Single and Multi-Person Face Recognition Using the Enhanced Eigenfaces Method

Preliminary Communication

Krešimir Romić
Josip Juraj Strossmayer University of Osijek,
Faculty of Electrical Engineering, Department of Software Engineering
Cara Hadrijana 10B, Osijek, Croatia
kresimir.romic@etfos.hr

Časlav Livada
Josip Juraj Strossmayer University of Osijek,
Faculty of Electrical Engineering, Department of Software Engineering
Cara Hadrijana 10B, Osijek, Croatia
clivada@etfos.hr

Andrija Glavaš
Josip Juraj Strossmayer University of Osijek,
Faculty of Electrical Engineering, Department of Software Engineering
Cara Hadrijana 10B, Osijek, Croatia
aglavas11@gmail.com

Abstract *– This research studies and analyzes the possibility of single-person and multi-person face detection and recognition. Face detection is performed by the Viola-Jones face detection method and recognition is performed by the Eigenfaces method. Unchanged face detection and recognition methods are explained and tested to their limits. Improvement in face recognition is achieved by observing the flaws of the Eigenfaces method and their enhancement.*

Keywords *– Eigenfaces, face detection, face recognition, Viola-Jones.*

1. INTRODUCTION

Face recognition was a field that was extensively studied in the past years but it is still an active area of research. Face recognition computer-based algorithms are being developed constantly.

There are quite different requirements on face recognition software itself, e.g., there are face recognition systems that do not have very high requirements on the accuracy but have a requirement referring to the overall speed of the system. On the other hand, there are systems where the primary requirement is recognition accuracy. Both of these applications face an obstacle in the fact that subjects do not cooperate with the recognition system. Images are not acquired in a controlled environment and because of that, illumination and appearance of subjects can change between images, e.g., an image taken outside on a sunny day versus an image taken inside with artificial lightning. Global research has shown that some face recognition algorithms have exceeded humans in face-matching tasks with images taken under varying illumination conditions [1].

There are numerous applications where the recognized subjects are still allies of the system, e.g., in automobiles or mobile phones, so it is reasonable to expect availability of adequate training data. Since we can expect a certain level of pose difference in our input data, this proposed method is based entirely on frontal face images.

This paper presents an enhanced Eigenfaces method for face recognition which uses the Viola-Jones method for face detection, an artificially expanded training set, as well as the CLAHE method for contrast enhancement.

This paper is organized as follows: Section 2 explains face detection and recognition methods, and Section 3 shows the application with proposed enhancements. Experimental results are presented in Section 4 and the conclusion is given in the last section.

1.1. RELATED WORK

According to the survey [2], the starting point in this research field is a face detection method introduced by Viola and Jones [3]. Many authors base their work on the Viola-Jones approach which is further modified and improved [4][5]. Survey [6] provides an overview of face detection methods with emphasis placed on detection in varying conditions which include variability in size, shape, color and texture.

The field of face recognition research can be roughly divided into two categories. The first category encompasses methods working in 3D space, and in the second category, these are methods working in 2D space. Face recognition methods in 3D require a 3D model as an input or recreate a 3D model from a 2D image. Significant examples of the latter example are given in [7] and [8]. With a sufficiently accurate 3D model, the 3D based techniques can achieve better results in applications than the 2D based techniques. Pertinent examples of 2D based face recognition methods are examples using Eigenfaces [9][10], Principal Component Analysis [11], or methods using sparse representation [12].

2. FACE DETECTION AND RECOGNITION

The main goal of face recognition software is to detect a single face or multiple faces in the image. There has been a plethora of methods for face detection but one of the most popular ones is the work of Viola and Jones [3][13]. In this article, the Viola-Jones algorithm is used in combination of Eigenfaces method for enhanced face detection.

2.1. VIOLA-JONES FACE DETECTOR

The Viola-Jones detection method was chosen in our research because of its high detection rate. This detector consists of three essential parts that allow its running in real time: the integral image, AdaBoost and the attentional cascade structure.

The *integral image* is an algorithm for cost-effective generation of the pixel intensities sum in a specified rectangle in an image [14]. Haar-like features are calculated by the Viola-Jones Integral image according to the following equation:

$$ii(x,y) = \sum_{x'\leq x, y'\leq y} i(x',y'), \quad (1)$$

where $i\ (x,y)$ and $i(x',y')$ are the integral image at pixel location and the original image, respectively. Calculation of the sum of a rectangular area inside the original image is quite efficient and requires four additions for any optional rectangle. This sum is calculated by using the following equation:

$$\sum_{(x,y)\in ABCD} i(x,y) = ii(A) + ii(D) - ii(B) - ii(C) \quad (2)$$

Five Haar-like patterns can be seen in Figure 1, where the size and position can vary with only one constraint, i.e., that the black and white rectangles have the same dimensions and border each other.

a) b) c) d) f)

Fig. 1. Five Haar-like patterns

For example, a 24 x 24 image has 43200 features of category a) and c), 27600 features of category b) and d) and 20736 features of category e).

a) b) c)

Fig. 2. Haar-like features application

Haar-like feature is applied to the original image, Figure 2a), that looks similar to the eye region which is darker than the upper cheeks, Figure 2b), and Haar-like feature that is similar to the eyebrow region which is darker at the ends because of the eyebrows is applied in Figure 2c).

The *AdaBoost* algorithm was first introduced by Freund and Schapire [15] and it is used for constructing strong classifiers as a combination of weak classifiers, as shown in the following equation:

$$H(x) = sign\left(\sum_t \alpha_t h_t(x)\right), \quad (2)$$

where $h_t(x)$ is a weak classifier, α_t is the weight and $H(x)$ is the final classifier made up of weak classifiers. Every weak classifier is configured to detect features not classified or misclassified in previous iterations [16].

The attentional cascade structure: The main idea of the attentional cascade structure is to build smaller boosted classifiers [13]. Every node is a collection of smaller, weaker classifiers and these nodes form a degenerated decision tree called a *cascade*. The number of classifiers increases with an increase in the level, which means that later stages have more classifiers. This happens because nodes try to pass all positive sub-windows to further stages but still reject some of the negative ones. An input sub-window passes a series of nodes and in every node it makes a decision whether to discard it or pass it onto the next stage. In this way, only a small amount of sub-windows will pass through to the next forthcoming stage. In the starting process, most of the negatives are rejected thus greatly improving efficiency.

2.2. EIGENFACES METHOD

The goal of this method is to find relevant information in the face image, code them and compare to the image database that consists of images coded in the same way [10]. A simple way of extracting the information contained in the face image is to capture the variations in the set of face images and use that information for coding and individual faces comparison. The main goal is to find elementary components of faces distribution, i.e., covariance matrix eigenvectors of the set of face images. Every part of the image contributes, more or less, to every eigenvector and in that way eigenvectors can be interpreted as ghost faces called *Eigenfaces*, as shown in Figure 3.

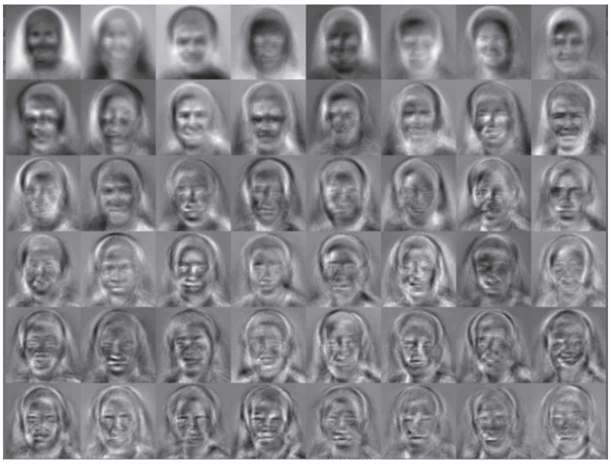


Fig. 3. Eigenfaces

Another crucial image called a *mean face* is calculated. This is achieved by calculating the median of all pixels in the face training set. This image contains information about all images in the training set, but only the information that is similar to all images. Eigenfaces and mean face are used because all faces can be reconstructed by using a small collection of Eigenfaces.

The Eigenfaces approach to face recognition involves the following steps:

1. The ready-state set of images for training,
2. Computing the Eigenfaces from the training set with retention of M images with the largest eigenvalues (M images define the *face space*), and
3. Calculating the distribution in the M dimensional Eigenfaces space.

After initialization, the following steps are used to recognize new faces in images:

1. For a new face image that needs to be recognized, a set of weights is calculated that are assigned to M Eigenfaces so that the image is projected on every Eigenface.
2. Determine whether the face is in the image or not by checking if the image is close to the face space.
3. If there is a face in the image, numerical values are assigned to eigenvalues.
4. Update the Eigenfaces and/or the numerical values (weights) assigned to faces.
5. If the new face is repeated several times, weights for this new face must be calculated and inserted into familiar faces.

3. FACE DETECTION AND RECOGNITION APPLICATION

This application for face detection and recognition was written in the MS Visual Studio 2010 framework in C++. It also uses the OpenCV library. The graphical interface of the application is shown in Figure 4.

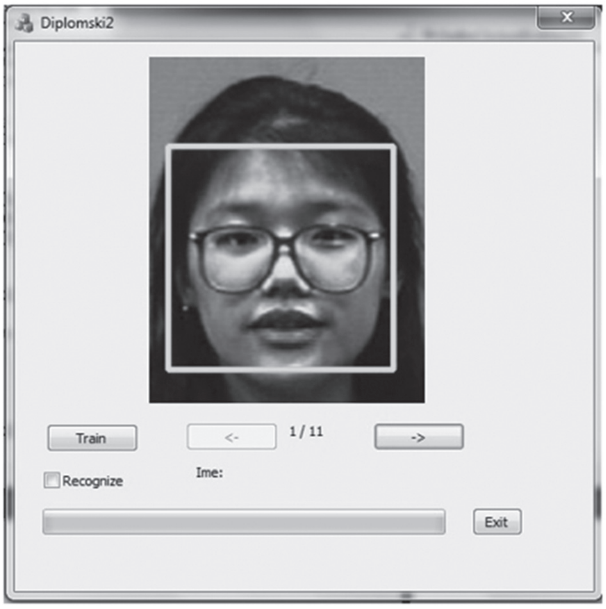


Fig. 4. Graphical interface of the application

For proper use of the face recognition application, it is necessary to train the application by known face images. Computational complexity of the learning process is *O(MNT)*, where *M*, *N* and *T* stand for the numbers of filters, examples and thresholds, respectively. For each round of training, the following must be performed:

- Evaluate each rectangle filter on each example,
- Select the best threshold for each filter,
- Select the best filter/threshold combination,
- Reweight examples.

3.1. RECOGNITION ACCURACY ENHANCEMENT

The basic way to increase the accuracy would be to increase the number of images used in the training set. In addition, it is possible to increase the number of software images from the existing images in the following way:

1. An automatic mirror image made by axially-symmetric transformation.

2. Resize, rotate, move a face in order to increase the variety of the images for training.
3. Add artificial noise to the image in order to make the system more resistant to noise.

Besides manipulation of set images, some preprocessing methods can further enhance face detection and recognition. Histogram equalization quite effectively solves the problem of poor contrast thus contributing to recognition accuracy.

The application incorporates a number of mechanisms so that detection and recognition accuracy could reach the optimal level. As can be seen in [17], the algorithm achieves the highest accuracy by combining the local brightness filter and contour highlighting. In order to solve the problem of over- and underexposure, CLAHE (Contrast Limited Adaptive Histogram Equalization) is used as the method for local histogram equalization. The second part is contour highlighting that draws contours on the original image for the purpose of increasing detection accuracy.

After the preprocessing stage, an elliptic mask is placed on the image so that the center of the ellipse is actually the center of the previously detected face quadrilateral. The ellipse covers unwanted parts of the image, such as the hair, ears and the background so after placing this mask the inner facial features stand out, as can be seen in Figure 5.

Fig. 5. Faces after preprocessing stage

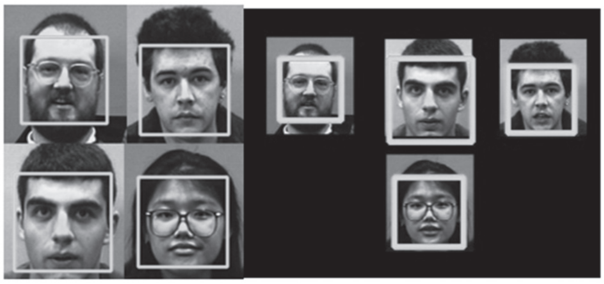
4. RESULTS

Application to face detection and recognition was tested on the face image database *faces94* [18]. The database consists of 153 people (133 men and 20 women), image resolution is 120x200 pixels. As this database is predominantly designed for face recognition, it was expected that in these cases the application will yield quite good results. For testing purposes, eight people were randomly selected and the application was trained with these faces. Every person had 20 face images with varying facial expressions. The ninth person that is unfamiliar to the application is added in order to test the application.

As is evident from the results shown in Table 1, the application has showed the utmost accuracy in detection and recognition. Images with one person in the image were used in this test, so it is necessary to try out the image in which there are several people. The database used has no group images of people but they can be digitally created, as can be seen in Figure 6.

Table 1. Face detection and recognition of one person in the image

Image	Face detected (Yes/No)	Face recognized (Yes/No)
1	Yes	Yes
2	Yes	Yes
3	Yes	Yes
4	Yes	Yes
5	Yes	Yes
6	Yes	Yes
7	Yes	Yes
8	Yes	Yes
9	Yes	No



a) b)

Fig. 6. The second stage testing image

Both images shown in Figure 6 are composed of the same faces but the main difference is image resolution. The resolution of the image in Figure 6a) is 352x400 and the resolution of the image in Figure 6b) is 805x558. The results of this test are shown in the following two tables.

Table 2. Results of detection and recognition in images with multiple faces – Fig. 6a)

Image	Face detected (Yes/No)	Face recognized (Yes/No)
1	Yes	Yes
2	Yes	Yes
3	Yes	Yes
4	Yes	Yes

Table 3. Face detection and recognition test performed in Fig. 6b)

Image	Face detected (Yes/No)	Face recognized (Yes/No)
1	Yes	Yes
2	Yes	No
3	Yes	No
4	Yes	No

The results in these tables show that although the images consist of the same faces, image resolution is the prevalent factor in face detection and recognition.

Images from the faces94 database have been an easy task for the program, so the photographs of a 2848x2134 resolution were taken and the persons in the images are relatively far from the camera (the facial area covers a relatively small portion of the image), Figure 7.

Fig. 7. Photographs for second-stage testing

Second-stage testing (testing the application on photographs) was performed on 9 images consisting of the faces seen in Figure 7. In all nine cases, the faces were detected but the faces were recognized only in eight images. One face was not recognized and the reason was brightness. Although the CLAHE algorithm should solve this problem, in some cases it only makes the situation worse.

The next test performs face detection and recognition in the image that consists of two faces trained on the previously mentioned faces.

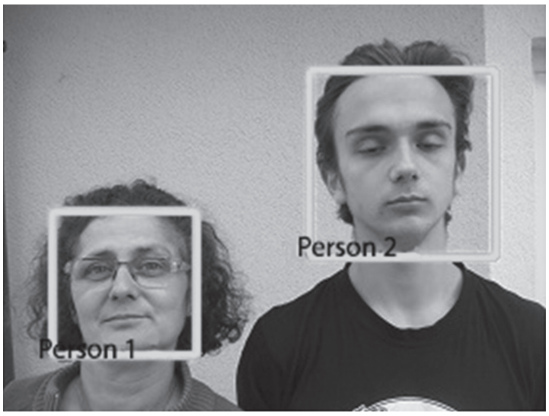


Fig. 8. Two-person image

The application successfully detected and recognized two persons in the image given in Figure 8. The photo was taken from the same position with the same brightness, at the same distance as the images from the training set. In these circumstances the application works flawlessly.

The ultimate experiment includes photos with three persons, Figure 9.

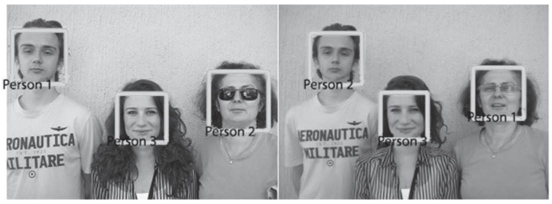


Fig. 9. Three-person image

The experiment was performed with seven different photographs with three persons from the database and the results are shown in Table 4.

Table 4. Face detection and recognition success with images with three persons

Image	Face detected (Yes/No)	Face recognized (Yes/No)	Face 3 detected/ recognized (Yes/No)
1	Yes/No	Yes/Yes	Yes/Yes
2	Yes/No	Yes/Yes	Yes/Yes
3	Yes/Yes	Yes/Yes	Yes/Yes
4	Yes/Yes	Yes/Yes	Yes/Yes
5	Yes/No	Yes/Yes	Yes/Yes
6	No/No	Yes/Yes	Yes/Yes
7	Yes/Yes	Yes/Yes	Yes/Yes

The application accuracy in case of a three-person image is slightly lower than in the previous cases but it is still quite high. The main reason for decreased accuracy is the difference in the distance from the camera between the images used for training and images used for detection and recognition.

5. CONCLUSION

In this paper, we researched face detection and recognition of photographed persons. The novel enhanced algorithm was tested on the image database *faces94*, on which it performed better than the unmodified Viola-Jones detection method. The method has been transferred to real-life photographs and tested to see what success it may yield on these images. The enhanced algorithm was tested to detect faces on images containing one, two and three persons.

The face detection rate referring to a one-person image is 100%, while due to discrepancy in the brightness of one image, the face recognition rate is 88.8%. When it comes to an image containing two persons, both the detection and the recognition rate are 100% for the current setup visible in Figure 8. In relation to an image containing three persons, detection and recognition for Faces 2 and 3 are 100%, whereas for Face 1 the detection rate is 85.7% and the recognition rate is 42.8% because of changing facial expressions and varying brightness.

The Eigenfaces method provides a practical solution that satisfies the basic needs for face detection and recognition. The method is quick, relatively simple and can be effective even in restricted environments, as used in this paper.

This enhanced Eigenfaces method proved that it is slightly better than the original Eigenfaces method due to CLAHE, contour highlighting, elliptic mask and other enhancements.

This application can be improved further by enhancing the mechanism for image alignment. This can be achieved by aligning the image by facial features (mouth, nose, eyes). It is also possible to improve detection by generating new images from the known images. New images would be generated by adding certain effects to older images, like facial rotation, axial symmetry, noise, etc. The idea of these processes is to increase the diversity of the training set to make it resistant to almost any possible situation.

6. REFERENCES:

[1] A. J. O'Toole, P. J. Phillips, F. Jiang, J. Ayyad, N. Penard, H. Abdi, "Face Recognition Algorithms Surpass Humans Matching Faces Over Changes in Illumination", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 9, Sep. 2007, pp. 1642–1646.

[2] C. Zhang, Z. Zhang, A survey of recent advances in face detection. Technical report, Microsoft Research, 2010.

[3] P. Viola, M. J. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision, vol. 57, no. 2, May 2004, pp. 137–154.

[4] S. Jeyaraman, M. Arun Kumar, S. Valarmathy, S. Deepak, S. Gandhiraj, "Face Recognition Using Viola-Jones Detector", International Journal of Advanced Research in Electronics and Communication Engineering, vol. 3, no. 11, Nov. 2014, pp. 1514–1517.

[5] Y.-Q. Wang, "An Analysis of the Viola-Jones Face Detection Algorithm", Image Processing On Line, vol. 4, Jun. 2014, pp. 128–148.

[6] M.-H. Yang, D. J. Kriegman, N. Ahuja, "Detecting faces in images: A survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, 2002, pp. 34–58.

[7] B. Amberg, A. Blake, A. Fitzgibbon, S. Romdhani, T. Vetter, "Reconstructing High Quality Face-Surfaces using Model Based Stereo", IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 2007, pp. 1–8.

[8] Y. Wang, J. Liu, and X. Tang, "Robust 3D Face Recognition by Local Shape Difference Boosting", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 10, Oct. 2010, pp. 1858–1870.

[9] M. A. Turk, A. P. Pentland, "Face recognition using eigenfaces", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Maui, HI, USA, 1991, pp. 586–591.

[10] M. Turk, A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, vol. 3, no. 1, Jan. 1991, pp. 71–86.

[11] W. Zhao, A. Krishnaswamy, R. Chellappa, D. L. Swets, and J. Weng, "Discriminant Analysis of Principal Components for Face Recognition", Face Recognition, pp. 73–85, Springer Berlin Heidelberg, 1998.

[12] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, Y. Ma, "Toward a practical face recognition system: robust alignment and illumination by sparse representation", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 2, Feb. 2012, pp. 372–386.

[13] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, vol. 1, pp. 511-518.

[14] F. C. Crow, "Summed-area Tables for Texture Mapping", Proceedings of the 11th Annual Conference on Computer Graphics and Interactive Techniques, New York, NY, USA, 1984, pp. 207–212.

[15] Y. Freund, R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", Journal of Computer and System Sciences, vol. 55, no. 1, Aug. 1997, pp. 119–139.

[16] Y. Freund, R. Schapire, N. Abe, "A short introduction to boosting", Journal-Japanese Society For Artificial Intelligence, vol. 14, no. 5, 1999, pp. 771–780.

[17] T. Heseltine, N. Pears, J. Austin, "Evaluation of Image Pre-Processing Techniques for Eigenface Based Face Recognition", The Proceedings of the Second International Conference on Image and Graphics, 2002, pp. 677–685.

[18] Libor Spacek, "Collection of Facial Images: Faces94", Available: http://cswww.essex.ac.uk/ mv/allfaces/faces94.html (accessed: 2015).