Vehicle-logo recognition based on the PCA

Qi Zheng[1], Hyo Jong Lee[1,2]
[1]Division of Electronics and Information Engineering
[2]Center for Advanced Image and Information Technology, Chonbuk National University

Abstract

Vehicle-logo recognition technology is very important in vehicle automatic recognition technique. The intended application is automatic recognition of vehicle type for secure access and traffic monitoring applications, a problem not hitherto considered at such a level of accuracy. Vehicle-logo recognition can improve Vehicle type recognition accuracy. So in this paper, introduces how to vehicle-logo recognition. First introduces the region of the license plate by algorithm and roughly located the region of car emblem based on the relationship of license plate and car emblem. Then located the car emblem with precision by the distance of Hausdorff. On the base, processing the region by morphologic, edge detection, analysis of connectivity and pick up the PCA character by lowing the dimension of the image and unifying the PCA character. At last the logo can be recognized using the algorithm of support vector machine. Experimental results show the effectiveness of the proposed method.

1. Introduction

Vehicle recognition is an important content of traffic system. With the rapid increase of the car, the vehicle identification research put forward higher request, through the license number and the models to identify whether the violations. The research on vehicle types, now confined in the recognition of large Lorries, medium-sized cars, small cars range [1-2]. The salient feature of the models is the vehicle logo, How to identify vehicle-logo is very important in vehicle recognition technique. Relevant vehicle-logo recognition algorithm: Based on energy enhancement and shape State of the positioning method to learn filter logo [3-4]; based on edge histogram fast car sign Identification method [5]. But vehicle-logo is small, there are a lot of decorations around vehicle-logo, and a lot of noises obstruct vehicle-logo positioning. In order to overcome interference, in this paper, according to the structure of the license plate and vehicle-logo relationship, first logo to rough positioning, then use Hausdorff distance invariance logo on accuracy position, On the base, processing the region by morphologic, edge detection, analysis of connectivity and pick up the PCA character by lowing the dimension of the image and unifying the PCA character. At last the logo can be recognized using the algorithm of support vector machine. Experimental results show the effectiveness of the proposed method.

2. Vehicle Logo Location

Vehicle-logo positioning technology is in the field of pattern recognition is a new topic, the number of research at home and abroad at present relatively little, available for reference material is relatively less. At present foreign study method of the commonly used within Based on texture consistency logo positioning method, Use the car emblem of the energy of the vertical edges car badge localization method, Based on the PCA and the same moment car badge localization method, Edge detection and morphology filter car badge positioning method, etc. In this paper, rough locating the vehicle-logo based on the position between the license plate and the vehicle-logo, on this basis used Hoff distance invariance logo on accuracy position.

2.1 Vehicle-logo coarse position

License logo and has a specific position, Logo is located in the plate region often just above, the head of the central area, this article use the two logo of characteristic of rough positioning [6-8].

(1) The vertical direction positioning

Logo and license plate in vertical direction meet the following position:

$$Y_2 = floor(y_2 - 1.3*(y_2 - y_1))$$

$$Y_1 = floor(y_1 - 1.8*(y_2 - y_1))$$

$$X_2 = x_2$$

$$X_1 = x_1$$

(2) The horizontal vehicle-logo positioning

Located in the upper part of the license plate logo，And the level of the center of the logo direction just and license plate horizontal center is coincide，Vehicle-logo level and the level of the license plate location meet the following relationship:

$$X_1 = (x_1 + (x + \frac{x_2}{2})) / 2.32$$

$$X_2 = (x_2 + (x + \frac{x_2}{2})) / 2.32$$

2.2 Based on the Hausdorff distance logo positioning

Hausdorff distance is mainly used in image registration areas, It does not need to build a point-to-point exactly corresponding relation, Only two calculation point set of similarity between degree. (Maximum distance)

The Hausdorff distances are defined as follows [9-11]:

Hypothesis space has two points set

$$A=\{a_1, a_2, ...a_p\}, B = \{b_1, b_2, ...b_p\}$$

Then set the Hausdorff distance between A B:

$$H(A,B) = \max(h(A,B), h(B,A)) \quad (1)$$

Type, h (A, B) and h (B, A) respectively for set A to set said B and set to set A one-way B Hausdorff distance

The mathematical expression as follows:

$$h(A,B)=\max_{a\in A}\min_{b\in B}\|a-b\| \quad (2)$$

$$h(B,A)=\max_{b\in B}\min_{a\in A}\|a-b\| \quad (3)$$

Compared to common distance, Hausdorff distance can real reaction objects relative position and shape. The shape of the car vehicle-logo most of elliptic shape and round, And there is some irregular shape the character.

3. Vehicle-logo feature extraction

3.1 Making vehicle-logo template library

This article through the collection and got a lot of search site for logo image. Such as masses, Audi, Citroen, Chevrolet, Buick, Toyota, Mazda, Peugeot, Honda, Hyundai, ford etc 20 template logo. The size of each template are 45 * 40 pixels, and after a gray histogram equalization and chemical processing, finally to extract logo template.

Fig.1 Vehicle-logo template figure

3.2 Vehicle-logo PCA feature extraction method

The principal component analysis is multivariate analysis of the PCA the oldest technical one, It comes from the theory of communication K-L transformation, 1901 by Pearson first proposed the main component analysis method Karhunan Loeve until 1963 to the method of the main problems in many of the changes, Finally it built the principal component analysis.

PCA algorithm is the main work calculation sample covariance matrix the eigenvalue and the eigenvalue vector, Vehicle-logo sample the size of the images are 45* 40 equals to the sample logo of the size of the sample matrix are n* d(n=40,d=45), The number of samples is m, So the sample covariance matrix S is a d*d square. Specific algorithm is as follows:

The principal component analysis algorithm is a concrete realization of can use the following steps to achieve[12-13]:

1) False for training practice sets

$$T=\{(x_1,y_1),(x_2,y_2),...,(x_i,y_i)\},x_i\in R^n,y_i\in\{-1,1\},i=1,2,...,n$$

And given the category that the ownership of the input value x, at the same time a given dimension reduction of dimension after less than the original dimension: d<n.

2) Construction set $\{x_1,x_2,\ldots,x_i\}$,calculation of that set covariance matrix

$$A=\frac{1}{I+1}\sum_{i=0}^{I}(x_i-\bar{x})(x_i-\bar{x})^T \quad (4)$$

x is a sample average and $x=\frac{1}{I+1}\sum_{i=0}^{I}x_i$ (5)

3) Covariance matrix of A with maximum D a characteristic value corresponding D a mutual orthogonal of characteristic vector unit $V_1, V_2,\ldots,V_d$.

4) Use for quantity $(V_1,V_2,\ldots,V_d.)$, to cast shadows of the group into torque array $V=\{v_1,v_2,\ldots,v_d\}$ and computing $x_i=V^T(x_i-\bar{x}),i=0,1,2,I$ and $\tilde{x}=\bar{x}$ and $\bar{x}....,\bar{x_I}$ is x and $x_1,x_2,\ldots,x_I$ dimension reduction of vector after.

Vehicle-logo sample repository contains 20 types of vehicle-logo each type of logo and contains 10 vehicle-logo samples are in these samples of different light under the background of different made selection of five of them as a sample training set another five as the test set. The training set the number of training for 20 * five samples of the dimension of 40 pixels * 45 pixel = 1800 pixels and then will a 1800 d row vector of a as sample matrix sample will last a sample matrix stored in a database, at the end of the day the sample size of 100 * 1800 matrix. Obviously if the training set directly into the support vector machine will bring a lot of computational complexity so need to use PCA algorithm will data dimension reduction processing here will get eigenvector of sample drop to 20 d that is each sample can use a 20 d characteristic vector to said.

3.3 Vehicle-logo PCA features normalization

The PCA dimension reduction Daniel had in the 20 D data is still not directly to identify classifier to data of standard processes will all the eigenvalues transform to provisions range.

Data normalization is the purpose of the characteristics of the attribute value will value scope of transformation to a particular range so as to eliminate the attribute value because of the different size range and influence the accuracy of classification algorithm

In the process of normalization of data in two of the most common method of big minimum normalization method zero mean the party method here the specifications minimum and maximum normalization method as normalization method formula is as follows [14-17]:

$$m'=\frac{m-\min A}{\max A-\min A}(new\_\max A-new\_\max A)+new\_\min A \quad (6)$$

4. Vehicle Logo Recognition

After PCA principal component mention, to take on quantity feature vector normalized extracting the need to identify the final logo feature vector below will extract characteristic vector into the support vector machine classifier

The sample of the 20 choose features of the first two components weight as the characteristic vector and Choose

the 18 class vehicle-logo samples，The category of the samples were 1 and 0.For the extraction of characteristic vector after normalized to support vector classifier for training，used the classifier for training produce test set tested. In 90 a logo image recognition, identify the correct number is 80 pieces, accuracy reached 89%, the experimental results are ideal.

5. Conclusion

The experimental data of color vehicle pictures all from network, the image size for 45 * 40 of the RGB JPEG format. After vehicle-logo coarse position and the Hausdorff distance logo positioning. We find vehicle-logo position, and then through the gray histogram equalization and chemical processing, finally to extract logo template Then use PCA logo library to dimensionality PCA features normalization, will be extracted vector into support vector classifier for training will produce of classifier training to the test set tested, then get the result. In 90 a logo image recognition, identify the correct number is 80 pieces, accuracy reached 89%, the experimental results are ideal.

In conclusion, verify the vehicle-logo recognition in the logo of precise positioning and principal component analysis in dimension reduction effectiveness in the process and support vector machine classification algorithm is feasible and high efficiency

Reference

[1] W ayne W, Justin W, EricK. Development of class models for model-based automatic target recognition. Proceedings of SPIE: The International Society for Optical Engineering, 1999, 3721: 6 5 0 -6 6 0.

[2] Wu W, An automatic system of vehicle number-plate recognition based on neural networks. Journal of Systems Engineering and Electronics, 2001, 12(2):63-72

[3] Li Guijun, Liu Zhengxi, You Zhisheng. Based on energy enhancement and morphological filtering logo positioning method, Application Research of Computers Computer Application Research 2005, 16(1):76-79

[4] Zhuang Yong, Yang Hongyu. A rapid vehicle-logo positioning method a method, Application Research of Computers Computer Application Research 2004, 41(6):1167-1171

[5] Luo Bin, Cao Gang. Based on edge histogram fast vehicles sign of recognition, Application Research Of Computers Computer Application Research 2004, 21(6):150-157

[6] Belhum eour P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Analysis and Machine Intelligence, 1997, 19(7): 711 -720.

[7] Zhao Y, Pinilla C, Valmori D, et al. Application of support vector machines for T-cell epitopes prediction [J].Bioinform,2003,19(15):1987-1984

[8] Brown M P S, Grundy W N, Lin D. Knowledge-based analysis of microarray gene expression Data by Using Support Vector Machines.Proc Nat`l Academy of Science, 2000,97:262-267.

[9] Dudoit S, Fridlyand J, Speed T P. Comparison of discrimination methods for the classification of tumors using gene expression data. J Am Statistical Assoc, 2002, 97:77-87.

[10] Howland P, Jeon M, Park H. Structure preserving dimension reduction for clustered text data based on the generalized singular value decomposition. J Matrix Analysis Applications, 2003, 25(1):165-179J.

[11] Furey T S, Cristianini N, Duffy N. Support vector machine classification and validation of cancer tissue samles using microarray expression data. Bioinformaticcs, 2000, 16:906-914.

[12] Zheng D, Zhao Y, Wang J. An efficient method of license plate location, Pattern Recognition Letters, 2005,26(15):3431-2438

[13] Nelson L J, License plate recognition, Advanced Imaging, 2002, 17(9):28-30.

[14] Rafael C G, Richard E W, Steven L E. Digital image processing using MATLAB. Beijing: Publishing House of Electronics Industry, 2005, 7:484.

[15] Wu W, An automatic system of vehicle number-plate recognition based on neural networks. Journal of Systems Engineering and Electronics, 2001, 12(2):63-72

[16] RAFAEL C G, RICHARD E W, STEVEN L E. Digital Image Processing Using Matlab, Beijing:Publishing House of Electronics Industry,2005:246－266

[17] Rafael C. Gonzalez Richard E. Woods. Digital Image Processing Second Edition. Beijing Publishing House of Electronics Industry 2006