We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



10

2D-3D Mixed Face Recognition Schemes

Antonio Rama Calvo[1], Francesc Tarrés Ruiz[1],
Jürgen Rurainsky[2] and Peter Eisert[2]
[1]*Department of Signal Theory and Communications*
Universitat Politècnica de Catalunya (UPC)
[2]*Image Processing Department*
Fraunhofer Institute for Telecommunications Heinrich-Hertz-Institut (HHI)
[1]*Spain*
[2]*Germany*

1.Introduction

Automatic recognition of people is a challenging problem which has received much attention during the recent years [FRHomepage, AFGR, AVBPA] due to its potential applications in different fields such as law enforcement, security applications or video indexing. Face recognition is a very challenging problem and up to date, there is no technique that provides a robust solution to all situations and different applications that face recognition may encounter.

Most of the face recognition techniques have evolved in order to overcome two main challenges: illumination and pose variation [FRVT02, FRGC05, Zhao03, Zhao06]. Either of these problems can cause serious performance degradation in a face recognition system. Illumination can change the appearance of an object drastically, and in the most of the cases these differences induced by illumination are larger than differences between individuals, what makes difficult the recognition task. The same statement is valid for pose variation. Usually, the training data used by face recognition systems are frontal view face images of individuals [Brunelli93, Nefian96, Turk91, Pentland94, Lorente99, Belhumeur97, Bartlett02, Moghaddam02, Delac05, Kim02, Schölkopf98, Schölkopf99, Yang02, Yang04, , Wang06, Yu06, Heo06]. Frontal view images contain more specific information of a face than profile or other pose angle images. The problem appears when the system has to recognize a rotated face using this frontal view training data. Furthermore, the appearance of a face can also change drastically if the illumination conditions vary [Moses94]. Therefore, pose and illumination (among other challenges) are the main causes for the degradation of 2D face recognition algorithms.

Some of the new face recognition strategies tend to overcome both challenges from a 3D perspective. The 3D data points corresponding to the surface of the face may be acquired using different alternatives: a multi camera system (stereoscopy) [Onofrio04, Pedersini99, structured light [Scharstein02, 3DRMA], range cameras or 3D laser and scanner devices [Blanz03, Bowyer04, Bronstein05]. The main advantage of using 3D data is that depth information does not depend on pose and illumination and therefore the representation of the object do not change with these parameters, making the whole system more robust.

Open Access Database www.intechweb.org

Source: Recent Advances in Face Recognition, Book edited by: Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett,
ISBN 978-953-7619-34-3, pp. 236, December 2008, I-Tech, Vienna, Austria

However, the main drawback of the majority of 3D face recognition approaches is that they need all the elements of the system to be well calibrated and synchronized to acquire accurate 3D data (texture and depth maps). Moreover, most of them also require the cooperation or collaboration of the subject making them not useful for uncontrolled or semi-controlled scenarios where the only input of the algorithms will be a 2D intensity image acquired from a single camera.

All these requirements can be available during the training stage of many applications. When enrolling a new person in the database, it could be performed off-line, with the help o human interaction and with the cooperation of the subject to be enrolled. On the contrary, the previous conditions are not always available during the test stage. The recognition will be in most of the cases in a semi-controlled or uncontrolled scenario, where the only input to the system will probably consist of a 2D intensity image acquired from a single camera. One possible example of these application scenarios are video surveillance or control access. This leads to a new paradigm using some mixed 2D-3D face recognition systems where 3D data is used in the training but either 2D or 3D information can be used in the recognition depending on the scenario. Following this concept, where only part of the information (partial concept) is used in the recognition, a novel method is presented in this work. This has been called Partial Principal Component Analysis (P^2CA) since it fuses the partial concept with the fundamentals of the well known PCA algorithms. Both strategies have been proven to be very robust in pose variation scenarios showing that the 3D training process retains all the spatial information of the face while the 2D picture effectively recovers the face information from the available data. Simulation results have shown recognition rates above 91% when using face images with a view range of 180^o around the human face in the training stage and 2D face pictures taken from different angles (from -90^o to $+90^o$) in the recognition stage.

2.State-of-the art face recognition methods

2.1 2D face recognition

The problem of still face recognition can be simply stated as: given a set of face images labelled with the person identity (learning set) and an unlabeled set of face images from the same group of people (the test set), identify each person in the test images. This problem statement is also known as person identification.

Different schemes and strategies have been proposed for the problem of face recognition. The categorization of these approaches is not easy and different criteria are usually used in literature. One popular classification scheme is the attending of the holistic/non-holistic philosophy of the methods[Zhao03, Zhao06]. Holistic methods try to recognize the face in the image using overall information, that is, the face as a whole. These methods are commonly referred as appearance-based approaches. On the contrary, non-holistic approaches are based on identifying particular features of the face such as the nose, the mouth, the eyes, etc. and their relations to make the final decision. Some recent methods try to exploit the advantages of both approaches at the same time and therefore they are classified as hybrid. The number of 2D face recognition algorithms is immense and they enclose a huge variety approaches so it would be impossible to make an exhaustive enumeration of all publications related with 2D face recognition.

Kanade's face identification system [Kanade73] was the first automated system to use a top-down control strategy directed by a generic model of expected feature characteristics of the

face. Later Brunelli et al. and Nefian have used correlation-based approaches [Brunelli93, Nefian96]. In this kind of methods, the image is represented as a bidimensional array of intensity values (I_T) and is compared with a single template (T) that represents the whole face. Nevertheless, one important date for face recognition was beginning of the 90's when Turk and Pendland implemented the Eigenfaces approach [Turk91, Pentland94], which is surely the most popular face recognition method. This was the beginning of the appearance-based methods for face recognition. After Eigenfaces, different statistical approaches have appeared that improve the results of Eigenfaces under certain constraints. For example, the most representative ones are Fisherfaces which is more robust towards illumination conditions [Belhumeur97], Kernel PCA [Kim02] and Independent Component Analysis [Bartlett98] which exploit higher-order statistics, or a recent two dimensional extension of the PCA [Yang04].

Another strategy that has been used to solve face recognition is neural networks [Kohonen88, Fadzil94, Lawrence97, Lin97, Haddadnia02 Palanivel03]. Neural networks approaches promise good performance but these have to be further improved and investigated mainly because of the difficulty of the training the system. One method that intends to solve the conceptual problems of conventional artificial neural networks,, and should be also mentioned, is Elastic Graph Matching [Lades93, Wiskott99]. Face recognition using Elastic Graph Matching (EGM) [Lades93] is based on the neural information processing concept, the Dynamic Link Architecture (DLA). In EGM, a face is represented by a set of feature vectors (Gabor responses) positioned on nodes of a rectangular grid placed on the image. Comparing two faces corresponds to matching and adapting a grid taken from one image to the features of the other image. Rotation in depth is compensated for by elastic deformation of the graphs.

Although all methods report encouraging and excellent results, the real fact is that approaches based on statistical appearance-based methods like Principal Component Analysis (PCA) and Elastic Graph Matching [Lades93, Wiskott99] are the algorithms which present the best face recognition rates [Zhang97]. For a more detailed survey about all the different methods, the author is addressed to the work of Zhao et al. [Zhao03 Zhao06].

2.2 3D face recognition approaches

The 3D structure of the human face intuitively provides high discriminatory information and is less sensitive to variations in environmental conditions like illumination or view-point. For this reason, recent techniques have been proposed employing range images, i.e. 3D data in order to overcome the main challenges of 2D face recognition: Pose and illumination. Next section a review of the most relevant approaches in 3D face recognition will be presented. In fact, the section has been divided into three main groups: Curvature-based algorithms, multimodal approaches and finally model-based methods.

2.2.1 Early work in 3D face recognition: curvature-based approaches

3D Face Recognition is a relatively recent trend although early work was done over a decade ago [Cartoux89, Lee90, Gordon91]. The first contributions to the field [Cartoux89, Lee90, Gordon91] were mainly based on the extraction of a curvature representation of the face from ranging images. Then a set of features were obtained or created from these curvature face representations and used to match the different faces. For example, in [Cartoux89], a profile curve is computed from the intersection of the face range image and the profile

plane. This profile plane is defined as the one that segments the face in two quasi-symmetric parts. In order to find it, Cartoux et al. propose an iterative method where the correspondence between points of the convex-concave representation of the face is analyzed. Very high recognition rates were reported (100%) for experiments carried out on a very small database of 18 face images which correspond to 5 persons. Another curvature-based method is the one presented in [Lee90]. An Extended Gaussian Image (EGI) is created from the parameterization to the Gaussian Sphere of the curvature maps (in fact, only the points which represent the convex regions of the face). To establish the correspondence between two different EGIs a region a graph-matching approach is applied. This graph-matching algorithm incorporates relational constraints that account for the relative spatial locations of the convex regions in the domain of the range image. Similar to the previous approach Gordon et al. [Gordon91] acquire range data from a laser scanner and parameterize it into a cylindrical coordinate representation. Afterwards, the *principal curvature* maps for this range data are computed and used for the segmentation of the face range image (four different regions: concave, convex and two types of saddle). Two different ways of matching the faces were proposed: The first one is a depth template-based approach, and the second one is a comparison between feature vectors composed of some fiducial points and their relationships. A recognition rate of 100% is reported for a database of 26 individuals and a probe set of 8 faces under 3 different views. Tanaka et al. [Tanaka98] extended the work of Lee [Lee90] and proposed a spherical correlation of the EGIs. Experiments are reported with a 100% recognition rate utilizing a set of 37 range images from the National Research Council of Canada database.

A more recent curvature-based approach [Feng06] extracts two sets of facial curves from a face range image. The authors present a novel facial feature representation, the affine integral invariant that mitigates the effect of pose on the facial curves. The authors claim that a human face can be characterized by 12 affine invariant curves, which are located near the face center profile, center and corner of eye regions. Each curve is projected onto a 8 dimensional space to construct a feature vector with a resulting performance by a 3-NN classifier of 92.57% recognition accuracy for a database of 175 face images.

The main drawback of these methods is that there are only few fiducial points (features) that can be reliably extracted from a 2D face image and would be insensitive to illumination, pose, and expression variations.

A recent and successful method which cannot be classified in this curvature-based category but that intends to perform recognition from geodesic distances between points of the face is the one presented by the Bronstein twins [Bronstein05]. The authors focused their research on a very robust system towards facial expression variations. Under the assumption that the facial skin does not stretch significantly, facial expressions can be modeled as *isometries*, i.e. a transformation that bends the surface such that the object (face) does not "feel" it. In other words, the main idea is to find a transformation which maps the geodesic distances between a certain numbers of sample points on the facial surface to Euclidean distances. This transformation is called the *canonical form*. The authors reported 100% recognition rate for a database of 30 subjects with big variations in facial expression even though two of the 30 subjects are the Bronstein twins.

A similar approach based also in geodesic distances detects surface creases, ravines and ridges [Antini06] which are less sensitive then the fiducial points needed for extracting the curves of previous approaches [Cartoux89, Lee90, Gordon91, Tanaka98, Feng06]. These

surface variations provide important information about the shape of 3D objects. Intuitively, these salient traits can be defined as those curves on a surface where the surface bends sharply. Then, a theory for modeling spatial relationships between curves in the 3D domain has been developed. Finally, a graph matching solution is proposed for the comparison between the spatial relationships computed on curves extracted from a template model and those of reference models. The results presented are worse than the ones shown by the Bronstein twins [Bronstein05] but it should be also mentioned that the database used for the experiments is composed of the double of subjects (61 persons).

2.2.2 2D+3D multimodal approaches

A second type of 3D Face Recognition approaches could be the so called multimodal algorithms. Basically, the general idea is to apply conventional statistical *appearance-based* methods (like PCA, LDA, ICA…) not only to texture but also to depth images [Tsalakanidou04, Chang05, Samani06]. It could be foreseen as two different face recognition experts (one for each modality) whose opinions are combined in a final stage in order to claim the identity of the person. The advantage of this category of 3D FR methods is that it adds depth information to conventional approaches without increasing too much the computational cost. These multimodal methods report generally better recognition rates for texture than for depth information when performing the recognition separately. Nevertheless, in all the cases an improvement of the recognition is reported when using both modalities together.

Tsalakanidou et al. [Tsalakanidou04] report on multimodal face recognition using color and 3D images. The input data of the system is a color frontal view face image with its corresponding frontal depth map (range image). The recognition algorithm is based on the application of PCA to the different color planes and to the range image individually. Experiments are carried out on a subset of 40 persons from the XM2VTS database. Again the best reported results show a 99% of accuracy for the multimodal algorithm which clearly outperforms the recognition rates of each modality (range and color) alone.

A similar, but most extended, 2D+3D proposal is presented by Chang et al [Chang05]. The study involves 198 persons in the gallery and either 198 or 670 time-lapse probe images. PCA-based methods are used separately for each modality and match scores in the separate face spaces are combined for multimodal recognition. The authors conclude that 2D and 3D have similar recognition performance when considered individually and that a simple weighting scheme combination of both modalities outperforms either 2D or 3D alone. Chang et al compare also the multimodal scheme with a 2D multi-image proposal, showing that again is better to use both modalities together rather than using multiple images of the same modality.

A more recent approach [Samani06] fuses both modalities before applying the statistical method. A range and an intensity image are obtained using two digital cameras in a stereo setup and both images are rearranged in a higher dimensional vector representation called *composite image*. The authors reported better results when using this combined depth-texture representation and then applying PCA, rather than applying PCA to each modality alone.

2.2.3 3D model-based approaches

One main drawback of all the methods reviewed so far (curvature-based and multimodal approaches) is that the input of the recognition stage of these approaches should maintain the same data format as the training images, i.e. if *frontal views* have been used during the

training stage then a depth and/or intensity *frontal image* is required in the recognition stage [Chang05]. Opposite to those two previous 3D face recognition categories, 3D model based approaches train the system with 3D data but then they perform the recognition using only one single intensity image. This last category can be enclosed in a 2D-3D mixed face recognition framework.

3D model-based approaches use complete morphable 3D models of a face to perform the recognition [Beymer95, Georghiades01, Ansari03, Blanz03, Lu05, Lu06]. These approaches build a 3D model (3D mesh) using some 3D scans by means of special 3D devices, structured light, or multiple camera in the training stage. Then the input data for the recognition stage is one simple 2D intensity image which can correspond to any pose of a face. The model-based approach try to fit this image to the 3D face model (a generic one of one for each person on the database) and then, it tries to recognize the person. In fact, fitting the 3D morphable model to test images can be used in two ways for recognition across different viewing conditions as stated in [Blanz03]:

Paradigm 1. After fitting the model, recognition can be based on model coefficients, which represent intrinsic shape and texture of faces, and are independent of imaging conditions. For identification, all gallery images are analyzed by the fitting algorithm, and the shape and texture coefficients are stored. Given a test image, the fitting algorithm computes the again the coefficients which are compared with all gallery data in order to find the nearest neighbor [Ansari03, Blanz03]

Paradigm 2. Three-dimension face reconstruction can be employed to generate synthetic views from gallery or probe images [Beymer95, Georghiades01, Lu05, Lu06]. The synthetic views are then transferred to a second viewpoint-dependent recognition algorithm, i.e. using some of the 2D Face Recognition methods presented in Section 2.1.

The most representative model-based approach using Paradigm 1 is the outstanding method presented by Blanz and Vetter [Blanz03]. The authors use an analysis-by-synthesis technique. First, a generic 3D morphable model is learned and created from a set of textured 3D scans (depth and texture maps) of heads. In fact, the morphable model is based on a vector space representation of faces. The authors, align 200 textured 3D scans of different persons using an optical-flow algorithm to a reference scan. Then using the 200 aligned textured 3D scans they apply PCA to shape and texture separately in order to construct the 3D morphable model. The fitting process of the model to an image optimizes shape coefficients, texture coefficients and 22 additional rendering parameters (camera parameters, illumination, viewpoint, etc), i.e. given an input image the fitting procedure minimizes a cost function that takes into account all these parameters. The goal of this fitting procedure can be defined as an analysis-by-synthesis loop, that tries to find the model and scene parameters such that the model, rendered by computer graphics algorithms, produces an image as similar as possible to the input image. The authors presented recognition rates on large databases (FERET and CMU-PIE) above the 95% and stated that their algorithm has been evaluated with 10 more face recognition systems in the Face Recognition Vendor Test 2002 [FRVT02] obtaining better results for non-frontal face images for 9 out of 10 systems.

On the other hand, one of the most recent examples of the second paradigm is the work presented by Lu et al. [Lu05, Lu06]. For each subject, a 3D face model is constructed by integrating several 2.5D face scans which are captured from different views. The authors considered a 2.5D surface as a simplified version of the 3D (x, y, z) surface representation that contains at most one depth value (z direction) for every point in the (x, y) plane. Each

3D model has also its associated texture map. The recognition stage consists of two components: Namely, surface matching and appearance-based matching. The surface matching component is based on a modified Iterative Closest Point (ICP) algorithm. This surface matching returns a small candidate list from the whole gallery that is used for appearance matching. This combination of surface and appearance matching reduces the complexity of the system. Three-dimensional models in the gallery are used to synthesize new appearance samples with pose and illumination variations and the synthesized face images are used in discriminant subspace analysis. Experimental results are given for matching a database of 200 3D face models with 598 2.5D independent test scans acquired under different pose and some lighting and expression changes.
The main drawback of 3D model-based approaches, independent of what paradigm do they use, is the high computational burden of the fitting procedure of the intensity image to the 3D generic model in an accurate manner that do not degrade recognition results. Thus, section 0 will introduce a novel 2D-3D face recognition framework which can be foreseen as a gap between pure 2D and pure 3D face recognition methods.
For a more detailed survey about all the different 3D Face Recognition methods, the author is addressed to the work. [Bowyer04, Chang05, Zhao03, Zhao06].

3. Partial principal component analysis: a 2D-3D mixed face recognition scheme

3.1.1 Fundamentals of the 2DPCA method

Like in the majority of face recognition methods, in 2DPCA [Yang04] the dimensionality of the face images is reduced through the projection into a set of M optimal vectors which composed the so called *feature space* or *face space*. The vectors representing the i^{h} individual are obtained as:

$$\mathbf{r}_k^{\,i} = (\mathbf{A}_i - \mu) \cdot \mathbf{v}_k \qquad k = 1,..,M \tag{1}$$

,where A_i is the mxn texture image representing individual i , μ is the *mean* image of the training set, and v_k are the M optimal projection vectors that maximize the energy of the projected vectors r_k averaged through the whole database. These vectors could be interpreted as unique signatures that identify each person. The projection described in Equation (1) is depicted in Fig. 1. Note that each vector r_k^i has m components where m is the dimension of the matrix A_i in the vertical direction (height of the face image).

$$\begin{bmatrix} r_{11} \\ r_{21} \\ \vdots \\ r_{m1} \end{bmatrix} = \quad \begin{bmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{n1} \end{bmatrix}$$

Fig. 1. Representation of a face in 2DPCA using only one vector (v_1)

The set of ortonormal vectors which maximize the projection of Equation (1) may be obtained as the solution to the following optimization problem: Find v_k for k=1,..,M such that $\xi = \sum_k \sum_l (r_k^l)^T \cdot r_k^l$ is maximum, where r_k^l is defined as the projection of image l through the vector v_k and l accounts for the number of images in the training set. The function to be maximized may be expressed as:

$$\xi = \sum_k \sum_l \left((A_l - \mu)\cdot v_k\right)^T \cdot \left((A_l - \mu)\cdot v_k\right) = \sum_k v_k^T \left(\sum_l (A_l - \mu)^T \cdot (A_l - \mu) \right) \cdot v_k \quad (2)$$

,which states that the optimum projection vectors may be obtained as the eigenvectors associated to the *M l*argest eigenvalues of the *n*x*n* non-negative definite Covariance matrix C_s

$$C_s = \sum_l (A_l - \mu)^T \cdot (A_l - \mu) \quad (3)$$

Therefore, a total of *M* feature vectors are available, with *n* (width of the face image) components each as depicted in Fig. 1. The image has been compressed to a total of *m*x*M* scalars with *M* always being smaller than *n*.

After computing the *face space*, at least one signature ($\mathbf{r}_k^i$) will be extracted for each person of the database by projecting one (or more) representative image (usually one frontal face image with a neutral expression) in the face space. During the recognition stage, when a new image is input to the system, the mean image is subtracted ant the result is projected into the face space resulting in a new signature r_k with k=1,…,M. The best match is found for the identity *i* that minimizes the Euclidean distance:

$$\min_i \left\{ \xi_k = \sum_{k=1}^{M} \sum_{l=1}^{n} \left(r_k(l) - r_k^{\,i}(l) \right)^2 \right\} \qquad i = 1,..,L \quad (4)$$

where L represents the number of individuals in the database.

The procedure is quite different from conventional PCA, since in 2DPCA the whole image is represented as a 2D matrix instead of a 1D vector like in PCA. Certainly, in PCA a scalar number is obtained when the vector image is projected to one eigenvector, whereas in 2DPCA, an *m*-dimensional vector ($\mathbf{r}_k$) is obtained, when the image (in matrix form) is projected to an eigenvector. It can seem that the 2DPCA approach demands more computational cost because it uses vectors instead of numbers to represent the projections. However, the number of eigenvectors {$\mathbf{v}_k$} needed in 2DPCA for an accurate representation is much lower than in PCA [Yang04].

The authors report extensive simulations of these algorithms in different data sets that include images of individuals under different expressions and taken in different dates (AR database) and compare the performance of the method with Fisherfaces, ICA, Eigenfaces and Kernel Eigenfaces. In most simulations 2DPCA has shown better or at least equal performance than the other alternatives. Probably, the main reason for the improvement is that the covariance matrix is of lower order and therefore can be better estimated using a reduced number of samples. Another inherent advantage of the procedure is the computational time for the feature extraction and the computation of eigenvectors is significantly below the PCA.

3.2 A novel mixed 2D-3D FR scheme: Partial Principal Component Analysis (P^2CA)

As already mentioned, the main objective is to implement a face recognition framework which takes advantage of 3D data in the training stage but then use either 2D or 3D in the recognition stage.

The most used definition of the term "3D face data" could be the data that represents the three-dimensional shape of a face from all the possible view angles with a texture map overlaid on this 3D depth map [Zhao06]. In this work, however, the main objective was to show the validity of the method; thus, we will refer to the multi-view texture maps (180º cylindrical texture representations) as the "3D data". Nevertheless, all the fundamentals and experiments explained below and in the following sections are also valid when using "complete 3D data" (depth and texture maps) as shown in [Onofrio06].

The 3D face information required in the training stage could be obtained by means of 3D laser scanners, by structured light reconstruction systems, by multicamera camera sets using stereoscopic techniques, or by simple morphing techniques. A simple approach based on semiautomatic morphing of 2D pictures taken from different view has been followed in this work for its simplicity since our main objective, as already stated, is a mixed 2D-3D face recognition framework and not an accurate 3D face reconstruction system.

Let us present the main idea of P^2CA which is based on the fundamentals of 2DPCA [Yang04] explained in the previous section. Now, let us take the *mxn* 180º texture map of the subject *i* and rotate it 90º as shown in Fig 2. Now we can reformulate Equation (1) by changing A_i with A_i^T so that the transposed texture image is represented by the M vectors $\mathbf{r}_k^i$. Now, unlike in the previous section, each vector $\mathbf{r}_k^i$ has *m* components where *m* is the dimension of the matrix A_i in the horizontal direction (width of the original texture map).

$$\mathbf{r}_k^{\,i} = \mathbf{A}_i^T \cdot \mathbf{v}_k \qquad k = 1,..,M \tag{5}$$

Note also, that the mean is not subtracted when computing the signature of individual *i*. This fact does not represent a problem and will lead only to the necessity of using one more eigenvector as can be mathematically demonstrated.

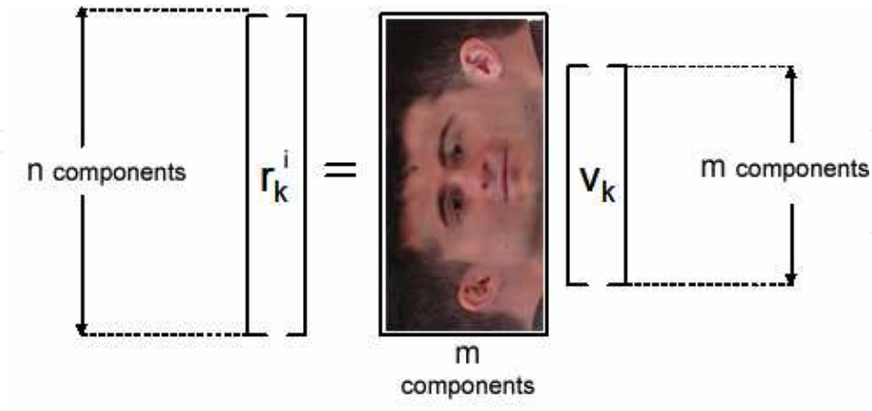


Fig. 2. Description of a cylindrical coordinate image by means of projection vectors (training stage)

During the recognition procedure if complete 3D data of the individual is available, the recognition stage is straightforward. In fact, it is only necessary to convert the 3D data to cylindrical coordinates (texture maps) and compute the resulting M vectors $\mathbf{r}_k$. The best

match is found for the identity i that minimizes the Euclidean distance formulated in Equation (4).

The main advantage of this representation scheme is that it can also be used when only partial information of the individual is available. Consider, for instance, the situation depicted in Fig 3, where it is supposed that only one 2D picture of the individual is available. Each of the 2D pictures of the subject (frontal and lateral) show a high correlation with the corresponding area of the cylindrical representation of the 3D image. The procedure for extracting the information of a 2D picture is illustrated in Fig 4.

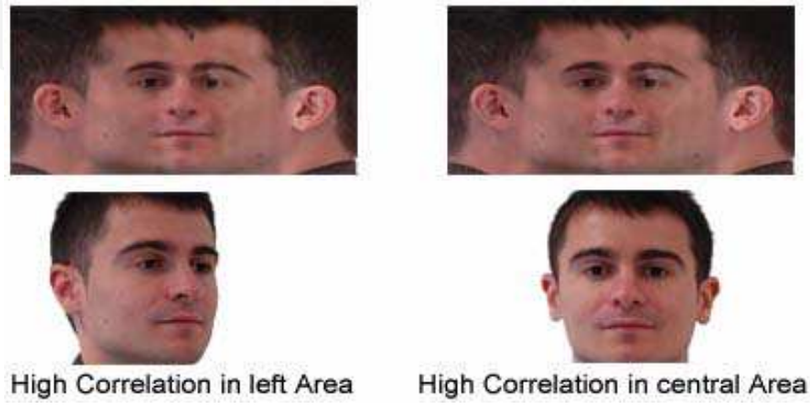


Fig. 3. Comparing 2D pictures with a cylindrical representation of the subject

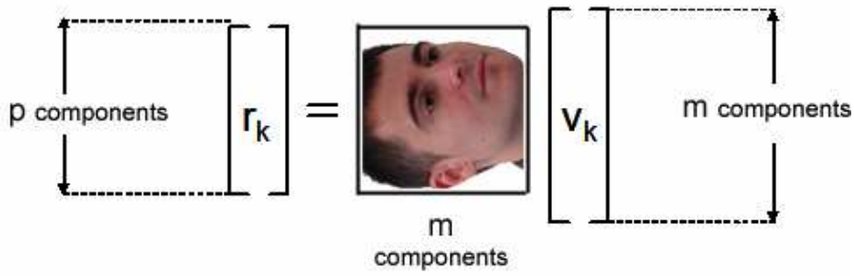


Fig. 4. Projection of a partial 2D picture through the vector set v_k (recognition stage)

In this case, mxp 2D picture will be represented by M vectors r_k with a reduced dimension p. However, it is expected that these p components will be highly correlated with a section of p components in the complete vectors r_k^i computed during the training stage. Therefore, the measure proposed below can be used to identify the partial available information (p components) through the vectors r_k^i:

$$\min_{(i,j)} \left\{ \sum_{k=1}^{M} \sum_{l=1}^{p} \left(r_k(l) - r_k^{\,i}(l+j) \right)^2 \right\} \qquad (6)$$
$$i = 1,..,L; \quad j = 0,..,n-p$$

The most outstanding point of P²CA is that the image projected in the n-dimensional space does not need to have dimension mxn (3D data) during the recognition stage so that only

partial information (2D data) can be used as illustrated in Fig 4. It is possible to use a reduced *mxp* ($p<n$) image which is projected to a smaller subspace. Fig 5 depicts the general block diagram of the P²CA technique explained above and its flexibility. The whole system is trained with the 3D data (texture maps in the bottom part of Fig 5) and the feature vectors r_k^i are computed for each individual of the database (a total of L individuals). The recognition procedure is illustrated in the top part of Fig 5. If the scenario permits to acquire 3D face images, these are projected to the *face space* obtained *m*-dimensional signatures which can be directly compared with the ones stored in the database using Euclidean Distance. On the other hand, if only one normal 2D face image is available, then the resulting *p*-dimensional signatures are compared with the database ones applying some correlation methods like the criteria defined in Equation (6). For more details about Partial Principal Component Analysis the reader is addressed to the work of Rama et al. [Rama05a, Rama05b, Rama06].

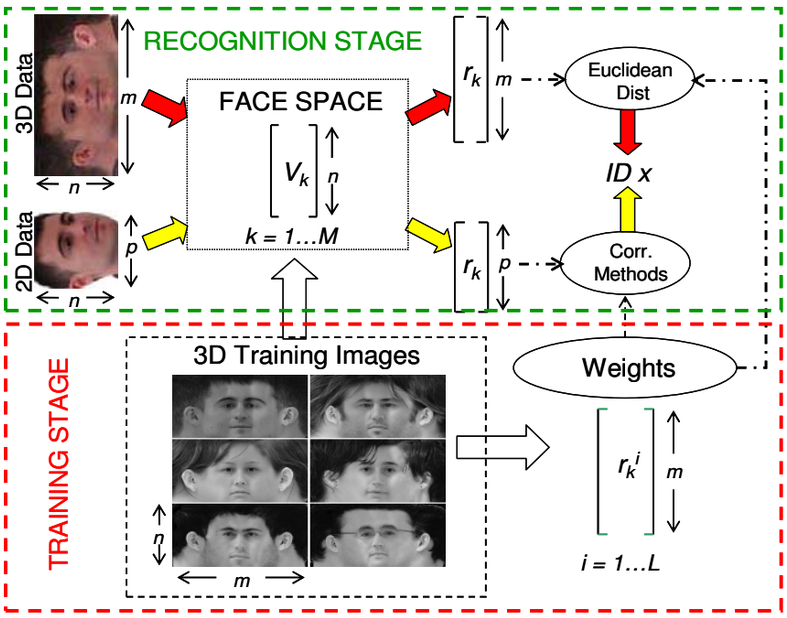


Fig. 5. General Block Diagram of P²CA Approach

4. Aligned texture map creation

4.1 Introduction

Multi view image analysis and synthesis is used for a variety of computer vision and graphics approaches. Combining or stitching of several captured views to one complete virtual representation can be achieved by interpolation between the real captured views. In the articles of Akimoto et al. and Soh et al. [Akimoto93, Soh95] the creation of virtual views for a face is described by using a cylindrical projection onto the real or almost real object surface. A multi resolution edge based approach to align real views are described in the article of Lee et al. [Lee00]. Cylindrical approximations of a human face is used for head motion tracking, like in the publication of Brown, Cascia et al. and in the article of Xiao et al. [Brown01, Cascia98, Xiao93]. Spherical approximation of the head, which is more reliable for head motions, like in the used video sequences is described in the conference publication of

Lui and Chen [Liu03]. A complete mosaicking scheme for the creation of a virtual face image used for face recognition purposes is described in the journal publication of Singh and Ross [Singh07]. A hierarchical registration model is used to align the single views before the stitching and blending via multiresolution splines tales place. Marker points placed on the human face are used for panoramic face mosaicking is described in the journal publication of Yang et al. [Yang]. The markers are removed from the finale virtual view representation.
We present an approach, which aims to use just a cylinder not only for the frontal part but for a complete ±90◦ projection of a face. There image analysis is compensated by the calibration data of the capture setup. The main advantage is the simple approximation of a human head in contrast to a detailed surface representation used for similar results.
For further processing of such face projections additional steps have to be applied, which depend on the purpose of usage. In the case of face recognition input images are aligned before analyzing them, e.g. the eye centers are very often used for this purpose and other face features are usually not received attention. In the conference publication of Kouzani and Sammut [Kouzani99] the advantage of a local face feature analysis is demonstrated. The importance of local aligned face features for recognition is described in the conference publication of Tsapatsoulis et al. [Tsapatsoulis98], where the local alignment is described as resizing. We present an approach for the local alignment of not only the eye center but for any face feature, e.g. eye corners, nose, mouth and chin. This method is based on a feature wire grid model rendered on the graphics card. The result is a face image with face features aligned to a reference, so that a component analysis does not reflects the differences of one face feature but the properties. The novelty of the face feature alignment is the handling of the face feature area by using affine transformation.

4.2 Aligned texture maps

The goal is it to create a cylindrical projection of a face, like the one shown in Fig 6 with a desired view angle range of ±90° using multi view images. Such image can be used as texture map for 3D models or pose invariant face recognition tasks. The creation of texture maps from one or multiple views are described in many publications. The common idea is to project the captured images onto a cylinder using a more or less detailed surface representation or to find registration data, which related one captured view to another. Blending rules at the overlapping areas define the quality of the synthetic image. Our approach is based on a simple approximation of a human head in contrast to a detailed surface representation used for similar results. Therefore a detailed reconstruction of the face is not required. Registration techniques require high frequency parts, which maybe not available for the complete face or hair. Marker points are not very likely, because of the more complicated capture and removal process. On the other hand, ghost edges are very likely, if the cylindrical approximation is not placed at the right position or the system calibration is not adequate.
In addition selected face features are aligned through several created texture maps. The importance of local aligned face features for recognition is described in the conference publication of Tsapatsoulis et al. [Tsapatsoulis98], where the local alignment is described as resizing. For the alignment of face features we propose a method consisting of a global alignment using 2D affine transformation and a local feature alignment based on textured wire grid deformation using graphics card rendering. The local alignment is also using an affine transformation, but this time of the area around the desired face feature. Feature

points are associated to one feature and connected to the used wire grid model, comparable to a 3D alignment. Using such transformed images for a PCA will reflect more the properties of face features and not the differences.



Fig. 6. Texture map created from an image set of nine images captured with

4.2.1 Stitching multiview images

The following paragraphs will explain the creation of cylindrical face projections from several different views. Our goal is to use a very simple approximation instead of a detailed object surface. We have used a cylinder as approximation of a human head. The combination of several views captured with a system as shown in Fig 7 to a synthetic image as shown in Fig 6 requires information about the surface of the captured object. Such

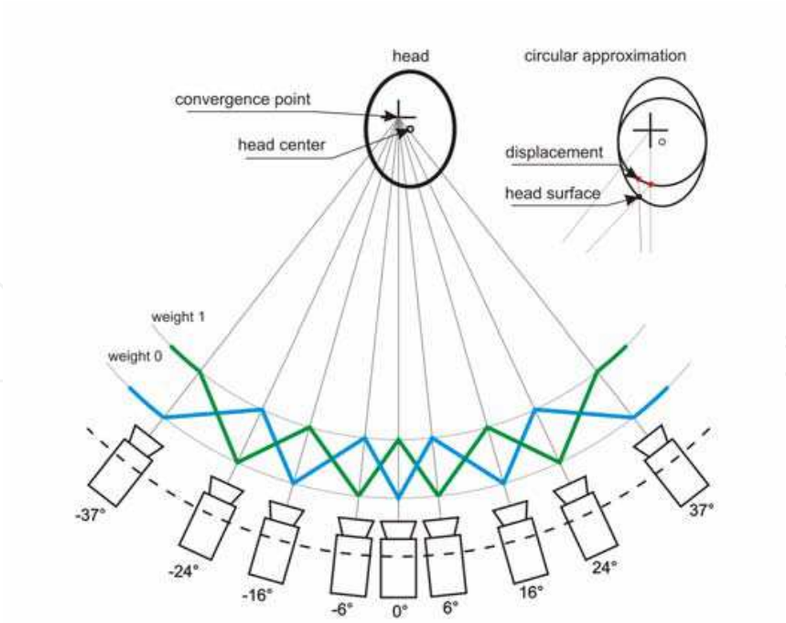


Fig. 7. Capture system for training data acquisition consists of nine cameras at different horizontal angles. (left) face surface point displacement due to the circular approximation. (middle) Blending rule for the combining of the input images

information can range from an exact 3D model over some points in 3D space to an approximation of the basic shape. If such surface information are not available, they can be extracted from the captured view images with a wide variety of methods. Triangulation only requires corresponding point pairs in two views. Surface point reconstruction for a couple of corresponding point pairs allows the definition of an approximation of the desired surface up to a detailed representation. If the captured images show a unique silhouette, a method known as shape from silhouette can be used to reconstruct a volume model. This method is limited to convex surfaces. Other methods use registration information, created from gradients or marker positions.

A cylindrical approximation is adapted to the captured head and positioned in the middle of the assumed head axis. Head dimensions are defined in the book of Farkas [Farkas95] with average values of 151mm for the width of a head and 197mm for the depth of a head. Similar numbers can also be found in the report of Young [Young93]. Considering a head with hair and two ears, a circular assumption is one possible approximation to the ellipse shaped head therefore suitable for the given problem. Besides the head size (radius for the cylinder), which can be taken from the mentioned statistical publications, there are other remaining unknown positions. Both offsets (x- and z-axis) have to be estimated.

Therefore the following equation describes the problem.

$$r^2 = \left(x - o_x\right)^2 + \left(z - o_z\right)^2$$

This non-linear equation can be solved by Levenberg-Marquadt approach as described in the book of Scales [Scales85]. We have used the 3D reconstructed locations of both eye pupil centers as well as left and right eye corner of both eyes as fitting locations. The radius was fixed to a slightly bigger value than given for the head depth by the statistical publications, because of the requirements for the resulting texture map (±90°). The determined cylinder offset parameter for the x- and z-axis are used to place the cylinder at the face surface with the highest impact for the alignment, like the eyes and mouth. Therefore, the edges of these features are aligned and the displacement error for other features are not visible, e.g. nose tip. The fitted cylinder and a face surface model are shown in Fig 8. Therefore, some face surface points are reconstructed using the calibration data and perspective projection. Due to the fact, that several views provide 2D locations of the same feature point, a multi view approach for the reconstruction is used for more reliable 3D feature point locations. The handling of outliers, which can be a result of the point correspondences or calibration data, is crucial at this step. The average of permuted reconstructed 3D point locations using two views as well as a closed-form solution for all views have the drawback of moving the solution to one or more outliers. We have used a combination of the permutation and closed-form solution, which takes advantage of the back projection error in all considered views and is therefore an iterative solution. The consideration of the minimum back projection error as measurement for the 3D point localization accuracy leads to reliable surface points and identifies possible problems with 2D point locations. The evaluation result can not only be used for more exact 3D point locations by excluding outliers, but also to define weights according to the quality of back projection and therefore to define the influence of each view to the to be created texture map.

The composition is mainly described by the projection of each view onto the object surface and in our case a cylindrical approximation. The perspective warping of images onto a cylindrical surface is described in the technical report of Szeliski [Szeliski04]. Afterwards a projection will

take place, in order to create the texture map. Sampling the cylinder surface with the desired resolution of the virtual view is used for this approach. This method allows a set of DOFs, like the horizontal and vertical resolution as well as defining a specific region of interest. A linear blending rule, like the one shown in Fig 7 is used to incorporate adjacent views. For each horizontal position along the circumference left and right views are selected according to the camera rotation around the y-axis. The angle differences between the view vector and the selected left and right views are converted to weights. This method is constrained on the assumption, that the views are now rotating around the head and cylinder center and not longer around the convergence point of all views. The 3D reconstructed face features show a very small offset along the x axis and a strong offset along the z axis, so that middle view still refers to the most frontal face view. Using a linear interpolation between adjacent views leads to a weight of 0.5 in the middle of these both views.

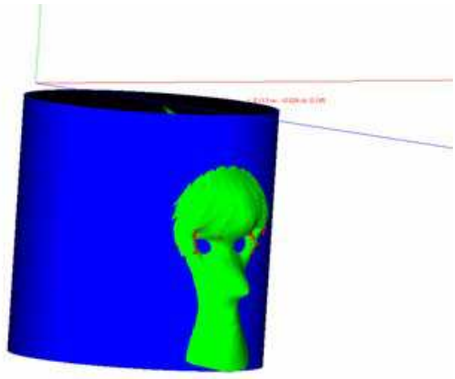


Fig. 8. Interpolated face surface model with a cylinder fitted at both eyes.

The result is a texture map shown in Fig 6, where nine views are incorporated to one cylindrical projected face image showing ±90°. At the top and the bottom of the created image some ghost parts are visible. These wrong stitched parts are caused by the cylindrical assumption and the difference to the real object surface. A more detailed 3D object would decrease such miss alignments of the regions below the chin.

4.3 Global and local alignment

In order to use the created texture maps for the mentioned pose invariant face recognition approach, an alignment has to take place. Common is to align the input images globally, by using the eye centers as reference points. With the publication of Tsapatsoulis et al. [Tsapatsoulis98] the positive influence of local aligned face features is described, if only a resizing takes place. We extend this idea to a complete affine transformation of the area around the desired face feature and applying the transformation under the support of graphics card rendering. Global alignment is based on 2D affine transformation and performed for each created texture map separably. The parameters for the transformation are determined by using selected face feature locations in all created virtual images. The average feature point location is used as reference and all transformations are calculated with respect to these data. The feature point locations used for the texture map creation are transformed, in order to use them for the alignment. The accuracy of these locations are good enough, compared a manual selections, considering the depth distance between approximation and real object surface.

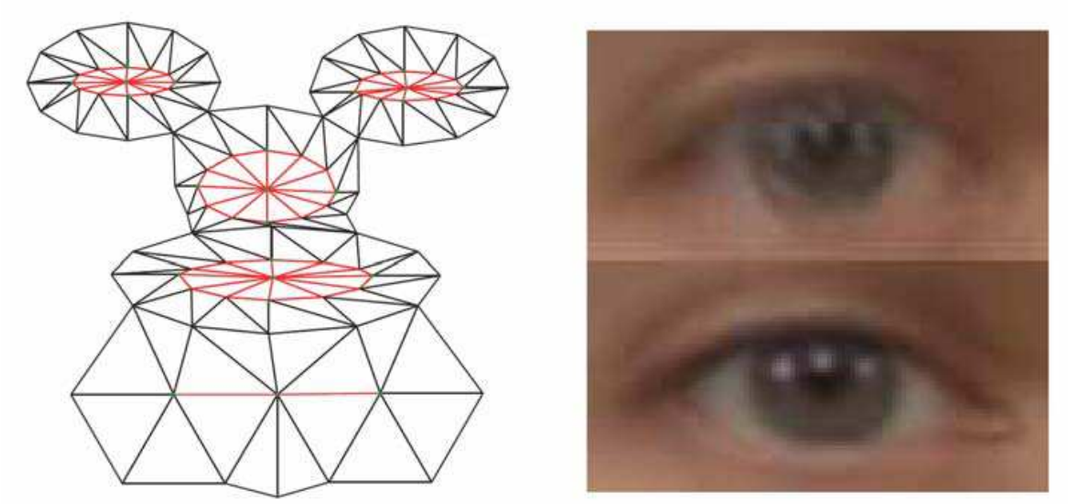


Fig. 9. (left) Adapted triangle mesh used for the local alignment of selected face features (both eyes, nose, mouth and chin). (right) Average images after (right top) global and (right bottom) local alignment of 30 images of right eyes from male persons.

The alignment result for 30 right eyes is shown in the upper part of Fig 9. In order to achieve a better alignment result, the regions around selected face features are aligned locally. The generic triangle mesh from Fig 9 is adapted and placed at the global transformed face feature locations and the associated texture information for each triangle is extracted. The face feature locations are transformed to the desired position and the associated textured triangles will perform the texture map interpolation. The feature region warping is supported by the graphics hardware by rendering the deformed triangle mesh.



Fig. 10. Average image of 25 texture maps of different persons (male and female mixture) after global and local alignment of selected face features (both eyes, nose, mouth and chin).

Using this approach leads to locally aligned virtual texture maps, as can be seen in Fig 10. The average image shows a very good definition in the locally aligned facial features. A closer look and direct comparison between just global and global plus local aligned right eyes is given in Fig 10, which shows again the alignment of 30 right eyes.

5. Experimental results

5.1 Dataset and experiment description

Two different databases have been created in order to show the performance of the Partial Principal Component Analysis (P²CA) approach and the improvement in the recognition accuracy when using the automatic approach for the creation of the virtual alignment texture maps.

The first one: The UPC face database [UPC-FaceDatabase] contains a total of 756 images corresponding to 28 persons with 27 pictures per person acquired under different pose views (0^o, $\pm30^o$, $\pm45^o$, $\pm60^o$ and $\pm90^o$) and three different illuminations (environment or natural light, strong light source from an angle of 45^o, and finally an almost frontal mid-strong light source). The images have been normalized to an output resolution of 122x100 pixels. The 180^o cylindrical training images have been created by manually morphing five images (0^o, $\pm45^o$and $\pm90^o$) that have been acquired in a different session than the rest of the pictures under environmental light conditions. This database is used for testing the P²CA approach towards pose and illumination variations and also for comparing this technique with the conventional Eigenfaces approach [Turk91].

The second dataset has been created using the method proposed above for the automatic creation of texture maps. Although this database contains only 20 different subjects it will be used to show the improvement of the local alignment method proposed. For each subject two different texture maps have been created: One using only global alignment, and another texture map using the global and local alignment process described in Section 4.3. So for this database, two different training sets have been created for the training stage of the face recognition system: Training set T1 using the global alignment maps; and training set T2 using the local alignment maps. These training sets will be also used as gallery sets. The test set is composed of a total of 9 different views for each subject (180 test images) acquired in a second session using the setup depicted in Fig 7.

A total of three different experiments have been performed: The first one tries to evaluate the P²CA technique under different assumptions using the UPC Face Database. The second experiment is focused on a comparative between the Partial Principal Component Analysis technique and the conventional *Eigenfaces* approach. Finally, the third experiment is carried out on the second dataset and it shows the improvement in the recognition rate when introducing the proposed automatic creation and alignment of the texture maps.

5.2 Results

In the first experiment we will identify which view angles present lower recognition accuracy using P²CA. For this reason, the test set of the UPC-FaceDatabase is divided in subsets depending on the angle view of the face image to obtain Table 1 and Table 2. The first Table represents the recognition results when using only the images under the neutral illumination for the evaluation, whereas the second Table considers the three different illuminations of the test set.

Angle	Recog.	Angle	Recog.	Angle	Recog.
-90º	82.14%	-30º	92.85%	+45º	96.43%
-60º	85.71%	0º	100%	+60º	82.14%
-45º	96.43%	+30º	82.14%	+90º	85.71%

Table 1. Recognition Accuracy of P²CA under natural light (neutral illumination)

Angle	Recog.	Angle	Recog.	Angle	Recog.
-90º	72.63%	-30º	73.80%	+45º	70.23%
-60º	71.42%	0º	75%	+60º	70.23%
-45º	73.80%	+30º	69.05%	+90º	70.23%

Table 2. Recognition under 3 different illumination conditions

Table 1 shows that the new technique is considerably robust towards changes in the pose of the face. On the other hand, when testing the images under the three different illuminations (especially the strong light source from an angle of 45º) the algorithm presents lower recognition accuracy as depicted in Table 2. Moreover, it can also be concluded that the positive pose variations present a lower recognition rate. The reason for this reduction of the recognition accuracy is that one of the three illuminations of the UPC-FaceDatabase corresponds to a very strong spotlight coming from the positive side

In the next experiment, we verify the robustness of P²CA in front of the conventional 2D strategies. Thus, Eigenfaces (PCA), and 2DPCA have been implemented. For the training of the conventional 2D strategies, 5 different face views for each subject have been used as training and gallery data.

Method \ Exp.	**Neutral illumination**	**3 illuminations**
PCA / Eigenfaces	72.22%	60.45%
2DPCA	75%	61.24%
P²CA	91.91%	72.9%

Table 3. Recognition Accuracy of the different algorithms

The results presented in Table 3 show that the novel 2D-3D mixed scheme (P²CA) outperforms its respective two dimensional approaches (PCA and 2DPCA) when varying pose.

Finally, in the third experiment we analyze the advantages of creating the virtual images with the global and the local alignment method. Table 4 summarizes the results for face recognition when using the datasets described above.

	Training Set T1 (global alignment)	**Training Set T2 (global + local alignment)**
Recognition Rate	93.33%	96.66%

Table 4. Improvement using the local alignment method

Although only 20 different persons are used in the experiments, results show that the local aligned virtual images present a slight improvement in the recognition rate. The improvement of this rate has been obtained for the 0^o and $\pm 8^o$ views since these views enclose all the face features used for the alignment.

6. Conclusions and future work

Face processing is one of the most active research fields as demonstrated by more than 1000 publications that have appeared in different conferences and journals in the last two years. Additionally, it is also a mature topic with more than 30 years. Recently, a new trend of 3D face recognition approaches showed an increase in the recognition rate if 3D data is available. Nevertheless, cost of the set-up, acquisition time and cooperation of the subjects are still some of the requirements for obtaining accurate 3D data that may not be available during the recognition stage. Thus, we have presented here a possible alternative following a mixed 2D-3D face recognition philosophy, i.e. the system is trained with 3D data but it can use either 2D or 3D data in the test stage. We have presented the extension of the 2D statistical PCA method [Turk91, Yang04] to a 2D-3D face recognition scheme (Partial Principal Component Analysis). However, this philosophy may be extended also to other face recognition statistical approaches like LDA or ICA with have shown a higher robustness in the presence of illumination variations. Additionally, we have presented an automatic approach for the creation of aligned virtual view images using nine different views. These aligned virtual view images are used as training data for the P²CA technique. The virtual view image is created by using a cylindrical approximation for the real object surface. The alignment is done by global and local transformations of the whole image and face features, respectively. Results show an improvement in the recognition rate when using the local alignment procedure proposed.

7. References

General

[3DRMA] 3D RMA Database http://www.sic.rma.ac.be/~beumier/DB/3d_rma.htm

[AFGR] *Proceedings of International Conference on Automatic Face and Gesture Recognition*

[AVBPA] *Proceedings of International Conference on Audio and Video-Based Person Authentication*

[FRGC05] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W.Worek, "Overview of the Face Recognition Grand Challenge," *IEEE Conf. Computer Vision and Pattern Recognition(CVPR)*, 2005

[FRHomepage] Face Recognition Homepage, in http://www.face-rec.org/

[FRVT02] P. J. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and J. Bone , "Face Recognition Vendor Test 2002: Evaluation Report". NISTIR 6965, National Institute of Standards and Technology, 2003. http://www.frvt.org

[Zhao03] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey" *ACM Computing Surveys*, Vol.35, Nº4, December 2003

[Zhao06] W. Zhao, and R.Chellapa, "Face Processing: Advanced modeling and methods", Academic Press, 2006

2D face recognition

[Bartlett02] M. S. Barlett, J.R. Movellan, T.J.Sejnowski, "Face Recognition by Independent Component Analysis", on *IEEE Trans. On Neural Networks*, vol 13, no. 6, Nov 2002

[Bartlett98] M. S. Bartlett, H. M. Lades, and T. J. Sejnowski, "Independent component representations for face recognition," in *Proc. SPIE Symp. Electon. Imaging: Science Technology – Human Vision and Electronic Imaging III*, vol. 3299, T. Rogowitz and B. Pappas, Eds., San Jose, CA, 998, pp. 528–539.

[Belhumeur97] R.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, "Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection" *in IEEE Transacations Pattern Analysis and Machine Intelligence*, vol 19, N. 7, July 1997

[Brunelli93] R. Brunelli and T. Poggio, "Face Recognition: Features vs Templates" *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.15, no. 10, pp 1042-1053. Oct 1993

[Delac05] Delac, K., Grgic, M., Grgic, S., Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set, *International Journal of Imaging Systems and Technology*, Vol. 15, No. 5, 2005

[Fadzil94] Mohamad Hani Ahmad Fadzil, Abu Bakar H. "Human Face Recognition Using Neural Networks" *ICIP (3) 1994*: 936-939. 1994

[Haddadnia02] J. Haddadnia, K. Faez, M. Ahmadi, "N-Feature Neural Network Human Face Recognition," *Proc. of 15th International Conference on Vision Interface*, Calgary, May 2002, pp. 300-307

[Heo06] J. Heo, Ma. Savvides, R. Abiantun, Ch. Xie, and B.V.K. Vijayakumar "Face Recognition With Kernel Correlation Filters On A Large Scale Database", in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Toulouse, May 14th- 19th, 2006

[Kanade73] T. Kanade, "Picture processing system by computer complex and recognition of human faces", Dept. of Information Science, Kyoto University, 1973

[Kim02] K. I. Kim, K. Jung, H.J. Kim, "Face Recognition Using Kernel Principal Component Analysis", *IEEE Signal Processing Letters*, Vol.9, No. 2, February 2002

[Kohonen88] T. Kohonen, "Self-Organization and Associative Memory". Berlin: Springer. 1988

[Lades93] M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R.P. Wurtz, W. Konen. "Distortion Invariant Object Recognition in the Dynamic Link Architecture" *IEEE Transactions on Computers* archive Volume 42 , Issue 3 Pages: 300 – 311. 1993

[Lawrence97] S. Lawrence, C. L. Giles, A. C. Tsoi and A.D. Back, "Face Recognition: A Convolutional Neural-Network Approach", *IEEE Transactions Neural Networks*, Vol. 8, No. 1, 1997, pp. 98-113. IDIAP--RR 03-20

[Lin97] S. H. Lin, S. Y. Kung, and L.J. Lin, "Face Recognition/Detection by Probabilistic Decision-Based Neural Network" *IEEE Transactions on Neural Networks*, 8(1):114-132.1997

[Lorente99] L. Lorente, L. Torres, "*Face recognition of video sequences in a MPEG-7 context using a global eigen approach*", *International Conference on Image Processing*, Kobe, Japan, October 25-29, 1999

[Moghaddam02] B. Moghaddam, Principal manifolds and probabilistic subspaces for visual recognition, *in IEEE Transactions Pattern Analysis and Machine Intelligence* 24 (2002), 780– 788.

[Moses94] Y. Moses, Y. Adani, and S. Ullman. "Face Recognition: The problem of compensating for changes in illumination direction", *in Proceedings European Conference on Computer Vision* pp. 286-2996, 1994

[Nefian96] A.V. Nefian. "Statistical Approaches To Face Recognition". Qualifying Examination Report. Georgia Institute of Technology. Dec 1996

[Palanivel03] S. Palanivel, B.S.Venkatesh and B.Yegnanarayana, "Real time face recognition system using Autoassociative Neural Network models," *in IEEE International Conference on Acoustics, Speech and Signal Processing*, Hong Kong, April 2003, pp. 833-836

[Pentland94] A. P. Pentland, B. Moghaddam, T. Starner and M. Turk, "*View-based and modular eigenspaces for face recognition*", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 84-91,1994.

[Schölkopf98] B.Schölkopf, A. Smola, K. Müller, „Non-linear Component Analysis as a Kernel Eigenvalue Problem", *Neural Computers*, vol. 10, pp. 1299-1319, 1998

[Schölkopf99] B.Schölkopf, A. Smola, K. Müller, "Kernel Principal Component Analysis" in *Advances in Kernel Methods-Support Vector Learning*, pp. 327-352, MIT Press, 1999

[Turk91] M. A. Turk, A. P. Pentland, "Face recognition using eigenfaces", *Proc. of the IEEE Comp. Soc. Conf. on CVPR*, pp. 586-591, Hawaii 1991

[Wang06] J. Wang, K. N. Plataniotis, A. N. Venetsanopoulos, "Selecting Kernel Eigenfaces For Face Recognition With One Training Sample Per Subject", ", in *IEEE International Conference & Multimedia Expo*, Toronto, Canada, July 9th-12th 2006

[Wiskott99] L. Wiskott, J.M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching" *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775—779. Revised version 1999

[Yang02] M.H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods", in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition (FG'02)*, pp. 215-220, May 2002.

[Yang04] J. Yang, D. Zhang, A.F. Frangi, and J.Yang, "Two-Dimensional PCA: A New Approach to Appearance-based Face Representation and Recognition", in *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Jan. 2004

[Yu06] J. Yu, Q. Tian, "Constructing descriptive and discriminant features for face classification" in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Toulouse, May 14th- 19th, 2006

[Zhang97] J. Zhang, Y. Yan, M. Lades, "*Face recognition: eigenface, elastic matching, and neural nets*", *Proceedings of the IEEE*, Vol. 85, No. 9, pp. 1423-1435, September 1997

3D-face recognition

[Antini06] G. Antini, S. Berretti, A. Del Bimbo, P. Pala, "3D Face Identification Based On Arrangement Of Salient Wrinkles", in *IEEE International Conference & Multimedia Expo*, Toronto, Canada, July 9th-12th 2006

[Beymer95] D. Beymer and T. Poggio , "Face Recognition from One model View", in *Proc. Fifth Int'l Conf. Computer Vision*, 1995

[Blanz03] V. Blanz, and T. Vetter, "Face Recognition based on fitting 3D morphable model", in *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(9):1063-1074, 2003

[Bowyer04] K. Bowyer, K. Chang, and P. Flynn, "A Survey of Approaches to 3D and Multi-Modal 3D+2D Face Recognition," *IEEE Intl.Conf. on Pattern Recognition*, 2004.

[Bronstein05] A. M. Bronstein, M. M. Bronstein, R. Kimmel, "Three-dimensional face recognition „ in *International Journal of Computer Vision* Vol.64/1, pp. 5-30, August 2005

[Cartoux89] J. Y. Cartoux, J. T. Lapreste, and M. Richetin. "Face authentification or recognition by profile extraction from range images." *In Workshop on Interpretation of 3D Scenes*, pages 194- 199, 1989

[Chang05] K. I. Chang, K. W. Bowyer, and P. J. Flynn, "An Evaluation of Multimodal 2D+3D Face Biometrics", in *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.27, pp 619-624, April 2005

[Feng06] S. Feng, H. Krim, I. Gu, M. Viberg, "3D Face Recognition Using Affine Integral Invariants" in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Toulouse, May 14th- 19th, 2006

[Georghiades01] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From Few to Many: Illumination Cone Models for Face Recognition" in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643-- 660, 2001

[Gordon91] Gaile G. Gordon. Face recognition based on depth maps and surface curvature. In *Geometric Methods in Computer Vision*, volume 1570 of *SPIE Proceedings*, pages 234- 247, 1991

[Lee90] John Chun Lee and E. Milios. Matching range images of human faces. *In Proc. IEEE International Conference on Computer Vision*, pages 722- 726, 1990

[Lu05] X. Lu, and A.K. Jain, "Integrating Range and Texture Information for 3D Face Recognition", in *Proc. IEEE WACV*, Breckenridge, Colorado 2005

[Lu06] X. Lu, and A.K. Jain, "Matching 2.5D face scans to 3D models", in *IEEE Transactions Pattern Analysis and Machine Intelligence*, Jan. 2006 Volume: 28, pp: 31- 43

[Onofrio06] D. Onofrio, A. Rama, F. Tarres, S. Tubaro, "P²CA: How Much Information is needed", *IEEE International Conference on Image Processing*, Atlanta, USA, October 2006

[Rama05a] A.Rama, F.Tarrés, D.Onofrio, and S. Tubaro, "Using partial information for face recognition and pose estimation", in *IEEE International Conference & Multimedia Expo*, Amsterdam, July 6th-8th 2005

[Rama05b] A.Rama, and F.Tarrés, "P²CA: A new face recognition scheme combining 2D and 3D information", in *IEEE International Conference on Image Processing*, Genoa, Italy, September 11th-14th 2005

[Rama06] A.Rama, F.Tarres, D.Onofrio, and S.Tubaro, "Mixed 2D-3D Information for pose estimation and face recognition" in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Toulouse, May 14th- 19th, 2006

[Samani06] A. Samani, J. Winkler, M. Niranjan, „Automatic Face Recognition Using Stereo Images", in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Toulouse, May 14th- 19th, 2006

[Tanaka98] Hiromi T. Tanaka, Masaki Ikeda, and Hisako Chiaki. Curvature-based face surface recognition using spherical correlation - principal direcions for curved object recognition. *In Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pages 372- 377, 1998

[Tsalakanidou04] F. Tsalakanidou, S. Malassiotis, and M. Strintzis, "Integration of 2D and 3D Images for enhanced face authentication". *Sixth International Conference on Automated Face and Gesture Recognition*, pages 266-271, May 2004

[UPC-FaceDatabase] "UPC Face Database" in http://gps-tsc.upc.es/GTAV

3D face modelling, reconstruction and texture map creation

[Akimoto93] Akimoto, T., Suenaga, Y.,Wallace, R.S.: Automatic Creation of 3D Facial Models. IEEE Computer Graphics and Applications 13 (1993) 16–22

[Ansari03] A. Ansari, M. Abdel-Mottaleb, "3D Face Modelling Using Two Ortogonal Views and a Generic Face Model", *in Proc. of Int. Conf. and Multimedia and Expo*, July 2003

[Brown01] Brown, L.M.: 3D Head Tracking Using Motion Adaptive Texture-Mapping. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Hawaii, IEEE Computer Society (2001) 998–1003

[Cascia98] Cascia, M.L., Isidoro, J., Sclaroff, S.: Head Tracking via Robust Registration in Texture Map Images. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA, USA, IEEE Computer Society (1998) 508

[Farkas95] Farkas, L.G.: Anthropometry of the Head and Face. 2nd edn. Raven Press (1995)

[Forster01] Forster, F., Lang, M., Radig, B., 2001. Real-time 3D and colour camera. In: Proc. Internat. Conf. on Augmented, Virtual Environments and 3D Imaging, Mykonos, 2001. pp. 45–48.

[Kouzani99] Kouzani, A., Sammut, K.: Quadtree principal component analysis and its application to facial expression classification. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Hawaii (1999) 835–839

[Lee00] Lee, W.S., Magnenat-Thalmann, N.: Fast head modeling for animation. Image and Vision Computing 18 (2000) 355–364(10)

[Liu03] Liu, X. and Chen, T., Geometry-assisted statistical modeling for face mosaicing, IEEE Transactions on Systems, in Proc. IEEE Int. Conf. Image Process, 2003, Volume 2, page 883-886

[Onofrio04] D. Onofrio, A. Sarti, and S. Tubaro, "Area Matching Based on Belief Propagation with Applications to Face Modeling", in *IEEE International Conference on Image Processing*, Singapore 2004

[Onofrio05] D. Onofrio, S. Tubaro, A. Rama, F. Tarres, "3D Face Reconstruction with a four camera acquisition system". *Proc. International Workshop on Very Low Bitrate Video Coding (VLBV05)*, Costa Rei, Sardinia, 15-16 September 2005

[Pedersini99] F. Pedersini, A. Sarti, S. Tubaro, "Multicamera Systems: Calibration and Applications," *IEEE Signal Processing Magazine, Special Issue on Stereo and 3D Imaging*, vol. 16, N. 3, pp. 55-65, May 1999

[Scales85] Scales, L.E.: Introduction to non-linear optimization. Springer-Verlag New York, Inc., New York, NY, USA (1985)

[Scharstein02] D. Scharstein, R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *Int'l J. Computer Vision*, vol. 47, N. 1, pp 7-42, 2002.

[Singh07] Singh, R. and Ross, A., A Mosaicing Scheme for Pose-Invariant Face Recoginition, IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, Volume 37(5), 2007

[Soh95] Soh, A.W.K., Yu, Z., Prakash, E.C., Chan, T.K.Y., Sung, E.: Texture Mapping of 3D Human Face for Virtual Reality Environments. International Journal of Information Technology 8 (2002) 54–65

[Szeliski04] Szeliski, R.: Image Alignment and Stitching: A Tutorial. Technical Report MSR-TR-2004-92, Microsoft Research (2004)

[Tsapatsoulis98] Tsapatsoulis, N., Doulamis, N., Doulamis, A., Kollias, S.: Face extraction from non-uniform background and recognition in compressed domain. In:

Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Volume 5, Seattle, WA, USA (1998) 2701–2704

[Xiao93] Xiao, J., Kanade, T., Cohn, J.F.: Robust Full-Motion Recovery of Head by Dynamic Templates and Re-registration Techniques. International Journal of Imaging Systems and Technology 13 (2003) 85–94

[Young93] Young, J.W.: Head and Face Anthropometry of Adult U.S. Civilians. Technical Report ADA268661, Civil Aeromedical Institute, Federal Aviation Administration (1993)

IntechOpen

IntechOpen



Recent Advances in Face Recognition

Edited by Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett

ISBN 978-953-7619-34-3
Hard cover, 236 pages
Publisher InTech
Published online 01, June, 2008
Published in print edition June, 2008

The main idea and the driver of further research in the area of face recognition are security applications and human-computer interaction. Face recognition represents an intuitive and non-intrusive method of recognizing people and this is why it became one of three identification methods used in e-passports and a biometric of choice for many other security applications. This goal of this book is to provide the reader with the most up to date research performed in automatic face recognition. The chapters presented use innovative approaches to deal with a wide variety of unsolved issues.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Antonio Rama Calvo, Francesc Tarrés Ruiz, Jürgen Rurainsky and Peter Eisert (2008). 2D-3D Mixed Face Recognition Schemes, Recent Advances in Face Recognition, Kresimir Delac, Mislav Grgic and Marian Stewart Bartlett (Ed.), ISBN: 978-953-7619-34-3, InTech, Available from: http://www.intechopen.com/books/recent_advances_in_face_recognition/2d-3d_mixed_face_recognition_schemes

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2008 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen