Face Recognition by Kernel Independent Component Analysis

T. Martiriggiano, M. Leo, T. D'Orazio, A. Distante,
CNR- ISSIA via Amendoda 122/D-I

70126 BARI ITALY

{leo,dorazio,martiggiano,distante}@issia.ba.cnr.it

Abtract. In this paper, we introduce a new feature representation method for face recognition. The proposed method, referred as Kernel ICA, combines the strengths of the Kernel and Independent Component Analysis approaches. For performing Kernel ICA, we employ an algorithm developed by F. R. Bach and M. I. Jordan. This algorithm has proven successful for separating randomly mixed auditory signals, but it has never been applied on bidimensional signals such as images. We compare the performance of Kernel ICA with classical algorithms such as PCA and ICA within the context of appearance-based face recognition problem using the FERET database. Experimental results show that both Kernel ICA and ICA representations are superior to representations based on PCA for recognizing faces across days and changes in expressions.

1 Introduction

Face recognition has become one of most important biometrics technologies during the past 20 years. It has a wide range of applications such as identity authentication, access control, and surveillance.

Human face image appearance has potentially very large intra-subject variations due to 3D head pose, illumination, facial expression, occlusion due to other objects or accessories (e.g., sunglasses, scarf, ect.), facial hair, and aging. On the other hand, the inter-subject variations are small due to the similarity of individual appearances. This makes face recognition a great challenge. Two issues are central: 1) what features to use to represent a face and 2) how to classify a new face image based on the chosen representation. This work focuses on the issue of feature selection. The main objective is to find techniques that can introduce low-dimensional feature representation of face objects with enhanced discriminatory power. Among various solutions to the problem (see [1] for a survey), the most successful are the appearance-based approaches, which generally operate directly on images or appearances of face objects.

Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Independent Component Analysis (ICA) are three powerful tools largely used for data reduction and feature extraction in the appearance-based approaches [2] [3] [4].

Although successful in many cases, linear methods fail to deliver good performance when face patterns are subject to large variations due to 3D head pose, illumination, facial expression, and aging. The limited success of these methods should be attributed to their linear nature. As a result, it is reasonable to assume that a better solution to this inherent nonlinear problem could be achieved using non linear methods, such as the so-called kernel machine techniques [5].

Yang [6], Kim *et al*. [7] investigated the use of Kernel PCA and Kernel LDA for learning low dimensional representations for face recognition. Experimental results showed that kernel methods provided better representations and achieved lower error rates for face recognition.

2 Overview of present work

In this paper, motivated by the success that ICA, Kernel PCA and Kernel DLA have in face recognition, we investigate the use of Kernel Independent Component Analysis (Kernel ICA) for face recognition. Kernel ICA combines the strengths of the Kernel and ICA approaches. Here, we employ an algorithm developed by F. R. Bach and M. I. Jordan [8]. This algorithm has proven successful for separating randomly mixed auditory signals. We use Kernel ICA to find a representation in which the coefficients used to code images are statistically independent, i.e., a factorial face code. Barlow and Atick discussed advantages of factorial codes for encoding complex objects that are characterized by high-order combinations of features [9], [10].

3 Experimental results

The face images employed for this research are a subset of the FERET face database. The FERET dataset contain images of 38 individuals. There are four frontal views of each individual: a neutral expression and a change of expression from one session, and a neutral expression and change of expression from a second session that occurred three weeks after the first. Examples of the four views are shown in fig. 1.

Fig. 1. Example from the FERET database of the four frontal image viewing conditions: neutral expression and change of expression from session 1; neutral expression and change of expression from session 2. Reprinted with permission from Jonathan Phillips.

The algorithms are trained on a single frontal view of each individual. The training set is comprised of 50% neutral expression images and 50% change of expression images. The algorithms are tested for recognition under three different conditions:

same session, different expression (Test Set 1); different day, same expression (Test Set 2); and different day, different expression (Test Set 3).

Face recognition performance is evaluated by the nearest neighbor algorithm. Coefficient vectors b in each test set were assigned the class label of the coefficient vector in the training set that was most similar as evaluated by the Euclidean distance measure δ_{euc} and the cosine similarity measure $\delta_{\cos}$, which are defined as follows:

$$\delta_{euc}(b_{test}, b_{train}) = \sqrt{\sum_i \left(b_{test_i} - b_{train_i}\right)^2}\ ,\ \delta_{\cos}(b_{test}, b_{train}) = \frac{-b_{test}{}^{T} \cdot b_{train}}{\|b_{test}\|\|b_{train}\|}$$

where $\|.\|$ denotes the norm operator.

Figures 2 and 3 report the face recognition performances with the Kernel ICA, ICA factorial code representations (for performing ICA, we employ the FastICA algorithm developed by A. Hyvärinen [12]) and PCA representations (the eigenface representation used by Pentland *et al.* [2]). In figure 2 and 3 the performances have been evaluated with the δ_{Euc} and the $\delta_{\cos}$ similarity measures, respectively.

There is a trend for the Kernel ICA and ICA representation to give superior face recognition performance to the PCA representation. The difference in performance is statistically significant for Test Set 2 and Test Set 3, when the test and training images differ not only in expression but also in lighting, scale and the date on which they were taken. Therefore, the high-order relationships among pixels, estimated by Kernel ICA and ICA, improve notably the performance when the face recognition is more difficult.

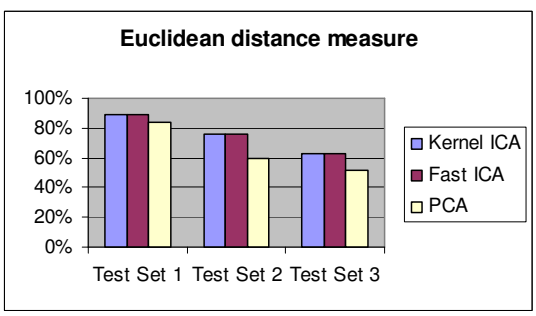

Fig.2. Recognition performance of the Kernel ICA, ICA factorial code representations and PCA representations corresponding to the δ_{Euc} similarity measure.

The lack of a substantial difference between the performances of the Kernel ICA and ICA algorithms, as found in their mono-dimensional applications, is probably due to the PCA preprocessing which is necessary in order to reduce the dimensionality of the data. In our opinion, the new orthogonal representation of the data provided by PCA precludes the kernel methods to improve their ability of represent the knowledge. In other words the evaluation of ICA produces the same results if it is applied directly

after PCA or after a further transformation of PCA in a non-linear space (kermel method).

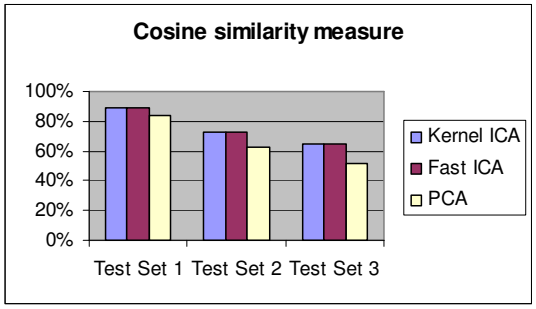

Fig 3. Recognition performance of the Kernel ICA, ICA factorial code representations and PCA representations corresponding to the $\delta_{\cos}$ similarity measure.

References

1. Zhao, W., Chellappa, R., Rosenfeld, A., Phillips, P.J.: Face Recognition: A literature survey. Technical Report CART-TR-948. University of Maryland, Aug. 2002.
2. Turk, M.A., Pentland A.P.: Eigenfaces for recognition. Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71–86, 1991
3. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711–720, 1997
4. Bartlett, M.S., Movellan, J.R., T.J., Sejnowski.: Face Recognition by Independent Component Analysis. IEEE Transactions on Neural Networks, vol. 13, NO. 6, November 2002.
5. Ruiz, A., López de Teruel P.E., : Nonlinear kernel-based statistical pattern analysis. IEEE Transactions on Neural Networks, vol. 12, no. 1, pp. 16–32, January 2001
6. Kim, K.I., Jung K., Kim H.J.: Face Recognition Using Kernel Principal Component Analysis. IEEE Signal Processing Letters, vol. 9, no. 2, February 2002
7. Yang, M.H.: Kernel eigenfaces vs. Kernel fisherfaces: Face Recognition using kernel methods. In Proc. 5th Int. Conf. Automat. Face Gesture Recognition, Washington, DC, May 2002, pp. 215-220
8. Bach, F.R., Jordan M. I., Kernel Independent Component Analysis, J. Machine Learning Res., vol. 3, pp. 1-48, 2002
9. Barlow, H.B.: Unsupervised learning Neural Comput., vol. 1, pp. 295-311, 1989
10. Atick, J.J.: Could information theory provide an ecological theory of sensory processing? Network, vol. 3, pp. 213-251, 1992
11. Phillips, P.J., Moon H., Rizvi S.A., Rauss P.J.: The FERET evaluation methodology for face-recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 10, pp. 1090–1104, 2000
12. Hyvärinen, A.: Fast and Robust Fixed-Point Algorithms for Independent Component Analysis. IEEE Transactions on Neural Networks 10(3):626-634, 1999