Face Recognition from Spatially-Morphed Video Sequences

R. Sebastião, Jorge A. Silva, and A.J. Padilha
Faculty of Engineering, University of Porto
arsebastiao@aibilu.pt, jsilva@fe.up.pt, padilha@fe.up.pt

Abstract. The aim of the present work is the recognition of human face visual information, in order to automatically control the access to restricted areas, granting access to authorized “clients” and barring the entrance to “impostors”. The vision system assembled performed the image acquisition, processing and recognition by first creating a database with a single view of each “client” and then by using multiple test images of each individual candidate to access. To get the test images, a video sequence was captured during the individual’s approach path to the camera. Because subjects presented themselves in a random pose before the camera, the synthesis of frontal views was incorporated, by using a view-morphing method. The modelling and the recognition were handled through the use of ICA methods. The identification of valid “clients” was fully successful. In order to check the rejection of “impostors”, a leave-one-out test was performed which gave promising results.

1 Introduction

Biometrics is the person identification science that allows the verification of an individual’s identity based on precise and careful measures of biological and physiological characteristics, such as fingerprints, hand geometry, iris and retina patterns, voice and face recognition and others.

In this work we intend to authenticate “clients” based on recognition of human face visual information. The current standard in face recognition is that of creating a database with several (sometimes, numerous) views of each individual’s face, corresponding to distinct poses, different emotional states, short-term variations in appearance caused by cosmetics or beard size, and the use of some accessories such as glasses, earrings and necklaces; these multiple views are intended to increase the individual’s recognition rate for the capture of a single test image. This work analysed a “reversed” paradigm, which creates a database with a single view of each individual, in a frontal pose, with neutral expression, and with a uniform background.

To get the test images a video sequence was captured during the individual’s approach path to the camera, and then three images (a larger number could be used, but three is sufficient for demonstration purposes) were extracted, with virtually no restrictions and with relatively poor quality.
Since subjects didn’t always present themselves in a front pose before the camera, the synthesis of front views was incorporated, by using a View-Morphing method, developed by Seitz and Dyer ([8], [9] and [10]).

Because image faces are complex and multidimensional it is necessary to construct models in order to reduce the data dimension, obtaining a simpler representation for them. Information theoretic approaches decompose face images into a small set of characteristic feature images, commonly called “eigenfaces”. Recognition is performed by projecting images into the subspace spanned by the eigenfaces, extracting the most relevant information in a face image, encoding it and then classifying the face by comparing one face’s code with a database of models encoded similarly.

Many authors have been using Principal Component Analysis (PCA) for face recognition ([1], [2], [11] and [12]). PCA consists of finding the principal components of the distribution of faces. These components are the eigenvectors of the covariance matrix of the database of facial images, each one accounting for a different amount of the variation among the images. Each face can then be represented as a linear combination of the eigenfaces (eigenvectors) or approximated using only the “largest” eigenfaces, ordered according to the associated eigenvalues.

However, PCA does not address higher-order statistical dependencies. As a generalization of PCA, Independent Component Analysis (ICA) accounts for the higher-order moments of the input; [1], [2] and [3] compare face recognition results under PCA and ICA, showing that ICA provides better performance. This technique represents the data \( X \) as a linear combination of nongaussian data components so that these are statistically independent, or as much independent as possible, according to the next equation:

\[
X = As
\]  

where \( s \) represents the independent components and \( A \) is the mixing matrix.

The ICA representation can be constructed under architectures I and II. The former considers images as random variables and pixels as observations, the second treats pixels as random variables and images as observations ([1] and [2]). Architecture II produces global features while architecture I produces spatially localized features that are only influenced by small parts of the image, leading to better object recognition results.

To find the independent components a quantitative measure of nongaussianity is needed; in fact, the ICA representation is based on maximizing objective functions, leading to an optimization problem. Hyvärinen et al., [6], present many contrast functions and explain the relations between them. The objective function optimization can be archived by an algorithm for maximizing it, such as the Newton method, the gradient method and the fixed-point algorithm, named FastICA, [4].

Based on reasons given in [7], we constructed the ICA representation considering architecture I and implemented it by the FastICA algorithm, using \( g(y) = \tanh(y) \) as the objective function. To construct the ICA representation the data need to be pre-processed by centring and whitening, [4]. The data, \( X \), is centred, i.e., made zero-mean, according to the following equation: