www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

VIDEO-BASED VEHICLE DETECTION AND CLASSIFICATION IN CHALLENGING SCENARIOS

Yiling Chen[1], GuoFeng Qin[2*]

School of Electronics and Information

Tongji University, Shanghai 201804, China

[1]yilingchen5210@163.com, [2*]gfqing@aliyun.com

Submitted: Apr. 28, 2014 ***Accepted: July 15, 2014*** ***Published: Sep. 1, 2014***

Abstract- In intelligent transportation system, research on vehicle detection and classification has high theory significance and application value. According to the traditional methods of vehicle detection which can't be well applied in challenging scenario, this paper proposes a novel Bayesian fusion algorithm based on Gaussian mixture model. We extract the features of vehicle from images, including shape features, texture features, and the gradient direction histogram features after dimension reduction. In vehicle classification part, we adopt fuzzy support vector machine, and design a novel vehicle classifier based on nested one-vs-one algorithm. Finally, experimental tests show excellent results of our methods in both vehicle detection and classification.

Index terms*:* Vehicle detection, Gaussian mixture, Bayesian fusion, fuzzy SVM, vehicle classification.

I. INTRODUCTION

With the rapid development of social economic, intelligent transportation system (ITS) emerges. Although ITS is developing fast, it is still far from our goal. Vehicle detection and classification in a video has become a potential area of research due to its numerous applications to video-based intelligent transportation systems. According to different purposes algorithms can be divided into target foreground detection algorithm, target tracking algorithm, target classification algorithm, etc. Foreground detection algorithm includes frame difference method, background difference method, and optical flow method [1]. Background difference method is the most widely used one among them, and how to acquire a real-time updating background model is the core of the algorithm. Common updated background model algorithm can be divided into the following categories: average background mode, W4 method, vibe method, Gaussian background model, non-parametric background model [2], [3], etc.

In research fields of vehicle classification, due to high-dimensional feature of the algorithm, it has been interdisciplinary combinations of machine learning and data mining. As a result, many algorithms are widely studied, including unsupervised k-means algorithm, support vector machine (SVM) [4], neural network [5], particle swarm optimization (PSO) [6], Naive Bayesian classification algorithm, Boosting classifier, cascade classifier, etc. Classic support vector machine was originally used to solve binary classification problems, and decompose an m-class problem into a series of two-class problems [7], [8]. In a one-vs-all algorithm [9], for example, a C-class problem is converted to C Binary classification problems. To improve the accuracy of the algorithm, researchers use continuous decision function instead of discrete decision function [10]. Regardless of one-vs-all algorithm or one-vs-one algorithm, both have inseparable portions. In order to solve this problem, Abe [11]-[13] proposed support vector machine (SVM) based on fuzzy decision function, Platt proposed directed acyclic graph based on decision tree [14], Pontil and Verri [15], Kijsirikul and Ussivakul [16] proposed self-adaptive directed acyclic graph, Cheong [17] and Fei [18] proposed binary tree algorithm. As for the performance comparison between multi-class algorithm, Hsu [19], Kikuchi [20], Rifkin [21] and Galar did some important work.

Since traditional background modeling method is unsuccessful in many cases such as shadow casted by vehicles in challenging scenario, this paper proposes a novel algorithm. A Bayesian fusion algorithm based on block Gaussian mixture background modeling is used to extract target and remove noise. Finally, morphological operation is used to process the target vehicles. We use Gradient direction histogram descriptors in detected images, and also use principal component analysis method to reduce the dimension. Hence, Shape features, texture features and the gradient direction histogram features are used as vehicle features. Furthermore, linear discriminate analysis method is used to process the vehicle features. Moreover, we use fuzzy support vector machine, and design a vehicle classifier based on nested one-vs-one algorithm. Finally, experimental tests analysis our proposed methods by comparison.

II. IMPROVED BAYESIAN FUSION ALGORITHM

a. Traditional Gaussian mixture model

In detection and segmentation process of moving target, how to obtain the background model, which can simulate real scenes accurately, is the key for background extraction algorithm. Wren et al. [22] created a three-dimensional Gaussian model for per-pixel color value in handling the problems. The literature in which the model is updated through Kalman filters established a similar model. In terms of the complex outdoor scenes, Friedman and Russell [23], Stauffer and Grimson [24] both proposed the Gaussian mixture background model. It needs to establish k Gaussian distribution model for each pixel when classical Gaussian mixture model (GMM) is created. Recent sampling value of each pixel is: $\{\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3 \cdots\cdots, \mathrm{X}_t\} = \{\mathrm{I}(\mathrm{x}_{0,}\ y_0, i) \,|\, 1 \le i \le t\}$, we use k Gaussian function to approximate the sampling value sequences of image pixels, and the probability of pixel values is

$$P(\mathrm{X}_t) = \sum_{i=1}^{K} \omega_{i,t} \frac{1}{(2\pi)\dfrac{n}{2} \,|\sum_{i,t}|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t-\mu_{i,t})^T \sum_{i,t}^{-1}(X_t-\mu_{i,t})} \qquad (1)$$

Where X_t is the pixel value at time t, K is the number of Gaussian mixture model, $\omega_{i,t}$ is the weight of Gaussian model at time t, which is satisfied with $0 \le \omega_{i,t} \le 1$ and $\sum_{i=1}^{K} \omega_{i,t} = 1$, $\mu_{i,t}$ is the mean vector of Gaussian model at time t, $\sum_{i,t}$ is the model's covariance matrix. If

$$|X_t - \mu_{i,t-1}| \le 2.5 * \sigma_{i,t-1} \qquad (2)$$

The model updates its parameters

$$\begin{aligned} \omega_{i,t} &= (1-\alpha)\omega_{i,t-1} + \alpha \\ \mu_{i,t} &= (1-\rho)\mu_{i,t-1} + \rho X_t \\ \sigma_{i,t}^2 &= (1-\rho)\sigma_{i,t-1}^2 + \rho(X_t - \mu_{i,t-1}) \cdot (X_t - \mu_{i,t-1})^T \end{aligned} \qquad (3)$$

After the weights are updated, weights usually should be normalized. If no model which can match with current pixel value occurs, new model will be recreated. And if the model is the biggest one, we use new model to replace the model which is the $\frac{\omega}{\sigma}$ of the smallest mode among whole current models. The equation for background modeling is as follows

$$B = \arg_b \min(\sum_{i=1}^{b} \omega_{i,t} > T) \qquad (4)$$

b. Improved block background modeling and updating method

Since traditional GMM generally requires the establishment of K Gaussian distribution model for each pixel, the paper improves the model. Without prejudice to the results of vehicle classification, we use block modeling instead of point modeling. Therefore, it can clearly overcome the GMM's shortcomings, such as complex computing process and heavy computation. Traditional GMM is modeled on isolated points, and the induction of block model can obviously improve the rate of target detection, as well as consider the spatial information among pixels. Block modeling needs to process video in block, then replace pixel value with each block mean value, so as to build background model for each pixel.

Recent sampling points of pixels $\{X_1, X_2, X_3 ...., X_t\}$, are from continuous distribution $p(x)$, the definition of Gaussian function estimation at any point is

$$p_n(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\frac{(x-x_i)^2}{\sigma^2}} \qquad (5)$$

Where $\int p_n(x)dx = 1$, x means the pixel to be detected. Video frame is divided into $L \times L$ square sizes from left to right and from top to bottom at time t. 0 is used to fill it if the border is not wide enough. Let x_t be the block mean, then $x_t = \frac{1}{L \times L}\sum_{x=1}^{L}\sum_{y=1}^{L} I(x, y, i)$, where $I(x, y, i)$ means the pixel value at time t, wherein

$$p_n(x) = \frac{1}{n}\sum_{i=1}^{n}\omega_{i,t}\frac{1}{(2\pi)\frac{n}{2}|\sum_{i,t}|^{\frac{1}{2}}}e^{-\frac{1}{2}(X_t-\mu_{i,t})^T\sum_{i,t}^{-1}(X_t-\mu_{i,t})} \qquad (6)$$

Literature [25] proposed the parameter updating equations, and this paper updated the mean equation as follows

$$\mu_{i,t} = (1-\rho)\mu_{i,t-1} + \rho \qquad (7)$$

$$\text{Learning rate: } \rho = \alpha / \omega_{i,t} \qquad (8)$$

Wherein

$$\sigma_{i,t}^2 = (1-\rho)\sigma_{i,t-1}^2 + \rho(\mathrm{x}_t - \mu_{i,t-1}) \cdot (\mathrm{x}_t - \mu_{i,t-1})^T \qquad (9)$$

For no matching models, their mean and variance don't change. In this case, weight is updated as follows

$$\omega_{i,t} = (1-\alpha) \cdot \omega_{i,t-1} \qquad (10)$$

Where the learning rate is generally set to $\rho = 0.002$.

c. Target detection based on Bayesian fusion model

Let w_1 represent the background pixel, and w_2 represent the foreground pixel. Let the prior probability of background and foreground be denoted as $p(w_1)$ and $p(w_2)$, respectively, and the probability density function of background and foreground as $p(x/w_1)$ and $p(x/w_2)$, respectively.

The two types of pixels are classified based on the principal of maximum posterior probability, using the following decision rules

$$\mathrm{x} \in \begin{cases} w_1 & p(w_1 / x) > p(w_2 / x) \\ w_2 & p(w_1 / x) < p(w_2 / x) \end{cases} \qquad (11)$$

Applying Bayesian formula to the operations of posterior probability, we can acquire

$$p(w_i / x) = \frac{p(x / w_i)\, p(w_i)}{\sum_{j=1}^{2} p(x / w_j)\, p(w_j)}, i = 1, 2 \qquad (12)$$

Image pixels can be viewed as a Gaussian distribution, so we can calculate the independent probability density estimation of each category. The sum probability of background and foreground pixels can be expressed

$$p(x / w_i) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\frac{(x - x_i)^2}{\sigma_i^2}}, i = 1, 2 \qquad (13)$$

Based on the posterior probability formula and Gaussian density function, if you want to detect the moving object from video images, as well as extract the target images, you need to calculate the posterior probability ratio between the foreground pixels and background pixels. Set the ratio to the threshold of foreground pixels and background pixels, so the discriminant formula is given by

$$g(x) = \begin{cases} 0 & T > 1 \quad background \\ 1 & else \quad foreground \end{cases} \qquad (14)$$

where

$$T = \frac{p(w_1 / x)}{p(w_2 / x)} \qquad (15)$$

Calculate the value of threshold T from（12）and（13）

$$T = \frac{p(w_1 / x)}{p(w_2 / x)} = \frac{\dfrac{p(x / w_1)\, p(w_1)}{\sum_{j=1}^{2} p(x / w_j)\, p(w_j)}}{\dfrac{p(x / w_2)\, p(w_2)}{\sum_{j=1}^{2} p(x / w_j)\, p(w_j)}} = \frac{p(x / w_1)\, p(w_1)}{p(x / w_2)\, p(w_2)} \qquad (16)$$

The prior probability of background pixels and foreground pixels can be calculated by the formula (16), then we can detect target from given images. The whole procedure of our proposed algorithm can be summarized as the following steps:

(1) Calculate the prior probability of foreground pixels $p(x/w_1)$ and background pixels $p(x/w_2)$, respectively, according to the formula of block Gaussian kernel density formula. We can get the ratio: $t_1 = \dfrac{p(x/w_1)}{p(x/w_2)}$.

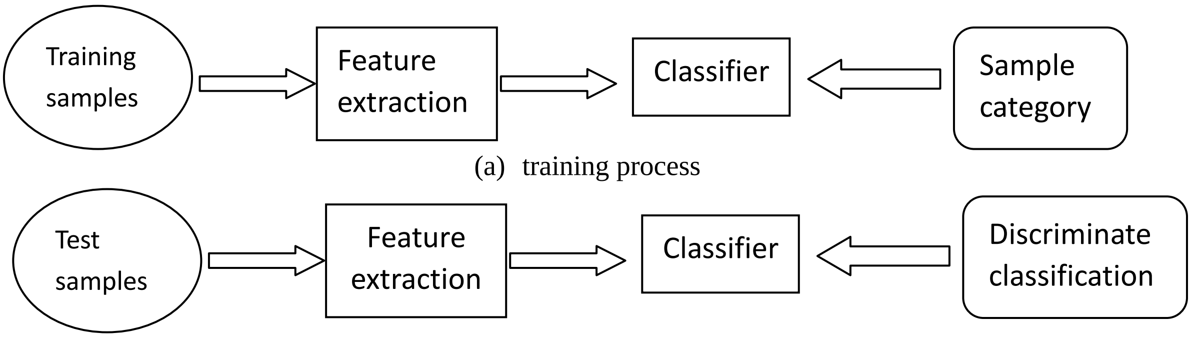
(2) Estimate the prior probability of background $p(w_1)$ and foreground $p(w_2)$, get the ratio $t_2 = \dfrac{p(w_1)}{p(w_2)}$.

(3) Get the threshold $T = \dfrac{p(w_1/x)}{p(w_2/x)} = t_1 t_2$.

(4) We can decide pixels according to the value of the threshold, T>1 is the background pixel, otherwise it belongs to pixels of moving vehicle to be detected.

III. IMPROVED MULTI-CLASSIFIER BASED ON SVM

General classification and identification problems can be described as follows: obtain the known samples, and take the known samples set as the training sample set, so the category of the training sample is known. According to the training process shown in figure 2.1(a), we conduct the classifier training process, so as to get the decision representation of the classifier. Then use the trained classifier to identify the classification of samples to be tested, the process is shown in figure 1(b).



(b) classification process

Figure 1. Process of training and classification

a. Fuzzy SVM

In practical application, a SVM tries to find the hyperplane leaving the largest possible fraction of points of the same class on the same side, while maximizing the distance of either class from the hyperplane. This paper proposes fuzzy SVM to deal with binary classification problems, which adopts fuzzy membership to measure its degree. Given a set of points which belong to either of two classes

$$(x_1, y_1), (x_2, y_2), \cdots, (x_l, y_l), \qquad x \in R^n, y \in \{-1, +1\}$$

Fuzzy into

$$(x_1, y_1, s_1), (x_2, y_2, s_2), \ldots, (x_l, y_l, s_l)$$

s_i is fuzzy membership, to measure the x_i degree of the category it belongs to. $\sigma \le s_i \le 1, i = 1, 2, \ldots, l$, small enough constant $\sigma > 0$. So the optimization case turns into

$$\min \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{l} (s_i \xi_i)$$

Which must satisfy with the following constraints:

$$\begin{cases} y_i (w \cdot \varphi(x_i) + b) \ge 1 - \xi_i \\ \xi_i \ge 0 \end{cases}, \qquad i = 1, 2, \ldots, \mathrm{l} \quad (17)$$

The optimization problem represented by correspondent Lagrange:

$$\begin{aligned} &\mathrm{L}(\mathrm{w}, \mathrm{b}, \xi, \alpha, \beta) \\ &= \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{l} (s_i \xi_i) - \sum_{i=1}^{l} \alpha_i (y_i (w \cdot \phi(x_i) + b) - 1 + \xi_i) - \sum_{i=1}^{l} \beta_i \xi_i \end{aligned} \qquad (18)$$

Find the saddle point of $\mathrm{L}(\mathrm{w}, \mathrm{b}, \xi, \alpha, \beta)$, and parameters must satisfy

$$\begin{cases} \dfrac{\partial L(w, b, \xi, \alpha, \beta)}{\partial w} = w - \sum_{i=1}^{l} \alpha_i y_i \varphi(x_i) = 0 \\ \dfrac{\partial L(w, b, \xi, \alpha, \beta)}{\partial b} = -\sum_{i=1}^{l} \alpha_i y_i = 0 \\ \dfrac{\partial L(w, b, \xi, \alpha, \beta)}{\partial \xi_i} = s_i C - \alpha_i - \beta_i = 0 \end{cases} \qquad (19)$$

substitute (19) into (18), we can obtain

$$\max_{\alpha} W(\alpha) = \max \{ \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \}$$

with constraints

$$\begin{cases} \sum_{i=1}^{l} \alpha_i y_i = 0 \\ 0 \le \alpha_i \le s_i C \end{cases}, \quad i = 1, 2, \ldots, l$$

For the binary classification problem, we can use the class center point to determine the value of s_i. Assume that the mean of class +1, -1 is x_+ and x_- respectively. So the radius of sample class +1 is

$$r_+ = \max_{x_i : y = 1} | x_+ - x_i |$$

the radius of sample class -1 is

$$r_- = \max_{x_i : y = -1} | x_- - x_i |$$

s_i is found to be

$$s_i = \begin{cases} 1 - |x_+ - x_i| / (r_+ + \delta), & if \quad y_i = 1 \\ 1 - |x_- - x_i| / (r_- + \delta), & if \quad y_i = -1 \end{cases}$$

$\delta > 0$, to avoid $s_i = 0$。

b. Nested algorithm based on one-vs-one strategy

Support vector machine (SVM) can be only used to solve the binary classification problem, and people often decompose a m-class problem into a series of two-class problems. In a one-vs-one algorithm, a C-class problem is frequently converted to C(C-1)/2 two-class classification problems. The ith classifier is trained by classifying the samples, which belong to the class i, into the positive class, as well as classfying other samples into negative class. Let the optimal hyperplane between class i and class j be denoted as

$$D_{ij}(x) = w_{ij}^T \Phi(x) + \mathrm{b}_{ij} = 0 \qquad (20)$$

$$D_{ij}(x) = -D_{ji}(x) \qquad (21)$$

When adopting the discrete discriminant function, for a given sample x, calculate

$$D_i(x) = \sum_{j \ne i, j=1}^{c} sng(D_{ij}(x)) \qquad (22)$$

Where

$$sng(D_{ij}(x)) = \begin{cases} 1, & D_{ij}(x) \geq 0 \\ 0, & D_{ij}(x) \geq 0 \end{cases} \tag{23}$$

Let

$$k = \arg\max_{i=1,\dots,C}(D_i(x)) \tag{24}$$

Then classify x into the class k. If many k values are equal, x is inseparable. In general, the algorithm exists inseparable portions, an example of which is shown in figure 2.

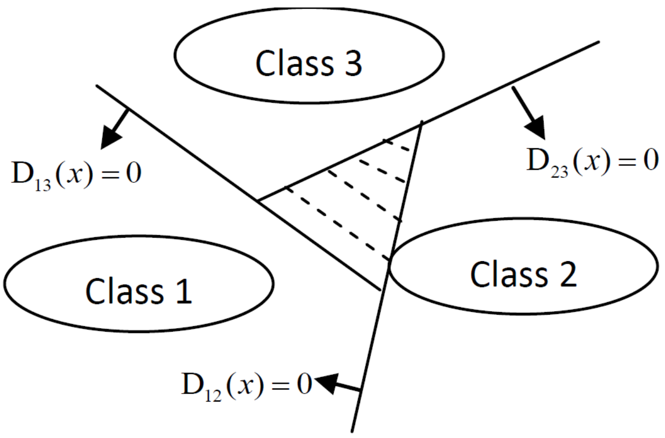


Figure 2. Distinguish results of one-vs-one algorithm

This paper describes a nested one-vs-one algorithm for three-kind classification issues, which aims at the drawback of one-vs-one algorithm in handling inseparable regional problems. The details as follows:

(1) For a C-class classification problem, we can construct C(C-1)/2 optimal hyperplanes based on one-vs-one strategy, where the optimal hyperplanes between class i and class j is given by (20). This paper aims at three-class vehicle classification problem, so we need to construct three optimal hyperplanes.

(2) For a given sample x, if k obtained by formula (24) is unique, sample x is classified into class k. Otherwise, we classify the samples into inseparable portions according to formula (24).

(3) If the number of samples in inseparable portions is not less than 3, we treat it as multi-classification problem, and use these samples to construct the hyperplane based on one-vs-one strategy.

(4) Repeat steps (2)-(3), until the inseparable portions only contain one or two class, even none.

(5) If the final inseparable region only contains samples which belong to only one class, we assign the region to this class. If the final region contains two classes, we make use of binary fuzzy

SVM to divide the portion, so as to classify the samples into correspondent classes.

IV. ALGORITHM IMPLEMENTATION AND RESULT INTERPRETATION

Videos used in this paper for vehicle detection and vehicle classification are traffic video streaming from the highway. In this paper, vehicle types are divided into three types: car (including sedan car, saloon car and jeep), van (including commercial vehicles and ordinary van) and bus, here we used C, V, B respectively. We extract more than 2000 key frames from the video streaming to build vehicle image library.

a. Experimental procedure

a.i Vehicle detection and key frame extraction

Preprocess the video image, then adopt the proposed improved Gaussian mixture algorithm which fused with Bayesian decision theory for vehicle detection. First step, use the block mixed Gaussian algorithm to set up background model, get the gray-scale probability distribution of each pixel point in the video images; Second step, extract the movement foreground with Bayesian decision theory; Third step, through comparison and analysis, determine whether the foreground is derived from the movement of the vehicle or the effects of noise, thus separateing the vehicle to be extracted and the background.

Select a fixed virtual line artificially in every frame of the input video, choose those frames as key frames, in which vehicle centroid goes across the virtual line. Then clip these key frames to make the vehicle as far as possible in the middle of the frame, so these images have the same size after clipping.

a.ii Feature extraction and dimensionality reduction

Extract the shape and texture features [26] of the clipping frames. Get 5 dimensional feature vector $v=(S,m,\sigma,\mu_3,e)$ which consists of the number of region pixels, the mean value, the standard deviation, the third order moment and the entropy. Then perform HOG feature [27] extraction on the frames. The frame is divided into cell of 8 x 8 pixel size, the number of usable blocks of the frame is (85/8-1)*(60/8-1)=9*6, among them, and x should be integer operations.

And use a (2 * 2 * 9) dimensional feature vectors to express each block. So the whole frame turned into a 9 * 6 * (2 * 2 * 9) = 1944 - dimensional feature vector.

Built a matrix using the frame HOG feature vector of the entire training samples, each line represents a sample. After dimensionality reduction with PCA [28] turns into 21 dimension vector, add five characteristics for shape and texture description, we get a 26 dimension feature vector, then calculate the within class scatter matrix S_w and the between class scatter matrix S_w by LDA algorithm [29], finally, take the first four feature vector of the matrix $S_w^{-1}S_b$ as the projection matrix of LDA feature subspace. So the training samples and test samples are both expressed in 4 dimension vector.

a.iii Selection of kernel functions and determination of parameters of SVM

Using support vector machine (SVM) for classification, the selection of kernel function is the key factor of the accuracy of the classification performance. There are four common kernel functions: Linear kernel, polynomial kernel, radial basis kernel (RBF kernel) and Sigmoid kernel. Among them, the radial basis kernel function corresponding to the nonlinear mapping can deal with the situation in which samples are nonlinear and separable. So this paper chose radial basis kernel function $K(x_i, x) = \exp(-\gamma \| x_i - x \|^2)$.

Since basic SVM classifier is binary classifier, in order to select the kernel function as soon as possible, we directly choose cars and buses which are quite different as test objects. In this paper, initially set γ∈ [2-10, 22], C∈ [1, 100], and then set parameters by using the open source software LibSVM toolkit of Chih-jen Lin, a professor at Taiwan university, to run in Matlab. Finally, we get the best parameter combination: (γ, C) = (0.001, 10).

a.iv Design of the three-class classifier

First, we construct 3 one-vs-all classifiers, namely (C, B + V) (cars as positive class, vans and buses as negative class), (B, C + V) and (V, C + B), choose radial basis kernel as the kernel function. Then, use the proposed nested algorithm based on one-vs-one strategy in this paper to classify vehicles.

The flow chart of video-based vehicle detection and classification algorithm is shown in figure 3.

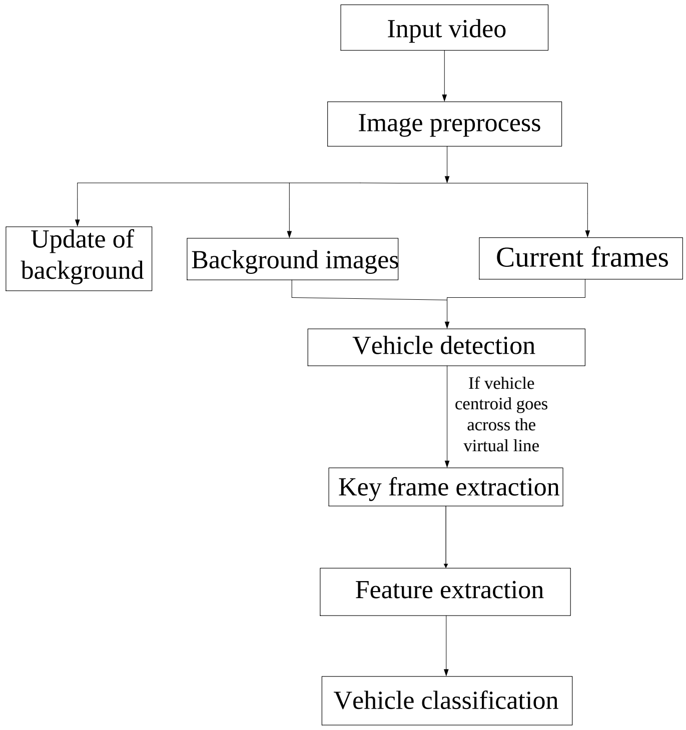


Figure 3. Process of synthesized algorithm for vehicle detection and classification

b. Experimental results and analysis

In order to prove the validity and advantages of the improved algorithm proposed in this paper, we implement the described algorithms by using OpenCV on VS2010 platform to simulate and analyze. In the experiment, there are two traffic video streaming in two difference scenes adopted as experimental data for the contrast test. Scene1 is the street mixed with cars and people, and Scene2 is an expressway during the peak period. Compared with different detection algorithm, the experiment adopts traffic flow algorithm based on virtual loop [30]. Figures 4 and 5 respectively show the tracking effect in Scene1 and Scene2. What's more, the statistical results of the two algorithms are shown in table 1.

(a) original image (b) background image (c) target tracking image

Figure 4. Effect drawings of the program running on Scene1

(a) original image (b) background image (c) target tracking image

Figure 5. Effect drawings of the program running on Scene1

Table 1: Statistical results of traffic flow

the recognition rate （%）	**Scene 1**	**Scene2**
Classic background difference algorithm	98.4	94.3
Our improved Bayesian fusion algorithm	98.6	98.1

Table 1 illustrates the recognition rate using different algorithms. It can be seen from Table 1 that our proposed detection algorithm is unsatisfactory in Scene1. Whereas in Scene2, aiming at complicated background and bad influence factors(such as shadows and occlusion), the improved vehicle detection algorithm can extract the moving object more accurately compared with traditional algorithm, and the recognition rate of our algorithm is improved greatly.

In order to prove the improved performance of our classification algorithm, we compared with the classification performance of different vehicle features as shown in figure 6.

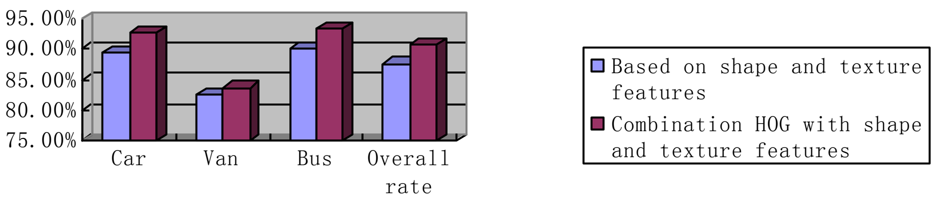


Figure 6. Comparison results of the recognition rate

Figure 6 shows that the vehicle characteristics including HOG features, as well as shape features and texture features, improves the recognition accuracy, thus it verified the feasibility and correctness of our proposed algorithm.

In order to prove the validity of our nested one-vs-one algorithm, we make comparison among our algorithm, one-vs-one algorithm, and directed acyclic graph （DAG）algorithm on different training sets and test sets. Numbers of experimental data, which are selected from vehicle image library, are respectively 500, 900, 1300, 1700, 2000, regarding 2/3 as training sets and the remain 1/3 as test sets. The experimental result is shown in figure 7.

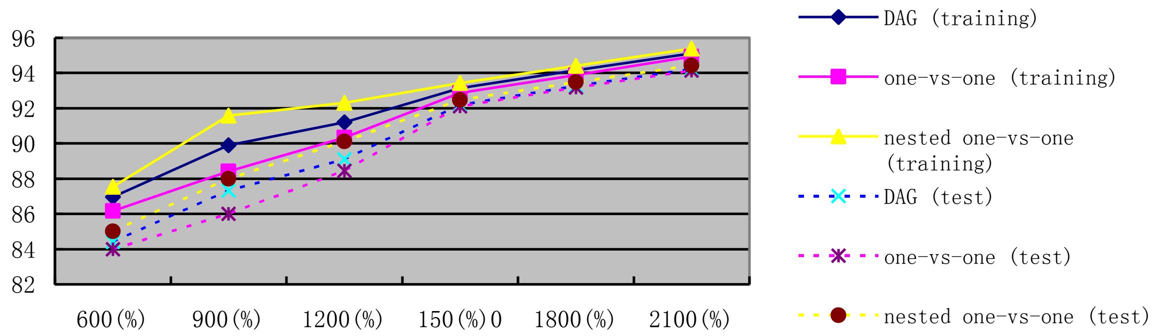


Figure 7. Classification accuracy of different algorithms

Figure 7 illustrates that for the whole date set, the accuracy of the classification performance of our nested one-vs-one algorithm is higher than both one-vs-one algorithm and directed acyclic graph algorithm, which is also verified the feasibility and correctness of our proposed nested one-vs-one algorithm based on fuzzy SVM.

V. CONCLUSIONS

For the background model, dynamic updating of background and the target extraction problem in complex background of the process of traffic video inspection, this paper puts forward to an improved detection method of moving vehicle based on Gaussian mixture algorithm and Bayesian decision theory. However, our proposed method is still computationally highly expensive and become unsuccessful in many cases such as occlusion among the vehicles, which leads to significant restriction to the detection and identification of vehicle; furthermore, the randomness of external environmental factors increases the difficulty of identification. In practical complex scenarios or bad weather (such as rain, snow, etc.), the classification result of vehicle is not very satisfying. In addition, this article failed to consider the color and the three-dimensional characteristics of the vehicles, etc. These deficiencies reveal the development direction of future research, such as the research of robust algorithm which was more adaptability [31], set up more complex mathematic model [32][33][34], Improve the accuracy and speed up the processing speed are also the research direction of this technology.

Acknowledgments: The researcher is supported by the national 863 program in Ministry of Science and Technology of the People's Republic of China. Project number:2013AA040302.

REFERENCES

[1] Castrillon F, Guin A, Guensler R. Comparison of Modeling Approaches for Imputation of Video Detection Data in Intelligent Transportation Systems[J]. Transportation Research Record: Journal of the Transportation Research Board, 2012, 2308(1):138-147.

[2] S.-C.Cheung, C.Kamath. Robust techniques for background subtraction in urban traffic video [J]. In Proe. Video Communication and Image Processing, 2004, 5308:881-892.

[3] Haritaoglu, D.Harwood, L.Davis. W4:Real-time surveillance of people and their activities [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000,22(8):809-830.

[4] Jin Mei-Shan, Meng Qing-hui,Huang Fei,Li. Research on the moving vehicle detection technology[C]. The IEEE 2nd International Conference on Consumer Electronics, Communications and Networks, 2012:2046-2048.

[5] Hai Wang, Yingfeng Cai, Long Chen, Yu-Bo Yuan. A vehicle detection algorithm based on deep belief network [J]. The scientific world journal, 2014.

[6] Ruochen Liu, Yangyang Chen, Licheng Jiao, Yangyang Li. A particle swarm optimization based simultaneous learning framework for clustering and classification [J]. Pattern Recognition, 2014, Vol.47(6).

[7] Cortes C, Vapnik V N. Support-vector networks. Machine Learning, 1995, 20:273-297.

[8] Vapnik V N. The nature of statistical learning theory. IEEE Transactions on Neural Networks, 1990, 10(5):988-999.

[9] Vapnik V N. The nature of statistical learning theory. London, UK: Springer-Verlag,1995.

[10] Vapnik V N. Statistical learning theory. New York: John Wiley & Sons, 1998

[11] Inoue T, Abe S. Fuzzy support vector machines for pattern classification. Proceedings of International Joint Conference on Neural Networks, 2001: 1449-1454.

[12] Kikuchi T, Abe S. Error correcting output codes vs. fuzzy support vector machines. Proceedings of Artificial Neural Networks in Pattern Recognition, Florence, Italy, 2003:192-196.

[13] Tsujinishi D, Abe S. Fuzzy least squares support vector machines for multiclass problems. Neural Networks, 2003, 16(5-6):785-792.

[14] Platt J C, Cristianini N, Shawe-Taylor J. Large margin DAGs for multiclass classification// Solla S A, Leen T K, MAuller K R. Advances in Neural Information Processing Systems. The MIT Press, 2000:547-553.

[15] Pontil M, Verri A. Support vector machine for 3-D object recognition. IEEE transactions on Pattern Analysis and Machine Intelligence, 1998, 20(6):637-646.

[16] Kijsirikul B, Ussivakul N. Multiclass support vector machines using adaptive directed acyclic graph. Proceedings of International Joint Conference on Neural Networks,2002:980-985.

[17] Cheong S, Oh S H, Lee S Y. Support vector machines with binary tree architecture for multi-class classification. Neural Information Processing-Letters and Reviews,2004,2(3):47-51.

[18] Fei B, Liu J B. Binary tree of SVM: a new fast multiclass training and classification algorithm. IEEE Transactions on Neural Networks, 2006,17(3):696-704.

[19] Hsu C W, Lin C J. A comparison of methods of multiclass support vector machines. IEEE Transactions on Neural Networks, 2002, 13(2):415-425.

[20] Kikuchi T, Abe S. Comparison between error correcting output codes and fuzzy support vector machines. Pattern Recognition Letters, 2005,26:1937-1945.

[21] Rifkin R, Klautau A. In defense of one-vs-all classification. Journal of Machine Learning Research, 2004,5:101-141.

[22] Wren C R, Azarbayejani A, Darrell T. Pfinder: real-time tracking of the human body[J]. IEEE Transactions on PAMI, 1997, 19(7):780-785.

[23] Friedman N, RussellS. Image segmentation in video sequences: a probabilistic approach[C]// Proceedings of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence. San Francisco, CA: Morgan Kaufmann Publishers, 1997:175-181.

[24] Stauffer C, Grimson W E L. Adaptive background mixture models for real-time tracking[C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Cambridge, MA: IEEE Computer Society, 1999:1063-1069.

[25] Wang Hong-zhi and Dong Ying. An improved image segmentation algorithm based on Otsu method[C]. Proceedings of the SPIE, 2007, 6625:662501-08.

[26] H. Moon, R. Chellappa, A. Rosenfeld. Performance analysis of a simple vehicle detection algorithm [J]. Image and Vision Computing, 2002, 20(no.1):1-13.

[27] Navneet Dalal, Bill Triggs. Histograms of oriented gradients for human detection. In: Anon. Proceedings of Conference on Computer Vision and Pattern Recognition. San Diego, California, USA.2005. New York: IEEE Computer Society Press, 2005. 556-893.

[28] Jae-Young Choi, Young-Kyu Yang. Vehicle detection from aerial images using local shape information[J]. Advances in Image and Video Technology, 2009:227-236.

[29] Xiao-Yuan Jing, David Zhang, Yuan-Yan Tang. An improved LDA approach[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2004, 34(no.5):1942-1951.

[30] Wang B, Jia K B. The traffic flow statistic algorithm based on video analysis[J]. Traffic information and security, 2010(001):20-25.

[31] Sang-Ho Kim, Kosuke Sekiyama, Toshio Fukuda. Pattern Adaptive and Finger Image-guided Keypad Interface for In-vehicle Information Systems. In International Journal on Smart Sensing and Intelligent Systems, v1, n3, p572-591, September 2008.

[32] A.-M. Cretu, P. Payeur. Visual Attention Model with Adaptive Weighting of Conspicuity Maps For Building Detection in Satellite Images. In International Journal on Smart Sensing and Intelligent Systems, v5, n4, p742-766, December 2010.

[33] G.Sengupta, T.A.Win, C.Messom, S.Demidenko and S.C.Mukhopadhyay, "Defect analysis of grit-blasted or spray printed surface using vision sensing technique", Proceedings of Image and Vision Computing NZ, Nov. 26-28, 2003, Palmerston North, pp. 18-23.
[34] Henry X. Liu, Jie Sun. Length-based vehicle classification using event-based loop detector data. Transportation Research Part C, 2014, Vol.38.
[35] Qianwei Zhou, Baoqing Li, Zhijun Kuang, Dongfeng Xie, Guanjun Tong, Liping Hu, Xiaobing Yuan. A quarter-car vehicle model based feature for wheeled and tracked vehicles classification. Journal of Sound and Vibration, 2013, Vol.332(26).