Combination of Face Classifiers for Person Identification

Bernard Achermann and Horst Bunke
Institut für Informatik und angewandte Mathematik
Universität Bern, Neubrückstrasse 10, CH-3012 Bern
email: ackerman,bunke@iam.unibe.ch

Abstract

In this paper we present a system for person identification which is based on the combination of several face classifiers. First, we shortly describe the individual face classifiers. Then, classifier combination is discussed. Finally, the impact of the combination is presented in a series of practical experiments which show a significant improvement of the recognition rate.

1. Introduction

During the last years the research activities in the area of face recognition have grown significantly. From the global point of view, two main directions can be distinguished, namely, the analysis of images of full faces and of profiles. In the survey papers of Samal/Iyengar [11] and Chellappa/Wilson/Sirohey [4] the different approaches to face recognition are exhaustively described. Despite the progress that has been achieved, there is still more research required in order to improve recognition accuracy and to deal with face images taken under more realistic conditions, because most of the approaches known until now rely on quite severe constraints regarding lighting, background, distance of the subject from the camera, head position and orientation, and others.

Another area of pattern recognition that has received significant attention recently is classifier combination and sensor fusion (Waltz/Llinas [14], Hall [6], Dasarathy [5]). The aim is to get more reliable and better recognition results by combining the outputs of different classifiers. Thus, more information is involved in the final decision process. Classifier combination and sensor fusion have been applied in the areas of optical character recognition and shown very promising results (e.g., Ho/Hull/Srihari [7], Lam/Suen [8] and Xu/Krzyzak/Suen [15]). For identification purposes (e.g., security systems) it is important that the decisions taken by the system are very reliable. Therefore, classifier combination has been proposed also in this area. For example, Brunelli/Falavigna/Poggio/Stringa [2, 3] combine face and speech recognition for person identification. In this paper we describe a different approach that combines full face and profile views of humans. In Figure 1 a graphical representation of our system architecture is shown.

In the remaining part of this paper we shortly introduce the individual classifiers. Then, their combination is described in greater detail. Next, experimental results are presented. Finally, some conclusions are given.

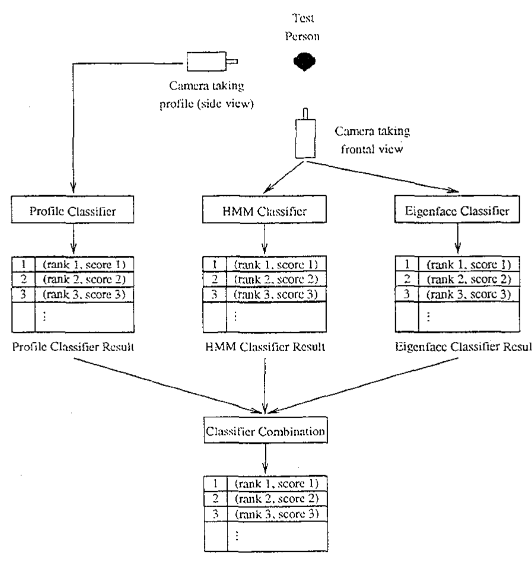

Figure 1. System architecture

1015-4651/96 $5.00 © 1996 IEEE
Proceedings of ICPR '96

2. Frontal View Classification

2.1. The Eigenface Classifier

This full face classifier is an implementation of the eigenface approach proposed by Turk and Pentland. For a detailed description of this method refer to [13].

2.2. The HMM Classifier

This classifier for full face images is based on hidden Markov models (HMMs). The approach we used is similar to the method of Samaria proposed in [12] and is in detail described in Nyffenegger's thesis [9]. An introduction to hidden Markov models is presented by Rabiner in [10].

Generally, HMMs are working on one-dimensional signals or feature vectors. Images instead contain two-dimensional information. In order to make HMMs applicable to images, we reduce the image information to one-dimensional vectors by applying a sliding window. This window is moving from the top of the image to the bottom and covers the whole width of the image. The step size is chosen so that two successive windows have a certain overlap. The reason for the overlap is to avoid the cutting of significant face features and bring some context information into the process. The intensity values in the sliding window are given (after some preprocessing steps) as feature vector to the HMM. In the experiments described in Section 5, the size of each image is normalized to a width of 40 pixels in the preprocessing phase. The height of the images is adjusted proportionally and, therefore, may vary. One window is of dimension 40×4, and the overlap between successive windows is 3 pixels.

The idea underlying this method is the following. Intuitively, a human face consists of a number of regions like forehead, eyes, nose, mouth, chin and so on. These regions remain identifiable for a human observer, even when the face image is cut into sliding windows as described above. With the HMM we try to make use of this property. A human face is represented by a linear left-right model consisting of five states as shown in Figure 2. These states correspond to the face parts ("1" for the forehead, "2" for the eyes, "3" for the nose, "4" for the mouth and "5" for the chin).

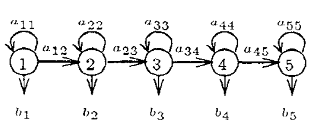

Figure 2. Linear left-right human face model

For every person in the database the parameters of the hidden Markov model are calculated in a training phase (cf. [10]). If a test image is presented, the probability of producing this image is computed by means of the Viterbi algorithm (cf. [10]) for every person (i.e., model) in the database. The classifier returns a ranking of the possible persons in ascending order of the score for each model. The score s is computed in dependence of the probabilities p of the models: $s = -2 \cdot \log p$.

3. Profile Classification

This classifier is based on the profile shape of human faces. More details of this method are described in our paper [16].

4. Classifier Combination

Generally, the aim of classifier combination is to combine the results of m individual classifiers $C_i (i = 1, \ldots, m)$ by means of a combination classifier K. The combiner is nothing else but an additional classifier taking the output of the classifiers C_i as input features, whereas the classifiers C_i are based on the data of the sensors, i.e., face images. There are several general strategies for classifier combination ([14, 6]):

1. The combination may be achieved by taking into account exclusively the top decision of each individual classifier.
2. The combination may depend on the ranks of the classes that are assigned by the individual classifiers.
3. The combination may be done by using the information of the score function of the involved classifiers.

If one wants to deal with the values of the score function of the individual classifiers, a transformation of the scores is required in most cases, because the values of these score functions may vary much. (They may represent, for example, Euclidean distances, probabilities a.s.o.) Therefore, it may be necessary to transform them to the same range of values. We implemented the following transformation schemes.

A *linear transformation* maps the interval I_s into $I_{s'}$ by interpolation:

$$s' = s'_{min} + \frac{s - s_{min}}{s_{max} - s_{min}} (s'_{max} - s'_{min}) \qquad (1)$$

Under a *logarithmic transformation* the score values are first linearly transformed into $I_{s''} = [0.0, 100.0]$ (normalization). Then the transformed score is computed as:

$$s' = \log(1 + s'') \quad (2)$$

The *exponential transformation* is defined similarly. After the normalization step the score is calculated as:

$$s' = \exp(s'') - 1 \quad (3)$$

A *logistic transformation* is done in two steps. First, a linear transformation to the interval $[0.0, 100.0]$ is computed. Then the new score is obtained by

$$s' = \frac{\exp(\alpha + \beta s)}{1 + \exp(\alpha + \beta s)} \quad (4)$$

The parameters α and β have to be determined empirically; α defines where the function intersects the x axis and β controls the slope.

In the following, the combination techniques that were implemented for our face recognition task are described.

- **Voting:** Here each classifier is regarded as an expert with one vote, i.e., it votes for its top choice. The resulting decision is depending on the majority of the votes. If the voting ends in a draw, the combination classifier is unable to decide for a certain class and rejects the input.

- **Ranking:** The simplest approach is to compute the sum of the ranks for every class in the combination set. The class with the lowest rank sum will be the first choice of the combination classifier.

- **Scoring:** If the score functions are directly comparable or if there exists a suitable transformation to make the involved classifiers comparable (as outlined above), score based strategies are a good way for classifier combination. Here, the score functions provide the combiner with some additional information which is not available in the voting and the rank based methods described above. Of course, the quality of the score function is of eminent importance in this case.

  In our system, we combine classifiers by simply computing the sum of the score functions. Ranking is done according to ascending scores.

5. Experimental Results

In the following sections we briefly describe our experiments. An exhaustive description of our experimental work may be found in [1].

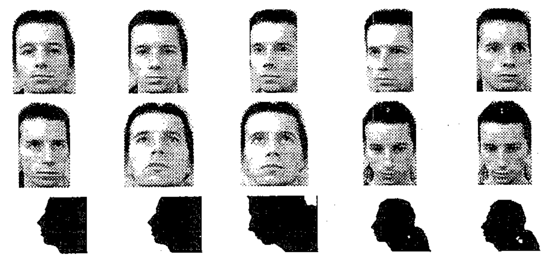

Figure 3. Example of one person in the database (full face images in the upper two rows, binarized profile images in the last row)

5.1. Data Collection

Since it was our intention to combine classifiers working with different views of human faces, we collected a small database which is divided into two parts:

1. **Frontal images**

   Included are 10 grey level images of each person with varying head positions (2 looking straight into the camera, 2 looking to the left, 2 looking to the right, 2 downwards, and 2 upwards). All images (original size 512×342) are cut such that exclusively the face part remains in the image, and they are normalized to a width of 40 pixels. The greylevel resolution is 8 bit.

2. **Profile images**

   We took 5 images of each person with varying size and orientation of the head. All images are of size 512×342 and binarized.

The lighting conditions during image acquisition were carefully controlled. The database contains images of 30 persons. An example of one person is shown in Figure 3.

The database was divided into a training and a test set. Those sets are disjoint. There exist five training and five test images per person for the frontal face classifiers, and three training and two test images for the profile classifier. Each profile test image of a person can be combined with each full face test image of the same individual. Therefore, there exist totally $30 \cdot 5 \cdot 2 = 300$ possible input combinations for the combiner.

All images used in our experiments can be obtained via anonymous ftp at `iamftp.unibe.ch` under `/pub/Images/FaceImages`.

5.2. Results of the Individual Face Classifiers

The results of the individual face classifiers are shown in Figure 4. We observe that the eigenface and

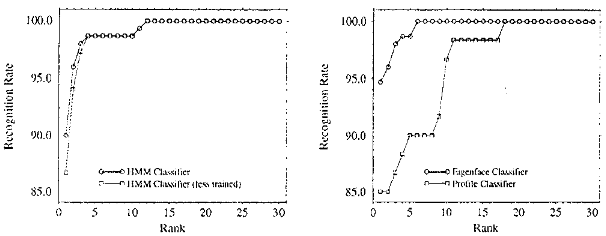

Figure 4. Recognition rates of the individual face classifiers

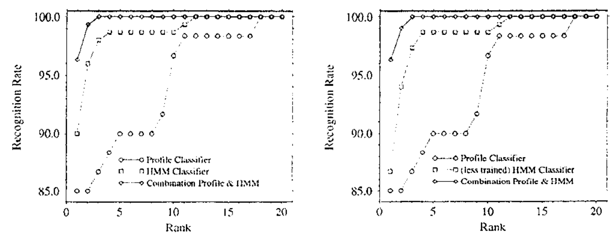

Figure 5. Recognition rates for combinations by score summation (logarithmic transformation): combination profile and HMM classifier (left side), combination profile and less trained HMM classifier (right side)

the HMM classifier yield very good recognition results (94.7% and 90.0% on the first rank), whereas the profile classifier reaches a recognition rate of 85.0%. This is not surprising, since a person's profile obviously contains less information than a frontal view.

During the training phase of the HMM classifier all training images are presented several times to the classifier. The number of such iterations has an impact on the recognition rate. Thus, there is an optimal number of training iterations. If less training iterations are executed the recognition rate deteriorates. In order to study the potential of classifier combination more thoroughly, we also used a suboptimally, i.e., less trained HMM classifier in addition to the optimally trained one. This classifier achieves a recognition accuracy of 86.7% on the first rank.

5.3. Combination of the Profile and the HMM Classifier

In our first experiment we combined the results of the profile classifier with the HMM classifier. The results are shown in Figure 5 (left part). The combination technique applied in this and all of the following experiments is score summation after a logarithmic transformation of the score values. This technique generally yielded better results than linear, exponential and logistic transformations combined with voting or ranking. On the first rank, we got a recognition rate of 96.3%, which is a significant improvement over the performance of the individual face classifiers (90.0% for the HMM classifier and 85.0% for the profile classifier).

The voting techniques and rank summation yielded no or only slight improvements over the individual classifiers. In most cases the voting techniques even resulted in a lower recognition rate than the single classifiers due to a large number of rejections.

In the second experiment we combined the results of the profile classifier with those of the less trained HMM classifier. The aim was to investigate how the combination scheme performs on classifiers with lower recognition rates (85.0% for the profile and 86.7% for the less trained HMM classifier). The result is shown in Figure 5 (right part). The performance proved to be very good (96.3%), actually identical on the first rank to the experiment with the optimally trained HMM classifier.

5.4. Combination of the Profile and the Eigenface Classifier

In the next experiment we combined the profile and the eigenface classifier. A graphical representation of the results is shown in Figure 6 (left part). The improvement of the recognition rate due to classifier combination that was observed in the previous experiments is also true for this case. The recognition rate of 97.7% on the first rank is even higher than in the first and second experiment.

We also ran some additional tests with other combination schemes. And again, the score summation clearly proved to be the best method. The other techniques (voting and ranking) yielded only marginal or even no improvement over the individual classifiers.

5.5. Combination of the HMM and the Eigenface Classifier

With this experiment we investigated the combination of the HMM and the eigenface classifier. In contrast with the experiments described before, it is a combination of classifiers working on only one single information source. I.e., this combination scheme relies exclusively on images of frontal faces. The result is shown in Figure 6 (right part). Though the HMM classifier (90.0%) and the eigenface classifier (94.7%) have high individual recognition rates already, the combination of both classifiers achieves a further significant improvement (98.7% on the first rank).

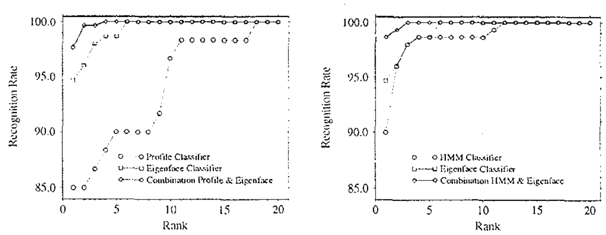

Figure 6. Recognition rates for combinations by score summation (logarithmic transformation): combination profile and eigenface classifier (left side), combination HMM and eigenface classifier (right side)

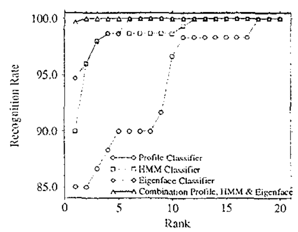

Figure 7. Recognition rate for the combination of profile, HMM and eigenface classifier by score summation (logarithmic transformation)

Among the other techniques, the rank summation also showed good results in this experiment, whereas the voting strategies are again inferior.

5.6. Combination of the Profile, the HMM and the Eigenface Classifier

In this experiment, finally, we combined all available face classifiers. The result is presented in Figure 7. With 99.7% we got the best recognition rate of all experiments. Obviously, the additional information obtained from three classifiers instead of only two makes the recognition even more reliable.

The same trend was observed when other combination techniques were applied, i.e., three individual classifiers give better results than two. If three classifiers are integrated even the simple voting techniques perform very well.

6. Conclusions

In this paper we have proposed to combine several individual classifiers in order to improve the recognition accuracy of face and person identification. Generally, different classifiers have their particular advantages and disadvantages. Therefore, a synergetic effect may be expected by a combination if the weak and the strong properties of each individual classifier are appropriately taken into account. In a number of experiments, we have demonstrated that even a simple classifier combination concept such as score summation can lead to a significant improvement of classification accuracy.

References

[1] B. Achermann and H. Bunke. Combination of Classifiers on the Decision Level for Face Recognition. Tech. Rep. IAM-96-002, IAM, Universität Bern, Jan. 1996.

[2] R. Brunelli and D. Falavigna. Person Identification Using Multiple Cues. *IEEE T-PAMI*, 17(10):955–966, Oct. 1995.

[3] R. Brunelli, D. Falavigna, T. Poggio, and L. Stringa. Automatic Person Recognition by Acoustic and Geometric Features. *Machine Vision and Applications*, 8:317–325, 1995.

[4] R. Chellappa, C. Wilson, and S. Sirohey. Human and Machine Recognition of Faces: A Survey. *Proc. of the IEEE*, 83(5):704–740, May 1995.

[5] B. Dasarathy. *Decision Fusion.* IEEE Computer Society Press, 1994.

[6] D. Hall. *Mathematical Techniques in Multi-Sensor Data Fusion.* Artech House, Inc., 1992.

[7] T. Ho, J. Hull, and S. Srihari. Decision Combination in Multiple Classifier Systems. *IEEE T-PAMI*, 16(1):66–75, Jan. 1994.

[8] L. Lam and C. Suen. Optimal Combinations of Pattern Classifiers. *Pattern Recognition Letters*, 16:945–954, 1995.

[9] C. Nyffenegger. Gesichtserkennung mit Hidden-Markov-Modellen. Master's thesis, IAM, Universität Bern, Sept. 1995.

[10] L. Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proc. of the IEEE*, 77(2):257–286, Feb. 1989.

[11] A. Samal and P. Iyengar. Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey. *Pattern Recognition*, 25(1):65–77, 1992.

[12] F. Samaria and S. Young. HMM-Based Architecture for Face Identification. *Image and Vision Computing*, 12(8):537–543, Oct. 1994.

[13] M. Turk and A. Pentland. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[14] E. Waltz and J. Llinas. *Multisensor Data Fusion.* Artech House, Inc., 1990.

[15] L. Xu, A. Krzyzak, and C. Suen. Methods of Combining Multiple Classifiers and their Application to Handwritten Numeral Recognition. *IEEE T-SMC*, 22(3):418–435, 1992.

[16] K. Yu, X.-Y. Jiang, and H. Bunke. Face Recognition by Facial Profile Analysis. In M. Bichsel, editor, *Proc. Int. Workshop on Automatic Face and Gesture Recognition, Zürich*, pages 208–213, June 1995.