A Hybrid Face Detector based on an Asymmetrical Adaboost Cascade Detector and a Wavelet-Bayesian-Detector

Rodrigo Verschae and Javier Ruiz del Solar
Department of Electrical Engineering, Universidad de Chile.
Email: {rverscha, jruizd}@ing.uchile.cl

Abstract. In this paper is proposed a hybrid face detector that combines the high processing speed of an Asymmetrical Adaboost Cascade Detector with the high detection rate of a Wavelet Bayesian Detector. This integration is achieved by incorporating this last detector in the middle stages of the cascade detector. Results of the application of the proposed detector to a standard face detection database are also presented.

Keywords : Face Detection, Adaboost, Wavelet-based Face Detection.

1. Introduction

Face Analysis (face recognition, face detection, face tracking, facial expression recognition, etc.) is a very lively and expanding research field. The increasing interest in this field is mainly driven by applications like access control to buildings and (computational) systems, identification for law enforcement, borders control, identification and verification for credit cards and ATM, human-machine interfaces, communications, multimedia, and very recently passive recognition of criminals (e.g. terrorists, hooligans) in public places or buildings (airports, train stations, stadiums, etc.).

Face detection is a key step in almost any computational task related with the analysis of faces in digital images. Given an arbitrary image, the goal of a face detection system is to find all contained faces and to determine the exact position and size of the regions containing these faces. When analyzing real-world scenes, face detection is a challenging task, which should be performed robustly and efficiently, regardless variability in scale, location, orientation, pose, illumination and facial expressions, and considering possible object occlusions.

Several approaches have been proposed for the computational detection of faces in digital images. A very comprehensive review can be found in [1][8]. Main approaches can be classified as: (i) feature-based, which uses low-level analysis (color, edges, …), feature analysis or active shape models, and (ii) image-based, which employs linear subspace methods, neural networks or statistical analysis [1]. Image-based approaches have shown a much better performance than feature based. Within the first category are to mention the seminal works of Sung and Poggio [5], based on a Gaussian modeling of PCA decompositions together with a neural classifier, and the one of Rowley *et al.* [3], which use a neural network composed by retinal-based receptive fields.

According to the current state-of-the-art (for references and exact performance information please see [1] or [8]), the following face detection systems have the higher recognition rates, tested using standard face databases: (i) SNoW (Sparse

Network of Winnows), proposed by Roth *et al*. [2], which uses sparse feature vectors, two linear processing neurons and the Winnow learning rule, (ii) Schneiderman and Kanade detector [4], which uses wavelet analysis together with a Bayesian classifier, and (iii) Yang *et al*. Detector [7], which uses Self-Organizing Map clustering and Fisher linear discriminant analysis.

On the other, the system proposed by Viola and Jones [6] outperforms previous systems in terms of processing speed, by keeping an acceptable recognition rate. This system uses simple, rectangular features (a kind of Haar wavelets), a cascade of filters that discard non-face images, the integral image for fast computation of these filters and asymmetrical Adaboost as a boosting strategy for the training of the detectors.

The aim of this paper is to propose a hybrid face detector that combines the high processing speed of the cascade detector of Viola and Jones with the high detection rate of the detector of Schneidermann and Kanade. This behavior is achieved by incorporating this last detector in the middle stages of the cascade detector. The detection results obtained with both detectors are fused using a new heuristic algorithm that merges overlapping face detection regions.

The paper is structured as follows. In section 2 the proposed hybrid face detector is described. In section 3 some face detection results using the proposed detector are presented. Finally, some conclusions of this work are given in section 4.

2. Proposed Hybrid Face Detector

For simplicity in the explanation of the proposed hybrid detector, in this paper we will call the Viola and Jones detector *cascade Adaboost detector*, and the Schneidermann and Kanade detector *Wavelet-Bayesian detector*.

The proposed hybrid detector takes as starting point the cascade Adaboost detector [6], a very fast face detector which detection rate can be increased by improving the detection of faces which usually present problems, like for example black people faces, low contrast faces, and asymmetrical illuminated faces.

For increasing the detection rate of the hybrid detector, the Wavelet-Bayesian detector is also included [4]. High processing speed is kept by applying this last detector inside the Adaboost cascade, after the first 3 Adaboost filters are applied. Afterwards, the detection results of both detectors are fused.

2.1 System Overview

The block diagram of the proposed detector is presented in figure 1. For detecting faces at different scales a multiresolution analysis is performed by scaling the input image by a factor of 1.2. This scaling is performed until images from about 24x24 pixels are obtained. This operation is carried out in the *Multiresolution Analysis Module*. In the *Window Extraction Module* for each of these scaled versions of the input image, all possible windows of 24x24 pixels are extracted. These windows are then processed by the face detectors.

For detecting the faces, each of the 24x24 windows is processed by the first three filters of the cascade Adaboost detector. When a window is classified as non-face by any of these three filters non-further processing is done. In the opposite case, it goes

further in the cascade. After the application of the third filter, the remaining windows are sent in parallel to the next part of the cascade (filters 4 to 21), which keep on discarding non-face windows, and to the Wavelet-Bayesian detector. This detector makes a histogram equalization of the received windows (*Pre-Proc. Module*) and a Wavelet Decomposition followed by a Bayesian Classification (*Wavelet Face Detection Module*). A switch before the Wavelet-Bayesian detector discards one window every two windows for speeding the processing (see fig. 1). After all windows are processed and classified as faces or non-faces, in the *Overlapping Detection Processing Modul*e all face windows are processed and fused for determining the size and position of the detected faces.
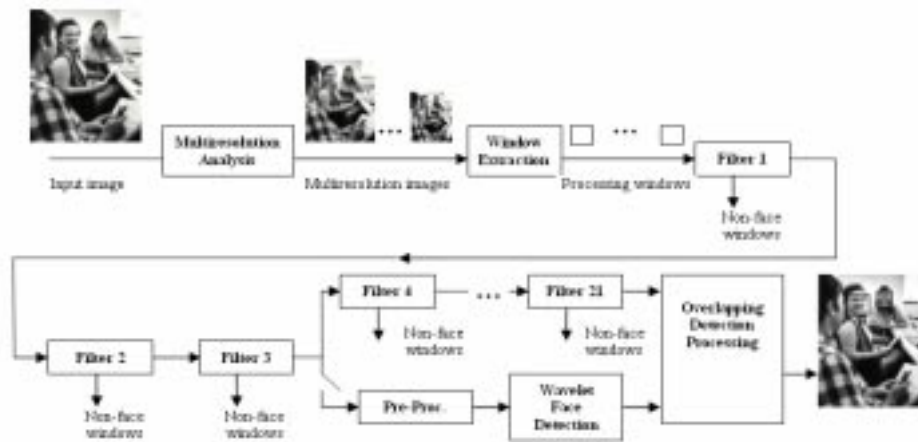

Fig. 1. Block diagram of the proposed hybrid face detector. Filters 1 until 21 correspond to the filters of the cascade Adaboost detector. Pre-Processing and Wavelet Face Detection modules correspond to the Wavelet Bayesian detector.

2.2 Cascade Adaboost Detector

The implemented cascade face detector detects frontal faces with small in-plane rotations and it is based mainly on [6]. This face detector corresponds to a cascade of filters that discard non-faces and let faces to pass to the next stage of the cascade. This architecture seeks to have a fast face detector, considering the fact that only a few faces are to be found an image, while almost all the image area correspond to non-faces. The fast detection is achieved in two ways: (i) having a small complexity in the first stages of the cascade (filters composed by few detectors, 2 to 5) and greater complexity in the later stages of the cascade (filters composed by many detectors, 100 to 400), and (ii) using simple features called rectangular features (the detectors), which are quickly evaluated using a representation of the image called integral image.

Each of the filters of the cascade is trained using an asymmetric version of Adaboost (see explanation in [6]), which gives more importance to errors occurring during the training process when classifying faces as non-faces than non-faces as faces. Adaboost sequentially trains and selects a small number of rectangular features.

The main problem of this face detector is the large training time, which can extend for weeks or even months.

The final cascade Adaboost detector implemented in this work has 21 layers and was trained in about a month. To train each layer 1338 face images were used, and non-faces images were collected from 8000 images that did not contain any face. All these training images were obtained mainly from Internet, especially from the google image searcher, and from personal images databases. For training the first 2 filters, 4000 and 3000 non-face images were randomly chosen from our dataset. For training the remaining filters of the cascade 1500 non-face images wrongly classified by the already trained cascade were collected (a kind of boostraping). For reducing the training time a randomly chosen subset of the set of rectangular features was used in each iteration of the Adaboost algorithm. Each time that a decision rule was trained, only 50% of the training examples (faces and non-faces) were employed. As a results of the training, the final number of detectors used at each of the 21 stages of the cascade was 2, 5, 20, 20, 50, 50, 100, 100, 100, 100, 100, 200, …, 200 and 400, respectively.

2.3 Wavelet-Bayesian Detector

The Wavelet-Bayesian detector use a Wavelet decomposition for extracting feature patterns, together with a naive Bayes classifier that first estimates the probability of that a given window correspond to a face or a non-face, using these patterns, and then it employs the ratio between these two values for determining if the window is a face or a non-face. As in [4], this is carried out using the a posteriori probabilities of different patterns occurring in faces and non-faces. These patterns are extracted from a two level wavelet transform [9].

Groups of 8 coefficients of the wavelet transform of the image were used, and each of the coefficients was quantized to 3 values. The quantization thresholds were chosen from 5 values (see details in [9]). These thresholds were chosen for maximizing the detection rate and diminishing the false positive rate. Six different types of groups of coefficients were used, which represent inter-orientation coefficients (LH-HL level 1 and LH-HL level 2), inter-frequency coefficients (LH level 1 – LH level2, HL level 1 – HL level 2) and intra-frequency coefficients (LH level 1, LH level 2, HL level 1, HL level 2).

For training the detector 800.000 faces were generated from 1500 faces, using variations in different scales, rotations and displacements. 2.000.000 non-faces were collected from images non-containing faces. All these training images were obtained mainly from Internet, especially from the google image searcher, and from personal images databases. For training this detector much more images were used than in the case of the cascade Adaboost detector, because the training of this last detector is very time consuming. During training 120.000 more non-faces were collected in cases were the detector has wrongly classified them (boostraping).

2.4 Overlapping Detection Processing

Face windows are processed and fused for determining the size and position of the detected faces. Overlapping detections were processed for filtering false detections

and for merging correct ones. All detections (detected face region) were separated in disjoint sets using the following heuristic. Considering the inscribed circumference of each square face region, two detections belonged to the same set if the sum of their circumference radius is smaller than 0.4 times the distance between their centers, and if each radius is not larger than twice the other. If a set contained only one element, this detection is discarded. Detections belonging to each set are merged by averaging the coordinates of the corners of all square face regions. Figure 2 shows an example of applying this heuristic on an image which has two close faces.
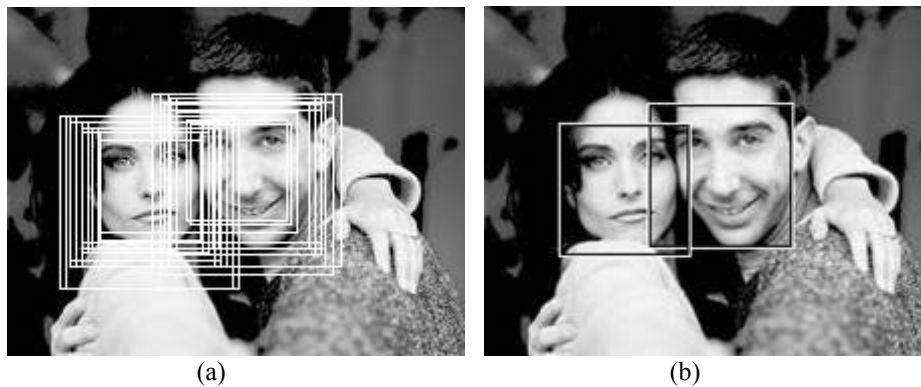

Fig. 2. Results from the use of the heuristic that separates and merges overlapping detections. (a) Overlapping detection, (b) Final detections.

3. Face Detection Results

In order to test the system we used the database used in [3] that correspond to the union of a database of Rowley, Baluja and Kanade and a database of Sung and Poggio. That is available in [10] and in commonly known as MIT+CMU face database. This database consists in 130 images containing 507 faces.

Figure 3 shows a graph comparing the detection results, ROC (Receiver Operating Characteristic) curve, for the hybrid detector, the cascade Adaboost detector and the Wavelet-Bayesian detector. This last detector was used together with three filters from the cascade for speeding the results. Table 1 shows the exact values of this evaluation. For obtaining the different points of the ROC curve for the hybrid detector and the Wavelet-Bayesian detector, different thresholds of the wavelet decomposition were used, while in the cascade Adaboost detector, filters at the end of the cascade were sequentially removed (see details in [9]).

From the presented results it can be notice that the hybrid detector has an increase of 2% to 5% in the detection rate over the cascade detector, improving the detection of black faces and faces with low contrast. This is shown with more details in figure 4. Results for the hybrid detector are presented in figures 4 (a) and (c) and results of the application of the cascade detector on the same images are presented in figures 4 (b) and (d). These results show how the hybrid detector detects black faces and low contrasts that the cascade detector does not detect.

Concerning processing time we can says that following. The time that took to evaluate one point of the ROC curve was about 19 minutes for the hybrid detector

detectors, while it was 10 minutes for the cascade detector. In contrast to that, the time for evaluating the Wavelet-Bayesian detector was more than 6 hours.
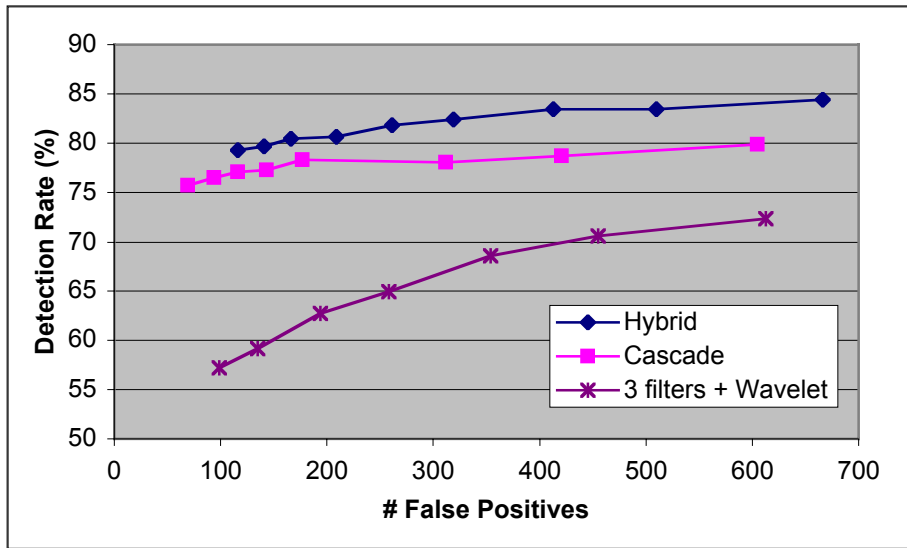

Fig. 3. Comparison ROC curves of Hybrid, Cascade and Wavelet detectors on MIT+CMU test sets.

Table 1. Evaluation of Hybrid, Cascade and Wavelet detectors on MIT+CMU test set.

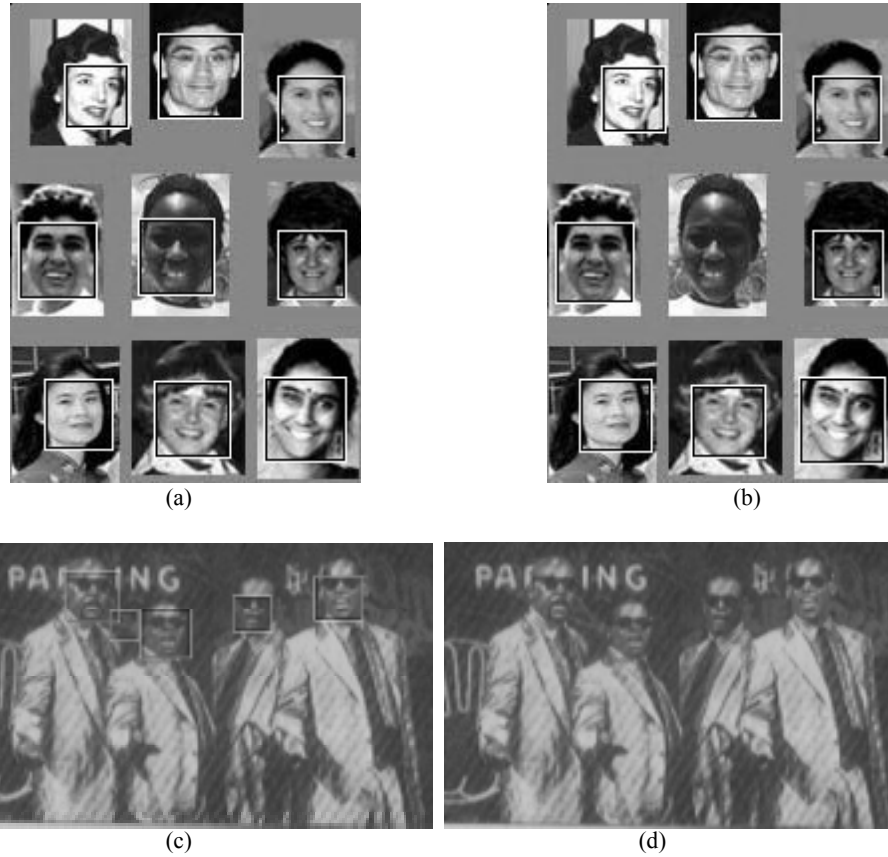| Hybrid Detector		Adaboost Cascade Detector		3 Filters + Wavelet-Bayesian Detector	
Detection Rate (%)	False Positives	Detection Rate (%)	False Positives	Detection Rate (%)	False Positives
79.29	116	75.73	69	57.19	99
79.68	141	76.52	94	59.17	135
80.47	166	77.12	116	62.72	194
80.67	209	77.31	143	64.89	258
81.85	261	78.3	177	68.63	354
82.44	319	78.1	312	70.61	455
83.43	413	78.69	421	72.38	613
83.43	510	79.88	605		
84.41	666				

Fig. 4. Detection results of the hybrid detector (a) and (c), and the Adaboost cascade detector (b) and (d).

4. Conclusions

In this article was proposed a hybrid face detector that combines the high processing speed of the cascade Adaboost detector with the high detection rate of the Wavelet-Bayesian detector. This integration was achieved by incorporating this last detector in the middle stages of the cascade detector. Results of the application of the proposed detector to a standard face detection database were presented. These results shown that the obtained detector has higher detection rate than the Adaboost detector, while keeping its high processing speed.

Currently we are working for improving the detection rates of the cascade Adaboost as well as the Wavelet-Bayesian detectors, which should improve the detection rate of the hybrid system. For doing that we are building better face database and optimizing the training of the cascade detectors. The use of methods for compensating the illumination, such as linear function subtracting [3], should also improve the performance of the final detector.

References

1. E. Hjelmås, B. K. Low, "Face detection: A survey", *Computer Vision and Image Understanding* 83, 236-274, 2001.
2. D. Roth, M. Yang, and N. and Ahuja, "A SNoW-based Face Detector", *Advances in Neural Information Systems* 12, MIT Press, 2000.
3. H. Rowley, S. Baluja, and T. Kanade, " Neural Network-Based Detection", *IEEE Trans. Pattern Anal. Mach. Intell*., Vol.20, No. 1, 23-28, 1998.
4. H. Schneidermann and T. Kanade, " A statistical model for 3D object detection applied to faces and cars", *IEEE Conf. on Computer Vision and Pattern Recognition*, Vol. 1, 746 – 751, 2000.
5. K. Sung and T. Poggio, "Example-Based Learning for Viewed-Based Human Face Deteccion", *IEEE Trans. Pattern Anal. Mach. Intell*., Vol.20, No. 1, 39-51, 1998.
6. P. Viola and M. Jones, "Fast and Robust Classification using Asymmetric Adaboost and a Detector Cascade", *Advances in Neural Information Processing System* 14, MIT Press, Cambridge, MA, 2002.
7. M. Yang, N. Ahuja, and D. Kriegman, " Mixtures of linear Subspaces for Face Detection", *Fourth IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 70 – 76, 2000.
8. M. Yang, D. Kriegman, N. Ahuja, "Detecting Faces in Images: A Survey", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No 1, pp. 34-58, 2002.
9. R. Verschae, "Detecting Faces in Image Databases", Thesis for the title of Electrical Engineer, Department of Electrical Engineering, Universidad de Chile, 2003 (in Spanish).
10. CMU+MIT face database. Available until March 2, 2003 on http://vasc.ri.cmu.edu/IUS/eyes_usr17/har/har1/usr0/har/faces/test/.