Face Recognition Based on Real AdaBoost and Kalman Forecast

Chao Yan[1], Yuanqing Wang[1], and Zhaoyang Zhang[2]

[1] Stereo Imaging Laboratory, NanJing University, Nan Jing, China
yanchao3756@sina.com, yqwang@shu.nju.cn
[2] Advanced Display Laboratory, Shanghai University, Shanghai, China
zhangzhaoyang@shu.edu.cn

Abstract. In this paper, a novel face recognition method based on Real AdaBoost algorithm and Kalman Forecast is implemented. Real AdaBoost algorithm can obtain great accuracy with machine learning. Meanwhile, Kalman Forecast is introduced to track human faces detected, making face detection more efficient. We tested our new method with many video sequences. The detection accuracy is 98. 57%, and the average processing time on a windows XP, PIV 2.4GHz PC was less than 20 ms for each 640*480-pixel image. So the proposed face recognition method is real-time.

Keywords: Real AdaBoost, Feature Calculation, Integral Image, Kalman Forecast.

1 Introduction

Face recognition means to obtain the location, size and position of all faces in an image [1]. This technology is used extensively, such as Teleconference, Security Check System, human-computer interaction and so on. Accuracy and speed are two standards of evaluating a face recognition algorithm. For accuracy, face recognition is a complicated and delicate process, face rotation, large-scale head motion and illumination variation will all make influence on the final effect of face recognition. For speed, more and more application of face recognition ask for higher speed than 30fps to fulfill the real-time requirement.

Owing to the importance of itself, face recognition has been recently bringing up a few excellent algorithms. Such as: Template Matching [2], Frame-to-frame Difference [3] , Support Vector Machine SVMs[4].The face recognition systems based on these algorithms behave pretty well in accuracy and recognition speed.

In 2001, Viola combined Discrete AdaBoost algorithm with waterfall structure, making a face recognition system. Then, Schapire improved the output of weak classifier to be continuous and obtained a face recognition system, which was based on Real AdaBoost algorithm and waterfall structure. This kind of face recognition system is one of the systems the accuracy of which are most high, meanwhile, the recognition speed of the system is almost much faster than any other face recognition system.

H. Deng et al. (Eds.): AICI 2011, Part III, LNAI 7004, pp. 489–496, 2011.

Even if the face recognition based on Real AdaBoost algorithm and waterfall structure has those advantages that have been mentioned, face detection for multiple users consumes considerable time. To solve this problem, Kalman Forecast is employed to predict face locations in current frame according to faces detected in the last frame. With Kalman tracking and marking faces detected, the whole detection process is quicker a lot.

The contribution of this paper includes: (1) Real AdaBoost algorithm has continuous output confidence-rate, so it could get classification border more accurately, making the accuracy of face recognition much higher. (2) To optimize recognition speed, Kalman Forecast is utilized to predict face locations, saving the whole detection time. This is very significant for multiple faces detection in real time. (3) We design active infrared illumination, and collect all images in the same illumination condition, so the effect caused by illumination variation is eliminated.

2 Real AdaBoost Algorithm

As implied by the name, boosting algorithm could switch weak classifiers to strong classifiers by integrating and machine learning. AdaBoost algorithm is one kind of Boosting algorithm, which could be adaptive. AdaBoost algorithm is able to adaptively adjust the weight of training samples, and selects the best weak classifiers, then integrates them to become a strong classifier, in which the different weak classifiers vote respectively.

2.1 Feature Calculation and Integral Image

Given limited information, the recognition based on features could code the condition of special areas, and the recognition based on features is much faster than the recognition based on pixels. So AdaBoost algorithm is an algorithm that is based on features. Haar features are used in our AdaBoost algorithm. Haar features are brought up by Paul Viola, which are a kind of simple rectangle features. The value of Haar feature is defined the sum of gray scale value of the black (white) pixels subtracts the sum of gray scale value of the white (black) pixels.

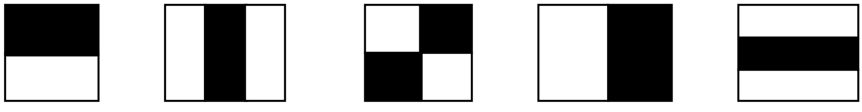


Fig. 1. Basic Haar feature

These five kinds of basic Haar features in Fig.1are used in our face detection phase, and they behave great for front faces and leaning faces less than 30 degree. Nevertheless, large degree face leaning will cause failure to our face detection. So some kinds of Haar features have been taking into our consideration in order that the large leaning faces could also be detected in our face detection phase. Some of these

Haar features are shown in Fig.2, and the classifiers with these Haar features are now during the process of machine learning.

Fig. 2. Some of improved Haar features

These Haar features could be placed in any size and on any position of images; every placing form could be called a feature. So the same recognition sub-window will have a lot of features. For example, a 24*24-pixel image has more than 160000 Haar features. So if we calculate their feature values directly, it will cost considerable time. So integral image should be used in calculating feature values. Value of every point of integral image is the sum of gray scale values of the point's top-left pixels. For example, the value of point A(x,y) in integral image is:

$$I(x, y) = \sum_{\substack{x' \le x \\ y' \le y}} i(x', y')$$

i(x',y') is gray scale value of point (x',y') in the formula. Besides, we define the sum of gray scale values of point (x,y) and its top points as b(x,y):

$$b(x, y) = \sum_{y' \le y} i(x, y').$$

And we define $b(x,0) = 0, I(0, y) = 0$.

As a result, we could calculate I(x,y) as follows:

$$I(x, y) = I(x-1, y) + b(x, y)$$
$$b(x, y) = b(x, y-1) + i(x, y)$$

With integral image, we could calculate feature values much faster. As is shown in Fig.3, the sum of gray scale values of area 1 could be got only with the values of points A,B,C,D.

$$I_1 = I_A + I_D - I_B - I_C$$

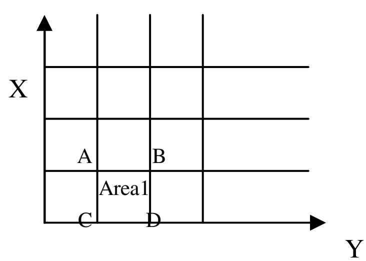


Fig. 3. Integral image

The calculating of rectangle Haar feature values, with integral image, is only related to the values of four rectangle's endpoints and the calculating is only concerning adding and subtracting. Therefore, integral image improves the speed of face recognition largely[5].

2.2 Introduction of Real AdaBoost Algorithm

(1) Given the set of training samples S={(x1,y1),(x2,y2),…… (xn,yn)}, weak classifier space H. In the set, x∈X is a sample vector; y=±1 is class label; n is the number of samples. The initialized sample weight is Dt(i)=1/n,i=1,2,……,n.

(2) For t=1,2,……,T(T is the number of features which are aimed to get.):

①Apply the following steps to every weak classifier in H:

1) Divide the sample space X to x1, x2, x3, …….., xn;

2) With the weight of training samples Dt, calculating:

$$W_k^i = P(x_i \in X_j, y_i = k) = \sum_{\substack{x_i \in x \\ y_i = k}} D_t(i),\ k = \pm 1$$

3) Under the division, set the output of weak classifier as:

$$\forall x \in X_j, h(x) = \frac{1}{2}\ln(\frac{W_{+1}^j + \varepsilon}{W_{-1}^j + \varepsilon}),\ j = 1,2,\ldots\ldots,\ m.$$

e is a tiny positive number.

4) Calculating the initialization factor:

$$Z = 2\sum_j \sqrt{W_{+1}^j W_{-1}^j}$$

② Select ht in weak classifier space to minimize Z:

$$Z_t = \min_{h \in H} Z,$$

$$h_t = \arg\min_{h \in H} Z,$$

③ Update the weight of training samples:

$$D_{t+1}(i) = D_t(i)\frac{\exp[-y_i h_t(x_i)]}{Z_t\%}$$

Zt is the initialization factor, to make Dt+1 a probability distribution.

(3) The final strong classifier is:

$$H(x) = sign[\sum_{t=1}^{T} h_t(x) - b]$$

b is threshold which is set manually, usually 0. Similarly, we define the confidence rate of H is:

$$conf_H(x) = \left| \sum_t h_t(x) - b \right|$$

3 Kalman Forecast

Simulating the motion situation of the "ready-for-detect" target in front of CCD camera, we suppose the motion of the target both in x and y axles are even-speed straight motion which is bothered by a random acceleration α[6]. α is a random variable, $a(t) \sim N(0, \sigma_\omega^2)$ .And the motion state vector of the "ready-for-detect" target is supposed as follows:

X(k)=[XMO(k),YMO(k),Vx(k),Vy(k)]T,

where XMO(k) and YMO(k) are the abscissa and ordinate of the "ready-to-detect" target; Vx(k) and Vy(k) are the speeds of the "ready-to-detect" target in x and y axles. The measure matrix is Y(k):

Y(k) =[XME(k),YME(k)]T,

where XME(k) and YME(k) are the measure abscissa and measure ordinate of the "ready-to-detect" target. So Kalman anticipation algorithm includes two models:

Motion state vector model:

X(k+1)=A(k)X(k)+W(k),

where A(k) is state transition matrix and W(k) donates system perturbation.

Measure vector model:

Y(k) =C(k)X(k)+M(k),

where C(k) is the state transition matrix from current to current measurement and M(k) represents measurement uncertainty. Because the motion of the target is supposed to be even-speed straight motion and Y(k) only involves position, these two models could be also described with the following two matrixes:

$$\begin{bmatrix} X_{MO}(k+1) \\ Y_{MO}(k+1) \\ Vx(k+1) \\ Vy(k+1) \end{bmatrix} = \begin{bmatrix} 1,0,t,0 \\ 0,1,0,t \\ 0,0,1,0 \\ 0,0,0,1 \end{bmatrix} \begin{bmatrix} X_{MO}(k) \\ Y_{MO}(k) \\ Vx(k) \\ Vy(k) \end{bmatrix} + W(k)$$

and

$$\begin{bmatrix} X_{ME}(k) \\ Y_{ME}(k) \end{bmatrix} = \begin{bmatrix} 1,0,0,0 \\ 0,1,0,0 \end{bmatrix} \begin{bmatrix} X_{MO}(k) \\ Y_{MO}(k) \\ Vx(k) \\ Vy(k) \end{bmatrix} + M(k)$$

.

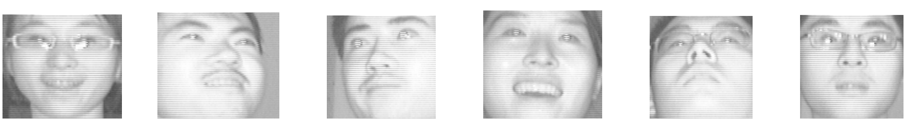
where t is the time interval between adjacent images; the state transition process noise covariance σ_ω^2 equals 1; M(k) is normally distributed as p(M)~N(0,R), and R represents measurement noise covariance.

4 Active Illumination

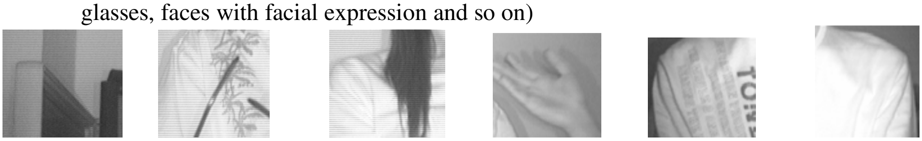
The illumination condition influences the effect of face recognition to a great extent because that face recognition is based on the calculating of gray on faces. And in the process of practical face recognition, the illumination condition always changes. Shiny illumination, dark illumination and shadow will all make the accuracy of face recognition descend to a great extent. Up to present, illumination variation is still one of the bottlenecks of the practical face recognition system[7]. So we need to compensate the illumination. Many solutions have been brought up for the illumination compensation, such as SFS method[8], lateral marking of face organs[9], and two-dimensional Gabor strengthening edge outline[10]. Some of these methods are directed against special images, and others have complicated calculating process, so they all have limitation respectively in the practical application.

With the research of infrared illumination condition, we discover that, in infrared illumination condition, the distribution of gray is much more even than that in natural light condition. So against variation of natural light, the variation of infrared illumination influences the effect of face recognition in a much smaller degree. And the active infrared illumination this paper brings up is a kind of illumination compensation based on hardware, which could make sure every image collected is in the same illumination condition. So the active infrared illumination not only compensates illumination condition greatly, but also extends the sphere of illumination compensation.

We use some hardware to supply infrared illumination and capture images, which contains the array of infrared LED, monitor, video capture card, video transmission line and PC. Array of infrared LED is responsible for providing infrared illumination in stable intensity; monitor is responsible for capturing images and video transmission line is responsible for transmitting the images to video capture card which is installed in PC. Then, with some relative program, PC will deal with the images.



(a) Examples of face samples(including front faces, a little degree leaning faces, faces with glasses, faces with facial expression and so on)



(a) Examples of non-face samples

Fig. 4. Examples of samples

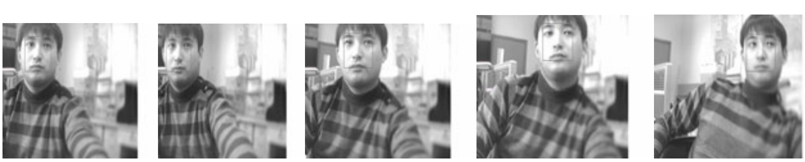


Fig. 5. Face recognition in practical video sequence (the black frames in the images are face area)

5 Experiment and Conclusion

In our experiment, we train a waterfall face recognition system which has eight layers [11]. Because in practical application, non-face samples are much more than face samples, we decide that the ratio of face number and non-face number is 1:13. The number of face samples in our experiment is 6000, and the number of non-face samples is 78000. Besides, to solve the problem that the non-face samples, which have more abundant texture, are much easier to be classified wrongly, we make sure the non-face samples which have abundant texture are of a certain proportion when we do our best to make non-face samples discrete. Some samples we use in our experiment are shown in Fig.4. Wrong-recognition rate of our face recognition system is less than two percent in our experiment. The practical face recognition in video sequence is shown in Fig.5. There is no miss-recognition of faces in our experiment, and rare wrong-recognition of non-faces.

In this paper, a novel face recognition method based on Real AdaBoost and Kalman Forecast has been brought up. Real AdaBoost algorithm has continuous confidence-rate, so it could raise the face recognition accuracy; with tracking human faces detected, Kalman Forecast can speed up the whole face recognition process; active illumination could provide infrared illumination in stable intensity, and basically eliminate the influence of illumination variation on face recognition. With the experiment, we prove that the new recognition method has great accuracy and it is real-time. At the same time, it is robust for illumination variation, a little degree leaning of human face and variation of facial expression. So the application of this novel face recognition based on Real AdaBoost and Kalman Forecast will have a pretty good future.

Acknowledgement. This research activity has been partially funded by Chinese National Natural Science Fund Commission, the Scientific Research Foundation of Graduate School of NanJing University and Research and Innovation Project for College Graduates of JiangSu Province.

References

1. Liang, L.: A Survey of Human Face Recognition. Chinese Journal of Computer 5, 449–458 (2002)
2. Zhangjie, Yang, X., Zhao, M.: Eyes detection based on Hough transform. Compute Engineering and Application 27, 43–44 (2005)
3. Yano, K., Lshihara, K., Btakikawa, M.: Detection of eye blinking from video camera with dynamic ROI fixation. IEEE International Conference on Systems, Man, and Cybernetics 6, 335–339 (1999)
4. Kumar, V.P., Poggio, T.: Learning based approach to real time tracking and analysis of faces (1999), `http://cbcl.mit.Edu/cbcl/publications/ai·publications`
5. Zhao, N.: Face Recognition Based on AdaBoost. Bachelor Thesis. Peking University (2005)

6. Yan, C., Wang, Y., Zhang, Z.: A Novel Real-Time Eye Detection in Human-Computer Interaction. In: 2010 Second International Conference on E-Learning, E-Business, Enterprise Information Systems, and E-Government (EEEE 2010), December 11, vol. 1, pp. 57–62.
7. Zheng, Q.: Facial illumination compensation based on composite transformation (February 2008)
8. Tsengyc, Chenyy, Panhk: A secure data hiding scheme forbinmy images. IEEE Transactions on Communications 50(8), 1227–1231 (2002)
9. Fridich, J., Goljan, M., Hogea, D.: Steganalysis of JPEG images: Breaking the F5 algorithm. In: Petitcolas, F.A.P. (ed.) IH 2002. LNCS, vol. 2578, pp. 310–323. Springer, Heidelberg (2003)
10. Fridich, J., Goljan, M., Du, R.: Recognition LSB steganography in color and gray-scale images. Magazine of IEEE Multimedia: Special Issue on Security 8(4), 22–28 (2001)
11. Viola, P., Jones, M.: Rapid Object Recognition using a Boosted Cascade of Simple Features. In: Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, Kauai Hawaii, USA, pp. 905–910 (2001)