Rotated Haar-Like Features for Face Detection with In-Plane Rotation

Shaoyi Du, Nanning Zheng, Qubo You, Yang Wu, Maojun Yuan, and Jingjun Wu

Institute of Artificial Intelligence and Robotics,
Xi'an jiaotong University, Xi'an, Shaanxi Province 710049, P.R. China
{sydu,nnzheng,qbyou,ywu,mjyuan,jjwu}@aiar.xjtu.edu.cn

Abstract. This paper extends the upright face detection framework proposed by Viola et al. 2001 to handle in-plane rotated faces. These haar-like features work inefficiently on rotated faces, so this paper proposes a new set of $\pm 26.565^{\circ}$ haar-like features which can be calculated quickly to represent the features of rotated faces. Unlike previous face detection techniques in training quantities of samples to build different rotated detectors, with these new features, we address to build different rotated detectors by rotating an upright face detector directly so as to achieve in-plane rotated face detection. This approach is selected because of its computational efficiency, simplicity and training time saving. This proposed method is tested on CMU-MIT rotated test data and yields good results in accuracy and maintains speed advantage.

1 Introduction

In recent decades, face detection techniques have developed very quickly, and the survey of it have been detailed in [1,2]. It could be said without question that the last decade witnessed the most fruitful researches in the field of face detection. First of all, Rowley et al. [3,4] had proposed neural network-based face detection which had greatly improved the accuracy of face detection. Secondly, with great accuracy, Viola and Jones had introduced a fast object detection based on a boosted cascade of haar-like features [5]. Thereafter more importance began to be attached to the research of haar-like features and AdaBoost. Viola had once advanced asymmetric AdaBoost for more robust classification [6]; Lienhart had extended the haar-like features to an efficient set of 45° rotated features and used discrete AdaBoost (DAB), real AdaBoost(RAB) and gentle AdaBoost(GAB) for face detection [7,8].

With the marked improvement of speed and precision in upright face detection, more and more scholars came to focus on the multi-view face detection. FloatBoost presented by Li was used to overcome the monotonicity problem of the sequential AdaBoost learning [9]. Jones had created a fourth type of rectangle filter and applied it to multi-view face detection [10]. Lin addressed MBHboost whose integration with a cascade structure is robust for face detection [11]. Wang had introduced RNDA to handle multi-view face and eye detection [12].

H. Zha et al. (Eds.): VSMM 2006, LNCS 4270, pp. 128–137, 2006.

The above-mentioned techniques had achieved somewhat success, but all of them are based on the framework of retraining samples. That is, those techniques averagely divided a 360-degree in-plane rotated face into 12 different classes, each covering 30 degrees. Within each rotation class, a large number of face samples had to be collected and a corresponding features set had to be designed so that face detectors for each class could be trained to cover the full 360-degree in-plane rotation.

This paper proposed a novel notion that rotated face detectors needn't the retraining of rotated face samples. For the reason of the correspondence of face detectors and their features, rotated face detectors are very likely to be obtained in the way of rotating efficient upright face detector. On the ground of this, with our trained upright face detector and designed set of $\pm 26.565°$ rotated haar-like features, a variety of rotated face detectors can thus be gained to fulfill the task of face detection with 360-degree in-plane rotation. This new proposed approach has been tested in CMU-MIT rotated test set. The test demonstrates our new rotated detectors do not add more computation but can get good results with great accuracy as well. Besides, training time is saved considerably.

This paper is organized as follows. In section 2, the haar-like features have been reviewed, and our rotated haar-like features with corresponding integral images are introduced. In section 3, the framework of in-plane rotated face detection has been presented. Following that is section 4 in which the proposed method is evaluated on CMU-MIT rotated test set and a conclusion is finally drawn in section 5.

2 Features and Its Computation

As we all know, features are important for classification problem. Haar-like features have been widely used in upright face detection for two reasons: 1) Their quantity is much larger than the number of the raw pixels and they can well represent the upright faces. 2) They are simple and can be calculated quickly. Nevertheless, the features only represent the face's horizontal and vertical local features, so they can't be used in rotated faces. For this reason, we have to extend the original haar-like features.

2.1 Previous Work

Haar-like features that can directly represent the local relation between the image regions [5,7,8] perform efficiently in the upright face detector, so they are used for our upright face detection. In this paper, we use the following haar-like features shown in Fig. 1.

The real value of haar-like features can be computed in the way of the pixel sum of the black rectangular region subtracted from that of the white region, and the pixel sum of each rectangular region feature can be obtained by integral images. Fig. 2 illustrates the computation.

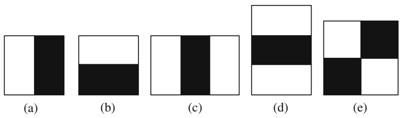

Fig. 1. Haar-like features for upright face detection

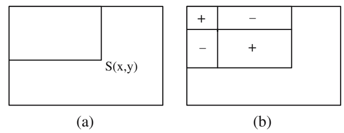

Fig. 2. The computation of rectangle feature by the integral image. (a) the upright integral image (b) calculation scheme of the pixel sum of upright rectangle feature.

The integral image at location (x, y) contains the sum of the pixels above and to the left of (x, y):

$$S(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y') \tag{1}$$

where $I(x', y')$ is the pixel value.

$$R(x, y) = R(x, y-1) + I(x, y) \quad S(x, y) = S(x-1, y) + R(x, y) \tag{2}$$

where $R(x, y)$ is cumulative column sum.

The merit of this computation is simple, fast and accurate. Namely, if the integral image is to be computed, we only need to scan the image once.

2.2 Design of Rotated Degrees

The face rotation is generally classified into two categories: in-plane rotation and out-plane rotation. There are mainly two approaches for in-plane rotation. The first is to rotate images for an upright detector to detect faces. The other approach is to design detectors with different degrees directly to detect faces. Compared with the former one conducting relatively bad in its slow speed, the latter has been adopted more often. Based on it, this paper extends the upright face detector to produce different face detectors with in-plane rotation.

Given that images are continuous signals, theoretically, when an efficient upright face detector is rotated to any arbitrary degree, it is supposed to detect faces in the corresponding position. Yet, the fact is not in that case. Since images are discrete digital signals, the arbitrary rotated features for face detectors are usually unable to be computed fast. Consequently, most researchers retrained samples for different detectors. On the contrary, this paper presents a novel technique which can save time without retraining samples any more.

To design rotated features corresponding to face detectors which needn't the retraining of samples, three rules should be followed:

1. Tilted integral images should be computed fast, through which the rotated haar-like features can be calculated.

2. As the upright face detection can detect in-plane rotated faces of $[-20°, 20°]$ and out-plane faces of $[-15°, 15°]$, to ensure that different rotated face detectors, which are obtained by rotating the upright face detector, can detect faces with

360-degree in-plane rotation, the rotated degree between the neighborhood haar-like features (see Fig. 3) ought to be less than $20° \times 2 = 40°$.

3. The above-mentioned rotated degree between the neighborhood haar-like features should be as large as possible, which can make the divided classes of in-plane rotation as few as they could.

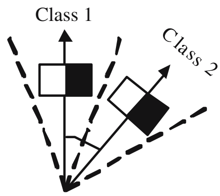

Fig. 3. Rotated degree between neighborhood haar-like features

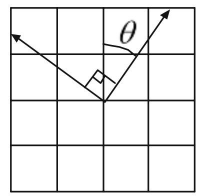

Fig. 4. Design of rotated degree for fast computation of tilted integral images

On the basis of rule 1, we can find only when the degree θ satisfies $\tan\theta = 1/n$, $(n = 1, 2, 3, \cdots)$ can the tilted integral images be computed fast in Fig. 4. But according to rule 2, we can get $n \geq 2$. Because of rule 3, n can only equal 2. Hence, the rotated degrees used in this paper are $\arctan 0.5 = 26.565°$ (clockwise) and $-\arctan 0.5 = -26.565°$ (anti-clockwise).

2.3 Rotated Haar-Like Features

In consideration of the design of $\pm 26.565°$ rotated degree, we divide the rotated faces into 12 new classes: $[-13.283°, 13.283°]$, $[\pm 13.283°, \pm 45°]$, $[\pm 45°, \pm 76.717°]$, $[\pm 76.717°, \pm 103.283°]$, $[\pm 103.283°, \pm 135°]$, $[\pm 135°, \pm 166.717°]$, $[166.717°, 193.2\ 83°]$, and the degrees θ of these classes' detectors are $0°$, $\pm 26.565°$, $\pm 63.435°$, $\pm 90°$, $\pm 116.565°$, $\pm 153.435°$, $180°$ respectively, see Fig. 5. The paper redefines the set of haar-like features with the angles $\pm 26.565°$, which are given in Fig. 6.

For the sake of description convenience, two definitions are given here: within-class haar-like features and between-class haar-like features. Within-class haar-like features are a set of features which are used for one detector to detect faces in one class, such as Fig. 6-1(a), 1(d), 2(a), 2(d), 3(a) for upright detector. Between-class haar-like features are the set of same type of features that can be used in different detectors respectively, so they can be transformed each other easily, such as Fig. 6-1(a), 1(b), 1(c).

It is obvious to see that within-class haar-like features Fig. 6-1(a), 1(d), 2(a), 2(d), 3(a) are used for an upright detector $[-13.283°, 13.283°]$. So are features 6-1(b), 1(e), 2(b), 2(e), 3(b) for $26.565°$ detector with the range $[13.283°, 45°]$ and 6-1(c), 1(f), 2(c), 2(f), 3(c) for $-26.565°$ detector with the range $[-45°, -13.283°]$. Rotating these features $\pm 90°$, $180°$ each, the 360-degree frontal face detectors can be obtained with ease.

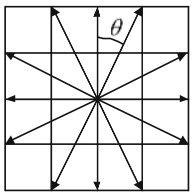

Fig. 5. The distribution of the degrees of 360-degree in-plane rotated detectors

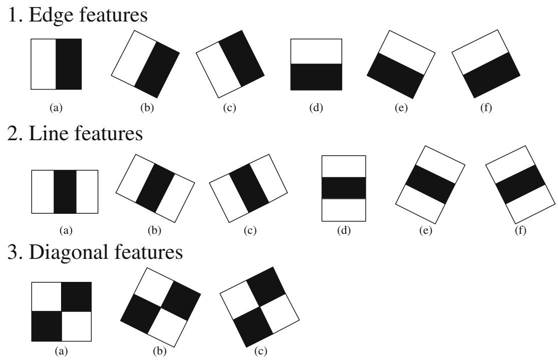

Fig. 6. The set of upright and ±26.565° rotated haar-like features

2.4 Fast Feature Computation

Similar to [7,8], we compute these features by new integral images as well. Firstly, we rotate the basic component, rectangle feature as Fig. 7. The rotated features can't be computed like section 2.1 any longer. To compute the tilted integral images fast, we address the following method.

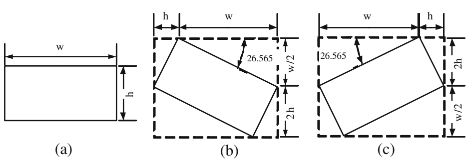

Fig. 7. The Set of rectangle features. (a) upright rectangle feature (b) 26.565° rotated rectangle features (c) −26.565° rotated rectangle features.

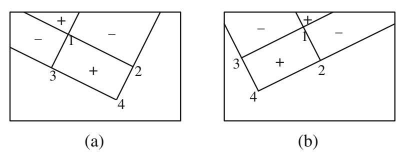

Fig. 8. Computation of rectangle features. (a) the pixel sum of 26.565° rectangle (b) the pixel sum of −26.565° rectangle.

In Fig. 8, the value of every rotated rectangle can be computed by the pixel sum of 4 corners, so our focus is the computation of integral image (see Fig. 9).

In Fig. 9(a), for each pixel point (x, y) , its integral image $S(x, y)$ can be calculated by iteration:

$$\begin{aligned} R(x,y) &= R(x+1,y-2) + I(x,y) \\ S(x,y) &= S(x-2,y-1) + R(x,y) + R(x-1,y-1) + R(x,y-1) \\ &\quad + R(x-1,y-2) + R(x,y-2) \end{aligned} \tag{3}$$

where $R(x, y)$, which is denoted as the arrow lines in the figure, is the pixel sum of tilted lines, and $I(x, y)$ is the pixel value.

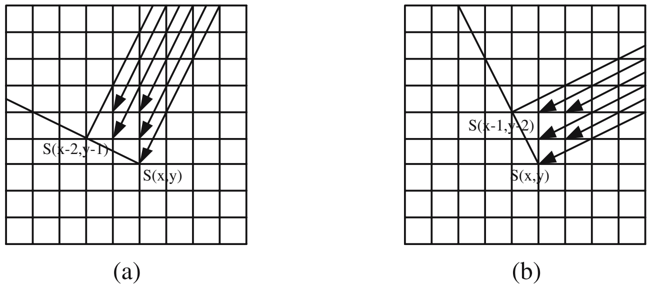

Fig. 9. The calculation schedule of rotated integral images. (a) 26.565° integral image (b) −26.565° integral image.

Fig. 9(b) lists the similar calculation:

$$\begin{aligned} R(x,y) &= R(x+2,y-1) + I(x,y) \\ S(x,y) &= S(x-1,y-2) + R(x,y) + R(x,y-1) + R(x+1,y-1) \\ &\quad + R(x,y-2) + R(x+1,y-2) \end{aligned} \tag{4}$$

With these equations, we can easily compute rotated integral images by only scanning images once, so the speed of the computation of the rotated rectangle features is fast.

3 The Framework of Face Detection with In-Plane Rotation

Based on the rotated haar-like features, the following gives the framework of face detection with in-plane rotation.

3.1 Gentle AdaBoost

Compared with other AdaBoost, Gentle AdaBoost (GAB) algorithm is of simplicity and stability [7,8], so GAB is used in this paper for training upright detector. In each update step, GAB uses the Newton method. The algorithm is listed as follows:

1. Initialize the weights: $d_i^0 = 1/N, i = 1, 2, \cdots, N$
2. Repeat for $t = 1, 2, \cdots, T$
   - A regression function $\hat{h}_t(x)$ can be obtained by minimizing a square loss function $E_d\{(y - h_t(x))^2\}$. This regression function can be calculated in the following way:

   $$\hat{h}_t(x) = P_t(y = 1|x) - P_t(y = -1|x)$$

   - Update the weights: $d_n^t \propto d_n^{t-1} \exp\{-y_i \hat{h}_t(x_i)\}, i = 1, 2, \cdots, N$
3. Output: $f_T(x) = \sum_{t=1}^{T} \hat{h}_t(x)$

3.2 Cascade Classifiers

In the design of detectors, the cascade classifiers, in which each stage is a strong classifier trained by AdaBoost, are always chosen. The cascade detectors can get rid of much background by only using a few features, so the detection can save much time. Its structure is shown in Fig. 10.

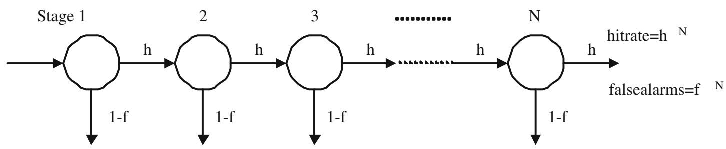

Fig. 10. The design of the detection cascade

Each stage is trained by GAB. Stage one only need a few features that can filter a large number of non-face windows. In the following stages, the number of the features increases very quickly, and the candidate faces are left fewer. After these processes, any efficient technique can be used in the post-process, such as adding the stage numbers or using other detectors.

3.3 The Framework of Face Detection with In-Plane Rotation

In this paper, the in-plane rotated faces are divided into 12 classes like section 2.3. For each class, we can use the new haar-like features to get the corresponding detectors by rotating the upright detector. These new detectors maintain the speed advantage because of the rapid computation of the ±26.565° rotated integral images. Our face detection system does not use any method about view estimation, because this paper focuses on the precision of our proposed method and the speed of each detector, but not the speed of the system which is to be researched in our future work. Fig. 11 illustrates the framework of the face detections in our present system.

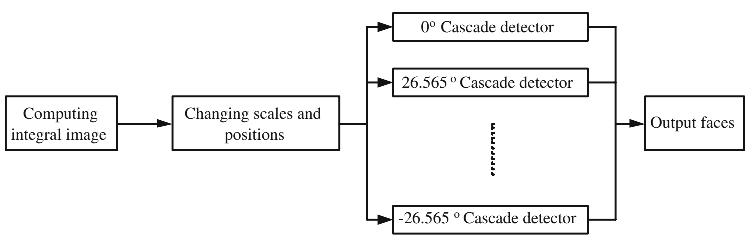

Fig. 11. The framework of face detection with in-plane rotation

4 Experimental Results

The rotated face detection technique is evaluated on CMU-MIT rotated test set [4] which consists of 50 images containing 223 frontal faces with in-plane rotations. We design two experiments to demonstrate our proposed method can get good accuracy. ROC curve detailed in [5] has been given in each experiment.

In the first experiment, the rotation degrees are classified into 12 classes and then 22-layer cascade classifier of the upright face detector is used. To evaluate our method's accuracy, we collect statistics about the number of faces in CMU-MIT for each class and test our method. Most face degrees are in the range [−13.283°, 13.283°], [13.283°, 45°] and [−13.283°, −45°] with 23, 21 and 155 faces respectively, then Fig. 12 gives ROC curves within the three classes.

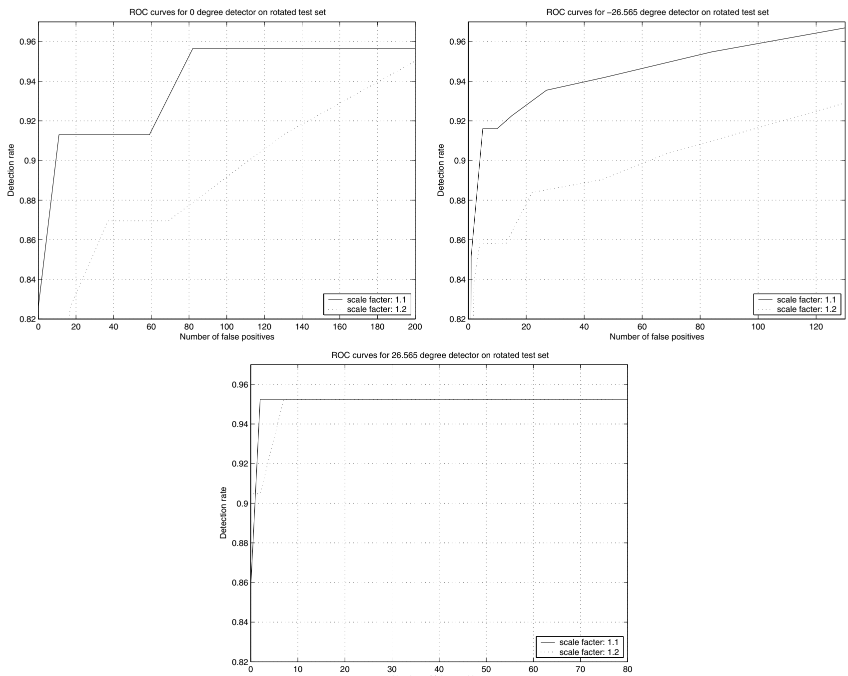

Fig. 12. ROC curve showing the performance of detecting the corresponding classes' faces. (a) ROC curve for [−13.283°, 13.283°] (b) ROC curve for [13.283°, 45°] (c) ROC curve for [−13.283°, −45°].

In the second experiment, to compare with other works on multi-view face detection, we use all detectors to detect the faces in 360°. Fig. 13 gives the ROC curve for a try-all-rotations detector. Rowley et al. had ever reported a detection rate of 89.2% with 221 false positives using a three-stage detector and Jones et al. had a result of 89.7% correct detections with 221 false positives, yet our result

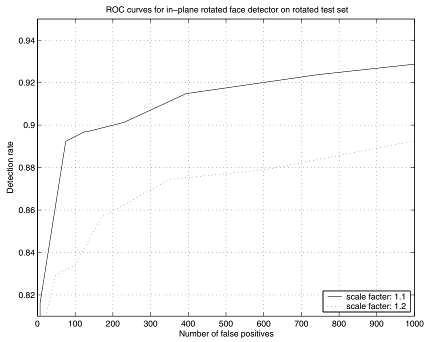

Fig. 13. ROC curve for a try-all-rotations detector on CMU-MIT rotated test set

is 90.1% with 221 false positives with scale factor 1.1 and 86.1% with scale factor 1.2, and the speed of scale factor 1.2 is much faster than that of scale factor 1.1.

The ROC curve of rotated detectors in the first experiment is found much close to that of upright detector, which demonstrates our good accuracy of rotated haar-like features. Moreover, it is in the same case with the time of rotated detectors and that of the upright detector, meaning the rotated detector does not add much computation in detecting process. In the second experiment, with a try-all-rotations detector tested on the CMU-MIT rotated test set, our result works appreciably superior to those of Rowley and Viola.

5 Conclusions

Distinct from the previous face detection techniques, this paper is not constrained with training quantities of samples to build different rotated detectors; instead, through extending upright face detector to different rotated detectors with a new set of haar-like features, face detection with in-plane rotation is achieved. The main contributions of this paper are: 1) The framework proposed by Viola can be extended to in-plane rotated face detection. 2) A new set of rotated haar-like features is described of great use to represent features of rotated faces. 3) A demonstration that without retraining time any more, various rotated face detectors can be obtained with good accuracy. Our future work will focus on view estimation based on using between-class harr-like features to improve the system speed further.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant No. 60021302, and the National High-Tech Research and Development Plan of China under Grant No. 2005AA147060. We thank Jianru Xue for helpful discussions.

References

1. Yang, M.H., Kriegman, D.J., Ahuja, N.: Detecting faces in images: A survey. IEEE Trans. Pattern Analysis and Machine Intelligence, **24** (2002) 34–58.
2. Hjelmas, E., Low, B.K.: Face detection: A survey. Computer Vision and Image Understanding **83** (2001) 236–274.
3. Rowley, H., Baluja, S., Kanade, T.: Neural network-based face detection. In IEEE Trans. Pattern Analysis and Machine Intelligence, **20** (1998) 23–38.
4. Rowley, H., Baluja, S., Kanade, T.: Rotation invariant neural network-based face detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (1998) 38–44.
5. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In proceedings of Conference on Computer Vision and Pattern Recognition, (2001) 511–518.
6. Viola, P., Jones, M.: Fast and robust classification using asymmetric adaBoost and a detector cascade. Advances in Neural Information Processing System 14. MIT Press, Cambridge, MA (2001) 1311–1318.
7. Lienhart, R., Maydt, J.: An extended set of haar-like features for rapid object detection. In Proceedings of the IEEE Conference on Image Processing. New York, USA (2002) 155–162.
8. Lienhart, R., Kuranov, A., Pisarevsky, V.: Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In DAGM 25th Pattern Recognition Symposium (2003) 297–304.
9. Li, S.Z., Zhu, L., Zhang, Z., Blake, A., Zhang, H., Shum, H.: Statistical learning of multi-view face detection. ECCV 2002: 7th European Conference on Computer Vision, **4** (2002) 67–81.
10. Jones, M.J., Viola, P.: Fast multi-view face detection. IEEE Conference on Computer Vision and Pattern Recognition (2003).
11. Lin, Y.Y., Liu, T.L.: Robust face detection with multi-class boosting. Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2005) 680–687.
12. Wang, P., Ji, Q.: Learning discriminant features for multi-view face and eye detection. IEEE Computer Society Conference, **1** (2005) 373–379.