen Lichi. Vehicle Make and Model Recognition Using Symmetrical SUER［C］// Advanced Video and Signal Based Surveillance (AVSS), Krakow, 2013 10th IEEE International Conference on, IEEE, 2013: 472－477.

[6] Sara Saravi, Eran A．Edirisinghe．Vehicle Make and Model Recognition in CCTV Footage[C]. Digital Signal Processing(DSP), 2013 18th International Conference on, IEEE, 2013: 1－6.

[7] M Saquib Sarfraz, et al．Bayesian Prior Models for Vehicle Make and Model Recognition[C]. Frontiers of Information Technology(FIT)，Proceedings of the 6th International Conference on, IEEE，2009.

[8] Zhang Hongbing, Li Hailing, Huang Xiaoting. Research and Implementation of Vehicle-Type Recognition Method Based on HOG Features of Vehicle Frontal Face[J]. Computer Simulation, 2015, 32(12): 119-123.

[9] Wan Yuan, Li Huanhuan, Wu Kefeng. Fusion with Layered Features of LBP and HOG for Face Recognition[J]. Journal of Computer-Aided Design & Computer Graphics, 2015，27(4):640-650

[10] Dai Jinbo, Xiao Xiao, Zhao Hongwe. Human face recognition based on low resolution local binary pattern[J]. Journal of Jilin University, 2013, 43(2): 435-438

[11] Zhang Jieyu, Zhao Hongping, Chen Shu. Face Recognition Based on Weighted Local Binary Pattern with Adaptive Threshold[J]. Journal of Electronics & Information Technology, 2014, 36(6): 1327-1333

[12] Shan C, Gong S, McOwan P W. Facial expression recognition based on local binary patterns: a comprehensive study[J]. Image and Vision Computing,2009,27(6): 803-816

[13] Heikkila M, Pietikainen M, Schmid C. Description of interest regions with local binary patterms[J]. Pattern Recognition, 2009, 42(3): 425-436.

[14] Déniz O, Bueno G, Salido J, et al. Face recognition using histograms of oriented gradients[J]. Pattern Recognitin Letters, 2011, 32(12): 1598-1