Analysis of face detection based on skin color characteristic and AdaBoost algorithm

Pei Li¹, Hongjuan Wang³, Yeli Li¹, Mengyang Liu¹

¹School of Information Engineering, Beijing Institute of Graphic Communication, Beijing, China
²School of New Media, Beijing Institute of Graphic Communication Beijing, China
Email:1365588043@qq.com, 229359989@qq.com, 1605872754@qq.com, 1367274860@qq.com

Abstract. For the face detection method based on skin color feature and AdaBoost algorithm, if one of them is used to detect the face, it can also catch the face to a certain extent. However, the detection rate and an error rate of this single method in its detection experiment can't achieve good results. Therefore, this paper combines the advantages of the two algorithms, combines the two approaches, and improves them. The main idea is to use the skin color features of face detection as pre-detection, and use the established skin color distribution Gaussian model to obtain candidate regions containing the skin color of the face, and then use a cascade classifier to detect the skin color regions. By using OpenCV and Visual Studio software, a lot of experimental statistics and analysis are carried out. The research shows that the improved algorithm is superior to the two algorithms in detection rate and false detection rate, and it can also achieve a good detection effect for the face in a complicated situation.

1. Introduction

In recent years, the application of computer technology in the customs inspection, security, security access control, military, and police security and other aspects began to increase excessively. With the increasing attention paid to biometric technology in various industries, people are increasingly demanding the functions of machinery and equipment, which means that machines are required to have the ability to perceive the outside world like human sensory organs. Among them, in the field of computer vision, the non-contact authentication, direct, friendly, and convenient features of face recognition technology make the research of face detection technology widely researched and developed, which has become a hot issue of study[1].

At present, students and office workers mainly use the way of swiping in and out of the campus or the company, which has the problems of low efficiency, forgetting to bring cards and losing cards. Face recognition has the characteristics of direct, friendly, and convenient, which makes it easier for people to enter and leave the campus or the company and other places. However, as a complex structure, the details of the face vary greatly. Due to the different skin color, state, appearance, shape, clothing, race, and environment of the detection object, the detection clarity and accuracy are affected differently. In the current face detection research, the main problems are: how to create a high-precision face detection distribution model in the face diversity situation; how to accurately retain the face area in the complex background; how to improve the recognition rate of the system in the face diversity situation. At present, there are four methods for face detection of image or image sequence:
knowledge-based methods, feature invariant approaches, template matching methods, and appearance-based methods.

2. Face detection method based on skin color feature

2.1. Feature extraction

Gray features include facial contour features, facial features (such as symmetry, etc.), facial gray distribution features, template features, etc. For gray features, the contour feature is an essential feature of the human head. In the process of face detection, craw[2] and others first match the approximate contour of the face with the template in the high rule image, then use Sobel[3] algorithm to extract the direction and position in the low rule image, and finally form a complete face contour.

The features of eyes, nose, and mouth in the core area of the face are darker and more evident than those of other features, so it has a unique gray distribution feature. The eye, nose, mouth, and other features of the facial core area are deeper and more evident than other features, so it is a unique gray distribution feature. According to its characteristics, it can be used to distinguish various parts of the face correctly. In these features, the elements of human eyes are more visible. In this premise, the images to be detected are screened, and the images that may have human faces are selected, and then the face is determined by the location algorithm. The gray level of facial region is also used as a template feature, and the central area of eyes, nose, and mouth is usually used as a standard template feature, and this method can eliminate most of the hair and both sides of the cheek in the positioning process.

In the low-resolution image, the facial features of the face will be confused with the background. Based on the research on the detection method of mosaic, Yang[4] et al. Established a hierarchical three-level detection system to determine the face candidate area, find the same features of the face in the candidate area in the low-resolution image, and then select the front from the candidate area with a rule and binarize the face area to further verify the mosaic method.

2.2. Establishment of skin color model

Skin color is one of the most commonly used features in face detection because skin color is an essential part of the face. It has relative stability, which is different from the background color of most images. The skin color model mainly describes skin color features. The skin color model uses the form of a similar table to express the color of the spatial pixels belongs to skin color, or to show the degree that the pixel color is identical to the skin color. In face detection, the selection of chroma space is significant for the collection and use of skin color model form. The color space is shown in the following formula:

\[ r = \frac{R}{R+G+B}, \quad g = \frac{R}{R+G+B} \]

Skin color model is generally divided into clustering model and Gaussian model: the "multi-model tracking of faces for video communications" established by J.L. crowley and F.berard[5] is to normalize
the RGB color of skin color area, and use the \((r, g)\) value of color histogram \(H(r, g)\) to obtain the threshold value of skin color variable.

RGB to "rg" space.

RGB to YUV (YCrCb) space, converting to "FI" space.

\[
\begin{align*}
\begin{pmatrix}
Y \\
U \\
V
\end{pmatrix}
&=
\begin{pmatrix}
0.299 & 0.587 & 0.114 \\
0.147 & 0.289 & 0.436 \\
0.615 & -0.515 & -0.100
\end{pmatrix}
\begin{pmatrix}
R \\
G \\
B
\end{pmatrix}
\end{align*}
\]  

\(F = \tan^{-1}\left(\frac{|V|}{|U|}\right)\)  

\(I = (0.596, -0.274, -0.322)\)

In "Detecting human faces in color images", M.H.Yang and N.A.Huang use the statistical principle to think that the color value of the skin color area of the face presents a Gaussian distribution state, so use the mean and variance of the Gaussian distribution to obtain the threshold value of the skin color variable. According to T.S.Jebara and A.Pentland\(^6\), "Parameterized structure from motion from 3D adaptive feedback tracking of faces", people of different races and countries have different colors. Because there are many clusters on color histogram, it is represented by the Gaussian mixture model.

\[F \sim N(\mu, \sigma)\]

\[
\mu = \frac{1}{N} \sum_{i=1}^{N} F_i, \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (F_i - \mu)^2
\]

\[N(\mu, \Sigma) \text{ where } \mu = (\bar{r}, \bar{g}) \quad F \sim N(\mu, \sigma)\]

\[
\bar{r} = \frac{1}{N} \sum_{i=1}^{N} r_i, \quad \bar{g} = \frac{1}{N} \sum_{i=1}^{N} g_i
\]

\[
\Sigma = \begin{pmatrix}
\sigma_{rr} & \sigma_{rg} \\
\sigma_{gr} & \sigma_{gg}
\end{pmatrix}
\]

Figure 2. One dimensional normal distribution skin color model.
2.3. Face region filtering
Because the skin color areas determined by the established skin color distribution model do not belong to the face area, we need to screen again here. For the background and noise, the existing white isolated points and some small areas, these are usually treated by morphology. The continuous open and close operations are used to eliminate some remote low spots and regions.

First, the skin color region is divided into various areas. We know that the ratio of length to width of the regular face is between 0.8 and 1.6. The rate of height to width of the skin color region that matches this ratio may be the face region, and the part that does not match is the impossible region. Through this kind of pre-knowledge screening, we can not only get the face area but also reduce the skin color area as much as possible, to reduce the possible face area. The following figure shows the detection process of face detection based on skin color features:

3. Face detection method based on AdaBoost algorithm

3.1. AdaBoost algorithm principle
AdaBoost, translated from the English "Adaptive Boosting," is proposed based on Boosting optimization[7]. The difference is that it can determine the weight of samples in the next training according to the previous sample training result[8], which has the characteristics of Adaptive. Adaboost algorithm uses offline learning classifiers for face detection and cascade detection based on Harr
rectangular feature integral graph. The algorithm has high computational efficiency and fast detection speed [9-10]. Viola-Jones [11] algorithm is a face detection algorithm based on Haar-like feature and AdaBoost algorithm classifier. Specifically, the process of AdaBoost iteration algorithm is as follows.

① Initialize data weight distribution. All samples are given the same weight first:

$$D_v = (w_{i1}, w_{i2}, \ldots, w_{i1}, \ldots, w_{iN})$$

Where: \(i=1,2,\ldots,N\)

② Train weak classifiers. Set \(M = 1, 2, \ldots M\) to represent the iteration rounds

Training and learning data sets with a correct value \(D_m\), training obtains a weak classifier.

$$G_m(x): x \rightarrow \{-1, +1\}$$

Calculate the classification error rate of \(G_m(x)\) training set

$$e_m = P(G_m(x_i) \neq y_i) = \sum_{i=1}^{N} w_{mi}I(G_m(x_i) \neq y_i)$$

The error rate \(e_m\) is the sum of the weights of the samples misclassified by \(G_m(x)\).

Calculate the coefficient of \(G_m(x)\)

$$\alpha_m = \frac{1}{2} \ln \frac{1-e_m}{e_m}$$

\(\alpha_m\) represents the importance of \(G_m(x)\) in the final classifier. It can be seen from the formula that when \(e_m \leq \frac{1}{2}\), \(\alpha_m \geq 0\), \(\alpha_m\) increases with the decrease of \(e_m\). Update the training set weight distribution for the next iteration.

$$D_{m+1} = (w_{m+1,1}, w_{m+1,2}, \ldots, w_{m+1,i}, w_{m+1,N})$$

Where: \(w_{m+1,i} = \frac{w_{mi}}{z_m} \exp \left(-\alpha_m y_i G_m(X_i)\right)\) \(i=1,2,\ldots,N\). It is the normalization factor that makes a probability distribution.

③ The strong classifier consists of a weak classifier.

Combined weak classifier:

$$f(x) = \sum_{i=1}^{N} \alpha_m G_m(x)$$

Final classifier (strong classifier):

$$G(x) = \text{sign}(f(x)) = \text{sign}(\sum_{i=1}^{N} \alpha_m G_m(x))$$

3.2. Implementation details

Viola et al. Put forward the concept of an integral image and calculated Haar-like eigenvalue by an essential graph. Haar-like feature is a kind of feature description operator. People usually use the feature-based method to describe the face. The reason is that it runs fast and can meet the needs of real-time detection. It is the first used for face description. Harr features are divided into edge features, center features, diagonal features, and linear features, which are integrated into feature templates. The template consists of black-and-white rectangles. It is defined that the sum of white squares minus the amount of black boxes is the characteristic value of the model. It can roughly reflect the general characteristics of facial features, as shown in the figure:
All rectangular features can be represented by the quintuple of $r = (x, y, w, h, \alpha)$, where $x$, $y$ are the base coordinates of the coordinate axis, $w$ is the width of the rectangular feature, and $h$ is high. $\alpha$ is the angle of rotation.

Integral image and Haar-like have nothing to do with each other. They are two completely independent concepts. In the process of face detection research, they are connected by "Rapid object detection using a boosted cascade of simple features" [12]. In face recognition, the Haar-like feature detection method needs to scan a large number of windows to calculate the relative image feature value each time, which brings a lot of burden to the computing system, thus significantly reducing the speed of training and detection. Through research, it is found that using an integral graph method can calculate the Haar feature faster. The function information can be obtained by traversing once:

$$g(u, v) = \sum_{i=0}^{u} \sum_{j=0}^{v} f(i, j)$$

Where $(u, v)$ represents the sum of all pixels on the top left of $(u, v)$ in the original image. With the help of an integral graph, the calculation of the sum of pixels in the rectangular areas can be completed quickly.

The weak classifier obtains the strong classifier after several iterations. For each feature $f$, a weak classifier $H(x, f, p, \theta)$ needs to be trained after the number and value of rectangular features are determined.

$$h(x, f, p, \theta) \begin{cases} 1 & \text{if } pf(x) < p \theta \\ 0 & \text{others} \end{cases}$$

$$e = \min(S^+ + (T^- - S^-), S^- + (T^+ - S^+))$$

The cascaded classifier adopts a degenerate decision tree structure, which concatenates several strong classifiers from the AdaBoost algorithm. The strong classifier itself has a high detection rate and a suitable error rate, to ensure that it can pass most of the face samples while excluding non-face samples. When the cascaded classifier is applied to face detection when it is judged as a face, it will be detected in the next strong classifier; if it is non-face, it does not need to be identified. The stronger the classifier, the higher the detection rate and false detection rate.

4. Face detection based on AdaBoost algorithm and skin color feature

The face detection method based on skin color features is equivalent to the face detection method based on AdaBoost algorithm. The former has a better detection rate in the face side and face expression change, but it is easy to receive the influence of light and complex environment, resulting in the phenomenon of false detection [13]; while the AdaBoost algorithm based face detection method has a smaller impact on the background than the skin color based face detection method, but it is easier to receive the shadow of face pose change Ring, thus affecting the detection rate. If using a single way of a face detection system, then the system in the detection rate and false detection rate is not very good. So in this paper, considering the advantages of the above two methods, the two approaches are combined and improved. The main idea is based on the skin color feature of face detection as a preface detection, get the candidate area containing the skin color of the face, and then use the cascaded classifier to detect the skin color area. The specific flow chart is as follows:
5. Experimental results and analysis
This article mainly uses the face detection method based on the combination of skin color features and AdaBoost algorithm to simulate and analyze the face color image. In this experiment, four aspects of a single face, multi-face, simple background, and sophisticated image, are tested, respectively. Some of the test results are as follows:

- (1) Original image
- (2) Skin color detection
- (3) AdaBoost detection
- (4) Detection in this paper

Figure 7. Algorithm comparison of single face
Figure 8. Algorithm comparison of multi face

Through the analysis of the experimental results, the preliminary results of each algorithm are statistically sorted out, and the statistical results are shown in the table below:

| Test method                  | Skin color detection | AdaBoost detection | Article method |
|-----------------------------|----------------------|--------------------|---------------|
| Total number of faces       | 600                  | 600                | 600           |
| detected                    |                      |                    |               |
| Number of faces detected    | 537                  | 513                | 565           |
| Number of undetected faces  | 89                   | 41                 | 32            |
| Detection rate(%)           | 89.5                 | 85.5               | 94.2          |
| Noise factor(%)             | 14.8                 | 6.8                | 5.3           |

Through the experimental analysis and the analysis of the result table, it can be seen that the detection rate of the face detection method based on skin color features is higher and faster. Still, its false detection rate is relatively reduced, because the algorithm is vulnerable to the influence of non-face skin color area and light. The AdaBoost algorithm is just the opposite of the skin color based detection method, which is mainly because the algorithm is easy to be affected by face pose and complex background. The improved algorithm based on the combination of the two algorithms proposed in this paper is better than the two single algorithms in detection rate and false detection rate. Based on ensuring the detection rate, it has a lower false detection rate, and to a certain extent, it also meets the speed of face detection.
Acknowledgments
This work was supported by the National Natural Science Foundation (Grant No. 11603004), Beijing Natural Science Foundation (Grant No. 1173010), Beijing Education Commission Project (Grant No. KM201710015004) and Beijing Institute of Graphic Communication Teachers Team Building Project (Grant No. 27170118003/006).

References
[1] Xu Xiaoyan. Overview of face recognition technology [J]. Electronic testing, 2015 (10): 30-35 + 45 J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73
[2] A. Craw, H. Ellis and J. Lishman, “Automatic extraction of face features.” Pattern Recognition Letters, vol. 5, pp. 183-187, 1987
[3] Yuan Chunlan, Xiong Zonglong, Zhou Xuehua, Peng Xiaohui. Image edge detection based on Sobel operator [J]. Laser and infrared, 2009,39 (01): 85-87
[4] Yang G Z, Huang T S. Human face detection in a complex background[J]. Pattern Recognition,1994
[5] Crowley, J.L, Berard, F. Multi-modal tracking of faces for video communications[P]. Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on,1997
[6] Yan He, Wang Peng, Yang Xiaolong, Dong Yingyan, Luo Cheng, Li Huan. A new ADA boost face detection method [J]. Journal of Chongqing University of Technology (NATURAL SCIENCE), 2019,33 (02): 106-110 + 161
[7] T.S. Jebara and A. Pentland, “Parameterized structure from motion fro 3D adaptive feedback tracking of faces,” Proc. Sixth IEEE Int’ l Conf. Computer Vision, pp. 128-135, 1998
[8] Wang Jian, Research on Face Detection Based on Gentle AdaBoost Algorithms [D]. University of Electronic Science and Technology, 2011:10-14
[9] Freund.Y.Boosting a Weak Learning Algorithm By Majority. Information and Computation,1995.121(2):256-285
[10] Yu Shiqi, Liu Ruiqi, translated. Gary Bradski, Adrian Kaebler. Learning OpenCv [M]. Beijing: Tsinghua University Press, 2009: 8-28
[11] Li Jingwei, Wang Aixia, Zhang Guangyuan, et al. Sergios Theodoridis, Konstantinos Koutroumbas. Pattern Recognition [M]. Third Edition. Beijing: Publishing House of Electronics Industry, 2006: 1-4
[12] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features[C]// IEEE Computer Society Conference on Computer Vision & Pattern Recognition. 2003
[13] Yuan Haifeng. Research on face detection based on CMYK-H-CbCr skin color segmentation and improved AdaBoost algorithm [D]. Nanjing University of Posts and Telecommunications, 2017.