Ensemble of texture descriptors for face recognition obtained by varying feature transforms and preprocessing approaches

Article *in* Applied Soft Computing · August 2017

DOI: 10.1016/j.asoc.2017.07.057

Loris Nanni
University of Padova

Alessandra Lumini
University of Bologna

Sheryl Brahnam
Missouri State University

Ensemble of texture descriptors for face recognition obtained by varying feature transforms and preprocessing approaches

Loris Nanni[1] Alessandra Lumini[2] Sheryl Brahnam[3]

[1]DEI, University of Padua, viale Gradenigo 6, Padua, Italy. loris.nanni@unipd.it;

[2]DISI, Università di Bologna, Via Sacchi 3, 47521 Cesena, Italy;

[3]Computer Information Systems, Missouri State University, 901 S. National, Springfield, MO 65804, USA.

Abstract

This paper presents a novel ensemble of descriptors for face recognition derived from the base Patterns of the Oriented Edge Magnitudes (POEM) descriptor. Starting from different texture descriptors recently proposed in the literature, namely, the base patterns of POEM and the Monogenic Binary Coding (MBC), we develop different ensembles by varying the preprocessing techniques, the subspace projections, and some parameters of the system. Our approach is tested on the FERET datasets and the Labeled Faces in the Wild (LFW) dataset. Our system performs well on both datasets, obtaining, to the best of our knowledge, one of the highest performance rates published in the literature on the FERET datasets with an average accuracy of 97.3%. We want to stress that our ensemble obtains outstanding results in both datasets without any supervised approach or transform. The main findings of our proposed system include the following: 1) significant improvement in performance can be obtained by simply varying the parameters of stand-alone descriptors; and 2) performance can be improved by combining different enhancement and feature transform techniques.

Keywords: Face recognition; ensemble of descriptors; patterns of oriented edge magnitudes; monogenic binary coding.

1. Introduction

Since 1960 face recognition has been extensively studied, and research in this area continues to gain momentum due to increasing governmental and business security needs. Face recognition applications can be divided into three main categories: 1) face verification, where the objective is to authenticate identity by comparing a face with a corresponding template; 2) face identification, where the goal is to match a face with an image already in the system; and 3) face tagging, which is a particular case of face identification that is growing in popularity. Biometric security, access management, criminal identification, and image sorting and retrieval are only a few of the numerous and profitable applications of face recognition.

Many face recognition techniques have been proposed over the last fifty years [1], including Principal Component Analysis, Discriminant Analysis, Local Binary Patterns (LBP), neural networks, Elastic Template Matching, Algebraic moments, and Gabor Filtering [2, 3]. One way to classify face recognition techniques is according to how they represent, or define, a face [1]. *Appearance based approaches* use global features, such as Eigenfaces [4] and other linear transformations. *Model based approaches* utilize information in both the shape (2D or 3D) and the texture of a face. *Geometry*, *or template-based*, *approaches* compare an input image with a set of templates that are constructed either by using statistical tools or by analyzing local facial features and their geometric relationships. Finally, there are techniques that use *Neural Networks*.

The goal of these techniques is to solve the following problem: given two face images, both of which can vary in acquisition time, pose, expression, lighting conditions, occlusions, and image quality, determine whether both images are of the same person or of two different people. Most state-of-the-art face recognition techniques perform well on full frontal views and when images are taken under optimal illumination conditions, but when pose, age, and expression change, or when environmental conditions are less than ideal, performance deteriorates. The difficulty lies in teasing out those specific features that indicate identity while ignoring those critical features that represent variations between two images that go beyond those of identity. If inadequate features are extracted for the task of identity recognition, then even the best classifier will be unable to achieve good recognition rates. Another crucial activity is accurate face detection, the goal of which is to detect faces within an image and return their precise locations and extent of coverage within the image frame via bounding boxes. This work does not explore the face detection task [5]; the proposed approach is robust against small amounts of error in the localization task.

To deal with image quality and localization problems, multi-biometrics has been proposed, which is the automated recognition of individuals involving the use of biometric fusion (i.e. multimodal, multi-instance, multisensorial, and multi-algorithmic).

A major focus of research in face recognition is thus centered on the design of face descriptors that are both discriminative and insensitive to 1) pose variations, 2) changes in facial expression, and 3) lighting conditions. Some recent developments in this area include the work of Pinto et al. [6], who designed V1-like features composed of a population of Gabor filters that are tolerant of view, lighting, and other image variations. Cao et al. [7] recently proposed a method that encodes the local micro-structures of a face into a set of more uniformly distributed discrete codes−a feature set which provides a good tradeoff between discriminative power and invariance. Patterns of Oriented Edge Magnitudes (POEM), proposed in [8-10], is an oriented spatial multi-resolution descriptor that captures rich information about the self-similarity structure of an image. Other encouraging results in unconstrained environments are obtained in [11] using a sparse representation of faces for person-specific verification. In [12], an efficient descriptor based on local feature extraction is proposed, called monogenic binary coding (MBC); it claims to significantly lower time and space complexity compared with other Gabor transformation based local feature methods. Monogenic signal representation decomposes an original signal into three components (amplitude, orientation, and phase). It encodes the local variation of such features and calculates the histogram of the extracted local features.

In the last couple of years, descriptors that are invariant to pose and illumination have been obtained by combining texture-based descriptors with other ideas. For example, in [13], an accurate 3D shape model is used to map images to a frontal view. In addition, discriminative models are obtained by comparing billions of faces. These models are able to handle aging, facial expressions, low light, and over-exposure. In [14] binary classifiers are used for face verification. One approach, based on 65 describable visual traits (manually labeled on the training set), recognizes the presence or absence of some describable aspects of visual appearance, e.g., gender and race. A second approach, based on "simile classifiers," removes the need of manual labeling by training binary classifiers to recognize the similarity of faces (or regions) to specific reference people. Both approaches use very simple, low-level features, such as image intensities in RGB and HSV color spaces, edge magnitudes, and gradient directions. However, these approaches require rectification to a common coordinate system using an affine warping to obtain pose invariance. An identity-preserving alignment is proposed in [15] using binary classifiers. In this approach, the training set of faces is used to perform an "identity-preserving" alignment (face warping was used to reduce differences in poses and expressions while preserving differences that indicate identity) as well as to train classifiers to recognize people.

Image blur is addressed in [16], where a novel blur-robust face image descriptor, based on multiscale Local Phase Quantization (MLPQ), is proposed. The MLPQ descriptor is computed regionally by adopting a component-based framework to maximize the insensitivity to misalignment, which is particularly frequent in the case of blurring. The regional features are then combined using kernel fusion. To increase insensitivity to illumination, the MLPQ representation is combined with the Multiscale Local Binary Pattern (MLBP) descriptor according to a supervised fusion that is based on Kernel Discriminant Analysis (KDA). The very impressive results reported in [17] on the LFW dataset are not comparable with the approach proposed in this work, however, since they are obtained using a supervised transform and a different testing protocol.

In this paper we start from some preliminary results that appeared in [18], showing that the performance of the POEM descriptor [8] (one of the most efficient and best performing descriptors recently proposed in the literature) can be boosted by designing an ensemble of classifiers based on the variation of preprocessing techniques and feature extraction parameters. The key additions to this journal version of the proposed approach are the following:

- We show that it is possible to further improve the performance of the ensemble using patterns perturbed by different feature transforms;
- We test a different texture descriptor, the Monogenic Binary Coding [12] (MBC). The scores based on MBC are included in our ensemble, and the recognition performance is better than that obtained by the stand-alone version of MBC;
- The developed fusion between the ensemble of POEM and MBC obtains, to the best of our knowledge, the highest mean accuracy in the FERET dataset.

The rest of this paper is organized as follows. Section 2 describes the original POEM and MBC descriptors and the method for designing ensembles that is proposed in this work. In section 3, we provide experimental results using two benchmark databases: the FERET datasets and the Labeled Faces in the Wild (LFW) dataset. Finally, in section 4, we summarize results and offer some suggestions for future explorations.

2. Ensemble of descriptors for face recognition

2.1 The POEM descriptor

The original POEM descriptor is based on characterizing edge directions or the local face appearance and shape using the distribution of local intensity gradients. It measures the edge/local shape information and the relation between the information in neighboring cells.

Extracting the POEM descriptors is a three step process:

Step 1: *Perform gradient computation and orientation quantization*: the gradient image is computed, and then the gradient orientation of each pixel is discretized over 0-π. To avoid problems due to image degradation, a soft assignment is also performed [9]. Thus, at each pixel *p*, the gradient is a *d*-dimensional vector (where *d* is number of discretized orientations) that codifies its soft assigned magnitude and its discretized direction: $(p) = [m_1(p), \dots, m_d(p)]$, where the *i*-th element takes the original magnitude if the discretized orientation of the current pixel belongs to this bin, or it is set to 0, otherwise.

Step 2. *Calculate the magnitude accumulation*: this step aims at incorporating gradient information from neighboring pixels by computing a local histogram of orientations within a local image patch, or cell *C*. In this way, each pixel carries information about the distribution of the edge direction of a local (overlapping) cell: the feature at each pixel *p* which is the center of the cell *C* is a vector: $\widehat{m}(p) = [\widehat{m}_1(p), \dots, \widehat{m}_d(p)]$, where $\widehat{m}_i(p) = \sum_{p_j \in C} m(p_j)$.

Step 3. *Calculate self-similarity*: the accumulated magnitudes are encoded across different directions using an LBP-based operator which is suited to encode self-similarity within a larger block *B*. Based on our previous results [18], Dense LBP (DLBP) [19] is used instead of standard LBP. Formally, at pixel *p*, for the discretized direction *i*, POEM feature is calculated as: $POEM_i(p) = \sum_{j=1}^{n} f(\widehat{m}_i(p) - \widehat{m}_i(p_j))$ where $p_j \in B$ are *n* surrounding pixels and *f*(*x*) is defined as: *f*(*x*)=1 if *x*>τ 0 otherwise (τ is a threshold slightly larger than 0 used to provide some stability in uniform regions). Then, the POEM feature set at the pixel *p* is the concatenation of these *d* unidirectional POEM values.

The result of the previous three steps is a set of "unidirectional" POEM maps. To incorporate spatial information, the POEM maps are divided into 8×8 non-overlapping blocks, and histograms are extracted from each block. The final POEM-HS descriptor is the concatenation of all unidirectional descriptors at different orientations (Figure 1).

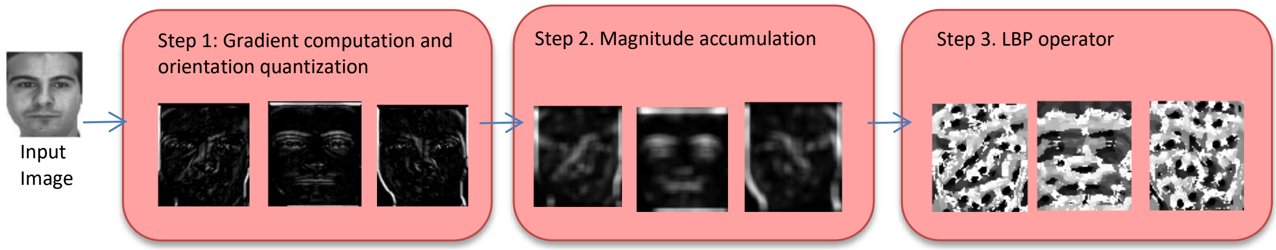

Figure 1. Schema of the POEM feature extraction process.

The POEM descriptor depends on a high number of parameters that should be tuned to the appropriate application: in our experiments the number of orientations discretized, the size of the cell, the size of the block, and the number of neighbors considered in LBP have been set to 3, 7, 5, and 8, respectively, according to the original paper [10].

2.2 The MBC descriptor

The Monogenic Binary Coding (MBC) is an efficient texture descriptor recently proposed in [12]. The monogenic signal is a rotation-invariant representation which extracts phase, amplitude, and orientation of the signal without using steerable filters [20], i.e. orientation-selective convolution kernels to extract multiple-orientation features. Thus, it requires much lower time complexity (e.g., only three convolutions on each scale) and space complexity (e.g., only three feature maps on each scale) than does the Gabor transformation.

Monogenic signal representation is defined as the combination of an image and its Riesz transform (a scalar-to-vector signal transformation which is a natural multidimensional extension of the Hilbert transform) [21], and it decomposes an original signal into three components: amplitude, orientation, and phase. Multiresolution Monogenic Signal Representation is obtained by performing band-pass filtering to an image before applying the Riesz transform by means of log-Gabor filters (three different resolutions are suggested in [12], corresponding to different scaling factors of the bandwidth).

Monogenic Binary Coding is based on encoding monogenic signal features by means of two complementary steps: the first step encodes the variation between the central pixel and its surrounding pixels in a local patch (monogenic local variation); and the second encodes the value of the central pixel itself (monogenic local intensity coding). Finally, a monogenic binary code (MBC) map is calculated as the concatenation of histograms from each of the amplitude, phase, and orientation components of the monogenic signal representation. As a final step, in [12], Linear Discriminant Analysis (LDA) [22] is used to reduce the histogram feature dimension while enhancing its discrimination capability. In this paper, we use only unsupervised feature transforms (not LDA since it is a supervised dimension reduction technique).

There are several parameters involved in different stages of MBC, e.g., multiscale log-Gabor filtering, subregion histogram computing, and feature combination by LDA. In our experiments all parameters have been set according to the original paper [12].

2.3 Designing an ensemble

In the last few years, there have been a number of proposals on how best to combine different classifiers for improving the performance of stand-alone methods. Combined approaches, or classifier ensembles, are usually based on combining the outputs of elementary classifiers for a given classification problem. The advantage of classifier ensembles over stand-alone methods are based on the following: ensemble systems are able to exploit the unique strengths of each of the individual classifiers, and they are less sensitive to parameter selection and optimization. Of course, in order for an ensemble to improve the average performance of stand-alone classifiers, two fundamental conditions are needed: the ensemble should not be composed of random classifiers, and they should be uncorrelated. Many ensemble techniques have been proposed in the field of biometric recognition [23], which can be categorized into four groups according to the level at which the information provided by the biometric traits is merged: sensor-level fusion, feature-level fusion, score-level fusion, and decision-level fusion. The present work proposes an ensemble based on score-level fusion and requires a single image acquisition. As stated above, one important requirement for achieving success in ensemble design is to maximize the independence of the classifiers involved in the fusion. A general approach to achieve this aim is to perturb the information given to the base classifiers. Existing approaches for the design of classifier ensembles are usually categorized into four categories [23]: *(i)* perturbation of the patterns, where each classifier is trained using different training sets or by weighting differently the training patterns (e.g. Bagging or Boosting); *(ii)* perturbation of features, where each classifier is trained using a different feature set (e.g. Random Forest or Random Subspace); *(iii)* perturbation of the classifiers by changing parameters or the base classifiers; and *(iv)* perturbation of the sensors when patterns are acquired using different sensors. Moreover, it should be noted that perturbation can take place at any step in the classification process: image preprocessing, feature transformation and selection, as well as matching.

As illustrated in Figure 2, our proposed system builds ensembles by using different preprocessing/filtering approaches and by perturbing POEM parameters:

- We propose a pattern perturbation approach based on the pre-processing of the input image using different enhancing methods. For the resulting pattern (an enhanced image), a descriptor is extracted and used to train a classifier; the three preprocessing techniques used in this work are detailed in section 2.4;
- As stated above we propose to extract two different descriptors (POEM and MBC), which are successively fused at the classifier level, thus performing feature perturbation;
- We propose a feature perturbation approach based on the use of different feature transforms to map descriptors in a reduced space: several feature transform are applied to the extracted feature vector, and each transformed vector is used to train a classifier. The four feature transform techniques used in this work are detailed in section 2.5.

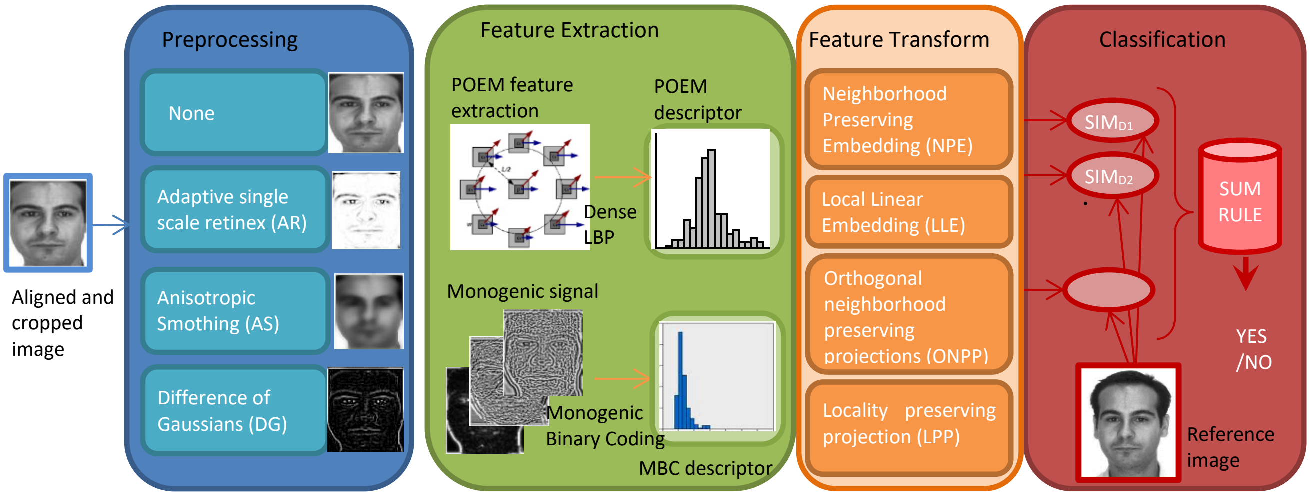

Figure 2. Schema of the proposed face recognition ensemble.

As depicted in Figure 2, the fusion is performed at score level by combining via sum rule the similarity scores obtained from different distance measures applied to the descriptors. The three distance measures used in this work are presented in section 2.6, while their application is detailed in section 3, together with the description of the experiments.

2.4 Preprocessing techniques

The problem of illumination variation can be solved in part by using recently developed image enhancement techniques before the feature extraction step. In this work we examined the following approaches:

- *Adaptive single scale retinex (AR)* [24]: this is a variant of the retinex technique, developed to improve scene detail and color reproduction in the darker areas of an image. AR uses the spatial information between each pixel and its surrounding pixels to normalize illumination. AR produced the best results in our experiments;
- *Anisotropic smoothing (AS)* [25]: this is a simple automatic image-processing normalization algorithm. AS estimates the illumination field and then compensates for it by enhancing the local contrast of the image in a way that is similar to human visual perception. AS has

demonstrated large performance gains for standard face recognition algorithms across multiple face databases [25];

- *Difference of Gaussians (DG):* this is a normalization technique that relies on the Difference of Gaussians filters to produce a normalized image. Before the feature extraction step, a band-pass filter is applied to an input image (note: the log transform is used before filtering, as in [26]).

2.5 Feature Transform

The feature reduction techniques used in our proposed approach are unsupervised; that is, while they do use the input data in the dimensionality reduction process (the FERET training set here), they make no use of the class labels. The most popular method for Unsupervised Dimensionality Reduction (UDR) is Principal Component Analysis (PCA) [3], which maps feature vectors into a smaller number of uncorrelated directions calculated to preserve the global Euclidean structure. The central idea behind PCA is to find an orthonormal set of axes pointing in the direction of maximum covariance in the data.

In this work we also test several of the following UDR techniques:

- *Locally linear embedding* (LLE) [27], which is a popular UDR method that aims at finding a low-dimensional embedding that optimally preserves the local neighborhood structure on the original data manifold. LLE represents the local geometry of a point by using its *k*-nearest neighbors (K-NNs). First, it calculates a weight matrix *W* based on local linear reconstructions of each point from its K-NNs with the sum-to-one constraint on the weight matrix. Subsequently, all embeddings are recovered to comply with the reconstruction weights by solving an optimization problem to minimize the reconstruction error;

- *Neighborhood preserving embedding* (NPE) [28], which is the linearization of the LLE. NPE employs a nearest neighbor search in the low-dimensional space to yield results as similar as possible to those in the high-dimensional space. More specifically, given a set of data points, first NPE constructs a weight matrix which represents the neighborhood relationship among the data points, and then an optimal embedding is determined such that this neighborhood structure can be preserved in the low-dimensional space.

The other two techniques we test are *locality preserving projections* (LPP) [29] and *orthogonal neighborhood preserving projections* (ONPP) [30]. Both these methods add linear constraints to some existing approach, thus forcing the embeddings to be linear combinations of the original features. For instance, LPP is such a linear variation of Laplacian eigenmaps [31], while ONPP is of LLE.

2.6 Distance functions

Several distance measures have been proposed in the literature to evaluate the similarity among descriptors. Based on the results published using POEM and MBC descriptors in [8, 12], we use the following distance functions in this work:

- *City block distance* (CBD): the City block distance between two points, $a,b \in \Re^n$, is calculated as: $CBD(a,b) = \sum_{k=1}^{n} |a_k - b_k|$. CBD is used to compare high dimensional POEM descriptors;
- *Chi-square distance* (CS): the Chi-Square Distance is usually employed as similarity measure for histograms. The Chi-Square Distance between two points, $a,b \in \Re^n$, is calculated as:

$$CS(a,b) = \frac{1}{2}\sum_{k=1}^{n} \frac{(a_k - b_k)^2}{a_k + b_k};$$

- *Angle distance* (AD): the angle distance between two points, $a,b \in \Re^n$, is calculated as the arccosine of the cosine distance:

$$AD(a,b) = \operatorname{acos}\left(\frac{\sum_{k=1}^{n} a_k b_k}{\|a\|\|b\|}\right).$$

3. Experimental results

3.1 Datasets

All experiments are performed using the FERET [32] and LFW [33] benchmark databases. The FERET database is divided into five datasets: Fa (1196 images), Fb (1195 images), Fc (194 images), Dup1 (722 images), and Dup2 (234 images). Fa is the gallery set, while the other sets are used for testing. Fb contains pictures taken on the same day and with the same camera and illumination conditions as used in Fa. Fc contains pictures taken on the same day as Fa but with different cameras and different illumination conditions. Dup1 and Dup2 contain pictures taken on different days: within the same year as Fa for Dup1 and more than one year for Dup2. The standard FERET evaluation protocol involves comparing images in the testing sets to each image in the gallery set in order to find the subject's image in the gallery. This is an identification task. In our experiments, all FERET grayscale images are aligned using the true eye positions and cropped to 110 × 110 pixels.

The LFW [33] database contains 13233 images of celebrities collected from the internet. LFW is challenging because the images vary greatly in terms of lighting, pose, age, and image quality. LFW provides two views of the database. View 1 is used for model selection purposes only and contains a training set of 2200 face pairs and a testing set of 1000 face pairs. View 2 is for performance reporting and contains ten non-overlapping sets of 600 face pairs. LWF protocol is a verification task and in each set the number (300) of true positives (i.e. pair of faces from the same individual) is equal to the number (300) of true negatives (pair of faces from different individuals). These sets can be used for 10-fold cross-validation algorithms and for parameters developed on View 1. In our experiments, LFW gray images are aligned automatically using the procedure described in [10] and cropped to 110×110 pixels.

3.2 Results

We test our systems on both databases using their official testing protocols. The performance indicator is the recognition rate (accuracy), which is the proportion of true classification results (both true positives and true negatives) in the population. For further comparisons, the MATLAB source of our best approach is freely available[1].

In the following, we report several sets of tests, each one designed to motivate our complete system as proposed in the previous section. In the first set of experiments, we evaluate the effect of the variations in the preprocessing step and the parameters of POEM, with and without the PCA subspace projection.

For the LFW dataset[2] we also use an artificial pose (as suggested in [34]), obtained by vertically flipping the image. Then, when we match two faces, we compare all pairs of images of the face (considering the artificial image as well), using the minimum among these distances as the matching score. In Table 1, a detailed description of the methods compared in our experiments is reported according to the following parameters:

- Preprocessing procedure (PreP): no preprocessing (NO), Adaptive single scale retinex (AR), Anisotropic smoothing (AS), and Difference of Gaussians (DG);
- Descriptors (Desc): POEM and MBC are the tested descriptors;
- Feature transform and distance measure (FT): city block distance (CBD) is used to compare high dimensional POEM descriptors (in the original code, the chi-square distance (CS) is used), while angle distance (AD) is used when the descriptor is reduced to a lower space by PCA or other unsupervised feature transforms. In this work, differently from [34], we use the

---

[1] http://www.dei.unipd.it/en/computer-science/data-mining-and-machine-learning + Pattern Recognition and Ensemble Classifiers

[2] Artificial pose generation has not been used in the FERET dataset due to lack of performance improvement.

same dimensionality parameter (D=500) and the same projection space (derived from the FERET training set) for all the experiments (that is, for both FERET and LFW). Moreover, when PCA is applied the square root normalization is performed before the matching, as in [34]; the other unsupervised feature transforms tested are NPE, LLE, ONPP, and LPP (see Section 2.5), all coupled to AD as the distance measure;

- Standalone/ensemble (SE): ensemble approaches are obtained by perturbing the preprocessing techniques (see Section 2.4) or by perturbing the feature transforms (see Section 2.5). The scores are fused by sum rule. We define the following in Table 1: SA, stand-alone method; EE, the perturbation of enhancements, EF the perturbation of feature transform, and EEF the perturbation of both enhancements and feature transform. The approach is named E_P if the perturbation is applied to POEM descriptor, or E_M if the perturbation is applied to MBC descriptor.

The first experiment is aimed at comparing different preprocessing techniques and ensembles based on the perturbation of preprocessing using the POEM descriptor. In Table 2 the accuracy obtained by our approaches and some reference works are reported. Please note that in order to test the robustness of our approach the same PCA projection matrix calculated in the FERET datasets is used in LFW. Due to computational issues, only a subset of the proposed approaches has been tested on LFW.

Name	PreP	Desc	FT	SE	Description
$\boldsymbol{POEM^{PCA}}$	-	POEM	PCA+AD	SA	Use of DenseLBP
$\boldsymbol{POEM_{LBP}^{PCA}}$	AR	POEM	PCA+AD	SA	Use of LBP (i.e. this method is equal to the original one)
$\boldsymbol{POEM_{AR}^{PCA}}$	AR	POEM	PCA+AD	SA	Use of DenseLBP and preprocessing AR
$\boldsymbol{POEM_{AS}^{PCA}}$	AS	POEM	PCA+AD	SA	Use of DenseLBP and preprocessing AS
$\boldsymbol{POEM_{DG}^{PCA}}$	DG	POEM	PCA+AD	SA	Use of DenseLBP and preprocessing DG
$\boldsymbol{POEM_{AR}^{PCA}}$	AR	POEM	PCA+AD	SA	PCA feature transform and preprocessing AR
$\boldsymbol{POEM_{AS}^{PCA}}$	AS	POEM	PCA+AD	SA	PCA feature transform and preprocessing AS
$\boldsymbol{POEM_{DG}^{PCA}}$	DG	POEM	PCA+AD	SA	PCA feature transform and preprocessing DG
$\boldsymbol{POEM_{AR}^{NPE}}$	AR	POEM	NPE +AD	SA	NPE feature transform and preprocessing AR

$\boldsymbol{POEM_{AS}^{NPE}}$	AS	POEM	NPE +AD	SA	NPE feature transform and preprocessing AS
$\boldsymbol{POEM_{DG}^{NPE}}$	DG	POEM	NPE +AD	SA	NPE feature transform and preprocessing DG
$\boldsymbol{POEM_{AR}^{LLE}}$	AR	POEM	LLE +AD	SA	LLE feature transform and preprocessing AR
$\boldsymbol{POEM_{AS}^{LLE}}$	AS	POEM	LLE +AD	SA	LLE feature transform and preprocessing AS
$\boldsymbol{POEM_{DG}^{LLE}}$	DG	POEM	LLE +AD	SA	LLE feature transform and preprocessing DG
$\boldsymbol{POEM_{AR}^{ONPP}}$	AR	POEM	ONPP +AD	SA	ONPP feature transform and preprocessing AR
$\boldsymbol{POEM_{AS}^{ONPP}}$	AS	POEM	ONPP +AD	SA	ONPP feature transform and preprocessing AS
$\boldsymbol{POEM_{DG}^{ONPP}}$	DG	POEM	ONPP +AD	SA	ONPP feature transform and preprocessing DG
$\boldsymbol{POEM_{AR}^{LPP}}$	AR	POEM	LPP +AD	SA	LPP feature transform and preprocessing AR
$\boldsymbol{POEM_{AS}^{LPP}}$	AS	POEM	LPP +AD	SA	LPP feature transform and preprocessing AS
$\boldsymbol{POEM_{DG}^{LPP}}$	DG	POEM	LPP +AD	SA	LPP feature transform and preprocessing DG
$\boldsymbol{E\_P_{ALL}^{PCA}}$	(AR, AS, DG)	POEM	PCA+AD	EE	Perturbation of enhancement with PCA
$\boldsymbol{E\_P_{ALL}^{NPE}}$	(AR, AS, DG)	POEM	NPE+AD	EE	Perturbation of enhancement with NPE
$\boldsymbol{E\_P_{ALL}^{LLE}}$	(AR, AS, DG)	POEM	LLE+AD	EE	Perturbation of enhancement with LLE
$\boldsymbol{E\_P_{ALL}^{ONPP}}$	(AR, AS, DG)	POEM	ONPP+AD	EE	Perturbation of enhancement with ONPP
$\boldsymbol{E\_P_{ALL}^{LPP}}$	(AR, AS, DG)	POEM	LPP+AD	EE	Perturbation of enhancement with LPP
$\boldsymbol{E\_P_{AR}^{ALL}}$	AR	POEM	ALL+AD	EF	Perturbation of feature transform
$\boldsymbol{E\_P_{ALL}^{ALL}}$	(AR, AS, DG)	POEM	ALL+AD	EEF	Perturbation of enhancement and feature transform
$\boldsymbol{MBC_{AR}^{NPE}}$	AR	MBC	NPE +AD	SA	NPE feature transform and preprocessing AR
$\boldsymbol{MBC_{AS}^{NPE}}$	AS	MBC	NPE +AD	SA	NPE feature transform and preprocessing AS
$\boldsymbol{MBC_{DG}^{NPE}}$	DG	MBC	NPE +AD	SA	NPE feature transform and preprocessing DG
$\boldsymbol{E\_M_{ALL}^{NPE}}$	(AR, AS, DG)	MBC	NPE +AD	EE	Perturbation of enhancement with NPE
$\boldsymbol{E\_M_{ALL}^{ALL}}$	(AR, AS, DG)	MBC	ALL+AD	EEF	Perturbation of enhancement and feature transform
POEM [10]	-	POEM	- + CS	SA	POEM descriptor (source code from ([10])
POEM+PCA [34]	-	POEM	PCA+AD	SA	code of [34] with fixed PCA dimension (500)

Table 1. Compared approaches.

	FERET Datasets					LFW Dataset
Method	**Fb**	**Fc**	**Dup1**	**Dup2**	**Average**	
$POEM_{LBP}^{PCA}$	98.5	100	90.4	89.3	94.5	74.9
$POEM^{PCA}$	98.7	99.0	88.1	83.8	92.4	-
$POEM_{AR}^{PCA}$	98.5	100	91.0	90.2	94.9	-
$POEM_{AS}^{PCA}$	98.5	99.5	90.6	91.7	95.1	-
$POEM_{DG}^{PCA}$	98.2	99.5	89.2	87.6	93.6	-
$E\_P_{ALL}^{PCA}$	98.8	100	94.2	94.0	96.7	76.6
POEM [10]	98.1	99.0	79.6	79.1	88.9	75.4
POEM+PCA [34]	99.6	99.5	88.8	85.0	93.2	82.7

Table 2. Accuracy obtained by our methods in FERET and LFW databases.

Examining Table 2, the following conclusions can be drawn:

- Our ensemble $E\_P_{ALL}^{PCA}$ boosts the performance of the base POEM descriptors in both datasets and outperforms [34] in the FERET datasets without any strong optimization (i.e. by using the same parameter settings for both the four FERET datasets and the LFW dataset);
- In LFW, the authors of [34] claim the highest results using an ad hoc projection matrix. In contrast, we use the unsupervised projection matrix for PCA that is constructed using training images (with no consideration of the labels) of the FERET dataset.

The second set of experiments, see Table 3, is aimed at evaluating post-processing techniques to reduce noise after feature extraction. For this, we test the different unsupervised feature transforms detailed in section 2.5. Each feature transform is not applied directly on the POEM features but rather on data projected onto a whitened PCA subspace (i.e. a subspace whose covariance is the identity matrix, so that the features are uncorrelated and all have variance 1). The final space after the feature transform has a dimension of 1000.

Method	FERET Datasets					LFW Dataset
	Fb	Fc	Dup1	Dup2	Average	
$\boldsymbol{POEM_{AR}^{NPE}}$	98.9	98.5	90.4	91.4	94.8	75.7
$\boldsymbol{POEM_{AS}^{NPE}}$	98.3	99.5	90.9	90.6	94.8	73.9
$\boldsymbol{POEM_{DG}^{NPE}}$	98.5	99.5	89.2	88.9	94.0	76.6
$\boldsymbol{E\_P_{ALL}^{NPE}}$	99.3	100	93.9	94.0	96.8	76.9
$\boldsymbol{POEM_{AR}^{LLE}}$	98.7	99.5	91.0	91.0	95.1	75.4
$\boldsymbol{POEM_{AS}^{LLE}}$	99.1	99.5	90.0	89.7	94.6	73.6
$\boldsymbol{POEM_{DG}^{LLE}}$	98.7	99.5	90.7	89.3	94.6	76.1
$\boldsymbol{E\_P_{ALL}^{LLE}}$	99.0	100	94.0	93.6	96.7	76.7
$\boldsymbol{POEM_{AR}^{ONPP}}$	98.6	98.5	90.3	92.7	95.0	75.8
$\boldsymbol{POEM_{AS}^{ONPP}}$	98.7	99.0	90.6	90.6	94.7	73.2
$\boldsymbol{POEM_{DG}^{ONPP}}$	98.5	99.5	89.6	89.7	94.3	76.5
$\boldsymbol{E\_P_{ALL}^{ONPP}}$	99.3	100	94.7	95.3	97.3	77.3
$\boldsymbol{POEM_{AR}^{LPP}}$	99.1	100	92.7	92.7	96.1	75.6
$\boldsymbol{POEM_{AS}^{LPP}}$	99.1	99.5	92.5	92.7	95.9	73.9
$\boldsymbol{POEM_{DG}^{LPP}}$	98.6	99.5	91.9	90.2	95.0	76.5
$\boldsymbol{E\_P_{ALL}^{LPP}}$	99.3	100	94.9	94.4	97.1	77.3
$\boldsymbol{E\_P_{AR}^{ALL}}$	99.1	99.5	92.4	93.2	96.1	77.3
$\boldsymbol{E\_P_{ALL}^{ALL}}$	99.3	100	94.7	94.9	97.2	77.1

Table 3. Performance obtained by different unsupervised feature transforms.

From the results reported in Table 3, the following conclusions can be drawn:

- For all the tested UDR techniques, the fusion among the three enhancement methods outperforms the stand-alone approaches;
- The best trade-off of performance/computation time is $\boldsymbol{E\_P_{AR}^{ALL}}$, where only the *AR* enhancement is applied, and the ensemble is built combining the scores obtained using different UDR (see Table 7 for time comparisons).

It is well known that combining different kinds of descriptors improves performance; so, as a third set of experiments (see Table 4), we combine our idea for building ensembles (based on the perturbation of enhancements) with the MBC descriptor [12], fixing NPE as the feature transform (to avoid a huge table), except for the ensemble where the different feature transforms are combined. The accuracy is reported in Table 4. As in previous tests, the ensemble outperforms the stand-alone approaches.

	FERET Datasets					LFW Dataset
Method	Fb	Fc	Dup1	Dup2	Average	
$\boldsymbol{MBC_{AR}^{NPE}}$	98.2	98.9	85.3	81.2	90.9	74.8
$\boldsymbol{MBC_{AS}^{NPE}}$	98.2	97.9	89.2	85.9	92.8	72.9
$\boldsymbol{MBC_{DG}^{NPE}}$	98.5	98.9	90.2	88.0	93.9	75.6
$\boldsymbol{E\_M_{ALL}^{NPE}}$	98.5	99.5	90.7	87.5	94.1	75.6
$\boldsymbol{E\_M_{ALL}^{ALL}}$	98.6	99.5	92.7	91.5	95.6	75.9

Table 4. Performance obtained by different approaches based on the MBC descriptor.

Notice that the results obtained by the MBC descriptor improve the performance reported in the original work [12], where the best reported performance without a supervised feature transform in the FERET dataset is lower (see Table 5, line: [12] (2012) no PCA).

Finally a comparison among our best approaches and a selection of the methods reported in the literature is reported in Table 5. We select for comparison $\boldsymbol{E\_P_{ALL}^{ALL}}$, $\boldsymbol{E\_M_{ALL}^{ALL}}$ and their fusion by the sum rule ($\boldsymbol{E\_P_{ALL}^{ALL}+E\_M_{ALL}^{ALL}}$).

In Table 5 we compare our approaches with the state-of-art for both FERET and LFW (several other approaches are reported at http://vis-www.cs.umass.edu/lfw/results.html for LFW and see [9] for results of other approaches using the FERET datasets). For both datasets, it is clear that in the last few years the performance of the state-of-the-art has greatly increased. Moreover, for a fair comparison with other approaches tested on LFW, we report only the methods that do not use outside training data in developing the recognition systems (considering not only classifier training but also alignment and feature extraction). In the last five years, because of the increase in the computational power of modern PCs, several approaches have made use of outside training data obtaining higher performance , e.g. [15] (93.3%), “face.com” (the face recognition system acquired by Facebook) [13] (91.3%) and [35] (92.6%). For recent updates see as well http://vis-www.cs.umass.edu/lfw/.

	FERET Datasets					LFW Dataset
Method	Fb	Fc	Dup1	Dup2	Average	
$E\_P_{ALL}^{ALL}$	99.3	100	94.7	94.9	97.2	77.1
$E\_M_{ALL}^{ALL}$	98.6	99.5	92.7	91.5	95.6	75.9
$E\_P_{ALL}^{ALL}+E\_M_{ALL}^{ALL}$	99.2	100	95.6	94.4	97.3	77.8
[18] (2013)	98.7	100	94.6	93.6	96.7	76.9
[36] (2004)	93.0	51.0	61.0	50.0	63.8	---
[37] (2005)	94.0	97.0	68.0	58.0	79.2	---
[38] (2005)	96.3	99.5	78.8	77.8	88.1	---
[39] (2007)	97.6	99.0	77.7	76.1	87.6	---
[40] (2007)	98.0	98.0	90.0	85.0	92.8	---
[41] (2010)	99.0	99.0	94.0	93.0	96.3	---
[12] (2012) no PCA	97.6	99	85.2	83.3	91.2	
[12] (2012)	99.7	99.5	93.6	91.5	96.1	
[9] (2013)	99.7	100	94.9	94.0	97.2	86.2
[42] (2007)	---	---	---	---	---	73.9
[43] (2008)	---	---	---	---	---	78.5
[6] (2009)	---	---	---	---	---	79.3
[44] [1][1](2009)	---	---	---	---	---	84.0
[17] (2010)	---	---	---	---	---	86.8
[45] (2011)	---	---	---	---	---	88.0
[46] (2011)	---	---	---	---	---	88.1

Table 5. Comparison among the proposed ensembles with state-of-the-arts approaches.

From the results reported in Table 5, it is clear that the performance using frontal and upright images, as in the FERET datasets, are very high, therefore, similar (or maybe better) than those scores obtained by human beings. This result is already reported in [47], where it is shown that the best algorithms tested on the FRGC and FRVT 2006 contests can outperform human raters. On LFW dataset, however, machine performance is lower than that obtained by human raters (above all for those systems not using outside training data). In [14] it is shown that human beings using cropped images obtained an accuracy of 97.5% in the LFW dataset.

Thus, it is clear that our ensemble can be considered a state-of-the-art for the FERET dataset, but it obtains a performance that is lower than the state-of-the-art in the LFW dataset. We want to stress, however, that our approach is texture based and does not use any outside training data beyond the alignment step. Moreover, it is a low-level face representation that can be easily combined with other approaches. It has already been shown in [9] that texture base approaches are better suited for upright images, as in the FERET datasets, than for more variable images such as those in the LFW dataset since they require very good alignments between the two images that should be matched. For this reason, we have not obtained good results in the LFW dataset using the FERET training set for the supervised feature transforms and region selection.

Even if our approach does not reach the state-of-the-art performance on the LFW dataset, it is still potentially useful as a component of an ensemble that improves the performance of different descriptors. As an example we combined the classification score from our best approach with a classification method based on the features (HLF) extracted in [14] (high level features that use binary classifiers to check for the presence or absence of the visual appearance of gender, race, and age, available at http://www.cs.columbia.edu/CAVE/databases/pubfig/download/lfw_attributes.txt). The classification approach based on HLF and a random subspace ensemble [48] of Support Vector Machine (SVM) [49] as the classifier obtained an accuracy of 81.6% on the LFW dataset. The fusion by sum rule (after the normalization of the scores) between HLF and our best ensemble obtains an accuracy of 85%, which strongly improved both stand-alone approaches.

For a final test, we report in Table 6 the results obtained using more enhancement techniques, using the three methods already detailed as well as the following six additional approaches:

1. Low-frequency discrete cosine transform based approach [50];
2. Oriented Local Histogram Equalization [51];
3. Multiscale retinex [52];
4. Isotropic smoothing normalization [53];
5. Photometric normalization [54];
6. Gradientfaces [55];

The new methods, whose results are reported in Table 6, are:

$\boldsymbol{E\_P_{9}^{ALL}}$, as in $\boldsymbol{E\_P_{ALL}^{ALL}}$ but using all nine enhancements;

$\boldsymbol{E\_M_{9}^{ALL}}$, as in $\boldsymbol{E\_M_{ALL}^{ALL}}$ but using all nine enhancements;

$\mathbf{2 * E\_P_9^{ALL}{+}E\_M_9^{ALL}}$, as in $\mathbf{E\_P_{ALL}^{ALL}{+}E\_M_{ALL}^{ALL}}$ but using all nine enhancements and the weighted sum rule ($\mathbf{E\_P_9^{ALL}}$ is assigned a weight of 2, and $\mathbf{E\_M_9^{ALL}}$ is assigned a weight of 1).

The reported results show that this ensemble based on nine enhancements improves the performance of the monogenetic methods in all the datasets, while POEM works slightly worse in the FERET datasets and better in the LFW dataset.

Hence, $\mathbf{2 * E\_P_9^{ALL}{+}E\_M_9^{ALL}}$ makes only five mistakes more than $\mathbf{E\_P_{ALL}^{ALL}{+}E\_M_{ALL}^{ALL}}$ in the four FERET datasets, and correctly classifies fifty-four images more than $\mathbf{E\_P_{ALL}^{ALL}{+}E\_M_{ALL}^{ALL}}$ in the LFW dataset.

	FERET Datasets					**LFW Dataset**
Method	Fb	Fc	Dup1	Dup2	Average	
$\mathbf{E\_P_{ALL}^{ALL}}$	99.3	100	94.7	94.9	97.2	77.1
$\mathbf{E\_M_{ALL}^{ALL}}$	98.6	99.5	92.7	91.5	95.6	75.9
$\mathbf{E\_P_{ALL}^{ALL}{+}E\_M_{ALL}^{ALL}}$	99.2	100	95.6	94.4	97.3	77.8
$\mathbf{E\_P_9^{ALL}}$	99.2	100	94.6	93.6	96.9	78.6
$\mathbf{E\_M_9^{ALL}}$	98.9	99.5	93.4	91.9	95.9	76.6
$\mathbf{2 * E\_P_9^{ALL}{+}E\_M_9^{ALL}}$	99.2	100	95.2	93.6	97.0	78.7
[18] (2013)	98.7	100	94.6	93.6	96.7	76.9

Table 6. Comparison among the proposed ensembles based on different enhancements.

4. Conclusion

In this paper we have shown that it is possible to improve the performance of a single descriptor (POEM) by building an ensemble obtained by perturbing some steps in the face recognition process. In particular, our experiments show that the most reliable approach for building an ensemble is to perturb the enhancement method.

The main drawback of our proposed system is the increase in computation time with respect to stand-alone methods. For example, considering $E\_P_{ALL}^{NPE}$, the time needed for the enhancement and the feature extraction steps is ~1 second, while the matching time is ~ 0.00013 seconds using an Intel i5 3.3GhZ processor with 8G of RAM and parallelized MATLAB code. Other computation times are reported in Table 7.

Method	seconds
$E\_P_{AR}^{ALL}$	1.05
$E\_P_{ALL}^{NPE}$	1.03
$E\_M_{ALL}^{NPE}$	1.45
[12]	0.43
GABOR [2]	0.088
POEM	0.036

Table 7. Comparison of computation times among different methods.

Some avenues for future explorations would include investigating the following: a) testing other feature transformations before the matching step; b) combining our proposed POEM-based approach with other descriptors; and c) testing other texture descriptors for representing the AM images of POEM.

References

[1] S. Muruganantham, T. Jebarajan, A comprehensive review of significant researches on face recognition based on various conditions, International Journal of Computer Theory and Engineering, 4 (2012) 7-15.
[2] A.-A. Bhuiyan, C.H. Liu., On Face Recognition Using Gabor Filters, World academy of science, engineering and technology, 28 (2007) 51-56.
[3] M. Lades, J.C. Vorbrueggen, J. Buhmann, J. Lange, C.v.d. Malsburg, R.P. Wuertz, W. Konen, Distortion invariant object recognition in the dynamic link architecture, IEEE Trans Comput, 42 (1993) 300-311.
[4] M.A. Turk, A.P. Pentland, Eigenfaces for recognition, J Cogn Neurosci, 3 (1991) 71-86.
[5] L. Nanni, A. Lumini, F. Dominio, P. Zanuttigh, Effective and precise face detection based on color and depth data, Applied Computing and Informatics, 10 (2014) 1-13.
[6] N. Pinto, J. DiCarlo, D. Cox, How far can you get with a modern face recognition test set using only simple features?, in: CVPR'09, 2009.
[7] Z. Cao, Q. Yin, X. Tang, J. Sun, Face recognition with learning-based descriptor, (2010) 2707-2714.
[8] N.-S. Vu, A. Caplier, Face recognition with patterns of oriented edge magnitudes, in: ECCV, 2010.
[9] N.-S. Vu, Exploring patterns of gradient orientations and magnitudes for face recognition, IEEE Transactions on Information Forensics and Security, 8 (2013) 295-304
[10] N.-S. Vu, H.M. Dee, A. Caplier, Face recognition using the POEM descriptor, Pattern Recognit Lett, 45 (2012) 2478-2488.
[11] Y. Liang, S. Liao, L. Wang, B. Zou, Exploring regularized feature selection for person specific face verification, in: ICCV, 2011.
[12] M. Yang, L. Zhang, S.C.K. Shiu, D. Zhang, Monogenic binary coding: An efficient local feature extraction approach to face recognition, IEEE Transactions on Information Forensics and Security, 7 (2012) 1738-1751.
[13] Y. Taigman, L. Wolf, Leveraging Billions of Faces to Overcome Performance Barriers in Unconstrained Face Recognition, ArXiv e-prints, (2011) http://arxiv.org/pdf/1108.1122.pdf.
[14] N. Kumar, A.C. Berg, P.N. Belhumeur, S.K. Nayar, Attribute and simile classifiers for face verification, in: International Conference on Computer Vision (ICCV), 2009.
[15] T. Berg, P.N. Belhumeur, Tom-vs-Pete Classifiers and Identity-Preserving Alignment for Face Verification, in: British Machine Vision Conference (BMVC), 2012.

[16] C.H. Chan, M.A. Tahir, J. Kittler, M. Pietikäinen, Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors, IEEE Trans Pattern Analysis Mach Intell, 35 (2013) 1164–1177.
[17] L. Wolf, T. Hassner, Y. Taigman, Similarity scores based on background samples, in: Asian Conference on Computer Vision (ACCV), 2009.
[18] L. Nanni, A. Lumini, S. Brahnam, M. Migliardi, Ensemble of patterns of oriented edge magnitudes descriptors for face recognition, in: International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'13), 2013.
[19] J. Ylioinas, A. Hadid, Y. Guo, M. Pietikäinen, Efficient image appearance description using dense sampling based local binary patterns, in: Asian Conference on Computer Vision, 2012.
[20] W.T. Freeman, E.H. Adelson, The design and use of steerable filters, IEEE Transactions on Pattern analysis and machine intelligence., 13 (1991) 891-906.
[21] M. Unser, D. Sage, D.V.D. Ville, Multiresolution monogenic signal analysis using the riesz-laplace wavelet transform, IEEE Transactions on Image Processing, 18 (2009) 2402-2418.
[22] P. Belhumeur, J. Hespanha, D. Kriegman, Eigenfaces vs. fisherfaces: Recognition using class specific linear projection, IEEE Trans Pattern Analysis Mach Intell, 19 (1997) 711-720.
[23] A. Lumini, L. Nanni, Overview of the combination of biometric matchers, Inf Fusion, 33 (2017) 71–85.
[24] Y.K. Park, S.L. Park, J.K. Kim, Retinex method based on adaptive smoothing for illumination invariant face recognition, Signal Processing, 88 (2008) 929-1945.
[25] R. Gross, V. Brajovic, An image preprocessing algorithm for illumination invariant face recognition, in: 4th International Conference on Audio-and Video-Based Biometric Personal Authentication, 2003.
[26] V. Štruc, N. Pavešic, Photometric normalization techniques for illumination invariance, in: V. Štruc, N. Pavešic (Eds.) Advances in Face Image Analysis: Techniques and Technologies, IGI Global, Hershey, PA, 2011, pp. 279-300.
[27] S. Roweis, L. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science, 290 (2000) 2323-2326.
[28] H. Xiaofei, D. Cai, S. Yan, H.-J. Zhang, Neighborhood preserving embedding, in: Tenth IEEE International Conference on Computer Vision (ICCV'2005), Beijing, China, 2005.
[29] Y. Xu, A. Zhong, J. Yang, D. Zhang, LPP solution schemes for use with face recognition, Pattern Recognition, 43 (2010) 4165-4176.
[30] E. Kokiopoulou, Y. Saad, Orthogonal neighborhood preserving projections, in: IEEE International conference on Data Mining, New Orleans, LA, 2005.
[31] M. Belkin, P. Niyogi, Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering, Nips, 14 (2001) 585-591.
[32] J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The feret evaluation methodology for face-recognition algorithms, IEEE Transactions of Pattern Analysis and Machine Intelligence, 22 (2000) 1090-1104.
[33] G.B. Huang, M. Ramesh, T. Berg, E. Learned-Miller, Labeled faces in the wild: a database for studying face recognition in unconstrained environments, in, University of Massachusetts, Amherst, 2007.
[34] N.-S. Vu, A. Caplier, Enhanced patterns of oriented edge magnitudes for face recognition and image matching, IEEE Transactions on Image Processing, 21 (2012) 1352-1365.
[35] C. Huang, S. Zhu, K. Yu, Large scale strongly supervised ensemble metric learning, with applications to face verification and retrieval, in, NEC, 2011.
[36] T. Ahonen, A. Hadid, M. Pietikainen, Face recognition with local binary patterns, in: ECCV'04, 2004, pp. 469-481.
[37] W. Zhang, S. Shan, W. Gao, X. Chen, 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’05, Beijing, 2005, pp. 786-791.
[38] W. Deng, J. Hu, J. Guo, Gabor-eigen-whiten-cosine: A robust scheme for face recognition, in: W. Zhao, S. Gong, X. Tang (Eds.) Second International Workshop, AMFG, Beijing, China, 2005, pp. 336-349.

[39] B. Zhang, S. Shan, X. Chen, W. Gao, Histogram of gabor phase patterns (HGPP): A novel object representation approach for face recognition, IEEE Transactions of Image Processing, 16 (2007) 57-68.
[40] X. Tan, B. Triggs, Fusing gabor and LBP feature sets for kernel-based face recognition, in: Analysis and Modelling of Faces and Gestures LNCS 4778, Springer, NY, 2007, pp. 235-249.
[41] S. Xie, S. Shan, X. Chen, J. Chen, Fusing local patterns of gabor magnitude and phase for face recognition, IEEE Transactions on Image Processing, 19 (2010) 1349-1361.
[42] E. Nowak, F. Jurie, Learning visual similarity measures for comparing never seen objects, in: Computer Vision and Pattern Recognition (CVPR), 2007.
[43] L. Wolf, T. Hassner, Y. Taigman, Descriptor based methods in the wild, Science, 6 (2008) 1-14.
[44] Y. Taigman, L. Wolf, T. Hassner, Multiple one-shots for utilizing class label information, in: British Machine Vision Conference (BMVC), 2009.
[45] H.V. Nguyen, L. Bai, Cosine similarity metric learning for face verification, in: Asian Conference on Computer Vision (ACCV), 2010.
[46] N. Pinto, D. Cox, Beyond simple features: a large-scale feature search approach to unconstrained face recognition, in: International Conference on Automatic Face and Gesture Recognition (FG), 2011.
[47] A.J. O'Toole, P.J. Phillips, A. Narvekar, Humans versus algorithms: Comparisons from the Face Recognition Vendor Test 2006, in: 8th International Conference on Automatic Face & Gesture Recognition FG '08, 2008, pp. 1-6.
[48] T.K. Ho, The random subspace method for constructing decision forests, IEEE Trans Pattern Analysis Mach Intell, 20 (1998) 832-844.
[49] M.A. Heartt, B. Schölkopf, S. Dumais, E. Osuna, J. Platt, Trends and controversies-Support vector machines, IEEE Intelligent Systems,, 13 (1998) 18-28.
[50] W. Chen, E. Meng-Joo, W. Shiqian, Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain, IEEE Trans. Syst. Man Cyber. B, 36 (2006) 458-466.
[51] P.H. Lee, S.W. Wu, Y.P. Hung, Illumination compensation using oriented local histogram equalization and its application to face recognition., IEEE Transactions on Image Processing, (2012) 1-10.
[52] D.J. Jobson, Z.u. Rahman, G.A. Woodell, A multiscale retinex for bridging the gap between color images and the human observation of scenes, IEEE Transactions on Image Processing, 6 (1997) 965-976.
[53] G. Heusch, F. Cardinaux, S. Marcel, Lighting normalization algorithms for face verification, IDIAP-com 05-03, 9 (2005).
[54] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, Analysis and Modelling of Faces and Gestures, LNCS 4778 (2007) 168-182.
[55] T. Zhang, Y.Y. Tang, B. Fang, Z. Shang, X. Liu, Face recognition under varying illumination using gradientfaces, IEEE Transactions on Image Processing, 18 (2009) 2599-2606.