Acta Technica 62 (2017), No. 6B, 497–508

Face detection based on feature similarity clustering and random forest regression ensemble learning algorithm

Xin Xiong[1], Kefeng Li[1]

Abstract. A comprehensive color model (CCM) extraction method is proposed for the shortcomings such as poor extraction effect of the common color model extraction method and less removal of the background. The Adaboost classifier algorithm is used for face detection. This method provides better effect of complex background removal, reduces the Haar-like feature number of the classifier and accelerates the detection rate by adjusting the amplification factor of the search window. The experimental results show that this method effectively improves the detection ratio, reduce the detection time, provides high adaptability and has higher practicality in face detection.

Key words. Face detection, Feature extraction, Face recognition, Color model, Classifier.

1. Introduction

Face detection is a key link in the face recognition system. With the development of e-commerce, the potential of face recognition as a means of biological status authentication gradually stands out, leading to higher and higher requirements for the performance of the recognition system and harsher requirements for the adaptability of the complex background environment. As the preliminary stage of all face researches, the quality of face detection performance has a direct impact on the performance of the whole face image application system, which enables it to be in a very importance position.

Exiting face detection methods consist of two types: knowledge-based method and learning-based method [1]. The AdaBoost algorithm in the learning-base method

[1]School of Electrical Information Engineering, Henan University of Engineering, Zhengzhou Henan, 451191, China

is an algorithm providing fact face detection. It self-adaptively adjusts the error rate of the hypothesis according to the feedback of weak learning. It improves the accuracy of detection while maintaining the detection efficiency. This method makes an outstanding contribution that it provides a stable and real-time object detection framework, making it the first real-time face detection algorithm.

Despite the fact that this algorithm has advantages in certain aspects over other ones, the slow detection rate and high false detection ratio limit its practical application. The present paper extracts the face color-like area in the image and removes the complex non-color background with the comprehensive color mode (CCM) extraction method, reducing the search area and number of feature and improving the detection rate and accuracy in the course of face detection.

2. Method of color extraction

Color is an important feature of the face. The strong stability and easy distinguishing from the background color are irrelevant to the size and the direction of it in the background. Various color differences are more affected by the brightness but less by the chromaticity [2]. For color extraction, a color model is generally established in the chroma space at first. Where it is a color pixel (or not) will be determined according to the fact that whether the pixel is on the relevant frequency band in the chroma space.

The present paper uses YCbCr space as the mapping space of color. Such space has the construction principle similar to the visual perception process of human: color points achieve very good aggregation in the space; it achieves separation and independent processing of brightness and chromaticity; and it can better restrict the color distribution area [3].

Literature [4] roughly classifies multiple color models into two types: parameter model and non-parameter mode. The former mainly includes Gaussian, Gaussian mixture and elliptic boundary color models, while the latter mainly includes Bayes classifier and statistical histogram models. Common color models are as follows:

2.1. Simple chroma space model

In the color space, the image to be detected is segmented according to the characteristics of distribution of the pixels in the chromaticity components and with the artificial given threshold.

Such algorithm achieves both simplicity and fast speed. However, this model is for the simple skin color detection only due to the fact it distinguishes color and non-color areas with simple thresholds and the image to be detected has poor robustness as a result of the effect of external conditions such as light, shadow, etc.

2.2. Gaussian model

Color distribution may be normal. The distribution parameter can be predicated or the mean and square of the components in the chroma space can be calculated with analysis and statistics of the single Gaussian approximation of skin color distribution. In doing so, the probability at which the pixel belongs t color will be calculated. The Gaussian model better represents color distribution and provides much higher skin color detection ratio and easies computation of the parameters compared to the simple color model; however, the rate is slower than that of the simple chroma space model.

The similarity face image obtained by the traditional YCbCr space-based Gaussian color model without the support of Y channel information cannot reflect the gray feature of the face. Literature [5] proposes a face color model that combines Gaussian color model and gray distribution. With the use of the self-adaptive updating and selection method in the color model parameters, it achieves improved robustness subject to changes as a result of the light.

2.3. Gaussian mixture model

To obtain a more accurate color model to represent the distribution color, the strength of Gaussian mixture model can be used for such improvement. It can be provided in any complex nature. With an increase of the model number, it will arbitrarily approximate to any continuous probability density distribution. In contrast, it will be more accurate if Gaussian mixture model is used. However, it is rarely in fast face detection as it provides the parameter that is difficult to be estimated and low rate.

2.4. Elliptic boundary color model

In YCbCr chroma space, the chromatic values Cb and Cr are always linearly dependent on the brightness value Y. Such dependence greatly affects face detection. For this reason, Hsu came up with a non-linear transformation method [6] to remove the dependence of chromaticity on brightness. He made statistics of the color points of the image in HHI image library. Fig. 1 is the diagram of distribution of the color points in $Cb'Cr'$ space after non-linear transformation.

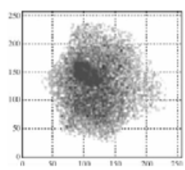


Fig. 1. Distribution of color point in $Cb'Cr'$ space

After transformation, the distribution of the pixels in the space is ellipse-like. The elliptic model may be used to describe its distribution as follows:

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1 \,.$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} Cb' & -c_x \\ Cr' & -c_y \end{bmatrix} .$$

Appropriate values of parameters in the formula can be determined according to the shape of distribution of color pixels in such space. After that, the pixels in the image can be determined to determine whether they are color pixels.

3. Adaboost face detection algorithm

In 1990, Schapire proved that when the samples were sufficient and adequate, the weak learning algorithm would be improved to be the strong learning algorithm so as to improve the accuracy, while it was necessary to search a strong learning algorithm that was unavailable, which developed the initial Boosting algorithm. Such algorithm requires pre-definition of the upper limit of weak learning algorithm error rate, which is hard to be achieved in practices, and it sets limits on the development of Boosting algorithm [7].

In 1995, Freund and Schapire proposed the Adaboost algorithm whose self-adaptability to continuously add weak classifiers. The strong classifier provided arbitrarily small error rate in the training set of the strong classifier obtained, which explored deep into the capability of weak classifier algorithm. This algorithm breaks the limitations of Boosting algorithm and provides improved complexity and practicality, making it widely researched and used by the researchers.

Viola and Jones proposed AdaBoost learning algorithm-based object detection method in 2001 [8]. Such method achieves the detection rate of 15 frames/second while obtaining high detection ratio, enabling the face research to be applied to practices. The key points are as follows:

(1) Haar-like feature is used in the detection.

(2) The integral image is used to calculate the value of Haar-like feature for acceleration purpose.

(3) The AdaBoost algorithm training area is used to distinguish the face and the non-face strong classifiers.

(4) The filter type cascade is used to connect strong classifier cascade for improved accuracy.

The core of Adaboost classifier is that it self-adaptively improves several weak classifiers to a strong classifier, during which the number of weak classifiers is continuously increased to achieve the goal of reduced final training error. The purpose is to concentrate the attention on key data by changing data distribution and to remove the unnecessary training data features by changing data distribution in order to reduce training data. The training process is shown in Fig. 2:

The focus of improvement of Adaboost algorithm is on algorithm convergence, feature extraction, weight updating and classifier relevance [3].

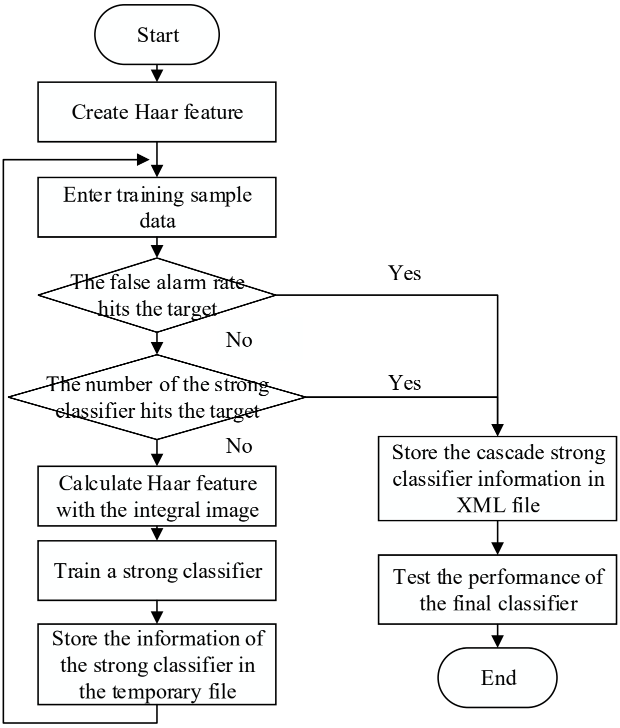


Fig. 2. Training flow chart of adaboost algorithm

4. Face detection by a combination of face color-like extraction and Adaboost classifier

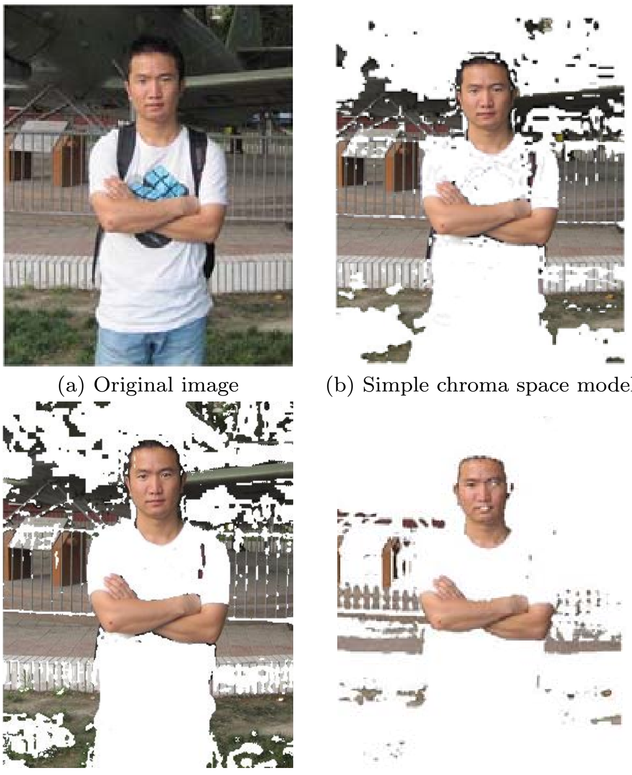
4.1. Principle introduction

The general procedures of face detection by color are that: the image in a color space is directly extracted from the color area and transformed into a binary image whose key part is highlighted by pre-processing such as denoising, dilation, corrosion, etc.; the false area is excluded by limitations including the area size, object aspect ratio to obtain the face area. Skin color detection can detect most of the faces in the image with high accuracy. However, it also detects the non-face color-like area with high false detection ratio.

The problem of Adaboost classifier is the vast number of features, long training time and low detection efficinecy. Researchers have thought of combining both detection methods. It means that the color model is first segmented from the color area including the skin area of other body parts and the background area similar to skin in the image, after which these areas are detected with the well trained AdaBoost cascade classifier as the input images to further remove the non-face area and achieve more accurate face positioning.

4.2. Improvement ideas

There is always deviation in the effect, such as poor effect of non-color background removal, mistaken deletion of darker or brighter color area, etc., although the non-color background can be removed with a single color model. Fig. 3 lists the processing effects of several single color models. Figures (b) and (c) clearly show the weak points of less background removal, while Figure (d) mistakenly deletes some color areas of ear and mouth although it provides good background removal effect. It should be noted that that in fact, the part mistakenly deleted in the image should be filled with black pixels for coloring is for easy observation of the effect.



(a) Original image (b) Simple chroma space model

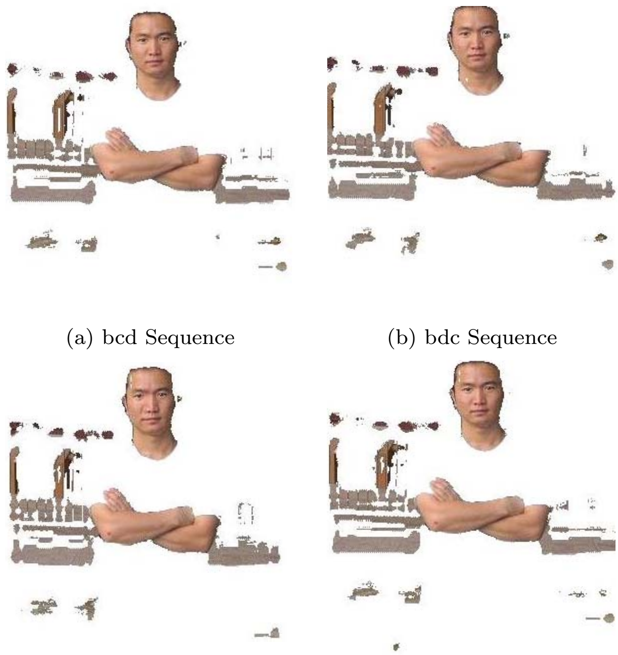
(c) Gaussian model (d) Elliptic model

Fig. 3. Extraction effect of a single model

Given these weak points in the single color model extraction method, the present paper proposes the use of CCM extraction method where the color is extracted with three different color models in turn, and the false area is excluded with the processing method of face detection by color in order to obtain a more accurate face-like area. The processing results as the input images are detected by the well trained AdaBoost classifier to further remove the non-face area and achieve more accurate face positioning.

4.3. Test analysis

The present paper compares the combination patterns in different sequences with the test comparison method and observes the test results. Mistaken deletion has existed at the elliptic extraction method is only used. Such method is not considered to be put in the first place when the comprehensive extraction sequence is combined. The rendering is shown in Fig. 4, where b denotes the simple chroma space model; c denotes Gaussian model; and d denotes the elliptic model.



(a) bcd Sequence (b) bdc Sequence

(c) cbd Sequence (d) cdb Sequence

Fig. 4. Rendering of the comprehensive extraction model in different sequences

With a comparison of Figures 3 and 4 and judgment from the background deletion and mistaken deletion, CCM extraction method provides significantly better extraction effect than that of a single extraction.

It is shown in Fig. 4 that there is little difference between the effects of CCM extraction method in various combination sequences, which is comprehensively evaluated by observation of the non-white pixels contained in the mistaken deletion area and the image. The non-white pixels contained in the images are as follows: (a) 4427, (b) 45135, (c) 44198, (d) 45586. With a careful observation, there is mistaken deletion in the forehead part in image (c). For this reason, the skin color extraction method in bcd sequence used in image (a) is used as the color extraction in this design. The design flow chart is shown in Fig. 5.

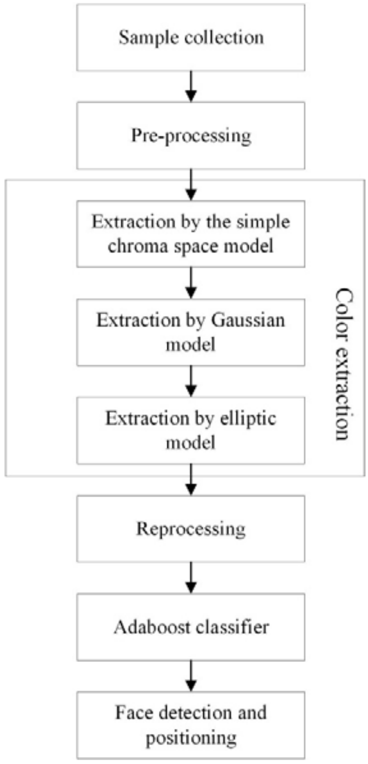


Fig. 5. Design flow chart

The Effect of reprocessing is shown in Fig. 6.

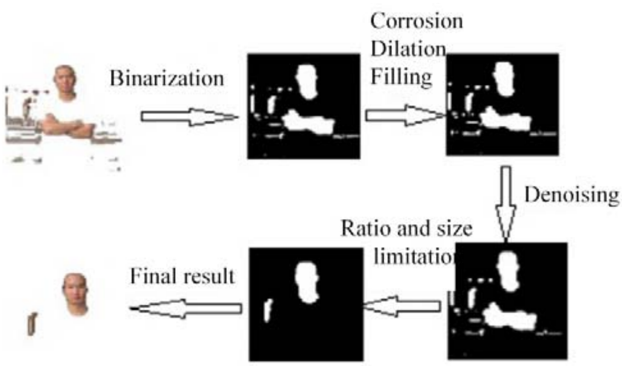


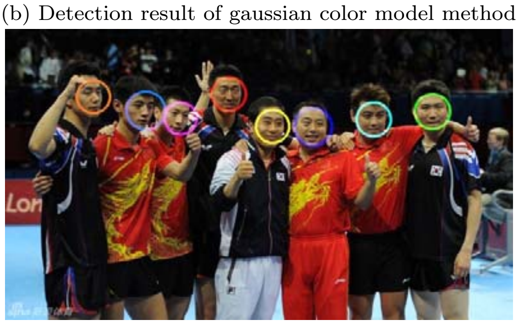
Fig. 6. Reprocessing rendering

5. Result analysis

To validate the effectiveness of the method referred to in the present paper, the author compared the skin color areas extracted by face detection by skin color, Gaussian skin color model extraction and CCM extraction and detected the face with Adaboost algorithm. The author also made statistics of the detection indicators of the detection results. A part of the tested images is the photos of daily life and parts are downloaded from the net. There are 80 images in total, including single-face photos and 50 multiple-face photos, contributing to 285 faces in total. Partial detection results are shown in Fig. 7. Table 1 lists the indicators of the detection.

(a) Detection result of skin color detection method

(b) Detection result of gaussian color model method



(c) CCM detection result

Fig. 7. Reprocessing rendering

Table 1. Indicator comparison of detection results

Detection algorithm	Total number of face (no.)	Number of detection (no.)	Detection ratio (%)	Number of false detection (no.)	False detection ratio (%)	Time (s)
Skin color detection	288	270	94.1	21	7.29	4.634
Gaussian model	288	266	92.4	13	4.51	7.681
CCM	288	278	96.5	8	2.78	5.412

As shown in the experimental results above, CCM extraction method provides better detection effect than that of the rest two methods, whether for a single-face photo or a multiple-face photo. In addition, CCM extraction method removes more background. In the event of face detection with Adaboost algorithm, the amplification ratio of the search box can be properly increased to reduce the detection time. With a comparison of the operation time, when the images processed by the models referred to in this paper are used in face detection, the detection time gains a decrease by 29.5% than the processing result of Gaussian model, which meets the speediness requirement.

6. Conclusion

The algorithm referred to in the present paper is mainly for skin color extraction and processing of an image and maximal removal of non-skin color part and other disturbances to reduce the complexity of face detection by Adaboost algorithm. The experimental results not only improve the detection ratio but also reduce the detection time, achieving the intended purpose. However, the author found in the test that given the different distribution of various image pixels, if the parameters are improperly selected upon skin color model extraction, missing detection will occur as it is easy to delete part of the face. Hence, the next work should focus on designing adaptive parameters. In addition, Adaboost algorithm provides much redundancy. Parameters in its decision tree remain to be improved. Further studies are required for much improved detection performance.

References

[1] Elyan E, Gaber M M: (2015) *A fine-grained Random Forests using class decomposition: an application to medical diagnosis*[J]. Neural Computing & Applications, 2015:1-10.

[2] Peters J, Verhoest N E C, Samson R, et al.: (2009) *Uncertainty propagation in vegetation distribution models based on ensemble classifiers*[J]. Ecological Modelling, 220(6):791-804.

[3] Laradji I H, Alshayeb M, Ghouti L: (2015) *Software defect prediction using ensemble learning on selected features*[J]. Information & Software Technology, 58:388-402.

[4] Dong Y, Du B, Zhang L: (2015) *Target Detection Based on Random Forest Metric Learning*[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 8(4):1830-1838.

[5] Liu Y, Dawant B M: (2015) *Automatic localization of the anterior commissure, posterior commissure, and midsagittal plane in MRI scans using regression forests.*[J]. IEEE Journal of Biomedical and Health Informatics, 19(4):1362-1374.

[6] Vladimir Svetnik, Andy Liaw, Christopher Tong, et al.: (2016) *Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling*[J]. Journal of Chemical Information & Computer Sciences, 43(6):1947-1958.

[7] Li S, Harner E J, Adjeroh D A: (2011) *Random KNN feature selection - a fast and stable alternative to Random Forests*[J]. BMC Bioinformatics, 12(1):1-11.

[8] Azadeh A, Saberi M, Anvari M, et al.: (2011) *An adaptive network based fuzzy inference system-genetic algorithm clustering ensemble algorithm for performance assessment and improvement of conventional power plants*[J]. Expert Systems with Applications, 38(3):2224-2234.

[9] Jo T, Cheng J: (2014) *Improving protein fold recognition by random forest*[J]. BMC Bioinformatics, 15(11):1-7.

[10] Shaheryar M, Khalid M, Qamar A M: (2013) *Rot-SiLA: A Novel Ensemble Classification Approach Based on Rotation Forest and Similarity Learning Using Nearest Neighbor Algorithm*[C]// International Conference on Machine Learning and Applications. IEEE Computer Society, 2013:46-51.

[11] Chan C W, Paelinckx D :(2008) *Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery*[J]. Remote Sensing of Environment, 112(6):2999-3011.

[12] Bian S, Wang W: (2007) *On diversity and accuracy of homogeneous and heterogeneous ensembles.*[J]. International Journal of Hybrid Intelligent Systems, 4:103-128.

[13] Saha I, Sarkar J P, Maulik U: (2015) *Ensemble based rough fuzzy clustering for categorical data*[J]. Knowledge-Based Systems, 77:114-127.

[14] Baik J, Ye Q, Zhang L, et al.: (2014) *Automated classification of oral premalignant lesions using image cytometry and Random Forests-based algorithms*[J]. Cellular Oncology, 37(3):193-202.

[15] Wu J, Rehg J M, Mullin M D : (2003) *Learning a Rare Event Detection Cascade by Direct Feature Selection*[J]. 2003:1523-1530.

Received May 7, 2017