Recognizing Faces from the Eyes Only

E. Hjelmås
Department of Informatics
University of Oslo
P.O. Box 1080 Blindern, N-0316 Oslo, Norway
Department of Electrical Engineering and Science
Gjøvik College
P.O. Box 191, N-2800 Gjøvik, Norway

J. Wroldsen
Department of Electrical Engineering and Science
Gjøvik College
P.O. Box 191, N-2800 Gjøvik, Norway

Abstract

The eyes are one of the most important facial features for recognizing human faces. Many face recognition systems today make use of local features (such as eyes) for identification or verification of individuals, but no system to our knowledge has studied performance when the only available information is the eyes. In this paper we show that we can obtain 85% correct classification on the popular ORL face database, when the features are extracted from the eye area only. We compare feature extraction from eigenfeatures and Gabor wavelets with features consisting of simple gray-level pixel values.

1 Introduction

Face recognition is a complex task which has recieved a great amount of attention in recent years, mostly due to its wide range of application in the area of biometric systems. Many different approaches have been proposed, and current systems exhibit very good performance on detection, identification and verification of human faces [10, 2]. However, one of the remaining problems is dealing with changes in faces over time. Since eyes are one feature of the face which are not greatly affected by certain typical face changes (e.g. facial hair like beard or mustasch), we address the problem of identifying subjects from the eyes only.

There is a significant amount of evidence from psychophysics which support the theory that humans make use of local features (such as the eyes) when recognizing other individuals [1]. And indeed, several current face recognition systems are based on local features as well as global (holistic) features. The eigenface approach, developed by Turk et al. [12], was further developed in [8] to include eigenfeatures such as eigeneyes, eigennose and eigenmouth. This improved performance of the system, and result were also surprisingly good when only the eigenfeatures were used. A very different system for object recogniton was developed by Lades et al. [5] and applied to face recognition in [13, 9]. This system is based entirely on local features, computed by a biologically motivated [3] wavelet transform. These systems are two of the best face recogniton systems today, with different advantages/disadvantages.

In this paper we investigate how well human faces can be recognized when the features are extracted from the eyes only. We explore performance with the eigenfeature technique [8], which is applicable for both local (eyes, nose, mouth, etc.) and global (entire face) feature extraction, and the gabor wavelet approach [5], which is a typical local feature extractor. We compare these techniques with classification directly from gray-level values. Recognizing faces from just the eyes is important work, because in many situation the eyes can be the most reliable (and perhaps the only) source of information. Our experimental results show surprisingly good performance on a small, but difficult dataset.

2 Dataset and Preprocessing

We use the face database from Olivetti Research Lab. This database has been widely used for testing face recognition systems, and the best reported result to our knowledge is 96% correct classification [6], when the database is divided into two equally large sets (for training and testing). The database contains 10 images of 40 subjects for a total of 400 images. The images are gray-level of size 92×112. We have manually located the center of the eyes in all the images, and then extracted

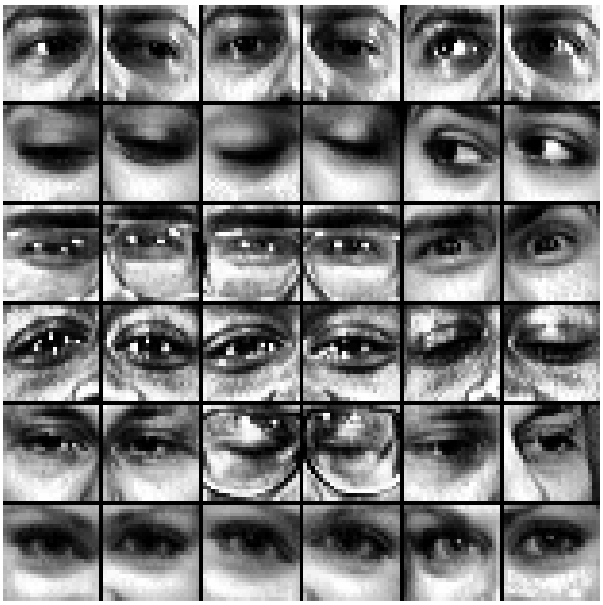

Figure 1: Examples of subtracted histogram equalized windows from the face image (same subject on each horizontal line).

a 25×25 window at this region (for both eyes). If this subwindow includes an area outside the original face image, we pad the image with the mean of the gray-level values in the full face image. The extracted eye-images show considerable variation with respect to open/closed eyes, expression, glasses (reflection in glasses) and gaze, making this a challenging dataset. Some example eye-images are shown in figure 1.

Before feature extraction, all images were histogram equalized. As an additional preprocessor for the eigenfeature approach, the mean image was computed from the training set and subtracted from all the training and testing images.

3 Feature Extraction

We only briefly outline the techniques we use here, since they are standard computer vision techniques and more appropriatly described in papers such as [11, 12, 5, 13]

3.1 Eigenfeatures

We do a principal component analysis of the matrix of vectorized training samples, separate for right eye and left eye. The resulting principal components (eigeneyes) span a subspace [4] which is a good representation for right and left eye images. Examples of these eigeneyes are given in figures 2 and 3, ordered according to eigenvalues. One important notice about this technique is that its aim is representation and not discrimination.

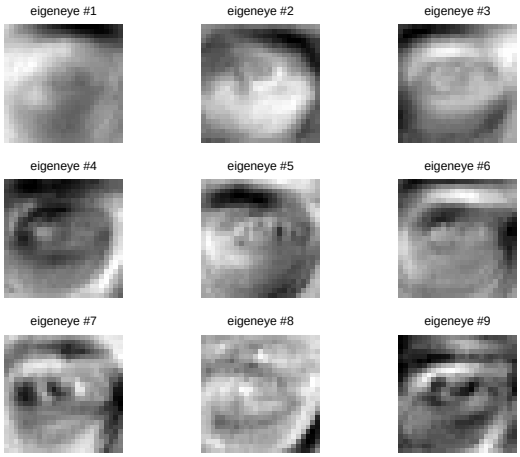

Figure 2: Right eye eigeneyes.

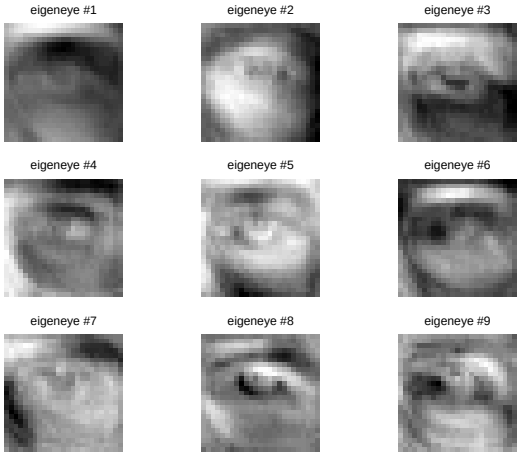

Figure 3: Left eye eigeneyes.

Thus, we should act with care when selecting the number of eigeneyes to keep for spanning the subspace (eigeneyes with small corresponding eigenvalues might still be important for discrimination).

One limitation which also needs to be pointed out in the way we apply eigeneyes in this paper, is that the training set is relatively small. Ideally, we would have a large separate set of eyes to compute a more general subspace of eigeneyes. We would also expect a degradation in performance due to the large variation in the dataset. We normalize the images by keeping the center point fixed, which gives us some invariance with respect to positioning of the eye, but a better way would perhaps be to locate the corners of the eye so we might be able to normalize with respect to rotation and size in addition to position.

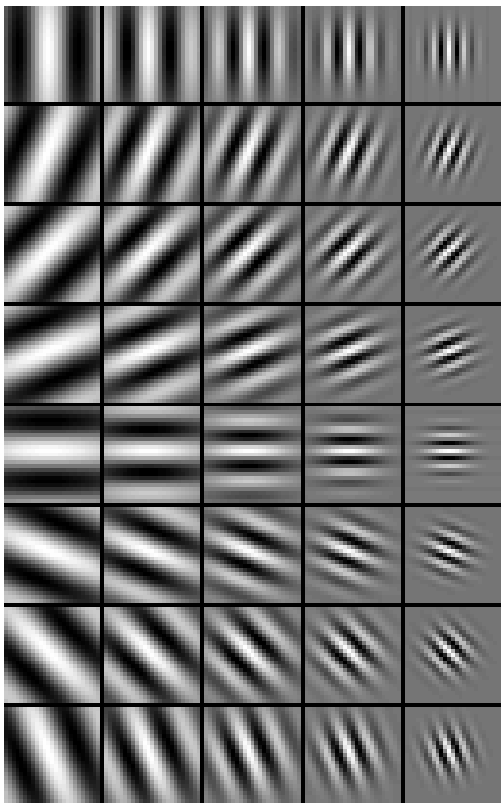

Figure 4: The Gabor wavelets.

3.2 Gabor Wavelets

Similar to the system of Lades et al. [5], we apply a wavelet transform based on the Gabor kernel

$$\psi_j(\vec{x}) = \frac{k_j^2}{\sigma^2} e^{-\frac{k_j^2 x^2}{2\sigma^2}} \left(e^{i\vec{k}_j \vec{x}} - e^{-\frac{\sigma^2}{2}} \right), \tag{1}$$

where

$$\vec{k}_j = \begin{pmatrix} k_\nu \cos\phi_\mu \\ k_\nu \sin\phi_\mu \end{pmatrix}, \quad k_\nu = 2^{-\frac{\nu+2}{2}}\pi, \ \phi_\mu = \mu\frac{\pi}{8}. \tag{2}$$

All the Gabor wavelets are created from this kernel by dilation and rotation. The 40 wavelets created from indices $\nu \in \{0, ..., 4\}$ (size) and $\mu \in \{0, ..., 7\}$ (orientation) are shown in figure 4. These wavelets are convolved with the image, and we keep the value of the center pixel. This provides us with a feature vector of 80 complex coefficients, but we only keep the magnitudes for classification.

4 Experimental Results and Discussion

For classification we apply the *nearest-neighbour* (NN) and the *minimum distance* (to class mean) (MD) classifier. As a measure of distance in feature space, a natural choice is the usual *Euclidian* metric (E)

$$d_e(\vec{x}, \vec{y}) = \sqrt{\sum_i (x_i - y_i)^2}. \tag{3}$$

However, for the eigenfeatures there is also a natural choice of metric in the *weighted Euclidian* metric (WE)

$$d_{we}(\vec{x}, \vec{y}) = \sqrt{\sum_i \lambda_i^{-1} (x_i - y_i)^2}. \tag{4}$$

since our goal is discrimination rather than representation. The weights for this metric is the inverse of the sequence of the eigenvalues, which means that if we consider the feature vectors in feature space as belonging to the same class, this metric is equivalent to the *Mahalanobis distance*. Since good results have also been obtained with other metrics [7, 10], we also use the *angle between vectors* metric (ABC)

$$d_{abc}(\vec{x}, \vec{y}) = cos^{-1}\left(\frac{\vec{x} \cdot \vec{y}}{||\vec{x}||||\vec{y}||} \right). \tag{5}$$

and the *city block distance* metric (CBD)

$$d_{cbd}(\vec{x}, \vec{y}) = \sum_i |x_i - y_i|. \tag{6}$$

Table 1 and 2 show the results from the nearest neighbour and minimum distance classifiers. We compare the feature extraction methods mentioned earlier (eigenfeatures (PCA) and Gabor), with simply using gray-levels (GL), either pixel values or pixel values after resizing the image (12×12 or 6×6). In these cases the feature vectors just consist of the pixel values of the vectorized image, e.g. for the simplest case (no resizing) we have a 625-dimensional feature vector (from a 25×25 image). All the scores are percentage correct classification averaged over 10 trials, when the dataset of 400 images is randomly divided into 2 equally sized sets (5 images for training and 5 images for testing for each subject). This methodology is equivalent to the approaches usually taken when working with the ORL database, which allows us to compare with other systems.

We observe from the experiments that we get surprisingly good results from just the eyes compared to using the whole face. The best reported results when using the whole face is significantly better of course, but from the sample images shown in figure 1 (observe the large variation), 85% correct classification is very satisfying. A bit surprising is that using simple gray-level values as features yields competitive performance. However, similar results for using the nearest neighbour on gray-level values from the entire face have been reporteded earlier [6]. We should take this as an indication that we

need further testing, specifically on a larger dataset, to validate our results.

The best results are obtained with the Gabor feature extraction method. However, we should note that there was a large degree of variation in the 10 trials (due to the process of randomly dividing the set into training and test set), so there is a question of the statistical significance of these results. But since the overall 3 best results all are from Gabor, we consider this a strong indication that this method is the best compared to our implementations of the eigenfeatures and gray-levels here.

As mentioned earlier, it is difficult to select the number of eigeneyes to use for spanning the subspace. We want to keep our feature space as low-dimensional as possible, which would initially lead us to discard some eigeneyes (and perhaps the ones with the smallest corresponding eigenvalues). We see from our results that this would lead to sub-optimal performance, since the best result obtained in our experiments when using this technique is when we keep all the 200 eigeneyes (84.4% with MD classifier and CBD metric).

It is also noteworthy to mention that there is a significant difference in performance for the different metrics. We cannot derive any general rule-of-thumb for this, but we can note that the weighted euclidian metric did poorly for large feature space with the nearest neighbour classifier. It is also the worst metric for the highest dimensions (PCA 100 and 200) with the minimum distance classifier. Both these cases is probably due to inaccurate estimated eigenvalues for the last principal components (eigeneyes), which is caused by a too small training set (from which the principal components were computed). Otherwise, the general observation is that one metric is clearly to prefer instead of another (which is typical for pattern recognition applications).

All the results we have presented here can be considered as preliminary since there are many potential parameter adjustments and classification strategies we can apply. For the eigenfeature case, we could either compute a subspace from the entire eye-area instead of separate eyes or we could try further normalization as mentioned earlier. There is no reason to believe that we are using the optimal Gabor wavelets either. We could further experiment with both size and rotation. Future work might also include looking into some strategy such as classifying from just one eye, and select that eye based upon some similarity criteria (to yield further robustness in cases where one eye might be completely covered).

5 Conclusion

We have presented in this paper an experimental study of recognizing faces from the eyes only. Performance for this approach is not as good as when using the entire face, but considering the large variation in the dataset, the results are good. Recognizing faces from the eyes only is important since the eyes are a central feature of the human face, and sometimes the only source of information in face recogniton applications.

Table 1: Results for the nearest neighbour classifier. The number after PCA indicates how many eigeneyes we keep (e.g. for "PCA 50" we keep the 50 eigeneyes with largest corresponding eigenvalues)

Method	NN			
	d_e	d_{we}	d_{abc}	d_{cbd}
GL 25×25	77.0	-	81.1	82.6
GL 12×12	80.8	-	82.2	83.4
GL 6×6	78.6	-	78.9	80.2
PCA 10	70.5	72.0	70.9	71.4
PCA 50	75.9	63.8	80.3	76.6
PCA 100	75.1	51.4	79.6	72.7
PCA 200	77.4	58.4	80.8	75.1
Gabor	84.6	-	84.6	85.2

Table 2: Results for the minimum distance classifier.

Method	MD			
	d_e	d_{we}	d_{abc}	d_{cbd}
GL 25×25	76.2	-	75.7	77.0
GL 12×12	76.7	-	76.5	77.3
GL 6×6	70.7	-	70.6	70.8
PCA 10	64.1	68.2	62.6	65.8
PCA 50	74.9	78.3	74.7	79.6
PCA 100	76.4	75.0	75.6	81.8
PCA 200	76.2	71.8	75.5	84.4
Gabor	77.4	-	76.7	74.5

Acknowledgments

We would like to thank the Olivetti Research Laboratory for compiling and maintaining the ORL face-database, Hermann Lia for useful discussions during the preparation of this work, and Wenyi Zhao for helpful comments on this manuscript. This work was supported by the Research Council of Norway and was done while E. Hjelmås was visiting Center for Automation Research, University of Maryland.

References

[1] V. Bruce. *Recognizing Faces.* Lawrence Erlbaum Associates, London, 1988.

[2] R. Chellappa, C. L. Wilson, and S. Sirohey. Human and Machine Recognition of Faces: A Survey. *Proceedings of the IEEE*, 83(5), 1995.

[3] R. L. De Valois and K. K. De Valois. *Spatial Vision*. Oxford University Press, New York, 1988.

[4] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, London, 2nd edition, 1990.

[5] M. Lades, J. C. Vorbrggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wrtz, and W. Konen. Distortion Invariant Object Recognition in the Dynamic Link Architecture. *IEEE Transactions on Computers*, 42(3), 1993.

[6] S. Lucas. Face Recognition with the Continuous N-tuple Classifier. In *Proceedings of the British Machine Vision Conference*, 1997.

[7] S. Lucas. The Continuous N-tuple Classifier and its Application to Face Recognition. *Electronics Letters*, 33:1676–1678, 1998.

[8] B. Moghaddam and A. Pentland. Face Recognition using View-Based and Modular Eigenspaces. In *Automatic Systems for the Identification of Humans, SPIE*, volume 2277, 1994.

[9] K. Okada, J. Steffens, T. Maurer, H. Hong, E. Elagin, H. Neven, and C. von der Malsburg. The Bochum/USC Face Recognition System and How it Fared in the FERET Phase III Test. In *Face Recognition: From Theory to Applications*. Springer-Verlag, 1998.

[10] P. J. Phillips, H. Moon, P. Rauss, and S. A. Rizvi. The FERET September 1996 Database and Evaluation Procedure. In *Proceedings of the First International Conference on Audio and Video-based Biometric Person Authentication*, 1997.

[11] L. Sirovich and M. Kirby. Low-Dimensional Procedure for the Characterization of Human Faces. *Journal of the Optical Society of America*, 4:519–524, 1987.

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

[13] L. Wiskott, J. M. Fellous, N. Krüger, and C. von der Malsburg. Face Recognition and Gender Determination. In *Proceedings International Workshop on Automatic Face- and Gesture-Recognition*, 1995.