An Automatic Human Face Detection Method *

Peter Peer, Franc Solina
Computer Vision Laboratory
Faculty of Computer and Information Science
University of Ljubljana
Tržaška c. 25, SI-1001 Ljubljana, Slovenia
Tel. +386 (61) 1768 381, Fax. +386 (61) 1264 630
e-mail: peter.peer@fri.uni-lj.si, franc@fri.uni-lj.si

Abstract

This article contains a proposal for an automatic human face detection method, that tries to join several theories proposed by different authors. The method is based on detection of shape features (eye pairs) and skin color. The method assumes certain circumstances and constraints, respectively. Therefore it is not applicable universally. Given the constraints, it is effective enough for applications where fast execution is required. Test results are given and at the end some directives for future work are discussed.

1 Introduction

Because of image-databases and "live" video information is growing more and more widespread and expansive, their intelligent or automatic examining is becoming exceptionally important. People, i.e. human faces, are one of most common and very specific objects, that we try to trace in images.

The system for automatic detection and recognition of human faces represents an example of such an intelligent examining. Basic algorithm of the system is evident from figure 1.

This article deals with first two steps of the algorithm, i.e. face and feature segmentation.

The purpose of automatic in-image face detection methods is obvious: their primary goal is to segment image into regions that contain human faces or its parts and into regions which can be - because they don't represent a human face neither any of its parts - neglected. Some possible applications for automatic face detection are:

*This work was supported by the Ministry of Science and Technology of Republic of Slovenia (Project J2-8829).

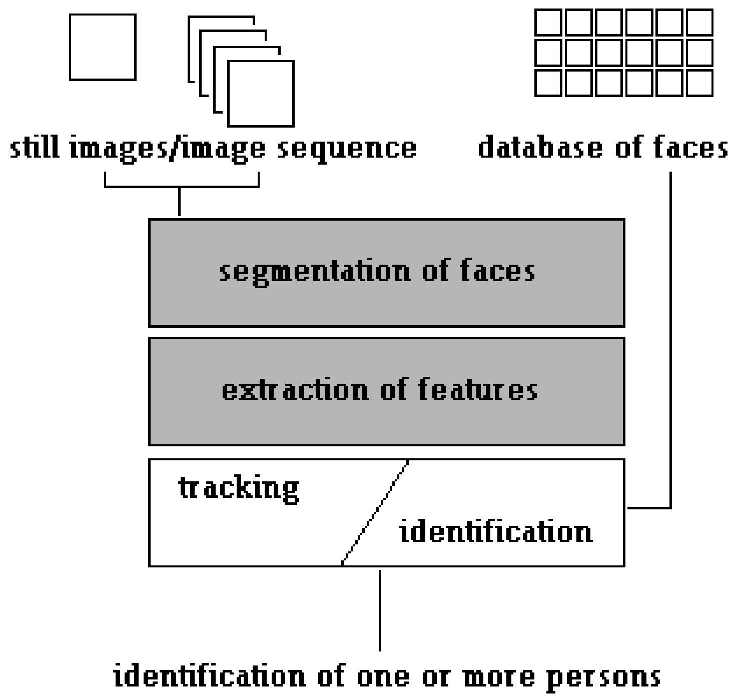

Figure 1: Machine face recognition algorithm

- supervision and security applications,
- video-conference applications,
- animation of facial expressions,
- remote camera control applications,
- etc.

The majority of already developed methods have generally at least one of these two problems:

- too high computational (time-, space-) complexity and/or
- too low effectiveness.

At this point it is necessary to emphasize the fact, that automatic face detection as well as most other automatic object-detection methods is a very pretentious task, especially because of significant sample variations, which can't be easily analytically described with parameters.

2 Face detection

There are a few distinct approaches to face detection [5, 7]:

- The *top-down model-based approach* assumes a different face model at different coarse-to-fine scales. For efficiency, the image is searched at the coarsest scale first. Once a match is found, the image is searched at the next finer scale until the finest scale is reached. In general, only one model is assumed in each scale (usually in the frontal-parallel view) and thus it is difficult to extend this approach to multiple views.

- The *bottom-up feature-based approach* searches the image for a set of facial features and groups them into face candidates based on their geometric relationship. Though this approach can be easily extended to multiple views, it is unable to work well under different image conditions because the image structure of the facial features vary too much to be robustly detected by the feature detectors.

- In *texture-based approach* faces are detected by examining the spatial distribution of the gray-level information in the subimage (using Space Gray Level Dependency (SGLD) matrices). This is again not easily extensible to multiple viewpoints.

- The *neural network approach* detects faces by subsampling different regions of the image to a standard-sized subimage and then passing it through a neural network filter. In general, the algorithm performs very well for frontal-parallel faces but performance deteriorates when extended to different views of the face. It is still not possible to extend the algorithm to detect faces in profile views.

- The *color-based approach* labels each pixel according to its similarity to skin color, and subsequently labels each subregion as a face if it contains a large blob of skin color pixels. It can cope with different viewpoint of faces but it is sensitive to skin color and the face shape.

- *Motion-based approaches* use image subtraction to extract the moving foreground from the static background. The face is then located by examining the silhouette or the color of the difference image. This approach will not work well when there are a lot of moving objects in the image.

- At *depth-based approach* primary facial features are localized on the basis of facial depth information. In the first step, pairs of stereo images containing frontal views are sampled from the input video sequence. Then point correspondences over a large disparity range are determined using a multiresolution hierarchical matching algorithm. Finally, the facial features are located based on depth information.

3 Proposed face detection method

The method in question combines two common approaches, one based on features and the other based on colors (see chapter 2). The two basic limitations of the method thus originate from constraints of already mentioned approaches:

- input image must have high enough resolution; the face must be big enough and
- it is sensitive to the complexion (i.e. fair).

The basic idea of the algorithm is as follows: find in image all regions, which contain possible candidates for an eye, then on the basis of geometric face characteristics try to join two candidates into an eye pair and finally, confirm or refuse the face candidate using complexion information.

The method was projected over a set of quite different pictures, i.e. the training set. The goal of the method was to reach maximum classification accuracy over the images, which meet the following demands and constraints, respectively (beside already mentioned two):

- real-time operation on - for the present - standard personal computer,
- plain background,
- uniform ambiental illumination,
- fair-complexion faces, which must be present in the image in their entirety (frontal position) and
- faces, turned round for at most 30 degrees.

The method's effectiveness was tested over an independent set of images, i.e. the testing set.

The basic principle of operation is shown on Fig. 2.

The method requires some thresholds, which play a crucial role for proper processing. They are set quite loosely (tolerantly), but they become effective as a sequence. All thresholds were defined experimentally using the training set.

Proposed face detection algorithm

Input: image in BMP format
Output: input image with denoted faces in it
Basic steps of algorithm are:

1. Firstly, completely unimportant colors are eliminated from image (those, which can't represent a face). All insignificant colors are replaced with white color.
2. Image is then converted into greyscale picture (PGM format).

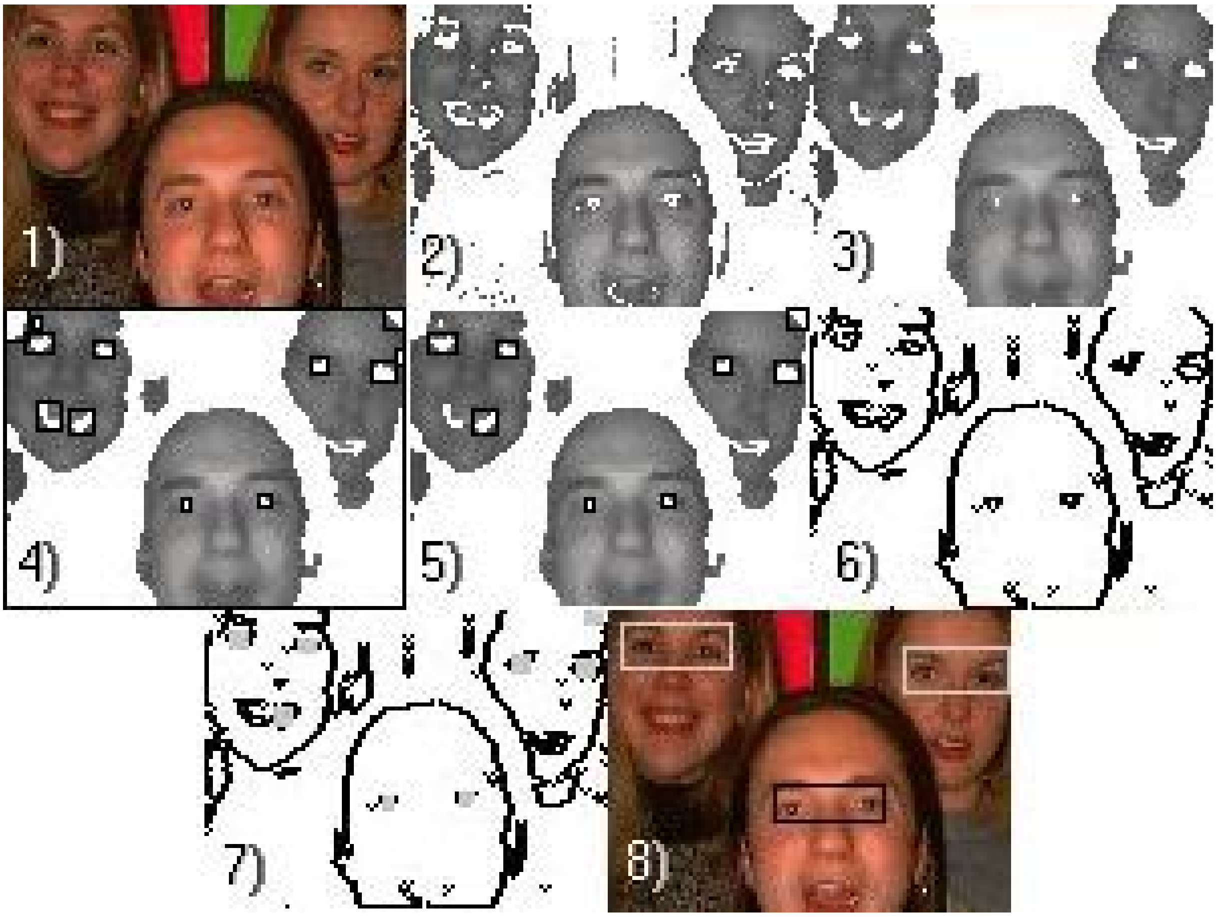

Figure 2: Basic principles of proposed face detection method: 1. input image, 2. eliminated insignificant colors, 3. image filtered with median filter, 4. segmented white regions, 5. eliminated insignificant regions, 6. traced edges, 7. best possible circles within regions of interest, 8. output image. Also next feature of the method is evident from the figure: because of greater distance of the faces in the background, the algorithm ascribe to the latter lower probability based on the fact, that they have lower resolution and worse illumination then the face in the foreground

3. It is filtered with a median filter[1].
4. With the help of a "Region growth" algorithm white regions are segmented.
5. Regions which can't comprise an eye are eliminated as well.
6. Edges are traced in the image with significant greys.

[1] The algorithm with median filter blurs unimportant white regions (like, for an instance, regions, that are consequences of different spots on the face), achieves greater compactness of interesting white regions and eliminates small skin-like regions (which are within white regions).

7. Within preserved regions the algorithm searches for circles (eye candidates) using Hough transform.

8. For each region the best possible circle is found.

9. Using geometric face characteristics the algorithm finds partner circles representing a pair of eyes.

10. For face candidate confirmation, color information of the whole face is used. With the help of this information the algorithm also (loosely) predicts the probability of a face.

11. Finally, using some heuristic rules, the algorithm tries to eliminate falsely classified faces.[2]

4 Results

In order to get reliable results regarding algorithm's effectiveness, this method was tested over an independent testing set. The set consists of images taken from two public image-databases (PICS, M2VTS - [8]), few realistic face animations and a couple of celebrity pictures.

A common feature of all images is their larger or lesser suitability for identification documents. Celebrity pictures are included in the test because they were taken by professional photographers.

All images of the testing set do not meet the required constraints, as stated in Chapter 1, but exactly these images illustrate the proposed method's drawbacks very well. Images of faces which are less likely to be detected are:

- images of faces turned away from frontal position

- images with a complex background.

There are five images of this kind in the testing set, in three out of these the method successfully detected a face. The fact, that in two images faces were turned towards profile position, tells us that the method tolerates small deviations (see Fig. 3b) - on the right side). In the third image the background wasn't plain (see Fig. 3d) - the first image). On the remaining two images the face was turned too much towards the profile position.

Table 1 presents testing results over the entire testing set (without the before mentioned two images that didn't meet the demands, stated in chapter 1), Figs. 3 and 4 however, illustrate the method's effectiveness. False positives could be eliminated in subsequent processing.

Generally speaking, glasses represent certain problem (e.g. compact frames or sunglasses), but hair-dress, moustache, beard and different mien don't!

[2]The 11 steps of the algorithm are the result of the conceptual design of the algorithm.

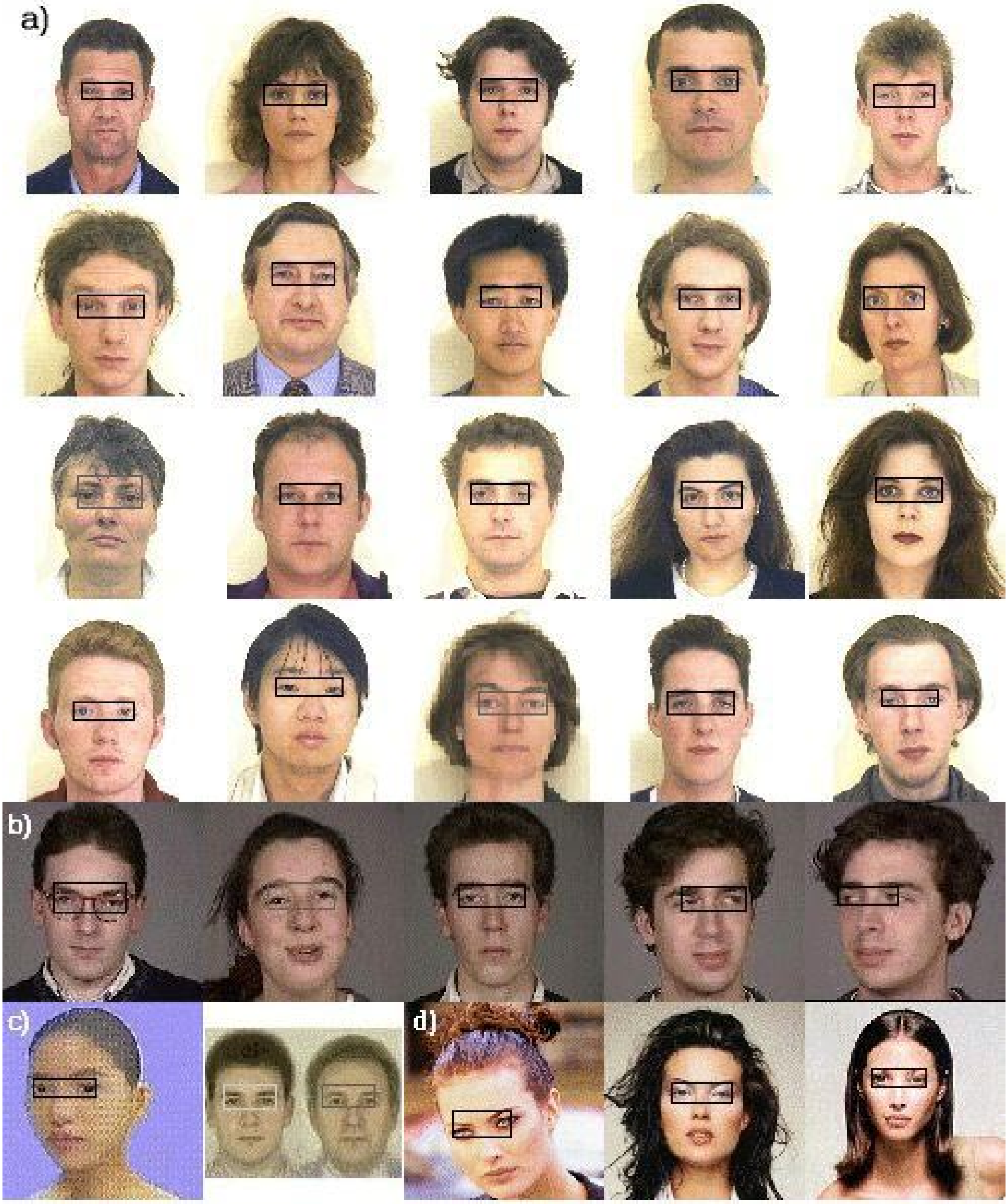

Figure 3: Some positive results from the test set: a) from PICS database, b) from M2VTS database, c) computer graphics face animations, d) images of professional photographs

complete test set - 44 images - 45 faces	
number of detected faces	number of hits
46	44

Table 1: Results of the method over the whole test set. There was 1 false positive and 1 false negative detected face

Figure 4: Negative results from the test set: on the left image the error is caused because of hair color, which corresponds to complexion, while on the right image the error is caused because an eye region and the background region align (the face is found on the basis of the combination eye-eyebrow)

5 Conclusion and future work

From these result we can conclude, that taking into account specific constraints, classification accuracy achieved by this method is very high. These constraints hold if we restrict ourselves on a particular image domain, which meets special requirements. In other words, the method can be effective enough for an application, which wouldn't suffer from algorithm constraints and which would require fast operation.

The proposed method could, for example, serve as a front part of a face recognition system. Falsely detected faces could be eliminated using correspondence to the database of already known faces. This was exactly the purpose of the method's conception.

We are planning to test the method over a larger set of images; also over images, that don't contain faces. Such tests should help to detect weak points of the algorithm and consequently give directions for algorithm upgrade.

References

[1] R. Feraud, O. Bernier, J.-E. Viallet, M. Collobert, D. Collobert, A Conditional Mixture of Neural Networks for Face Detection, Applied to Locating and Tracking an Individual Speaker, *CAIL'97*, Kiel, Germany, pp. 464-471, 1997.

[2] T. S. Jebara, A. Pentland, Parametrized Structure from Motion for 3D Adaptive Feedback Tracking of Faces, *IEEE CVPR Proceedings*, pp. 144-150, 1997.

[3] E. Osuna, R. Freund, F. Girosi, Training Support Vector Machines: an Application to Face Detection, *IEEE CVPR Proceedings*, pp. 130-136, 1997.

[4] H. A. Rowley, S. Baluja, T. Kanade, Neural Network-Based Face Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23-38,1998.

[5] K. Sobottka, I. Pitas, A novel method for automatic face segmentation, facial feature extraction and tracking, *Signal processing: Image communication*, 12, pp. 263-281, 1998.

[6] K.-K. Sung, T. Poggio, Example-Based Learning for View-Based Human Face Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 39-51, 1998.

[7] K. C. Yow, R. Cipolla, Feature-based human face detection, *Image and Vision Computing*, no. 15, pp. 713-735, 1997.

[8] Internet:

- *http://pics.psych.stir.ac.uk/* - PICS Image Database,
- *http://www.tele.ucl.ac.be/M2VTS/* - M2VTS Face Database.