Research on Moving Pedestrian Identification Based on Bayesian Classifier

Wu Hua-zhi

(School of Automotive Engineering, Shanghai University of Engineering Science, Shanghai, China)

Abstract: *Pedestrian identification systemplays an important role in vehicle safetyassistant driving technology. In order toidentify the movement of pedestriansmore quickly and easily, motion pedestrian identification is carried out based on the Bayesian classifier. The identificationprocess was realized in two steps. Firstly, Bayesian classifier is designed based on the HOG feature. Secondly, the image captured by the camera is processed by taking the five-frame difference methodto extract the identification target from the pre-processed images and the HOG feature of the identification target is extracted as the input of the classifier to identify the target. The results show that this method can achieve the goal of identifying pedestrians quickly and simply.*
Key words: *pedestrian identification; moving target; HOG features; Bayesian classifier*

I. Introduction

The development of road traffic by leaps and bounds is accompanied by soaring traffic accidents, which emerges as a severe problem for traffic management, with a lot of accidents caused by pedestrians' failure to observe traffic regulations and crossing the driveway, leading to a proportion of more than 90% of pedestrians death toll caused by this kind of accidents in the total pedestrians death toll in traffic accidents[1]. Hence pedestrian identification technology has always been a very active realm in research on intelligent traffic and intelligent monitoring control systemand it has attracted increasing attention by the researchers. Among a host of sensors of detecting pedestrians, vision sensor is widely applied due to its ability to collect substantive rich and useful information.

Presently the pedestrians detection techniques comprise method based on shape information[2] , method based on template matching [3] and method based on statistical learning [4] . The method based on shape information utilizes pedestrians' shape features to recognize them, yet the polytrophic of pedestrians' shape leads to high difficulty for detection. The method based on template matching is matching the temples with characteristics of images to be analyzed, which is complicated and has very high requirements on the details and accuracy of pedestrian temples.The method based on statistical learning is to learn changes of human body from enormous sample set, which is widely popularized and can be applied with better detection effect.

The method based on statistical learning mainly comprises test method based on such classifiers as support vector machine(SVM), neural network, AdaBoost and Bayes, etc. Being a statistical learning method based on norm of minimizing structural risk [5], SVM classifier is extensively used in various classification problems. Oren M [6] first suggested SVM for pedestrian detection by combining SVM and Haar wavelet characteristics and established a pedestrian detection system to detect the pedestrian in static images. Apart from this,combination betweenGabor characteristics [7], shape description characteristics [8] , characteristics of histogram of oriented gradients [10] and SVMhas been successfully applied to pedestrians detection system. Neural network conceals characteristics of pattern in some pre-set network parameters via learning to describe complicated characteristics. Such characteristics make it widely applied in the domain of pedestrian detection technique. Zhao Liang et al.[10] used gradient magnitude of image as characteristics and adopted trilaminar feed forward neural network as classifier. The reason to adopt gradient characteristic is to eliminate the influence of rays, also to solve difficulty in selecting threshold in binarizing.Single classifiers, neither SVM classifier nor neural network classifier, cannot deliver satisfactory effect in some instances. Cascading multiple classifiers can get powerful classifier system. Therefore, combining weak classifiers into a strong classifier, i.e. AdaBoost cascade classifier, is a typical combination mode of multiple classifiers. In recent years, combining multiple common classifiers together through specific logical mode into a new classifier system has become a widely used research technique [11]. Besides, in recent years the directions in the research by many researchers also include classification method of combining characteristics of histogram of oriented gradients and scale invariant feature transform (SIFT) with AdaBoost due to the former two's invariant scale [12] . Significant drawbacks of AdaBoost classifier are that algorithmscomplicated, and the training of classifier is very trivial and time-consuming. As a reasoning method under uncertainty, Bayes classifier is based on probability and statistic theories and has a solid mathematical foundation, so it is a simple and convenient algorithm and has high accuracy in identification result. In a lot of domains, the researches of Bayesian theory have made achievements, which can be roughly divided into two categories: one is to use Bayesian theory in conjunction with other identification methods [13] ; another is to

---

directly use Bayes classifier to classify and recognize [14] . The former has shortcomings of complicated algorithm and low instantaneity. Therefore, the paper adopts the latter. Firstly a Bayes classifier using HOG characteristics is designed, then moving target is extracted from the preprocessed image using five-frame difference method, followed by extracting HOG characteristics of moving target and inputting HOG characteristics to trained Bayes classifier for identification and classification.

II. Design of Bayes classifier

2.1 Principle of Bayes classifier

Bayes' classification is an inducing classifier through training of training set and using classifier to classify the unclassified data. Representative Bayes classifiers include naive Bayes classifier, Bayes network classifier and tree augmented naive Bayes classification model TAN, etc. The paper adopts naive Bayes classifier. The basic idea of naive Bayes classification is: defining the probability for the classification to which the sample belongs to by calculating posterior probability to expertly connect prior probability and posterior probability of event, then using prior information and sample data information to define posterior probability of event, lastly classifying the object as the classification with the largest posterior probability. Naive Bayes classification is based on a simple assumption, i.e. for the given sample set, the probability distribution of each classification is known, and properties have a mutually independent influence on classification.

The paper divides moving targets into two classifications, i.e. pedestrians and non-pedestrians, represented by C_1 and C_0. Assume the eigenvector of sample to be classified X is $\{x_0, x_1\}$, calculate the probability of X belonging to each classification $P(C_i, X)$ (i=0,1). Then if $P_1(C_i, X) > P_0(C_i, X)$, the sample to be classified X is a pedestrian, otherwise the sample is non-pedestrian. Thereinto, the calculation formula of $P(C_i, X)$ is as follows:

$$P(\mathrm{C}_i \mid X) = \frac{P(\mathrm{X}|\mathrm{C}_i)P(\mathrm{C}_i)}{P(\mathrm{X})} \tag{1}$$

As P(X) are all the same for all C_i, so the paper lets P(X)=1, hence it is just needed to compare the numerator $P(\mathrm{X}|\mathrm{C}_i)P(\mathrm{C}_i)$. $P(\mathrm{C}_i)$ can be obtained by training set, and this value is equal to the proportion of the sample whose training set classification is C_i. The paper takes $P(C_0) = P(C_1) = 0.5$, then what need to obtain is the value of $P(\mathrm{X}|C_i)$ via training set. $P(\mathrm{X}|\mathrm{C}_i)$ Means that the represented probability density function is conditional, i.e. given the sample to be measured belongs to classification C_i, the probability density for appearing of eigenvalue X of sample to be measured is also the distribution rule of eigenvalue X in sample of classification C_i. Now it is needed to select the form of conditional probability density function. Previous research shows that most of the statistical data in many problems satisfies normal distribution rule, plus its simple formulas and fewer parameters, it is very suitable for statistical analysis. So the paper selects normal distribution density function as the statistical calculation form of conditional probability density function.

2.2 Training of Bayes classifier

In training of Bayes classifier, the keys are methods of characteristics selection and extraction. The former has an impact on accuracy of classification result of classifier, while the latter influences the time of each training of classifier. For selection of pedestrians' characteristics, the commonly used pedestrians' characteristics include SIFT characteristics, Haar-like characteristics and HOG characteristics, etc. Thereinto, SIFT is a local histogram descriptor which is unrelated to size and gyration of image. Haar-like characteristics are calculated from a kind of templates composed of black and white rectangular blocks, including edge characteristics, linear characteristics, central characteristics and diagonal characteristics, describing change of grey scale of image. HOG characteristic is gradient orientation histogram characteristic which is an operator combining local characteristics and distributional characteristics, able to express edge information of human body and has strong robustness for directional shift. The paper selects HOG characteristics as the training characteristics of Bayes classifier, as the classification objects of the paper are the objects moving horizontally in front of vehicles on road, mainly including pedestrians and vehicles. Thus HOG characteristics are enough to differentiate them. For HOG characteristics, the paper adopts rectangular structure, whose units of composition are blocks each of which is composed of four cells. And each cell is composed of pixels of 4 * 8. The HOG character extraction can be divided into five steps:

1) Input image

The input image is the picture in sample database or picture of ROI in pedestrians partition, read it in characteristics extraction window, with image size being 32*64;

2) Gradient calculation

For each pixel (x,y) in the given image I, we solved its horizontal gradient and vertical gradient respectively, and solved its gradient and gradient direction:

Gradient in horizontal direction:

$$G_x(x, y) = [-1\ 0\ 1] * \mathrm{I}(x, y) \tag{2}$$

Gradient in vertical direction:

$$G_y(x, y) = [-1\ 0\ 1]^T * I(x, y) \qquad (3)$$

So the gradient size of pixel (x,y) is:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \qquad (4)$$

Gradient direction is:

$$\theta(x,y) = \frac{G_y(x,y)}{G_x(x,y)} \qquad (5)$$

3) Vote the gradient magnitude for all pixels in each cell to the gradient orientation for corresponding pixel to form the gradient histogram characteristics of one cell. The gradient orientation histogram of one cell is obtained by voting on Y axis with gradient magnitude, with gradient sampling direction as X axis, on the basis of gradient magnitude and gradient orientation. The commonly used scope of sampling direction for gradient orientation is 0-360°and 0-180°. The experiment shows that the scope of 0-180° for human detection, which ignores negative and positive polarity of degrees, can get better effect; while sampling direction decides detection performance and complexity of calculation. Fewer sampling directions cannot describe well the characteristics of pedestrians, while too many sampling directions cannot realize better detection performance but will increase calculation load. The experiment shows that 9 sampling directions (one every 20°) can both satisfy the requirements on pedestrian detection performance and instantaneity.

4) Normalize gradient histogram characteristics in one block. The paper adopts following normalization method: L2-normalization:

$$v^* = \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}} \qquad (6)$$

Where v represents eigenvector before normalization,$\|v\|$represents its k order norm (k=1,2), ε represents a very small constant (its value does not influence the result, but is for the purpose of preventing denominator from being 0). Normalizing blocks aims to eliminate the influence of illumination and shadow. Besides, blocks can mutually overlap.

5) Take a block composed of 2*2 cells, i.e. composed of 8*16 pixels. The sliding block traverses window, sliding step is a unit. Then there are altogether 49 blocks for a 32*64 image. Connect the vectors of these blocks to generate a 1784-dimension vector, which is just HOG characteristic. In extracting sample's characteristics, the paper adopts Grubcut algorithm to extract the target in sample picture, which can dramatically reduce the training time of classifier each time, thereby simplifying algorithm, and improving accuracy of algorithm. Fig.1 below is the original drawing of some samples and concrete effect drawing after the training target is extracted.

Fig.1. Original sample and target extraction.

Bayes classifier inputs Train Samples, Test Samples and Class Number. Train Samples extracted HOG characteristics of 228 pictures, Class Number is 2, being respectively for pedestrians and non-pedestrians.After designing the Bayes classifier, its validity is experimented on using pictures of pedestrians and non-pedestrians, with result showing this classifier has accurate identification result. The identification result is as shown in Fig.2 below:
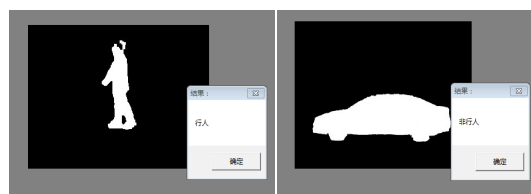

Fig.2. Bayesian classifier test results.

III. Extraction of moving target

The object of identification in this paper is the pedestrians crossing the street in front of the vehicles. To separate the pedestrians from background, five-frame difference method is used to extract moving targets by dint of the characteristic of pedestrians being moving. There are three concrete steps: image preprocessing, five-frame difference and morphological processing.

3.1 Image preprocessing

In normal condition, the original image obtained by sensor has poor quality due to influence of noise, illumination and other factors, which entails corresponding processing for improvement. The image preprocessing operation can effectively filter noise, enhance useful information and enable following treatment to go on wheels. For characteristics of image collected by camera being darkish and having noise, the paper equalizes image's greyscale and makes medium filtering, with processing result as shown in Fig.3 below:

(a) the original image (b) the preprocessed image
Fig.3. Original image and pre-processed image.

3.2 Five-frame difference method

Five-frame difference method was put forward based on traditional three-frame difference method which though has advantages of easiness to be realized and low time complexity, etc.Yet with target profile tending to be discontinuous and prone to have a lot of holes, plus failure to extract complete target information in scenes of changing illumination or background disturbance, etc. The five-frame difference method adopted in this paper mainly has following 2 steps:

1) Select adjacent images of five frames $f_1(x,y)$, $f_2(x,y)$, $f_3(x,y)$, $f_4(x,y)$, $f_5(x,y)$, as shown in figure 2, subject them to medium filtering and grey scale equalization respectively, remove Gaussian noise and salt and pepper noise to get corresponding images of five frames $I_1(x,y)$, $I_2(x,y)$, $I_3(x,y)$, $I_4(x,y)$, $I_5(x,y)$.

2) Take the third frame image as current frame, calculate frame difference from the 2 front frames and 2 rear frames.

$$d_1(x,y) = |I_3(x,y) - I_1(x,y)| \tag{7}$$

$$d_2(x,y) = |I_3(x,y) - I_2(x,y)| \tag{8}$$

$$d_3(x,y) = |I_3(x,y) - I_4(x,y)| \tag{9}$$

$$d_4(x,y) = |I_3(x,y) - I_5(x,y)| \tag{10}$$

Binarize the difference results $d_i(x,y)(i=1,2,3,4)$ obtained respectively, with threshold of binarization being taken as 0.1. To eliminate target overlapping, make AND operation for above result of binarization$h_i(x,y)$:

$$g_{14}(x,y) = h_1(x,y) \&\& h_4(x,y) \tag{11}$$

$$g_{23}(x,y) = h_2(x,y) \&\& h_3(x,y) \tag{12}$$

Where "&&" represents"AND"operation for image, to inhibit the holes caused by "AND" operation, make "OR" operation for$g_{14}(x,y)$and $g_{23}(x,y)$to get the targe profile of the third frame image:

$$B_3(x,y) = g_{14}(x,y) \| g_{23}(x,y) \tag{13}$$

The result of above two steps is relatively complete moving target, as shown in Fig.5 below:
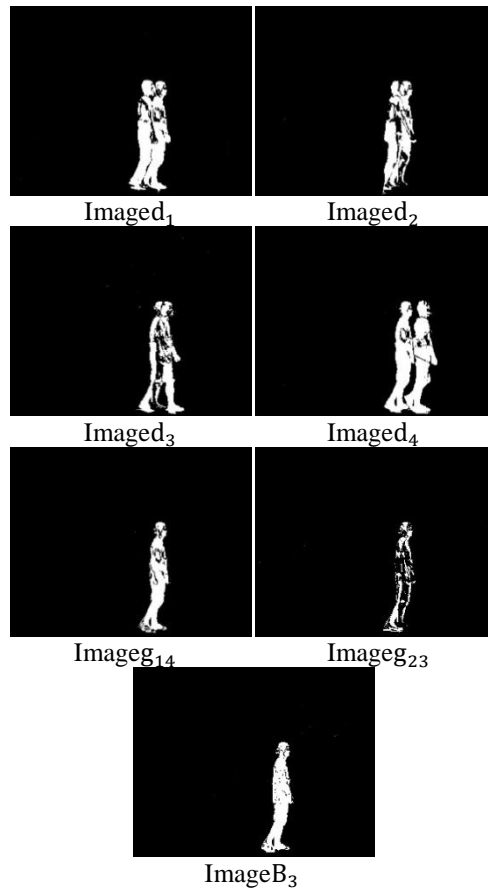
Fig.4. Five pictures of the original image.

Imaged$_1$ Imaged$_2$

Imaged$_3$ Imaged$_4$

Imageg$_{14}$ Imageg$_{23}$

ImageB$_3$

Fig.5. Five - frame difference method renderings.

3.3 Morphological processing

        The five-frame difference method separates moving target from background, yet the moving target obtained still has isolated points and disturbance of holes, so we need to subject the image to morphological processing to eliminate disturbance. Morphological processing mainly comprises expanding, corroding, opening and closing, etc. According to the obtained difference image, the paper first uses expansion algorithm to form a complete connected area in the target, then uses corrosion algorithm to restore the expanded target to original size. The final effect drawing of moving target obtained is as shown in Fig.6.

Fig.6.Morphologically processed image.

After obtaining a complete moving target, the method used in Bayes classifier design in preceding context is used to extract HOG characteristics of this moving target.

IV. Pedestrians identification based on Bayes classifier

The experiment was done on a computer with Windows 7 OS, and Visual Studio 2010 was used to program code. The experiment used 50 groups of continuous five-frame test images, which were classified by Bayes classifier. The recognized pedestrians were framed out using red frame in original drawing by image projection algorithm. The part of test result is as shown in Fig.7 below:
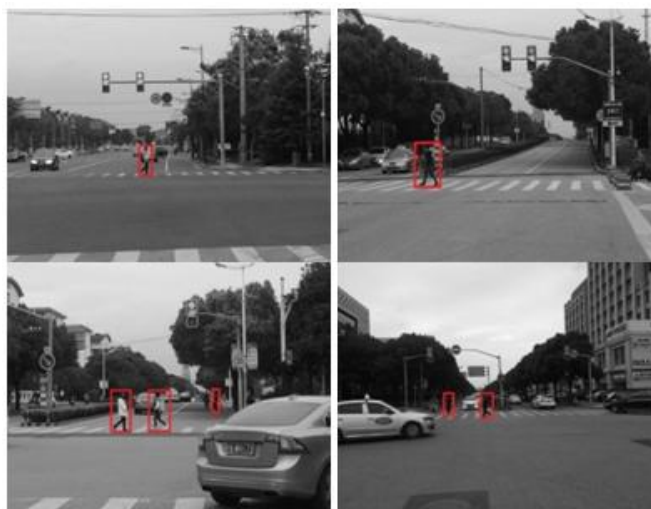

Fig.7. Bayesian classifier identification results.

In the test pictures, there are altogether 127 pedestrians crossing the street. The identification result is 123 recognized, 4 undetected, and 1 place wrongly detected. The reason for 4 being undetected is due to too large wind, causing the large swinging of branches and leaves of trees on the roadside, while the swinging branches and leaves overlap with pedestrians, leading to poor effect of five-frame difference, hence 4 undetected. The reason for wrong detection is because the non-pedestrians sample database is dominated by vehicle pictures, with fewer categories of non-pedestrian object, leading to wrong detection, which is the content of future research on this algorithm.

V. Conclusion

This paper utilizes pedestrians' kinetic characteristics to use five-frame difference method to separate pedestrians from background. This method can significantly lower the influence of weather and pedestrians' clothes on the paper's algorithm, and can detect pedestrians in real time in complicated scenes such as changing illumination condition. On the other hand, the Bayes classifier adopted by this paper classifies according to maximum posterior probability decision rule to realize an accuracy rating of up to 97% in identification of pedestrians crossing the street. This classification is provided with adequate robustness to ignore the defects in naive Bayes probability model. Of course, in the algorithm as described in this paper, the non-pedestrian sample database of Bayes classifier needs richer categories, and instantaneity of algorithm should be further improved.

References

[1]   GuoQiang. Based on the characteristics of pedestrian detection technology research. Jilin University, 2011.
[2]   Li Chunming. Pedestrian detection under camera movement. Computer Engineering and Design, 2013 (09), 3333-3337.
[3]   Guo Lie. Pedestrian detection method based on the combination of typical features of human body. Automotive Engineering, 2014 (04), 496-502.
[4]   Jiang Ailian, Yang Xingtong. Pedestrian detection based on AdaBoost-SVM cascade classifier. Computer Engineering and Design, 2013 (07), 2547-2550.
[5]   GavrilaD.Pedestriandetectionfromamovingvehicle.ProceedingsofEuropeanConferenceComputerVision.Dublin,Ireland:LectureNotes in Computer Science, 2000, 37-49.
[6]   Oren M, Papageorgiou C, Sinha P, et al. Pedestrian detection using wavelet templates. Proceedings of IEEE Conference on Computer Vision and Pattern Identification, San Juan, Puerto Rico. USA: IEEE, 1997, 193-199.
[7]   Cheng Hong, Zheng Nanning, Qin Junjie. Pedestrian detection using sparse Gabor filter and support vector machine. Proceedings of IEEE Intelligent Vehicles Symposium, Vienna, Austria. USA: IEEE, 2005, 583.
[8]   Dai Congxia, ZhengYunfei, Li Xin. Layered representation for pedestrian detection and tracking in infrared imagery. Proceedings of IEEE Conference on Computer Vision and Pattern Identification, San Diego, CA, USA. USA: IEEE, 2005, 3-13.

[9]   Gao W, Ai H Z, Lao S. Adaptive contour features in oriented granular space for human detection and segmentation. Proceedings of IEEE Conference Computer Vision and Pattern Identification. Miami, Florida, USA. USA: IEEE Computer Society, 2009, 1786-1793.

[10]   Zhao Liang, Thorpe C. Stereo and neural network-based pedestrian detection. IEEE Transactions on Intelligent Transportation Systems (S1524-9050), 1(3), 2000, 148-154.

[11]   Paul viola, Michael Jones. Robust Real-time Object Detection. International Journal of Computer Vision, (S0920-5691), 57(2),2001, 87-112.

[12]   Shashua A, Gdalyahu Y, Hayun G. Pedestrian detection for driving assistance systems: single-frame classification and system level performance. Proceedings of IEEE Intelligent Vehicles Symposium, Parma, Italy. USA: IEEE, 2004, 1-6.

[13]   Li Peng. A face detection method based on neural network and Bayesian decision.Application Research of Computers. 2007 (08), 198-200.

[14]   Dong Liyan. Image classification based on Bayesian classifier. Journal of Jilin University. 2007 (03), 249-253.