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

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

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

An Integrative Approach to Face and Expression Recognition from 3D Scans

Chao Li
Florida A&M University
USA

1. Introduction

Face recognition, together with fingerprint recognition, speaker recognition, etc., is part of the research area known as 'biometric identification' or 'biometrics', which refers to identifying an individual based on his or her distinguishing characteristics. More precisely, biometrics is the science of identifying, or verifying the identity of, a person based on physiological or behavioral characteristics (Bolle et al., 2003). Biometric characteristics include something that a person is or produces. Examples of the former are fingerprints, the iris, the face, the hand/finger geometry or the palm print, etc. The latter include voice, handwriting, signature, etc. (Ortega-Garcia et al., 2004).

Face recognition is a particularly compelling biometric approach because it is the one used every day by nearly everyone as the primary means for recognition of other humans. Because of its natural character, face recognition is more acceptable than most other biometric methods. Face recognition also has the advantage of being noninvasive.

Face recognition has a wide range of potential applications for commercial, security, and forensic purposes. These applications include automated crowd surveillance, access control, mug shot identification (e.g., for issuing driver licenses), credit card authorization, ATM machine access control, design of human computer interfaces (HCI), etc. Especially, the surveillance systems rely on the noninvasive property of face recognition systems.

According to the different purposes in applications, face recognition scenarios can be classified into the following two:

- Face verification: ("Am I who I say I am?") is a one-to-one match that compares a query face image against a gallery face image whose identity is being claimed.

- Face identification: ("Who am I?") is a one-to-many matching process that compares a query face image against all the gallery images in a face database to determine the identity of the query face. In the identification task, it is assumed that the person is in the database. The identification of the query image is done by choosing the image in the database that has the highest similarity score with the query image.

According to the format of the data, analyzed face recognition methods can be classified as 2D face recognition, 3D face recognition and infrared face recognition modalities. The infrared face recognition is commonly combined with other biometrics technologies.

Most of the face recognition methods developed until recently use 2D intensity images obtained by photographic cameras as the data format for processing. There are two reasons for the interest in 2D face recognition. First of all, human beings can recognize a person from his or her picture alone, which means that the picture contains enough information about a person's identity for a human being. Second, a picture is very easy to obtain in terms of acquisition and cost. This was even more important decades ago, when other imaging techniques, such as 3D imaging, were not well developed.

Varying levels of success have been achieved in 2D face recognition research. A short overview is given in the following paragraph. More detailed and comprehensive surveys can be found in (Chellappa et al., 1995; Zhao et al., 2003). Most proposed techniques fall into two lines of research. The first one is appearance based (view based) face recognition. This approach represents a face image as a high dimensional vector, i.e., a point in a high dimensional vector space. Statistical techniques are then used to analyze the distribution of the face image in the vector space, and features are extracted in the vector space using linear decomposition like Principal Component Analysis (PCA) (also called 'the Eigen Face' method proposed in (Turk and Pentland, 1991)), Independent Component Analysis (ICA) (Bartlett, et al., 1998) or non linear methods like Kernel Principal Component Analysis (KPCA) (Cholkopf et al., 1998). The second group of approaches is model-based. The model-based face recognition scheme is aimed at constructing a model of the human face, which can capture the facial variations of individuals. Exemplar approaches in this category include Feature-based Elastic Bunch Graph Matching (Wiskott et al. 1997) and Active Appearance Model (AAM).(Edwards et al., 1998)

In 2002, the Face Recognition Vendor Test (FRVT 2002) was held. It was an independently administered technology evaluation sponsored by the Defense Advanced Research Projects Agency (DARPA), the National Institute of Standards and Technology (NIST) and other agencies. The primary objective of FRVT 2002 was to provide performance measures for assessing the ability of automatic face recognition systems to meet real-world requirements. FRVT 2002 measured the performance of the core capabilities of face recognition technology, all based on 2D face recognition. Ten participants were evaluated. (Phillips et al., 2002)

FRVT 2002 showed that one of the most challenging tasks for modern face recognition systems is recognizing faces in non-frontal imagery. Most face recognition systems performed well when all of the images were frontal. But, as a subject became more and more off angle (both horizontally and vertically), performance decreased. Additionally FRVT 2002 also showed that the variation in the structure of lighting had a great effect on performance (Phillips et al., 2002).

From these observations, the following conclusions can be drawn. Although 2D face recognition has achieved considerable success, certain problems still exist. Because the 2D face images used not only depend on the face of a subject, but also depend on the imaging

factors, such as the environmental illumination and the orientation of the subject. These two sources of variability in the face image often make the 2D face recognition system fail. That is the reason why 3D face recognition is believed to have an advantage over 2D face recognition. Although the first main rationale for 2D face recognition mentioned above, i.e. "a picture contains enough information about the identity of a person", may be true for a human being, for an artificial system, without the biological knowledge of how the human vision system works, 2D face recognition is still a very difficult task.

With the development of 3D imaging technology, more and more attention has been directed to 3D face recognition. In (Bowyer et al., 2004), Bowyer et al. provide a survey of 3D face recognition technology. Some of the techniques are derived from 2D face recognition, such as the use of PCA in (Hesher et al., 2003; Chang et al. 2003) to extract features from faces. Some of the techniques are unique to 3D face recognition, such as the geometry matching method in (Gordon, 1991) and the profile matching proposed in (Cartoux et al. 1989; Nagamine et al. 1992).

Most of the 3D face recognition systems treat the 3D face surface as a rigid surface. But actually, the face surface is deformed by different expressions of the subject. So, systems which treat the face as a rigid surface are significantly challenged when dealing with faces with expressions. In (Chang et al., 2005), experiments using Iterative Closest Point (ICP) and PCA methods were performed on the recognition of faces with expression. The authors found that expression changes do cause performance declines in all the methods. (Chang et al., 2005)

Therefore, expression has become a big challenge in 3D face recognition systems. Up to now, only some methods address the facial expression issue in face recognition. In (Bronstein et al., 2003), the authors present a 3D face recognition approach based on a representation of the facial surface, invariant to isometric deformation by facial expression. In (Lu and Jain, 2005), both rigid registration and non-rigid deformations caused by expression were calculated. Iterative Closest Point (ICP) was used to perform rigid registration. For non-rigid deformation, the thin plated spline (TPS) was applied. For the purpose of face matching, the non-rigid deformations from different sources were identified, which was formulated as a two-class classification problem: intra-subject deformation vs. inter-subject deformation. The deformation classification results were integrated with the matching distance of rigid registration to make the final decision. In (Chang et al., 2005), the author tried to extract the area that deforms least with different facial expressions and used this area as the feature for every subject. Then ICP and PCA methods were applied for the matching.

In our research, we want to tackle the expression challenge in 3D face recognition from a different point of view. Because the deformation of the face surface is always related with different expressions, an integrated expression recognition and face recognition system is proposed. In section 2, a model on the relationship between expression and face recognition is introduced. Based on this model, the framework of integrated expression recognition and face recognition is proposed. Section 3 explains the acquisition of the experimental data used and preprocessing performed. Section 4 outlines our approach to 3D face expression

recognition. Section 5 explains the process used for 3D face recognition. Section 6 describes the experiments and the results obtained. Section 7 presents our discussion and conclusion.

2. Relationship between expression recognition and face recognition

From the psychological point of view, it is still not known whether facial expression recognition information directly impacts the face recognition process in human beings. Some models suggest there is not relationship between face recognition and facial expression recognition (Bruce and Young, 1986). Other models support the opinion that a connection exists between the two processes (Hansch and Pirozzolo, 1980).

One of the experiments that support the existence of the connection between facial expression recognition and face recognition was reported in (Etcoff and Magee, 1992). The authors found that people are slower in identifying happy and angry faces than they are in identifying faces with neutral expression. Also, in (Hay et al., 1991) experiments show that people are slower in identifying pictures of familiar faces when they exhibit uncommon facial expressions.
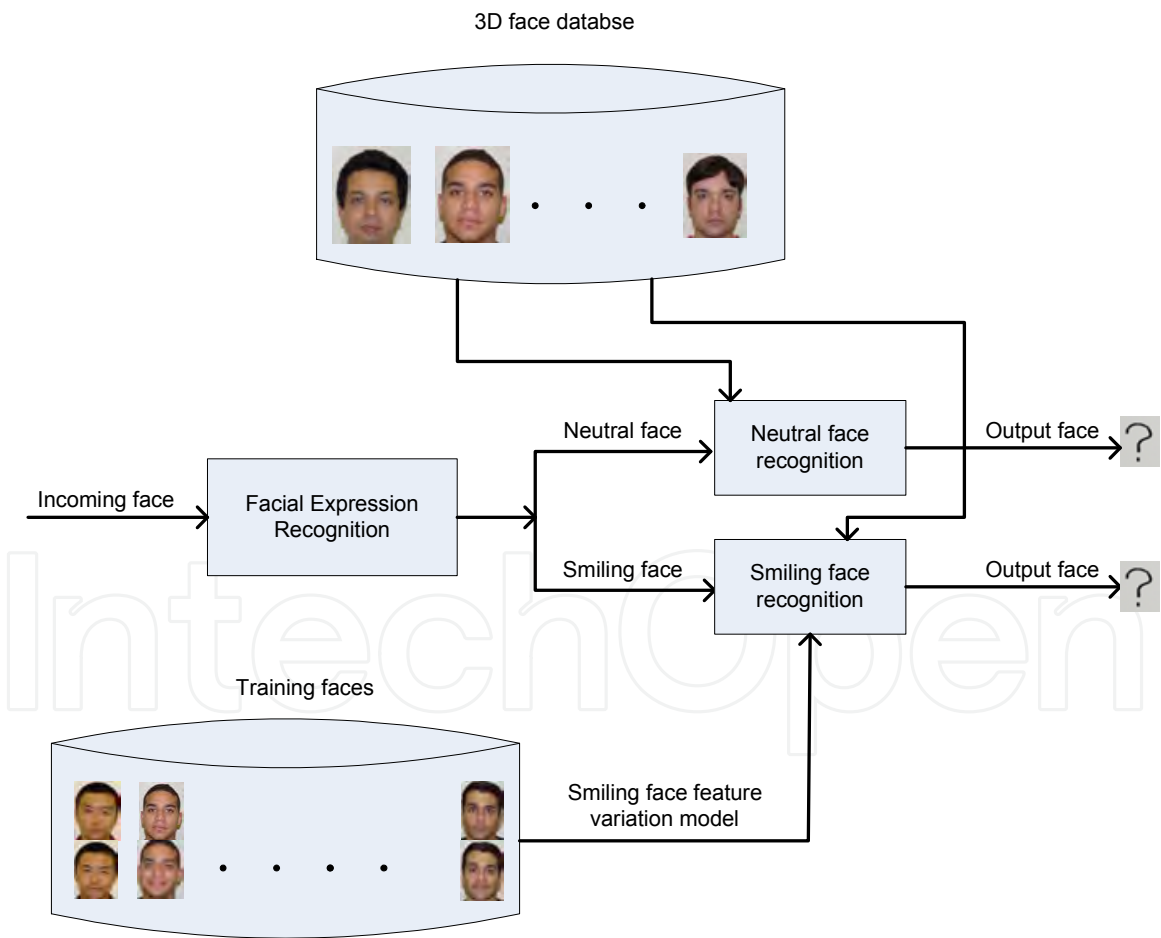

Fig. 1. Simplified framework of 3D face expression

Our proposed framework is based on the assumption that the identification of the facial expression of a query face will aid an automated face recognition system achieve its goal.

The incoming 3D range image is first processed by an expression recognition system to find the most appropriated expression label for it. The expression label could be one of the six prototypical expressions of the faces, which are happiness, sadness, anger, fear, surprise and disgust (Ekman and Friesen, 1971). In addition, the face could also be labeled as 'neutral'. Therefore, the output of the expression recognition system will be one of the seven expressions. Our framework proposes that a different face recognition approach be used for each type of expression. If the expression label determined is neutral expression, then the incoming 3D range image is directly passed to a neutral expression face recognition system, which uses the features of the probe image to match those of the gallery images and get the closest match. If the expression label determined is other than neutral expression, then for each of the six prototypical expressions, a separate face recognition subsystem should be used. The system will find the right face by modeling the variations of the face features between the neutral face and the expressional face. Because recognition through modeling is a more complex process than the direct matching for the neutral face, our framework aligns with the view that people will be slower in identifying happy and angry faces than they will be in identifying faces with neutral expression. Figure 1 shows a simplified version of this framework, it only deals with happy (smiling) expressions in addition to neutral. Smiling is the most common (non-neutral) expression displayed by people in public.

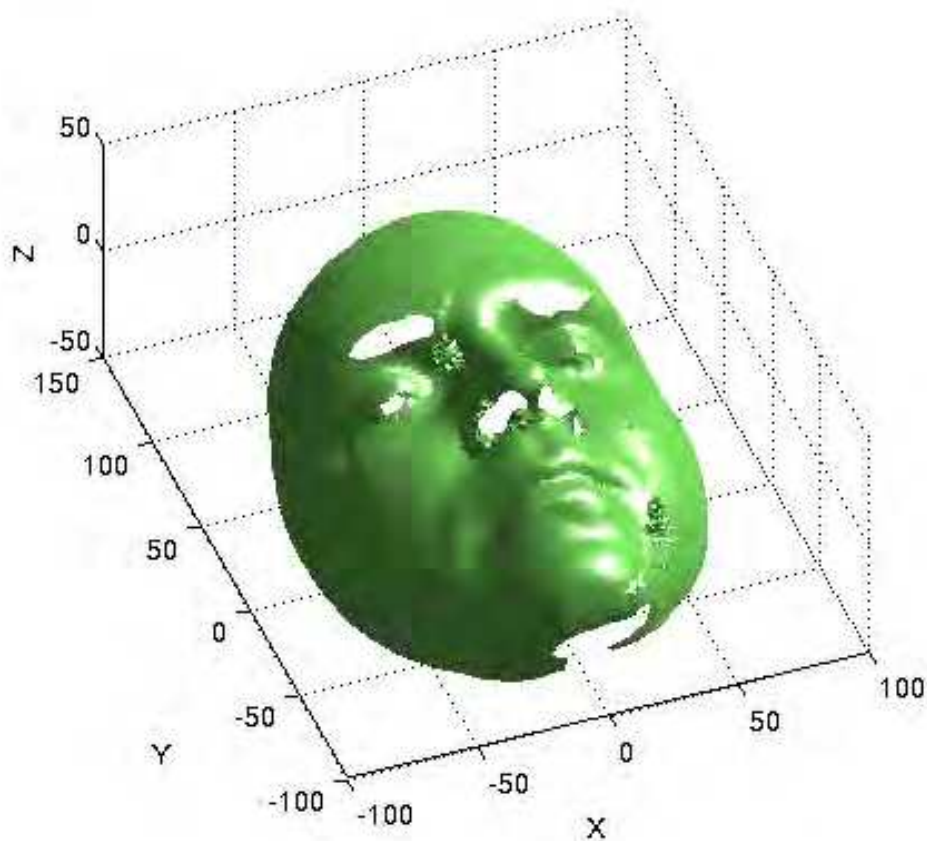## 3. Data acquisition and preprocessing

Fig. 2. 3D face surface acquired by the 3D scanner

To test the performance of our framework, a database including images from 30 subjects, was built. In this database, we included faces with the most common expression i.e., smiling, as well as neutral faces from the same subjects. Each subject participated in two data acquisition sessions, which took place in two different days. In each session, two 3D scans were acquired. One was neutral expression; the other was a happy (smiling) expression. The 3D scanner used was a Fastscan 3D scanner from Polhemus Inc. The accuracy of this scanner is specified as 1mm. The resulting database contains 60 3D neutral scans and 60 3D smiling scans of 30 subjects. Figure 2 shows an example of the 3D scans obtained using this scanner.

In 3D face recognition, registration is a key pre-processing step. It may be crucial to the efficiency of matching methods. In our experiment, a method based on the symmetric property of the face is used to register the face image. In converting the 3D scan from triangulated mesh format to a range image with a sampling interval of 2.5 mm(e.g., Fig 2), trilinear interpolation was used (Li and Barreto, 2005). Unavoidably, the scanning process will result in face surfaces containing unwanted holes, especially in the area covered by dark hair, such as the eye brows. To circumvent this problem, the cubic spline interpolation method was used to patch the holes. An example of the resulting 3D range image is shown in Fig 3.
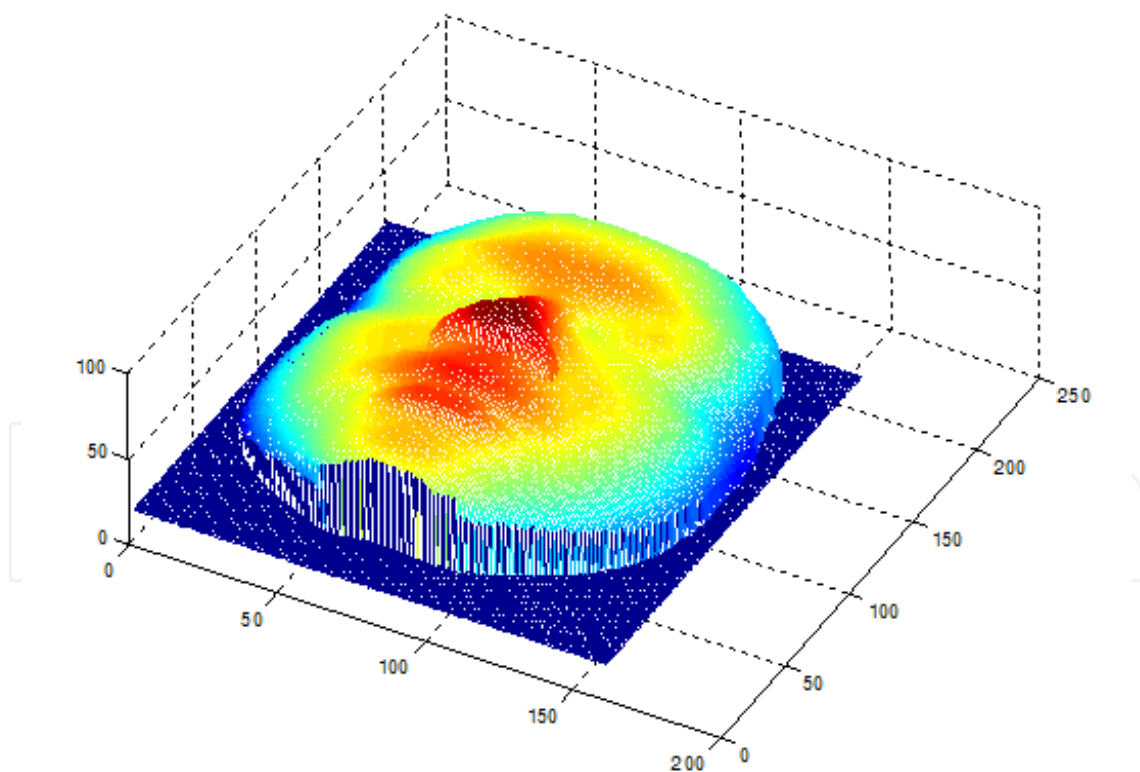

Fig. 3. Mesh plot of the converted range image

4. 3D expression recognition

Facial expression of the face is a basic mode of nonverbal communication among people. The facial expression of another person is often the basis on which we form significant opinions on such characteristics as friendliness, trustworthiness, and status. The facial expressions convey information about emotion, mood and ideas.

In (Ekman and Friesen, 1971), Ekman and Friesen proposed six primary emotions. Each possesses a distinctive content together with a unique facial expression. These prototypical emotional displays are also referred to as basic emotions. They seem to be universal across human ethnicities and cultures. These six emotions are happiness, sadness, fear, disgust, surprise and anger. Together with the neutral expression, these seven expressions also form the basic prototypical facial expressions.

Facial expressions are generated by contractions of facial muscles, which result in temporally deformed facial features such as eye lids, eye brows, nose, lips and skin textures, often revealed by wrinkles and bulges. Typical changes of muscular activities for spontaneous expressions are brief, usually between 250ms and 5s. Three stages have been defined for each expression, which are onset (attack), apex (sustain) and offset (relaxation). In contrast to these spontaneous expressions, posed or deliberate expressions can be found very commonly in social interactions. These expressions typically last longer than spontaneous expressions.
Automatic facial expression recognition has gained more and more attention recently. It has various potential applications in improved intelligence for human computer interface, image compression and synthetic face animation. "Face expression recognition deals with the classification of facial motion and facial feature deformation into abstract classes that are purely based on visual information." (Fasel and Luettin, 2003).

As in face recognition, most contemporary facial expression recognition systems use two-dimensional images or videos as data format. Therefore, the same challenges exist for the face expression recognition as for face recognition, i.e. 2D formats are dependent on the pose of the subjects and on the illumination of the environment. In this respect this paper fills the gap by proposing a facial expression recognition system that uses three dimensional images or range images. 3D range images have the advantage of invariance with respect to subject alignment and illumination. In addition, the deformed features resulting from expressions are easy to extract from 3D range images.

In our experiment, we sought to recognize social smiles, which were posed by each subject, in their apex period. Smiling is the easiest of all expressions to find in photographs and is readily produced by people on demand. The smile is generated by the contraction of the Zygomatic Major muscle. The Zygomatic Major originates in the cheek bone (Zygomatic arch) and inserts in muscles near the corner of the mouth. This muscle lifts the corner of the mouth obliquely upwards and laterally, producing a characteristic "smiling expression". So the most distinctive features associated with a smile are the bulge of the cheek muscle and the uplift of the corner of the mouth, as can be seen in Fig 4. The line on the face generated by a smiling expression is called the nasal labial fold (smile line).
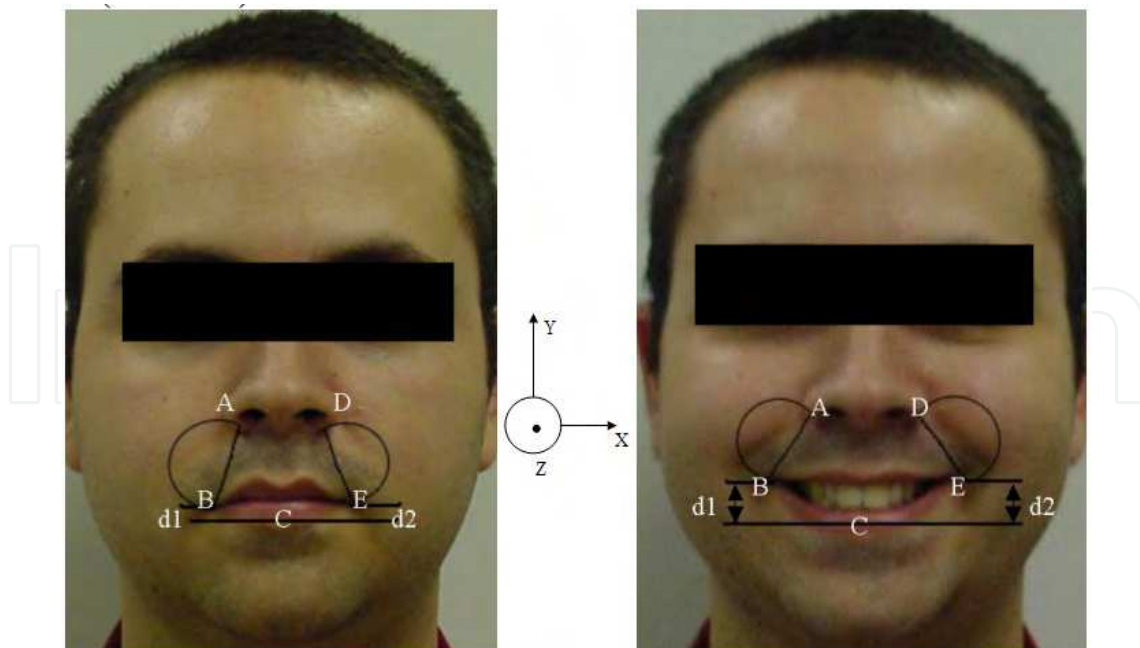
Fig. 4. Illustration of the features of a smiling face

The following steps are followed to extract the features for the smiling expression:

1. An algorithm is developed to obtain the coordinates of five characteristic points A, B, C, D and E in the face range image as shown in Figure 4. A and D are at the extreme points of the base of the nose. B and E are the points defined by the corners of the mouth. C is in the middle of the lower lip.
2. The first feature is the width of the mouth BE normalized by the length of AD. Obviously, while smiling the mouth becomes wider. The first feature is represented by *mw.*
3. The second feature is the depth of the mouth (The difference between the Z coordinates of points BC and EC) normalized by the height of the nose to capture the fact that the smiling expression pulls back the mouth. This second feature is represented by *md.*
4. The third feature is the uplift of the corner of the mouth, compared with the middle of the lower lip d1 and d2, as shown in the figure, normalized by the difference of the Y coordinates of points AB and DE, respectively and represented by *lc.*
5. The fourth feature is the angle of AB and DE with the central vertical profile, represented by *ag.*
6. The last two features are extracted from the semicircular areas shown, which are defined by using AB and DE as diameters. The histograms of the range (Z coordinates) of all the points within these two semicircles are calculated.

The following figure shows the histograms for the smiling face and the neutral face of the above subject.
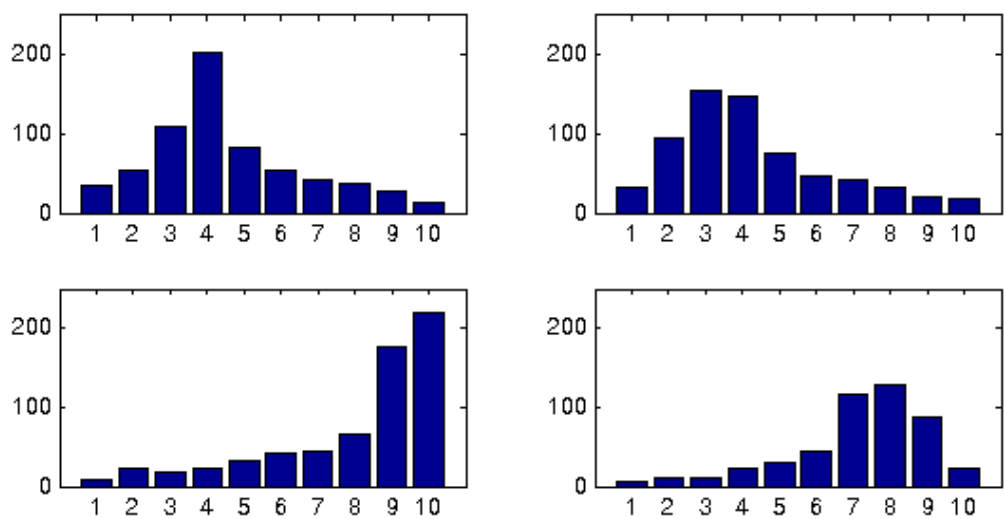
Fig. 5. Histogram of range of cheeks for neutral and smiling face

The two figures in the first row are the histograms of the range values for the left cheek and right cheek of the neutral face image; the two figures in the second row are the histograms of the range values for the left cheek and right cheek of the smiling face image.

From the above figures, we can see that the range histograms of the neutral and smiling expressions are different. The smiling face tends to have large values at the high end of the histogram because the bulge of the cheek muscle. On the other hand, a neutral face has large values at the low end of the histogram distribution. Therefore two features can be obtained from the histograms: One is called the 'histogram ratio', represented by *hr*, the other is called the 'histogram maximum', represented by *hm*.

$$hr = \frac{h6 + h7 + h8 + h9 + h10}{h1 + h2 + h3 + h4 + h5} \tag{1}$$

$$hm = i \quad ; \quad i = \arg\{\max(h(i)\} \tag{2}$$

In summary, six features, i.e. *mw*, *md*, *lc*, *ag*, *hr* and *hm* are extracted from each face for the purpose of expression recognition.

After the features have been extracted, this becomes a general classification problem. Two pattern classification methods are applied to recognize the expression of the incoming faces.

    1.   Linear discriminant classifier: (Linear Discriminant Analysis-LDA)

LDA tries to find the subspace that best discriminates different classes by maximizing the between-class scatter matrix S_b, while minimizing the within-class scatter matrix S_w in the projective subspace. S_w and S_b are defined as follows,

$$S_w = \sum_{i=1}^{L} \sum_{\vec{x}_k \in X_i} (\vec{x}_k - \vec{m}_i)(\vec{x}_k - \vec{m}_i)^T \qquad (3)$$

$$S_b = \sum_{i=1}^{L} n_i (\vec{m}_i - \vec{m})(\vec{m}_i - \vec{m})^T \qquad (4)$$

Where $\vec{m}_i$ is the mean vector for the individual class X_i, and n_i is the number of samples in class X_i, $\vec{m}$ is the mean vector of all the samples. L is the number of classes.

The LDA subspace is spanned by a set of vectors W, satisfying

$$W = \arg\max \left| \frac{W^T S_b W}{W^T S_w W} \right| \qquad (5)$$

2.   Support Vector Machine (SVM):

Support vector machine is a relatively new technology for classification. It relies on preprocessing the data to represent patterns in a high dimension, typically much higher than the original feature space. With an appropriate nonlinear mapping to a sufficiently high dimension, data from two categories can always be separated by a hyperplane (Duda et al., 2001). In our research, the Libsvm program package (Chang and Lin, 2001) was used to implement the support vector machine.

5. 3D face recognition

5.1 Neutral face recognition

In previous related work, we have found that the central vertical profile and the face contour are both discriminant features for every person (Li et al., 2005a). Therefore, for neutral face recognition, the same method as in (Li et al., 2005b) is used. In this approach, the results of central vertical profile matching and contour matching are combined. The combination of the two classifiers improves the performance noticeably. The final similarity score for the probe image is the product of ranks for each of the two classifiers. The image which has the smallest score in the gallery will be chosen as the matching face for the probe image.

5.2 Smiling face recognition

For the recognition of faces labeled as 'smiling' by the expression recognition module, the probabilistic subspace method proposed by (Moghaddam and Pentland, 1995) is used. The following paragraphs provide an outline of this method and the relative principal component analysis (PCA).

Subspace methods are commonly used in computer vision, including face recognition. A raw 2D image can be represented as a vector in a high dimensional space. In most cases, however, the information which needs to be extracted has a much lower dimension. That is where subspace methods such as *principal component analysis* (PCA), or the previously

introduced *linear discriminant analysis* (LDA), can be applied to cope with the problem of reducing excessive dimensionality in the data to be analyzed.

PCA

Unlike LDA, which seeks a set of features that results in the best separation of each class, PCA seeks a projection that best represents the data in a least-square sense. In PCA, a set of vectors are computed from the eigenvectors of the sample covariance matrix C,

$$C = \sum_{i=1}^{M} (\vec{x}_i - \vec{m})(\vec{x}_i - \vec{m})^T \qquad (6)$$

where $\vec{m}$ is the mean vector of the sample set. The eigen space Y is spanned by k eigenvectors $u_1 \; u_2 \; ..... \; u_k$, corresponding to the k largest eigen values of the covariance matrix C.

$$\vec{y}_i = (\vec{x}_i - \vec{m})^T \left[\vec{u}_1 \vec{u}_2 .... \vec{u}_k \right] \qquad (7)$$

The dimensionality of vector $\vec{y}_i$ is $K \; (K<<M)$.

Probabilistic subspace method

In (Moghaddam and Pentland, 1995) (Moghaddam and Pentland, 1997), B. Moghaddam et al. presented an unsupervised technique for visual learning, which is based on density estimation in high dimensional spaces using an eigen decomposition. The probability density is used to formulate a maximum-likelihood estimation framework for visual search, target detection and automatic object recognition. Using the probabilistic subspace method, a multi-class classification problem can be converted into a binary classification problem.

Let Δ represents the difference between two vectors in a high dimensional subspace.

$$\Delta = I1 - I2 \qquad (8)$$

Δ belongs to the intrapersonal space in the high dimensional subspace if I1 and I2 are two different instances of the same subject; Δ belongs to the interpersonal or extrapersonal space if I1 and I2 are instances from different subjects. $S(\Delta)$ is defined as the similarity between I1 and I2. Using Bayes Rule,

$$S(\Delta) = P(\Omega_I \mid \Delta) = \frac{P(\Delta \mid \Omega_I)P(\Omega_I)}{P(\Delta \mid \Omega_I)P(\Omega_I) + P(\Delta \mid \Omega_E)P(\Omega_E)} \qquad (9)$$

$P(\Delta \mid \Omega_I)$ and $P(\Delta \mid \Omega_E)$ are the likelihoods of intrapersonal space and extrapersonal space. The likelihood function can be estimated by traditional means, i.e. maximum likelihood estimation or Parzen window estimation if there are enough data available. In most cases,

because of the high dimensionality of the subspace, training data are not sufficient. Subspace density estimation is another choice, which is the case in our experiment. $P(\Omega_I)$ and $P(\Omega_E)$ are *a priori* probabilities for intrapersonal and extrapersonal subspace. Thus, according to the maximum *a posteriori* (MAP) rule, if $P(\Omega_I | \Delta)$ is greater than $P(\Omega_E | \Delta)$, the two images are considered to be different instances of the same subject, otherwise, they belong to two subjects.

Another method based only on Ω_I can be used to simplify the computation. This maximum-likelihood (ML) similarity measure ignores extrapersonal variations.

$$S^{'}(\Delta) = P(\Delta | \Omega_I) \qquad (10)$$

In (B.Moghaddam 1995), it was found that the Ω_I density in (10) carries greater weight in modeling the posterior similarity used for MAP recognition. The extrapersonal Ω_E, on the other hand serves a secondary role and its accurate modeling is less critical. By dropping the Ω_E likelihood in favor of an ML similarity, the results typically suffer only a minor deficit in accuracy as compared to $S(\Delta)$.

Subspace density estimation

Given the high dimensionality of Δ, traditional methods are not suitable for the purpose of probability density estimation. An efficient subspace density estimation method proposed in (B.Moghaddam 1995; B.Moghaddam 1997) was used. The vector space of R^N is divided into two complementary subspaces: DIFS (Difference in Feature Space), F, and DFFS (Difference from Feature Space), $\overline{F}$, as show in the figure.
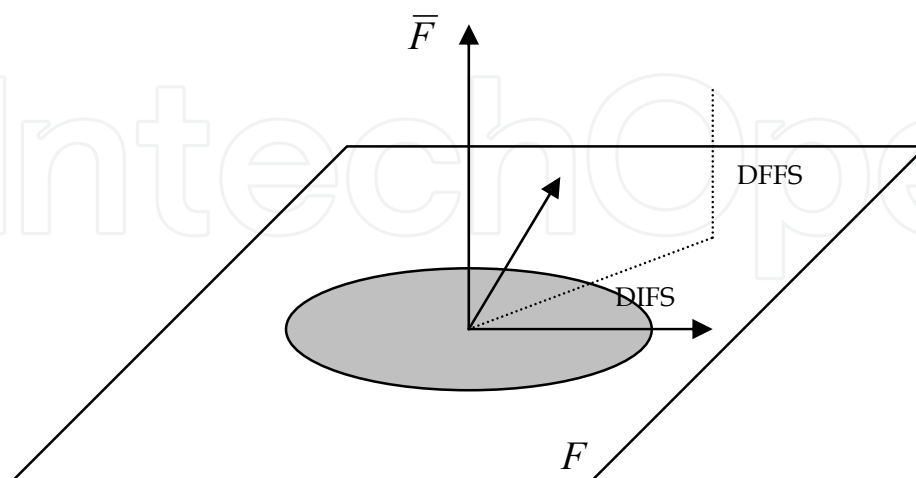
Fig. 6. The principal subspace F and its orthogonal complement $\overline{F}$ for a Gaussian density

F is spanned by the first M (M<<N) eigen vectors corresponding to the largest M eigen values of principal component decomposition results.

As derived in(B. Moghaddam 1995), the complete likelihood estimate can be written as the product of two independent marginal Gaussian densities

$$\hat{P}(\Delta \mid \Omega) = \left[\frac{\exp\left(-\frac{1}{2} \sum_{i=1}^{M} \frac{y_i^2}{\lambda_i}\right)}{(2\pi)^{\frac{M}{2}} \prod_{i=1}^{M} \lambda_i^{1/2}} \right] \left[\frac{\exp\left(-\frac{\varepsilon^2(\Delta)}{2\rho}\right)}{2\pi\rho^{(N-M)/2}} \right] = P_F(\Delta \mid \Omega)\hat{P}_{\bar{F}}(\Delta \mid \Omega; \rho) \quad (11)$$

where $P_F(\Delta \mid \Omega)$ is the true marginal density in *F*, $\hat{P}_{\bar{F}}(\Delta \mid \Omega; \rho)$ is the estimated marginal density in the orthogonal complement $\bar{F}$, y_i are the principal components and $\varepsilon^2(\Delta)$ is the PCA residual . From (B.Moghaddam 1995), the optimal value for ρ is the average of the $\bar{F}$ eigen values.

$$\rho = \frac{1}{N-M} \sum_{i=M+1}^{N} \lambda_i \quad (12)$$

In the experiment for smiling face expression recognition, because of the limited number of subjects (30), the central vertical profile and the contour are not used directly as vectors in a high dimensional subspace. Instead, they are down sampled to a dimension of 17 for the analysis. The dimension of subspace F is set to be 10, which contains approximately 97% of the total variance. The dimension of complementary subspace $\bar{F}$ is 7. In this case also, independent ranks are computed for the central profile and the contour of each gallery face. The overall rank is found by sorting the product of these two ranks and is used to determine the final recognition result.

6. Experiments and Results

One gallery and three probe databases are formed for the evaluation of our methods. The gallery database has 30 neutral faces, one for each subject, acquired in the first data acquisition session. Three probe sets are constituted as follows and used in experiments 2 and 3.

Probe set 1: 30 neutral faces acquired in the second session.
Probe set 2: 30 smiling faces acquired in the second session.
Probe set 3: 60 faces, constituted by probe set 1 and probe set 2.
The following experiments are undertaken:

Experiment 1: Testing the expression recognition module

The leave-one-out cross validation method is used to test the expression recognition classifier. Every time, the faces collected from 29 subjects in both data acquisition sessions are used to train the classifier and the four faces of the remaining subject collected in both

sessions are used to test the classifier. The results shown below are the average of the 30 recognition rates. Two classifiers are used. One is the linear discriminant classifier; the other is a support vector machine classifier. They have similar performance of over 90% recognition rate.

Methods	LDA	SVM
Expression recognition rate	90.8%	92.5%

Table 1. Expression recognition rate

Experiment 2: Testing the neutral and smiling recognition modules separately

In the first two sub experiments, probe faces are directly fed to the neutral face recognition module. In the third sub experiment leave-one-out cross validation is used to verify the performance of the smiling face recognition module. In each cycle, 29 subjects' faces from both acquisition sessions are used for the training and the remaining subject's smiling face from the second session is used as testing face.

    2.1 Neutral face recognition: probe set 1 is used. (neutral face recognition module is used)

    2.2 Neutral face recognition: probe set 2 is used. (neutral face recognition module is used)

    2.3 Smiling face recognition: probe set 2 is used. (smiling face recognition module is used)
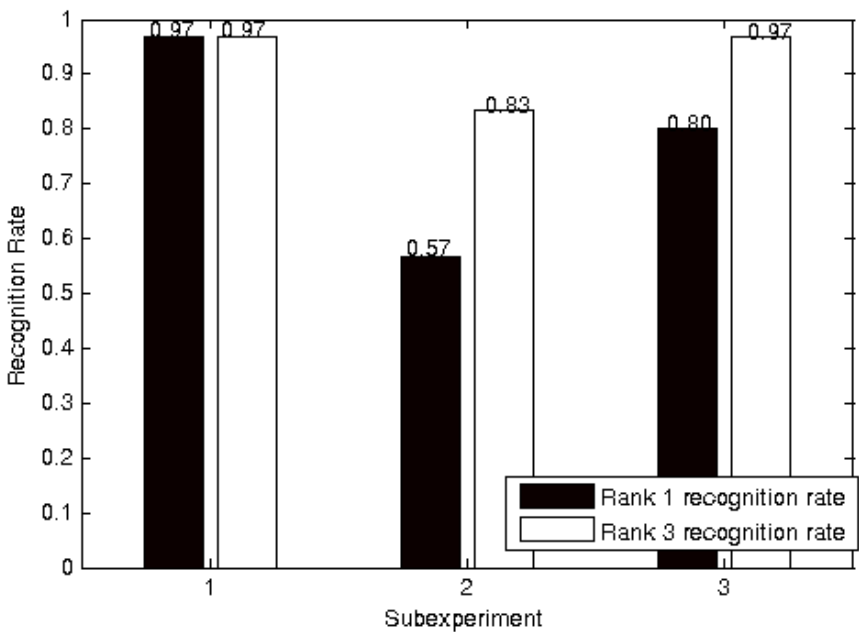

Fig. 7. Results of Experiment 2

From Figure 7, it can be seen that when the incoming faces are all neutral, the algorithm which treats all the faces as neutral achieves a very high recognition rate. On the other hand,

if the incoming faces are smiling faces, then the neutral face recognition algorithm does not perform well, only 57% rank one recognition rate is obtained. In contrast, when the smiling face recognition algorithm is used to deal with smiling faces, the recognition rate can go back as high as 80%.

Experiment 3: Testing a realistic scenario

This experiment emulates a realistic situation in which a mixture of neutral and smiling faces (probe set 3) must be recognized. Sub experiment 1 investigates the performance obtained if the expression recognition front end is bypassed, and the recognition of all the probe faces is attempted with the neutral face recognition module alone. The last two sub experiments implement the full framework shown in Figure 1. (Faces are first sorted according to expression and then routed to the appropriate recognition module.) In 3.2 the expression recognition is performed with the linear discrimant classifier, while in 3.3 it is implemented through the support vector machine approach.

3.1 Neutral face recognition: probe 3 is used. (Probe 3 is treated as neutral faces.)
3.2 Integrated expression and face recognition: probe 3 is used. (Linear discriminate classifier for expression recognition.)
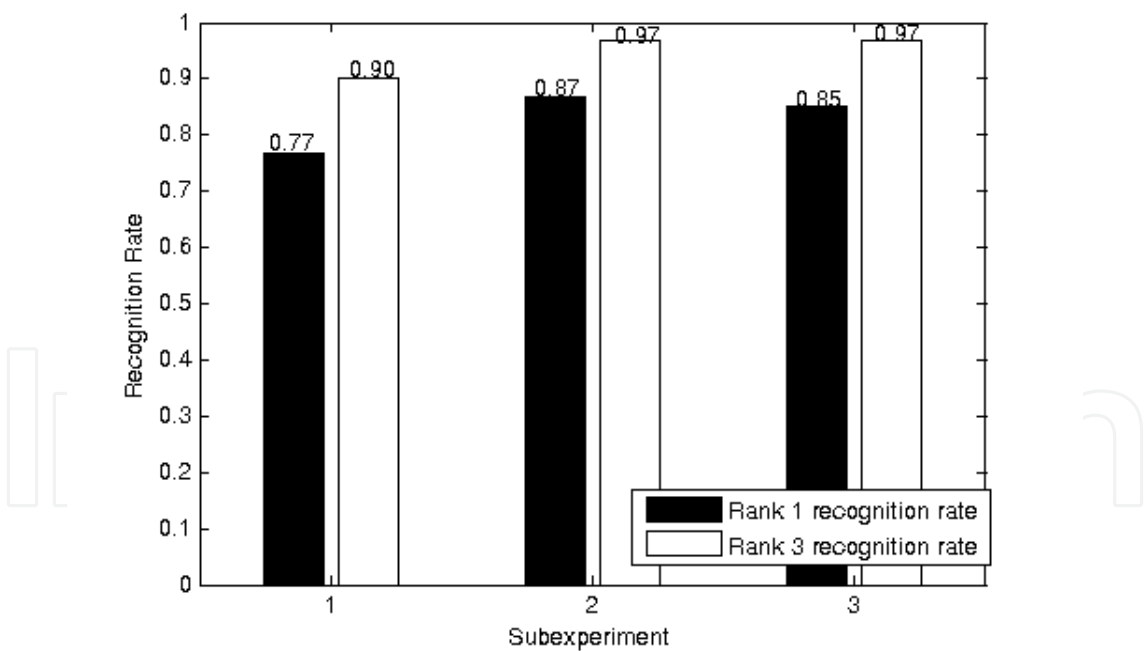3.3 Integrated expression and face recognition: probe 3 is used. (Support vector machine for expression recognition.)

Fig. 8. Results of Experiment 3

It can been seen in Figure 8, that if the incoming faces include both neutral faces and smiling faces the recognition rate can be improved about 10 percent by using the integrated framework proposed here.

7. Discussion and Conclusion

7.1 Discussion

Experiment 1 was aimed at determining the level of performance of the Facial Expression Recognition Module, by itself. Using the leave-one-out cross validation approach, 30 different tests were carried out (Each using 29 x 2 neutral faces and 29 x 2 smiling faces for training). The average success rate in identifying the expressions of the face belonging to the subject not used for training, in each case, was 90.8% with LDA and 92.5% when SVM was used. This confirms the capability of this module to successfully sort these two types of faces (neutral vs. smiling). Both algorithms were applied on the six facial features obtained from the range images (*mw*, *md*, *lc*, *ag*, *hr* and *hm*). Using these features, the actual choice of algorithm used to separate neutral from smiling faces did not seem to be critical.

Experiment two was carried out to test one of the basic assumptions behind the framework proposed (Figure 1). That is, a system meant to recognize neutral faces may be successful with faces that are indeed neutral, but may have much less success when dealing with faces displaying an expression, e.g., smiling faces. This differentiation was confirmed by the high rank-one recognition (97%) achieved by the Neutral Face Recognition Module for neutral faces (probe set 1) in subexperiment 1, which was in strong contrast with the much lower rank-one recognition rate (57%) achieved by this same module for smiling faces (probe set 2), in subexperiment 2. On the other hand, in the third subexperiment we confirmed that a module that has been specifically developed for the identification of individuals from smiling probe images (probe set 2) is clearly more successful in this task (80% rank-one recognition).

Finally, Experiment 3 was meant to simulate a more practical scenario, in which the generation of probe images does not control the expression of the subject. Therefore for all three subexperiments in Experiment 3 we used the comprehensive probe set 3, including one neutral range image and one smiling range image from each of the subjects. In the first subexperiment we observe the kind of results that could be expected when these 60 probe images are processed by a "standard" Neutral Face Recognition Module alone, which is similar to several of the contemporary approaches used for 3D face recognition. Unfortunately, with a mix of neutral and smiling faces this simple system only achieves a 77% rank-one face recognition (much lower than the 97% obtained for probe set 1, made up of just neutral faces, in Experiment 2). This result highlights the need to account for the possibility of a non-neutral expression in 3D face recognition systems. On the other hand, in sub experiments two and three we apply the same mixed set of images (Probe set 3) through the complete process depicted in our proposed framework (Figure 1). That is, every incoming image is first sorted by the Facial Expression Recognition Module and accordingly routed to either the Neutral Face Recognition Module or the Smiling Face Recognition Module, where the identity of the subject is estimated. The right-most four columns in Figure 8 show that, whether using the linear discriminant analyzer or the support vector machine for the initial expression sorting, the rank-one face recognition levels achieved by the overall system are higher (87%, 85%).

In reviewing the results of these experiments, it should be noted that all the experiments involving smiling faces are done using the leave-one-out cross validation method because of

the size of the database. Therefore the results displayed are the average, not the best one. For simplicity of implementation, the training samples for the expression recognition system and the smiling face recognition systems are the same faces. In a real application, we would select the training samples to make the best classifier for expression recognition and the identification of faces with a type of expression separately. Considerable performance improvement might be achieved in this way.

7.2 Conclusion

In this paper we have presented an alternative framework proposed to enhance the performance of 3D face recognition algorithms, by acknowledging the fact that the face of a subject is a deformable surface that undergoes significant changes when the subject displays common expressions. Our main proposition is that, instead of ignoring the possibility of significant facial changes due to expressions, 3D face recognition systems should account for the potential presence of an expression in their probe images. In particular our suggested framework requires the development of two new functional modules, in addition to a "standard" face recognition module for neutral faces:

- A Facial Expression Recognition Module, capable to "tag" an incoming probe image with an appropriate expression label, and route it to an appropriate "specialized" face recognition classifier (matched to the expression found in the probe face), where the identity of the subject will be estimated.
- "Specialized" Face Recognition Classifiers that are trained to identify faces with expressions other than "neutral"

In this work we have developed the framework for the simplest case in which we consider only neutral and "smiling" faces, as one very common form of expression, frequently displayed by people in public.

Our experimentation with this implementation of the framework, using 3 sets of probe images has revealed that:

- It is possible to implement an appropriate module for the sorting of neutral vs. smiling 3D face images, based on classification of six facial features we have defined, and utilizing either the linear discriminant analysis or the support vector machine approaches (Experiment 1).
- While a contemporary "neutral" face classifier is capable of achieving a good performance, (97% rank-one recognition), when identifying neutral 3D faces, the performance of this same classifier is much weaker (57% rank-one recognition) when dealing with otherwise comparable "smiling" faces. (Experiment 2).
- It is feasible to develop a "specialized" classifier that will identify "smiling" faces with a reasonable level of success (80% rank-one recognition), which is clearly higher than the performance of the "neutral" face classifier for the same scenario (Experiment 2).
- A system that follows the complete framework proposed (Figure 1) is better able to identify subjects from a mixture of neutral and smiling 3D faces (87% and 85% rank-one recognition) than a standard 3D face recognition system (77% one-rank

recognition) that relies on the assumption that the subjects are expressionless during the capture of the probe images (Experiment 3).

The work reported in this paper represents an attempt to acknowledge and account for the presence of expression on 3D face images, towards their improved identification. In comparison with other methods that pursue similar goals [1, 19, 32], the method introduced here is computationally efficient. Furthermore, this method also yields as a secondary result the information of the expression found in the faces.

Based on these findings we believe that the impact of expression on 3D face recognition and the development of systems that account for it, such as the framework introduced here, will be keys to future enhancements in the field of 3D Automatic Face Recognition.

8. References)

Bartlett, M., Lades, H., Sejnowski, T., 1998. Independent Component Representations for Face Recognition. Proceedings of SPIE.

Bolle, R., Connell, J., Pankanti, S., Ratha, N., Senior, A., 2003. Guide to Biometrics, Springer.

Bowyer, K., Chang, K., Flynn, P., 2004. A Survey of Approaches to 3D and Multi-Modal 3D+2D Face Recognition. Proceedings of IEEE International Conference on Pattern Recognition. pp. 358-361.

Bronstein, A., Bronstein, M., Kimmel, R., 2003. Expression-invariant 3D face recognition. Proceedings of Audio & Video-based Biometric Person Authentication (AVBPA), pp. 62-69.

Bruce, V., Young, A., 1986. Understanding face recognition. British Journal of Psychology **77**: 305-327.

Cartoux, J., LaPreste, J., Richetin, M.,1989. Face authentication or recognition by profile extraction from range images. Proceedings of the Workshop on Interpretation of 3D Scenes. pp.194-199.

Chang, C, Lin, C., 2001. LIBSVM: a library for support vector machines.

Chang, K., Bowyer, K., Flynn, P., 2003. Multimodal 2D and 3D biometrics for face recognition., 2003. IEEE International Workshop on Analysis and Modeling of Faces and Gestures. pp.187-194.

Chang, K., Bowyer, K., Flynn, P., 2005. Effects of facial expression in 3D face recognition. SPIE 5779: Biometric Technology for Human Identification II.

Chellappa, R., Wilson, C., Sirohey, S., 1995. Human and Machine Recognition of Faces: A Survey. Proceedings of the IEEE **83(5)**: 705-740.

Cholkopf, B., Smola, A., Muller, K., 1998. Nonlinear Component Analysis as a Kernel Eigenvalue Problem. Neural Computation **10**(5): 1299-1319.

Duda, R., Hart, P., Stork D., 2001 Pattern Classification.

Edwards, G., Cootes, T., Taylor, C., 1998. Face Recognition Using Active Appearance Models. Proceedings of ECCV, pp. 581-695.

Ekman, P., Friesen, W., 1971. Constants across cultures in the face and emotion. Journal of Personality and Social Psychology **17**(2): 124-129.

Etcoff, N., Magee, J., 1992. Categorical perception of facial expressions. Cognition **44**: 227-240.

Fasel, B., Luettin, J., 2003. Automatic facial expression analysis: a survey. <u>Pattern Recognition</u> **36**: 259-275.

Gordon, G., 1991. Face recognition based on depth maps and surface curvature. Geometric Methods in Computer Vision, SPIE 1570: pp. 1-12.

Hansch, E., Pirozzola, F., 1980. Task Relevant Effects on the Assessment of Cerebral Specialization for Facial Emotions. Brain and Language **10**: 51-59.

Hay, D., Young A., Ellis, A.,1991. Routes through the face recognition system. Q. J. Exp. Psychol A-Human Exp. Psy. **43**: 761-791.

Hesher, C., Srivastava, A., Erlebacher G., 2003. A novel technique for face recognition using range images. Seventh International Symposium on Signal Processing and Its Application.

Li, C., Barreto, A., 2005. Profile-Based 3D Face Registration and Recognition. Lecture Notes on Computer Science **3506**: 484-494.

Li, C., Barreto, A., Zhai, J., Chin C., 2005. Exploring Face Recognition Using 3D Profiles and Contours. Proceedings of IEEE SoutheastCon 2005. pp.576-579.

Lu, X., Jain, A., 2005. Deformation Analysis for 3D Face Matching. 7th IEEE Workshop on Applications of Computer Vision (WACV'05).

Moghaddam, B., Pentland, A., 1995. Probabilistic Visual Learning for Object Detection. Proceedings of International Conferrence of Computer Vision (ICCV' 95), pp.786-793.

Moghaddam, B., Pentland, A., 1997. Probabilistic Visual Learning for Object Representation. IEEE Trans. on Pattern Analysis and Machine Intelligence **19**(7): pp. 696-710.

Nagamine, T., Uemura, T., Masuda, I., 1992. 3D facial image analysis for human identification. Internaitonal Conference on Pattern Recognition (ICPR,1992).pp. 324-327.

Ortega-Garcia, J., Bigun J., Reynolds, D., Gonzales-Rodriguez, J., 2004. Authentication gets personal with biometrics. IEEE Signal Processing **21**(No.2): 50-61.

Phillips, J., Grother, P., Blackburn, D., Bone, M., Michaels, R., 2002. Face Recognition Vendor Test 2002.

Turk, M., Pentland, A., 1991. Eigenfaces for Recognition. Jounal of Cognitive Neuroscience **3**(1): 71-86.

Wiskott, L., Fellous, J., Krger, N., Von der Malsburg, C., 1997. Face Recognition by Elastic Bunch Graph Matching. IEEE Trans. on Pattern Analysis and Machine Intelligence **19**(7): 775-779.

Zhao, W., Chellappa, R., Rosenfeld, A., 2003. Face recognition: a literature survey. ACM Computing Survey **35**: 399-458.

Face Recognition

Edited by Milos Oravec

This book aims to bring together selected recent advances, applications and original results in the area of biometric face recognition. They can be useful for researchers, engineers, graduate and postgraduate students, experts in this area and hopefully also for people interested generally in computer science, security, machine learning and artificial intelligence. Various methods, approaches and algorithms for recognition of human faces are used by authors of the chapters of this book, e.g. PCA, LDA, artificial neural networks, wavelets, curvelets, kernel methods, Gabor filters, active appearance models, 2D and 3D representations, optical correlation, hidden Markov models and others. Also a broad range of problems is covered: feature extraction and dimensionality reduction (chapters 1-4), 2D face recognition from the point of view of full system proposal (chapters 5-10), illumination and pose problems (chapters 11-13), eye movement (chapter 14), 3D face recognition (chapters 15-19) and hardware issues (chapters 19-20).

How to reference

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

INTECH
open science | open minds