

RESEARCH

Open Access



# KCCA-based technique for profile face identification

Malek Nadil<sup>1,2\*</sup>, Feryel Souami<sup>2</sup>, Abdenour Labeled<sup>1</sup> and Hichem Sahbi<sup>3</sup>

## Abstract

During the last two decades, satisfactory results have been obtained for face identification techniques based on frontal pose. However, face identification from uncontrolled pose remains a challenging open problem in biometric recognition. Recently, pose invariant techniques that exploit either 3D scans or 2D images of the same face to generate the corresponding 3D model have emerged. Even if they tolerate pose variability and lead to high identification scores, they have the drawback to be computationally intensive and/or require the cooperation of the individual to be identified. Hence, they are not appropriate for the interesting real-time application of video surveillance. In this paper, we propose a profile face identification method based on correspondence mapping of 2D frontal face images. Kernel canonical correlation analysis (KCCA) is used to learn changeover from the profile pose to the frontal one. To show the effectiveness of our approach, tests are performed on FERET database according to a protocol referred to as leave one out-like protocol (LOOLP). These tests demonstrate that it leads to enhanced scores comparatively to other 2D-based methods proposed in the literature.

**Keywords:** Profile face identification, Frontal face, KCCA, LBP, Biometry

## 1 Introduction

Biometric security is an active research area with applications in several domains such as public space surveillance. Numerous biometric features like fingerprints, iris, and hand geometry are widely used in the surveillance systems [1]. These features allow a unique description of a person. Nevertheless, in public space surveillance area, a person is hard to be identified from such features through video flows and faces are rather used as “biometric” features to recognize/identify a person. Face recognition techniques have evolved theoretically and technologically since the 1970s but still suffer from individuals’ appearance variability due to lighting conditions, posture, facial expressions, and aging [2]. Pose variation in uncontrolled acquisition systems is one of the most challenging problems for face recognition. The literature proposes several methods for pose variability problems.

They can be categorized according to specific criteria: 2D-based and 3D-based models or single image-based and multiple image-based methods [3]. In this work, the 2D-based/3D-based model categorization is considered, and 2D images are used for face recognition in our approach. A profile-frontal face recognition method using a regression based on a canonical correlation analysis of face components is proposed. These components are represented by the local binary pattern (LBP) descriptor [4]. The transformation between profile face image (90° angle pose) and frontal face image (0° angle pose) is known to be nonlinear, and this nonlinearity is taken into account by introducing kernel functions. For the evaluation of the proposed approach, we use a leave one out-like protocol (LOOLP) on the FERET [5] and SCface [6] databases. The present paper is organized as follows. In Section 2, we present a state of the art in face recognition with pose variability methods. We introduce our overall approach in Section 3. In Section 4, we describe the preprocessing steps required by the geometric normalization of faces. Section 5 is devoted to the LBP descriptor used for feature extraction. In Section 6, the method used to carry out the correspondence between profile and frontal poses is developed. In Section 7, we present the dataset used

\*Correspondence: nad.malek@gmail.com

<sup>1</sup>Ecole Militaire Polytechnique, Computer Science Department, BP 17, 16111 Algiers, Algeria

<sup>2</sup>Université des Sciences et Technologies Houari Boumedienne, BP 32 EL ALIA, 16111 Algiers, Algeria

Full list of author information is available at the end of the article

for learning and tests and summarize the scores of validation tests of the proposed approach. Comparisons of the method with those of the literature are provided by Section 8.

## 2 Face recognition with pose variability

This paper deals with the problem solving of pose variability using a 2D/3D-based method. Before introducing our approach, an overview of 2D-based methods proposed in the literature is first given and then short description of 3D-based methods is presented.

### 2.1 2D image techniques

The techniques using 2D images are subdivided into three categories [3]: real view-based matching, pose transformation in image space, and pose transformation in feature space.

1. Real view-based matching: Consists of representing the individuals in the gallery with different rotations and look for the individual to be identified in this gallery, once his pose has been determined. In [7], a method with a gallery of 15 images per individual which covers pose variability with  $\pm 40^\circ$  in yaw and  $\pm 20^\circ$  in tilt has been proposed. Identification process is typically a template matching algorithm with templates around the eyes, nose, and mouth, where the only difference is that it matches a probe face image with gallery face images in similar poses. These methods are easy from an algorithmic point of view but are difficult to exploit practically since they need images of faces in several orientations.
2. Pose transformation in image space: One way to overcome the shortcoming of collecting numerous images per person, in order to cover different orientations, is to create artificial sights from existing images. Beymer and Poggio [8] proposed parallel deformation to generate virtual sights from a single example view using feature-based 2D warping [9]. Their algorithm allows generating new faces, covering  $-30^\circ$  to  $30^\circ$  rotations in yaw and  $-15^\circ$  to  $15^\circ$  rotations in tilt. Another approach based on Active Shape Models (ASM) [10] has been proposed in [11] to generate new poses of faces. However, this technique performs well only with small angles shifting. As extension to ASM-based approach, the Active Appearance Models (AAM) [12] have been used to take simultaneously account of shape and texture variations of the face. Numerous techniques for pose invariance identification based on AAMs have been proposed in the literature [13–15]. It is stated that even if they give better results than those based on ASM, their performance is proven only for small orientations shifts.

3. Pose transformation in feature space: Tolerance to pose variation can also be achieved in feature space. The most known techniques use kernel functions to map the images of faces in a space of higher dimension. Thus, non-separable distributions because of pose variations may become linearly separable. We can cite the works of Liu [16] and Xie and Lam [17] where kernel principal components analysis (KPCA) is used and the works of Huang et al. [18] and Yang et al. [19] where kernel Fisher discriminant analysis (KFDA) is used. The idea of using a regression model to extract frontal views from non-frontal ones has been exploited. For instance, Chai et al. [20] proposed a local linear regression (LLR) to create virtual frontal view from single horizontally rotated views. Prince et al. [21] proposed a linear statistical model referred to as tied factor analysis model, to describe pose variations on face images. The main point in this approach is to find features which are invariant to orientation. It provided high scores, even for large variations in the angle of view, outperforming methods proposed earlier.

### 2.2 3D model techniques

3D-based face recognition techniques can be subdivided into two categories: firstly, methods based on 3D scans of the face [22, 23] that perform well regardless the pose variation but require very specific and expensive devices together with a large amount of time to generate the 3D model of the face. In this case, the cooperation of the person to be identified is needed. Consequently, they are not adequate for real-time recognition like in video surveillance. Secondly, methods that use 2D images for a 3D recognition where the cooperation of individuals is not needed. Among these methods, we have:

1. Generic shape-based approaches: They use 3D shapes to represent the face. For instance, the method proposed in [24] uses a cylinder to map the face in different orientations, and then the frontal pose of the face is generated. Another method [25] uses an ellipsoid to warp the texture of the face. These techniques are fast but perform well only if pose varies in a small range.
2. 3D face reconstructions: They can be subdivided into two subcategories, feature-based methods and image-based methods. Features extracted from 2D images (e.g., edges and corners), or intensity of pixels, can be used to construct the 3D shapes. Among methods which use features extracted from face, we can find the Lee and Ranganath method [26] that presents a

composite 3D deformable face model for pose estimation and face synthesis based on a template deformation which maintains connectedness and smoothness. Jiang et al. [27] used facial features to efficiently reconstruct personalized 3D face models from a single frontal face image for recognition. Their method is based on the automatic detection of facial features on the frontal views using Bayesian shape localization. Image-based 3D face reconstructions carefully study the relationship between image pixel intensities and its corresponding shape/texture properties. In this context, Blanz and Vetter [28, 29] proposed a suitable face recognition system using 3D morphable model and Georgiades et al. [2] proposed illumination cone models which successfully performs face recognition under pose and lighting variations, using the techniques of photometric stereo. Finally, stereo vision [30] techniques where 3D face models are reconstructed from 2D face images in different poses can also be applied.

### 2.3 2D model versus 3D technique

It is worth recalling that 3D-based techniques of face recognition outperform their 2D-based counterparts and may cover large variations of pose, since the human head is a 3D object and the changes of its appearance lay in a 3D space. However, this score enhancement is achieved at the expense of more processing overheads. Moreover, the recognition results of 2D techniques, given the variation of pose, vary from a technique to another, and the literature does not propose a unified protocol of evaluation for such techniques. In this paper, the proposed method is based on 2D images and is evaluated with respect to LOOLP.

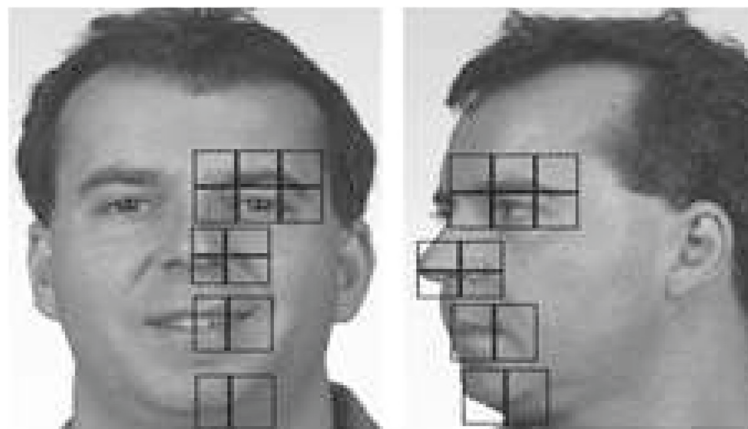
### 3 Description of the approach

The proposed approach for face identification relies on the idea of learning a transformation which maps the profile faces onto their corresponding frontal faces using canonical correlation analysis (CCA). We recall that the researchers who introduced the use of the CCA in face identification across pose differences are A. Li, S. Shan, X. Chen, and W. Gao [31]. In their work, the mapping between different orientations ( $20^\circ$ ,  $40^\circ$ , and  $60^\circ$ ) and frontal pose is built using gray level patches of faces. For  $60^\circ$  orientation, which is close to the orientation considered in our contribution ( $90^\circ$ ), a true identification score of 65 % has been obtained on a gallery of forty faces. However, we can improve this score considerably by taking into account the nonlinear nature of the mapping between poses. This can be achieved through appropriate kernel functions leading to kernel-CCA. The choice of the kernel is not straightforward where several tests have to be performed in order to choose the appropriate kernel model and then tune its parameters.

Moreover, instead of raw pixels, we describe the face using LBP face components: eye, nose, mouth, and chin. In the sequel, our contribution is presented in more details.

In this work, the CCA learns the mapping between frontal and profile faces represented by their main components: eye, nose, mouth, and chin. Given the fact that the transformation from frontal to profile face is not linear, faces are not described by the whole image, but rather by only their components.

To describe the face main components, the local binary pattern descriptor is chosen because of its high performance in frontal face identification and its invariance against illumination changes [32–34]. Figure 1 shows the



**Fig. 1** Main parts of the face considered for the CCA mapping between frontal and left profile. The figure shows the components describing the faces to be used for the learning of the CCA-based mapping between frontal and left profile poses

components describing the faces to be used for the learning of the CCA-based mapping.

In our study, we have considered the transformation to be learned for frontal-left profile. However, if the transformation frontal-right profile is required instead, components represented in Fig. 2 are used, and then we operate a flip to transform the right profile into a left profile.

Face components are split into regions as illustrated in Fig. 1. For each region, a histogram is calculated on the extracted LBP descriptor. Then, the obtained histograms are concatenated in a unique vector of features that describes frontal or profile faces.

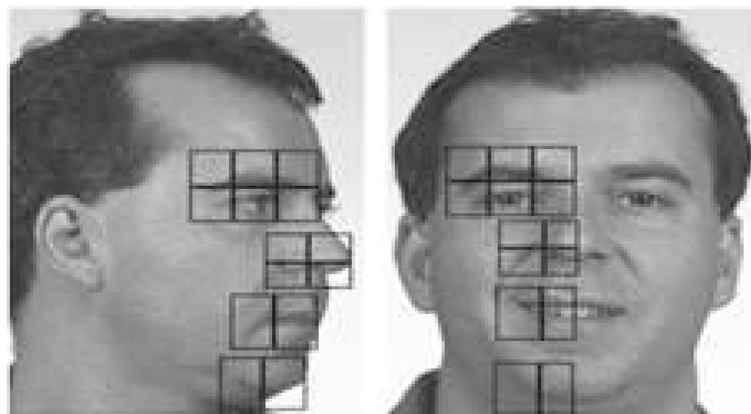
Generally, face identification methods use patches corresponding to regions of the same size but with different weights. The more relevant is the region, the higher is its weight. For instance, the eye region is assigned a higher weight than that of the chin. Inspired by this idea, in our work the relevance of a region is represented by the size of its patch given the fact that two patches with different sizes lead to histograms with the same number of bins. Hence, the weight of a patch is inversely proportional to its size. As shown in Fig. 1, the patches around the eye are smaller than those surrounding the chin since the former contains more relevant information than the latter.

Once the vectors are constructed, the CCA transforms the feature space where profile faces lay into the feature space where frontal faces lay.

As the derived vector of features is of high dimension, a dimensionality reduction helps to make the CCA applicable [35]. In this paper, the principal components analysis (PCA) is used for dimensionality reduction. Therefore, we use a kernel-PCA [16] followed by a CCA, resulting in a kernel canonical correlation analysis (KCCA) [36] which allows us to consider the nonlinearity of the transformation between the components of frontal and profile faces.

In order to evaluate our method, we experiment it on FERET database. Indeed, this database contains subsets for different orientations. We used the fa subset to constitute the dataset of the frontal images of faces ( $0^\circ$  orientation) and the pl subset for profile images ( $90^\circ$  orientation). Figure 3 illustrates examples of images taken from a pose subsets of FERET database.

The effectiveness of the transformation between frontal and profile poses is demonstrated through a leave one out-like protocol (LOOLP). We recall that in the conventional leave one out (LOO) procedure [37], with a dataset of size  $n$  in hand, at each validation step,  $n - 1$  individuals are used to train a classifier and the remaining  $n$ th individual is used for its test and henceforth until all the individuals have been considered for test. The testing protocol of our approach is inspired by the LOO with a slight difference. For the learning step, two subsets of size  $m$  from fa and pl for the same individuals are considered. In each iteration,  $m - 1$  faces (from both datasets) are used to learn the CCA transformation frontal-profile. However, for the remaining  $m$ th individual, only its frontal pose is considered and added to the frontal gallery (another subset of fa), and then the profile face of this  $m$ th individual together with the frontal faces of the augmented gallery are projected onto the canonical space for comparison. The Euclidean distance is then used to measure the proximity of this individual to those of the frontal faces. This operation is repeated until all the  $m$  individuals have been considered for test. This validation approach leads to a more reliable estimate of the generalization error of identification than those proposed in the previous works. Indeed, in these approaches, the learning and testing datasets are fixed in prior, which can lead to a biased estimate of the generalization error. However, in our evaluation, all examples are assigned the same weights and have the same contribution to both of learning and testing steps. It is also worth



**Fig. 2** Main parts of the face considered for the CCA mapping between frontal and right profile. The figure shows the components describing the faces to be used for the learning of the CCA-based mapping between frontal and right profile poses



**Fig. 3** Examples of images from FERET database corresponding to different poses. The figure illustrates examples of images taken from a pose subsets of FERET database

noting that these approaches have been tested on galleries of small sizes (about 100 faces). It is stated that the difficulty of pose-independent identification task grows with the size of the gallery [21]. In the present work, the validation is conducted on a gallery of 600 individuals. Figure 10 summarizes the way our approach is evaluated. The procedure is repeated until all the individuals of the learning dataset are considered.

#### 4 Geometric normalization

Before feature extraction step, a geometric normalization is performed in order to ensure that faces are in a similar scale, orientation, and position. Normalization procedure differs from frontal to profile faces. In the case of frontal pose, we detect the eyes in face image, whereas for profile images, we detect the nose and chin since the relative positions of these parts are generally stable versus varying facial expression [38]. Face component detection is achieved manually. We annotate these face components for each database image. Having annotated eyes in the frontal dataset and nose and chin in the profile dataset, we perform the geometric normalization as described below [38, 39]:

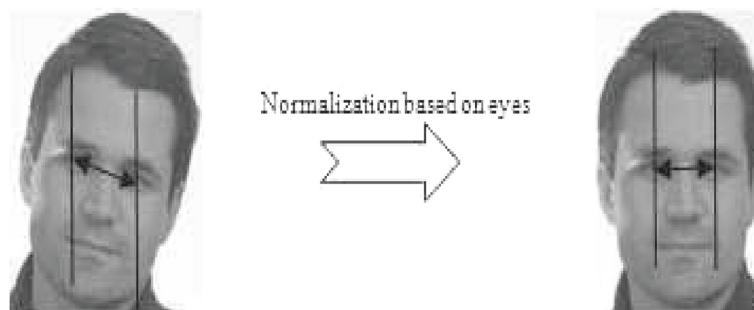
**Normalization for 0° orientation:** Images are automatically normalized thanks to the following steps:

1. Each image is rotated until the line joining the centers of eyes becomes horizontal.
2. Images are rescaled in order to get the same distance between the centers of eyes for all images.

Figure 4 shows the normalization for the 0° orientation.

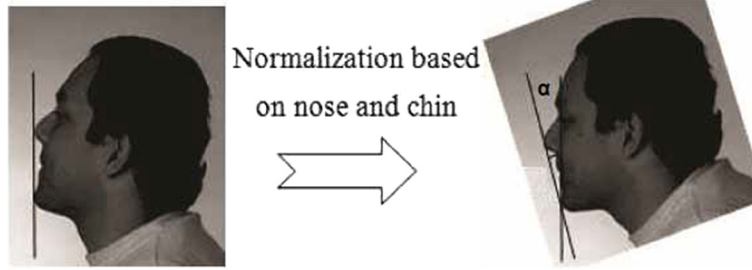
**Normalization for 90° orientation:** For 90° orientation, according to the tangent-based profile normalization technique [39], normalization consists of tilting by an angle  $\alpha$  the line joining the tip of the nose and the chin, from the vertical axis (Fig. 5).

In the case where the profile faces are not at the same scale, an operation of rescaling is necessary. Since this operation cannot be performed on 90° orientation, the learning normalized frontal database is used to maintain the distance between the eye and the chin in profile pose for all persons, as illustrated in Fig. 6. This step is added to be sure that the obtained scores are not biased by eventual normalization inaccuracies.



**Fig. 4** Image normalization for faces with 0° orientation. Geometric normalization is performed as described in figure for the 0° orientation





**Fig. 5** Image normalization for faces with 90° orientation. Geometric normalization is performed as described in figure for the 90° orientation

Rotation angle  $\alpha$  is determined empirically on the base of a series of preliminary tests where  $\alpha$  is varied between 18° and 22° with a step of 1°. These tests revealed that  $\alpha = 20^\circ$  is an appropriate choice for this task. Effectively, for this value of  $\alpha$ , all the rotated faces are in normalized profile pose.

To identify a profile face in practical situations, the image is tilted such that the tangent joining the tip of the nose and the chin, forms an angle  $\alpha = 20^\circ$  with the vertical axis. Then, to put it at the same scale as the profile faces of the learning database, the reference distance between the eyes and chin, used for rescaling, is calculated on the average face of the normalized frontal learning database (Fig. 7).

## 5 Local binary pattern

The LBP have been proposed by Ojala et al. [4] to characterize textures in images. For a pair  $(c, n)$ , with  $c$  a central pixel,  $n = (n_1, \dots, n_s)$  a set of pixels sampled from the neighborhood of  $c$ , the LBP operator assigns a 0 to each neighbor pixel in  $n$  that is smaller than the central pixel  $c$ , a 1 to each neighbor larger than  $c$ , and interpreting the result as a number in base 2 (Fig. 8). Consequently, if a neighborhood of  $S$  pixels is considered, there are  $2^S$  possible LBP values.

$$b = \sum_{i=1}^s 2^{i-1} I(c, n_i) \quad (1)$$

and

$$I(c, n_i) = \begin{cases} 1 & \text{if } c < n_i \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

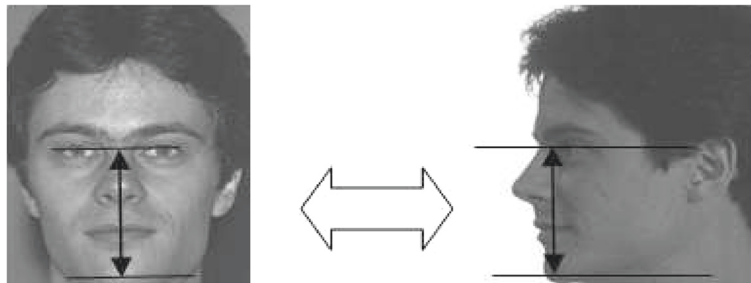
## 6 Kernel canonical correlation analysis

In our work, KCCA is obtained using KPCA followed by CCA.

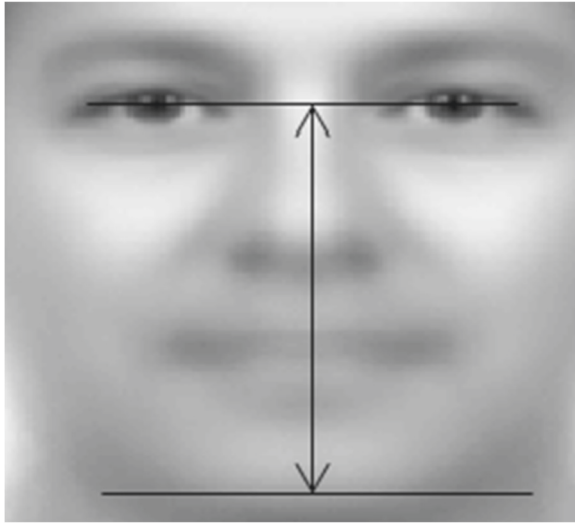
### 6.1 Canonical correlation analysis

Canonical correlation analysis (CCA) is suited to put in correspondence two sets of measurements. CCA takes advantage of the correlations between the response variables to improve predictive accuracy [40]. Given  $N$  pairs of samples  $(x_i, y_i)$  of  $(X, Y)$ ,  $i = 1, \dots, N$ , where  $X \in \mathbb{R}^m$ ,  $Y \in \mathbb{R}^n$ . The mean of both  $X$  and  $Y$  is zero. The goal of CCA is to learn a pair of directions  $w_x$  and  $w_y$  to maximize the correlation between the two projections  $w_x^T X$  and  $w_y^T Y$ , where  $T$  denotes the transpose, i.e., to maximize:

$$\rho = \frac{E[w_x^T X Y^T w_y]}{\sqrt{E[w_x^T X X^T w_x] E[w_y^T Y Y^T w_y]}} \quad (3)$$



**Fig. 6** Scaling for frontal and profile faces. This step consists to set frontal and profile faces to the same scale. Explicitly, the distance between the eye and the chin in profile face must be equal to the distance between the eyes and the chin in the frontal face



**Fig. 7** Reference distance between the eyes and chin used for profile rescaling. The reference distance between the eyes and chin calculated on the average face of the normalized frontal learning database, used for rescaling profile faces in practical identification situations

where  $E[f(x, y)]$  denotes the empirical expectation of the function. The covariance matrix of  $(X, Y)$  is

$$\begin{aligned} C(X, Y) &= E\left(\begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}^T\right) \\ &= E\left(\begin{pmatrix} C_{xx} \\ C_{xy} \end{pmatrix} \begin{pmatrix} C_{yx} \\ C_{yy} \end{pmatrix}^T\right) \end{aligned} \quad (4)$$

where  $C_{xx}$  and  $C_{yy}$  are within-sets covariance matrices;  $C_{xy}$  and  $C_{yx}$  are between-sets covariance matrices. Hence,  $\rho$  can be rewritten as:

$$\rho = \frac{w_x^T C_{xy} w_y}{\sqrt{w_x^T C_{xx} w_x C_{yy} w_y}} \quad (5)$$

Let :

$$A = \begin{pmatrix} 0 & C_{xy} \\ C_{xy} & 0 \end{pmatrix}, B = \begin{pmatrix} C_{xx} & 0 \\ 0 & C_{yy} \end{pmatrix} \quad (6)$$

It can be shown that the solution  $W = (w_x^T, w_y^T)$  amounts to the extremum points of the Rayleigh quotient [41]:

$$r = \frac{W^T A W}{W^T B W} \quad (7)$$

The solution  $w_x$  and  $w_y$  can be obtained as solutions of the generalized eigen-problem:

$$A W = B W \lambda \quad (8)$$

As a subspace learning method, CCA is inclined to overfit to the training data, especially when the sample size is small [35]. Here, we add a dimensionality reduction step (like PCA) before applying the CCA. To introduce a nonlinear generalization of CCA based on a kernel formulation to take account the nonlinearity of transformation between the profile face image and frontal one, KPCA is used instead of PCA. The transformation of input data (frontal or profile faces) is performed by a mapping from the original input space to a high-dimensional feature space.

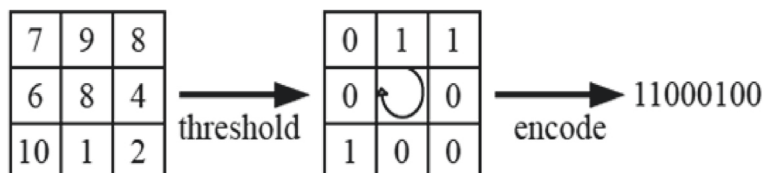
## 6.2 Kernel principal component analysis

Standard PCA only allows linear dimensionality reduction. However, if the data has more complicated structures which cannot be well represented in a linear subspace, standard PCA will not be very helpful. Fortunately, kernel PCA allows us to generalize standard PCA to nonlinear dimensionality reduction [42].

Assume we have a nonlinear transformation  $\phi(x)$  from the original  $D$ -dimensional feature space to an  $M$ -dimensional feature space, where usually  $M \gg D$ . Then each data point  $x_i$  in dataset  $\{x_i\}$ , where  $i = 1, 2, \dots, N$ , is projected to a point  $\phi(x_i)$ . We can perform standard PCA in the new feature space, but this can be extremely costly and inefficient. To simplify the computation, kernel methods can be used [43].

First, we assume that the projected new features have zero mean:

$$\frac{1}{N} \sum_{i=1}^N \phi(x_i) = 0 \quad (9)$$



**Fig. 8** LBP operator

The covariance matrix of the projected features is  $M \times M$ , calculated by

$$C = \frac{1}{N} \sum_{i=1}^N \phi(x_i) \phi(x_i)^T \quad (10)$$

Its eigenvalues and eigenvectors are given by

$$Cv_k = \lambda_k v_k \quad (11)$$

where  $k = 1, 2, \dots, M$ . From Eqs. (10) and (11), we have

$$\frac{1}{N} \sum_{i=1}^N \phi(x_i) \left\{ \phi(x_i)^T v_k \right\} = \lambda_k v_k \quad (12)$$

which can be rewritten as

$$v_k = \sum_{i=1}^N a_{ki} \phi(x_i) \quad (13)$$

Now by substituting  $v_k$  in Eqs. (12) with (13), we have

$$\frac{1}{N} \sum_{i=1}^N \phi(x_i) \phi(x_i)^T \sum_{j=1}^N a_{kj} \phi(x_j) = \lambda_k \sum_{i=1}^N a_{ki} \phi(x_i) \quad (14)$$

Left multiplying  $\phi(x_i)^T$  to both sides of the equation above, we get

$$\frac{1}{N} \sum_{i=1}^N k(x_i, x_i) \sum_{j=1}^N a_{kj} k(x_i, x_j) = \lambda_k \sum_{i=1}^N a_{ki} k(x_i, x_i) \quad (15)$$

where

$$k(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (16)$$

We can use the matrix notation

$$k^2 a_k = \lambda_k N k a_k \quad (17)$$

where

$$k_{i,j} = k(x_i, x_j) \quad (18)$$

and  $a_k$  is the  $N$ -dimensional column vector of  $a_{ki}$ :

$$a_k = [a_{k1} \ a_{k2} \ \dots \ a_{kN}]^T \quad (19)$$

$a_k$  can be solved by

$$k a_k = \lambda_k N a_k \quad (20)$$

and the resulting kernel principal components can be calculated using

$$y_k(x) = \phi(x)^T v_k = \sum_{i=1}^N a_{ki} k(x, x_i) \quad (21)$$

The power of kernel methods is that we do not have to compute  $\phi(x_i)$  explicitly. We can directly construct the kernel matrix from the training dataset  $\{x_i\}$  [44]. Two commonly used kernels are the polynomial kernel

$$k(x, y) = (x^T y)^d \quad (22)$$

and the Gaussian kernel

$$k(x, y) = \exp \left( -\frac{\|x - y\|^2}{2\sigma^2} \right) \quad (23)$$

## 7 Experiments

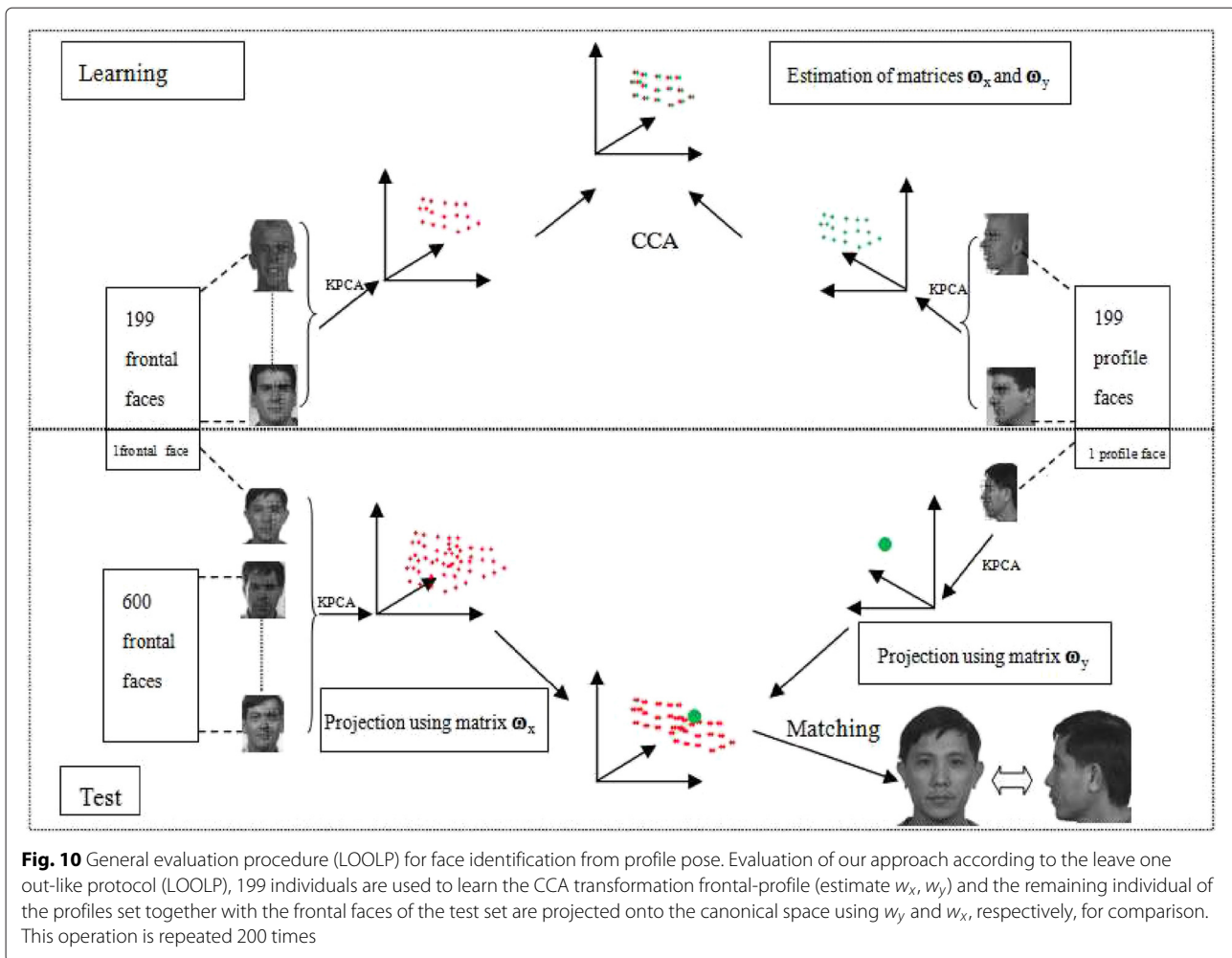
To evaluate our approach according to LOOLP, we test it on FERET database considered as a reference in face identification area. Since it contains faces in different orientations, it is used for pose-invariant identification. It is known to be difficult because the faces with different orientations have not necessarily been acquired during the same time.

We take from fa subset, 800 frontal faces subdivided into 200 for learning and 600 for testing. Profile faces



**Fig. 9** Main regions of the face selected for the CCA mapping. The figure shows the regions describing the faces to be used for the learning of the CCA-based mapping





of the same individuals used for learning (200) are taken from the subset pl. Consequently, we get 200 pairs of faces (frontal and profile) for learning step and 600 frontal (as gallery) to evaluate the approach. We recall that 2D-based pose-invariant identification methods proposed in the literature do not follow a rigorous protocol of testing, as is the case for frontal identification. In fact, each author adopts his own protocol and uses reference datasets (gallery) that contain about 100 individuals. Unlike the previous works, we have augmented the gallery to 600 faces with one image per individual, in order to give more credibility to our tests. For the frontal parts of each face, we manually annotate (click on) the eyes, the nose, the mouth, and the chin and for the profile, we click on the nose, one eye, the mouth, and the chin. These parts are used for geometric normalization of the faces and for feature extraction where 64-bin histograms are calculated for each region of interest in the LBP images, and then histograms of the 14 regions (Fig. 9) of the face are concatenated in a vector to describe it. Hence, we get a

descriptor of size  $p = 896$  ( $64 \times 14$ ) for frontal faces and also  $q = 896$  ( $64 \times 14$ ) for profile faces.

It is well known that a condition for CCA applicability is to have  $(p + q) < 199$  which is the number of faces in both of the frontal and the profile bases in the learning step. This condition is necessary, because if  $(p + q) > 199$ , a slight perturbation in the two bases (frontal and profile) will affect drastically the results of identification. In

**Table 1** Comparison of identification scores provided by CCA based on different kernel types

Kernel	Score (over 200)
Linear	109
Polynomial degree 2	120
Polynomial degree 3	120
Polynomial degree 4	118
Polynomial degree 5	116
Gaussian $\sigma_1 = \sigma_2 = 0.5$	140

**Table 2** Selection of the relevant features

Components (regions)	Score (/200)
Eye, nose, mouth, chin (1, ..., 14)	140
Eye, nose, mouth (1, ..., 12)	90
Eye, nose, chin (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14)	111
Eye, mouth, chin (1, 2, 3, 4, 5, 6, 11, 12, 13, 14)	84
Nose, mouth, chin (7, 8, 9, 10, 11, 12, 13, 14)	54
Part of eye, nose, mouth, chin (2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14)	150
Eye, left part of nose, , mouth, chin (1, 2, 3, 4, 5, 6, 7, 9, 11, 12, 13, 14)	112
Eye, right part of nose, mouth, chin (1, 2, 3, 4, 5, 6, 8, 10, 11, 12, 13, 14)	115
Eye, nose, left part of mouth, chin (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14)	125
Eye, nose, right part of mouth, chin (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14)	127
Eye, nose, mouth, left part of the chin (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13)	119
Eye, nose, mouth, right part of the chin (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14)	126

our work, we conform to this condition by a dimensionality reduction based on the KPCA for both poses with linear, polynomial, or Gaussian kernel. To proceed with the LOOLP, 199 individuals are used to learn the CCA transformation frontal-profile (estimate  $w_x$ ,  $w_y$ ) and the remaining individual of the profiles set together with the frontal faces of the test set are projected onto the canonical space using  $w_y$  and  $w_x$ , respectively, for comparison. Euclidean distance is used to measure the proximity of this

individual to those of the frontal faces. This operation is repeated 200 times (Fig. 10).

Scores of correct identification are summarized in Table 1. It can be noticed that the Gaussian kernel provides the best score with a correct identification rate of 70 % (140/200).

In order to improve the identification score, the Gaussian kernel with  $\sigma = 0.5$  is used for both frontal and profile faces, associated to a feature selection that consists of testing the discriminative power of a subset of the regions of interest. Table 2 gives the list of the retained components for our experiments according to Fig. 9.

We get the best score while we use the components of the face from which we remove regions 1 and 4 of the eye. In order to select the optimal values of the Gaussian kernel parameters  $\sigma_1$  and  $\sigma_2$  for frontal and profile, respectively, we vary them between 0.01 and 1, with a 0.01 step size. The highest score of 153/200 is obtained for  $\sigma_1 = 0.41$  and  $\sigma_2 = 0.39$ .

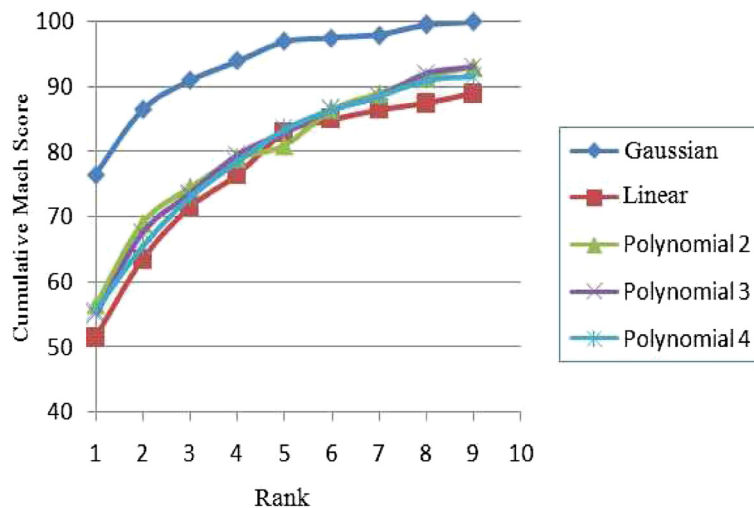
Figure 11 representing the cumulated match scores for different kernels, shows that the Gaussian kernel overclasses the others.

Figure 12 depicts the identification rate versus the selected canonical space dimensionality for a gallery of 200 individuals, and Fig. 13 shows the identification rate versus the size of the gallery used for test.

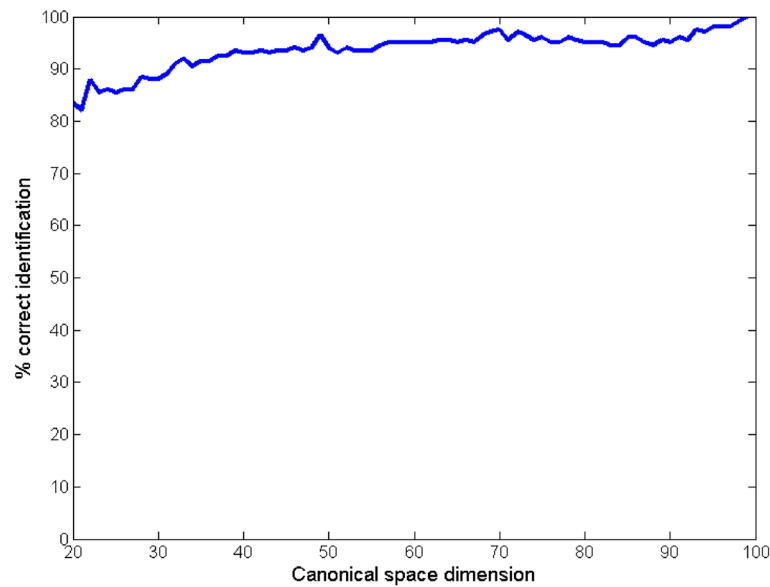
## 8 Discussion

Table 3 sums up the scores obtained by methods proposed in the literature for 90° pose deviation from the gallery of frontal faces.

It can be noticed that galleries used to test these methods contain at most 200 individuals (column 2 of Table 3 gives size in brackets). Indeed, the highest score (92 %



**Fig. 11** Cumulative match curve for profile face identification. The figure shows that the Gaussian kernel overclasses the others

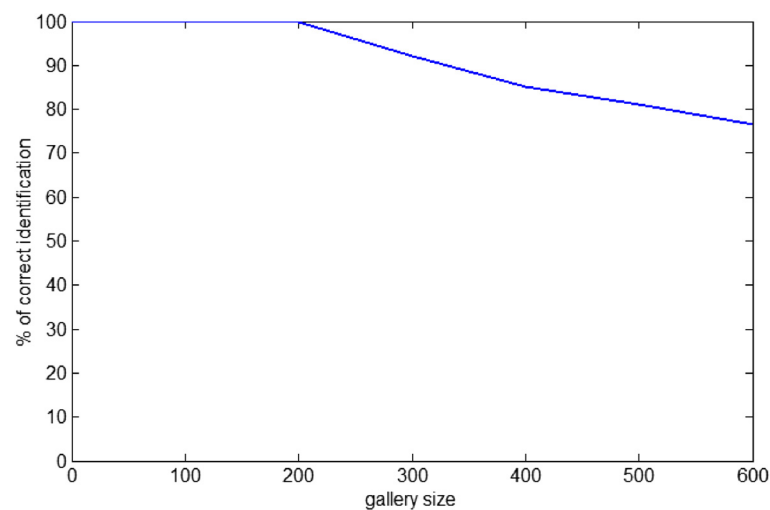


**Fig. 12** Identification rate versus canonical space dimension

of correct identification) is obtained by Prince et al. [21], on a gallery of only 100 individuals. However, the authors underline the fact that if there are more than 100 individuals, there are more people to confuse the probe with, and the task becomes harder [21]. In our study, we obtain 100 % of correct identification even for a gallery of much larger size (200 individuals). The score decreases to 76.5 % for a gallery of 600 individuals. Unlike the approach proposed by Prince et al. [21], which requires 14 landmarks, our technique needs only four landmarks. It is worth noting that though these four landmarks may be detected automatically, the manual annotation is adopted to be

sure that identification errors are inherent to choices of the approach rather than to eventual erroneous landmark detection. Reducing the number of landmarks from 14 to 4, allows not only a gain in execution time but an increase in the identification rate, which is advantageous for a practical use.

In order to confirm the results obtained for FERET database, the approach is also tested on another recent database according to LOOLP. This database is SCface [6] which contains static images of human faces taken in uncontrolled indoor environment using five video surveillance cameras of various qualities. Database contains



**Fig. 13** Identification rate versus the size of the gallery used for test

**Table 3** Comparison of profile face identification studies (90°)

Study	Database	% Correct
Zhang and Samaras [45]	CMU PIE (68)	55
Wallhoff et al. [46]	Mugshot (100)	60
Wallhoff and Rigoll [47]	FERET+ Mugshot (200)	42
Kanade and Yamada [48]	CMU PIE (34)	40
Prince et al. [21]	FERET (100)	92
Prince et al. [21]	XM2VTS (100)	91
Proposed method	FERET (200)	100
Proposed method	FERET (600)	76.5

130 individuals pictured in different orientations. In the present paper, the 0° and 90° orientations are considered.

One hundred pairs of images (frontal and profile) are used to learn the transformation and the remaining 30 frontal images as gallery for test. For feature extraction, the same steps as for FERET database are performed. The condition for CCA applicability is  $(p + q) < 99$  which is the number of faces in both of the frontal and the profile bases of the learning step ( $p$  and  $q$  are the components to retrieve for frontal and profile faces). The best score is 52 %, obtained with a Gaussian kernel. The decrease of the score may be explained by the fact that only 49 components are retrieved for both the frontal and the profile databases. If one considers as good identification, the situation where the individual to be identified ranks among the five first individuals found, the score increases to reach 73 %.

We consider that this score is satisfactory given the fact a reduced number of components has been retrieved.

## 9 Conclusions

In this paper, we proposed a method for identification of profile faces using frontal faces as references. It is based on a KCCA and consists of learning the transformation between frontal and profile faces. Since most of the existing databases of faces contain only frontal faces, and it is often difficult to have simultaneously the frontal and the profile face of any individual, this transformation may have a central role in practice. In fact, if a face is detected in a scene, it will be represented in the frontal space, thanks to the transformation learned by CCA, and then, compared to the frontal faces. Due to the nonlinearity of the transformation between the frontal and the profile pose, we were constrained to not represent the global face but to restrict to its main components (eyes, nose, mouth, chin) and to use kernel functions. The validation of our method has been achieved according to a leave one out-like protocol (LOOLP), based on a gallery of 600 individuals. This way of validation guarantees a more accurate estimate of the generalization error of the approach. For

the transformation between 0° and 90°, a score of 100 % on a 200-face gallery that overpasses those published in the literature, is obtained. This score decreases to 76.5 % for a gallery of 600 faces. In a future work, we will generalize the method to other orientations (other than 90°) in order to get a totally pose-invariant face identification system.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Ecole Militaire Polytechnique, Computer Science Department, BP 17, 16111 Algiers, Algeria. <sup>2</sup>Université des Sciences et Technologies Houari Boumedienne, BP 32 EL ALIA, 16111 Algiers, Algeria. <sup>3</sup>TELECOM ParisTech, Signal and Image Processing department, 46, rue Barrault, 75013 Paris, France.

Received: 30 September 2014 Accepted: 21 June 2016

Published online: 08 July 2016

## References

1. JL Wayman, AK Jain, D Maltoni, D Maio, *Biometric systems: technology, design and performance evaluation*. (Springer, London, 2005)
2. AS Georgiades, PN Belhumeur, DJ Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 643–660 (2001)
3. X Zhang, Y Gao, Face recognition across pose: a review. *Pattern Recognit.* **42**(11), 2876–2896 (2009)
4. T Ojala, M Pietikäinen, D Harwood, A comparative study of texture measures with classification based on feature distributions. *Pattern Recognit.* **29**, 51–59 (1996)
5. PJ Phillips, H Wechsler, J Huang, P Rauss, The FERET database and evaluation procedure for face recognition algorithms. *Image Vision Comput.* **16**(5), 295–306 (1998)
6. M Grgic, K Delac, S Grgic, Sface—surveillance cameras face database. *Multimedia Tools Appl. J.* **51**(3), 863–879 (2011)
7. D Beymer, in *Proceedings of the IEEE Conference on CVPR*. Face recognition under varying pose, (Seattle, WA, USA, 1994), pp. 756–761
8. D Beymer, T Poggio, in *Proceedings of the International Conference on Computer Vision*. Face recognition from one example view, (Cambridge, MA, USA, 1995), pp. 500–507
9. T Beier, S Neely, Feature-based image metamorphosis. *Proc. SIGGRAPH 92* (Computer Graphics). **26**, 35–42 (1992)
10. TF Cootes, CJ Taylor, DH Cooper, J Graham, Active shape models-their training and application. *Comp. Vision Image Underst.* **61**(1), 38–59 (1995)
11. D Gonzalez-Jimenez, JL Alba-Castro, Toward pose-invariant 2-d face recognition through point distribution models and facial symmetry. *IEEE Trans. Inf. Forensic Secur.* **2**(3), 413–429 (2007)
12. TF Cootes, GJ Edwards, CJ Taylor, Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 681–685 (2001)
13. D Shan, R Ward, Face recognition under pose variations. *J. Franklin Inst.* **343**(6), 596–613 (2006)
14. T Vetter, Synthesis of novel views from a single face image. *Int. J. Comput. Vision.* **28**(2), 103–116 (1998)
15. F Kahraman, B Kurt, M Gokmen, in *Proceedings of the IEEE Conference on CVPR*. Robust face alignment for illumination and pose invariant face recognition, (Minneapolis, MN, USA, 2007), pp. 1–7
16. C Liu, Gabor-based kernel pca with fractional power polynomial models for face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(5), 572–581 (2004)
17. X Xie, KM Lam, Gabor-based kernel PCA with doubly nonlinear mapping for face recognition with a single face image. *IEEE Trans. Image Process.* **15**(9), 2481–2492 (2006)
18. J Huang, PC Yuen, WS Chen, JH Lai, Choosing parameters of kernel subspace LDA for recognition of face images under pose and illumination variations. *IEEE Trans. Syst. Man Cybern. B Cybern.* **37**(4), 847–862 (2007)
19. J Yang, AF Frangi, J Yang, D Zhang, Z Jin, Kpca plus lda: A complete kernel Fisher discriminant framework for feature extraction and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(2), 230–244 (2005)

20. X Chai, S Shan, X Chen, W Gao, Locally linear regression for pose-invariant face recognition. *IEEE Trans. Image Process.* **16**(7), 1716–1725 (2007)
21. SJD Prince, J Warrell, JH Elder, FM Felisberti, Tied factor analysis for face recognition across large pose differences. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(6), 970–984 (2008)
22. KW Bowyer, C Kyong, P Flynn, A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. *Comput. Vision Image Underst.* **101**(1), 1–15 (2006)
23. A Scheenstra, A Ruifrok, RC Veltkamp, A survey of 3D face recognition methods. *Proc. Int. Conf. Audio Video Based Biom. Pers. Authentication.* **3546**, 891–899 (2005)
24. Y Gao, MKH Leung, W Wang, SC Hui, Fast face identification under varying pose from a single 2-d model view. *IEE Proc. Vision Image Signal Process.* **148**(4), 248–253 (2001)
25. X Liu, T Chen, Pose-robust face recognition using geometry assisted probabilistic modeling. *Proc. IEEE Conf. CVPR.* **1**, 502–509 (2005)
26. MW Lee, S Ranganath, Pose-invariant face recognition using a 3D deformable model. *Pattern Recog.* **36**(8), 1835–1846 (2003)
27. D Jiang, Y Hu, S Yan, L Zhang, H Zhang, W Gao, Efficient 3D reconstruction for face recognition. *Pattern Recog.* **38**(6), 787–798 (2005)
28. V Blanz, T Vetter, in *Proceedings of SIGGRAPH*. A morphable model for the synthesis of 3D faces, (New York, NY, USA, ACM Press, 1999), pp. 187–194
29. V Blanz, T Vetter, Face recognition based on fitting a 3D morphable model. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(9), 1063–1074 (2003)
30. CD Castillo, DW Jacobs, in *Proceedings of the IEEE Conference on CVPR*. Using stereo matching for 2-d face recognition across pose, (Minneapolis, MN, USA, 2007), pp. 1–8
31. A Li, S Shan, X Chen, Y Gao, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition - CVPR*. Maximizing intra-individual correlations for face recognition across pose differences, (Miami, FL, USA, 2009), pp. 605–611
32. T Ahonen, A Hadid, M Pietikainen, Face description with local binary patterns: application to face recognition. *IEEE Trans. Pattern. Anal. Mach. Intell.* **28**(12), 2037–2041 (2006)
33. D Maturana, D Mery, A Soto, in *Proceedings of the XXVIII International Conference of the Chilean Computer Science Society, IEEE CS Society*. Face recognition with local binary patterns, spatial pyramid histograms and naive Bayes nearest neighbor classification, (Santiago, TBD, Chile, 2009), pp. 125–132
34. D Huang, C Shan, M Ardabilian, Y Wang, L Chen, Local binary patterns and its application to facial image analysis: a survey. *IEEE Trans. Syst. Man Cybern. Appl. Rev.* **41**(6) (2011)
35. M Reiter, Enhanced multiple output regression based on canonical correlation analysis with applications in computer vision. PhD thesis. (Graz University of Technology, Graz Austria, 2010)
36. T Melzer, M Reiter, H Bischof, Appearance models based on kernel canonical correlation analysis. *Pattern Recog.* **36**(9), 1961–1971 (2003)
37. S Arlot, A Celisse, A survey of cross-validation procedures for model selection. *Stat. Surv.* **4**, 40–79 (2010)
38. G Gordon, in *Proceedings of the International Workshop on Face and Gesture Recognition*. Face recognition from frontal and profile views, (Zurich, Switzerland, 1995), pp. 47–52
39. G Pan, L Zheng, Z Wu, in *7th IEEE Workshop on Applications of Computer Vision*. Robust metric and alignment for profile-based face recognition: an experimental comparison, (Breckenridge, CO, USA, 2005), pp. 117–122
40. L Breiman, JH Friedman, Predicting multivariate responses in multiple linear regression. *J. R. Statist. Soc.* **59**(1), 3–54 (1997)
41. M Borga, Learning multidimensional signal processing. linköping studies in science and technology, dissertations. PhD thesis. (Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1998)
42. B Schölkopf, A Smola, K Müller, in *Advances in Kernel Methods, Support Vector Learning*. Kernel principal component analysis, (MIT Press, Cambridge, MA, USA, 1999), pp. 327–352
43. B Schölkopf, A Smola, K Müller, Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.* **10**, 1299–1319 (1998)
44. KQ Weinberger, F Sha, LK Saul, in *Proceedings of the 21st International Conference on Machine Learning*. Learning a kernel matrix for nonlinear dimensionality reduction, (ACM, New York, NY, USA, 2004), pp. 839–846
45. L Zhang, D Samaras, in *Proceedings of ECCV Int'l Workshop Biometric Authentication Workshop*. Pose invariant face recognition under arbitrary unknown lighting using spherical harmonics, (Prague, Czech Republic, 2004)
46. F Wallhoff, S Muller, G Rigoll, in *Proceedings of Second IEEE ICCV Workshop Recognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems*. Hybrid face recognition systems for profile views using the mugshot database, (2001), pp. 149–156
47. F Wallhoff, G Rigoll, in *Proceedings on 8th Intern. Fall Workshop Vision Modelling and Visualization, VMV 2003, Munich, Germany*. Synthesis and recognition of face profiles, (2003)
48. T Kanade, A Yamada, in *Proceedings of IEEE Int'l Symp. Computational Intelligence in Robotics and Automation*, vol. 2. Multi-subregion-based probabilistic approach toward pose-invariant face recognition, (2003), pp. 954–959

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)