UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

ENHANCEMENTS BY WEIGHTED FEATURE FUSION, SELECTION AND ACTIVE SHAPE MODEL FOR FRONTAL AND POSE VARIATION FACE RECOGNITION

TESIS PARA OPTAR AL GRADO DE DOCTOR EN INGENIERÍA ELÉCTRICA

LEONARDO AUGUSTO ANÍBAL CAMENT RIVEROS

PROFESOR GUÍA:
DR. CLAUDIO PÉREZ FLORES

MIEMBROS DE LA COMISION:
DR. KEVIN BOWYER
DR. PABLO IRARRÁZAVAL MENA
DR. JORGE SILVA SÁNCHEZ

SANTIAGO DE CHILE
2015

Mejoras por Fusión de Características, Selección y Modelo de Forma Activo para Reconocimiento de Rostros Frontal y de Variación de Pose

Resumen:

Reconocimiento de rostros es una de las áreas más activas en visión computacional debido a la amplia variedad de aplicaciones posibles en identificación de personas, control de acceso, interfaces hombre-computador, búsqueda en video, entre otras. Identificación de rostros es un problema de de búsqueda de uno a n, donde el la imagen de rostro capturada es comparada con n muestras en una base de datos. En este trabajo se propone un nuevo método para un reconocimiento de rostro robusto. La metodología está dividida en dos partes, la primera se enfoca a reconocimiento robusto a iluminación, gesticulación y pequeñas variaciones en edad y la segunda en variación de pose. El algoritmo propuesto está basado en características Gabor; las cuales han sido ampliamente estudiadas en identificación de rostros debido a sus buenos resultados y robustez. En la primera parte se propone un nuevo método para identificación de rostros que combina normalización local (local normalization) para la etapa de compensación de iluminación, ponderación de las características Gabor basada en entropía en la etapa de extracción de características, y mejoras en el clasificación por Borda count a través de un umbral que elimina los bajos puntajes de los jets en el proceso de votación. Para probar y comparar los resultados del método propuesto se usaron las bases de datos FERET, AR y FRGC 2.0. Los resultados en estas bases de datos muestran mejoras significantes respecto a los previamente publicados, alcanzando el mejor desempeño en FERET y AR. El método propuesto muestra una robustez significante a pequeñas variaciones de pose. El método fue probado asumiendo ruido en la detección de ojos para probar el desempeño a alineaciones de cara inexactas. Los resultados muestran que el método propuesto es rubusto hasta errores de 3 pixeles en detección de ojo para una cara en una imagen de 251×203 pixeles. Sin embargo, la identificación de rostros es fuertemente afectada cuando las imágenes de prueba son muy diferentes de las enroladas, como sucede cuando existe variación de pose. En la segunda parte de este trabajo se propone un nuevo método 2D basado en características Gabor el cual modifica la grilla en que las características Gabor son extraidas usando un modelo de mallas que modela las deformaciones producidas en el rostro al variar la pose. El método fue probado en las bases de datos FERET y CMU-PIE. El método propuesto logra el la taza de clasificación más alta en la base de datos FERET con un método de reconocimiento de rostros en 2D. El desempeño en CMU-PIE está entre los más altos publicados. Se realizaron exaustivos experimentos para distintas combinaciones del método, incluyendo pruebas rostros en dos poses enrolados.

Palabras clave: Reconocimiento de rostros, identificacion de rostros, Borda count, ponderación por entropía, Gabor, LBP, FERET, AR, FRGC 2.0.

ENHANCEMENTS BY WEIGHTED FEATURE FUSION, SELECTION AND ACTIVE SHAPE MODEL FOR FRONTAL AND POSE VARIATION FACE RECOGNITION

Abstract:

Face recognition is one of the most active areas of research in computer vision because of its wide range of possible applications in person identification, access control, human computer interfaces, and video search, among many others. Face identification is a one-to-n matching problem where a captured face is compared to n samples in a database. In this work a new method for robust face recognition is proposed. The methodology is divided in two parts, the first one focuses in face recognition robust to illumination, expression and small age variation and the second part focuses in pose variation. The proposed algorithm is based on Gabor features; which have been widely studied in face identification because of their good results and robustness. In the first part, a new method for face identification is proposed that combines local normalization for an illumination compensation stage, entropy-like weighted Gabor features for a feature extraction stage, and improvements in the Borda count classification through a threshold to eliminate low-score Gabor jets from the voting process. The FERET, AR, and FRGC 2.0 databases were used to test and compare the proposed method results with those previously published. Results on these databases show significant improvements relative to previously published results, reaching the best performance on the FERET and AR databases. Our proposed method also showed significant robustness to slight pose variations. The method was tested assuming noisy eye detection to check its robustness to inexact face alignment. Results show that the proposed method is robust to errors of up to three pixels in eye detection. However, face identification is strongly affected when the test images are very different from those of the gallery, as is the case in varying face pose. The second part of this work proposes a new 2D Gabor-based method which modifies the grid from which the Gabor features are extracted using a mesh to model face deformations produced by varying pose. Also, a statistical model of the Borda count scores computed by using the Gabor features is used to improve recognition performance across pose. The method was tested on the FERET and CMU-PIE databases, and the performance improvement provided by each block was assessed. The proposed method achieved the highest classification accuracy ever published on the FERET database with 2D face recognition methods. The performance obtained in the CMU-PIE database is among those obtained by the best published methods. Extensive experimental results are provided for different combinations of the proposed method, including results with two poses enrolled as a gallery.

Keywords: Face recognition, face identification, Borda count, entropy-like weighting, Gabor, LBP, FERET, AR, FRGC 2.0.

Este trabajo lo dedico a mi familia, a mis profesores, a mis compañeros y a todas aquellas personas que me ayudaron de alguna manera en el desarrollo de mi trabajo.

Les agradezco del mismo modo por haberme motivado a seguir, a superarme y a lograr las metas que me propuse. Me siento muy contento, he crecido como persona y como profesional y me siento capaz, con perseverancia, de enfrentar las nuevas etapas de mi vida.

Gracias a todos.

Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos FONDECYT 1080593 y 1120613, FONDEF D08I-1060 y el programa de Becas de Doctorado Nacional de CONICYT.

Contents

List of Tables

List of Figures

Abbreviations and Acronyms

2D: Two Dimensions

3D: Thre Dimensions

ASM: Active Shape Models

AAM: Active Appearance Models

BTH: Borda count Threshold

CLM: Constrained Local Model

DCT: Discrete Cosine Transform

E: Entropy (from the proposed method)

EBGM: Elastic Bunch Graph Matching

FERET: Face Recognition Technology

FRGC: Face Regognition Grand Challenge

FFT: Fast Fourier Transform

FLP: Fusing Local Patterns

GD: Grid Deformation

GFC: Gabor Fisher Classifier

HEC: Hierarchical Ensemble Classifiers

HGPP: Histogram of Gabor Phase Patterns

ICA: Independent Component Analysis

LBP: Local Binary Pattern

LDA: Linear Discriminant Analysis

LGBP: Local Gabor Binary Pettern

LGBPHS: Local Gabor Binary Pattern Histogram Sequence

LGBPWP: Local Gabor Binary Pattern Whitening PCA

List of Figures

LGT: Local Gabor Textons

LGXP: Local Gabor Xor Pattern

LLGP: Learned Local Gabor Pattern

LMG: Local Matching Gabor

LMGEW: Local Matching Gabor Entropy Weighted

LN: Local Normalization

LXP: Local Xor Pattern

P: Probabilistic Model

PCA: Principal Component Analysis

PDM: Point Distribution Model

PIE: Pose, Illumination and Expression

SRC: Sparse Representation-based Classification

WHT: Walsh-Hadamard transform

XOR: Exclusive Or

Chapter 1

Introduction

Contents

1.1 Definition of the Problem

Face identification is a one-to-n matching problem where the goal is to identify a person based on the face image, i.e., the captured face needs to be compared to n samples in the database [Phillips 2000]. Typical applications for face identification are used in immigration, access control, and law enforcement, agencies that try to answer the question, "Who is this person?". Face recognition is one of the most popularly studied topics in image analysis because of its wide range of possible applications, such as in surveillance, access control, information security, content-based video search, human computer interfaces, electronic marketing and advertising, and entertainment [Zhao 2003, Sharkas 2008, Serrano 2009, Chellappa 2011]. In spite of more than twenty years of intense research in face recognition, many real world situations are still a challenge. Uncontrolled conditions such as illumination changes, varying gestures, pose, and occlusions still present unsolved problems [Chellappa 2011].

1.2 Face Recognition Approaches

Face recognition has been approached in many different modalities. This work centers the literature review on those methods that are more popular and have yielded the highest face recognition performance on available face databases. The approaches covered in this section are Holistic methods, Local methods, Sparse methods and 3D methods.

1.2.1 Holistic Approaches

Holistic approaches use the image as a vector data and in some cases they reduce the data dimension by feature selection and by frequency component discrimination. Among the most widely used methods for face recognition based on feature extraction are Eigenfaces [Turk 1991], a holistic method that uses Principal Component Analysis (PCA) to project the image data vector into a reduced space, maximizing the variance of the data; and Fisherfaces [Belhumeur 1997], which is based on Linear Discriminant Analysis (LDA), maximizes the distance between classes and minimizes the distance between prototypes within each class; and methods based on Independent Component Analysis (ICA) [Bartlett 2002, Sharkas 2008]. Other methods for face recognition are based on frequency space as a discrete cosine transform (DCT) [Podilchuk 1996, Hafed 2001], that compares DCT–based feature vectors from different images, and the Walsh–Hadamard transform (WHT) [Faundez-Zanuy 2007], a low complexity algorithm that compares WHT-based feature vectors.

1.2.2 Local Approaches

Local feature-based methods for face recognition have shown improved robustness to changes in illumination, expression and occlusion. The Local binary pattern (LBP) method was proposed in [Ahonen 2006] where the face image is divided into square windows and a binary 1 is generated in the code whenever a pixel exceeds the value of the central pixel; otherwise a 0 is generated into the code.

Local Gabor methods are the widest used Local approaches. 2D Gabor wavelets [Bianconi 2007, Gabor 1946, Kyrki 2004, Lades 1993] (see details in section 2.1.3) have been used to extract local features achieving outstanding results in face recognition. Among the methods based on Gabor Wavelets are the Elastic Bunch Graph Matching (EBGM) method [Wiskott 1997], Gabor Fisher Classifier (GFC) [Liu 2002], Local Gabor Binary Pattern His-

togram Sequence (LGBPHS) [Zhang 2005], Histogram of Gabor Phase Patterns (HGPP) [Zhang 2007], Local Gabor Textons (LGT) [Lei 2007], Learned Local Gabor Pattern (LLGP) [Xie 2009], Local Gabor Binary Pattern Whitening PCA (LGBPWP) [Nguyen 2009], Local Matching Gabor method (LMG) [Zou 2007], Hierarchical Ensemble Classifiers (HEC) [Su 2009] and Fusing Local Patterns (FLP) [Xie 2010].

LMG is a simple but effective method for face recognition that extracts Gabor features from a grid of points, in which the distance between point is the wavelength of the Gabor filter. LMG treats each local features as an independent weak classifier combining the classification scores with Borda count voting system [Ho 1994].

In [Su 2009] the face was divided into patches without overlap, and then the best patches were selected and weighted with an LDA strategy in a greedy search. Finally, the local scores of the patches were combined with a global score obtained from the low frequency components of the FFT (Fast Fourier Transform) applied to the whole face, including its external boundary. Magnitude and phase Gabor features were combined in [Xie 2010]. The LBP operator was used on the Gabor magnitude features and the LXP operator (Local XOR Pattern) on the Gabor phase features. Then the face was divided into regions, and histograms of each region were computed on LGBP and LGXP features. Every region dimensionality was reduced using LDA, and finally the regions were compared with cosine distance.

1.2.3 Sparse Code

Sparse representation (SRC) methods [Wright 2009, Wagner 2009] characterize the testing image as a sparse linear combination of the training images, usually solving an optimization function with L1 or L2 norms. This method is very robust to occlusions and noise. Another important aspect of this method is that no information is lost as in the case of methods based on feature extraction, although it can be combined with PCA, LDA, Gabor or others. Nevertheless this method [Wright 2009, Wagner 2009] requires several enrolled images with different pose which may not be available in practical applications. A sparse Correntropy method which shows more robustness and efficiency with recognition of occluded and corrupted face images is proposed in [He 2011].

1.2.4 Face Recognition under Pose Variation

Face recognition under varying pose continues to be an area of active research. There are various approaches to solving this problem using 2D as well as 3D methods. The literature review of this work is focused on 2D methods because they are widely used and are applicable in real time.

1.2.4.1 Active Shape Models (ASM)

There are several methods that use 2D techniques to perform face recognition across pose. A method that performs frontalization by dividing the face into different components is presented in [Du 2009]. Several methods use Active Appearance Models (AAM) to frontalize the face [González-Jiménez 2007, Heo 2008, Gao 2009] and perform the match with a frontal face from a gallery set.

1.2.4.2 Statistic and Change of Space

In [Ashraf 2008] the image is divided into non-overlapping patches and then a statistically aligned model is built for each patch to perform a warping in the region. Using the same idea of patches, the image is divided into non-overlapping patches and a statistical model is constructed on each one at the score level [Vu 2009]. This method models how the matching score varies when the input face pose is at a certain angle. A face recognition Gabor-based method using a regressor with a coupled bias-variance trade-off is proposed in [Li 2012]. In this method, a statistical model is built at the score level, as in [Vu 2009]. A method that finds the sets of projection directions for different poses in the latent space is proposed in [Sharma 2012].

1.2.4.3 3D Approaches

Some methods use a single face as input and build a 3D model called the 3D Morphable Model [Blanz 2003, Romdhani 2005, Paysan 2009]. The 3D Morphable Model is based on a vector space representation of faces built using vectors of shape and texture. The parameters of the models are computed using a set of Eigen vectors obtained previously by training with images from 3D scans. A fully automatic face frontalization method using a 3D model was

introduced in [Asthana 2011b]. It works for poses varying up to ±45° on the yaw axis and ±30° on the tilt axis. In [Ding 2012] an automatic method was developed to find correspondences between 2D facial feature points and a 3D face model. The 3D face model built was then rotated to generate the frontal view.

1.2.5 Proposed method

This work takes as a starting point the local matching Gabor (LMG) method [Zou 2007], because of the simplicity, flexibility, robustness and performance of the method. The work is divided in two principal parts. The first part focuses on improving frontal face recognition. The second part takes the frontal face recognition and extends the methodology to a pose face recognition. The work was evaluated in FERET [Phillips 1998], AR [Martinez 1998], FRGC 2.0 [Phillips 2005, Phillips 2006] and PIE [Sim 2003].

A number of improvements to the LMG method [Zou 2007] are proposed [Perez 2011a, Perez 2011b, Cament 2014]. One is by using local normalization (LN) as a preprocessing stage of the LMG. Another is based on weighting Gabor jets by an entropy-like measure between the given face and the enrolled faces. An additional improvement is in the Borda count classification stage where it is proposed to use a threshold to eliminate low score jets that act as noisy inputs to the classifier. Also a combination of entropy weighting jets and threshold Borda count is assessed. Results are compared to recently published papers in FERET, AR, and FRGC 2.0 [Phillips 2005, Phillips 2006] face databases. Results are excellent in frontal and near frontal images, illumination changes and small occlusions.

The proposed pose face recognition method performed competitively with other published methods on face occlusions and in the presence of noise. Nevertheless, for face poses with increasing angles out of the face plane, the face normalization step loses the correction effect and, as in most 2D face recognition methods, performance declines significantly. For example, methods reaching near 100% for frontal face recognition may drop by up to 40% with pose variations ±60° [Ashraf 2008]. The proposed method modified the LMG alignment by including a type of ASM [Cootes 1992] called CLM [Saragih 2011]. A de-

formation field is built using the mesh obtained with CLM and Gabor jets are repositioned according to a face mesh. Because of local changes in the 2D face image with varying pose, a local statistical model [Vu 2009] is added to compensate deformations. The face is divided in non-overlapping patches, where a Borda count is computed in each one instead the whole image. Then a statistical model of each Borda count patch is obtained using a training database. The FERET database has pose variation between ±60°, and has been used in many recent publications of face recognition across pose [Vu 2009, Paysan 2009, Sarfraz 2010, Asthana 2011b, Ding 2012, Li 2012, Sharma 2012]. The proposed method reached the highest classification performance published to date in the FERET database with 2D face recognition methods. The CMU-PIE database has pose variations near ±90° and also has been used in most recent publications of face recognition across pose [Paysan 2009, Sarfraz 2010, Asthana 2011b, Asthana 2011a, Castillo 2011, Ding 2012, Sharma 2012, Li 2012]. The performance of the proposed method on the CMU-PIE database is among those that reached the highest classification performance.

1.3 Structure of the Document

The thesis is organized as follows. Methodology is presented in chapter 2, and it is divided in two parts. The first part focuses on frontal face recognition in section 2.1 and the second part focuses on face recognition across pose variations in 2.2. Results are presented in chapter 3, divided in frontal face recognition in section 3.2 and face recognition under pose variation in section 3.3. Finally a discussion and conclusions of the work can be found in chapter 4.

Chapter 2
Methodology

Contents

LMG face recognition method [Zou 2007, Perez 2011a, Cament 2014] is mainly composed of:

- Image alignment (section 2.1.1).
- Gabor feature extraction (section 2.1.3).
- Classification (section 2.1.4).

The description of the method is presented in the following sections.

2.1 Frontal Face Recognition

2.1.1 Image Alignment

Image alignment is an important stage for face recognition. A significant number of papers have focused directly on face recognition on the assumption that the face has already been localized [Perez 2003, Perez 2005, Zhang 2005, Zhang 2007, Zou 2007, Xie 2009, Nguyen 2009]. Others have focused on face and eye localization in which faces are cropped and aligned [Campadelli 2009, Perez 2007, Perez 2010a].

The alignment module on this work performs face normalization through image rotation, displacement, and resizing to locate the eyes at a fixed position. The normalized image has the eyes fixed at positions (67, 125) and (135, 125) within a 203×251 pixels image. In this work as well as in [Zhang 2005, Zhang 2007, Xie 2009, Nguyen 2009, Zou 2007] it is assumed that eye detection is performed previously and classification performance depends mainly on feature extraction and classification modules. Nevertheless, eye localization for frontal faces such as those present in the FERET database have reached very precise results [Campadelli 2009, Perez 2010a].

2.1.2 Illumination Compensation through Local Normalization

The image with Local Normalization features (LN) is used as a preprocessing stage to the input image in the LMG method. The local normalization [Xie 2006] is an illumination compensation method that makes uniform the mean and variance of an image around a local neighborhood. Let $I(x, y)$ the image, its LN image $I_{LN}^{n\times n}(x, y)$ using an $n \times n$ window neighborhood is given by (2.1),

$$I_{LN}^{n\times n}(x, y) = \frac{I(x, y) - I_{mean}^{n\times n}(x, y)}{I_{std}^{n\times n}(x, y) + 0.01}. \tag{2.1}$$

$I_{mean}^{n\times n}(x, y)$ denotes the mean of a $n \times n$ neighborhood around the pixel (x, y), and $I_{std}^{n\times n}(x, y)$ is the standard deviation in the $n \times n$ neighborhood. Figure 2.1(a) shows a face image and Figure 2.1(b) shows the LN of the image. In this work a 9×9 neighborhood is used, because of its performance compared to others.

Figure 2.1: (a) Original image, (b) local normalization image.

2.1.3 Gabor Feature Extraction

2.1.3.1 Gabor Wavelet

Gabor wavelets have proved to be a very successful tool for face recognition purposes, as previously mentioned in the introduction section. It has been shown that 2D Gabor functions model simple cells in the visual cortex in mammalian brains [Lee 1996, Marčelja 1980, Daugman 1980, Daugman 1985]. Gabor wavelets are capable of extract lines or borders (features) at different orientations and also different sizes or scales. The Gabor kernel is given by $\Psi_{\mu,\nu}$ in equation (2.2). $\vec{z}$ represents the pixel coordinate, μ is the orientation (columns on Figure 2.2), ν the scale (rows on Figure 2.2), f the step in frequency, and k_{max} is the maximum spatial frequency. Figure 2.2 shows 2D Gabor filters for eight orientations ($0 \leq \mu \leq 7$) and five scales ($0 \leq \nu \leq 4$).

$$\begin{gathered}\Psi_{\mu,\nu}\left(\vec{z}\right) = \frac{\left|\vec{k}\right|^2}{\sigma^2} exp\left(-\frac{\left|\vec{k}\right|^2 |\vec{z}|^2}{2\sigma^2}\right)\left[exp\left(i\vec{k}\cdot\vec{z}\right) - exp\left(-\frac{\sigma^2}{2}\right)\right] \\ \vec{z} = (x,y)^T,\ \vec{k} = \frac{k_{max}}{f^\nu}\left[cos\left(\pi\frac{\mu}{8}\right), sin\left(\pi\frac{\mu}{8}\right)\right]^T.\end{gathered} \tag{2.2}$$

2.1.3.2 Gabor Features

A Gabor feature is obtained by finding the convolution between the image $I\left(\vec{z}\right)$ and the kernel $\Psi_{\mu,\nu}\left(\vec{z}\right)$, i.e. $G_{\mu,\nu}\left(\vec{z}\right) = I\left(\vec{z}\right) * \Psi_{\mu,\nu}\left(\vec{z}\right)$. A Gabor feature has both a real and an imaginary part, and in this work the magnitude and phase features are used separately. Magnitude $M_{\mu,\nu}\left(\vec{z}\right)$ and phase $P_{\mu,\nu}\left(\vec{z}\right)$ are computed by (2.3) and (2.4),

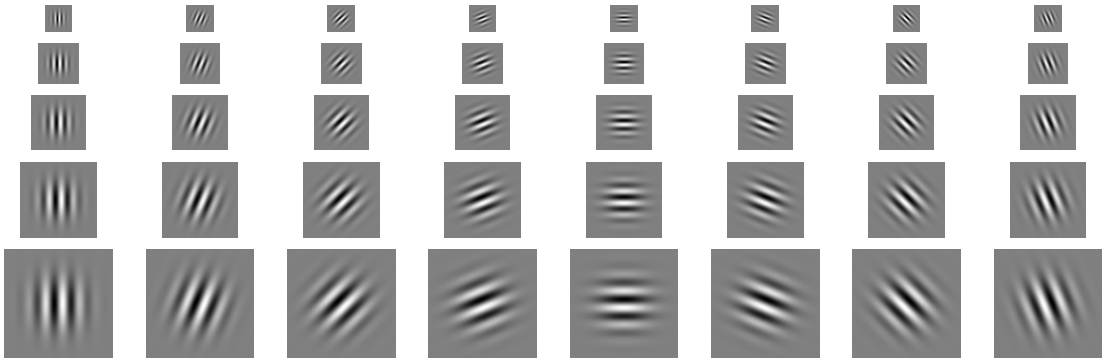


Figure 2.2: Imaginary part of Gabor filters for five scales ($0 \leq \nu \leq 4$) and eight orientations ($0 \leq \mu \leq 7$). Each row shows a determined scale, from $\nu = 0$ to $\nu = 4$ and each column shows a fixed angle, from $\mu = 0$ to $\mu = 7$.

$$M_{\mu,\nu}\left(\vec{z}\right) = \sqrt{Im\left(G_{\mu,\nu}\left(\vec{z}\right)\right)^2 + Re\left(G_{\mu,\nu}\left(\vec{z}\right)\right)^2}, \tag{2.3}$$

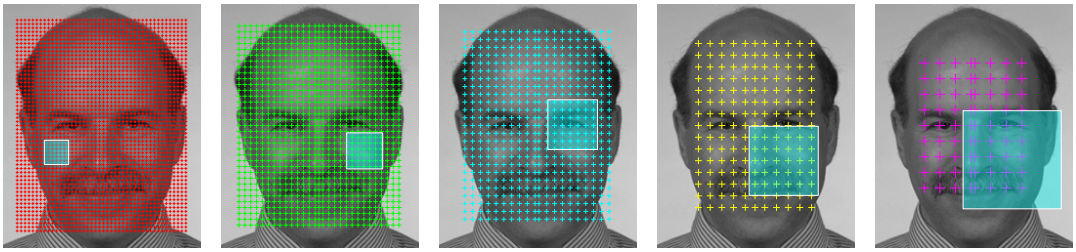
$$P_{\mu,\nu}\left(\vec{z}\right) = \arctan\left(\frac{Im\left(G_{\mu,\nu}\left(\vec{z}\right)\right)}{Re\left(G_{\mu,\nu}\left(\vec{z}\right)\right)}\right). \tag{2.4}$$

2.1.3.3 Gabor Feature Grids

Gabor features are not extracted on the whole image, because it is computationally expensive and it is not necessary for the similarity between continuous pixels and Gabor robustness to small spatial variations [Zou 2007, Perez 2011a]. Gabor features are computed on a grid of points, in which each point is separated from the next one by the wavelength of the filter frequency [Zou 2007]. In this work five scales ($0 \leq \nu \leq 4$) and eight orientations ($0 \leq \mu \leq 7$) for the Gabor filters are used as in [Zou 2007, Su 2009, Xie 2010]. For the five scales used in this work the wavelengths are $4, 4\sqrt{2}, 8, 8\sqrt{2}, 16$, and the location points where the Gabor features are computed can be seen in Figure 2.3. LMG uses the magnitude of the Gabor feature (2.3). A *jet* is defined as a vector of a Gabor feature magnitude in which each element represents one of the eight orientations, that is normalized to euclidean norm equal to one (2.5),

$$\vec{J}_{\nu}\left(\vec{z}\right) = \left[\frac{M_{0,\nu}(\vec{z})}{\sqrt{\sum_{\mu=0}^{7} M_{\mu,\nu}(\vec{z})^2}}, \quad \ldots, \quad \frac{M_{7,\nu}(\vec{z})}{\sqrt{\sum_{\mu=0}^{7} M_{\mu,\nu}(\vec{z})^2}}\right]^T. \tag{2.5}$$

Therefore, 5 sets of jets corresponding to different scales are computed for each image, making a total of N jets. The spatial information is preserved implicitly because the image was transformed to locate the eyes in a fixed position and the jets are computed in the same positions in every image. The jet computation is the same as was previously used in [Zou 2007, Cament 2014], and it results in a total of 4,172 jets.



(a) (b) (c) (d) (e)

Figure 2.3: Image from the FERET database showing the grids used to compute the Gabor jets. The five grids correspond to five spatial scales (a) $\nu = 0$, (b) $\nu = 1$, (c) $\nu = 2$, (d) $\nu = 3$ and (e) $\nu = 4$. The + signs represent the spatial position on the grid where the Gabor jet is computed. The white squares represent the size of the Gabor kernel for each spatial scale in pixels of (a) 25×25, (b) 37×37, (c) 51×51, (d) 71×71 and (e) 101×101.

2.1.4 Classification

Local Matching Gabor classification is performed in two stages; a comparison of jets, in which each jet of the target image is compared to each jet of the gallery image and a classification stage using Borda count voting [Ho 1994].

2.1.4.1 Gabor jet comparison

The jets of the target image are compared with the N jets of the M gallery images in the database, creating an $M \times N$ comparison matrix C. The comparison of a pair of jets is performed by cosine distance $C_{i,j} = \frac{J_j^T \cdot J_{i,j}^G}{\|J_j^T\| \cdot \|J_{i,j}^G\|}$, where J_j^T is the j-th jet of the test image, and $J_{i,j}^G$ is the j-th jet of the i-th gallery image.

Each jet is considered as an independent classifier that is combined using the voting method Borda count [Ho 1994, Kittler 1998]. Borda count is applied on all vectors formed with every column of matrix C. The Borda count consists of ranking the comparison values, assigning 0 to the lowest value, 1 to the second

lowest value, up to $M-1$ for the highest value. This ranking matrix is denoted as O, and in the Borda count, the score of each candidate is given by the sum of rankings of all classifiers, $B_i = \sum_{j=1}^{N} O_{i,j}$.

The cosine distance is employed to compare a Gabor jet of a test image with the corresponding Gabor jet from the database image. Thus, an $N \times M$ comparison matrix is built, where N is the number of enrolled images and M is the number of Gabor jets. Gabor jet comparison is called the score of the jet $S_{i,j}$.

2.1.4.2 Borda count classification scheme

Borda count is a voting method where every voter ranks all the candidates in order of preference. Then, a sum of all voter ranks is performed to obtain the candidates scores. Table 2.1 shows an example of a voting where voter ranks all candidates and the last column shows the final score obtained for each candidate. Borda count has been used in face recognition as a method that combines weak classifiers such as Gabor jets, i.e., each Gabor jet is a voter. A total of 4,172 jets are used in [Zou 2007, Perez 2011a], where every jet is considered an individual classifier. The combination of jet scores is carried out with Borda count. In this context, for a given jet, every subject score is a voter. Borda count score yields values between 0 to N-1 to each subject jet, in order to score value (0 for minimum and N-1 to maximum). Then the jets are sorted, and the sorted indexes $O_{i,j}$ are added for every subject $B_i = \sum_{j=1}^{M} O_{i,j}$, and the subject with the highest score B_i is the test image identity, as shown in Figure 2.4. Table 2.1 shows an example of a voting where the voter ranks all candidates and the last column shows the final score obtained for each candidate.

Table 2.1: Example of Borda count voting method.

	Voters					
Candidates	V1	V2	V3	V4	V5	Sum
Candidate 1	0	3	2	1	0	6
Candidate 2	2	2	3	2	1	**10**
Candidate 3	1	0	1	3	2	7
Candidate 4	3	1	0	0	3	7

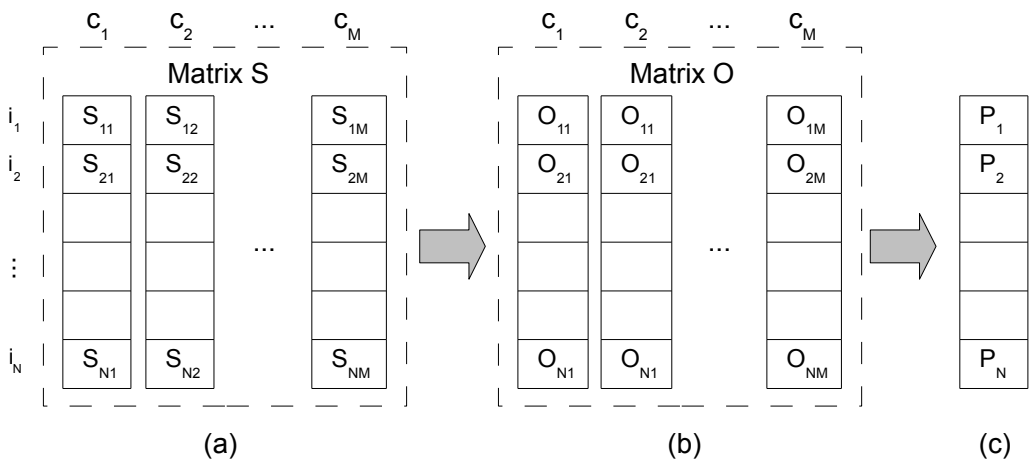


Figure 2.4: Illustration of Borda count computation method. (a) Matrix S with dimensions $N \times M$, where $S_{i,j}$ is the score of the jth Gabor jet comparison with the ith image of the gallery. (b) Matrix with sorted columns O with dimension $N \times M$ with values $N-1$ to 0 in each column. (c) Borda count score for the ith gallery element as the sum of all column in the ith rows of O

2.1.5 Local matching Gabor entropy-like weighted with LN features (LMGEW//LN-BTH)

There are two parallel stages in LMGEW//LN-BTH [Perez 2011a, Cament 2014]. The first stage computes the entropy-like weighted vector (Figure 2.5(a)). The second stage computes the Borda count from the LN input image I_{LN} (Figure 2.5(b)). A Borda count threshold (BTH) is used to eliminate from Gabor jet comparison matrix with low values that act as noise in the identification procedure. A block diagram of LMGEW//LN-BTH is shown in Figure 2.5.

2.1.5.1 Entropy-like weighted vector

The proposed entropy-like measure estimates the variability of the similarity of the Gabor feature in the position (x, y, λ) of the target image with respect to the same feature in the whole gallery set. Each Gabor feature is represented by a jet. Entropy has its maximum value when probabilities of different states are the same, and lower values when probabilities are different. The states are the comparison values between jets from the target image and the gallery set. Therefore, entropy provides quantitative information to determine if a jet can be used to distinguish faces in the database. Jets located in the face regions with

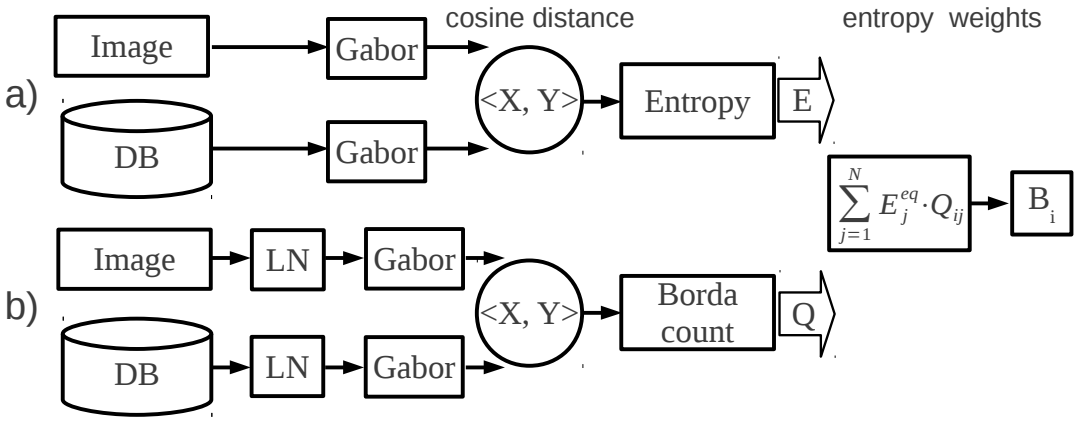


Figure 2.5: Block diagram of LMGEW//LN-BTH. (a) Computation of the entropy-like weighted vector. (b) Computation of the Borda count using LN in a preprocessing stage.

low entropy should be given greater weight in the final face recognition score. Probabilities are estimated using the histogram of the values resulting from the jets comparisons (C matrix). The values of C are quantized in K bins and the histogram for each jet along the database (rows of C) is computed. The values of the histograms are normalized by the number, N, of faces in the gallery, and are used as the probability $P_{j,k}$ that the comparison value $C_{i,j}$ of the j-th jet is within the bin k.

The entropy-like value of the j-th jet E_j is computed for each jet as shown in (2.6). Because jets with lower entropy-like value must be emphasized, the inverse of this value ($E_j^{-1} = 1/E_j$) is computed and normalized to the range [0,1]. Finally, E_j^{-1} is equalized to distribute the values uniformly, calling E^{eq} the equalized vector.

$$E_j = -\sum_{k=1}^{K} P_{j,k} \cdot \log_2 P_{j,k} \tag{2.6}$$

Figure 2.6 (a-e) shows an example of entropy weights, E^{eq}, computed for each spatial frequency. The example shows that features around the face, near the mouth and ears, have lower weights while features around the nose and eyes have higher weights, in this way indicating features that contribute most in the face identification process.

In order to visualize the entropy-like values the following experiment is performed. $\bar{E}^{eq}$ is computed for each target image of a train set, and an average E^{eq}

is computed. Then using a threshold, those jets with 25% of the highest value were selected An example of the selection is shown in Figure 2.6 (f-j).

It can be observed that the most important selected features are located around the eyes and nose. At lower spatial scales, features around the mouth are also selected. At higher spatial scales the chin and the lower part of the cheeks are selected. It can also be seen that features on the forehead and external area around the face are not selected.

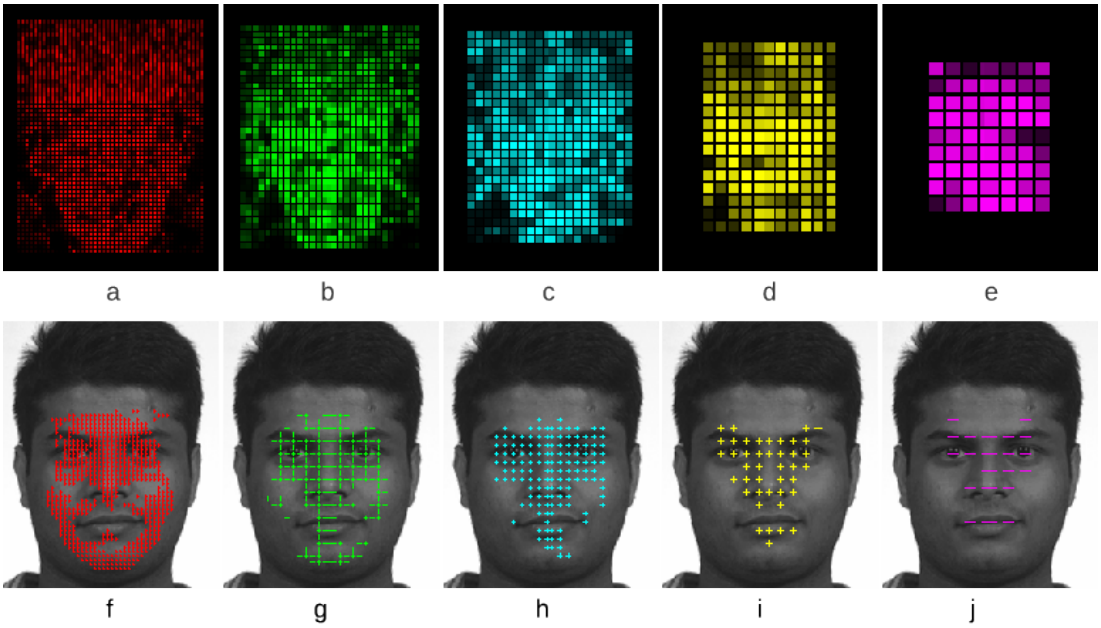


Figure 2.6: (a-e) Computed equalized weights for jets at five spatial frequencies (a) ν =0, (b) ν =1, (c) ν =2, (d) ν =3 and (e) ν =4. The intensity color is proportional to the jet weight, i.e., a 0 weight is assigned to gray intensity 0, and a weight of 1 to the gray intensity 255. (f-j) Jets selected by the highest weight values (25%) for each spatial frequency; (f) ν =0, (g) ν =1, (h) ν =2, (i) ν =3 and (j) ν =4.

2.1.5.2 Borda count enhancement by threshold (BTH)

In the standard Borda count [Zou 2007], all jets contribute to the final score even for jets with very low inner product value. By using a threshold T_h, over the inner product value between the candidate face and the gallery image, it is possible to eliminate very small scores which may act as noise for the Borda count computation. In this proposal the jet scores under the threshold are set to zero. If $C_{ij} < T_h$, then $O_{i,j} = 0$, otherwise the score value is maintained. Using

a training set from any database it is possible to obtain an optimum T_h. The rank-1 identification rate must be computed varying the T_h threshold and the value that produces the maximum rate is selected.

2.1.5.3 Borda count ranking matrix for LN images

In this stage, Gabor features of $I_{LN}(x, y)$ are computed using 5 scales and 8 orientations, and a comparison matrix C_{LN} and ranking matrix O_{LN} are obtained. A modified ranking matrix Q_{LN} is created by (2.7), where T_h is a threshold that eliminates noisy values. The BTH suffix is added to the method name when the matrix Q is used instead of O,

$$Q_{i,j}^{LN} = \begin{cases} O_{i,j}^{LN}, & C_{i,j}^{LN} \geq T_h \\ 0, & C_{i,j}^{LN} < T_h. \end{cases} \tag{2.7}$$

Finally, Q_{LN} is weighted with the entropy-like vector E^{eq}. An identification score $B_i = \sum_{j=1}^{N} E_j^{eq} \cdot Q_{i,j}^{LN}$ is obtained, and the highest score represents the person's identity.

2.1.6 Fusion of LMGEW//LN-BTH with LGBP and LGXP

2.1.6.1 Review of LGBP and LGXP

The method proposed in [Xie 2010] combines local binary pattern (LBP) applied to the magnitude component of the Gabor features (LGBP), and the local XOR operator (LXP) applied to the phase component of the Gabor features (LGXP) for face recognition. As shown in Figure 2.7 the method consists of the extraction of LGBP and LGXP [Xie 2010] features from the image. The image is divided into blocks and each block into sub-blocks. For every sub-block a histogram is computed and for each block the sub-block histograms are concatenated. Then LDA is applied to reduce the histogram dimensionality, selecting the most significant features. LGBP and LGXP are combined at the score or feature level using $S = w \cdot S_{LGBP} + (1 - w) \cdot S_{LGXP}$, where S_{LGBP} is the LGBP score, S_{LGXP} the LGXP score, and w the weight relation between LGBP and LGXP features.

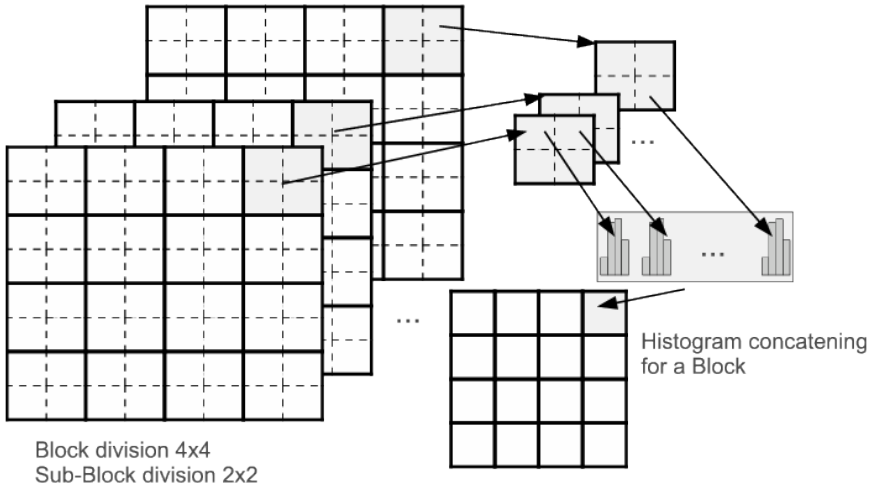


Figure 2.7: Division of blocks and sub-blocks of LGBP.

2.1.6.2 Fusion with LMG

Fusion of different methods is performed at the score level as follows. In the HEC algorithm [Su 2009] global features C_G are combined with local features C_L at the score level as $HEC = w \cdot C_G + (1 - w) \cdot C_L$. In [Xie 2010] LGBP and LGXP features are combined at the score level in the same way as HEC as shown in section 2.1.6.1. A combination between LMGEW//LN-BTH and LGBP or LGXP [Cament 2014] is done in the same way as in previous methods as shown in (2.8), where S_1 denotes the identification score of LMGEW//LN-BTH, and S_2 the score of LGBP or LGXP, depending on which kind of feature is used,

$$S = w \cdot S_1 + (1 - w) \cdot S_2. \tag{2.8}$$

The simulated example shown in Table 2.2 has matching scores for two methods (Method 1 and Method 2) of three matches: AA (comparing two individuals of the same class match), AB (comparing one individual to another), and AC (comparing one individual to another). It can be observed that match AA is not the best for both methods, but nevertheless reaches the highest score in the fusion of both methods. This result can be explained when each different method (method 1 and method 2) uses different features to reach the score. Then both methods can combine the best features to reach the score for their fusion. In this case they will be combined synergistically reaching the highest scores because each method solves different features better.

Table 2.2: A simulated example to illustrate the benefit of the fusion of two different methods. The scores are shown for two methods and their fusion for the comparison of one individual with him/herself (AA), with individual B (AB) or with individual C (AC).

Match	AA	AB	AC
Method 1	0.6	0.7	0.4
Method 2	0.6	0.3	0.7
Fusion	0.6	0.5	0.55

2.2 Pose Variation Face Recognition

As for frontal face recognition, this model consists of three main modules [Cament 2014]: image alignment, feature extraction through Gabor jets computation, and classification using a statistical model, as shown in Figure 2.8.

In the first module, an ASM is used to align the face and to determine its shape with the goal of extracting features from corresponding points relative to the gallery image. Also, LN is used in the normalized image, as was shown in frontal face recognition section.

The second module performs the Gabor jet computation to extract face features. The spatial position for the computation of each Gabor jet is defined using the ASM adjusted to the face pose. A grid of Gabor jets is placed over the face using the eye positions as references as described in [Zou 2007]. The grid is deformed using the ASM adjusted to the face. In this way, the position of the Gabor jets with respect to the face features (eyes, nose, mouth, etc.) is closer to the original position in the frontal face (see Figure 2.11).

The third module carries out face classification. At this stage, entropy weights are used to weight each jet. A Borda count method is used for classification and a local statistical model is employed to address pose change [Vu 2009]. This local statistical model learns how the face texture varies as the face changes in pose. The model divides the face into several regions and the texture variation is computed for each region as the pose changes.

2.2.1 Alignment and Facial Shape Adjustment

2.2.1.1 Active Shape Model Adjustment

A deformable model is adjusted to the face in order to take pose changes into account. A linear shape model is used [Matthews 2004]. This type of deformable model uses a linear transformation in a vector space, which takes the main con-

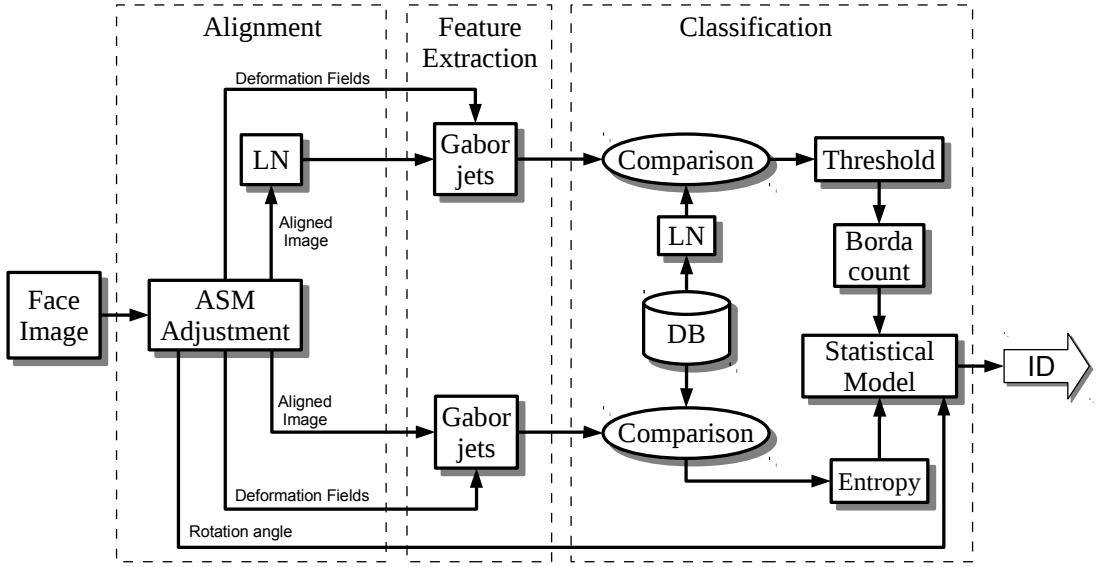


Figure 2.8: The proposed method consists of three main modules: image alignment based on ASM, feature extraction through Gabor jets computation, and classification using entropy weights and statistical model matching.

figurations of the model in a set of examples into account. In particular, a linear model called ASM is used [Cootes 1992] as reported in this study. This model uses a Point Distribution Model (PDM), where the shape of the object is defined by a vector with the coordinates of a set of points:

$$S = [x_1, y_1, z_1, x_2, y_2, z_2 \ldots, x_l, y_l, x_l]^T . \tag{2.9}$$

The shape S of a new object can be expressed as a mean shape S_0 linearly deformed by an affine transformation and by the linear combination of basis vectors that represent the principal modes of variation of the object in a set of examples:

$$S = \mathcal{T}(s, R, T; S_0 + \Psi_i p). \tag{2.10}$$

where p is a vector of parameters for the basis $\Psi = [S_1|S_2|\ldots|S_n]$ of n vectors, and $\mathcal{T}(s, R, T; \cdot)$ is a rigid transformation that performs a rotation by the rotation matrix $R(\alpha, \phi)$, scaling s, and translation by the vector $T = [t_x, t_y]^T$, i.e., if $\mathcal{T}(s, R, T; \cdot)$ is applied to a point $[x, y]^T$:

$$\mathcal{T}(s, R, T; [x, y]^T) = sR \begin{bmatrix} x \\ y \end{bmatrix} + T. \tag{2.11}$$

The basis Ψ is built by a PCA of a set of training shapes. Before performing the PCA, the training shapes are normalized by an iterative Procrustes analysis [Goodall 1991] to remove global variations, such as translation and rotation, and therefore taking into account only the non-rigid local variations in the PCA. The PCA provides a set of orthogonal modes of variation or eigenvectors, and its eigenvalues. The eigenvalues represent the importance of the corresponding eigenvector within the modes of variation across the training shapes. Among the eigenvectors, those with larger eigenvalues are selected to build the basis Ψ. Thus, a model that can generate large variations in shape using few parameters is obtained.

In order to fit the deformable model to the image, a particular type of ASM named Constrained Local Model (CLM) is used [Cristinacce 2006, Saragih 2011]. This CLM uses patches where local features are computed to search landmarks in the image. The position of each landmark is related to a point of the PDM. A set of training images is used to build a model that incorporates local features associated with each anatomical landmark. If the shape S is fitted to a test image by the ASM, the parameters of $\mathcal{T}(s, R, T; \cdot)$ and the vector of parameters, p, of the basis Ψ, are optimized (2.11). The optimization minimizes the difference between features computed in the patches at each point of S and the features of the model built using the training images. The image features are extracted from patches of the raw gray image, the image gradient, and local binary patterns (LBP).

The features computed in the test images are compared with the features of the training images by using the normalized correlation coefficient. Then, the method uses a non-parametric approach based on the mean-shift [Fukunaga 1975] mode seeking algorithm to optimize the position of the points in the PDM by minimizing the difference between the training and test features. The result is a shape fitted to the facial features (Figure 2.9). For a completely automatic face recognition system, pose and fiducial landmarks should be obtained automatically. However, for experimentation purposes, it is assumed them to be known beforehand, a common practice followed in several previous studies [Sharma 2012, Li 2012, Vu 2009, Zou 2007, González-Jiménez 2007, Ashraf 2008, Blanz 2003]. The shapes adjusted by the ASM that did not fit correctly for large pose variations were adjusted manually to evaluate the expected performance of the proposed method.

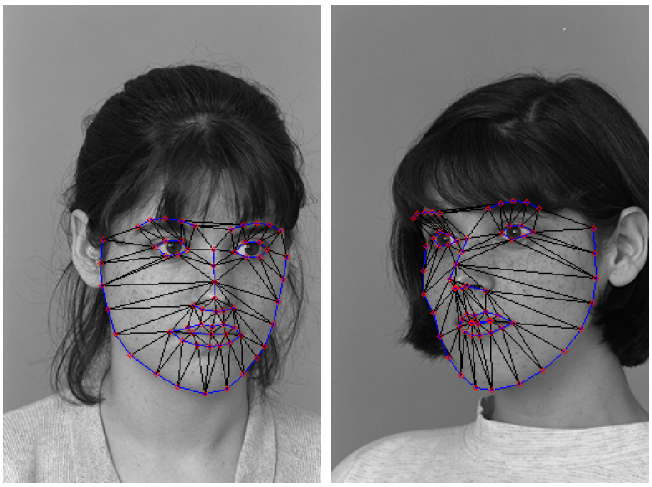


Figure 2.9: Example of the mesh used to detect facial features by using the CLM method of Saragih et al. [Saragih 2011].

2.2.1.2 Face Alignment

Before computing the Gabor features, the images are normalized to a size of 300×400 pixels using the shapes adjusted to the face and a mean face shape. Figure 2.10 shows the face mean shape (in blue lines) that is positioned on the face as follows: The vertex of the mean shape representing the position of the columella's base of the nose is positioned on the horizontal center of the image and at 1/3 of the height of the image. The columella is the tissue that links the nasal tip to the nasal base, and separates the nares. It is the inferior margin of the nasal septum. The eyes are localized on the same horizontal line and their separation between centers is set to 68 pixels. The separation between the eyes controls the scale of the aligned mean shape and has the same scale used in [Zou 2007, Cament 2014]. In the alignment step, the images are rotated to set the eyes on the same horizontal line, scaled such that the width of the adjusted shapes is equal to the width of the aligned mean shape, and translated so that the center of the triangle formed by the center of each eye and the vertex at the columella of the nose coincide between the mean shape and the shapes adjusted to the faces. This alignment allows comparisons between gallery and test faces in similar positions, improving the performance of the method. If the face pose has a large rotation (yaw rotation) and one side is occluded, the vertex at the upper part of the nose is used instead of the occluded eye. Figure 2.10 shows a normalized image of the FERET database with the adjusted shape in green and

the mean shape in blue.

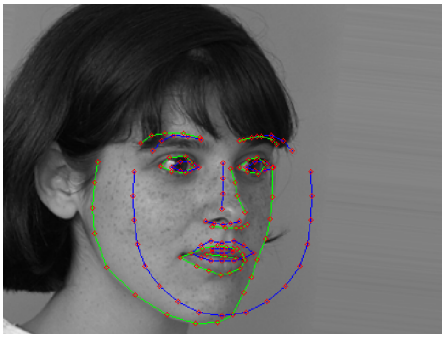


Figure 2.10: Image from the FERET database normalized before using the face recognition method. The image is aligned using a mean face shape (blue) and the shape adjusted to the face (green). The image is rotated to place the eyes on the same horizontal line, scaled such that the width of the adjusted shape is equal to the width of the mean shape, and translated so that the center of the triangle formed by the center of each eye and the vertex at the columella of the nose coincide.

2.2.2 Gabor Features Extraction with Grid Deformation

Gabor jets are computed on selected points, using five different grids. Frontal LMG used the grids defined in [Zou 2007], where the images are normalized to 203×251 pixels with the eyes in fixed positions (67, 125) and (135, 125). Figure 2.3 shows an example of these normalized images with the five grids superimposed. The grids indicate the positions where features are extracted by the Gabor jets at five different spatial resolutions. The same Gabor filters as used for frontal face recognition (details in section 2.1.3) are applied on these images.

Figure 2.11(a) shows an image of the FERET database with the mean face shape and the grid with $\nu = 3$ superimposed. Then, the grids are deformed by using the adjusted shape to change the positions at which the jets are computed in each image. A field of deformations $D(x, y) = [x', y']$ is computed using the Thin Plate Splines method [Bookstein 1989] on the vertex of the model and adjusted shapes. This deformation field is the spatial transformation between the mean shape and the adjusted shapes, and it is used to deform the grid in such a way that the jets are computed at the same relative position with respect to the adjusted shapes of each image. Therefore, the positions of the points in the deformed grids are $p'_\nu(i) = D(p_\nu(i))$. Figure 2.11(b) shows the grid with $\nu = 3$ deformed using this method. Figures 2.11(c) and 2.11(d) show deformed grids in a frontal and

rotated face, respectively. As these figures show, the positions of the Gabor jets have the same relative position with respect to the face features in both images. If one side of the face is occluded because of the rotation, only the vertices and jets on the side of the face that is exposed are used to compute the deformation field and the Gabor features, respectively. The method assumes that a gallery is available with faces of the persons to be identified. For each face in the gallery, the Gabor jets are computed off-line and stored in a database for later on-line identification.

2.2.3 Classification

A similar classification scheme than the one shown in section 2.1.4 is used for pose variation case. It is used the same comparison with cosine distance between jets, Borda count as a classifier combiner, BTH and entropy weighting. An Statistical model per pose was added to the classification of pose variation face recognition.

2.2.3.1 Statistical Model

Although the position where the jets are computed is corrected by deformation fields, the image is still 2D. Therefore, the Gabor features computed in images of the same person but with varying pose are different, mainly for large rotation angles. This variation of the Gabor features with respect to the angle can be modeled as follows: First, the probability distribution functions of the scores $\mathcal{S}_j$ with respect to the rotation angle are computed in different blocks defined in the grids. A training set of images is used to compute this probability. Then, the estimated probability is used to correct the scores $\mathcal{S}_j$ computed in an input image depending on its rotation angle.

The image is divided in $V \times H$ blocks as shown in Figure 2.12, and a local probability model is built for each one. Instead of using all jets in the Borda count method, the Borda count is computed for each block individually in this approach. Thus, for the r-th block the comparison matrix C^r is extracted and the ranking matrix $O^r = BC(C^r)$ is computed. The matrix C^r is a subset of the comparison matrix C, that contains the inner product of the jets inside the r-th block. Note that if LN images are used, the matrices C and O are replaced by C^{ln} and O^{ln}. The identification score of the j-th gallery image in the r-th block $\mathcal{S}_j^r = \sum_{i=1}^{M_r} O_{ji}^r$, where M_r is the number of jets in the r-th block, is used as the input of the statistical model. The training and recognition are performed as follows:

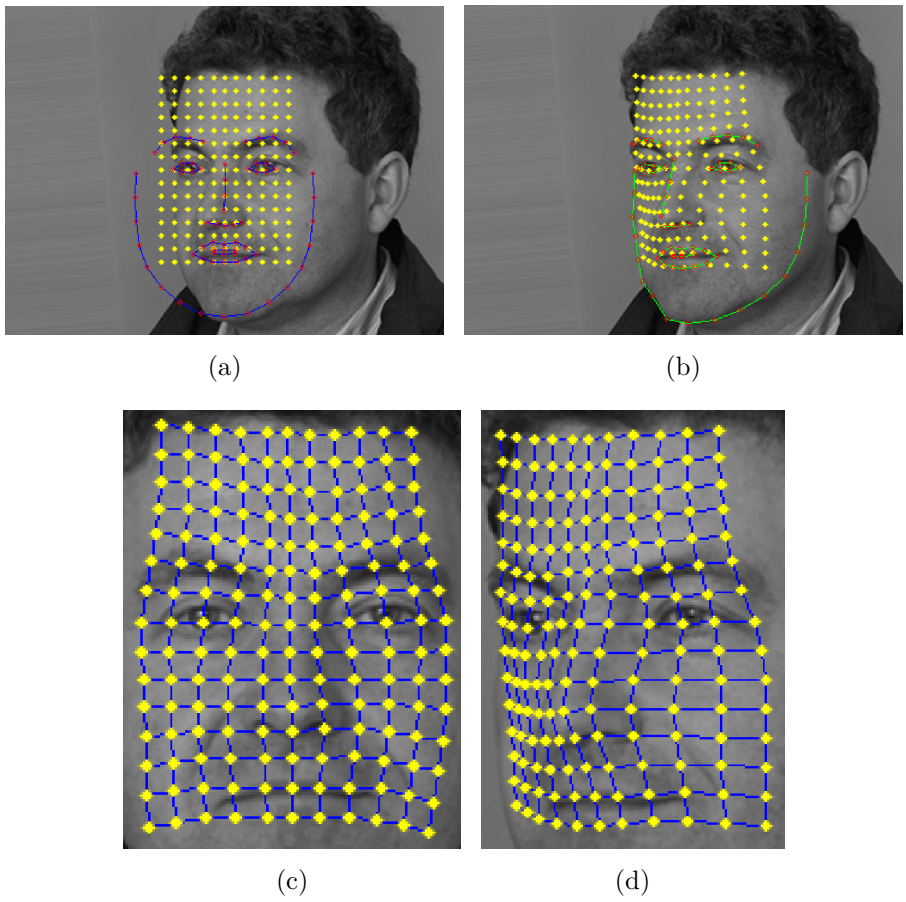


Figure 2.11: Example of the deformation of the grids used to compute the Gabor jets. (a) Initial grid (yellow) and mean face shape (blue). The grid is set as if the mean face shape were a face normalized. (b) Deformed grid (yellow) and adjusted face shape (green). The original grid is deformed to have the same relative position with respect to the adjusted shape that the original grid has with respect to the mean shape. By using this deformation, the position of the Gabor jets has the same relative position with respect to the face features in different images. Figures (c) and (d) show the position of the Gabor jets in a frontal and rotated face, respectively.

Training The prior distribution, $\mathcal{P}$, of the identification scores $\mathcal{S}^r$ (score for a given comparison in the r-th block) with respect to a pose and matching label is computed for each block r. The pose label is the rotation angle ϕ_p, and the matching label w can take two values: *same* if the input and gallery images to be compared are of the same person, and dif if they are of different persons. Thus, the prior distribution of the r-th block given a determine type of match and an

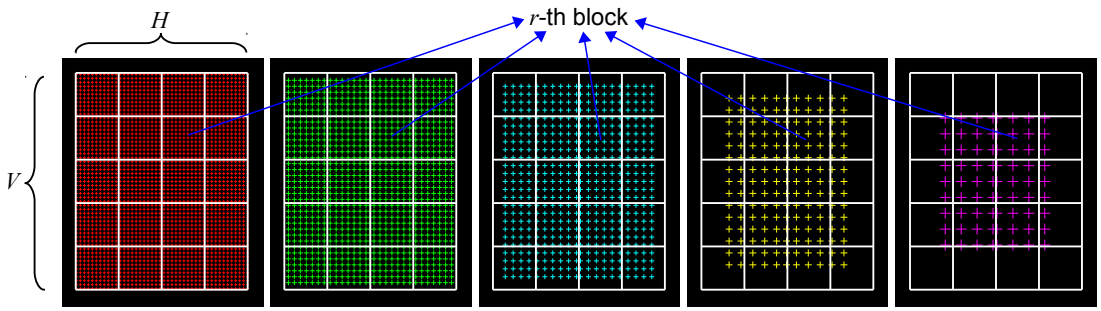


Figure 2.12: Division in Blocks for the Statistical Model. The five spatial scales used are shown. The white lines represent the different blocks and the + signs represent the spatial position where a Gabor jet is computed.

angle is:

$$\mathcal{P}\left(\mathcal{S}^r|w,\phi_p\right), w \in \{same, dif\}. \tag{2.12}$$

The prior distribution is modeled by using a normal distribution as in [Vu 2009]:

$$\mathcal{P}\left(\mathcal{S}^r|w,\phi_p\right) = \frac{1}{\sqrt{2\pi}\sigma^r_{w,p}} \exp\left[-\frac{1}{2}\left(\frac{\mathcal{S}^r - \mu^r_{w,p}}{\sigma^r_{w,p}}\right)^2\right], \tag{2.13}$$

where $\mu^r_{w,p}$, and $\sigma^r_{w,p}$ are the mean and the standard deviation of the scores $\mathcal{S}^r$ for the r-th block, and across all the training images with rotation angle ϕ_p and matching label w. To obtain these values the training database must contain a gallery set and test sets labeled with angles ϕ_p. The Borda count scores $\mathcal{S}^r$ are computed for all test images, and the labels w and ϕ_p are assigned to compute $\mu^r_{w,p}$ and $\sigma^r_{w,p}$.

Recognition To perform the recognition for the r-th block the Bayes rule is used. The probability of comparing the same person given the r-th score and an angle is:

$$\mathcal{P}\left(same|\mathcal{S}^r,\phi_p\right) = \frac{\mathcal{P}\left(\mathcal{S}^r|same,\phi_p\right)\mathcal{P}\left(same\right)}{\mathcal{P}\left(\mathcal{S}^r|same,\phi_p\right)\mathcal{P}\left(same\right) + \mathcal{P}\left(\mathcal{S}^r|dif,\phi_p\right)\mathcal{P}\left(dif\right)}. \tag{2.14}$$

If the entropy is used in the classification, the weighting is not applied to each jet as in the LMGEW method. In this case, an entropy is computed for each block. The entropy of the r-th block, E^r, is computed as the mean of the entropy

of all jets inside the block. To use $\mathcal{P}$ in the computation of the identification score of the j-th gallery image with respect to the input image, the rotation angle, ϕ_p, of the input image is also needed. This angle can be considered as known for each input image, or computed by using the shape S adjusted to the face. Finally, the identification score of the j-th gallery image with respect to the input image is computed as:

$$\mathcal{S}_j(\phi_p) = \frac{1}{V\ H} \sum_{r=1}^{V\ H} E^r \mathcal{P}\left(same|\mathcal{S}^r, \phi_p\right). \tag{2.15}$$

Chapter 3

Experiments and Results

Contents

A number of experiments were performed with the objective of assessing the proposed face recognition method. The most commonly used databases for face recognition evaluation were selected to compare results obtained in this work with those published in the literature. Section 3.1 contains descriptions of the databases used in this work. Section 3.2 describes the experiments for frontal face recognition and section 3.3 describes the experiments for face recognition across pose variation.

3.1 Face databases

3.1.1 The FERET database

Figure 3.1: Examples of face images taken from frontal view (top) and different view points (bottom) from FERET database. Face poses are shown only for left side rotation, right side rotation is analogous.

The FERET database [Phillips 1998] is one of the most widely used benchmarks for face identification methods. The FERET database has a large number of images with different gestures, illuminations, and a significant amount of time between pictures taken. Frontal FERET database is organized into 5 sets of images: the gallery is Fa, and the test sets are Fb, Fc, Dup1, and Dup2. In the Fa set there are 1,196 face images of different people. In the Fb set there are 1,195 images of people with different gestures. Fc has 194 images with different illuminations. In Dup1 there are 722 images taken between 0 and 34 months of difference from those of Fa. The Dup2 set has 234 images taken at least 18 months after the Fa set. The Fa set contains one image per person and is the Gallery set, while Fb, Fc, Dup1, and Dup2 are called test sets. Figure 3.1 shows images from different sets of the FERET database: (a) a neutral image from the Fa set, (b) an image with a different expression from the Fb set, (c) an image with an illumination change from the Fc set, and (d) an image taken with several months difference from the Dup1 set. To test faces with varying pose, 8 sets from ba through bi were taken to investigate the pose angle effects. Specifically, bf-bi are symmetric analogues of bb-be, while ba is the frontal face. The sets are: bb, bc, bd, be, bf, bg, bh, and bi with pose angles 0°, 60°, 40°, 25°, 15°, −15°, −25°, −40°, −60° respectively. Each of these sets has 200 images of 200 different people. Face examples for FERET database, bb, bc, bd, be, ba, bf, bg, bh and bi sets, are shown in Figure 3.1.

3.1.2 The AR face database

The AR database [Martinez 1998] contains frontal face images of men and women (60 females and 76 males) with different conditions of illumination, expression, and occlusion. Pictures were taken in two different sessions, with 13 pictures per session. It is called Session 1, S1, and Session 2, S2. Seven of the thirteen images contain illumination changes and gestures. In three images the person is wearing sunglasses, and in three other images, the person is wearing a scarf. Figure 3.2 shows images from the AR database: (a) a neutral face image, (b) an expression variation, (c) an illumination variation, (d) a face with sunglasses, and (e) a face with a scarf.

3.1.3 The FRGC 2.0 database

The FRGC 2.0 database [Phillips 2005, Phillips 2006] contains more than 50,000 images that are divided into training and testing partitions. The images are

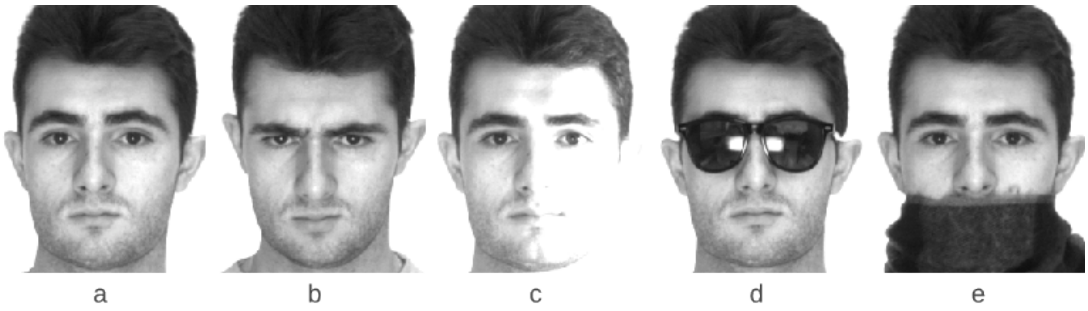


Figure 3.2: AR database example images. (a) Neutral face image, (b) expression variation, (c) illumination variation, (d) face with sunglasses and (e) face with scarf.

high resolution, with people looking frontally at the camera. There are images with and without controlled illumination. The FRGC 2.0 database has a total of six experiments related to different subsets in the database. Nevertheless, Experiments 1, 2, and 4 are the only ones for 2D face recognition; the rest of the experiments are for 3D face recognition. Experiment 1 contains controlled illumination images. Experiment 2 compares groups of images of the same individual, using the same images as in Experiment 1. Experiment 4 uses images under uncontrolled illumination, unfocused images, and some with small pose variations. Figure 3.3 shows examples of the FRGC 2.0 database; both (a) and (b) are gallery samples, (c) is a controlled image from Experiment 1, and (d) and (e) are uncontrolled images from Experiment 4.

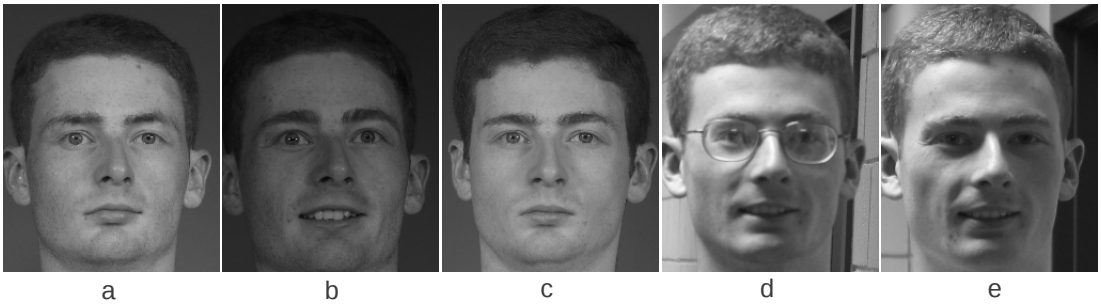


Figure 3.3:]
Examples of face images from the FRGC 2.0 database: (a) and (b) gallery images, (c) controlled images, (d) and (e) uncontrolled images.

3.1.4 The PIE face database

The CMU-PIE database [Sim 2003] is also one of the most commonly used databases for face recognition across pose variation. This database is composed of images of 68 different people. The images were acquired in 13 different poses, under 43 different illumination conditions, with 4 different expressions, and while the subjects were talking. To test the proposed face recognition method, the 13 poses of each subject were used, corresponding to cameras 22, 02, 25, 37, 05, 09, 27, 07, 29, 11, 31, 14 and 34. Images with frontal illumination (flash 11) were used. Face images for PIE database, 22, 02, 25, 37, 05, 09 and 27 sets, are shown in Figure 3.4; Oposites sets, 07, 29, 11, 31, 14 and 34, are analogs to those showed in Figure 3.4.

Figure 3.4: Examples of face images taken from different view points for PIE database. Images c25 and c09 taken with the camera located at an angle above the face. The other face images taken with a camera at the same height of the face. Face poses are shown only for left side rotation, right side rotation is analogous.

3.2 Experiments and Results for Frontal Face Recognition

3.2.1 Literature results

The results on face recognition in the literature are summarized in this section. In 2009 [Serrano 2009] provided an up to date summary highlighting the best

results reached in face recognition with the Local Gabor methodology. Results of more recent literature review and proposed methods on face recognition are summarized in Tables 3.1, 3.2 and 3.6 for different databases with frontal face cases. There are also shown results for faces with small pose variations; those that are rotated by ±15 and ±25 degrees in the FERET database are shown in Table 3.3. The first columns of each table shows the face identification rates for different subsets of different databases. The last column in Tables 3.1, 3.2, and 3.3 shows the total number of errors on all subsets of the databases.

Table 3.1: Rank-1 face recognition rate on different subsets of FERET database for different face recognition methods published in the literature.

Methods	Accuracy (%)				Errors
	Fb	Fc	Dup1	Dup2	Total
LMG [Zou 2007][a]	99.5	99.5	85.0	79.5	163
LGBPWP [Nguyen 2009][a]	98.1	98.9	83.8	81.6	185
LLGP [Xie 2009][a]	99.0	99.0	80.0	78.0	209
HGPP [Zhang 2007][a]	97.5	99.5	79.5	77.8	231
LGBPHS [Zhang 2005][a]	98.0	97.0	74.0	71.0	286
HEC [Su 2009][a]	99.00	99.00	92.00	88.00	99
LGBP+LGXP [Xie 2010][a]	99.00	99.00	94.00	93.00	77
LMGEW [Perez 2011a]	99.75	**100.00**	87.40	85.47	128
LMGEW-BTH [Perez 2011a]	99.75	**100.00**	90.17	88.03	102
LMGEW//LN-BTH [Cament 2014]	99.83	**100.00**	92.66	89.74	79
LGBP [Cament 2014] [b]	99.08	98.45	85.60	77.35	171
LGXP [Cament 2014] [b]	97.66	96.39	79.64	75.64	239
LMGEW//LN+LGBP [Cament 2014]	**99.92**	**100.00**	**95.57**	**93.59**	**48**
LMGEW//LN+LGXP [Cament 2014]	**99.92**	**100.00**	94.74	91.88	58

[a] results extracted from original source.
[b] results obtained with our implementation of the method.

3.2.2 Local Matching Gabor Entropy Weighted (LMGEW)

Table 3.1 summarizes the results obtained with the different methodological enhancements proposed in this work [Perez 2011a, Cament 2014]. LMG [Zou 2007] is shown on the first line for comparison. The first four columns of Table 3.1 show

Table 3.2: Rank-1 face recognition rate on the AR database.

Gal S1	Accuracy (%)			Errors
Methods	S2-Neutral	Scarves	Sunglasses	Total
LGBPHS [Zhang 2005][a]	—	98	80	—
SRC [Wright 2009][a]	95.7	93.5	**97.5**	84
LMG-BTH	99.00	98.83	89.50	77
LMGEW-BTH	99.29	99.00	95.33	39
LMGEW//LN-BTH	99.29	98.83	95.00	45
LGBP [Cament 2014]	99.71	97.83	65.83	220
LGXP [Cament 2014]	97.86	94.33	73.67	213
LMGEW//LN+LGBP [Cament 2014]	**100.00**	**99.83**	95.50	29
LMGEW//LN+LGXP [Cament 2014]	**100.00**	99.67	96.33	**24**

[a] results extracted from original source.

Table 3.3: Rank-1 face identification rate published in the literature and proposed methods on different subsets with varying pose of the FERET database.

	Accuracy (%)				Errors
Methods	bd	be	bf	bg	Total
LMG [Zou 2007][a]	81.0	97.0	98.0	79.5	89
LMGEW//LN-BTH	96.0	99.0	99.5	96.5	18
LGBP [Cament 2014]	86.5	98.0	97.5	88.5	60
LGXP [Cament 2014]	73.5	95.5	96.0	65.5	139
LMGEW//LN+LGBP [Cament 2014]	98.0	99.0	99.5	96.5	14
LMGEW//LN+LGXP [Cament 2014]	97.5	99.0	99.5	96.0	16

[a] results extracted from original source.

the results in percentage for the four different subsets of the FERET database, Fb, Fc, Dup1 and Dup2. The last column shows the total number of errors in all subsets of the FERET database. In the case of LMG [Zou 2007], the total number of errors in all subsets is 163. For LMGEW the total error was reduced from 163 to 128, which represents a 21% total error reduction.

The results on the AR database can be observed in Table 3.2. The database contains 76 images from males and 60 from females. Sets of 50 male and 50 female images are built. A gallery set is used with 7 images per person from session 1, containing 1 neutral image, 3 images with expressions and 3 images with different illumination. The set called S2-Neutral is identical to gallery set but with images from session 2. The set Sunglasses has 6 images per person using

sunglasses from both sessions and the set Scarves also 6 images per person using a scarf. Table 3.2 also shows a comparison between the proposed method and other previously published results [Wright 2009, Zhang 2005]. The LMGEW method improves in the Neutral AR subset from 98.86% to 99.00% when compared to LMG. For the Sunglasses subset the proposed method improved from 89.33% to 93.83% and for the Scarves subset from 96.67% to 97.67%. These results show that the LMGEW is better than LMG in all cases. Comparing with other methods published in the literature only for the Sunglasses subset, the method SRC [Wright 2009] is better than LMGEW.

3.2.2.1 Cascade LMGEW-BTH

The results obtained for the combined LMGEW-BTH enhancements are presented in Table 3.1. Figure 3.5 shows the face recognition accuracy for the training set of the FERET database using different thresholds in the interval 0.60-0.95. For this reason a threshold 0.84 is chosen, and LMGEW-BTH method obtained a total number of 102 errors compared to 163 using LMG. This represents a 37% improvement relative to the LMG method.

Table 3.2 shows results for the cascade LMGEW-BTH in the AR database. The last row shows a significant improvement for all AR subsets compared to LMG and LMGEW. If the results between the LMGEW-BTH and LMGEW are compared, the results in the Neutral subset improved from 99.00% to 99.29%, in the Sunglasses subset improved from 93.83% to 96.00% and in the Scarves subset from 97.67% to 99.50%.

3.2.2.2 Robustness to threshold selection for LMGEW-BTH

In section 3.2.2.1 a threshold range between 0.8 and 0.9 was determined from FERET training set. The robustness to threshold selection is tested by choosing three different thresholds in that range, 0.82, 0.85 and 0.88 for the FERET and the AR database. Table 3.4 shows the results for the FERET database and Table 3.5 for the AR database. For the FERET database there are slightly better results for threshold 0.85, but similar to those obtained for threshold 0.88 in the Dup2 subset. In the AR database there are no differences in the Neutral subset for the three thresholds, and differences are less than 1% among the other subsets for threshold 0.88. Therefore, results for AR and FERET databases show that if the threshold is chosen for any of the three values within the range 0.8 to 0.9 there are no significant differences in the results.

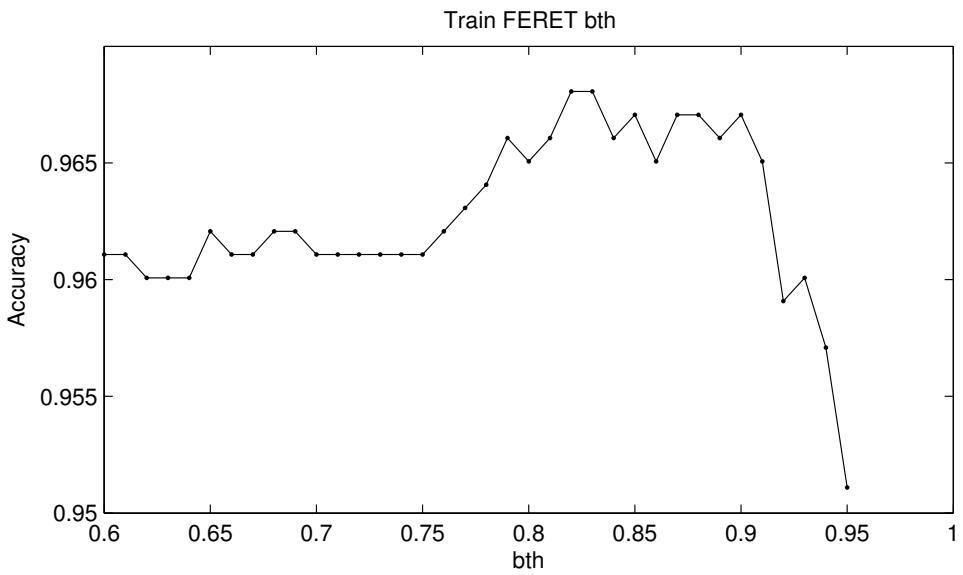


Figure 3.5: Identification rate on training FERET database varying T_h on LMGEW.

Table 3.4: Face recognition rate for three different thresholds for the LMGEW-BTH method in the FERET database.

	Accuracy (%)				Errors
Methods	Fb	Fc	Dup1	Dup2	Total
LMGEW	99.75	100.00	87.40	85.47	128
LMGEW-BTH-[0.82] [Perez 2011a]	99.75	100.00	89.47	86.32	111
LMGEW-BTH-[0.85] [Perez 2011a]	99.75	100.00	89.89	88.03	104
LMGEW-BTH-[0.88] [Perez 2011a]	99.83	100.00	89.34	86.32	111

* results extracted from original source.

3.2.2.3 Effect of error in eye localization for the LMGEW-BTH method

The LMG method is based on eye localization to align all faces. Therefore, it is important to assess the effect of eye localization error on face recognition for the LMG method. Gaussian noise is added to coordinates of the eyes for LMGEW-BTH method using a threshold of 0.88. Then, the image was aligned and cropped using the new eyes coordinates. Ten different standard deviation values are used for the Gaussian noise to generate errors from 1 to 10 pixels and a uniform probability for the angle between 0° and 360°. Figure 3.7 (a) shows the face recognition rate in FERET database and Figure 3.7 (b) for AR database.

Table 3.5: Face recognition rate for three different thresholds for the LMGEW-BTH method in the AR database.

	Accuracy (%)		
Methods	S2-Neutral	Sunglasses	Scarves
LMGEW	99.00	93.83	97.67
LMGEW-BTH-[0.82] [Perez 2011a]	99.29	95.33	99.00
LMGEW-BTH-[0.85] [Perez 2011a]	99.29	95.33	99.00
LMGEW-BTH-[0.88] [Perez 2011a]	99.29	96.00	99.50

* results extracted from original source.

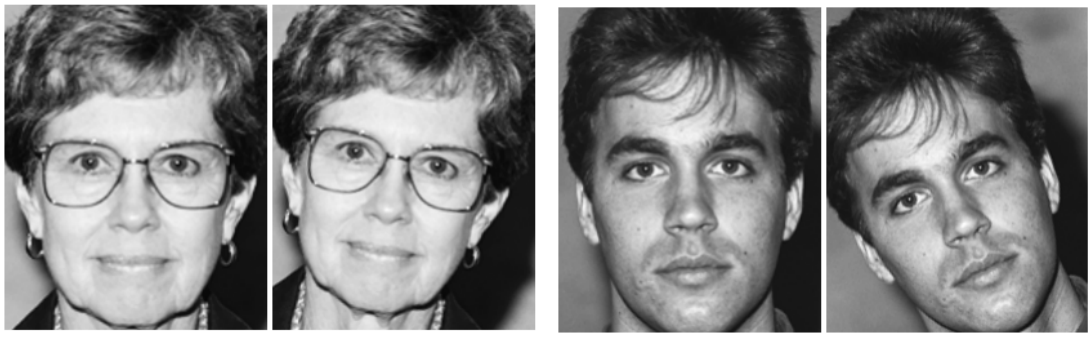
In the FERET database the image dimension is 256×384 pixels, and faces are about 150×150. For the FERET subsets Fb and Fc (expression and illumination changes), face recognition is very good up to a noise of 3 pixels and then decrease strongly. For Dup1 and Dup2 sets (time difference), the recognition rate decrease starting at 1 pixel of noise. Figure 3.6 shows some examples for alignment with noise. Figure 3.6 (a) and (b) show the difference between a well aligned image and a 5 pixels noise image. Figure 3.6 (c) and (d) show the difference of a well aligned image and a 10 pixels noise image. In the AR database the image dimension is 768×576 and the face size is about 400×400, almost three times larger than those of the FERET database. For S2-Neutral and Scarves subsets, the recognition rate is over 95% until a 7 pixels noise variance. For Sunglasses subset, recognition rate is over 90% until noise variance of 5 pixels. The three subsets do not show a fall in face recognition accuracy until a noise level over 3 pixels. Eye localization is an important issue but there are several methods that perform very precise eye localization [Perez 2010a].

3.2.3 Local Matching Gabor Entropy Weighted with Local Normalization (LMGEW//LN)

For all the experiments the Borda count threshold (BTH) used is 0.85, obtained in a previous section.

3.2.3.1 LMGEW//LN-BTH Results using the FERET database

The results obtained with LMGEW//LN-BTH and those published previously are presented in Table 3.1. The LMGEW//LN-BTH method obtained a total number of 79 errors with the FERET database. This result is better than all



a) b) c) d)

Figure 3.6: Face images with Gaussian noise in eyes position. (a) Image aligned without noise, (b) the image aligned with noise of 5 pixels in both eyes, (c) another image well aligned, (d) image aligned with noise of 10 pixels.

the best previously published results [Zou 2007]. Even though LMGEW//LN-BTH has the best overall results, most of the improvement was produced in the Fb and Fc subsets. The total number of errors was reduced by 22.5% on images compared to the LMGEW//LN-BTH method. HEC [Su 2009] and LGBP+LGXP [Xie 2010] obtained better results in the Dup2 subset, and LGBP+LGXP also obtained better results in subset Dup1. The LGBP+LGXP method reached 94% accuracy while LMGEW//LN-BTH obtained 92.66% in Dup1, and in Dup2 they reached 93% and 89.74%, respectively.

3.2.3.2 LMGEW//LN-BTH Results using the AR face database

Table 3.2 shows results for the LMGEW//LN-BTH and previously published methods in the AR database. As in [Wright 2009] 50 men and 50 women which each had 26 images available are randomly selected. A gallery set composed of 7 images of Session 1 (neutral, expression, and illumination variation images), and five test subsets are used. Subset S2-Neutral is similar to the gallery set, but contains images from Session 2. Subset Scarves contains faces occluded with scarves for Session 1 and Session 2, and subset Sunglasses contains faces occluded with sunglasses.

Results are almost the same as those of LMGEW-BTH, and better than the previously published identification methods. Only SRC has greater accuracy in the Sunglasses subset of AR, due to its robustness to occlusions.

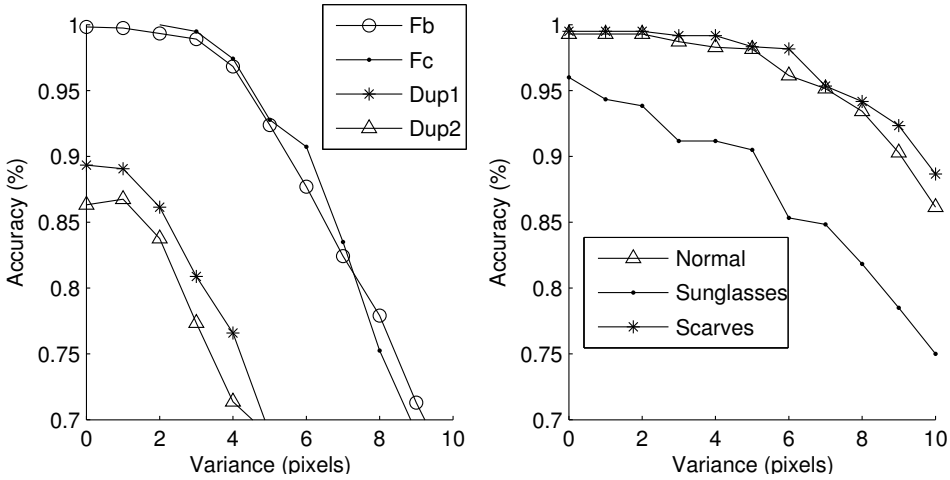


Figure 3.7: Rank-1 face recognition rate in the FERET (a) and AR (b) databases when a Gaussian noise with variance between 1 and 10 pixels is added over the eyes coordinates.

3.2.3.3 LMGEW//LN-BTH Results using the FRGC 2.0 database

For each experiment in FRGC 2.0 database, the Receiving Operator Characteristic (ROC) curves are constructed. ROC1 compares images taken within semesters, ROC2 within a year, and ROC3 during different semesters. Since the proposed method was designed for the face identification problem, the original is adapted from verification experiments of the FRGC 2.0 database to the identification problem using the Rank-1 method. Images of Experiments 1 and 4 from the ROC3 partition are used. For Experiment 1, the same gallery set containing 7,572 images and a sub-set of the test set without impostor images with a total of 6,512 images are used. For Experiment 4 the gallery set was formed with 7,512 gallery images and 3,256 test images.

Results of LMG methods are shown in Table 3.6. Experiment 1 shows a small difference between best and worst results, while the proposed method obtained the highest accuracy. In Experiment 4 the best score, 89.71%, was reached with the LMGEW-BTH, and the second best, 89.71%, was achieved using LMGEW//LN-BTH.

Table 3.6: Rank-1 face recognition rate on experiments 1 and 4 on the FRGC 2.0 database.

Methods	Accuracy (%)	
	Exp1	Exp4
LMG [Perez 2011b]	99.08	86.15
LMG-BTH [Perez 2011b]	99.23	89.10
LMGEW-BTH [Perez 2011a]	99.62	89.71
LMGEW//LN-BTH [Cament 2014]	99.69	88.57

3.2.3.4 Effect on face identification resulting from error in eye localization with the LMGEW//LN-BTH method

The LMG and derived methods require eye localization to align faces. Therefore it is important to assess the effect of eye localization error on face identification for the LMG method. Noise is added to the eye coordinates for the LMGEW//LN-BTH method. The error distance radius varied between 1 and 10 pixels, and the angle between the center and the computed position between 0° and 360°, both, radius and angle, using a uniform probability. Then, the images were aligned and cropped using the new eye coordinates. Figure 3.8 shows examples of alignment with noise in eye localization. Figures 3.8 (a) and (b) show the difference between a well-aligned image and an image with 4 pixels of noise. Figures 3.8 (c) and (d) show the difference between a well- aligned image and an image with 10 pixels of noise. Figure 3.9 shows the face identification rate in the FERET database for subsets Fb, Fc, Dup1, and Dup2.

In the FERET database the image dimension is 256×384 pixels, and faces vary approximately between 75×95 to 180×260 pixels, which means that an error of 10 pixels in both eyes is significant. Results for the FERET subsets Fb and Fc (expression and illumination changes), show that face identification is almost unaffected by noise levels of up to 4 pixels. For errors greater than 4 pixels, the identification rate falls significantly with increasing eye localization error. For the Dup1 and Dup2 sets (time difference), the identification rate decreases significantly with more than 2 pixels of noise.

3.2.4 Entropy-like weighting on LGBP and LGXP features

LGBP and LGXP are tested computing entropy-like values for each local classifier, i.e., for each block. The feature vector of each block on the test image is

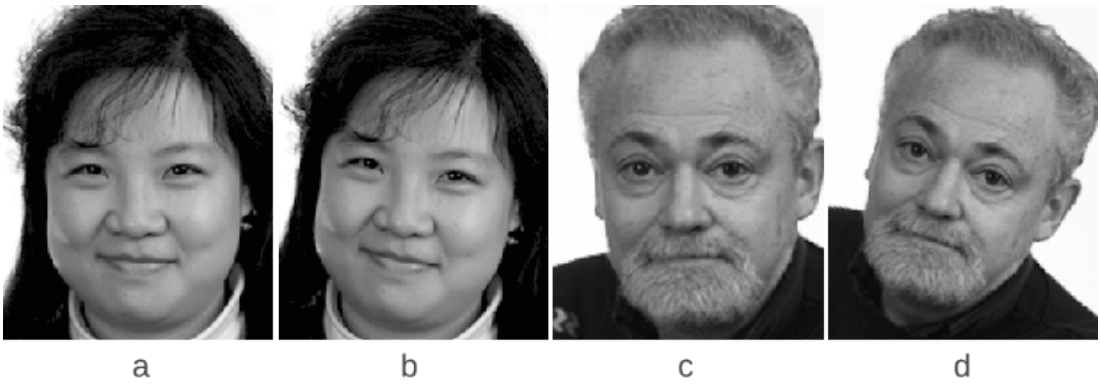


Figure 3.8: Face images with Gaussian noise in eye position. (a) Image aligned without noise, (b) the image aligned with noise of 4 pixels in both eyes, (c) another well aligned image, (d) image aligned with noise of 10 pixels.

compared with the same vector of the gallery image. Then, a comparison matrix is built containing comparisons between all blocks of each test image and the blocks of each gallery image. With this matrix, entropy-like values are computed and used to weight with the same strategy described in section 3.1. The results for Fb, Fc, Dup1, and Dup2 are 99.16%, 97.94%, 86.29%, 79.06% using LGBP; and 97.66%, 95.88%, 79.64%, 76.50% using LGXP. Results showed small improvements (0.5% for LGBP, and 0.1% for LGXP) in face recognition when the entropy strategy was applied to LGBP and LGXP compared to the case where the entropy strategy was not applied. A possible explanation for these results is that the local features are already significantly reduced by the application of LDA over the histograms of LGBP and LGXP features. Therefore, the entropy strategy does not improve results in LGBP and LGXP as it does in the case of the LMG method where features are not reduced.

3.2.4.1 Results of the fusion of methods LMGEW//LN-BTH+LGBP and LMGEW//LN-BTH+LGXP

The LGBP and LGXP methods proposed in [Xie 2010], and LMGEW//LN-BTH [Cament 2014] were combined, resulting in an improvement of face identification rates in all subsets of the FERET and AR databases. A fusion of the methods LMGEW//LN-BTH+LGBP and LMGEW//LN-BTH+LGXP is done at the score level as shown in (2.8). S_1 denotes the identification score of LMGEW//LN-BTH, S_2 the score of LGBP or LGXP.

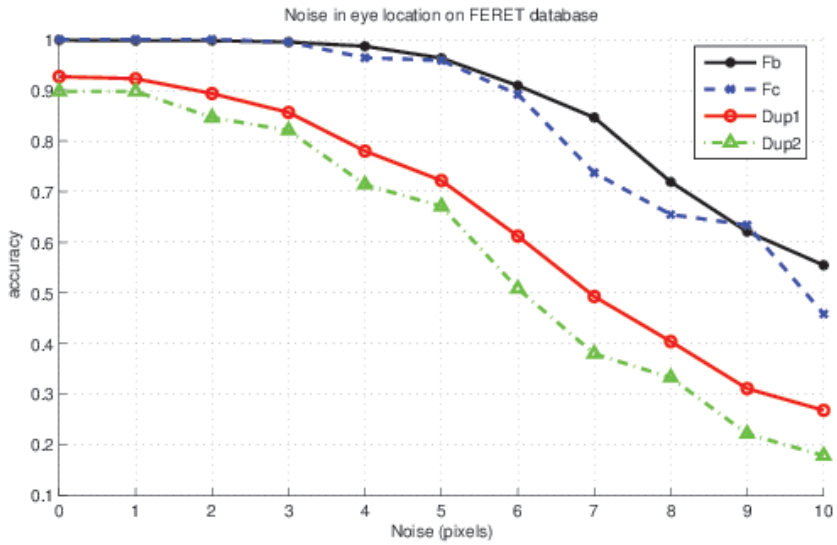


Figure 3.9: Rank-1 face identification rate using the FERET database when noise between 1 and 10 pixels is added to the eye coordinates.

Before fusing both methods, the ranges for the LMG, LGBP and LGXP scores were normalized to $[0, 1]$. LGBP and LGXP take values in the range $[-2, 16]$ while LMG takes values between $[100, 600]$. Therefore, the normalization was $S^{norm}_{LGBP} = (S_{LGBP} + 2)/18$, for LGBP and LGXP. For LMG the normalization was $S^{norm}_{LMG} = (S_{LMG} + 100)/500$. The weight w was varied from 0 to 1 as shown in Figure 3.10. It can be observed that for values of w between 0.5 and 0.85 the fusion achieved the best results compared with any of the methods applied independently. The weight $w = 0.75$ is chosen and used for all the following experiments.

Fusion of the LMGEW//LN-BTH+LGBP and LMGEW//LN-BTH+LGXP methods using the FERET database yielded the same results in Fb and Fc with 99.92% and 100% accuracy respectively. Dup1 was improved by 1.57% and Dup2 by 0.59% compared with LGBP+LGXP. The total number of errors decreased from 77 in LGBP+LGXP to 48 in LMGEW//LN-BTH+LGBP which is a 38% improvement relative to previously published methods.

Fusion of LMGEW//LN-BTH+LGBP and LMGEW//LN-BTH+LGXP methods in the AR database also improved the best results obtained previously with LMG-based methods for all sub-sets. For the S2-Neutral subset both fusions reached 100% accuracy, improving the 99.29% performance of LMGEW-BTH. In the Scarves subset both fusions improved compared with previous methods, and

LMGEW//LN-BTH+LGBP reduced the error rate by 1%. In the Sunglasses subset, the fusion LMGEW//LN-BTH+LGXP method improved the results but could not reach the best results obtained by SRC [Wright 2009].

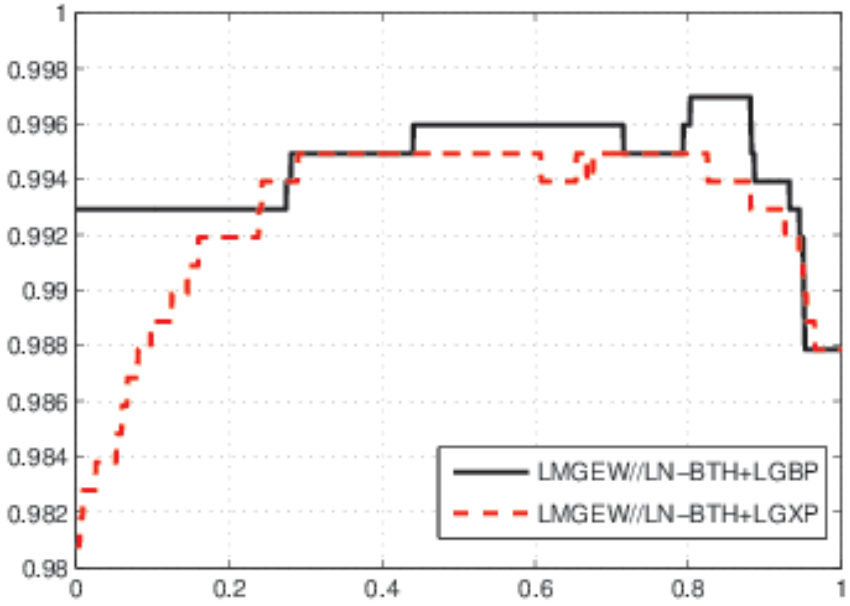


Figure 3.10: Results for the fusion LMGEW//LN-BTH+LGBP and LMGEW//LN-BTH+LGXP as a function of the weight w in FERET database.

3.2.5 Results of face identification with small pose variations

LMGEW//LN-BTH, LGBP, LGXP methods were tested and combined with the subsets bd, be, bf, and bg with pose variations of the FERET database as shown in Figure 3.1. Table 3.3 shows the results of face recognition for small pose variations, and also the best previously published results.

The proposed methods improve these results significantly for faces with small pose variations. LMGEW//LN-BTH showed a 80% in the error rate improvement compared to previous LMG results, reducing the total number of errors from 89 to 18. Results improved by 0.5% and 1.5% for $\pm 15°$ face pose rotation on the subsets be and bf, respectively. Results also improved 2.5% and 5% for subsets bd and bg, respectively, which have $\pm 25°$ face rotations. Fusion improved the

results slightly, reducing the errors by 4 and 2 with LMGEW//LN+LGBP and LMGEW//LN+LGXP, respectively. These results show improvements in using LMGEW//LN-BTH and fusion with LGBP and LGXP for face identification with small pose variations.

3.3 Experiments and Results for Pose Variation Face Recognition

3.3.1 Grid Deformation

The effectiveness of an ASM to deform the grid and adjust the Gabor filters to pose variation was measured by the following experiments:

- Face identification performing only a face alignment, without using the grid deformation, as described in section 2.2.1.2 and shown in Figure 2.11(a).
- Face identification performing grid deformation, as described in section 2.2.2 and shown in Figure 2.11(b).

Both experiments were performed without any additional improvement of the LMG method so that only the improvement due to the grid deformation was measured. Then, the experiments were repeated using two different variants of the proposed method to measure the improvement achieved by using each of them. The first variant includes a Borda count threshold (BTH) and an Entropy weighting (E). The BTH eliminates very small value scores that act as noise for the Borda count computation. The E weighting is used to emphasize the Gabor jets with the best performance. The second variant, in addition to BTH and E, includes the Local Normalization (LN) introduced in [Perez 2010b], which is used to compensate for illumination conditions.

Table 3.7 shows the Rank-1 results on the FERET database using only face alignment, or face alignment with grid deformation, and three different conditions. Both methods were tested without any other improvement of the LMG method, and then using the variants of the LMG which include BTH, E, and LN. As shown, the best results reach 100% for face pose near frontal, $-25°$ to $25°$. However, for larger pose changes, $40°$ or more, the identification results decrease significantly and differences among different variants of the proposed method become evident. For example, when only alignment is used in images with pose variation of $60°$, the correct identification rate is 34.5%. Using grid deformation,

results improve to 59.5%. As shown in Table 3.7 the results obtained with grid deformation are significantly better than those obtained with only alignment. This improvement can be explained because with the grid deformation, the extracted features are closer to the features extracted in the frontal face enrolled in the gallery. This effect is especially noticeable in faces with large pose variation. If BTH, LN, and E are combined with face alignment and grid deformation, the recognition rate improves significantly achieving the highest results. This improvement can be seen in Table 3.7, particularly for the $\pm 40^\circ$ and $\pm 60^\circ$ rotation angles. For example, the results using grid deformation (GD) increase 10% for the $\pm 40^\circ$ (GD+LN+BTH+E), and 32% for the $\pm 60^\circ$, if LN, BTH and E are used.

Table 3.8 shows the results of the same experiments described above but carried out using the PIE database. In same way as in the FERET database, the results are 100% for the near frontal poses (05, 09, 07, 29). However, for larger rotation angles the performance decreases. Results obtained using grid deformation are always better than those with alignment only. The best performance was obtained using grid deformation with BTH and E, in this database. The performance did not improve when LN was used. For example, the performance obtained using grid deformation increased approximately 43.5% for the largest rotation angles (sets 22 and 34) when BTH and E were used, but it increased only 28.2% if BTH, E and LN were used.

3.3.2 Statistical Model

The statistical model was trained for each block. Each of the FERET eight pose sets (bb, bc, bd, be, bf, bg, bh, bi) contains 200 images, including the gallery set. Each set was partitioned into two subsets, the first to train the model, and the second for testing. The training subsets were built by randomly choosing 100 images of the corresponding set, and the remaining 100 images were used to test the model. A cross-validation was performed exchanging the training and test sets, i.e., training with the second partition and testing with the first one. In the CMU-PIE database, an analogous procedure was carried out. Each of the 13 pose sets of the PIE database contains 68 images; therefore, 34 were randomly selected for the training set and 34 for the testing set to perform the cross-validation. The combinations of grid deformation, LN, BTH, and E with best performance in the previous experiments (section 3.3.1), were used to evaluate the performance of

the statistical model in each database. The method was tested using different numbers of blocks for the statistical model (see Figure 2.12).

In order to measure the performance of the statistical model (P) in combination with the other methods used in the system, a set of experiments similar to those presented in section 3.3.1 were performed. Face identification performing only a face alignment, or also using grid deformation, was carried out using P, and combinations of BTH, E, and LN.

Table 3.9 shows the results of the classification performance on the FERET database for different arrays of blocks used to test the statistical model. The best result was reached with an array size of 20×25, therefore this size was selected for further testing of the statistical model. Table 3.10 shows analogous results on the CMU-PIE database. The best result for the PIE database was achieved with an array size of 8×10, accordingly this size was selected for further testing on the PIE database.

Table 3.11 shows the results of face identification using only face alignment, or face alignment with grid deformation, with P and combinations of BTH, E, and LN in the FERET database. Results show that the use of the statistical model improves the recognition rate significantly, as can be seen by comparison with those in Table 3.7. The recognition rates reach 100% for near frontal images, and nearly 99% for $\pm 40°$ pose angles. For $\pm 60°$ pose angles the improvement is roughly 16% of the best result shown in Table 3.7.

Table 3.11 shows that the results using GD are always better than those using A. The best result was achieved by the combination GD+P+BTH+LN, with a mean performance of $96.4 \pm 0.2\%$. However, this performance was very close to that obtained by the combinations GD+P+BTH, GD+P+BTH+E and GD+P+BTH+LN+E. The results show that the improvement in the performance is similar using only BTH, or combinations of BTH, E and LN.

Table 3.12 shows the results of using P and combinations of BTH, E and LN on the CMU-PIE database. In the same way as on the FERET database, the results improved significantly by using the statistical model. Considering the experiments with better average results in Tables 3.8 and 3.12, the performance improvement in the largest rotation angles (sets 22 and 34) is 53% when the statistical model is used. Also the results using GD are always better than when using A.

The best performance was achieved using GD+P on the CMU-PIE database, with a mean performance of $86.1 \pm 1.4\%$ (Table 3.12). The performance of combinations GD+P+BTH and GD+P+BTH+E were close to this result, nevertheless,

the combinations using LN did not improve the results.

3.3.3 Gallery Sets with Different Pose Angles

In order to measure the robustness of the method when images of rotated faces are used as gallery, different pose sets were enrolled. The sets bc, bd, be, bf, bg and bh of the FERET database were enrolled as gallery. The sets bb and bi were not used as gallery because one side of the face is usually not visible in these images due to the large rotation ($\pm 60°$). In the CMU-PIE database, the sets of cameras 37, 05, 09, 27, 07, 29 and 11 were used as gallery, and the sets 22, 02, 25, 31, 14 and 34 were not used because of their large rotation. This experiment also allows comparison with other published methods in which the results are presented using rotated faces as galleries.

Table 3.13 shows the results of enrolling sets of images with various pose angles, between $-40°$ and $40°$, as gallery in the FERET database. Results show that the highest recognition performance was reached consistently for faces with pose closest to the enrolled face angle. Even if faces with a small pose variation are used as gallery, the recognition rate of faces with large pose variation improves significantly. Using galleries with just $\pm 15°$, the average result in the test sets with $\pm 60°$ of rotation in the same direction is 94.6%.

Table 3.14 shows the results of using images with different pose angles as gallery in the CMU-PIE database. Also, the highest recognition rates were consistently reached for faces with pose angle closest to the enrolled faces. Using the sets of cameras 37 and 11 as gallery, the average results in the sets with largest rotations is 86.1%.

3.3.4 Comparison with other Methods

The results of the proposed method were compared with those published previously with pose variations on the FERET database. Table 3.15 shows the performance of 7 different 2D methods and the proposed method on the FERET database. Comparing the results, it can be concluded that the proposed method reaches the highest overall classification performance.

Table 3.16 shows the performance of 5 different 3D methods and the proposed method on the FERET database. Some methods consider the set with $15°$ of rotation as the gallery. To perform an equivalent comparison with these methods, the sets with $15°$ and $-15°$ of pose were used both individually and together as gallery. The proposed method outperformed all the methods that use the set with

0° of pose as gallery. Only one of the methods that uses the set with 15° of pose as gallery had a performance that exceeded that of the proposed method, when sets with 15° or −15° of pose are used as gallery individually [Paysan 2009] the proposed method gets a comparable performance. Nevertheless, the sets in which the proposed method is surpassed are only those with ±60° of face rotation. If both sets together (15° and −15°) are used as gallery, the method achieves the best performance. To use two sets as gallery, the rotation angle of the test face is computed by using the shape adjusted to the face, and then the comparison is carried out using the gallery set with the closest rotation angle.

Table 3.17 shows the results of 10 different 2D methods and the proposed method on the CMU-PIE database. The proposed method is compared using the set with frontal faces (27) and also two sets with rotated faces (37 and 11) as gallery. Since few methods show results in just some sets, the mean performance of the proposed method is also shown taking only the sets used for each method into account. By using the frontal faces as gallery, the proposed method outperforms 6 of the 10 methods. This overall performance is caused by poor recognition in the sets with large rotation angles, although the performance in near frontal faces is 100%. If two sets of rotated faces are used as gallery, the performance of the proposed method is increased, surpassing that of all the comparison methods.

Table 3.18 shows the performance of 5 different 3D methods and the proposed method on the CMU-PIE database. The performance of the proposed method is shown using one and two sets as gallery. If only frontal faces are used in the gallery, the method outperforms only 2 of the comparison methods. Nevertheless, when two sets are used in the gallery, the method outperforms 4 of the comparison methods.

Table 3.7: Rank-1 results on the FERET database. The results of using only alignment (A), or grid deformation (GD) to compute the position of the Gabor jets were tested. These two methods were tested without including any other improvement, and using two variants of the LMG method. A first variant of the proposed method includes Borda count with threshold (BTH) and Entropy weighting (E) of the Gabor jets. A second variant additionally includes a Local Normalization (LN) for illumination compensation.

Method \Set	bb (60°) [%]	bc (40°) [%]	bd (25°) [%]	be (15°) [%]	bf (−15°) [%]	bg (−25°) [%]	bh (−40°) [%]	bi (−60°) [%]	Mean [%]
A	34.5	84.0	96.5	99.5	100.0	98.5	74.0	36.5	77.9
GD	59.5	88.5	98.5	99.5	100.0	98.0	86.5	53.5	85.5
A+BTH+E	43.5	95.0	99.5	100.0	100.0	99.5	88.5	47.0	84.1
GD+BTH+E	71.0	94.5	99.0	99.5	100.0	99.5	94.5	67.5	90.7
A+LN+BTH+E	46.0	97.5	99.5	100.0	100.0	100.0	90.5	48.0	85.2
GD+LN+BTH+E	77.5	97.5	100.0	100.0	100.0	100.0	95.5	71.5	92.8

Table 3.8: Rank-1 results on the CMU-PIE database. The results of using only alignment (A). or grid deformation (GD) to compute the position of the Gabor jets were tested. These two methods were tested without including any other improvement. and using two variants of the LMG method. A first variant of the proposed method includes Borda count with threshold (BTH) and Entropy weighting (E) of the Gabor jets. A second variant additionally includes a Local Normalization (LN) for illumination compensation.

Method \Set	22 [%]	02 [%]	25 [%]	37 [%]	05 [%]	09 [%]	07 [%]	29 [%]	11 [%]	31 [%]	14 [%]	34 [%]	Mean [%]
A	25.0	54.4	51.5	95.6	100.0	100.0	100.0	100.0	69.1	55.9	38.2	22.1	67.6
GD	26.5	64.7	60.3	98.5	100.0	100.0	100.0	100.0	97.1	69.1	61.8	20.6	74.9
A+BTH+E	30.9	70.6	67.6	100.0	100.0	100.0	100.0	100.0	80.9	54.4	51.5	25.0	73.4
GD+BTH+E	42.6	79.4	76.5	98.5	100.0	100.0	100.0	100.0	98.5	72.1	57.4	25.0	79.2
A+LN+BTH+E	27.9	73.5	64.7	100.0	100.0	100.0	100.0	100.0	75.0	58.8	44.1	26.5	72.5
GD+LN+BTH+E	33.8	70.6	75.0	98.5	100.0	100.0	100.0	100.0	98.5	69.1	67.6	26.5	78.3

Table 3.9: Performance of the statistical model on the FERET database using different array sizes of $H \times V$ blocks. The grid deformation, LN, BTH, and E were used in the tests.

Arrays of Block divisions \Set	bb (60°) [%]	bc (40°) [%]	bd (25°) [%]	be (15°) [%]	bf (−15°) [%]	bg (−25°) [%]	bh (−40°) [%]	bi (−60°) [%]	Mean [%]
4×5	87.0±1.4	99.2±0.2	100.0±0.0	99.5 ±0.0	100.0±0.0	100.0±0.0	97.8±0.4	81.9±0.7	95.7±0.3
8×10	87.4±0.7	99.5±0.2	100.0±0.0	99.9 ±0.2	100.0±0.0	100.0±0.0	97.9±0.5	83.0±1.4	95.9±0.4
12×15	87.5±0.6	99.5±0.0	100.0±0.0	99.8 ±0.3	100.0±0.0	100.0±0.0	97.7±0.5	84.3±0.9	96.1±0.3
16×20	88.6±0.9	99.5±0.0	100.0±0.0	99.7 ±0.2	100.0±0.0	100.0±0.0	97.9±0.3	82.0±1.4	96.0±0.4
20×25	89.3±0.6	99.3±0.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	97.9±0.5	84.2±1.3	96.3±0.3
24×30	88.5±1.0	99.5±0.0	100.0±0.0	100.0±0.2	100.0±0.0	100.0±0.0	97.9±0.5	81.5±1.2	95.9±0.3
28×35	87.9±1.2	99.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.5±0.0	82.4±1.4	96.0±0.3
32×40	88.0±1.2	99.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.1±0.3	82.3±1.9	96.0±0.4
36×45	87.4±1.0	99.4±0.2	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.2±0.3	81.0±1.3	95.8±0.4
40×50	87.2±0.9	99.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.3±0.4	81.5±1.8	95.8±0.4
44×55	86.6±1.2	99.4±0.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	97.8±0.5	81.6±1.0	95.7±0.4

Table 3.10: Performance of the statistical model on the CMU-PIE database using different array sizes of $H \times V$ blocks. The grid deformation, BTH and E were used in the tests.

Arrays of N° Blocks \Set	22 [%]	02 [%]	25 [%]	37 [%]	05 [%]	09 [%]	07 [%]	29 [%]	11 [%]	31 [%]	14 [%]	34 [%]	Mean [%]
4×5	61.0±4.3	82.2±2.5	88.1±2.4	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	71.8±3.6	62.4±3.7	43.2±2.5	84.1±1.6
8×10	64.1±3.9	89.9±2.2	91.6±2.2	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	76.0±3.7	74.3±3.5	46.2±4.9	86.8±1.7
12×15	58.8±3.3	89.1±1.2	90.7±2.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.9±0.5	72.8±3.0	69.0±3.1	43.7±4.0	85.3±1.5
16×20	59.6±1.9	88.5±2.3	91.2±2.9	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.6±0.7	75.4±3.7	69.0±4.1	42.8±4.1	85.5±1.6
20×25	56.9±3.1	88.4±1.9	88.2±1.8	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.1±0.8	78.1±3.8	65.4±3.6	36.8±3.7	84.4±1.6
24×30	57.5±2.7	88.1±2.5	87.2±1.8	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.0±0.7	77.5±3.5	64.4±4.1	39.1±4.4	84.4±1.7
28×35	61.3±3.6	89.3±2.7	88.2±2.2	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.4±0.8	79.3±3.6	65.3±3.4	35.9±4.6	84.9±1.7
32×40	57.6±4.8	87.5±1.4	86.8±2.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.0±0.7	79.0±3.7	62.8±5.1	34.4±3.7	83.9±1.8
36×45	54.1±5.0	86.6±1.6	87.1±3.5	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.6±0.7	78.2±3.2	61.6±3.6	38.4±5.3	83.8±1.9
40×50	54.3±3.4	88.1±1.6	86.6±3.4	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.6±0.7	74.9±3.3	59.6±2.7	32.9±3.5	83.0±1.5
44×55	51.8±2.2	86.3±1.0	84.9±1.7	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	99.4±0.8	76.3±4.2	59.1±3.5	34.9±5.2	82.7±1.5

Table 3.11: Rank-1 results of the proposed methods in FERET database, using only alignment (A), or alignment with grid deformation (GD), with the statistical model (P) and combinations of a Borda count with threshold (BTH), Entropy weighting (E), and Local Normalization (LN).

Method \Set	bb (60°) [%]	bc (40°) [%]	bd (25°) [%]	be (15°) [%]	bf (−15°) [%]	bg (−25°) [%]	bh (−40°) [%]	bi (−60°) [%]	Mean [%]
A+P	56.2±1.2	96.2±0.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.0±0.6	56.9±1.7	88.0±0.5
GD+P	83.2±1.4	98.3±0.4	100.0±0.0	99.6 ±0.2	100.0±0.0	100.0±0.0	97.9±0.3	82.9±1.7	95.2±0.5
A+P+BTH	60.7±1.7	97.9±0.2	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.7±0.5	55.7±2.1	88.7±0.6
GD+P+BTH	87.4±1.1	98.9±0.2	100.0±0.0	99.7 ±0.2	100.0±0.0	100.0±0.0	99.0±0.3	83.4±1.0	96.0±0.4
A+P+BTH+E	58.8±1.3	97.7±0.4	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	96.1±1.0	55.3±1.6	88.5±0.5
GD+P+BTH+E	87.5±1.4	98.7±0.3	100.0±0.0	99.7 ±0.2	100.0±0.0	100.0±0.0	98.6±0.4	82.9±1.0	95.9±0.4
A+P+BTH+LN	58.8±1.7	98.7±0.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.7±0.6	59.9±1.1	89.1±0.5
GD+P+BTH+LN	89.0±0.8	99.5±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.1±0.4	84.6±0.8	96.4±0.2
A+P+BTH+LN+E	56.1±1.5	98.0±0.4	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	95.5±0.6	57.8±0.9	88.4±0.4
GD+P+BTH+LN+E	89.3±0.6	99.3±0.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	97.9±0.5	84.2±1.3	96.3±0.3

Table 3.12: Rank-1 results of the proposed methods in CMU-PIE database. using only alignment (A). or alignment with grid deformation (GD). with the statistical mode (P) and combinations of a Borda count with threshold (BTH). Entropy weighting (E). and Local Normalization (LN).

Method \Set	22 [%]	02 [%]	25 [%]	37 [%]	05 [%]	09 [%]	07 [%]	29 [%]	11 [%]	31 [%]	14 [%]	34 [%]	Mean [%]
A+P	38.4±2.2	79.7±2.4	74.6±2.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	91.9±1.7	62.5±5.0	51.6±4.2	31.8±2.8	77.5±1.7
GD+P	56.6±3.5	86.9±2.9	85.6±1.7	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	80.7±2.8	76.3±2.1	46.8±4.3	86.1±1.4
A+P+BTH	40.4±3.1	82.1±2.3	76.0±2.5	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	87.8±2.3	67.4±5.0	58.1±3.0	34.6±1.7	78.9±1.7
GD+P+BTH	61.2±3.6	90.9±2.3	92.2± 2.2	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	79.0±4.0	74.3±1.9	47.5±3.7	87.1±1.5
A+P+BTH+E	37.6±1.6	80.0±1.7	75.4±2.9	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	87.6±3.4	65.3±4.3	52.9±3.2	34.6±2.4	77.8±1.6
GD+P+BTH+E	64.1±3.9	89.9±2.2	91.6±2.2	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	76.0±3.7	74.3±3.5	46.2±4.9	86.8±1.7
A+P+BTH+LN	40.9±4.1	81.3±2.2	74.4±4.3	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	90.0±2.1	66.3±5.5	54.0±2.7	35.1±4.9	78.5±2.1
GD+P+BTH+LN	55.3±4.2	87.4±2.2	89.0±2.2	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	80.1±3.6	77.6±2.8	45.1±5.5	86.2±1.7
A+P+BTH+LN+E	37.8±5.1	80.1±1.7	74.0±3.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	87.4±2.4	63.1±6.2	53.8±3.0	33.4±3.5	77.5±2.1
GD+P+BTH+LN+E	53.7±3.5	86.3±1.7	87.9±2.9	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	80.4±3.9	74.9±3.1	47.9±5.0	85.9±1.7

Table 3.13: Rank-1 performance for the method using galleries with different pose angles in FERET database. The GD, P, BTH and LN were used in the tests.

Yaw rotation of the gallery \Set	bb (60°) [%]	bc (40°) [%]	bd (25°) [%]	be (15°) [%]	ba (0°) [%]	bf (−15°) [%]	bg (−25°) [%]	bh (−40°) [%]	bi (−60°) [%]	Mean [%]
40°	98.5±0.0	gallery	99.5±0.0	99.5 ±0.3	99.2 ±0.0	99.0 ±0.7	92.8 ±1.4	77.0±1.6	34.4±1.6	87.5±0.7
25°	99.3±0.4	99.5±0.0	gallery	100.0±0.0	100.0±0.0	100.0±0.3	99.3 ±1.4	93.5±2.2	50.3±2.2	92.7±0.8
15°	96.8±0.4	99.5±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.2	99.1 ±0.4	96.5±0.9	66.2±0.9	94.8±0.3
0°	89.0±0.8	99.5±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.0	98.1±0.4	84.6±0.8	96.4±0.2
−15°	69.2±1.4	99.3±0.3	100.0±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.3	99.8±0.6	93.5±0.6	95.2±0.4
−25°	49.0±1.5	93.5±1.3	98.8 ±0.3	99.3 ±0.4	100.0±0.0	100.0±0.0	gallery	99.8±0.3	98.0±0.3	92.3±0.5
−40°	25.5±1.4	76.8±1.2	89.0 ±0.9	96.3 ±1.0	96.4 ±0.3	99.0 ±0.0	99.5±0.0	gallery	99.5±0.2	85.2±0.6

Table 3.14: Rank-1 performance for the method using galleries with different pose angles in the CMU-PIE database. The GD, P and BTH were used in the tests.

Gallery set \Set	22 [%]	02 [%]	25 [%]	37 [%]	05 [%]	09 [%]	27 [%]	07 [%]	29 [%]	11 [%]	31 [%]	14 [%]	34 [%]	Mean [%]
37	97.5±0.7	99.7±0.6	99.9± 0.5	gallery	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	98.2±1.5	77.8±3.5	18.4±4.4	19.4±3.4	12.2±2.9	76.9 ±1.5
05	88.1±2.4	99.4±0.8	99.1± 0.8	100.0±0.0	gallery	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	97.8±0.8	41.0±5.1	39.9±4.1	24.3±3.3	82.5±1.4
09	58.5±2.6	94.6±1.0	94.6± 1.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	80.6±2.7	72.1±3.2	43.7±4.2	87.0±1.2
27	61.2±3.6	90.9±2.3	92.2± 2.2	100.0±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.0	100.0±0.0	79.0±4.0	74.3±1.9	47.5±3.7	87.1±1.5
07	30.7±4.2	63.1±2.4	61.2± 3.8	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	96.6±0.7	39.1±4.2	43.4±3.5	14.7±5.4	70.7±2.0
29	27.4±3.0	55.6±4.1	60.4± 4.8	96.9±1.5	100.0±0.0	100.0±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	92.2±1.7	85.7±1.7	60.6±3.3	81.6±1.7
11	16.3±3.5	30.7±3.5	30.0± 2.6	74.9±2.3	92.1±2.1	95.0±0.8	98.7±0.5	95.3±1.2	100.0±0.0	gallery	87.2±2.0	83.1±2.0	72.4±2.5	73.0±1.9

Table 3.15: Comparison of results with other 2D methods with face pose variations using the FERET database.

Method \Set	bb (60°) [%]	bc (40°) [%]	bd (25°) [%]	be (15°) [%]	ba (0°) [%]	bf (−15°) [%]	bg (−25°) [%]	bh (−40°) [%]	bi (−60°) [%]	Mean [%]
[Ashraf 2008]*	37.0	61.5	85.5	94.0	gallery	97.0	88.5	67.0	40.0	71.3
[Gao 2009]	44.0	81.5	93.0	97.0	gallery	98.5	91.5	78.5	52.5	79.6
[Gross 2004]*	20	38	46	-	gallery	-	44	45	8	33.5
[Li 2012]*	86.9	95.7	99	97.9	gallery	95.9	96	90.9	78.1	92.5
[Sarfraz 2010]*	78	90	95	-	gallery	-	87	84	80	85.7
[Sharma 2012]	70.0	82.0	94.0	95.0	gallery	96.0	94.0	85.0	79.0	86.8
[Vu 2009]*	76.5	90.0	99.5	100.0	gallery	100.0	99.0	95.5	87.0	93.4
Ours	89.0±0.8	99.5±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.0	98.1±0.4	84.6±0.8	96.4±0.2
* Approximated values extracted from a graph.										

Table 3.16: Comparison of results with 3D methods with face pose variations using the FERET database.

Method \Set	bb (60°) [%]	bc (40°) [%]	bd (25°) [%]	be (15°) [%]	ba (0°) [%]	bf (−15°) [%]	bg (−25°) [%]	bh (−40°) [%]	bi (−60°) [%]	Mean [%]
[Asthana 2011b]	-	90.5	98.0	98.5	gallery	97.5	97.0	91.9	-	95.6
[Blanz 2003]	94.8	95.4	96.9	99.5	gallery	97.4	96.4	95.4	90.7	95.8
[Ding 2012]	89.5	97.0	98.5	99.0	gallery	98.5	98.0	94.5	78.0	94.1
Ours	89.0±0.8	99.5±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.0	98.1±0.4	84.6±0.8	96.4±0.2
[Paysan 2009]	97.4	99.5	100.0	gallery	99.0	99.5	97.9	94.8	83.0	96.4
[Romdhani 2005]	92.7	99.5	99.5	gallery	96.9	99.5	95.8	89.6	77.1	93.8
Ours (15°)	96.8±0.4	99.5±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.2	99.1 ±0.4	96.5±0.9	66.2±0.9	94.8±0.3
Ours (−15°)	69.2±1.4	99.3±0.3	100.0±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.3	99.8±0.6	93.5±0.6	95.2±0.4
Ours (±15°)	96.8±0.4	99.5±0.0	100.0±0.0	gallery	100.0±0.0	gallery	100.0±0.3	99.8±0.6	93.5±0.6	98.5±0.3

Table 3.17: Comparison of results with other 2D Methods with face pose variations using the CMU-PIE database.

Method \Set	22	02	25	37	05	09	27	07	29	11	31	14	34	Mean	Our mean	Our mean (c37, c11)
	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
[Asthana 2011a]	-	-	-	89.4	95.5	98.5	gallery	98.5	100.0	88.1	-	-	-	95.0	100.0±0.0	98.8±0.3
[Castillo 2011]	62	98	88	99	97	100	gallery	99	99	97	91	93	60	90.2	87.1±1.5	94.1±0.9
[Chai 2007]	-	-	-	82.4	98.5	98.5	gallery	98.5	100.0	89.7	-	-	-	94.6	100.0±0.0	98.8±0.3
[González-Jiménez 2007]	-	-	-	62	93	-	gallery	-	93	75	-	-	-	80.8	100.0±0.0	100.0±0.0
[Gross 2004]	39*	59*	57*	89*	94*	95*	gallery	88*	56*	89*	56*	70*	48*	70.0	87.1±1.5	94.1±0.9
[Kanade 2003]	50	100	85	100	100	100	gallery	100	100	100	81	81	25	85.2	87.1±1.5	94.1±0.9
[Li 2012]	73.5	100*	88,3*	100*	100	100*	gallery	100*	100*	100	70.6*	85.4*	73.5*	90.9	87.1±1.5	94.1±0.9
[Prince 2008]	91	-	-	-	100	-	gallery	-	-	-	-	-	-	95.5	80.6±1.8	98.8±0,4
[Sarfraz 2010]	36*	78*	80*	86*	88*	92*	gallery	100*	91*	86*	81*	85*	45*	79.0	87.1±1.5	94.1±0.9
[Sharma 2012]	79	88	85	100	100	100	gallery	100	100	100	91	97	85	93.8	87.1±1.5	94.1±0.9
Ours	61.2±3.6	90.9±2.3	92.2± 2.2	100.0±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.0	100.0±0.0	79.0±4.0	74.3±1.9	47.5±3.7	87.1±1.5	-	-
Ours (c37 and c11)	97.5±0.7	99.7±0.6	99.9± 0.5	gallery	100.0±0.0	100.0±0.0	100.0±0.0	95.3±1.2	100.0±0.0	gallery	87.2±2.0	83.1±2.0	72.4±2.5	94.1±0.9	-	-

* Approximated values extracted from a graph.

Table 3.18: Comparison of results with 3D methods with face pose variations using the CMU-PIE database.

Method \Set	22	02	25	37	05	09	27	07	29	11	31	14	34	Mean	Our mean	Our mean (c37, c11)
	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]	[%]
[Asthana 2011b]	-	-	-	97	100	100	gallery	98.5	100	98.5	-	-	-	99.0	100.0±0.0	98.8±0.3
[Blanz 2003]*	79.5	-	-	-	97.5	-	gallery	-	-	-	-	-	-	88.5	80.6±1.8	98.8±0.4
[Ding 2012]	-	85.5	-	100	100	100	gallery	100	100	100	-	77.6	-	95.4	95.7±0.5	96.4±0.6
[Paysan 2009]*	75.7	-	-	-	96.1	-	98.9 (gallery)	-	-	-	-	-	-	90.2	80.6±1.2	99.2±0.2
[Romdhani 2005]*	76	-	-	-	98	-	100 (gallery)	-	-	-	-	-	-	91.3	80.6±1.2	99.2±0.2
Ours	61.2±3.6	90.9±2.3	92.2± 2.2	100.0±0.0	100.0±0.0	100.0±0.0	gallery	100.0±0.0	100.0±0.0	100.0±0.0	79.0±4.0	74.3±1.9	47.5±3.7	87.1±1.5	-	-
Ours (c37 and c11)	97.5±0.7	99.7±0.6	99.9± 0.5	gallery	100.0±0.0	100.0±0.0	100.0±0.0	95.3±1.2	100.0±0.0	gallery	87.2±2.0	83.1±2.0	72.4±2.5	94.1±0.9	-	-

* Values computed with ambient light (set "light").

Chapter 4
Discussion and Conclusions

Face recognition is an important topic in a wide variety of applications ranging from surveillance to selective marketing. Several recent studies have shown the predominance of local matching approaches in face recognition results. The work in this thesis was divided in two parts, the first one in frontal face recognition and the second one in face recognition under pose variation. For frontal face recognition, we proposed a local matching Gabor method using a preprocessing step with local normalization. The method uses entropy-like weighted Gabor features to weight the jet scores. Experiments were done on the FERET, AR, and FRGC 2.0 databases to test and compare the obtained results with those published previously.

Results of the methodological improvements proposed in this work show that it is possible to improve face recognition rate by weighting jets with entropy as in LMGEW. Weighting jets by entropy produces significant improvements in face recognition in the FERET database subsets Dup1 and Dup2 where images from the same person were taken further apart in time. The proposed LMGEW method reduces the total number of errors in the FERET database subsets Fb, Fc, Dup1 and Dup2 from 163 to 128, a 21% improvement. It was also shown that the combination of the proposed methods in cascade LMGEW-BTH reached an overall of 102 errors in the FERET database which represents a 37% error reduction compared to the best previously published results. On the AR database LMGEW-BTH achieved a reduction in error rate compared to LMG, of method 38% in the Normal subset, 63% in the Sunglasses subset and 85% in the Scarves subset. The threshold BTH is the only parameter that is chosen within a range in the proposed method. Choosing the T_h in the range 0.8 to 0.9 yields high accuracy on the results of the LMGEW-BTH method. Adding noise to the eye position ground truth, produces a small error in face recognition if the error in eye localization is of few pixels.

LMGEW//LN-BTH showed good identification performance improving results of LMG and LMGEW. The method was assessed on different databases

with varying illumination, small pose variations (up to 25°), changing expressions, and using images taken on different dates and locations. The proposed method shows robustness to illumination changes and expressions as shown in the results on the FERET, subsets Fb, Fc, the AR Normal subset and FRGC Exp.1, with results over 99%. Robustness to illumination and spatial changes is due to the Gabor jets' capacity to deal with small spatial perturbations. The LN method, capable of eliminating the noise introduced by illumination changes, plays an important role. Improvements can be observed across all subsets of the FERET database, although they are most evident in Dup1 and Dup2 which contain images with age difference compared to the gallery face images. The error was reduced from 15% to 7.3% in Dup1 and from 20% to 10.3% in Dup2. This is a reduction of 52.3% in the error rate for Dup1 and 48.5% for Dup2. In addition, for pose variation of 25° the improvement in face recognition increased from 80.3% to 96.3%, i.e., a 16% increase. This improvement could be due to the entropy-like weights on the Gabor jets. The entropy-like function measures the consistency of each jet in one image compared to the same jet in the entire database. The weights are larger on face zones that are similar to each other on the frontal view of the face, while weights are smaller on other parts of the face.

LMGEW//LN-BTH, LGBP and LGXP individually have the highest scores on face identification, and they were combined in this thesis. Results on fusion methods suggest that different methods for face identification improve different features, and combining these features improves the overall results. Different methodologies applied to face recognition extract and emphasize different aspects of the face image and therefore, fusing them has a synergistic effect as shown in Fig. 3.10.

Real face identification applications require automatic face detection and eye localization, which reduce performance in face recognition systems. The effect of adding noise to the eye localization on face identification was tested. Noise from 1 to 10 pixels was added in each eye and the results indicate that up to 4 pixels of error do not have a significant effect on Fb and Fc subsets. In Dup1 and Dup2, the effect is significant with over 2 pixels of error in eye localization. Real-time eye localization systems [Perez 2010a] tolerate only less than 3 pixels of error.

The LMG is a well-established method for frontal face identification with excellent performance. Nevertheless, as with most face identification methods, LMG performance declines significantly for large geometric deformations such as those produced by varying pose. In the second part of this study, a new method based on LMG is proposed using ASM to correct the position where

Gabor jets are computed with pose changes. Also, the method incorporates a statistical model of the Borda count scores computed by using the Gabor jets. The method includes illumination compensation by Local Normalization, and an entropy weighting of the Gabor jets to emphasize those features most relevant for identification.

The method was tested using the FERET and CMU-PIE databases, which are the most widely used international databases for face identification across pose. The FERET database has a subset with pose variations ($+60°$, $+40°$, $+25°$, $+15°$, $+0°$, $-15°$, $-25°$, $-40°$, $-60°$) with 200 faces for each pose angle. The proposed method was tested enrolling each pose angle as the gallery set. Consistently the highest recognition performance was reached for faces with pose closest to the enrolled face angle. Also, even if a gallery with a small pose angle was used, the recognition rates of faces with high rotation improved significantly. The mean recognition rate of faces with $\pm 60°$ of rotation increased from 86.8% to 95.2% if a gallery set with only $15°$ of rotation in the same direction was used. The mean improvement in the recognition rate of the proposed method compared to the classical LMG on the FERET database for pose variation went from 77.9% to 96.4%.

As expected, the improvement was larger for significant face pose variations. For example, the mean performance of the classical LMG in faces with $\pm 60°$ pose variation improved from 35.5% to 86.8% (144.5% improvement). The CMU-PIE database has 13 subsets with different pose variations, each with 68 faces. The recognition rate of faces with large rotation angles also improved significantly if a gallery with pose angle was used in this database. The recognition rate in the set with the largest rotation (set 22), increased from 61.2% to 97.5% when set 37 is used as gallery. Compared to the classical LMG, the mean performance in the recognition rate of the proposed method on the CMU-PIE database increased from 67.6% to 87.1% when a gallery of frontal faces was used. The recognition rate improvement for the sets with largest rotation angles (22 and 34) was 130.8%.

Results on the FERET database were compared to those of 12 different 2D and 3D state-of-the-art methods published previously. Obtained results were significantly better than all previous 2D methods which used the set with $0°$ of pose as gallery. The best 2D method [Vu 2009] achieved a mean performance of 93.4%, while the proposed method reached 96.4%. This difference was greater in faces with large pose variation angles. For example, the mean performance obtained in [Vu 2009] for faces with $\pm 60°$ of pose was 81.8% while the proposed method results reached 86.8%. Only one of the 3D methods obtained a better

mean result than the proposed method on the FERET database [Paysan 2009]. This method uses the set with 15° of pose as gallery, and the proposed method yields better results if the sets with ±15° of pose are used together as gallery. The 3D methods require a greater number of computations which can be interpreted as a disadvantage compared to the proposed method.

In the CMU-PIE database, the proposed method was compared with 10 different 2D and 5 different 3D previously published methods. Using the set with frontal faces as gallery the performance of the proposed method outperforms 6 of the 2D methods. Nonetheless, if two sets with rotated faces (37 and 11) are used as galleries, the mean result of the proposed method surpasses the performance of all previously published methods. If frontal faces are used in the gallery, the mean performance of the proposed method is better than the performance of 2 of the 3D methods on the CMU-PIE database. Using the two sets with rotated faces as galleries, the mean performance of the proposed method surpasses the results of 4 of the 5 methods.

The proposed method has a more consistent performance for the same rotation on the two tested databases (FERET and CMU-PIE) than the other methods because methods that show highest performance on CMU-PIE have lower performance on the FERET database. The cause of this may be that the proposed method is more robust to changes in the general conditions regarding the image capture procedure. This can be observed by comparing the sets with similar rotations in the FERET and CMU-PIE databases.

Chapter 5
Publications

The following publications resulted from the work of this thesis:

- Perez CA, **Cament LA**, Castillo LE, *Methodological improvement on local Gabor face recognition based on feature selection and enhanced Borda count*, Pattern Recognition, 44(4), pp.951-963, 2011.
- Perez CA, **Cament LA**, Castillo LE, *Local matching Gabor entropy weighted face recognition, IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011)*, pp.179-184, 2011.
- **Cament LA**, Castillo LE, Perez JP, Galdames FJ, Perez CA, *Fusion of Local Normalization and Gabor Entropy Weighted Features for Face Identification*, Pattern Recognition, 47(2), pp.568-577, 2014.

Bibliography

[Ahonen 2006] T. Ahonen, A. Hadid and M. Pietikainen. *Face Description with Local Binary Patterns: Application to Face Recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pages 2037–2041, 2006.

[Ashraf 2008] Ahmed Bilal Ashraf, Simon Lucey and Tsuhan Chen. *Learning Patch Correspondences for Improved Viewpoint Invariant Face Recognition*. In IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), pages 1–8, Anchorage, Alaska, USA, 23-28 June 2008.

[Asthana 2011a] Akshay Asthana, Michael Jones, Tim Marks, Kinh Tieu and Roland Goecke. *Pose Normalization via Learned 2D Warping for Fully Automatic Face Recognition*. In British Machine Vision Conference (BMVC), pages 127.1–127.11, University of Dundee, Dundee, Scotland, UK, 29 August - 2 September 2011.

[Asthana 2011b] Akshay Asthana, Tim K. Marks, Michael J. Jones, Kinh H. Tieu and M. V. Rohith. *Fully Automatic Pose-Invariant Face Recognition via 3D Pose Normalization*. In The 2011 IEEE International Conference on Computer Vision (ICCV 2011), pages 937–944, Barcelona, Spain, 6-13 November 2011.

[Bartlett 2002] M. S. Bartlett, J. R. Movellan and T. J. Sejnowski. *Face recognition by independent component analysis*. IEEE Transactions on Neural Networks, vol. 13, no. 6, pages 1450–1464, 2002.

[Belhumeur 1997] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman. *Eigenfaces vs. Fisherfaces: recognition using class specific linear projection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pages 711–720, 1997.

[Bianconi 2007] F. Bianconi and A. Fernández. *Evaluation of the Effects of Gabor Filter Parameters on Texture Classification*. Pattern Recognition, vol. 40, no. 12, pages 3325–3335, 2007.

Bibliography

[Blanz 2003] V. Blanz and T. Vetter. *Face recognition based on fitting a 3D morphable model*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 9, pages 1063–1074, September 2003.

[Bookstein 1989] Fred L. Bookstein. *Principal Warps: Thin-Plate Splines and the Decomposition of Deformations*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 6, pages 567–585, June 1989.

[Cament 2014] Leonardo A. Cament, Luis E. Castillo, Juan P. Perez, Francisco J. Galdames and Claudio A. Perez. *Fusion of local normalization and Gabor entropy weighted features for face identification*. Pattern Recognition, vol. 47, pages 568–577, 2014.

[Campadelli 2009] P. Campadelli, R. Lanzarotti and G. Lipori. *Precise eye and mouth localization*. International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), vol. 23, no. 3, pages 359–377, 2009.

[Castillo 2011] C. D. Castillo and D. W. Jacob. *Wide-baseline stereo for face recognition with large pose variation*. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 537–544, Colorado Springs, Colorado, USA, 21-15 June 2011.

[Chai 2007] Xiujuan Chai, Shiguang Shan, Xilin Chen and Wen Gao. *Locally Linear Regression for Pose-Invariant Face Recognition*. IEEE Transactions on Image Processing, vol. 16, no. 7, pages 1716–1725, 2007.

[Chellappa 2011] Rama Chellappa, Jie Ni and Vishal M. Patel. *Remote identification of faces: Problems, prospects, and progress*. Pattern Recognition Letters, vol. 33, no. 14, pages 1849–1859, 2011.

[Cootes 1992] T. F. Cootes and C. J.Taylor. *Active Shape Models - 'Smart Snakes'*. In British Machine Vision Conference (BMVC), pages 266–275, Leeds, UK, 22-24 September 1992.

[Cristinacce 2006] David Cristinacce and Tim Cootes. *Feature detection and tracking with constrained local models*. In British Machine Vision Conference (BMVC), pages 929–938, Universities of Edinburgh and Heriot-Watt, Edinburgh, Scotland, UK, 4-7 September 2006.

[Daugman 1980] J.G. Daugman. *Two-dimensional spectral analysis of cortical receptive field profiles*. Vision Research, vol. 20, no. 10, pages 847–856, 1980.

[Daugman 1985] J.G. Daugman. *Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters*. Journal of the Optical Society of America, vol. 2, no. 7, pages 1160–1169, 1985.

[Ding 2012] Liu Ding, Xiaoqing Ding and Chi Fang. *Continuous Pose Normalization for Pose-Robust Face Recognition*. IEEE Signal Processing Letters, vol. 19, no. 11, pages 721–724, November 2012.

[Du 2009] Shan Du and Rabab Ward. *Component-wise pose normalization for pose-invariant face recognition*. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '09), pages Pages 873–876, Taipei, Taiwan, 19-24 April 2009.

[Faundez-Zanuy 2007] M. Faundez-Zanuy, J. Roure, V. Espinosa-Duro and J. A. Ortega. *An efficient face verification method in a transformed domain*. Pattern Recognition Letters, vol. 28, no. 7, pages 854–858, 2007.

[Fukunaga 1975] Keinosuke Fukunaga and Larry D. Hostetler. *The estimation of the gradient of a density function, with applications in pattern recognition*. IEEE Transactions on Information Theory, vol. 21, no. 1, pages 32–40, January 1975.

[Gabor 1946] D. Gabor. *Theory of Communication*. Journal of Institute for Electrical Engineering, vol. 93, no. III, 26, pages 429–457, 1946.

[Gao 2009] Hua Gao, Hazim Kemal Ekenel and Rainer Stiefelhagen. *Pose Normalization for Local Appearance-Based Face Recognition*. In Third International Conference on Advances in Biometrics (ICB '09), volume 5558 of *Lecture Notes in Computer Science*, pages 32–41, Alghero, Italy, 2-5 June 2009.

[González-Jiménez 2007] Daniel González-Jiménez and José Luis Alba-Castro. *Toward Pose-Invariant 2-D Face Recognition Through Point Distribution Models and Facial Symmetry*. IEEE Transactions on Information Forensics and Security, vol. 2, no. 3, pages 413–429, September 2007.

[Goodall 1991] Colin R. Goodall. *Procrustes methods in the statistical analysis of shape*. Journal of the Royal Statistical Society. Series B (Methodological), vol. 53, no. 2, pages 285–339, 1991.

[Gross 2004] R. Gross, I. Matthews and Simon Baker. *Appearance-Based Face Recognition and Light-Fields*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 4, pages 449–465, April 2004.

[Hafed 2001] Z. M Hafed and M. D. Levine. *Face Recognition Using Discrete Cosine Transform*. International Journal of Computer Vision, vol. 43, no. 3, pages 167–188, 2001.

[He 2011] Ran He, Wei-Shi Zheng and Bao-Gang Hu. *Maximum Correntropy Criterion for Robust Face Recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 8, pages 1561–1576, 2011.

[Heo 2008] Jingu Heo and Marios Savvides. *Face recognition across pose using view based active appearance models (VBAAMs) on CMU Multi-PIE dataset*. In 6th International Conference on Computer Vision Systems (ICVS'08), volume 5008 of *Lecture Notes in Computer Science*, pages 527–535, Santorini, Greece, 12-15 May 2008.

[Ho 1994] Tin Kam Ho, J. J. Hull and S. N. Srihari. *Decision combination in multiple classifier systems*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 1, pages 66–75, 1994.

[Kanade 2003] T. Kanade and A. Yamada. *Multi-subregion based probabilistic approach towards pose-invariant face recognition*. In IEEE International Symposium on Computational Intelligence in Robotics Automation, volume 2, pages 954–959, Kobe, Japan, 16-20 July 2003.

[Kittler 1998] J. Kittler, M. Hatef, R. P. W. Duin and J. Matas. *On combining classifiers*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 3, pages 226–239, 1998.

[Kyrki 2004] V. Kyrki, J.-K. Kamarainen and H. Kälviäinen. *Simple Gabor feature space for invariant object recognition*. Pattern Recognition Letters, vol. 25, pages 311–318, 2004.

[Lades 1993] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurtz and W. Konen. *Distortion invariant object recognition in the dynamic link architecture*. IEEE Transactions on Computers, vol. 42, no. 3, pages 300–311, 1993.

[Lee 1996] Tai Sing Lee. *Image Representation Using 2D Gabor Wavelets*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 10, pages 959–971, October 1996.

[Lei 2007] Z. Lei, S. Li, R. Chu and X. Zhu. *Face recognition with Local Gabor textons*. In International Conference on Biometrics IAPR 2007, volume 4642, pages 49–57, 2007.

[Li 2012] Annan Li, Shiguang Shan and Wen Gao. *Coupled Bias-Variance Tradeoff for Cross-Pose Face Recognition*. IEEE Transactions on Image Processing, vol. 21, no. 1, pages 305–315, January 2012.

[Liu 2002] Chengjun Liu and H. Wechsler. *Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition*. IEEE Transactions on Image Processing, vol. 11, no. 4, pages 467–476, 2002.

[Martinez 1998] A.M. Martinez and R. Benavente. *The AR Face Database*. Technical Report 24, CVC Technical Report, 1998.

[Marčelja 1980] S. Marčelja. *Mathematical Description of the Responses of Simple Cortical Cells*. Journal of the Optical Society of America, vol. 70, no. 11, pages 1297–1300, 1980.

[Matthews 2004] Iain Matthews and Simon Baker. *Active Appearance Models Revisited*. International Journal of Computer Vision, vol. 60, no. 2, pages 135–164, 2004.

[Nguyen 2009] H. V. Nguyen, L. Bai and L. Shen. *Local Gabor Binary Pattern Whitened PCA: A Novel Approach for Face Recognition from Single Image Per Person*. In International Conference on Advances in Biometrics IAPR 2009, volume 5558, pages 269–278, 2009.

[Paysan 2009] Pascal Paysan, Reinhard Knothe, Brian Amberg, Sami Romdhani and Thomas Vetter. *A 3D Face Model for Pose and Illumination Invariant Face Recognition*. In 6th IEEE International Conference on Advanced

Video and Signal Based Surveillance (AVSS 2009), pages 296–301, Genova, Italy, 2-4 September 2009.

[Perez 2003] C. A. Perez, C. A. Salinas, P. A. Estevez and P. M. Valenzuela. *Genetic design of biologically inspired receptive fields for neural pattern recognition*. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 33, no. 2, pages 258–270, 2003.

[Perez 2005] C. A. Perez, G. D. Gonzalez, L. E. Medina and F. J. Galdames. *Linear versus nonlinear neural modeling for 2-D pattern recognition*. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 35, no. 6, pages 955–964, 2005.

[Perez 2007] C. A. Perez, V. A. Lazcano and P. A. Estevez. *Real-Time Iris Detection on Coronal-Axis-Rotated Faces*. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 37, no. 5, pages 971–978, 2007.

[Perez 2010a] C. Perez, C. Aravena, J. Vallejos, P. Estévez and C. Held. *Face and Iris Localization using Templates Designed by Particle Swarm Optimization*. Pattern Recognition Letters, vol. 31, no. 9, pages 857–868, 2010.

[Perez 2010b] Claudio A. Perez, Luis E. Castillo and Leonardo A. Cament. *Illumination compensation method for local matching Gabor face classifier*. In International Symposium on Optomechatronic Technologies (ISOT), pages 1–5, Toronto, ON, Canada, 25-27 October 2010.

[Perez 2011a] Claudio A. Perez, Leonardo A. Cament and Luis E. Castillo. *Local Matching Gabor Entropy Weighted Face Recognition*. In IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011), pages 179–184, March 2011.

[Perez 2011b] Claudio A. Perez, Leonardo A. Cament and Luis E. Castillo. *Methodological improvement on local Gabor face recognition based on feature selection and enhanced Borda count*. Pattern Recognition, vol. 44, pages 951–963, 2011.

[Phillips 1998] P. J. Phillips, H. Wechsler, J. Huang and P. Rauss. *The FERET database and evaluation procedure for face recognition algorithms*. Image and Vision Computing, vol. 16, no. 5, pages 295–306, 1998.

[Phillips 2000] P. J. Phillips, Hyeonjoon Moon, S. A. Rizvi and P. J. Rauss. *The FERET evaluation methodology for face-recognition algorithms*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 10, pages 1090–1104, 2000.

[Phillips 2005] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, Jin Chang, K. Hoffman, J. Marques, Jaesik Min and W. Worek. *Overview of the face recognition grand challenge*. In IEEE Conference on Computer Vision and Pattern Recognition CVPR 2005, volume 1, pages 947–954, 2005.

[Phillips 2006] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer and W. Worek. *Preliminary Face Recognition Grand Challenge Results*. In International Conference on Automatic Face and Gesture Recognition FG 2006, pages 15–24, 2006.

[Podilchuk 1996] C. Podilchuk and Xiaoyu Zhang. *Face recognition using DCT-based feature vectors*. In IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP 1996, volume 4, pages 2144–2147, 1996.

[Prince 2008] Simon J. D. Prince, James H. Elder, Jonathan Warrell and Fatima M. Felisberti. *Tied Factor Analysis for Face Recognition across Large Pose Differences*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 6, pages 970–984, June 2008.

[Romdhani 2005] Sami Romdhani. *Face image analysis using a multiple features fitting strategy*. PhD thesis, University of Basel, Faculty of Science, Basel, Switzerland, 2005.

[Saragih 2011] Jason M. Saragih, Simon Lucey and Jeffrey F. Cohn. *Deformable Model Fitting by Regularized Landmark Mean-Shift*. International Journal of Computer Vision, vol. 91, no. 2, pages 200–215, January 2011.

[Sarfraz 2010] M. S. Sarfraz and Olaf Hellwich. *Probabilistic learning for fully automatic face recognition across pose*. Image and Vision Computing, vol. 28, no. 5, pages 744–753, May 2010.

[Serrano 2009] Á. Serrano, I. M. de Diego, C. Conde and E. Cabello. *Recent Advances in Face Biometrics with Gabor Wavelets: A Review*. Pattern Recognition Letters, vol. 31, no. 5, pages 372–381, 2009.

[Sharkas 2008] M. Sharkas and M. Abou Elenien. *Eigenfaces vs. Fisherfaces vs. ICA for Face Recognition; A Comparative Study.* In IEEE International Conference on Signal Processing ICSP 2008, pages 914–919, 2008.

[Sharma 2012] Abhishek Sharma, Murad Al Haj, Jonghyun Choi, Larry S. Davis and David W. Jacobs. *Robust pose invariant face recognition using coupled latent space discriminant analysis.* Computer Vision and Image Understanding, vol. 116, no. 11, pages 1095–1110, November 2012.

[Sim 2003] T. Sim, S. Baker and M. Bsat. *The CMU pose, illumination, and expression database.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 12, pages 1615–1618, 2003.

[Su 2009] Yu Su, Shiguang Shan, Xilin Chen and Wen Gao. *Hierarchical Ensemble of Global and Local Classifiers for Face Recognition.* IEEE Transactions on Image Processing, vol. 18, no. 8, pages 1885–1896, 2009.

[Turk 1991] M. A. Turk and A. P. Pentland. *Face recognition using eigenfaces.* In IEEE Conference on Computer Vision and Pattern Recognition CVPR 1991, pages 586–591, 1991.

[Vu 2009] Ngoc-Son Vu and Alice Caplier. *Efficient Statistical Face Recognition across pose using Local Binary Patterns and Gabor wavelets.* In IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS '09), pages 1–5, Washington, DC, USA, 28-30 September 2009.

[Wagner 2009] A. Wagner, J. Wright, A. Ganesh, Zihan Zhou and Yi Ma. *Towards a practical face recognition system: Robust registration and illumination by sparse representation.* In IEEE Conference on Computer Vision and Pattern Recognition CVPR 2009, pages 597–604, 2009.

[Wiskott 1997] L. Wiskott, J.-M. Fellous, N. Kuiger and C. von der Malsburg. *Face recognition by elastic bunch graph matching.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pages 775–779, 1997.

[Wright 2009] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and Yi Ma. *Robust Face Recognition via Sparse Representation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 2, pages 210–227, 2009.

[Xie 2006] X. Xie and K.-M. Lam. *An efficient illumination normalization method for face recognition.* Pattern Recognition Letters, vol. 27, pages 608–617, 2006.

[Xie 2009] S. Xie, S. Shan, X. Chen, X. Meng and W. Gao. *Learned local Gabor patterns for face representation and recognition.* Signal Processing, vol. 89, pages 2333–2344, 2009.

[Xie 2010] Shufu Xie, Shiguang Shan, Xilin Chen and Jie Chen. *Fusing Local Patterns of Gabor Magnitude and Phase for Face Recognition.* IEEE Transactions on Image Processing, vol. 19, no. 5, pages 1349–1361, 2010.

[Zhang 2005] Wenchao Zhang, Shiguang Shan, Wen Gao, Xilin Chen and Hongming Zhang. *Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition.* In IEEE International Conference on Computer Vision ICCV 2005, volume 1, pages 786–791, 2005.

[Zhang 2007] Baochang Zhang, Shiguang Shan, Xilin Chen and Wen Gao. *Histogram of Gabor Phase Patterns (HGPP): A Novel Object Representation Approach for Face Recognition.* IEEE Transactions on Image Processing, vol. 16, no. 1, pages 57–68, 2007.

[Zhao 2003] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld. *Face Recognition: A Literature Survey.* ACM Computing Surveys (CSUR), vol. 35, no. 4, pages 399–458, 2003.

[Zou 2007] Jie Zou, Qiang Ji and G. Nagy. *A Comparative Study of Local Matching Approach for Face Recognition.* IEEE Transactions on Image Processing, vol. 16, no. 10, pages 2617–2628, 2007.