Face Gender Recognition Using Neural Networks and DCT Image’s Coefficient Selection

Marco GRASSI a,1, Marcos FAUNDEZ-ZANUY b, Ondrej SMIRG c, Honza MIKULKA c,

a D.I.B.E.T. Università Politecnica delle Marche, Ancona, Italy
(http://www.3medialabs.org,m.grassi@univpm.it)
b Escola Universitària Politécnica de Matarò (Adscrita a la UPC) 08303 Matarò
(Barcelona), Spain
(http://www.eupm.es/veu/faundez@tecnocampus.com)
c Brno University of Technology, UTKO, Parkynova 118, 612 00, Czech Republic
{smirg,mikulka}@tecnocampus.com

Abstract. Gender recognition plays an important role for a wide range of application in the field of Human Computer Interaction. In this paper, we propose a gender recognition system based on 2D Discrete Cosine Transform and Neural Networks. In particular, a discriminability criterion is used to select the DCT coefficients that make up the biometric template. Experimental results show how the proposed approach leads to a significant enhancement of recognition performances.

Keywords. Gender Recognition, Frequency Selection, Neural Networks, Biometrics

Introduction

Gender recognition plays an important role for a wide range of applications in the field of Human Computer Interaction also for commercial purposes. Gender recognition may in fact find application in areas such as advertising. An example might be an advertising screen, which uses a camera for detecting the gender of a human who looks at it and displays the content accordingly.

Gender recognition represents a very common task for humans. Since the first moment of life we learn to recognize our mother and father and during all existence we constantly perform gender recognition, often without being aware of it. By the way, notwithstanding this continuous training, gender recognition remains a difficult task due to the high variability of the characteristics that identify a male or a female. In several situations, the humans error rate in gender recognition can be surprisingly high. This is the case, for example, of the results of a recent online gender recognition test [1] in which...
the average of correct recognition rate was of 74%. Therefore, it’s not surprising that gender recognition constitutes a difficult and challenging task for machines.

Algorithms for gender recognition mainly focus on the entire human body to exploit the whole amount of available information. However, there are several situations in which the human silhouette is not shown completely or people are so close to the camera that only their face can be acquired. These cases can be tackled using algorithms for gender recognition from face. These algorithms are in general slightly less efficient because the used input data (faces) contain less information than the whole body. On the other side, they commonly provide the advantage of allowing real time recognition.

In the last decade significant advances have been achieved on biometrics, especially on face recognition [2] due in particular to the increased computational power of the computers. Such advantages can be exploited also for gender recognition that de facto represents a specific area of face recognition.

In this paper, we propose a gender recognition system based on 2D Discrete Cosine Transform and Neural Networks. In particular, a discriminability criterion is used to select the DCT coefficients that make up the biometric template. Experimental results show how the proposed approach leads to a significant enhancement of recognition performances.

1. Gender recognition

Gender recognition belongs to the category of pattern recognition. This implies that a large amount of methods exist. Usually, a pattern recognition system consists of three main blocks, as shown in Fig. 1: information source, feature extraction and classifier. In our case an RGB camera is the source of information, used for acquiring the face image. The image from camera is transformed to grayscale and continues to next block. Because the image contains a large amount of redundant information, system must include a block for the selection of relevant information that best describes the pattern. Features extraction is therefore a very important block of the pattern recognition system.

There are two main approaches for features extraction:

- **Statistical approaches** are based on statistical analysis to choose the most important characteristics of the image. For the statistical approach is necessary to have a sufficient number of input data that adequately characterizes the pattern. The main problem is that the images contain a large amount of redundant infor-
mation; hence the image is usually transferred from a high dimensional space into a small set of measurements. Dimensionality reduction algorithms using statistical methods search for the most common features in the input data. Typically the Karhunen-Loeve transform (KLT) is applied with a simplified algorithm known as eigenfaces [3]. However, eigenfaces algorithm is a suboptimal approximation to KL transform. Nowadays, with the improvements on computational speed and memory capacities, it is possible to compute the KLT directly. Generalized eigenfaces algorithm is called Principal Component Analysis (PCA) and is used in speech processing for example. Another dimensionality reduction method is based on DCT. DCT transforms data from the spatial to the frequency space. DCT is applied, between the others, in JPEG compression algorithm. JPEG algorithm uses the fact that the largest amount of relevant information is concentrated in the low frequencies. Both algorithms are used very often for face recognition [4].

- Geometry-features-based methods try to identify the position and relationship between face parts, such as eyes, nose, mouth etc., and the extracted parameters are measures of textures, shapes, sizes, etc. of these regions. Geometrical methods are used less frequently because they requires studies of the geometric variations of each recognized groups [5].

For the classifier is important to choose the most important characteristics of each group. This is particularly true for gender recognition in which only two groups exist: male and female.

2. Feature Extraction

DCT and PCA are the most widely used algorithms for face recognition. These algorithms have been published with different types of classifiers, and achieved very good results in face recognition. PCA is considered the norm among the features extractors and is implemented in many libraries, as for example OpenCV (Open Source Computer Vision) [6].

The Discrete Cosine Transform (DCT) is an invertible linear transform and is similar to the Discrete Fourier Transform (DFT). The original signal is converted to the frequency domain by applying the cosine function for different frequencies. After the original signal has been transformed, its DCT coefficients reflect the importance of the frequencies that are present in it. In fact, the very first coefficient refers to the signal’s lowest frequency, and usually carries the majority of the relevant (the most representative) information from the original signal [7]. The last coefficients represent the component of the signal with the higher frequencies. These coefficients generally represent greater image detail or fine image information, and are usually more noisy. DCT has an advantage when compared with DFT: the coefficients are real values while DFT produces complex values [8]. The 2D-DCT (two dimensional DCT) used in this work is the DCT-II. The DCT-II definition is shown in Equations (1) and (2).

\[
X[k, l] = \frac{2}{N} c_k c_l \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x[m, n] \cos \left( \frac{(2m + 1)k \pi}{2N} \right) \cos \left( \frac{(2n + 1)l \pi}{2N} \right)
\]  

(1)
Figure 2. Discriminability weight of DCT coefficients using average absolute energy vs. inter/intra variance measure for BiosecurID database

where:

\[
    c_i, c_j = \begin{cases} 
        \sqrt{\frac{1}{2}} & \text{for } k = 0, l = 0 \\
        1 & \text{for } k = 1, 2, \ldots, a - 1 \text{ and } l = 1, 2, \ldots, b - 1
    \end{cases}.
\]  

Feature extraction represents a crucial point for the overall performances of a gender recognition (or more in general for a biometric) system. Therefore several approaches and techniques have been applied to improve this task. In a previous paper [11] we explored the application of a genetic algorithm for selecting the most discriminant training images for the PCA algorithm. In the following section, we introduce a selection mechanism for the most discriminant coefficients of DCT.

3. Frequency selection

As told above, typically, biometric template based on DCT transform are obtained selecting a top left square region of the transformed image. In order to improve this selection mechanism which is only based on energy, we proposed a novel discriminability criteria based on a measure of the inter and intra class variation of the frequencies [9]. The goal is to pick up those frequencies that yield a low intra-class variation and high inter-class variation. Defined \( P \) and \( F \) respectively the number of people inside the database and the number of face image for each of them, \( I_{p,f}(f_1, f_2) \) the f-th transformed image of the p-th person and \( m(x, y) \) is the average value of each frequency obtained from the whole training subset images, we propose the following measure:

\[
    M_1(x, y) = \frac{\sigma^2_{\text{inter}}(x, y)}{\sigma^2_{\text{intra}}(x, y)}
\]

where,
\[ \sigma^2_{\text{inter}}(x, y) = \frac{1}{P \times F} \sum_{p=1}^{P} \sum_{f=1}^{F} (I_{p,f}(x, y) - m(x, y))^2 \] (4)

\[ \forall p = 1, ..., P \] is the variance of every frequency is the variance of each frequency evaluated over the whole training subset.

\[ \sigma^2_{\text{intra}}(x, y) = \frac{1}{P} \sum_{p=1}^{P} (\sigma_{p,f}(x, y)) = \frac{1}{P} \sum_{p=1}^{P} \frac{1}{F} \sum_{f=1}^{F} (I_{p,f}(x, y) - m(x, y))^2 \] (5)

\[ \forall p = 1, ..., P \] is the average of the variance of each frequency for each person.

Experimental results conducted over the AR database showed how the application of the proposed frequency selection criterion leads to a significant improvement in biometric recognition performances both in identification and in verification, both using Radial Basis Function (RBF) and Multilayer Perceptron (MLP) neural network [9].

4. The BiosecurID Database

In this paper, we have used our own database of faces to train and test the proposed method [10]. This database contains photos from faces of 185 female and 215 male, each person have 4 sets of pictures and each set contains 4 photo with different poses, totaling 6400 photos. The faces were photographed at different moments (different sets), with varying lighting, facial expressions (eyes closed/opened, smiling/not smiling), facial poses and facial details (with/without glasses, with/without beard), among other types.
Figure 5. Correct recognition rate (%) by means of RBF neural network classifier varying the number of hidden neurons with and without frequency selection for different training sets if images of variations. The images are in grayscale, with 256 different levels and a dimension of 92x112 pixels. This resolution was chosen for correspondence with the AT&T and ORL database of faces. Examples of used images are shown in Fig. 3.

5. Gender recognition system

In our system, for gender recognition from gender images, the frequency selection algorithm presented above has been employed to select the most discriminant features used to compose the biometric template, then an RBF (Radial Basis Function) neural network is employed for classification.

The RBF neural network has an input layer of 100 neurons, which is equal to the number of coefficients of the DCT template and an output layer of 2 neurons, one for each gender as shown in Fig. 3. In our simulation we consider an hidden layer with a number of neurons varying from 1 to 25 and a spread varying between 0 and 4. Different image training sets have been experimented using a different number of persons (same number of male and female) and for each of those the first image of each image set. System performances have been tested on the same set of 640 images belonging to 20 males and 20 females (not used for training) using all the 16 available images. Net has

<table>
<thead>
<tr>
<th>Method</th>
<th>Train Images</th>
<th>Test Images</th>
<th>Correct Recognition Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>1600/1600</td>
<td>100/100</td>
<td>73 %</td>
</tr>
<tr>
<td>PCA with genetic algorithm</td>
<td>1600/1600</td>
<td>1600/1600</td>
<td>85.9 %</td>
</tr>
<tr>
<td>DCT &amp; RBF</td>
<td>640/640</td>
<td>360/360</td>
<td>90.78 %</td>
</tr>
<tr>
<td>DCT (selected frequencies) &amp; RBF</td>
<td>640/640</td>
<td>360/360</td>
<td>93.28 %</td>
</tr>
</tbody>
</table>

Table 1. Result comparison.
been trained so that the "male" neuron respond with 1 to an input male image and -1 to a female image and vice versa for the female neuron. In the tests, the highest value of the output identifies the detected gender.

We compared system performances using different DCT templates of 100 coefficients with and without frequency selection. Experimental results show in Fig. 5 highlight how the application of the proposed frequencies selection method leads to a significant improvement of the recognition results. Better performances are also obtained in comparison with PCA and PCA with a genetic algorithm for images selection that we test in a previous work [11], as show in Table 1.

6. Conclusions

In this paper, we proposed a gender recognition system based on 2D Discrete Cosine Transform and Neural Networks. In particular, a discriminability criterion is used to select the DCT coefficients that make up the biometric template. Experimental results show how the proposed approach leads to a significant enhancement of recognition performances leading to a correct recognition rate up to 93%.

7. Acknowledgments

This work has been supported by FEDER and MEC, TEC2009-14123-C04-04.

References