Improved AdaBoost Algorithm for Robust Real-Time Multi-face Detection

Xin-Chao Zhao[1], Jia-Zheng Yuan[2*], Hong-Zhe Liu[1], Jian-She Zhou[3]

[1] Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China.
[2] Institute of Computer Technology, Beijing Union University, Beijing 100101, China.
[3] Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing 100101, China.

* Corresponding author: email: jiazheng@buu.edu.cn

Abstract: Face detection is the basis for research topics such as face recognition, facial expression recognition and face attribute analysis and it plays a crucial role in the field of computer vision. Traditional methods are difficult to meet the needs of robust real-time multi face detection because of some influencing factors such as head pose, image scene, illumination condition and so on. In this paper, we introduce an intelligent vision algorithm that is able to detect human face from complex scene and filter out all the non-face but face-like images. The human face is detected in real-time environment using the approach called Adaboost-based Haar-Cascade Classifier, and the real human face detection is improved from single-face detection to multi-face detection. In addition, variable head poses are taken into account, such as pitch, roll, yaw, etc. Furthermore, the real-time experiments proved the effectiveness and robustness of the algorithm for human detection we have proposed.

Key words: Face detection, adaboost algorithm, strong classifier, real-time, multi-pose.

1. Introduction

In recent years, as an important biometric technology, face recognition has attracted much attention in many fields, such as machine learning, image processing, pattern recognition and artificial intelligence, etc. [1]-[6]. In general, face detection is an image processing process. For any given image or video frame, through some of its searching strategies to determine whether it contains any faces or not, and if faces are included, output the position, size and amount of the faces [7]. As the first step as well as the most important step of face recognition, it has become an important research topic in the field of computer vision [8], [9].

In this paper, we focus on one of the significant problems, face detection. We propose a new intelligent vision algorithm that can detect existing multiple faces in real time from various complex scenes robustly. Face detection performance is directly related to its practical application effect in a real environment [10], [11]. For example, in human-computer interaction system, the inaccurate face detection results will deteriorate the effect of human-computer interaction and user experience. Therefore, a robust real-time multi-face detection algorithm with high accuracy is indispensable for the research of face recognition, facial expression recognition, and face attribute analysis, etc. [12], [13].

2   Relate Work

In recent literatures, many experts did research in the intelligent interaction robot. Bernhard Froba et al focused on developing intelligent system of face tracking of mobile robot [16] using Kalman filter [14-15]. Kwang Ho An et al. paied attention to use relatively small amounts of critical rectangle features selected and trained by Adaboost learning algorithm, and they detected the initial position, size and view of a face correctly [17]. Paul Viola et al presented a robust real-time face detection with the integral image method and AdaBoost training algorithm [18]. The X. Xie etc. presented a real-time tracking algorithm on eye feature tracking [19]. Compared with the previous approaches, we have modified and improved the AdaBoost face algorithm [22], [23] for rapidly multi-face detection in the sequence image frames [20-21], and proposed a scheme that is effective and robust for the problems of variation of scene and head poses. This approach not only improves the face detection accuracy, in the meantime, retains the real-time detection speed.

3   Proposed Method

3.1   Original Method

Various algorithms have been proposed for face detection in the real-time application [24] such as robot system, but the deficiencies existed that often limit its application have also been frequently discussed [25].
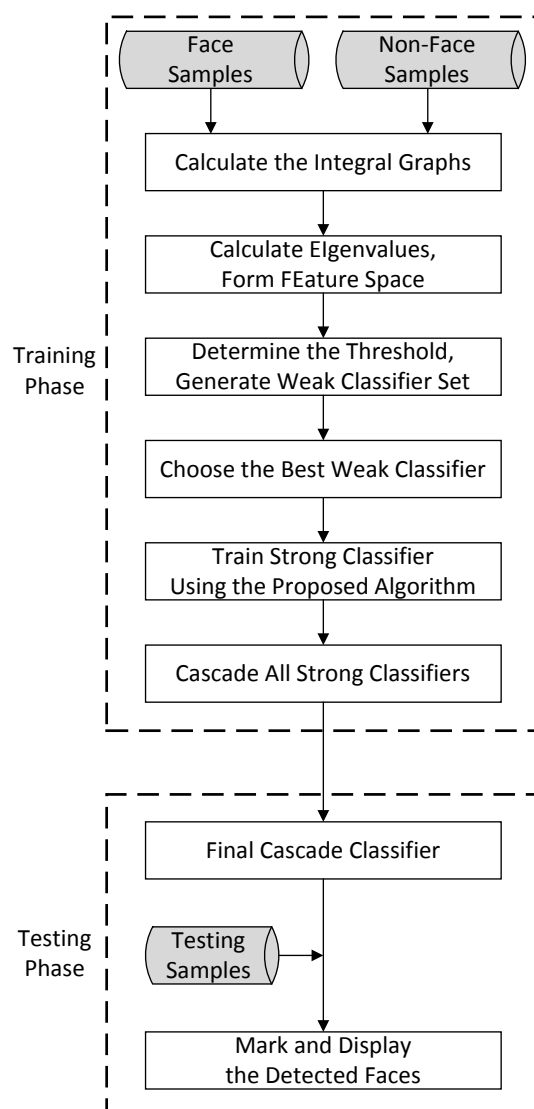

Fig. 1. The whole system flow chart.

- **Scene problem**

Firstly, how the detector responses when a new object enters the field of view scene and he should be detected. In addition, when the surrounding scene similar to the target characters to be detected, with a strong confusion, the detector will do what response at this time.

- **Head pose problem** [26], [27]

In general, there are two possible causes of failure of the detector. One is that the target characters leave the field of view scene, this is a normal situation. While the other is that the head pose of the person is beyond the normal range, but characters themselves are still in the range of the scene, and this is an abnormal case. According to our survey and analysis, the head pose is mainly divided into three categories: pitch (around the X axis direction), yaw (around the Y axis direction), and roll (around the Z axis direction).

Based on the above problems, we proposed the algorithm of face detection to resolve those drawbacks. The whole system flow chart is illustrated schematically in Fig. 1.

3.2 Improved Method

The AdaBoost algorithm is adopted to perform the face detection in image sequences [28]. We denoted it as Global AdaBoost Face Detection algorithm (GAFD), because of its huge time consuming and running loading to the system, and it is executed in low frequency, especially when intelligent robot system running in real-time. The proposed diagram is shown in Fig. 2.

For fast multi-face detection, we improved the AdaBoost algorithm in order to reduce its running time by utilizing the state of the tracked face. We call the algorithm as Local AdaBoost Face Detection algorithm (LAFD). It is presumed that there is single face in the image and its previous state of the tracked face is presented [29]. The tracker tracks and predicts the new state of the face between the sequential image frames using the Kalman Filter [14], [15]. Partial image which is called region of interest (ROI) is obtained for LAFD, according to the prediction of the tracker. The LAFD is launched in high frequency up to real-time. The tracker controller controls the running timing, while GAFD and LAFD maintains the states of tracked faces.
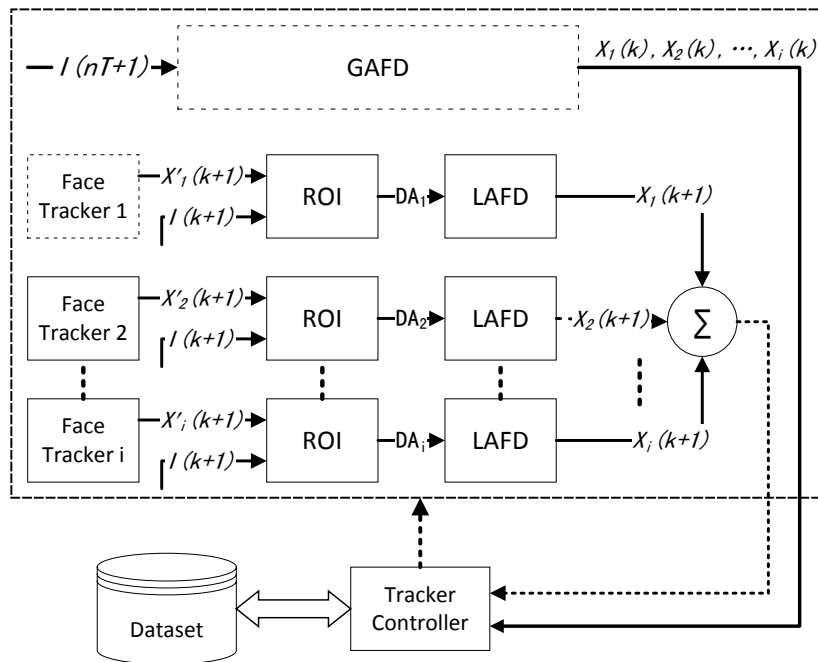

Fig. 2. The algorithm flow chart.

Some improvements are proposed in this paper, aiming at overcoming the defects of the traditional AdaBoost algorithm. The improved algorithm is mainly to provide face relevance, which is used for the ROI selection. This approach reduces the redundant information to minimum which is unnecessary for face detection, among selected ROI. A certain number of samples, labeled positive or negative, are selected as training set, then the algorithm is used for feature selection in ROI. The weak classifiers are boosted into a stronger classifier by the following steps:

(1) The given sample images of $(x_i, y_i)_{i=1}^N$ as input, where N is the number of samples, $y_i = 0$, 1 represent negative and positive samples, respectively.

(2) Initialize the weights $\omega_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0$, 1, where m and l are the number of the negative and positive samples, respectively.

(3) Normalize the weights ω_t for t = 1, $\cdots$, T, so that it is a probability distribution.

$$\omega_{t,i} = \frac{\omega_{t,i}}{\sum_{j=1}^t \omega_{t,j}}$$

(4) To optimize the weight error:

$$\varepsilon_t = \min_{f,p,\theta} \sum_i \omega_i |h(x_i, f, p, \theta) - y_i|$$

where a weak classifier $h(x, f, p, \theta)$ consists of a feature f, a threshold θ and a polarity p indicates the direction of the inequality:

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

where x is a 24 × 24 pixel sub-window of an image.

(5) Then select the best classifier with respect to the weighted error:

$$h_t(x) = h(x, f_t, p_t, \theta_t)$$

where f_t, p_t and θ_t are the minimizers of ε_t.

(6) Defining $\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$, and update the weights:

$$\omega_{t+1,i} = \omega_{t,i} \beta_t^{1-e_i}$$

where $e_i = 0$ if sample x_i is classified correctly and $e_i = 1$ otherwise.

(7) By defining $\alpha_t = \log \frac{1}{\beta_t}$, the final strong classifier is:

$$C(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_n h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, & other \end{cases}$$

Through the above steps, we can get an optimal stronger classifier, which can be used to detect human face from complex scene and filter out all the non-face but face-like images.

4 Experimental Result

To evaluate the efficiency and accuracy of the algorithm proposed in this paper, we used two dataset to carry out our experiments. One is MIT+CMU frontal face dataset and CMU profile face dataset that is the same as Viola and Jones [30], and the other is our training set. The approach is tested using 6500 560×420 pixel color images, which were collected from photos, videos on the Internet. One or more faces are contained in each image, with different lighting conditions and complex scenes. The faces are different in size, pose, location and facial expression. Specifically, most of the faces are multi-pose faces which include rotated frontal and profile faces.

We experimentally verified the algorithm on a testing host of Intel Core 2 Dual 2.8 GHz CPU and 4G RAM.

Fig. 3 shows a face detection result for single face image. We only marked the face regions detected using a rectangle bounding box as small as possible. Fig. 4 shows a result for multi-face detection in the images with complex and confusing background and different expressions of targets. Fig. 5 shows a result for multi-face detection. The pose, orientation, position and size of the head of the people in the candidate images are various and different.
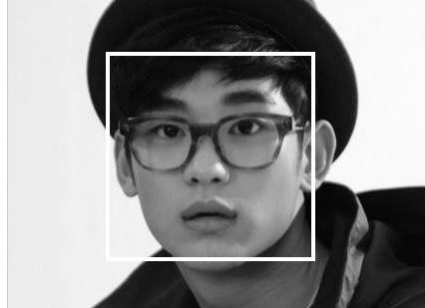

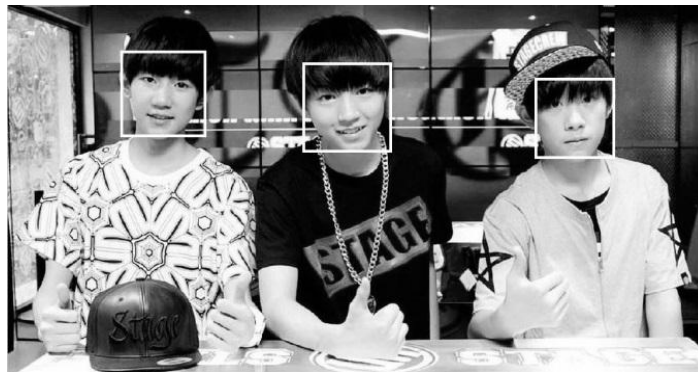Fig. 3. Face detection result for single face image.

Fig. 4. Face detection result for multi-face image with confusing scene.

Fig. 5. Face detection result for multi-face image with varied pose.

5   Experimental Analysis

We compared the performance of our method with the previous work on the MIT+CMU dataset and our own dataset, and the results are shown in Table 1. The results of various face detection experiments indicated that our algorithm can achieve a higher correct rate (up to 96.3%), about 10 percentage points higher than the traditional Adaboost method [22], and about 7 percentage points higher than the improved

Adaboost method [31]. Furthermore, the average speed of detection is also faster than the above two about 25ms and 22ms, respectively.

The reason that our algorithm can improve the accuracy of detection, while maintaining a relatively low false detection rate is that we have considered the effects of a variety of complex scenes and variable head poses on the final detection results when designing algorithms and processing samples, which can improve the accuracy and robustness of the algorithm. Therefore, our results are relatively better than others. And also, our method shows a better detection rate in the case that more precise detection of facial features are required.

Table 1. Comparison Results of Face Detection

Detection Algorithm	Testing Dataset	Missing Rate(%)	Correct Rate(%)	Average Speed(ms)
Viola Jones	MIT+CMU	31.3	72.3	121
	Our Dataset	35.6	68.2	135
[22] Adaboost	MIT+CMU	21.8	87.5	76
	Our Dataset	24.1	84.6	84
[31] Improved Adaboost	MIT+CMU	18.8	89.2	69
	Our Dataset	22.7	87.4	73
Our Method	MIT+CMU	6.9	93.8	51
	Our Dataset	5.2	96.3	48

Furthermore, it performs at a speed of 26 frames/s speed, the frame's size is 560 x 420 pixel, and the configuration of the testing computer is Intel Core 2 Dual 2.8 GHz CPU and 4G RAM, which can provide good inputs for face recognition and facial expression recognition. Therefore, the proposed method has the advantage over others not only in face detection, but also in face recognition and analysis of face properties [32].

6   Conclusion and Future Work

In this paper we implemented a novel methodology for robust multi-face detection in image sequences, intended for HRI in IER applications. Experimental results have confirmed the effectiveness and the increased computational efficiency of the proposed approach and proved that the characteristic advantages are maintained, leading to merits that combine efficiency, accuracy and robustness simultaneously.

We intend to make full use of the proposed approach to support natural interaction with autonomously navigating robots that guide visitors in exhibition centers and museums. In particular, the proposed method will provide input for the recognition and analysis of facial expressions that people would utilize when engaged in various conversational states.

An interesting problem is whether we can use fewer training samples in our algorithm to achieve comparable detection results. In general, the tracking performance does not degrade (if there are no distractions). However, for some challenging cases, the detector succeeded, but drifted in others due to cluttered background. This implies that training samples should be selected more carefully when fewer samples are used. Therefore, problem about how to select representative training samples is nontrivial and may be our future work.

Moreover, the future work includes extension of the face detection algorithm to handle stereo vision by exploiting epipolar constraints. so that our approach will be widely employed in an integrated system for naturalistic human-robot interaction.

Acknowledgement

References

[1] Zhu, J., & Chen, Z. (2016). Real time face detection system using adaboost and haar-like features. *Proceedings of the 2016 International Conference on Information Science & Control Engineering*.

[2] Sun, S., Xu, Z., Wang, X., & Huang, G. (2015). Real-time vehicle detection using haar-SURF mixed features and gentle adaboost classifier. *Proceedings of the 2016 The 27th Chinese Control and Decision Conference*.

[3] Zhu, J., & Chen, Z. (2015). Real time face detection system using adaboost and haar-like features. *Proceedings of the International Conference on Information Science and Control Engineering*.

[4] Burcu, K. S., Sümeyya, İ., & Yaşar, B. (2016). The realization of face detection and fullness detection in medium by using haar cascade classifiers. *Proceedings of the Signal Processing and Communication Application Conference*.

[5] Erik, H., & Boon, K. L. (2001). Face detection: A survey. *Computer Vision and Image Understanding*..

[6] Yang, M.-H., Kriegman, D. J., & Ahuja, N. (2002). Detecting faces in images: A survey. *IEEE Trans. Pattern Analysis and Machine Intelligence*.

[7] Hjelmas, E., & Low, B. K. (2001). Face detection: A survey. *Computer Vision and Image Understanding*.

[8] Ge, S. S., & Fua. C. H. (2005). Queues and artificial potential trenches for multi-robot formations. *IEEE Transactions on Robotics*, *21(4)*, 646-656.

[9] Ge, S. S. (2007). Social robotics: Integrating advances in engineering and computer science. *Computer Telecommunications & Information Technology*. Mae Fah Luang University Chang Rai.

[10] Darijan, M., Tomislav, H., & Slobodan, R. (2016). Two-stage cascade model for unconstrained face detection. *First International Workshop on Sensing, Processing and Learning for Intelligent Machines*.

[11] Mykoniatis, K., Angelopoulou, A., Akbas, A. S., & Hancock, P. A. (2016). Multi-method modeling and simulation of a face detection robotic system. *Proceedings of the Annual IEEE Systems Conference (SysCon)*.

[12] Meyer, G. P., Alfano, S., & Do, M. N. (2016). Improving face detection with depth. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*.

[13] Dahal, B., Alsadoon, A., Prasad, P. W. C., & Elchouemi, A. (2016). Incorporating skin color for improved face detection and tracking system. *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*.

[14] Kalman, R. E. (2015). A new approach to linear filtering and prediction problems.

[15] Greg, W., & Gary, B. (1995). An introduction to the kalman filter. University of North Carolina at Chapel Hill.

[16] Bernhard, F., & Christian, K. (2004). Face tracking by means of continuous detection. *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*.

[17] Kwang, H. A., Dong, H. Y., Sung, U. J., & Myung, J. C. (2005). Robust multi-view face tracking. *Proceeding of the 2005 IEEE/IRSJ International Conference on Intelligent Robots and Systems*.

[18] Viola, P., & Jones, M. (2001). Robust real-time object detection. *International Journal of Computer Vision*, *57(2)*, 87-89.

[19] Xie, X. D., Sudhakar, R., & Hanqi, Z. (1995). Real-time eye feature tracking from a video image sequence using Kalman filter. *IEEE Transactions on Systems, Man and Cybernetics*, *25(12)*, 1568-1577.

[20] Aniruddha, D. (2016). A contour based procedure for face detection and tracking from video. *Proceeding of the International Conference on Recent Advances in Information Technology*.

[21] Pranti, D., & Nachamai, M. (2016). Detection of faces from video files with different file formats. *Proceeding of the International Conference on Microelectronics, Computing and Communications*.

[22] Wu, B., Ai, H., Huang, C., & Lao, S. (2004). Fast rotation invariant multi-view face detection based on real adaboost. *Proceeding of the IEEE International Conference on Automatic Face & Gesture Recognition*.

[23] Abualkibash, M., Mahmood, A., & Moslehpour, S. (2015). A near real-time, parallel and distributed adaptive object detection and retraining framework based on AdaBoost algorithm. *Proceeding of the High Performance Extreme Computing Conference*.

[24] Abualkibash, M., Mahmood, A., & Moslehpour, S. (2015). A near real-time, parallel and distributed adaptive object detection and retraining framework based on AdaBoost algorithm. *Proceeding of the High Performance Extreme Computing Conference*.

[25] Cheney, J., Klein, B., Jain, A. K., & Klare, B. F. (2015). Unconstrained face detection: State of the art baseline and challenges. *Proceeding of the International Conference on Biometrics*.

[26] Wei, L. L., & Liu, M. Y. (2016). Multi-pose face detection research based on adaboost. *Proceeding of the 2016 Eighth International Conference on Measuring Technology and Mechatronics Automation*.

[27] Wei, L., & Liu, M. (2016). Multi-pose face detection research based on adaboost. *Proceeding of the Eighth International Conference on Measuring Technology and Mechatronics Automation*.

[28] Blais, F. (2004). Review of 20 years of ranges sensor development. *SPIE - The International Society for Optics and Photonics*.

[29] Bernhard, F., & Christian, K. (2004). Face tracking by means of continuous detection. *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*.

[30] Putro, M. D., Adji, T. B., & Winduratna, B. (2015). Adult image classifiers based on face detection using Viola-Jones method. *Proceeding of the International Conference on Wireless and Telematics*.

[31] Ma, S., & Du, T. (2010). Improved adaboost face detection. *Proceeding of the International Conference on Measuring Technology & Mechatronics Automation*.

[32] Liu, W., Lv, J., Yu, B., & Shang, W. (2015). Multi-type road marking recognition using adaboost detection and extreme learning machine classification. *Proceeding of the 2015 IEEE Intelligent Vehicles Symposium (IV)*.

Xin-Chao Zhao is currently a master of Beijing Key Laboratory of Information Service Engineering, Beijing Union University. His research interests are in the areas of digital image processing, deep learning and pattern recognition.

Jia-Zheng Yuan recived the PhD degree in computer science from Beijing Jiaotong University, China. He is a professor Institute of Computer Technology, Beijing Union University, China. His research interests are in digital image processing, navigation and positioning, digital museum.

Hong-Zhe Liu recived her Ph.D. in computer science from Beijing Jiaotong University in 2012. She is now an associate   professor in Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing, China. Her current research interests include digital museum, semantic calculating and image processing.

Jian-She Zhou received the Ph.D. Degree from Wuhan University in 2002.   He is now a professor in Capital Normal University, Beijing, China. His research interests include linguistics, logic, and language intelligence.