

TWO-STAGE CLASSIFICATION APPROACH FOR HUMAN DETECTION IN CAMERA VIDEO IN BULK PORTS

Chao MI^a

Zhiwei ZHANG^b

Xin HE^b

Youfang HUANG^a

Weijian MI^a

^a)Container Supply Chain Tech. Engineering Research Center, Shanghai Maritime University, No.1550
Haigang Ave, Shanghai 201306, China

^b)Logistics Engineering College, Shanghai Maritime University, No.1550 Haigang Ave, Shanghai
201306, China

ABSTRACT

With the development of automation in ports, the video surveillance systems with automated human detection begun to be applied in open-air handling operation areas for safety and security. The accuracy of traditional human detection based on the video camera is not high enough to meet the requirements of operation surveillance. One of the key reasons is that Histograms of Oriented Gradients (HOG) features of the human body will show great different between front & back standing (F&B) and side standing (Side) human body. Therefore, the final training for classifier will only gain a few useful specific features which have contribution to classification and are insufficient to support effective classification, while using the HOG features directly extracted by the samples from different human postures. This paper proposes a two-stage classification method to improve the accuracy of human detection. In the first stage, during preprocessing classification, images is mainly divided into possible F&B human body and not F&B human body, and then they were put into the second-stage classification among side human and non-human recognition. The experimental results in Tianjin port show that the two-stage classifier can improve the classification accuracy of human detection obviously.

Keywords: Human Detection, Histograms of Oriented Gradients, Support Vector Machine, Classification

INTRODUCTION

According to the safety and security requirements of open-air unmanned handling operation areas of bulk ports, the pedestrians without permission are forbidden entering those areas because of the high risk in the port handling process. Recently, the machine vision technology begun to be applied in some areas with high security and efficiency requirements such as the automated handling fields of bulk terminals [1][2]. Comparing to the large quantity of researches on port handling technology, the research of machine vision technology in automatic port is only just the beginning [3][4]. Although we have done some research on human detection algorithm, the detection accuracy is not enough to meet the practical necessity, which are in need of further research.

At the earliest, template matching was the popular approach to recognize the human in the images. Template matching is to match the feature of detection image and the template of human contour to conduct human detection [5]. Songmin Jia et al. proposed varying scale template matching (VSTM) to detect human from disparity image based on head-shoulder method [6]. But the diversity of human postures and clothing caused problems in template matching and led to a large amount of calculation.

After that, because of the low accuracy and high calculation consumption, the template matching has been eliminated. The research on human detection is now categorized under 2 steps – feature extracting and classifier designing. Navneet Dalal and Bill Triggs firstly proposed Histograms of Oriented

Gradients (HOG) approach in 2005[7]. Since geometric and optical deformation affects a little on the features and some subtle body actions can be ignored, HOG is suitable for human detection. It relies on contrast between human contour and background. Thus, Hong Tian et al. proposed a multiplex cascade classifier which can filter out most of irrelevant background [8]. Hai-Miao Hu et al. applied a joint global-local information algorithm to suppress the background interference and enrich the description of pedestrian [9]. To a great extent, these researches reduced the influence of background and improved the detection accuracy.

However, 3780 HOG features needs to be extracted in each sub-image with the consideration of the impact of each possible factor and then Support Vector Machine (SVM) classifier is used to make a binary decision, which results in low efficiency. A year after the proposition of the above method, Qiang Zhu et al. introduced integral image into the calculation of HOG features to solve huge computation of HOG feature extraction [10]. Variable-size blocks were used to replace the multi-scale detection in traditional method and the cascade classifier was used to select a small subset of HOG features. Compared with Navneet Dalal's method, its accuracy is lower than Navneet Dalal's algorithm in most cases because integral histograms of orientated gradients abandons the original Gaussian weighted tri-linear interpolation and cannot avoid the influence of overlapping. Moreover, the considered feature set is not abundant enough. In our previous research, we presented an improved HOG algorithm with a combined classifier to reduce the calculation consumption [1]. Human detection efficiency is improved, but the accuracy is still equal to the traditional ways, which cannot meet the precision requirements in practical application.

Currently, most scholars study on the human detection mainly based on HOG proposed by Navneet Dalal [11]. Marco Pedersoli et al. proposed presents a multiresolution cascade of HOG that can highly reduce the computational cost without affecting accuracy [12]. HOG can obtain abundant feature sets by extracting shape and motion information of human body, which can entirely describe the contour features with a strong robustness [13]. HOG feature has been widely used in human detection, as a feature based on edge and shape of object, which is a kind of fuzzy edge direction oriented feature. Therefore, HOG feature can ignore some subtle changes of human body to get a better robustness than other methods such as template matching. But the large difference between human standing postures, such as front & back standing posture or side standing posture, will lead to the final HOG features for classifiers are less specific.

At present, the classifiers widely used in human detection are the Adaptive Boosting (Adaboost) cascade classifier and SVM classifier. The traditional single classifier is trained by HOG features combined of different human body postures can only screen a few effective features. For most classifiers, they are usually used to find those specific features among the whole features from samples. SVM is a machine learning method based on Vapnik-Chervonenk is dimension and Structural Risk Minimization of statistical learning theory [14] [15].

SVM can classify the samples by the specific features with classification ability. Less sample training can obtain a SVM classifier with preferable performance. In terms of Adaboost classifier, this classifier selects some important specific feature vectors as weak classifiers to train a strong classifier [16] [17]. Then it establishes a statistical method for human body so that it can detect human body from candidate targets. Some researchers tried to combine the single classifiers to classify the human body based on HOG. In our previously researches and Lie Guo's researches, Adaboost and SVM were combined to classify the human body, the performances were better than the traditional single classifier, but these combined classifiers did not consider the human postures [1][18]. Therefore, there will be room for improvement.

The major work of this paper is that a two-stage SVM classifier is chosen to replace traditional SVM classifier to improve the detection accuracy. As the research proposed by Huimin Qian, humans have a variety of postures in images [19]. Front or back standing (F&B) humans with relatively large effective areas have rich contour features and are easy to be classified, while side standing (Side) humans have less features and are more complex, comparatively. If the classifier is trained like the traditional method using mixed samples, it may cause classification conflict which will lead to unsatisfactory effects. Therefore, at the beginning of classification, detection images should be divided into three categories, front and back standing posture, side standing posture and non-human. In the first stage, the samples are mainly divided into possible F&B human body and not F&B human body, and then they were put into the second-stage classification among Side human and non-human recognition. Finally the detection results are merged together.

TWO-STAGE HUMAN CLASSIFICATION ALGORITHM

HOG FEATURE EXTRACTION ALGORITHM

In fact, it has been found that HOG descriptor has many advantages compared with other feature description methods. 1) HOG represents the structural feature of edge, so that it can describe local shape information and depict contour feature of human body to a great extent [20]; 2) Quantization of position and direction in the space can suppressed the influence of translation and rotation to some extent; 3) Since HOG is operated on the local cell unit of images, it can maintain good invariance on geometric and photometric deformation[21]. The main idea of HOG is to calculate the statistics data of local gradient direction in the images proposed by Navneet Dalal and Bill Triggs [7]. The whole image is divided into small areas (called cells), and histograms are created for each cell according to gradient direction. The combination of these histograms shows the detected object's feature. In order to improve accuracy, local histogram can combine small cells into a large area (called block), and then normalize in block units to ensure better illumination or shadow invariance [22].

The HOG feature is extracted by the following steps:

1. Standardization of color space is needed for images because the environment of open-air handling operation is complex and the illumination is strong. Image noise can be suppressed using square root compression gamma correction operations. The standardized pixel values of three channels, $R_{x,y}$, $G_{x,y}$ and $B_{x,y}$, can be obtained by Eq.1;

$$\begin{cases} R_{x,y} = \sqrt{r_{x,y}} \\ G_{x,y} = \sqrt{g_{x,y}} \\ B_{x,y} = \sqrt{b_{x,y}} \end{cases} \quad (1)$$

Where, r_{xy} , g_{xy} and b_{xy} are the pixel original values of red channel, green channel and blue channel, respectively, where (x, y) is the pixel coordinates.

2. The intensity and direction of gradient for each point need to be calculated after standardized gradation; One-dimensional discrete template $[-1, 0, 1]$ and $[-1, 0, 1]^T$ are used both in the horizontal and vertical directions, the gradient magnitude $G_{x,y}$ and gradient direction $\theta_{x,y}$ of the pixel point (x, y) are calculated according to the following Eq.2 and Eq.3:

$$G_{x,y} = \sqrt{G_x^2(x, y) + G_y^2(x, y)} \quad (2)$$

$$\theta_{x,y} = \tan^{-1} (G_y(x, y) / G_x(x, y)) \quad (3)$$

Where, $G_x(x, y)$ and $G_y(x, y)$ are the gradient magnitude of pixel point (x, y) in the x axis (horizontal direction) and y axis (vertical direction), which can be calculated by Eq.4:

$$\begin{cases} G_x(x, y) = f_{x+1,y} - f_{x-1,y} \\ G_y(x, y) = f_{x,y+1} - f_{x,y-1} \end{cases} \quad (4)$$

Where, $f_{x,y}$ is the value of pixel point (x, y) .

3. According to Fig.1, the whole image is scanned by a block, which is divided into 4 cells. $[0^\circ, 180^\circ]$ is divided into 9 bins as shown in Fig.2. And then the gradient direction of each point in the unit is discretized into the 9 bins. The gradient direction of each pixel is assigned to the neighboring bins by a certain ratio using a tri-linear interpolation method [23]. 36 dimensional HOG feature vector can be extracted from each block.

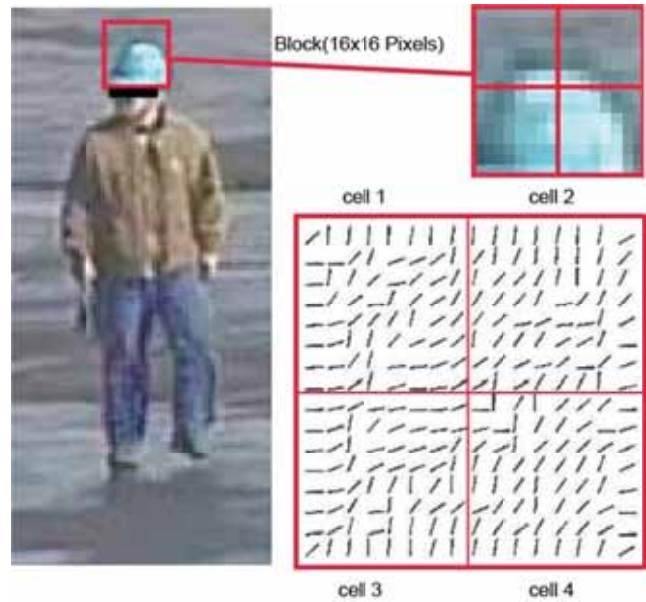


Fig. 1 HOG descriptors schematic.

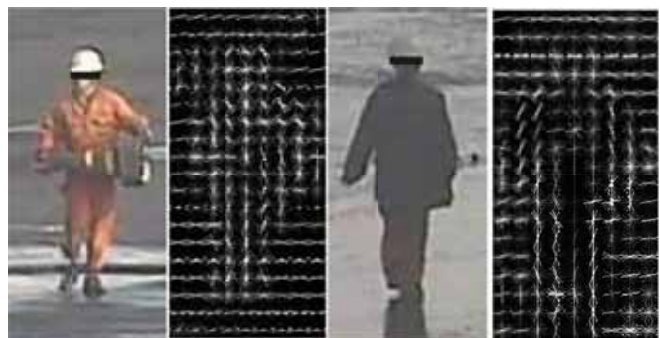


Fig. 2 Bin's schematic diagram.

4. L2-Norm is used in each block for normalization to reduce the effect of illumination and shape variation of edges.
5. Finally, the entire 105 blocks are connected in series to form a 3780 dimensional feature vectors for each detection window.

In the above steps, HOG feature extraction of each block is the foundation to extract the HOG feature of the whole image. Here, tri-linear interpolation is usually adopted to process the local block image. For analyzing its principle, it can be easy to know that HOG feature is a feature based on edge direction and shape feature of human body.

(a)



(b)



Fig. 3 HOG features of human

However, HOG feature based on edge direction information limits the performance of traditional classifier for human detection. As shown in Fig.3, in the case of different posture, effective features from human body also show different specific characteristics. As shown in Fig.3 (a), the head and shoulders of F&B human body basically keeps invariant geometric features, including relative position, angle and size. When standing or walking, F&B human bodies are basically not cross their legs. HOG features extracted by F&B samples are nearly symmetrical with more obvious specific characteristics. For side human samples as shown in Fig.3 (b), the shoulders overlap each other and the legs also overlap or cross, which means that the effective features from side humans are more confusion and lack symmetry. Obviously, as described in the above HOG extraction process, HOG features can obtain the specific characteristics of human body based on contour features and shape features of human. Thus, under this situation, it is necessary to add one-stage classifier for distinguishing different postures of the human.

TWO-STAGE CLASSIFICATION APPROACH

High-dimensional feature vectors in detection window have been calculated in previous section. The next step is to make judgment according to the particularity of human contour.

Over the past ten years, SVM has been widely used for human detection because of its generality and good performance [24]. On the other hand, the selection of classification feature has a great impact on SVM classification effect. However, as mentioned in previous section, HOG feature will represent

large different specific characteristics when extracted by different human body postures. Hence, a novel classification needs to be applied to improve SVM classification accuracy on the basis of assuring classification efficiency.

Within the scope of the detection, human body will not always show F&B postures in cameras. Due to the effect of many factors, such as dust, illumination, color change and body postures, it is very difficult to design classifier to obtain an ideal detection result by using single classifier. One of the reasons is the long training time and the poor detection performance of SVM algorithm for various human postures. The major purpose of this paper is to improve the traditional classification method by replacing single classifier with a two-stage classifier based on classifiers fusion using the idea of SVM classifier. The detailed concept is shown as follows:



Fig. 5 The sample sets of F&B and side postures.

Two-stage classifier firstly processes human contour samples and divides targets into three types: possible F&B and side and non-human as shown in Fig.4. Firstly, the pre-classified images of each posture should concurrently be made a binary decision: F&B human or not F&B human by F&B classifiers. Then, the second classifier can make the rest image divided into two types: side human or non-human. Finally, the detection results of F&B and side human body are merged to generate candidate areas with people inside. The sample sets of F&B and side postures are shown in Fig.5.

SVM is a supervised learning algorithm, which you need to import the training sample sets with a label. Here, the calculated HOG features set will be the initial 3780 dimensional feature vector space. The detailed training process of classifiers is shown as follows:

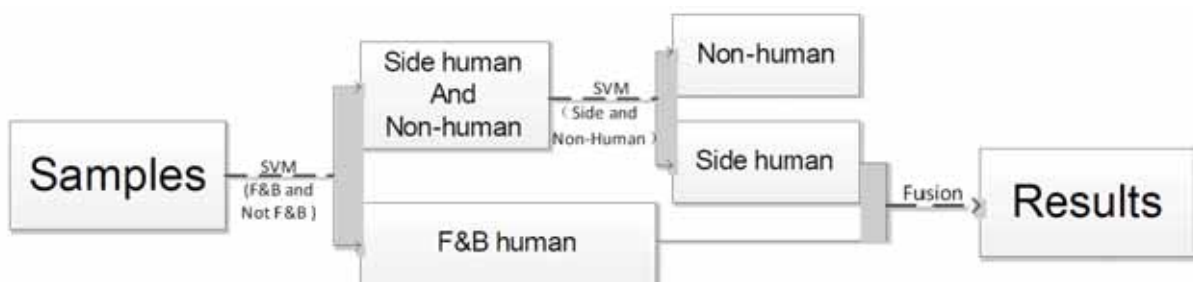


Fig. 4 Two-stage classification approach flowchart.

1. Preprocessing of samples

First, the samples are divided into three types: F&B, side and non-human. Using radial basis function (RBF) as Eq.5, the 3780 dimensional feature space of samples will be mapped to a higher-dimensional feature space, which is linearly separable.

$$K(x, x_c) = e^{-\frac{\|x-x_c\|^2}{(2\sigma^2)}} \quad (5)$$

Where, \mathbf{x} is a feature vector in the feature space and \mathbf{x}_c is the position of the center of kernel function. The width parameter σ of function can control the radial scope of impact from kernel function. Thus, when \mathbf{x} far away from \mathbf{x}_c , the function value will be very small.

In order to find the two hyper planes in the labeled sample feature space, objective function and constraint function of the two-stage classifier is shown in Eq.6.

$$\begin{cases} \min \frac{1}{2} \|\mathbf{w}\|^2, \min \frac{1}{2} \|\mathbf{w}'\|^2 \\ s. t : y_i \cdot (\mathbf{w}^T \mathbf{x}_i + b) - 1 \geq 0 \quad (i = 1, 2, \dots, l) \\ (\mathbf{w}'^T \mathbf{x}'_i + b') - 1 \geq 0 \quad (i = 1, 2, \dots, l') \end{cases} \quad (6)$$

\mathbf{x}_i (\mathbf{x}'_i) and \mathbf{w} (\mathbf{w}') represent the feature vector and a combination of feature vectors when training F&B (side) classifier. l (l') means the number of samples. By given samples and labels, \mathbf{w} (\mathbf{w}') can be calculated by Eq.7:

$$\mathbf{w} = \alpha_1 \mathbf{x}_1 y_1 + \alpha_2 \mathbf{x}_2 y_2 + \dots + \alpha_n \mathbf{x}_n y_n \quad (7)$$

In Eq.7, \mathbf{x} represents the HOG feature vector and y means the label value of \mathbf{x} . α and n are Lagrange multipliers and the size of sample sets. To avoid that the results might be controlled by several points of samples space, slack variables ζ and punishment factor c need to be imported to increase fault tolerance of the two-stage classifier.

$$\begin{cases} \min \frac{1}{2} \|\mathbf{w}\|^2 + c \sum_{i=1}^l \zeta_i, \min \frac{1}{2} \|\mathbf{w}'\|^2 + c' \sum_{i=1}^{l'} \zeta'_i \\ s. t : y_i \cdot (\mathbf{w}^T \mathbf{x}_i + b) - 1 \geq 0 \quad (i = 1, 2, \dots, l) \\ (\mathbf{w}'^T \mathbf{x}'_i + b') - 1 \geq 0 \quad (i = 1, 2, \dots, l') \end{cases} \quad (10)$$

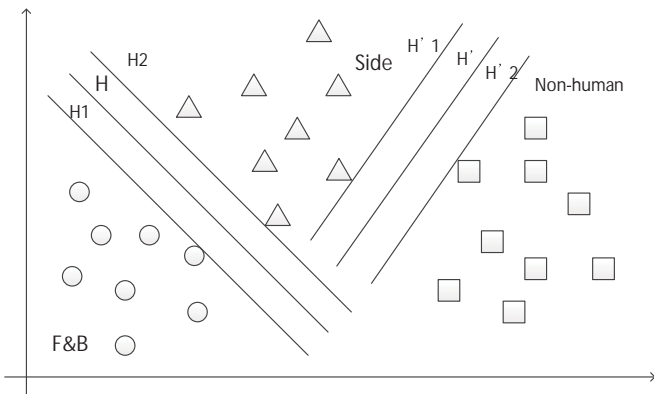


Fig. 6 Two-stage classifier.

Therefore, as shown in Fig.6, the objective functions of the

two-stage classifier will be used to find the two hyper planes H and H' for training F&B classifier and side classifier.

2. Training F&B classifier and side classifier

For F&B classifier, the sample feature \mathbf{x}_i will be labeled as $y_i = +1$, which belongs to F&B sample features, otherwise $y_i = -1$. For side classifier, if the sample feature \mathbf{x}'_i belongs to the side sample features, it will be labeled as $y'_i = +1$, otherwise $y'_i = -1$.

Then, SVM parameters needs to be determined, such as, the type of SVM classifier, kernel function type, punishment factor and the other main parameters.

After that, using these HOG feature vectors with labels respectively, SVM classifier is trained to get the first-stage and second-stage initial classifiers. As shown in Fig.6, through F&B as positive samples and the others as negative samples, the support vectors and hyper plane H will be got by training the first-stage classifier, which can separate two features spaces. In the same way, the hyper plane H' will be obtained by training side as positive samples and non-human as negative samples..

3. Training final two-stage classifier

Trying to detect human in the negative samples through initial classifiers, there will be many erroneous detection results, which can be termed "hard example". In order to improving detection accuracy, HOG feature from these hard examples will be combined with the original negative samples as new negative samples.

Based on the new samples space, new two-stage classifier with a higher accuracy can be trained. Repeating the step 3), the final two-stage classifier with appropriate detection accuracy will be trained.

EXPERIMENTAL RESULTS AND ANALYSIS

The system has conducted several field tests in operation field in terminals of Tianjin Port. Fig.7 shows the video captures of field tests.

Representative images are selected to build 3 image sets of different types and complexity to perform evaluation for the proposed human detection system. The main characteristics of the 3 test sets are shown in table 3, which are F&B test set and side test set and various mixed postures test set including F&B, side etc.

Table 1 The testing sample sets.

Test set	Sample size	Number of person	Posture
1	453	453	F&B
2	365	365	Side
3	398	866	Mixed



Fig. 7 The experimental results of grain terminal.

The experimental results are shown in the table 2 below, among which the two-stage classifier is the novel method introduced in Section 2.2. “FDR” means “false detection rate”.

Table 2 The experimental results.

Test set	Classifier	True	Undetected	False	FDR	Accuracy
F&B 1	F&B (first-stage SVM)	448	0	5	1.10%	98.90%
	Side (second-stage SVM)	214	223	16	3.53%	47.24%
	Two-stage SVM	447	0	6	1.32%	98.68%
	SVM	426	9	18	3.97%	94.03%
Side 2	F&B (first-stage SVM)	186	164	15	4.11%	50.96%
	Side (second-stage SVM)	334	14	17	4.66%	91.51%
	Two-stage SVM	341	12	12	3.29%	93.42%
	SVM	335	13	19	5.21%	91.78%
Mixed 3	F&B (first-stage SVM)	653	174	39	4.50%	75.40%
	Side (second-stage SVM)	346	477	43	4.97%	39.95%
	Two-stage SVM	842	6	18	2.07%	97.23%
	SVM	789	28	49	5.70%	91.11%

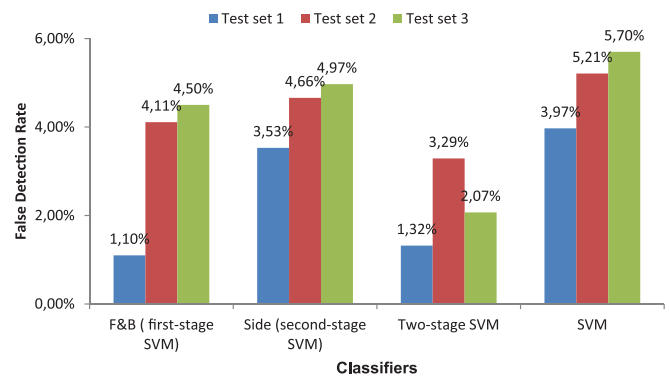


Fig. 8 Comparison of five classifiers about false detection rate.

The false detection rate of the two-stage classifier to the 3 test sets are 1.32%, 3.29 % and 2.07%, has been a lot lower than single SVM classifier. It is because that the classification method in this paper is the results of the combination of single classifiers.

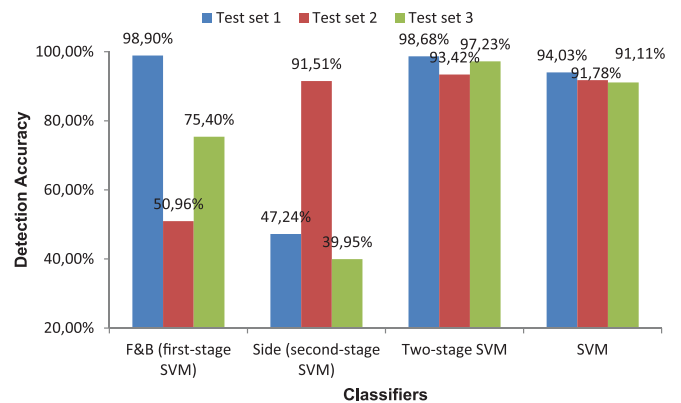


Fig. 9 Comparison of five classifiers about detection accuracy.

The F&B (first-stage) classifier has very good detection results for F&B images and its detection accuracy reached 98.90% while the side classifier it correctly detect the test set 2 at a rate of 91.51%. From the data above, it is found that the effect of F&B detection is better than side detection mainly because the HOG features of F&B posture in the images are

more specific to detect and easier to classify so that a better F&B classifier can be trained easier. As for side posture, there are large variation and great difference between these training sets, thus it is difficult to obtain a high-performance universal classifier for side-posture human.

According to the traditional classification method, we can get a SVM classifier trained with the mixed human postures. As shown in Fig.8 and Fig.9, the two-stage classifier based on SVM algorithm has a greater improvement of adaptive capacity for human postures. Comparing to our previous studies, the sample sets are all collected in the same background, but this paper improved the detection accuracy of nearly 3% due to the preprocessing of building different sample sets for different human postures.

This paper proposes a novel two-stage classifier which is made up of two single SVM classifiers for two types human postures. It can improve the detection accuracy of F&B posture significantly and make the detection accuracy of side postures flat with traditional approach. Its accuracy for the three test sets reaches 98.68%, 93.42% and 97.23%, which has a great progress compared to traditional SVM classifier.

CONCLUSION

According to the existing problems of traditional HOG based classification method, a novel two-stage classification approach is proposed. Two-stage SVM is selected to replace traditional classifier to improve classification accuracy. At the first stage of classification, the samples are mainly divided into possible F&B human body and not F&B human body, and then they were put into the second-stage classification among side human and non-human to detect human. Finally the detection result is integrated with the use of disjoint sets.

The experimental tests of operation field in bulk terminal of Tianjin Port show that single traditional SVM algorithm has a large limitation for F&B, side and complicated posture human body test sets. And it lacks adaptability to human postures. The first-stage classification has a greater detection accuracy than the second-stage classification, which means F&B posture human has more specific HOG features than side human. The two-stage classification method proposed in this paper has a detection accuracy of more than 97%, which is higher than using traditional SVM. In conclusion, the improved classification approach purposed in this paper is more feasible and advanced.

ACKNOWLEDGEMENT

This research was supported by “Scientific Research Innovation Project” of Shanghai Municipal Education Commission (No. 14ZZ140), the “Ph.D. Innovation Program” of Shanghai Maritime University (No. 2014ycx040), and the “Yang Fan Plan for Shanghai Youth Science and Technology Talent” of the Science and Technology Commission of Shanghai Municipality (No.15YF1404900).

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests regarding the publication of this paper.

REFERENCES

1. Mi, Chao; He, Xin; Liu, Haiwei; Huang, Youfang; Mi, Weijian, “Research on a Fast Human-Detection Algorithm for Unmanned Surveillance Area in Bulk Ports”, *Mathematical Problems in Engineering*, 2014, DOI: 10.1155/2014/386764
2. Mi C, Shen Y, Mi W, et al. “Ship Identification Algorithm Based on 3D Point Cloud for Automated Ship Loaders”. *Journal of Coastal Research*, 2015, 73(sp1): 28-34, DOI: 10.2112/SI73-006.1
3. Bian Zhicheng, Yang Yongsheng, Mi Weijian, Mi Chao. “Dispatching electric AGVs in Automated Container Terminals with long travel distance”. *Journal of Coastal Research*, 2015, 73(sp1): 75-81, DOI: 10.2112/SI73-014.1
4. Xiaoming Yang, Ning Zhao, Zhicheng Bian, Jiaqi Chai, and Chao Mi. “An intelligent storage determining method for inbound containers in container terminals”. *Journal of Coastal Research*, 2015, 73(sp1): 197-204, DOI: 10.2112/SI73-035.1
5. Yung-Chi Lo; Po-Yen Lee ;Shyi-Chyi Cheng. “Space-time template matching for human action detection using volume-based Generalized Hough transform”. *Image Processing (ICIP), 2011 18th IEEE International Conference on*. pp.2097 – 2100(2011), DOI: 10.1109/ICIP.2011.6116021
6. SongminJia; ShuangWang ;Lijia Wang et al. “Robust human detecting and tracking using varying scale template matching”. *Information and Automation (ICIA), 2012 International Conference on*. pp. 25 – 30(2012), DOI: 10.1109/ICInfA.2012.6246776
7. Navneet Dalal and Bill Triggs, 2005. “Histograms of oriented gradients for human detection”.. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, pp.886- 893(2005), DOI: 10.1109/CVPR.2005.177
8. Hong Tian, Zhu Duan, Ajith Abraham, Hongbo Liu.” A novel multiplex cascade classifier for pedestrian detection”. *Pattern Recognition Letters*, Volume 34, Issue 14, 15 October 2013, Pages 1687-1693, DOI: 10.1016/j.patrec.2013.04.015
9. Hai-Miao Hu, Xiaowei Zhang, Wan Zhang, Bo Li.” Joint global-local information pedestrian detection algorithm for outdoor video surveillance”. *Journal of Visual Communication and Image Representation*, Volume 26, January 2015, Pages 168-181, DOI: 10.1016/j.jvcir.2014.11.009

10. Qiang Zhu; Yeh, M.-C. ; Kwang-Ting Cheng, et al. "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients". *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol.2*, pp.1491 – 1498(2006), DOI: 10.1109/CVPR.2006.119
11. Cristina Conde, Daniela Moctezuma, et.al."HoGG: Gabor and HoG-based human detection for surveillance in non-controlled environments" *Neurocomputing*, Volume 100, 16 January 2013, Pages 19-30, DOI: 10.1016/j.neucom.2011.12.037
12. Marco Pedersoli, Jordi González, Andrew D. Bagdanov, Xavier Roca. "Efficient discriminative multiresolution cascade for real-time human detection applications". *Pattern Recognition Letters*, Volume 32, Issue 13, 1 October 2011, Pages 1581-1587, DOI: 10.1016/j.patrec.2011.06.019
13. Ninomiya, H; Ohki, H.; Gyohten, K. et al. "An evaluation on robustness and brittleness of HOG features of human detection". *Frontiers of Computer Vision (FCV), 2011 17th Korea-Japan Joint Workshop on*. pp. 1 – 5(2011), DOI: 10.1109/FCV.2011.5739746
14. Ying Tan; Jun Wang. "A support vector machine with a hybrid kernel and minimal Vapnik-Chervonenkis dimension". *Knowledge and Data Engineering, IEEE Transactions on. Vol.16*, pp 385 – 395(2004), DOI: 10.1109/TKDE.2004.1269664
15. QiuXintao; Fu Dongmei ; Yang Tao. "A novel approach to optimize the objective function based on VC dimension and structural risk minimization". *Control Conference (CCC), 2011 30th Chinese*. pp. 3226 – 3230(2011)
16. Ando, H. Fujiyoshi, H. "Human-Area Segmentation by Selecting Similar Silhouette Images Based on Weak-Classifer Response". *Pattern Recognition (ICPR), 2010 20th International Conference on*. pp. 3444 – 3447(2010). DOI: 10.1109/ICPR.2010.841
17. Jiayuan Yu. "The application of BP-Adaboost strong classifier to acquire knowledge of student creativity". *Computer Science and Service System (CSSS), 2011 International Conference on*. pp.669 – 2672(2011). DOI: 10.1109/CSSS.2011.5974999
18. Lie Guo, Ping-Shu Ge, Ming-Heng Zhang, Lin-Hui Li, Yi-Bing Zhao." Pedestrian detection for intelligent transportation systems combining AdaBoost algorithm and support vector machine" *Expert Systems with Applications*, Volume 39, Issue 4, March 2012, Pages 4274-4286, DOI: 10.1016/j.eswa.2011.09.106
19. Huimin Qian, Yaobin Mao, Wenbo Xiang, Zhiquan Wang." Recognition of human activities using SVM multi-class classifier". *Pattern Recognition Letters*, Volume 31, Issue 2, 15 January 2010, Pages 100-111, DOI: 10.1016/j.patrec.2009.09.019
20. Kanitkar, A. Bharti, B. ;Hivarkar, U.N. "Vision based preceding vehicle detection using self shadows and structural edge features". *Image Information Processing (ICIIP), 2011 International Conference on*. pp.1 – 6(2011), DOI: 10.1109/ICIIP.2011.6108922
21. Man Zhang; Zhenan Sun ;Tieniu Tan. "Deformed iris recognition using bandpass geometric features and lowpass ordinal features". *Biometrics (ICB), 2013 International Conference on*. pp.1 – 6(2013), DOI: 10.1109/ICIIP.2011.6108922
22. Yu Cheng; ZhigangJin; CunmingHao. "Illumination normalization based on 2D Gaussian illumination model". *Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on. Vol.3*, pp.V3-451 - V3-455(2010). DOI: 10.1109/ICACTE.2010.5579870
23. Ying Bai; Dali Wang. "On the Comparison of Trilinear, Cubic Spline, and Fuzzy Interpolation Methods in the High-Accuracy Measurements". *Fuzzy Systems, IEEE Transactions on. Vol.8*, pp. 1016 – 1022(2010), DOI: 10.1109/TFUZZ.2010.2064170
24. Corinna Cortes; Vladimir Vapnik. "Support-vector networks". *Machine Learning. Vol.20, Issue 3*, pp 273-297(1995), DOI: 10.1007/BF00994018

CONTACT WITH THE AUTHORS

Chao MI
 Youfang HUANG
 Weijian MI
 e-mail: kingdows@gmail.com

Container Supply Chain Tech. Engineering Research
 Center
 Shanghai Maritime University, No.1550 Haigang Ave,
 Shanghai 201306
CHINA

Zhiwei ZHANGb
 Xin HEB

Logistics Engineering College
 Shanghai Maritime University, No.1550 Haigang Ave,
 Shanghai 201306
CHINA