

## Face Detection Based on Feature Tailoring and Skin Color Space

Wenbo Jiang & Xiaoming Xie

*College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, China*

**ABSTRACT:** This paper is used to solve the time-consuming problem of training samples in Adaboost algorithm and propose an improved FTAdaboost algorithm based on feature tailoring. In the beginning, this paper is used to make all samples have the same weight, train them once and tailor the features before the first reflection point of the error rate curve which have high error rate and poor classification ability, then reduce the number of samples and save training time. According to the distribution of facial organs, the algorithm determines whether the specified area meets the characteristics of skin-color space, then eliminates the influence of wrong facial images. The experimental results show that the algorithm based on feature tailoring can shorten the training time significantly and the detection with the skin-color space can decrease the error rate to some extent.

**Keywords:** fatigue detection; FTAdaboost; classifier; skin color space

### 1 INTRODUCTION

In recent years, with the rapid development of computer image processing technology, the popular research of computer vision field, face detection had been studied by lots of scholars, and there had been many effective methods which have been widely used in many fields [1].

In numerous face detection algorithms, the machine learning method with high detection accuracy rate has been widely used. Nowadays, the Adaboost algorithm is a powerful representative in these machine learning algorithms. The algorithm is firstly proposed by Freund and Schapire [2], and the core idea is to train different classifiers (weak classifiers) on the same training set, then combine these weak classifiers to a stronger classifier (strong classifier) at last without any prior knowledge of weak classifiers [3]. The theory proves that if the classification ability of each weak classifier is better than random guessing, and when the number of weak classifiers tends to be infinite, the error rate of the last strong classifier will tend to zero. Viola and Jones proposed the concept of integral graphs and cascade detection based on the theory, they greatly improved the speed of Haar computing features and face detection [4-5].

Although the Adaboost algorithm has very high detection accuracy and real-time ability, there also exist some obvious flaws such as too much training time. Because of the diversity of the Haar features, even an image with only 20\*20 size has 78460 features. And in the training process of Adaboost algorithm, each feature needs to be trained to a weak classifier, which takes lots of time. When the number of training samples is large, the training time may even be several weeks long [5].

To solve this problem, many domestic and foreign scholars had made lots of rigorous study. Baumann

proposed a SEAdaboost [6], an algorithm which is based on the facial symmetry. The author thought that one weak classifier is produced in each training round, and according to the facial symmetry, there must exist a weak classifier with small error rate in the symmetrical region of the trained feature. If the symmetry is well controlled, each training round will produce two symmetrical weak classifiers, then, the training time will be decreased by half. Jia Huixing proposed a method based on the dynamic samples tailoring [7]. After each training round, he found samples which have small weight and tailored them, then he continued until the error rate tends to be at a high level, and then he expanded the sample capacity. When the threshold is reasonable, the training speed of algorithm can get an obvious promotion.

Besides, the Adaboost algorithm based on Haar features uses gray images as the training samples, reflects the characters by feature value. In some instances, the error rate is strictly controlled, excluding the wrong detection windows that will have great influence on the result. To save more training time and reduce the error rate, this paper proposed an improved FTAdaboost algorithm based on feature tailoring and skin-color space detection method, cut off a part of features which can't represent the face characters well, decreased the influence of the classifiers which have poor classification ability on the last strong classifier, and reduced the training time. At the same time, it was combined with the skin-color space method at the step of face detection in color images, computed the specified region of the detected face window whether meet the skin-color characteristics or not to verify the candidate face windows in a further step. The experiment result shows that the aforementioned method can reduce the error rate to some extent.

## 2 FTADABOOST BASED ON FEATURE TAILORING

### 2.1 Specific steps of FTAdaboost

The procedure of the proposed algorithm based on feature tailoring is shown as follows:

(1) Given:  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , where  $y_i = 1$  represents the face samples, the total number is  $m$ ,  $y_i = 0$  represents the non-face samples, the total number is  $l$ , and  $m + l = n$ .

(2) For  $t = 1, 2, \dots, T$  ( $T$  is the number of classifiers)

a) Initialization: ( $t \leq 2$ )

When  $t = 1$  to all samples:

$$w = 1/n, \sum_{i=1}^n w_{1,t} = 1 \quad (1)$$

When  $t = 2$  to all samples after feature tailoring:

$$w_{i,t} = \begin{cases} \frac{1}{2m} \\ \frac{1}{2l} \end{cases}, \sum_{i=1}^n w_{1,t} = 1 \quad (2)$$

b) Normalize the weights as follows:

$$q_{t,i} = \frac{w_{t,i}}{\sum_j w_{t,j}} \quad (3)$$

c) To each Haar feature  $j$ , train a weak classifier  $h_j$

$$h_j = \begin{cases} 1 & p_j f_j < p_j \theta_j \\ 0 & \text{others} \end{cases} \quad (4)$$

d) Compute the error rate of classification  $\varepsilon_j$  as follows:

$$\varepsilon_j = \sum_i q_{t,i} |h_j(x_i) - y_i| \quad (5)$$

e) Update the weights as follows:

$$w_{t+1,i} = w_{t,i} \beta^{1-\varepsilon_i} \quad (6)$$

Where  $i = 1, 2, \dots, n$ ,  $\beta = \varepsilon_t / (1 - \varepsilon_t)$ . If sample  $x_i$  is correctly classified, then  $\varepsilon_i = 0$ , or  $\varepsilon_i = 1$ .

f) Feature tailoring: ( $t = 1$ ):

Rank the obtained error rate of the first training round in descending order; make the first inflection point as the threshold; tailor the features before the threshold. To the tailored features.

$$\text{threshold} = \min(i), \text{where } r_{i+1} - r_i \neq 0 \quad (7)$$

(3) Compose the final strong classifier as follows:

$$h(x) = \begin{cases} 1, \sum_{t=1}^T \alpha_t h_t \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, \text{others} \end{cases} \quad (8)$$

Where,  $\alpha_t = \ln(1 / \beta_t)$ .

### 2.2 Algorithm analysis

The core idea of the algorithm in this paper proposed is: As shown in Figure 1, not all Haar features can reflect the face characteristics significantly, only a small part has good representativeness. Among the other features, there are some features which have poor ability in classifying face images and non-face images. If the features have poor classification ability that can be moved in the previous training round, then the training time can be saved a lot. It can be seen from (4), the higher the error rate is, and the worse the classification ability will be. Therefore, tailoring the features with high error rate and poor classification can decrease the number of total features, save the training time effectively without the effect of performance of the final strong classifier.



Figure 1. Haar features which can reflect face characteristic significantly

According to the experiment data of several face databases, we give every sample the same weight in the initialization, then train once and get the curve of error rate which has two obvious inflection points. Compared with other features, the features before the first inflection point have high error rate, and they

can't classify the face and non-face significantly. So, we choose the first inflection point as the threshold based on the distribution characteristics of error rate to tailor. Figure 2 shows the error rate curve of three face database after initialization. To different database or different count of samples, the error rate curve may be different, but compared with the tailoring with fixed number, it has an adaptability.

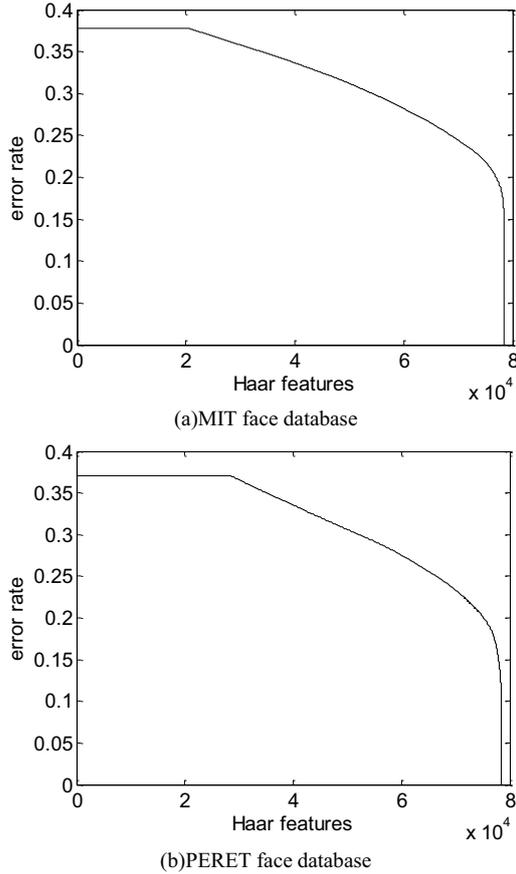


Figure 2. Error rate curve after initialization

### 3 FACE DETECTIN WITH SKIN COLOR SPACE

The Adaboost algorithm may produce some wrong face windows in color images detecting to reduce the wrong face windows and combine the skin-color space method to verify the candidate face windows. Usually, the color images obtained from camera have the format of RGB. In RGB color space, the three basic colors (R, G, B) not only represent color but also represent luminance. They will easy affect the skin-color segmentation as a result of the effect of complexity of light around [8]. In order to separate the color information and the luminance information, a color space conversion must be needed. We chose YCrCb color space in this paper because of the independence of Cr,

Cb and Y component. So, we can verify the face windows based on the distribution of skin-color space. Equation (9) shows the conversion matrix from RGB to YCrCb:

$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.018 \end{pmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (9)$$

In frontal face images, the distribution of skin has obvious characteristics. The chroma of eye, eyebrows, lips and beard has big difference with the skin-color region, and the rest of the area can represent the skin-color region. The samples trained in this paper exclude the effect of hair and ear. In  $20 \times 20$  image, the skin-color region can be shown in Figure 3:



Figure 3. Skin-color region in frontal face image

It can be seen that the distribution of skin-color has obvious rule. The skin-color area can be computed through (10) and (11):

$$k = \text{floor}(w/3) \quad (10)$$

$$\begin{aligned} S_{skin} = & S(x=(w-k)/2+1 \sim (w+k)/2, y=0 \sim (w-k)/2) \\ & + S(x=0 \sim w, y=(w-k)/2+1 \sim (w+k)/2) \\ & + S(x=0 \sim k, y=(w+k)/2+1 \sim w) \\ & + S(x=w-k \sim w, y=(w+k)/2+1 \sim w) \end{aligned} \quad (11)$$

$w$  is the width of the image;  $k$  is the width between eyes;  $x$ ,  $y$  is the coordinates in the image. When the detecting window is enlarged in proportion, the skin-color region is also enlarged which makes it in a certain degree of adaptability.

According to the experiment result from Douglas Chai [9], after computing the skin-color region, we remove the effects of eyes, lips and other non-skin region under different luminance and resolution, the distribution of skin-color in color face images is like this: The Cr component color information distributes in [123,175], and the Cb component color information distributes in [93,133]. Then, the formula for verification can be obtained as follows:

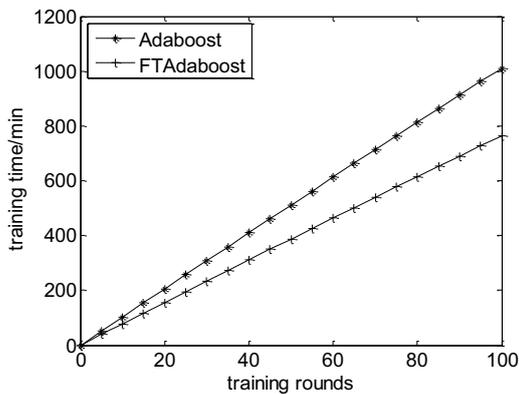
$$F(x) = \begin{cases} 1 & Cr(x) \in [123,175] \& Cb(x) \in [93,133] \\ 0 & \text{others} \end{cases} \quad (12)$$

$x$  represents the candidates face windows when

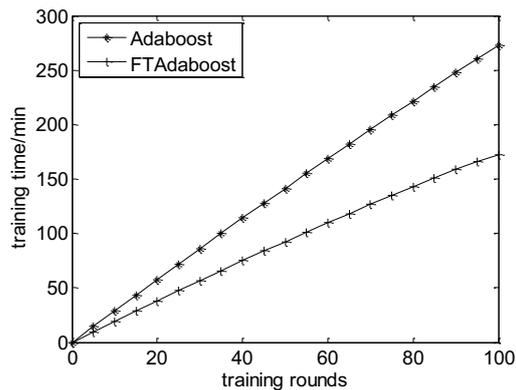
detecting and computing the skin-color region as shown in Figure 3. If  $F(x) = 1$ , then  $x$  is determined to face, otherwise, it is excluded.

#### 4 EXPERIMENT RESULT AND ANALYSIS

In order to test the performance of the Adaboost algorithm based on feature tailoring, the experiment use two test data to get the comparison result. The first data consists of MIT face database which has 2429 face images and 4000 non-face images, the second data consists of PERET database which has 1000 tailored face images without hair and ears, and 850 non-face images. Among them, 500 images are used as training samples and the other 500 images are used in detection, the non-face images are collected from internet. We give a uniform size of  $20 \times 20$  and train 100 times to get the final strong classifier. The PC is used 3.3GHz dual core CPU, 8G memory and the comparison of the training time is shown in Figure 4:



(a) Training time of MIT database



(B) Training time of PERET database

Figure 4. Comparison of training time

The training result shows that when train 100 times, the numbers of the tailored features in MIT and PERET database are 20191 and 28478. Compared

with traditional algorithm, the improved FTAdaboost algorithm almost saves 24.4% and 36.3% of training time and effectively improves the training efficiency.

The accuracy test uses PIE database, YALE database and PERET database. The PIE database and YALE data base contain 3329 and 165 front faces in different light condition, the PERET database contains 500 front images in normal light condition. The detection result is shown in Table 1 and Table 2:

Table 1. Comparison of detection accuracy of MIT

| Detection Method | Face Images |      |        | Training Time/s |
|------------------|-------------|------|--------|-----------------|
|                  | PIE         | YALE | PERE T |                 |
| FTAd-boost       | 1579        | 118  | 448    | 45760.538924    |
| Adaboost         | 1483        | 114  | 439    | 60563.913534    |

Table 2. Comparison of detection accuracy of PERET

| Detection Method | Face Images |      |       | Training Time/s |
|------------------|-------------|------|-------|-----------------|
|                  | PIE         | YALE | PERET |                 |
| FTAd-boost       | 1569        | 93   | 457   | 10326.355576    |
| Adaboost         | 1403        | 81   | 455   | 16210.919272    |

It can be seen from Table 1 and Table 2 that compared with the traditional Adaboost algorithm, the FTAdaboost algorithm has an improvement in accuracy. As the training samples are in normal light condition, the detection is sensitive to light. Some face images in PIE and YALE with little light are not detected, but in PERET database, the face images are tailored in normal light condition, so the detection accuracy is high. However, the experiment results are persuasive enough.

The test with skin-color space needs color images, so the test set consists of 200 images which are all collected from internet and totally have 981 faces. Test these color images in different algorithms, the detection result is shown in Table 3:

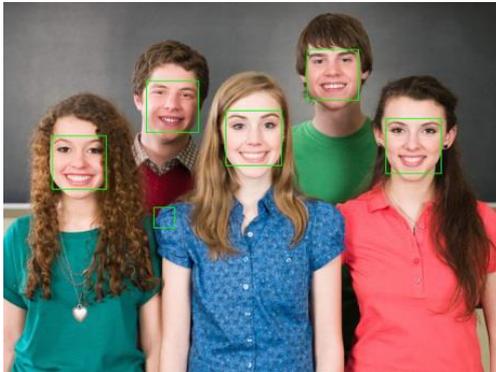
Table 3. Comparison of detection results in different algorithms

| Detection Method                | Faces | Wrong Faces | Accuracy | Error Rate |
|---------------------------------|-------|-------------|----------|------------|
| Adaboost                        | 810   | 174         | 82.6%    | 17.7%      |
| FTAdaboost And Skin-color space | 823   | 42          | 83.9%    | 4.3%       |

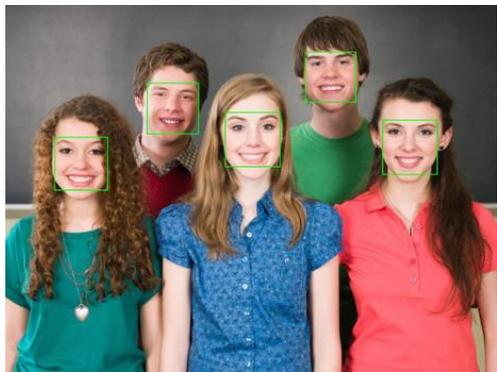
Figure 5 shows some detection results in color images.

The result shows that detection with skin-color space have the same accuracy as compared with the traditional algorithm. However, with lower error rate, it may be a great improvement in the situation which

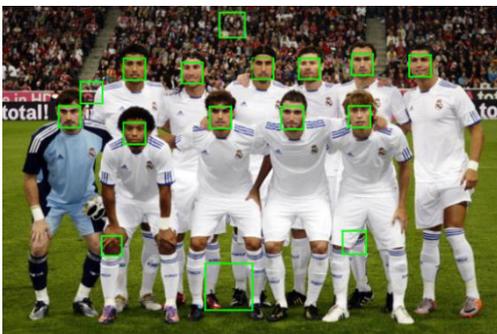
needs low error rate.



(a)



(b)



(c)



(d)

Figure 5. Comparison of detection results in color images

## 5 CONCLUSION

In this paper, we propose an improved FTAdaboost algorithm based on feature tailoring aiming at reduce the training time. In the beginning, we give all samples the same weight, and then we train them once, get the curve of error rate, and find the first inflection point as the tailoring threshold, and tailor the features before the threshold to reduce the number of features to save training time. At the same time, we combine the skin-color space method in color images detection and compute the specified region whether meets the characteristic of face skin-color distribution. Then we verify the candidate face windows and decrease the error rate to some extent. The experiment results show that the improved FTAdaboost algorithm has an improvement in detection accuracy, and it can save the training time significantly. Combining the skin-color space method at the step of face detection in color images, it can reduce the error rate to some extent, meet the real-time characteristic, and lay the foundation for practical applications in the future.

## REFERENCES

- [1] Liang Luhong, Ai Haizhou, Xu Guangyou & Zhang Bo. 2002. A Survey of Human Face Detection. *Chinese Journal of Computers*, 25(5): 449-458.
- [2] Freund Y. & Schapire R E. 1999. A Short Introduction to Boosting. *Journal of Japanese Society for Artificial Intelligence*, 14(5): 77-780.
- [3] Li Shaowen. & Wang Jiangbo. 2013. Research on driver fatigue detection system. *Computer Engineering and Applications*, 49(15): 253-258.
- [4] Paul Viola. & Machael Jones. 2001. Rapid object detection using a boosted cascade of simple feature. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (2): 511-518.
- [5] Paul Viola. & Machael Jones. 2001. Rapid object detection using a boosted cascade of simple feature. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (2): 511-518.
- [6] Baumann F, Ernst K, Ehlers A. & Rosenhahn B. 2010. Semmetry enhanced Adaboost. *International Conference on Advances in Visual Computing*.
- [7] Jia Huixing. & Zhang Yujin. 2009. Fast Adaboost Training Algorithm by Dynamic Weight Trimming. *Chinese Journal of Computers*, 32(2): 336-341.
- [8] Liu Zhenguang. & Liu Jie. 2007. Research on Face Detection Algorithm Based on Complexional Segmentation. *Computer Engineering*, 33(4): 179-181.
- [9] D. Chai & A. Bouzerdoum, 2000. "A Bayesian Approach to Skin Color Classification in YCbCr Color Space," TENCON 2000. Proceedings, *IEEE*, Kuala Lumpur Malaysia, 2: 421-424.