International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.2 (2016), pp.35-46
http://dx.doi.org/10.14257/ijsip.2016.9.2.04

Research on a Target Object Locating Method

Zhao Lei[1,a] and Ren Hong-e[1*]

[1]Northeast Forestry University, Harbin, China
[a]154372732@qq.com, [*]renhonge163@163.com

Abstract

Target detection has become a common problem to be solved in computer field. Face detection is an important research direction in target detection and a key link in face information processing as well as the premise and foundation of automatic face recognition. Face detection property influences the result of face information recognition directly. This paper improves current face detection algorithms according to traditional AdaBoost algorithm and Sa-AdaBoostSVM algorithm. The experimental result shows that the algorithm in this paper has good robustness for multi-pose and multi-expression face.

Key words: *AdaBoost algorithm, face detection, Haar-like feature, multi-pose*

1. Introduction

Target detection has become one of the most common problems to be solved in computer field. In most cases, we only pay attention to the information of key parts in target object, such as text or eye and face information in grabbed image. Therefore, improving target detection algorithms and the efficiency of target detection constantly is always a striving direction of researchers. Currently, face detection in the research field of target detection has become a research hotspot. This research is the research foundation of biological feature recognition, video surveillance and control system, human-computer interaction, safety monitoring and controlling, information retrieval and video conference etc. and plays a vital role in the development of fields above. Face detection is a key step of face information processing and the foundational work of automatic face recognition[1, 2, 4, 5].

Current face detection algorithms can be divided into two categories: detection methods based on face feature information and detection methods based on face image. [2, 6, 7] The former methods directly use face information such as skin color and geometric structure of face. Such methods mostly use classical theory of pattern recognition. The latter methods do not use face information directly; instead, they regard face detection problem as a common pattern recognition problem. Images to be detected are input directly as a system and feature extraction and analysis are not required in the middle. Training algorithm is used directly to divide learning samples into face category and non-face category. It can be judged whether the detection area is face only through the comparison of both categories and possible human face area during face detection.

In allusion to target detection problem, researchers have made attempts with multiple learning algorithms. Papageorgiou used support vector machine based on redundant wavelet feature to establish detection algorithm. Schneiderman put forward a Bayes classifier based on multi-scale wavelet transform for detection. Rowly used neural network method in face detection system. Viola and Jones put forward AdaBoost classifier based on cascade structure composed of multi-stage classifiers. The classifier on

[*] Corresponding author

ISSN: 2005-4254 IJSIP

each stage uses AdaBoost algorithm for training, thus realizing the comprehensive improvement of face detection speed and property. Due to property and speed advantages, AdaBoost algorithm has become a mainstream face detection algorithm in recent years. However, it also has certain problems, such as over-learning problem, which influence its generalization ability. The algorithm in this paper is mainly used to process driver image. Therefore, this paper puts forward an improved driver face detection method based on AdaBoost algorithm and further improves its training convergence speed and driver face feature detection property under the premise of complicated background of driving vehicles [3, 7, 8].

2. AdaBoost Face Detection Algorithm

AdaBoost face detection algorithm contains the main work in three aspects: (1) give Haar-like feature which can be obtained rapidly through picture integral image; (2) use AdaBoost algorithm to select a few features with strong ability of judgment for face detection classification; (3) put forward a cascade structure model and integrate various weak classifiers into a strong classifier which can exclude non-face areas rapidly and improve the detection speed of the algorithm [6, 9, 10].

2.1. Haar-like Feature

The earliest Haar-like rectangle feature library was put forward by PapageorgiousC *et. al.*, Rapid face recognition system of ViolaP et al. uses three-type and five-form rectangle features, namely 2 rectangle feature, 3 rectangle feature and 4 rectangle feature. These features can also be divided into three categories: linear feature, edge feature, point feature (central feature) and diagonal feature, as shown in Figure 1.

The calculation of Haar-like feature is realized through integral image. The integral image $II(x,y)$ of image $I(x,y)$ is defined as:

$$II(x,y) = \sum_{c=1}^{x}\sum_{r=1}^{y} I(c,r) \tag{1}$$

Its integral image $II(x,y)$ can be obtained rapidly through traversal of image $I(x,y)$ with formulas (2) and (3).

$$S(x,y) = S(x,y-1) + I(x,y) \tag{2}$$

$$II(x,y) = II(x-1,y) + S(x,y) \tag{3}$$

Here, $S(x,y)$ refers to the cumulative sum of gray level of pixel points in a row or column. $S(x,-1) = 0$ and $II(-1,y) = 0$.

The sum of gray level of pixel points in rectangular area in image can be calculated rapidly and conveniently through integral image. Then, Haar-like feature of image can be obtained rapidly. Therefore, the sum of gray level in rectangular area D is:

$$Sum(D) = II(x4,y4) + II(x1,y1) - II(x2,y3) - II(x3,y3) \tag{4}$$

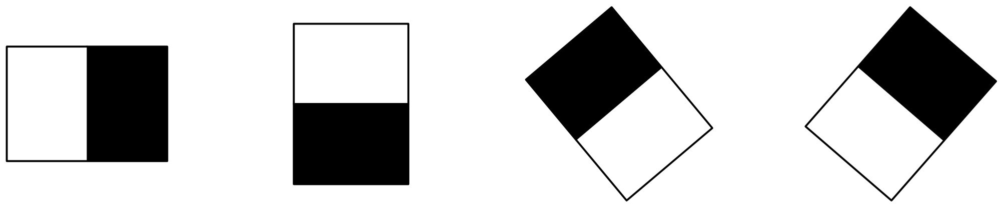
First, it is assumed that an image is W×H pixel and we use $r = (x,y,w,h,a)$ to represent a rectangle in the window and $0 \le x, x+w \le W$, $0 \le y, y+h \le H$, $x,y \ge 0$, $w,h > 0$, $\alpha \in \{0^\circ, 45^\circ\}$.

If the pixel sum of this rectangle is RecordSum(r), a Haar-like feature of an image can be defined as:

$$feature_1 = \sum_{i \in \{1,...,N\}} \omega_i \cdot RecordSum(r_i) \quad (5)$$

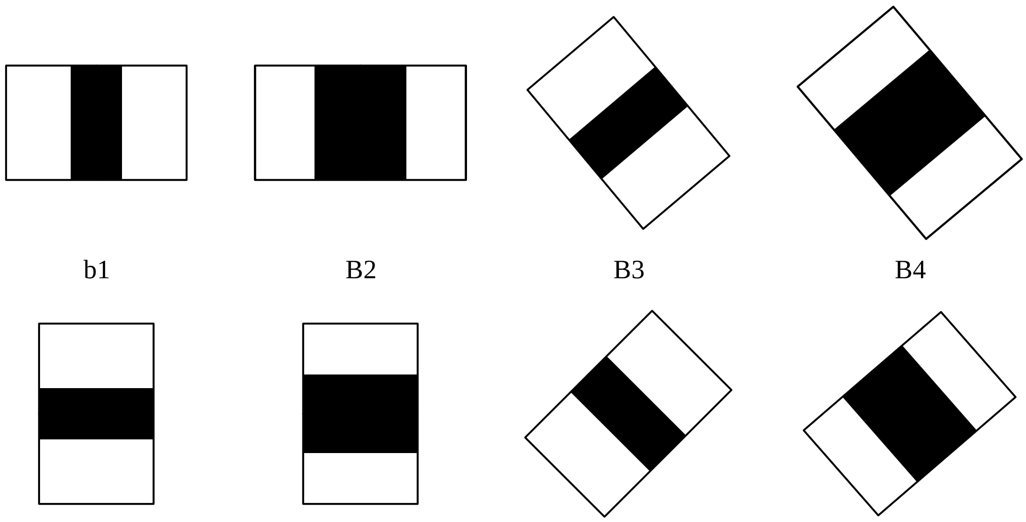
Where, ω_i refers to the weight of the ith rectangle, $RecordSum(r_i)$ refers to the sum of all pixel values in the ith rectangle, and $\{1,\ ...,\ N\}$ refers to the number of rectangles constituting feature1. Let's take 36×36 image for example. The number of all types of Haar-like feature is calculated with formula (5), as shown in Table 1.

For each child window of image, many Haar-like features of different sizes will be produced. Its huge quantity is far beyond the number of pixels in the child window of image. Though these features can be obtained rapidly with the method of integral image, its calculation is time-consuming.



A1 A2 A3 A4

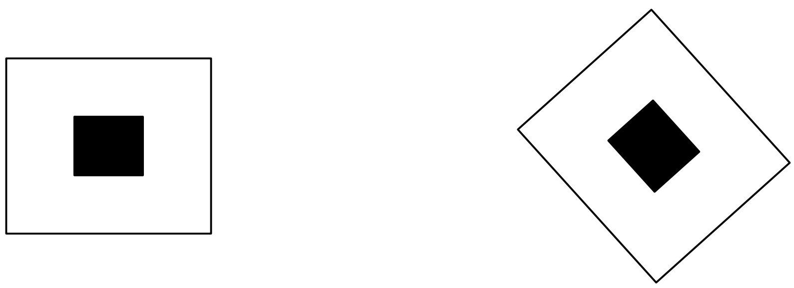
(A) Edge Features



b1 B2 B3 B4

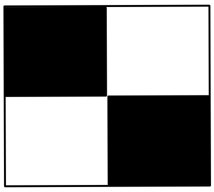
B5 B6 B7 B8

(b) Linear Feature



C1 C2

(c) Point Features (Center Features)



D1

(D) Diagona Features

Figure 1. Haar-like Feature

Table 1. Number of Various Haar-like Features of 36×36 Image

Feature type	Quantity
A1，A2	39150
A3，A4	14680
B1，B2	28760
B3，B4	16800
B5，B6	6500
B7，B8	5300
C1，C2	11200
D1	8658
Sum	131048

2.2. AdaBoost Algorithm

Boosting, also called as reinforcement learning or improvement method, is an important integrated study technique which can enhance weak learner with prediction accuracy slightly higher than random estimate into strong learner with high prediction accuracy. This provides an effective new thought and method for the design of learning algorithm when it is very difficult to establish a strong learner directly. [4] The basic process of AdaBoost algorithm will be introduced with driver face image locating as an example:

(1) A series of training samples $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ are given, where $y_i = 0$ means that it is a negative sample (*i.e.*, non-face), $y_i = 1$ means that it is a positive sample (face), x_i refers to Haar-like feature vector obtained for each sample image and n refers to the total number of trained samples.

(2) Initialized sample weight: $w_i^1 = 1/N$ where $i = 1, ..., N$.

(3) The following operations are conducted for t=1,…,T:

1) Normalize the weight of all samples:

$$q = \frac{\omega_{t,i}}{\sum_{j=1}^{n} \omega_{t,j}} \tag{6}$$

2) Train a weak classifier $f_t(x) \in \{1, -1\}$;

3) Calculate the sum of sample weights with errors of categorization by this weak classifier, *i.e.*,

$$err_t = E_w[I_{y+f_t(x)}],\ c_t = \log((I - err_t) / err_t) \tag{7}$$

4) Update trained sample weight:

$$\omega_i^{t+1} = \frac{\omega_i^t \exp\{-\alpha_t y_i h_t(x_i)\}}{C_t} \tag{8}$$

Where C_t is a normalized constant and $\sum_{i=1}^{N} \omega_i^{t+1} = 1$.

5) Final output:

$$f(x) = sign(\sum_{t=1}^{T} \alpha_t h_t(x) - Th) \tag{9}$$

The above is the basic thought of AdaBoost. Each iteration produces a new weak classifier (just as judge) and each weak classifier judges (vote for) a sample. Finally, voting results are summed and the judgment conclusion is made. The weight of samples with correct classification will decrease before the next iteration and the weight of samples without correct classification will increase greatly. The influence of judgment result of the sample is improved in the next iteration and the judgment accuracy rate is gradually improved. According to this thought, this paper puts forward many derivative algorithm types with further research.

2.3. Cascade Structure Classifier

Cascade classifier means that the classifier is divided into many stages and the next stage can be entered only when the previous one is passed. In this way, many simple picture windows can be excluded rapidly in first stages, thus saving time to detect those areas more like the target. The structure of this cascade detector is shown in Figure 2.

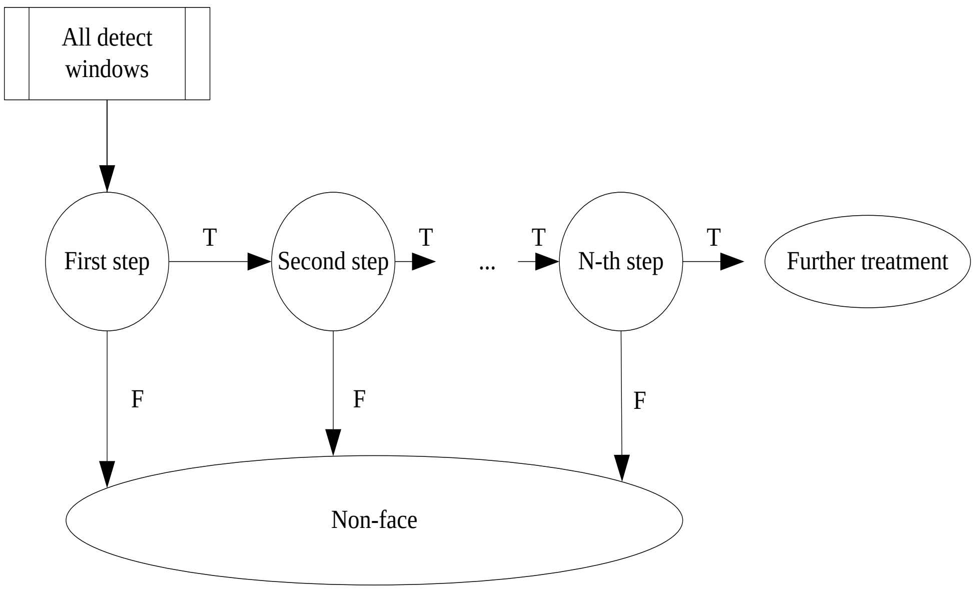


Figure 2. Cascade Structure Diagram

In this cascade structure, classifiers at the former stages use a few features and have a simple structure. They are required to exclude those non-face samples with great differences from face samples and have high detection rate. Generally, the detection rate of the former classifiers should reach 100%. The latter classifiers use many features and have a complicated structure so as to distinguish those non-face samples similar to face samples. This cascade structure makes a lot of non-face samples excluded in simple classifiers and only a few samples can enter the latter complicated classifiers. Therefore, the computing speed of the algorithm is improved to a great extent.

2.4. AdaBoost Algorithm improved in this Paper

This paper further improves AdaBoost algorithm based on paper [1]. In allusion to the asymmetry of positive and negative samples in face detection, the strategy for algorithm improvement is as below: when a weak classifier h_t has strong classification ability especially strong recognition ability of face sample, its evaluation coefficient a_t should increase. During the t^th^ cycle, assuming that the sum of weights of face samples correctly recognized by weak classifier h_t is p_t under error rate ε_t, the value of a_t is adjusted as below:

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right) + k\cdot\exp(p_t) \qquad (10)$$

Where

$$p_t = \sum_{y_i=1, h_t(x_i)=1} \omega_i^t \qquad (11)$$

It is assumed that training sample set

$S = \{(x_1,y_1),(x_2,y_2),(x_3,y_3),...,(x_n,y_n)\}$, $y_i \in \{1,2,...,K\}$. $h_t(x,l)$ refers to the confidence level of label l output by $h_t(x)$, $l = 1,2,...,K$.

$\forall x_i \in S$, when it is in section j of S divided by $h_t(x)$, $P_r[y_i = l] = W_l^{jt} \Big/ \sum_{k=1}^{K} W_k^{jt}$

where $W_k^{jt} = \sum_{i:(y_i=k)\wedge(x_i\in S_j)} \omega_i^t$.

If $h_t(x)$ are mutually independent, the following K classification AdaBoost algorithm can be obtained by direct multiplication of probabilities with the same label, the output of the label corresponding to the maximum probability according to Bayes statistical inference, the use of logarithms and the mission of public items of each label.

(1) Initialized weight: $\omega_i^l = 1/m$, i=1,2,…,m.

(2) The following operations are conducted for t=1,2,…,T:

1) Train weak classifier based on training set S with ω_i^t :

a. Divide S, $S = S_1 \cup S_2 \cup ... \cup S_n$. When $i \neq j$, $S_i \cap S_j = \varnothing$.

b. Count the cumulative sample weight with label k in S_j, $W_k^{j,t} = \sum_{i:(y_i = k) \wedge (x_i \in S_j)} \omega_i^t$, k=1,2,…;K=1,2,…,m;j=1,2,…,n.

c. Define $h_t(x,l)$: $\forall x_i \in S_j$ and make $h_t(x,l) = \ln(W_l^{j,t})$, l=1,2,…,K; j=1,2,…,n.

d. Select $h_t(x)$, minimize $Z_t = K \sum_{j=1}^{n} \sqrt[K]{\prod_k W_k^{j,t}}$ and select $h_t(x)$, *i.e.*, $h_t(x) = \arg\min_{h \in H} Z_t$.

2) Adjust sample weight, $\omega_i^{t+1} = \frac{\omega_i^t}{Z_t} \exp\times\left(-\ln(W_{y_i}^{j,t}) + \frac{1}{K}\sum_{k=1}^{K}\ln(W_k^{j,t})\right)$.

(3) Strong classifier: $H(x) = \arg\max_l f(x,l)$, where $f(x,l) = \sum_{t=1}^{T} h_t(x,l)$.

If weak classifiers are mutually independent, the analysis above has shown that the algorithm is equivalent to Bayes statistical inference. Whether the adjustment of sample weight and the selection of weak classifier can make the selected weak classifiers mutually independent is the key to the effectiveness of the method. Weight adjustment formula of the algorithm is analyzed. When $W_{y_i}^{j,t} = \max_k W_k^{j,t}$, $W_{y_i}^{j,t} = \max_k W_k^{j,t}$. Therefore, this adjustment coefficient reflects higher weight of wrongly classified samples. The selection strategy of weak classifier reflects the feature of focusing on wrongly classified samples. Their combination can make newly selected weak classifiers mutually independent to the greatest extent.

3. Experimental Result and Analysis

To verify the speed and performance of the algorithm in this paper, the experiment uses AT&T face training sample library, MIT face training sample library, ORL face training sample library and face pictures from the network. 2000 face samples are selected from these sample libraries, including samples of the same person under different light conditions, right face, non-right face and shaded face. The same face sample is processed with different methods and experimental results are compared. The experimental environment is Core i5 quad-core processor, 10G memory and Windows7 64-digit operating system. MATLAB7.0 is used as operation platform. Table 2 shows the experimental result.

According to the experimental result, the algorithm in this paper has higher recognition rate and lower omission rate in face detection compared to traditional AdaBoost and SA-AdaBoostSVM algorithms. The method in this paper can improve the training convergence speed of face detection classifier, obviously improve the speed and property of right face detection and achieve right face detection for input images. Figures 3-7 respectively show face detection effects.

Table 2. Comparison of Face Detection with the Method in this Paper, Traditional AdaBoost Algorithm and Sa-AdaBoostSVM Algorithm

	Recognition rate (%)	Omission rate (%)
Algorithm in this paper	93.41	2.02
Traditional AdaBoost algorithm	92.16	6.75
SA-AdaBoostSVM algorithm	93.12	2.18

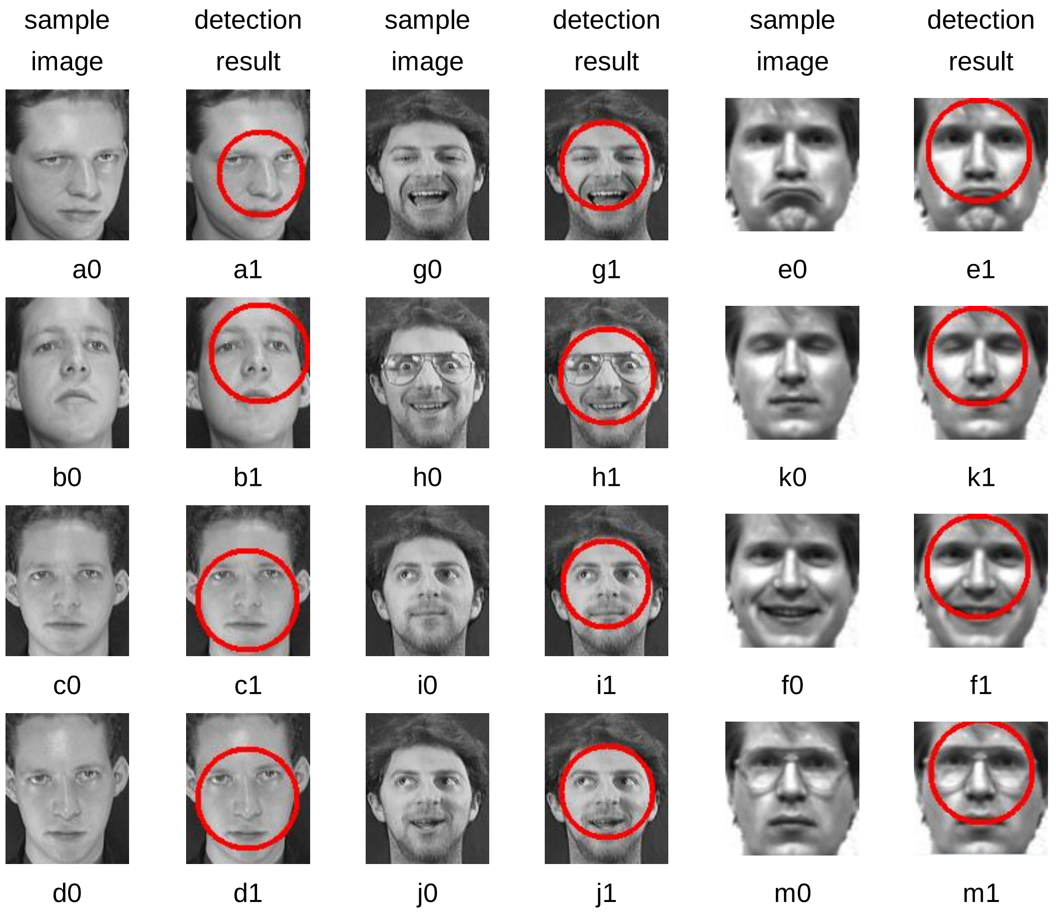


Figure 3. Multi-Pose and Multi-Expression Face Locating Effect in Sample Libraries (AT&T and Yale Face Libraries)

n0 multi-pose n1 detection result

Figure 4. Face Locating Effect under Driving Environment (Multi-Pose Life Photo)

tp0 front p1 detection result

Figure 5. Face Locating Effect under Driving Environment (Front, Muti-Face)

q0 original image q1 detection result

Figure 6. Face Locating Effect under Complicated Background Environment (life photo)

r0 original image r1 detection result

Figure 7. Multi-Face Locating Effect (Life Photo)

Table 3. Comparison of Time Spent in Three Methods

Sample number	Time spent (ms)			Sample number	Time spent (ms)		
	Method in this paper	Traditional AdaBoost algorithm	SA-AdaBoost SVM algorithm		Method in this paper	Traditional AdaBoost algorithm	SA-AdaBoost SVM algorithm
a0	1.48611	2.18957	1.57845	i0	2.82591	2.95617	2.95324
b0	2.56458	2.75984	2.70368	j0	2.73240	2.84781	2.86148
c0	2.57609	3.01894	2.68715	k0	2.42487	2.68125	2.56374
d0	1.82721	1.95148	1.86741	m0	2.27445	2.49216	2.36802
e0	2.33100	2.56921	2.41695	n0	2.61715	2.96540	2.81523
f0	2.26988	2.36800	2.20954	p0	2.65149	2.85247	2.71497
g0	2.08681	2.26125	2.16412	q0	15.7542	21.8756	18.6570
h0	3.03840	3.52198	3.19851	r0	22.8803	27.9654	23.4781

4. Conclusion

In allusion to the slow training speed of AdaBoost and SA-AdaBoostSVM algorithms and their excessive dependence of detection result on samples, this paper uses feature reduction to improve the training speed and introduces sample expansion to improve the detection efficiency. The algorithm in this paper has good effect and low omission rate in the detection of single face and has an ideal effect in the detection of multi-post and multi-expression faces and faces under complicated background. Therefore, this algorithm has certain robustness for facial gestures and expression change. However, it does not have robustness for rotated and side faces. This is a problem to be solved in follow-up work.

Acknowledgment

This work was supported by the key projects of Heilongjiang natural science foundation (ZD201203/C1603).

References

[1] G. Song, "The Research of Several Key Technologies of Face Detection", Harbin Engineering University, HeiLongJiang, **(2011)**.

[2] Q. Haifeng, "Researeh of Driver's Drowsy Detection System Based on Face Detection", Guangdong University of Technology, Guangdong, **(2008)**.

[3] L. Chuang, D. X. Qing, W. Y. Shou, "A Revised AdaBoost Algorithm AD-AdaBoost", Chinese Journal of Computers, vol. 30, no. 1, **(2007)**, pp. 103-109.

[4] C. Ying and M. Qi-Guang, "Advance and Prospects of AdaBoost Algorithm", Acta automatic sinica, vol. 39, no. 6, **(2013)**, pp. 745-758.

[5] X. xin and H. Xiao-li, "Long distance face detection based on Adaboost algorithm", Computer engineering and design, vol. 36, no. 4, **(2015)**, pp. 983-1021.

[6] F. Zhong-liang, "Analysis and Improvement on Real AdaBoost Algorithm", Journal of University of Electronic Science and Technology of China, vol. 41, no. 4, **(2012)**, pp. 545-551.

[7] L. Sheng-wen, B. Su-su, "Face recognition based on PCA+AdaBoost algorithm", Computer Engineeirng and Applications, vol. 46, no. 4, **(2010)**, pp. 170-173.

[8] L. Zhang-li, “Face Location Algorithm in Video Sequences with Complex Background” , Computer Science, vol. 35, no. 6, **(2008)**, pp. 255-279.

[9] C. Yanxiang, “Safety belt detection system based on AdaBoost”, Electronic Measurement Technology, vol. 38, no. 4, **(2015)**, pp. 123-127.

[10] A. Heng-xuan, “A face detection based on AdaBoost algorithm realize on the Android platform”, Electronic Design Engineering, 2014, vol. 22, no. 8, **(2014)**, pp. 126-129.

Authors

Zhao Lei received his BS in Computer science and technology from Northeast Forestry University, Harbin, China, in 2005. He got his MS in Computer application technology from Northeast Forestry University, Harbin, China, in 2009. He got his ph.d in Forestry Engineering Automation from Northeast Forestry University, Harbin, China, in 2013.He is an Associate Professor in the School of Hei LongJiang International University. His research interests include pattern recognition and intelligent control.

Honge REN received the Ph.D. degree from Northeast Forestry University, China, in2009. She is currently professor of information and computer engineering college at school of Northeast Forestry University, supervisor of Dr. Her main research interests include the pattern recognition and intelligent control.

