Chapter 17
Face Recognition Using 3D Images

I.A. Kakadiaris, G. Passalis, G. Toderici, E. Efraty, P. Perakis, D. Chu, S. Shah, and T. Theoharis

17.1 Introduction

Our face is our password—face recognition promises to revolutionize the way we identify individuals in a nonintrusive and convenient manner. Even though research in face recognition has spanned over nearly three decades, only 2D systems, with limited adoption to practical applications, have been developed so far. The primary reason behind this is the low accuracy of 2D face recognition systems in the presence of: (i) pose variations between the gallery and probe datasets, (ii) variations in lighting, and (iii) variations in the presence of expressions and/or accessories. The above conditions generally arise when noncooperative subjects are involved, which is the very case that demands accurate recognition.

Face recognition using 3D images was introduced in order to overcome these challenges. It was partly made possible by significant advances in 3D scanner technology. However, even 3D face recognition has faced significant challenges which have hindered its adoption for practical applications. The main problem of 3D face recognition is the high cost and fragility of 3D scanners. Over the last seven years, our research team has focused on exploring the usefulness of 3D data and the development of models for face recognition (under the general name URxD).

In this chapter, we present advances that aid in overcoming the challenges encountered in 3D face recognition. First, we present a fully automatic 3D face recognition system, UR3D, which has been proven to be robust under variations in expressions. The fundamental idea of this system is the description of facial data using an Annotated Face Model (AFM). The AFM is fitted to the facial scan using

I.A. Kakadiaris (✉) · G. Passalis · G. Toderici · E. Efraty · P. Perakis · D. Chu · S. Shah · T. Theoharis
Computational Biomedicine Lab, Department of Computer Science, University of Houston, Houston, TX 77204, USA
e-mail: ioannisk@grip.cis.upenn.edu

G. Passalis · P. Perakis · T. Theoharis
Computer Graphics Laboratory, Department of Informatics and Telecommunications, University of Athens, Ilisia 15784, Greece

S.Z. Li, A.K. Jain (eds.), *Handbook of Face Recognition*,
DOI 10.1007/978-0-85729-932-1_17,

a subdivision-based deformable model framework. The deformed model captures the details of an individual's face and represents this 3D geometry information in an efficient 2D representation by utilizing the model's parametrization. This representation is analyzed in the wavelet domain and the associated wavelet coefficients define the metadata that are used for comparing the different subjects. These metadata are both compact and descriptive. This approach that involves geometric modeling of the human face allows greater flexibility, better understanding of the face recognition issues, and requires no training.

Second, we demonstrate how pose variations are handled in 3D face recognition. The 3D scanners that are used to obtain facial data are usually nonimmersive which means that only a partial 3D scan of the human face is obtained, particularly so in noncooperative, practical conditions. Thus, there are often missing data of the frontal part of the face. This can be overcome by identifying a number of landmarks on each 3D facial scan thereby allowing correct registration with the AFM, independent of the original pose of the face. For nonfrontal scans, missing data can be added by exploiting facial symmetry, assuming that at least half of the face is visible. This is achieved by improving the subdivision-based deformable model framework to allow symmetric fitting. Symmetric fitting alleviates the missing data problem and facilitates the creation of geometry images that are pose invariant. Another alternative to tackle the missing data problem is to attempt recognition based on the facial profile; this approach is particularly useful in recognizing car drivers from side view images. In this approach, the gallery includes facial profile information under different poses, collected from subjects during enrollment. These profiles are generated by projecting the subjects' 3D face data. Probe profiles are extracted from the input images and compared to the gallery profiles.

Finally, we demonstrate how the problems related to the cost of 3D scanners can be mitigated through hybrid systems. Such systems employ 3D scanners for the enrollment of subjects, which can take place in a few specialized locations, and 2D cameras at points of authentication, which can be multiple and dispersed. It is practical to adopt this approach if hybrid systems can improve the accuracy of a 2D system. During enrollment, 2D+3D data (2D texture and 3D shape) are used to build subject-specific annotated 3D models. To achieve this, an AFM is fitted to the raw 2D+3D data using a subdivision-based deformable framework. A geometry image representation is then extracted using the parametrization of the model. During the verification phase, a single 2D image is used as the input to map the subject-specific 3D AFM. Given the pose in the 2D image, an Analytical Skin Reflectance Model (ASRM) is then applied to the gallery AFM to transfer the lighting from the probe to the texture in the gallery. The matching score is computed using the relit gallery texture and the probe texture. This hybrid method surpasses the accuracy of 2D face recognition system in difficult datasets.

17.1.1 3D Face Recognition

In recent years, several 3D face recognition approaches have been proposed that offer increased accuracy and resilience to pose and illumination variations when compared to 2D approaches. The limitations of 2D approaches were highlighted in the Face Recognition Vendor Test 2002 study. However, the advantages of 3D face recognition were not evident since most 3D approaches had not been extensively validated due to the non-availability of 3D databases. This is evident in the surveys of the 3D face recognition field given by Bowyer et al. [8], Chang et al. [13] and Scheenstra et al. [57]. To address this issue, NIST introduced the Face Recognition Grand Challenge and Face Recognition Vendor Test 2006 [21] and released two publicly available multimodal (3D and 2D) databases, FRGC v1 and FRGC v2.

On FRGC v1, a database that contains over 900 frontal scans without any facial expressions, Pan et al. [46] reported 95% rank-one recognition rate using a PCA approach, while Russ et al. [56] reported a 98% verification rate. Our approach achieved a 99% rank-one recognition rate [29].

On FRGC v2, a database that contains over 4000 frontal scans with various facial expressions, Chang et al. [11, 12] examined the effects of facial expressions using two different 3D recognition algorithms. They reported a 92% rank-one recognition rate. The same rank-one recognition rate (92%) was also reported by Lu et al. [40]. In their approach, a Thin Plate Spline (TPS) was used to learn expression deformation from a control group of neutral and non-neutral scans. Husken et al. [28] presented a multimodal approach that uses hierarchical graph matching (HGM). They extended their HGM approach from 2D to 3D but the reported 3D performance was poorer than the 2D equivalent. The fusion of the two approaches, however, provided competitive results, a 96.8% verification rate at 0.001 False Acceptance Rate (FAR), compared to 86.9% when using the 3D only. Al-Osaimi et al. [1] used a PCA subspace, referred to as the expression deformation model, to analyze facial deformations from 3D data. They reported an average (over ROC I, II and III experiments) verification rate of 94.2% at 0.001 FAR. Maurer et al. [43] also presented a multimodal approach tested on the FRGC v2 database, and reported a 87% verification rate at 0.01 FAR. In our initial work on this database [49], we analyzed the behavior of our approach in the presence of facial expressions. The improvements presented in our subsequent work [30] allowed us to overcome the shortcomings of this approach. Our method, using only 3D data, achieved 97% rank-one recognition and an average (over ROC I, II and III experiments) verification rate of 97.1% at 0.001 FAR.

17.1.2 3D Face Recognition from Partial Scans: UR3D-PS

Even though the majority of the 3D face recognition approaches focus on full frontal scans, there are several approaches that focus on partial scans (that are prone to missing data). Lu et al. [38, 39, 41], in a series of studies, presented methods to locate

the positions of the corners of the eyes and mouth, and the tips of the nose and chin, based on a fusion scheme of shape index on range maps and the "cornerness" response on intensity maps. They also developed a heuristic method based on cross-profile analysis to locate the nose tip more robustly. Candidate landmark points were filtered out using a static (nondeformable) statistical model of landmark positions. Although they report a 90% rank-one matching accuracy in an identification experiment, no claims where made with respect to the effects of pose variations.

Dibeklioglu et al. [17, 18] introduced a nose tip localization and segmentation method using curvature-based heuristic analysis to enable pose correction in a face recognition system that allows identification under significant pose variations. However, their system cannot handle facial scans with yaw rotations greater than 45°. Additionally, even though the Bosphorus database that was used consisted of 3396 facial scans, the date were obtained from only 81 subjects.

Blanz et al. [5, 6] presented an approach in which a 3D Morphable Model was fitted on 3D facial scans, which is a well-established approach for producing 3D synthetic faces from scanned data. However, face recognition testing was validated on the FRGC database that consists of frontal facial scans, and on the FERET database that contains faces under pose variations which do not exceed 40°. Bronstein et al. [10] presented a face recognition method that is capable of handling missing data. This was an extension of their previous approach [9] where they deformed the face by embedding it into a multi-dimensional space. Such an approach preserves only the intrinsic geometries of face. Since facial expressions are mainly extrinsic geometries, the result is an expression invariant representation (canonical form) of the face. They reported high recognition rates, but on a limited database of 30 subjects. Also, the database did not contain side scans. Furthermore, the scans that contained missing data were derived synthetically by randomly removing certain areas from frontal scans. In Nair and Cavallaro's [45] work on partial 3D face matching, the face was divided into areas and only certain areas were used for registration and matching. This approach was based on an assumption that the areas of missing data can be excluded. Using a database of 61 subjects, they showed that using parts of the face rather than the whole face, yields higher recognition rates. This approach, as well as their subsequent work on 3D landmark detection, cannot be applied to missing data resulting from pose self-occlusion, especially when holes exist around the nose region. Lin et al. [36] introduced a coupled 2D and 3D feature extraction method to determine the positions of eye sockets using curvature analysis. The nose tip was considered as the extreme vertex along the normal direction of eye sockets. The method was used in an automatic 3D face authentication system but was tested on only 27 datasets with various poses and expressions. Mian et al. [44] introduced a heuristic method for nose tip detection and used it in a face recognition system. The method is based on a geometric analysis of the nose ridge contour projected on the x–y plane. It is used as a preprocessing step to crop and pose correct the facial data. Even though it allows up to 90° roll variation, this approach requires yaw and pitch variation less than 15°, thus limiting the applicability to near frontal scans. Perakis et al. [50] presented methods for detecting facial landmarks and used them to match partial facial data. Local shape and curvature analysis were used to locate

candidate landmark points (eye inner and outer corners, mouth corners, and nose and chin tips). The points were identified and labeled by matching them with a statistical facial landmark model. The method addresses the problem of extreme yaw rotations and missing facial areas, and its face recognition accuracy was validated against the FRGC v2 and UND Ear databases.

17.1.3 3D-aided 2D Face Recognition

The literature in 3D and 2D+3D Face Recognition has rapidly increased in recent years. An excellent survey was presented by Bowyer et al. [8]. The approach proposed by Riccio and Dugelay [55] uses geometric invariants on the face to establish a correspondence between the 3D gallery face and the 2D probe. Some of the invariants were manually selected. This algorithm does not use the texture information registered with the 3D data from the scanner, and hence, does not take full advantage of the input data. Blanz and Vetter [5] employed a morphable model technique to acquire the geometry and texture of faces from 2D images. Wang et al. [67] used a spherical harmonic representation [2] with the morphable model for 2D face recognition. Toderici et al. [61] proposed a method referred to as UR2D that uses 2D+3D data to build a 3D subject-specific model for the gallery. In contrast, Wang's method uses a 2D image to build a 3D model for the gallery based on a 3D statistical morphable model. Yin and Yourst [69] used frontal and profile 2D images to construct 3D shape models. In comparison to these methods, the UR2D method is able to more accurately model the subject identity as it uses both 2D and 3D information. Smith and Hancock [58] presented an approach for albedo estimation from 2D images also based on a 3D morphable model. The normals of the fitted model were then used for the computation of shading, assuming a Lambertian reflectance model. Biswas et al. [3] proposed a method for albedo estimation for face recognition using two-dimensional images. However, their approach was based on the assumption that the image does not contain shadows, and does not handle specular light. The relighting approach of Lee et al. [34], also suffers from the self-shadowing problem. Tsalakanidou [62] proposed a relighting method designed for face recognition but this approach produced images with poorer visual quality when compared to more generic methods, especially when specular highlights over-saturate the images.

17.1.4 3D-aided Profile Recognition

The use of face profile for identification had attracted research interest even before the arrival of the associated computer technologies [22]. The methods for recognition using the profile curve can be classified into one of two categories: *landmark-based* methods [27, 32, 37, 68] or *global* methods [23, 31, 47, 70]. Landmark-based methods rely on the attributes associated with a set of fiducial points, and recognition uses similarity metrics based on those attributes. Global methods consider each

profile as a geometric object and introduce a similarity metric between homogeneous objects: all regions of a profile are treated equally.

Harmon et al. [27] defined 17 fiducial points; after aligning two profiles based on the selected landmarks, the matching was achieved by measuring the Euclidean distance of the feature vectors derived from the outlines. A 96% recognition rate was reported. Wu et al. [68] used a B-spline to locate six landmarks and extracted 24 features from the resulting segments. Liposcak and Loncaric [37] used scale-space filtering to locate 12 landmarks and extracted 21 distances based on those landmarks. The Euclidean distance between the vectors of features was used for the identification.

Bhanu and Zhou [70] proposed curvature-based matching using a dynamic warping algorithm. They reported a recognition rate of almost 90% on the University of Bern Database that consisted of 30 subjects. Gao and Leung [24] introduced a method to encode profiles as attributed strings and developed an algorithm for attributed string matching. They reported nearly 100% recognition rate on the Bern database. Pan et al. [47] proposed a method that uses metrics for the comparison of probability density functions on properly rotated and normalized profile curves. Gao et al. [23, 25] proposed new formulations of the Hausdorff distance. Initially, their method was extended to match two sets of lines, while later, it was based on weighting points by their significance. In both cases, they applied their distance metric to measure the similarity of face profiles.

All these methods were designed for standard profiles only and use 2D images as gallery. Kakadiaris et al. [31] introduced the use of a 3D face model for the generation of profiles under different poses for the gallery. Modified directional Hausdorff distance of the probe profile to the gallery profile was used for identification. In addition, four different profiles under various rotation angles were used to introduce robustness to pose.

An important step in the implementation of a fully automatic system suitable for unconstrained scenarios is developing an accurate profile extractor. The majority of profile-based identification approaches do not sufficiently address this issue: instead they use manual extraction [31, 47] or very basic thresholding methods based on the assumption of indoor controlled illumination and a uniform background [7, 37, 70]. More efficient methods have been applied for near-frontal face extraction and feature localization. Among the most powerful are the methods based on the Active Shape Model (ASM), originally proposed by Cootes et al. [15]. These methods are based on recovering parameters of a statistical shape model, when a local minimum of the matching energy is found based on a search in local neighborhoods of the shape points. During the last decade, numerous modifications for the ASM have been proposed [26, 42]. The ultimate goal for most of these algorithms is alignment, therefore the shape is mostly defined by sparse set of common face landmarks visible on the frontal view, enforced by only a few additional points. For the contour extraction task, points should be densely sampled in order to approximate the curve accurately. Another known shortcoming of the ASM approach is the sensitivity to initialization, which is especially critical for ridge-like shapes.

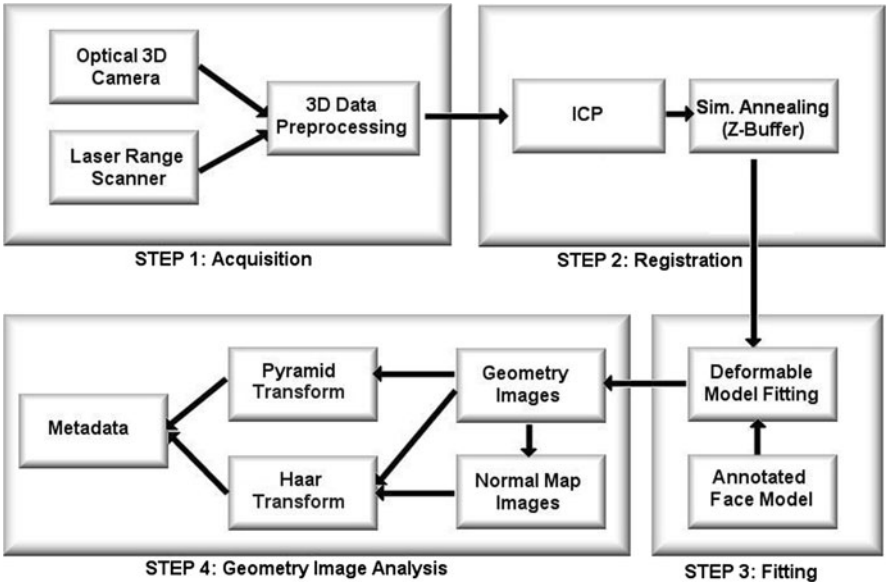


Fig. 17.1 Overview of the UR3D 3D face recognition method

17.2 3D Face Recognition: UR3D

The UR3D 3D face recognition method is reviewed in this section [30]. It is a purely geometric approach as it does not require any statistical training. The AFM is deformed to capture the shape of the face of each subject. This approach represents the 3D information in an efficient 2D structure by utilizing the AFM's UV parameterization. This structure is subsequently analyzed in the wavelet domain and the spectral coefficients define the final metadata that are used for comparison among different subjects.

This method has the following steps (Fig. 17.1):

1. *Acquisition:* Raw 3D data are acquired from the sensor and converted to a polygonal representation using sensor-dependent preprocessing.
2. *Registration:* The data are registered to the AFM using a two-phase approach.
3. *Deformable Model Fitting:* The AFM is fitted to the data using a subdivision-based deformable model framework.
4. *Geometry Image Analysis:* Geometry and normal map images are derived from the fitted AFM and wavelet analysis is applied to extract a reduced coefficient set as metadata (Fig. 17.2).

A detailed explanation of each step can be found at [30].

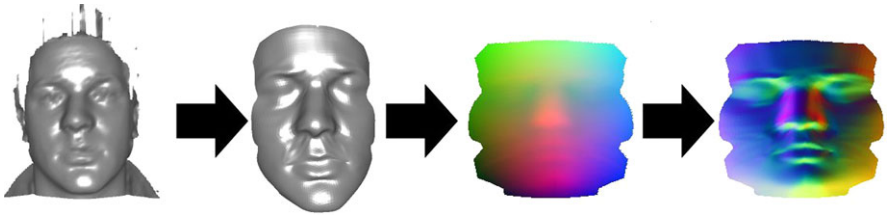


Fig. 17.2 From left to right: Facial scan → Fitted AFM → Extracted geometry image → Computed normal image

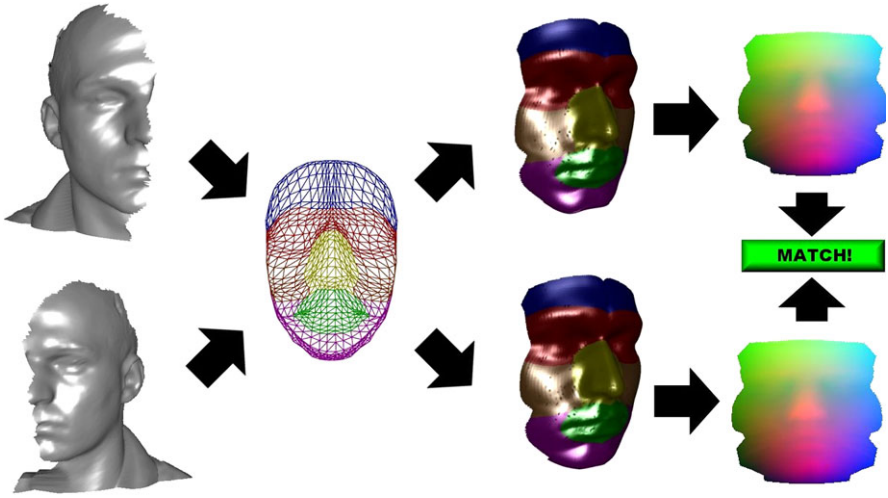


Fig. 17.3 Interpose matching using the proposed method (left to right): Opposite side facial scans with extensive missing data, Annotated Face Model (AFM), resulting fitted AFM of each scan (facial symmetry used), extracted geometry images

17.3 3D Face Recognition for Partial Scans: UR3D-PS

UR3D is focused on 3D frontal facial scans and does not handle extensive missing data. In this section, the focus is shifted to 3D partial scans with missing data (such as side facial scans with large yaw rotations). The goal is to handle both frontal and side scans seamlessly thus producing a biometric signature that is pose invariant and hence, making the method more suitable for real-world applications.

The main idea of the proposed method is presented in Fig. 17.3. It allows matching among interpose facial scans and solves the missing data problem by using facial symmetry. To this end, a registration step is added that uses an automated 3D landmark detector to increase the resiliency of the registration process to large yaw rotations (common in side facial scans). Additionally, the subdivision-based deformable model framework is extended to allow symmetric fitting. Symmetric fitting alleviates the missing data problem as it derives geometry images from the AFM that

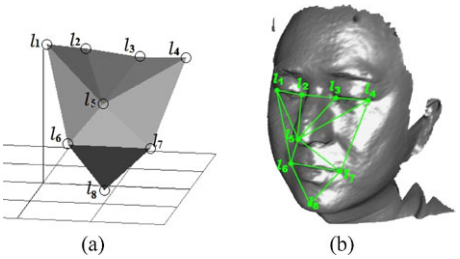


Fig. 17.4 Depiction of: **a** landmark model as a 3D object; and **b** landmark model overlaid on a facial scan

are pose invariant. Compared to the method presented in the previous section all other steps, except the registration and fitting step, remain unchanged. However, to make interpose matching more accurate, frontal facial scans are handled as a pair of independent side facial scans (left and right).

17.3.1 3D Landmark Detection

The proposed method (UR3D-PS) employs an improved version of the 3D landmark detection algorithm presented in [51]. Candidate interest points are extracted from the facial scans and are subsequently identified and labeled as landmarks by using a Facial Landmark Model (FLM). A set of 8 anatomical landmarks is used: right eye outer corner (l_1), right eye inner corner (l_2), left eye inner corner (l_3), left eye outer corner (l_4), nose tip (l_5), mouth right corner (l_6), mouth left corner (l_7) and chin tip (l_8) (Fig. 17.4). Note that at least five of these landmarks are always visible on side facial scans. The model with the entire set of eight landmarks will be referred to as FLM8 while the models with the reduced sets of five landmarks (left and right) will be referred to as FLM5L and FLM5R, respectively.

To create each FLM a mean shape is computed from a manually annotated training set. One hundred and fifty frontal facial scans with neutral expressions are randomly chosen from the FRGC v2 database as the training set. *Procrustes Analysis* [14, 19, 59] procedure is used to align the landmarks shape and calculate the mean shape. Subsequently, the variations of each FLM are analyzed by applying Principal Component Analysis (PCA) to the aligned landmark shapes. Aligned shape vectors form a distribution in the nd dimensional shape space, where n is the number of landmarks and d the dimension of each landmark. As described by Cootes et al. [16, 59], we can decompose this distribution and select the most significant eigenvectors of the eigenspace (*principal components*). We incorporated 15 eigenvalues (out of 24) in FLM8, which represent 99% of total shape variations of the frontal landmark shapes. Similarly, we incorporated 7 eigenvalues (out of 15) in FLM5L and FLM5R, which represent 99% of total shape variations of the left and right side landmark shapes.

The FLMs are used to detect landmarks in each facial scan as follows (depicted in Fig. 17.5):

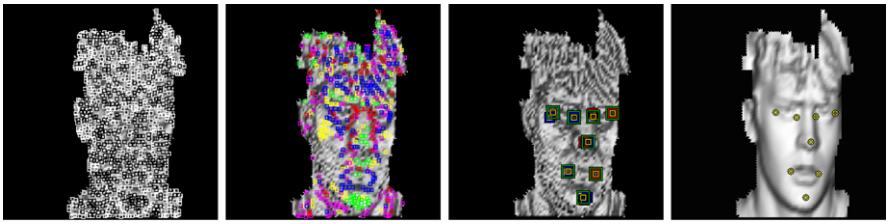


Fig. 17.5 Results of landmark detection and selection process: **a** shape indexes maxima and minima; **b** spin image classification; **c** extracted best landmark sets; and **d** resulting landmarks

- *Extract candidate landmarks by using the Shape Index map*. After computing shape index values on a 3D facial scan, mapping to 2D space is performed to create the shape index map. Local maxima (Caps) are candidate landmarks for nose tips and chin tips and local minima (Cups) for eye corners and mouth corners. The most significant subset of points for each group (Caps and Cups) is retained.
- *Classify candidate landmarks by using Spin Image templates*. Candidate landmarks from the previous step are classified and filtered according to their relevance with five Spin Image templates. The similarity between two spin image grids P and Q is expressed by the normalized linear correlation coefficient:

$$S(P, Q) = \frac{N \sum p_i q_i - \sum p_i \sum q_i}{\sqrt{[N \sum p_i^2 - (\sum p_i)^2][N \sum q_i^2 - (\sum q_i)^2]}}$$

  where p_i, q_i denotes each of the N elements of spin image grids P and Q, respectively.
- *Label Landmarks*. Using the classified candidate landmarks, feasible combinations of five landmarks are created. Subsequently, the rigid transformation that best aligns these combinations with the corresponding FLMs is computed. If the result is not consistent with FLM5L or FLM5R then the combination is filtered out. If it is consistent, the landmarks are labeled by the corresponding FLM and the combination is considered a possible solution. Possible solutions also include combinations of eight landmarks that are created from fusing two combinations of five landmarks (FLM5L and FLM5R) and are consistent with FLM8.
- *Select Final Solution*. The optimal solution (landmark combination) for each of the FLM5R, FLM5R and FLM8 is selected based on the distance from the mean shape of the corresponding FLM. To select the final solution the three optimal landmark combinations are compared using a normalized Procustes Distance that takes into consideration the shape space dimensions.

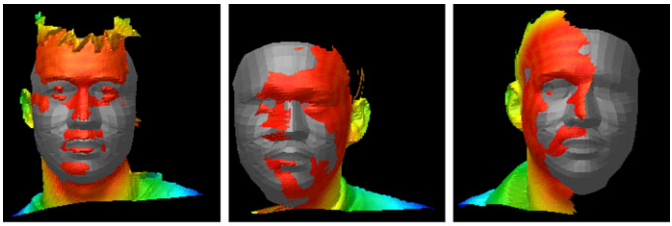


Fig. 17.6 AFM (*gray*) and facial scans (color coding: *red* means low registration error, *blue* means high registration error) superposed after registration (the scans): **a** frontal scan; **b** 45° left side scan; and **c** 60° right side scan

17.3.2 Partial Registration

Side facial scans with missing data cannot be registered robustly using the registration module of UR3D. To compute a rough but robust registration between the AFM and frontal or side facial scans (Fig. 17.6), the detected 3D landmarks are used. The Procrustes distance between a set of landmark points $\mathbf{x}$ on the scan and the corresponding landmark points $\mathbf{x}_0$ on the AFM is minimized in an iterative approach. If $\mathbf{T}$ translates $\mathbf{x}$ so that its centroid is at the origin (0,0,0), $\mathbf{T}_0$ translates $\mathbf{x}_0$ so that its centroid is at the origin (0,0,0), and $\mathbf{R}$ is an optimal rotation that minimizes the Procrustes distance of $\mathbf{x}$ to the reference shape $\mathbf{x}_0$, then, the final transformation to register a facial scan with vertices $\mathbf{v}_i$ to the AFM is:

$$\mathbf{v}_i' = \mathbf{T}_0^{-1} \cdot \mathbf{R} \cdot \mathbf{T} \cdot \mathbf{v}_i$$

and pose is estimated from $\mathbf{R}$. The landmark set detected on a facial scan (frontal, right or left) determines which of the FLM8, FLM5R and FLM5L will be used. However, in practice when a frontal scan is detected, we do not use the FLM8, but we consider it as a pair of side scans (and compute two independent registrations using FLM5R and FLM5L).

To fine-tune the registration we use Simulated Annealing. Note that for side scans, only one half of the model's z-buffer is used in the objective function. The other half is excluded as it would have been registered with areas that may contain missing data. The landmark detection algorithm effectively substitutes the ICP in the registration process. Therefore, the Simulated Annealing algorithm is only allowed to produce limited translations and rotations and cannot alleviate registration errors caused by erroneous landmark detection.

17.3.3 Symmetric Deformable Model Fitting

We have modified the fitting module of UR3D to incorporate the notion of *symmetric fitting* in order to handle missing data. The framework can now handle the left

and right sides of the AFM independently. The idea is to use the facial symmetry to avoid the computation of the external forces on areas of possible missing data. The internal forces are not affected and remain unmodified to ensure the continuity of the fitted surface. As a result, when fitting the AFM to facial scans classified as left side (from the previous step), the external forces are computed on the left side of the AFM and mirrored to the right side (and vice versa for right side scans). Therefore, for each frontal scan, two fitted AFMs are computed: one that has the left side mirrored to the right and another that has the right side mirrored to the left. The method derives geometry and normal images from the deformed AFMs as described in the previous section.

17.4 3D-aided Profile Recognition: URxD-PV

Until recently, research in profile-based recognition was based on comparison of *standard profiles*—the contours of side view images with yaw very close to $-90°$. Research in 3D–3D face recognition has indicated that the profile information contains highly discriminative information [35, 48, 69], where the term "profile" is often associated with the facial area along the symmetry axis of the 3D face model. However, neither approach is capable of accurate modeling of a silhouetted face profile, as observed in a 2D image because (i) the face is not perfectly symmetric, (ii) the face is almost never at yaw equal to $-90°$ with respect to the sensor, and (iii) if the distance between camera and object is not sufficiently large, perspective projection needs to be considered (based on imaging sensor parameters). Note that, in this paper, the term "profile" always indicates the silhouette of nearly side view head images for clarity of presentation.

The central idea of our approach is the use 3D face models to explore the feature space of a profile under various rotations. An accurate 3D model embeds information about possible profile shapes in the probe 2D images, which allows flexibility and control over the training data. We suggest that sufficient sampling in the pose space, which corresponds to nearly side-view face images, provides robustness for a recognition task. Specifically, we propose to generate various profiles using rotations of a 3D face model. The profiles are used to train a classifier for profile-based identification. Two different types of profiles are employed in our system: (i) *3D profiles*—those generated synthetically through 3D face models to be used as training data, and (ii) *2D profiles*—those extracted from 2D images of side-view faces.

The schematic illustration of the profile-based face recognition system is depicted in Fig. 17.7 and includes *Enrollment* and *Identification* phases. The algorithmic solutions for the entire 3D-aided profile-based recognition framework including profile modeling, landmark detection, shape extraction, and classification are provided in [20].

In our approach, we treat the profile as an open curve $\mathcal{C}$, it may be described by a pair of arc-length parameterized 1D functions $Y_{\mathcal{C}}(l)$ and $X_{\mathcal{C}}(l)$, where $l \in [0, 1]$. A set of k landmarks is defined by their coordinates on a parametric curve: $\{0 = \nu^1 < \cdots < \nu^k = 1\}$. The set contains both anatomical landmarks (e.g., "chin") and

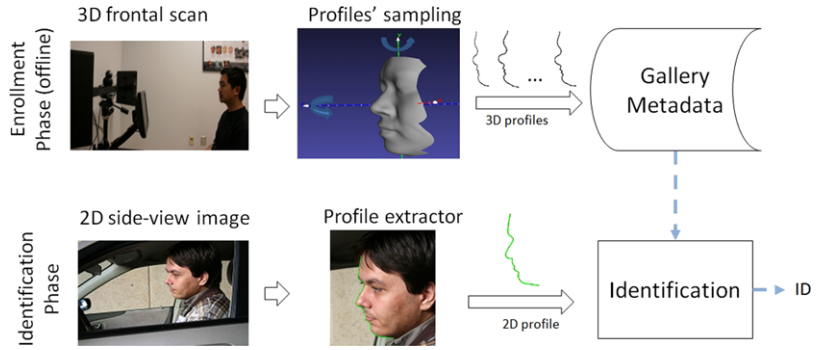


Fig. 17.7 Enrollment and identification phases of the proposed integrated profile-based face recognition system

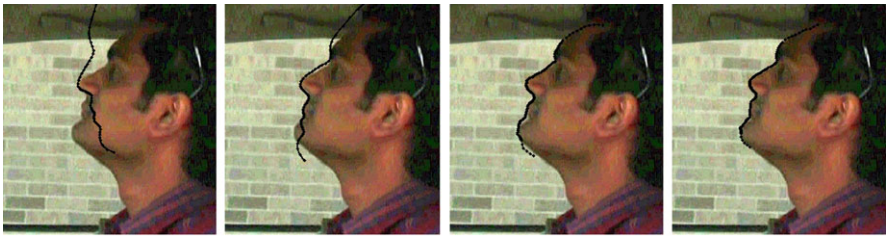


Fig. 17.8 Propagation of profile search. Depiction of **a** initial profile; **b** after two iterations; **c** after 5 iterations; and **d** final result

pseudo-landmarks (e.g., "middle of the nose"). We approximate functions $Y_{\mathscr{C}}(l)$ and $X_{\mathscr{C}}(l)$ by a finite set of points and obtain an equivalent *n-points shape model* as follows:

$$\mathbf{v} = [x_1, y_1, x_2, y_2, \ldots, x_n, y_n]^{\mathrm{T}} \in \mathbb{R}^{2n}. \tag{17.1}$$

The positions of the points are obtained through uniform arc-length sampling of the curve between a predefined subset of the landmarks. The sampling pattern is consistent for all profiles and, therefore, the coordinates of these landmarks always preserve their indices.

17.4.1 Profile Extraction from 2D Images

The profile extraction is based on the Active Shape Model paradigm developed by Cootes et al. [15]. It uses an iterative approach to gradually improve the fit of a given instance of the n-point shape to an image. Figure 17.8 depicts the shape propagation of the extractor. In the classical ASM framework, a manually labeled set of training images is needed. This dataset is used for the construction of the statistical

shape model, also known as Point Distribution Model (PDM). The same dataset is employed to design the features to guide the local search, typically by computing a pixel's likelihood belonging to a shape. In our case, no such labeled training set is available. Therefore, the main difference of our ASM framework from the classical approach is the fact that we only use available 3D shape information to guide the search of 2D profiles in the images and the local search is guided by features derived from color segmentation and edge detection operators.

17.4.2 Identification

During the identification phase, the matching scores between the probe profile and every profile in the gallery are computed. The decision is made according to the nearest neighbor rule. We propose to employ a modified Hausdorff distance as the matching score. For two finite point sets $\mathscr{M} = \{m_1, \ldots, m_n\}$ and $\mathscr{T} = \{t_1, \ldots, t_n\}$ with associated weights $\{w_1^{\mathscr{M}}, \ldots, w_n^{\mathscr{M}}\}$ and $\{w_1^{\mathscr{T}}, \ldots, w_n^{\mathscr{T}}\}$, the distance is defined as:

$$\frac{1}{n}\max\left(h_{\mathscr{M}} \sum_{m_i \in \mathscr{M}} \min_{t_j \in \mathscr{T}} \|m_i - t_j\| \sqrt{w_i^{\mathscr{T}} w_j^{\mathscr{M}}}, h_{\mathscr{T}} \sum_{t_i \in \mathscr{T}} \min_{m_j \in \mathscr{M}} \|t_i - m_j\| \sqrt{w_j^{\mathscr{M}} w_i^{\mathscr{T}}}\right)$$

where $\mathscr{M}$ and $\mathscr{T}$ are probe and gallery n-point shapes and $h_{\mathscr{M}}$ and $h_{\mathscr{T}}$ are normalization factors of the distance between predefined landmarks to eliminate scale influence. The set of weights for a probe profile reflects the accuracy of a shape extractor (all equal 1 for manually extracted profiles). The set of weights for a gallery profile reflects prior knowledge about the discriminative properties of the various regions.

A single face profile is a weak biometric, primarily because of pose uncertainty and inaccuracies in the acquisition and extraction stages. If the sequence of frames is available, we can compensate for these uncertainties by fusing the results of recognition from multiple frames. Our assumption is that, by using video frames acquired at a low frame rate, we will be able to accumulate evidence from more poses.

17.4.3 Integration

The complete framework of the automatic profile-based profile recognition system is illustrated in Fig. 17.9. During the *Enrollment phase* (E) the raw data from each subject is converted to metadata and stored in the database as follows:

E1. Acquire a facial shape with a 3D scanner and convert it to a polygonal mesh representation.
E2. Align and fit the 3D data to a common reference model.

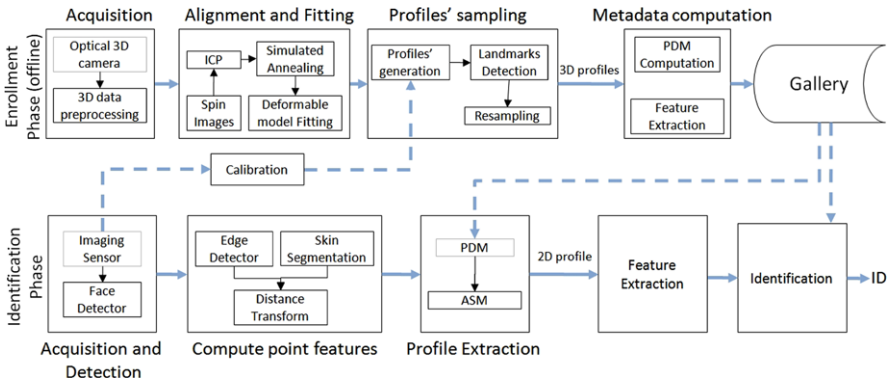


Fig. 17.9 The components of an integrated 3D-aided profile recognition system (URxD-PV)

E3. Generate multiple synthetic profiles by sampling a predefined range of rotation angles and locate a set of anatomical landmarks on them.
E4. Derive a set of features based on the profile geometry and landmark locations from profiles and store them as metadata to be used in the identification phase.

During the *Identification phase* (I), the profile is extracted from a 2D image, and its metadata is matched with the gallery metadata as follows:

I1. Acquire an image and compute a tight region of interest (*ROI*) that contains the face.
I2. Compute a set of features for each pixel in the ROI, which will be used to guide the shape extraction procedure.
I3. Extract the profile shape using the modified Active Shape Model.
I4. Extract features (varies depending on classifier) from the profile shape.
I5. Match/Classify the features.

17.5 3D-aided 2D Face Recognition: UR2D

We have developed a 3D-aided 2D face Recognition system (UR2D). Table 17.1 summarizes the different choices for the different modules in that UR2D system.

17.5.1 3D + 2D Enrollment

The UR2D enrollment method employs the Annotated Face Model (AFM) proposed by Kakadiaris et al. [30] to generate geometry images (regularly sampled 2D images with three channels) which encode geometric information (x, y and z components of a vertex in R^3). There are seven channels for the geometry images—three channels for representing the actual geometry of the face, three for representing the texture

Table 17.1 Variations of the UR2D system

Method name	Gallery Data	Probe Data	Geometry Image	Relighted	Distance metric	Score Normalization
UR3D	3D	3D	X		CWSSIM	MAD
UR2D-V-1	3D + 2D	2D	X	X	GS	E
UR2D-V-2	3D + 2D	2D	X	X	CWSSIM	E
UR2D-V-3	3D + 2D	2D	X		CWSSIM	E
L1	2D	2D				
UR2D-V-4	2D	2D			GS	E

Algorithm 17.1: Enrollment with 3D data

Input: 3D facial mesh, 2D facial image, subject ID.

1. Preprocess the 3D facial mesh.
2. Register AFM to the 3D facial mesh.
3. Fit AFM to 3D facial mesh.
4. Lift texture from the 2D facial image based on the fitted AFM.
5. Compute visibility map.
6. Store the fitted AFM, texture and visibility map in the gallery as metadata for subject ID.

information, and one for the visibility map. For practical purposes, all experiments use a resolution of 256×256.

Specifically, the algorithm first fits the AFM to the input 3D data [30]. Once the fitting is complete, the AFM is represented as a geometry image. For each vertex in the geometry image, the algorithm computes the closest point on the data. The texel corresponding to this point in the data is used to create the corresponding texture image for the fitted AFM. Additionally, a visibility map is computed (Algorithm 17.1). If the closest point on the data does not have a valid texel assigned (i.e., if the 3D point was not visible to the 2D image sensor), the value one (1) is assigned to the corresponding location in the visibility map. Otherwise, it is assigned a value of zero. The enrollment pipeline is depicted in Fig. 17.10.

17.5.2 2D Authentication

In the authentication stage (Algorithm 17.2), the input is a 2D image. Seven fiducial landmarks (two eye inner corners, two eye outer corners, nose tip, and two nose corners) are detected using PittPatt [54]. Once the pose is estimated (using these landmarks and their corresponding locations on the AFM), the texture is mapped onto the AFM (Fig. 17.11). An analytical skin reflectance model (described in the

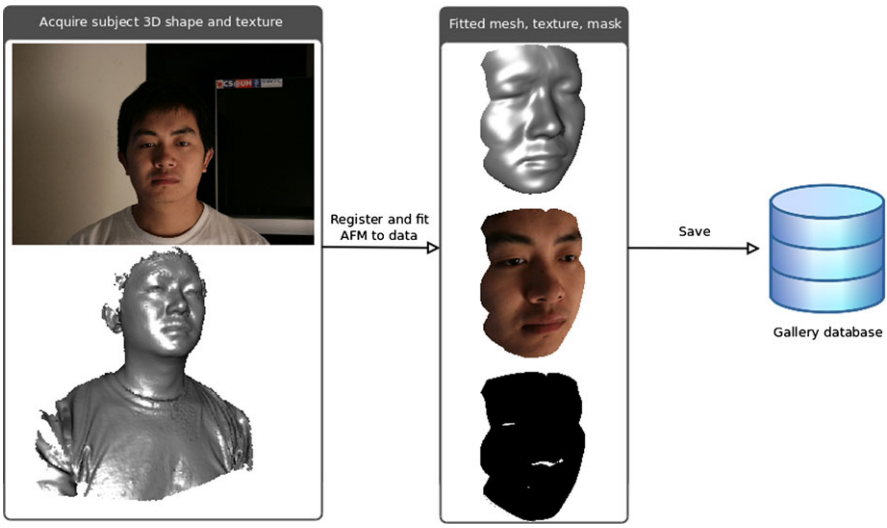


Fig. 17.10 Depiction of the enrollment procedure for the UR2D algorithm. The *first column* lists the input data while the *second column* list the fitted AFM with texture on the top and without texture on the bottom

Algorithm 17.2: Authentication with 2D data

Input: 2D facial image and claimed subject ID.

1. Retrieve "claimed ID" AFM from the gallery.
2. Locate the seven landmarks on the 2D facial image.
3. Register the AFM to the 2D facial image using the corresponding landmarks (Fig. 17.11).
4. Compute the visibility map.
5. Bidirectionally relight the enrollment 2D facial texture to match the probe 2D facial texture.
6. Compute the CWSSIM and GS scores between the relit texture and the probe texture.
7. Threshold the score to make an ACCEPT/REJECT decision.

next section) is used to bidirectionally relight the *gallery* texture using the stored AFM mesh, in order to match the illumination of the probe texture (Fig. 17.12).

17.5.3 Skin Reflectance Model

In the case when test data is of sufficient resolution, a bidirectional surface scattering reflection distribution function (BSSRDF) should be used to model the skin reflectance. However, in most recognition systems, we deal with data of rather low

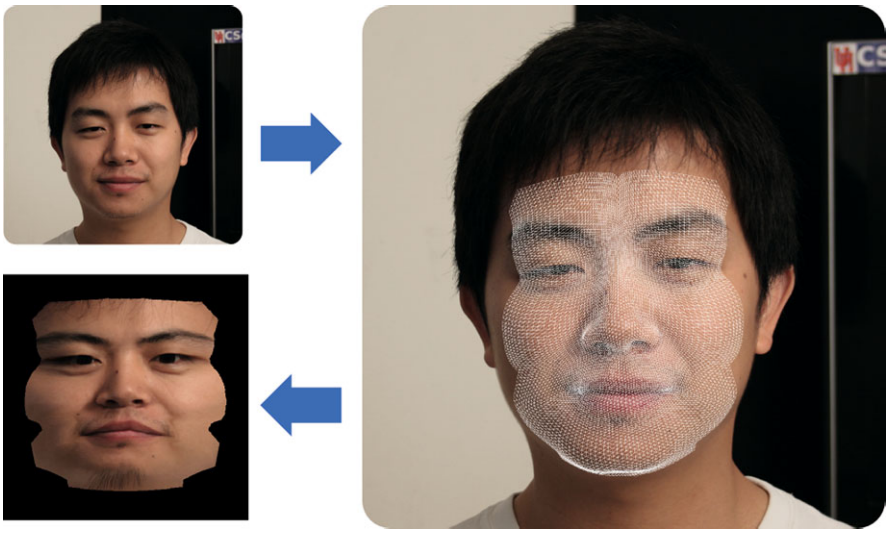


Fig. 17.11 Converting raw 2D images to textures in the geometry image space: Raw 2D image → Fitted AFM of the same subject registered and superimposed over the image → Image converted to texture in geometry image space. The conversion is done by matching a set of landmarks on the AFM and on the 2D image

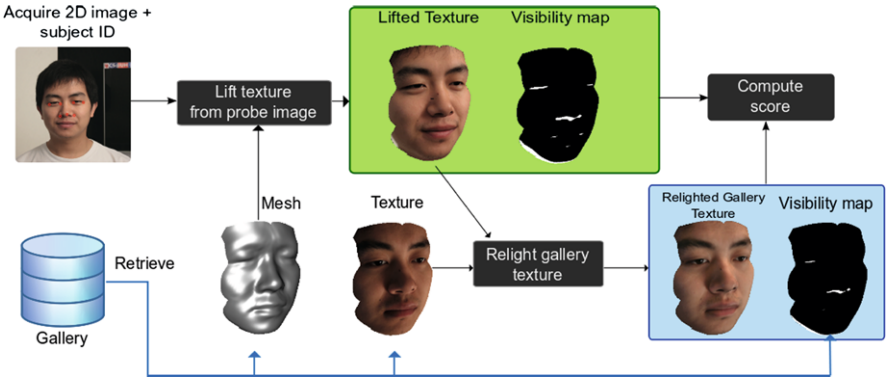


Fig. 17.12 The authentication phase of the 3D-aided 2D face recognition system

resolutions, thus, it is safe to employ a hybrid bidirectional reflectance distribution function (BRDF) to model skin reflectance. The ASRM uses the Lambertian BRDF to model the diffuse component and the Phong BRDF to model the specular component. The Lambertian BRDF is the simplest, most widely used, physics-based model for diffuse reflectance. The model assumes that the surface is equally bright from all directions. The intensity of the light at a surface point is proportional to the angle between the surface normal and the incident light directions (denoted as θ) $I_d = E \cos\theta$, where E is the intensity of the light source. The Lambertian BRDF

Fig. 17.13 Optimization for relighting (textures are in geometry image space): **a** M'_T: texture of subject A; **b** M_T: texture of subject B; **c** texture difference between subjects (before optimization); **d** texture difference between subjects (after optimization); **e** subject A with illumination of subject B ($I'_s + (I'_d + I'_a)\frac{M_T - I_s}{I_d + I_a}$)

does not take into account the specular reflections caused by the oily layer of the skin. The BRDF proposed by Phong [53] can be used to accommodate for this. The intensity of the specular reflection at a surface point is $I_s = E\cos^n\phi$, where ϕ is the angle between the view vector and the reflected light and n is a parameter that controls the size of the highlight. Note that each facial area has different specular properties, therefore the use of a specular map based on the annotation of the AFM is required [30].

17.5.4 Bidirectional Relighting

The illumination parameters and the ASRM can be optimized in two different ways: estimating the albedo [4, 60] and transferring illumination (relighting). In both cases, the UR2D algorithm represents the texture in the AFM's UV space.

Generally, the texture M_T is the result of the lighting applied on the unknown albedo M_A and is given by: $M_T = I_s + (I_d + I_a) \cdot M_A$, where I_a is the ambient component, I_d the diffuse component and I_s the specular component (assuming white specular highlights). Solving this equation for the albedo yields: $M_A = \frac{M_T - I_s}{I_d + I_a}$. However, for many practical applications, the albedo itself is not required, and is used only as an intermediate step for relighting. Thus, when possible, the user should use bidirectional relighting without first estimating the albedo. This means that the optimization directly estimates the parameters for two lights (one that *removes* the illumination from the *gallery* image and one that *adds* the illumination from the *probe* image). The goal is to match the illumination conditions of a *gallery* texture to that of a *probe* texture. The following metric is minimized:

$$D = \left| M'_T - I'_s - \left(I'_d + I'_a\right)\frac{M_T - I_s}{I_d + I_a} \right|, \tag{17.2}$$

where I_a, I_d, and I_s are the parameters of the light illuminating the *gallery*; I'_a, I'_d and I'_s are the parameters of the second light illuminating the *probe*, while M'_T is the target texture. This process is depicted in Fig. 17.13. The relighting method is bidirectional, meaning that *probe* and *gallery* textures can be interchanged.

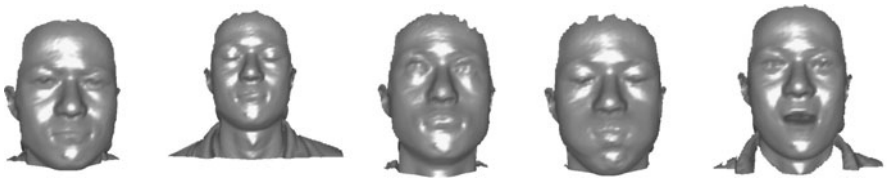


Fig. 17.14 Facial scans with various expressions for a subject from the FRGC v2 database

Table 17.2 Verification rates of our method at 0.001 FAR using different transforms on the FRGC v2 database

	ROC I	ROC II	ROC III
Fusion	97.3%	97.2%	97.0%
Haar	97.1%	96.8%	96.7%
Pyramid	95.2%	94.7%	94.1%

To improve the performance under low lighting conditions, instead of computing the difference in the RGB color space, a Hue-Saturation-Intensity (HSI) model can be used with the intensity weighed twice the amount of hue and saturation.

The equations above describe an ASRM for a single point light and the objective function to be minimized. The ASRM can be implemented as a Cg shader to greatly speed up the relighting process. For self-shadowing the shadow mapping technique can be used [60]. To model multiple point lights, the contribution of each light's ASRM must be summed. A full implementation of the ASRM on consumer level graphics hardware is able to bidirectionally relight a texture to a target within, on average, five seconds. Distance metrics and normalization methods are discussed in detail at [61].

17.6 Experimental Results

17.6.1 3D Face Recognition

For validation purposes, we have used the FRGC v2 [52] database, containing 4007 3D frontal facial scans of 466 persons. Figure 17.14 shows some examples of 3D facial scans from this database.

The performance is measured under a verification scenario. In order to produce comparable results, we use the three masks provided along with the FRGC v2 database. These masks, referred to as ROC I, ROC II and ROC III, are of increasing difficulty, respectively. The verification rates of our method at 0.001 False Acceptance Rate (FAR) are presented in the Table 17.2 [30]. The results are also presented using Receiver Operating Characteristic (ROC) curves (Fig. 17.15). The verification rate is measured for each wavelet transform separately, as well as for their weighted fusion. The average verification rate (over ROC I, II and III) was 97.16% for the fusion of the two transforms, 96.86% for the Haar transform and

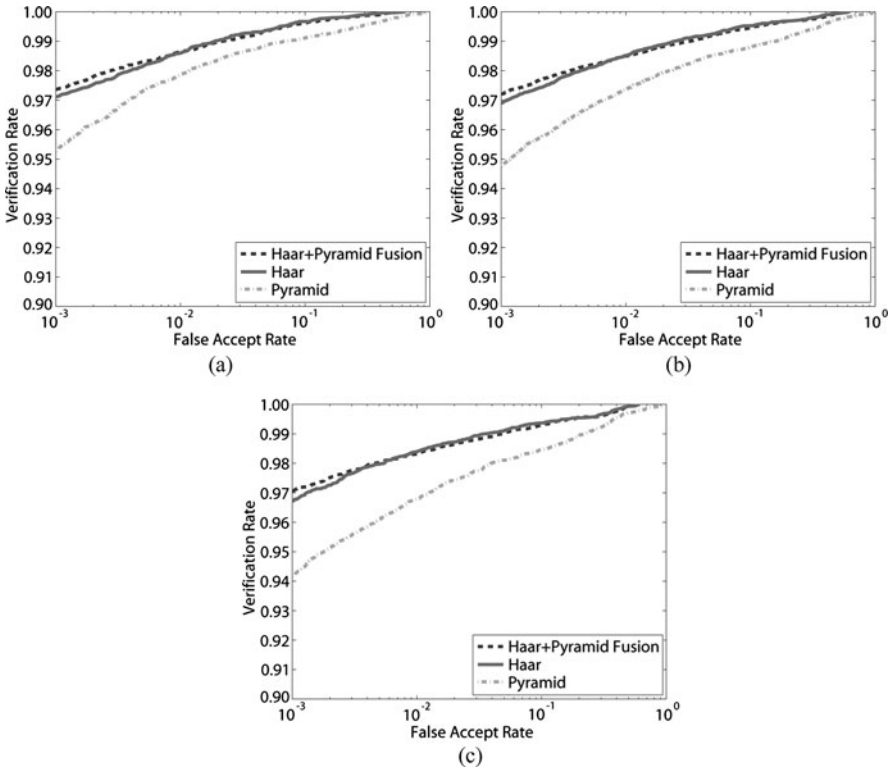


Fig. 17.15 Performance of the proposed method using the Haar and Pyramid transforms as well as their fusion on the FRGC v2 database. Results reported using: **a** ROC I, **b** ROC II, and **c** ROC III

94.66% for the Pyramid transform. Even though the Pyramid transform is computationally more expensive it is outperformed by the simpler Haar wavelet transform. However, the fusion of the two transforms offers more descriptive power, yielding higher scores especially in the more difficult experiments (ROC II and ROC III).

17.6.2 3D Face Recognition for Partial Scans

For interpose validation experiments, we combined the frontal facial scans of the FRGC v2 database with side facial scans of the UND Ear Database [63], collections F and G (Fig. 17.16). This database (which was created for ear recognition purposes) contains left and right side scans with yaw rotations of 45°, 60° and 90°. Note that for the purposes of our method, these side scans are considered partial frontal scans with extensive missing data. We use only the 45° side scans (118 subjects, 118 left and 118 right) and the 60° side scans (87 subjects, 87 left and 87 right). These data define two collections, referred to as UND45LR and UND60LR, respectively. For

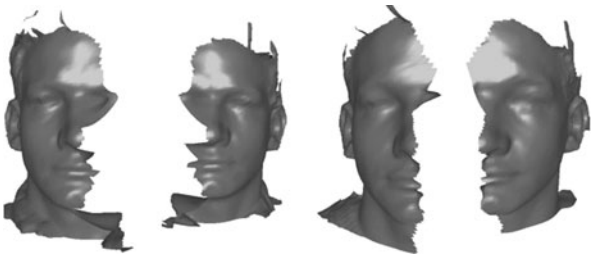


Fig. 17.16 Left and right side facial scans from the UND Ear Database

Table 17.3 Rank-one recognition rate of our method for matching partial scans

	Rank-one Rate
UND45LR	86.4%
UND60LR	81.6%
UND00LR	76.8%

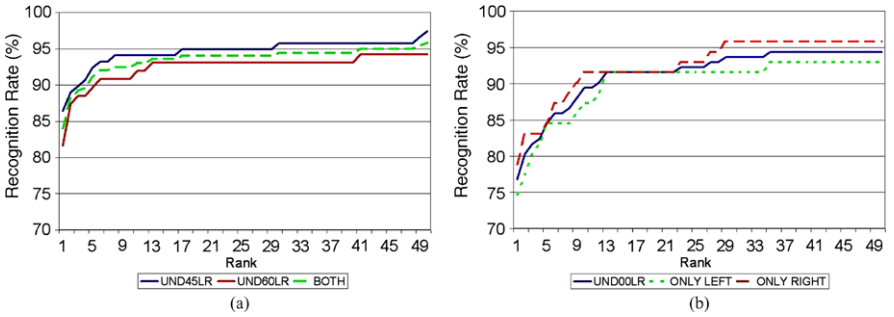


Fig. 17.17 **a** CMC graphs for matching left (gallery) with right (probe) side scans using UND45LR, UND60LR and the combination of the two; **b** CMC graphs for matching frontal (gallery) with left, right and both (probe) side scans using UND00LR

each collection, the left side scan of a subject is considered gallery and the right is considered probe. A third collection, referred to as UND00LR, is defined as follows: the gallery set has one frontal scan for each of the 466 subjects of FRGC v2 while the probe set has a left and right 45° side scan from 39 subjects and a left and right 60° side scan from 32 subjects. Only subjects present in the gallery set were allowed in the probe set.

We evaluated the performance of our method under an identification scenario using partial scans of arbitrary sides for the gallery and probe sets. Our method can match any combination of left, right or frontal facial scans with the use of facial symmetry. For each of the three collections, the rank-one recognition rates are given in the Table 17.3 while the Cumulative Match Characteristic (CMC) graphs are depicted in Fig. 17.17.

In the cases of UND45LR and UND60LR, for each subject, the gallery set contains a single left side scan while the probe set contains a single right side scan.

Therefore facial symmetry is always used to perform identification. As expected, the 60° side scans yielded lower results as they are considered more challenging compared to the 45° side scans (see Fig. 17.17(a)). In the case of UND00LR, the gallery set contains a frontal scan for each subject, while the probe set contains left and right side scans. This scenario is very common when the enrollment of subjects is controlled but the identification is uncontrolled. In Fig. 17.17(b) the CMC graph is given (UND00LR's probe set is also split in left-only and right-only subsets). Compared to UND45LR and UND60LR, there is a decrease in the performance of our method in UND00LR. One could argue that since the gallery set consists of frontal scans (that do not suffer from missing data), the system should perform better. However, UND00LR has the largest gallery set (it includes all of the 466 subjects found in the FRGC v2 database) making it the most challenging database in our experiments with partial scans.

17.6.3 3D-aided Profile Recognition

In our experiments, we employ data from the face collection from the University of Houston [66] that contains 3D data that was acquired with a 2-pod 3dMD™ system and side-view 2D images. The acquisition environment includes both controlled (indoor, stable background) and uncontrolled (driver) scenarios. The contents of the probe cohorts P1a and P1b are single side-view images of the driver in standard and arbitrary non-standard poses, corresponding to the gallery of 50 subjects. The probe cohorts P2a and P2b are video sequences of 100 frames each, from the same scene in visual and infrared spectrum, respectively. Each sequence corresponds to one of 30 subjects in the gallery.

In the first experiment, we validate recognition performance of the system on the single-frame and the multi-frame cohorts. The CMC curves for each type of pose for the single-frame cohort are depicted in Fig. 17.18(a). The results depicted in Fig. 17.18(b) are assessing the performance of profile recognition on visual spectrum and infrared sequences.

We observe that recognition is higher for the nearly standard profiles (rank-1 recognition rate is 96%), than for nonstandard profiles (78%). This effect may be attributed to the fact that standard profiles contain more discriminative information. The drop in performance for the infrared sequence (89%) with comparison to visual spectrum sequence (97%) is attributed to the fact that it corresponds to smaller face size (about 500 pixels for P2a and only 140 pixels for P2b).

For the gallery profile sampling, we consider angles in the range $[-110°, -70°]$ for yaw and $[-25°, 25°]$ for roll. We do not create profiles for different pitch angles because they correspond to only the in-plane rotations and do not influence the geometry of the profile. The resolution of sampling is 5°. To demonstrate the sensitivity of the algorithm to the predefined range of gallery sampling angles, we compare recognition results based on the original gallery to the results based on wider or narrower ranges, where each range is reduced by 5° from each side. The

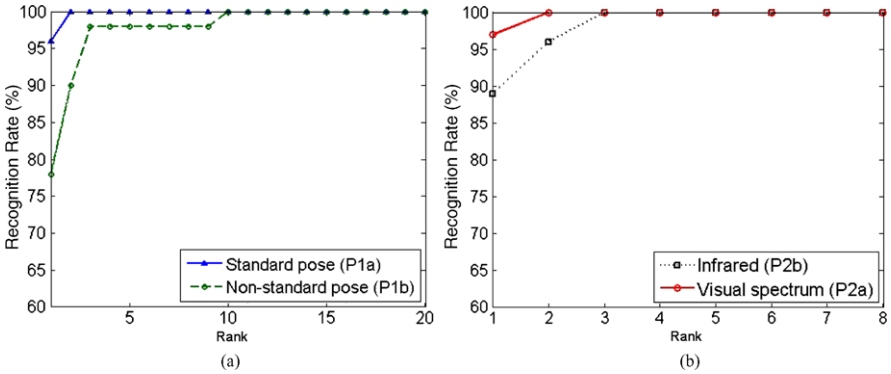


Fig. 17.18 Recognition results on side-view single-frame and multi-frame images: **a** performance on single-frame cohorts, and **b** performance on multi-frame cohorts

outcome of this comparison is depicted in Figs. 17.19(a, b) separately for standard and nonstandard poses. In a similar manner, Figs. 17.19(c, d) depict the influences of angular sampling density on recognition by comparison of the current sampling density of 5° to the alternative sparser sampling densities of 10° and 20°. These experiments were applied on P1a and P1b cohorts to examine the influence on standard and nonstandard poses.

The results show a clear tendency for the widely sampled pose domain to be more robust on non-standard poses. For instance, rank-1 recognition is 78% for wide region (current settings), 76% for slightly narrower region and only 56% for the sampling region with 10° reduced from each side. On the other hand, narrow sampled pose domain regions will slightly outperform if we consider only nearly standard poses. For instance, sampling in the narrow region results in 98% rank-1 recognition as compared to 96% recognition for other settings (wide and moderate). However, even in this case, sampling only a single point corresponding to standard pose (ultra-narrow) is worse than other options and results in 92% rank-1 recognition for nearly standard poses and only 56% for nonstandard poses. Unlike the area of sampling region, the frequency of sampling has less influence on the performance.

17.6.4 3D-aided 2D Face Recognition

Database UHDB11 [64] The UHDB11 database was created to analyze the impact of the variation in both pose and lighting. The database contains acquisitions from 23 subjects under six illumination conditions. For each illumination condition, the subject is asked to face four different points inside the room. This generated rotations on the Y axis. For each rotation on Y, three images are acquired with rotations on the Z axis (assuming that the Z axis goes from the back of the head to the nose, and that the Y axis is the vertical axis through the subject's head). Thus, each

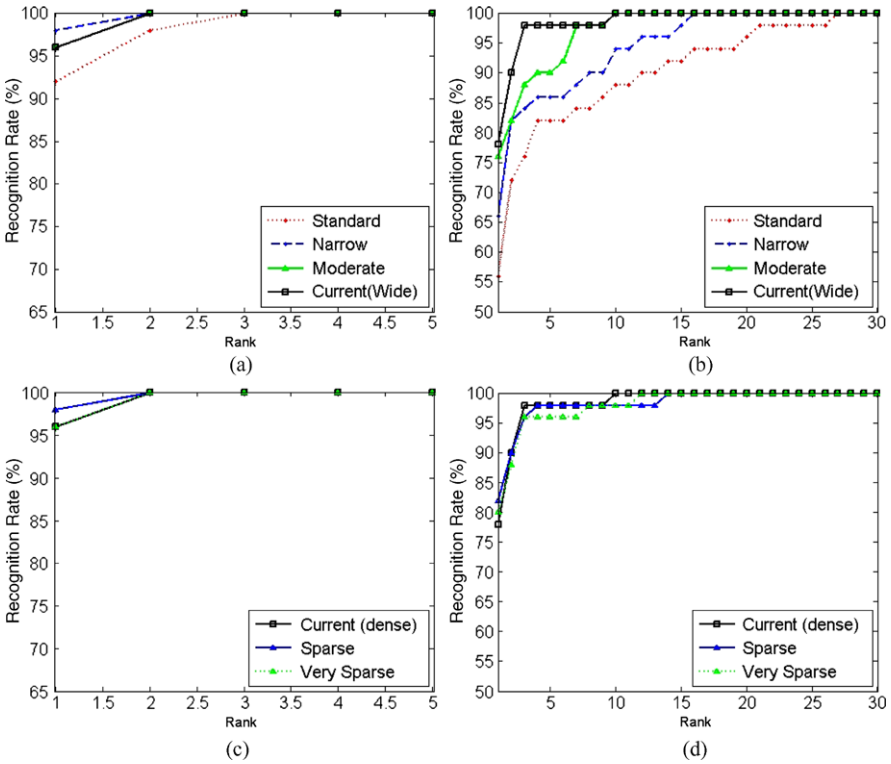
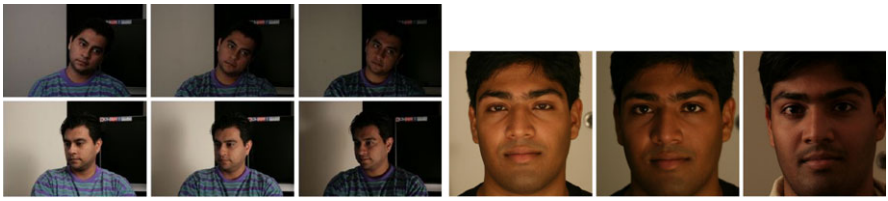


Fig. 17.19 Recognition results using various sampling domains: **a**, **c** cohort P1a (nearly standard poses), and **b**, **d** cohort P1b (nonstandard poses)

subject is acquired under six illumination conditions, four Y rotations, and three Z rotations. For each acquisition, the subject 3D mesh is also acquired concurrently. Figure 17.20(a) depicts the variation in pose and illumination for one of the subjects from UHDB11. There are 23 subjects, resulting in 23 gallery datasets (3D plus 2D) and 1,602 probe datasets (2D only).

Database UHDB12 [65] The 3D data were captured using a 3dMD™ two-pod optical scanner, while the 2D data were captured using a commercial Canon™ DSLR camera. The system has six diffuse lights that allow the variation of the lighting conditions. For each subject, there is a single 3D scan (and the associated 2D texture) that is used as a gallery dataset and several 2D images that are used as probe datasets. Each 2D image is acquired under one of the six possible lighting conditions depicted in Fig. 17.20(b). There are 26 subjects, resulting in 26 gallery datasets (3D plus 2D) and 800 probe datasets (2D only).

Authentication We performed a variety of authentication experiments. We evaluated both relighting and unlighting. In case of unlighting, both gallery and probe images were unlit (thus, becoming albedos). In the case of relighting, the gallery



(a) UHDB11 (b) UHDB12

Fig. 17.20 Examples images from database UHDB11 and database UHDB12 with variation of lighting and pose

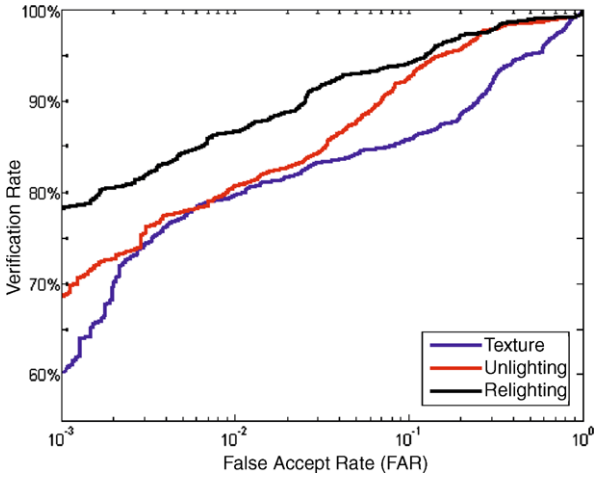


Fig. 17.21 ROC curve on authentication experiment on UHDB12 (varying illumination)

image was relit according to the probe image. The results for UHDB12 (using the UR2D algorithm, the CWSSIM metric and Z-normalization) are summarized using a Receiver Operating Characteristic (ROC) curve (Fig. 17.21). Note that face recognition benefits more from relit images than from unlit images. It achieves a 10% higher authentication rate at 10^{-3} False Accept Rate (FAR) than unlighting. The performance using the raw texture is also included as a baseline. Even though these results depend on the UHDB12 and the distance metric that was used, they indicate clearly that *relighting is more suitable for face recognition than unlighting*. The reason behind this is that any unlighting method produces an albedo for which the ground truth is not known; Therefore, the optimization procedure is more prone to errors.

UHDB11 was employed to assess the robustness of the 3D-aided 2D face recognition approach with respect to both lighting and pose variation. Figure 17.22 depicts the ROC curve for UHDB11 for four different methods: (i) 3D-3D: Using the UR3D algorithm where both the gallery and probe are 3D datasets (shape only no texture) [30]; (ii) 2D-3D(BR_GI, GS): The UR2D algorithm using bidirectionally relit images, GS distance metric, and E-normalization; (iii) 2D-3D(BR_GI, CWSSIM): The UR2D algorithm using bidirectionally relit images, CWSSIM distance

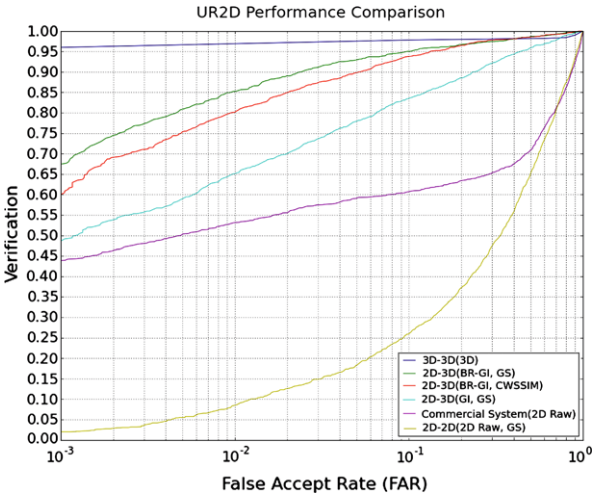


Fig. 17.22 ROC curve for an authentication experiment using data from UHDB11 (varying illumination and pose). Note that the Equal Error Rate which the 3D-aided 2D face recognition algorithm achieves is half that of the leading commercial product available at this time

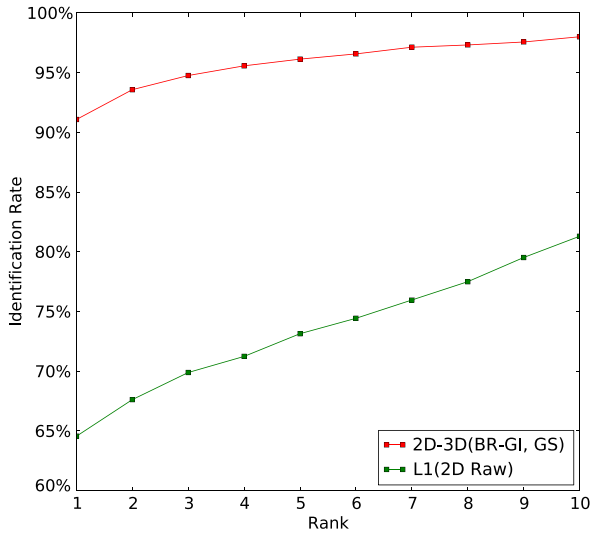


Fig. 17.23 Identification performance of the 3D-aided 2D face recognition approach versus the performance of a leading commercial 2D face recognition product

metric, and E-normalization; (iv) 2D-3D(GI, GS): The UR2D algorithm using raw texture from the geometry images, GS distance metric, and E-normalization; (v) 2D-2D(2D Raw, GS): Computing the GS distance metric for the raw 2D data, and E-normalization; (vi) L1(2D Raw, GS): Results from the L1 IdentityToolsSDK [33]. Note that the UR2D algorithm(BR_GS, GS) outperforms one of the best commercial products.

2D-3D Identification Experiment The UHDB11 database is also used in an identification experiment. The results are provided in a Cumulative Matching Characteristic (CMC) curve on 23 subjects of UHDB11 (Fig. 17.23). From these results

it is evident that the UR2D algorithm outperforms the commercial 2D-only product throughout the entire CMC curve.

17.7 Conclusions

While the price of commercial 3D systems is dropping, to tap into the wealth of 2D sensors that are already economically available, we would need to employ a 3D-aided 2D recognition technique. These 3D-aided 2D recognition methods can provide promising results without the need for an expensive 3D sensor at the authentication site. The effectiveness of these methods with relighting process have been demonstrated and it has been proven to provide robust face recognition under varying pose and lighting condition.

Acknowledgements This work was funded in part by the US Army Research Laboratory award DWAM80750, the UH Eckhard Pfeiffer Endowment, and the Central Intelligence Agency under the DISA ENCORE contract DCA200-02-D-5014 with Unisys Corporation serving as the primary on behalf of the Government. Any opinions, findings, conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of our sponsors.

References

1. Al-Osaimi, F., Bennamoun, M., Mian, A.: An expression deformation approach to non-rigid 3D face recognition. Int. J. Comput. Vis. **81**(3), 302–316 (2009)
2. Basri, R., Jacobs, D.: Lambertian reflectance and linear subspaces. IEEE Trans. Pattern Anal. Mach. Intell. **25**(2), 218–233 (2003)
3. Biswas, S., Aggarwal, G., Chellappa, R.: Robust estimation of albedo for illumination-invariant matching and shape recovery. IEEE Trans. Pattern Anal. Mach. Intell. **1**(8), 884–899 (2008)
4. Biswas, S., Aggarwal, G., Chellappa, R.: Robust estimation of albedo for illumination-invariant matching and shape recovery. IEEE Trans. Pattern Anal. Mach. Intell. **31**, 884–899 (2009)
5. Blanz, V., Vetter, T.: Face recognition based on fitting a 3D morphable model. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1063–1074 (2003)
6. Blanz, V., Scherbaum, K., Seidel, H.-P.: Fitting a morphable model to 3D scans of faces. In: Proc. 11th IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, 14–20 October 2007
7. Bottino, A., Cumani, S.: A fast and robust method for the identification of face landmarks in profile images. WSEAS Trans. Comput. **7**, 1250–1259 (2008)
8. Bowyer, K., Chang, K., Flynn, P.: A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition. Comput. Vis. Image Underst. **101**(1), 1–15 (2006)
9. Bronstein, A., Bronstein, M., Kimmel, R.: Three-dimensional face recognition. Int. J. Comput. Vis. **64**(1), 5–30 (2005)
10. Bronstein, A., Bronstein, M., Kimmel, R.: Robust expression-invariant face recognition from partially missing data. In: Proc. European Conference on Computer Vision, pp. 396–408, Graz, Austria (2006)
11. Chang, K., Bowyer, K., Flynn, P.: Adaptive rigid multi-region selection for handling expression variation in 3D face recognition. In: Proc. IEEE Workshop on Face Recognition Grand Challenge Experiments, pp. 157–164, San Diego, CA, 20–25 June 2005

12. Chang, K., Bowyer, K., Flynn, P.: Effects on facial expression in 3D face recognition. In: Proc. SPIE Biometric Technology for Human Identification II, vol. 5779, pp. 132–143, Orlando, FL (2005)
13. Chang, K., Bowyer, K., Flynn, P.J.: An evaluation of multi-modal 2D+3D face biometrics. IEEE Trans. Pattern Anal. Mach. Intell. **27**(4), 619–624 (2005)
14. Cootes, T., Taylor, C.: Statistical models of appearance for computer vision. Technical report, University of Manchester, October 2001
15. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models—their training and application. Comput. Vis. Image Underst. **61**(1), 38–59 (1995)
16. Cootes, T., Taylor, C., Kang, H., Petrovic, V.: Modeling facial shape and appearance. In: Handbook of Face Recognition, pp. 39–63. Springer, Berlin (2005)
17. Dibeklioglu, H.: Part-based 3D face recognition under pose and expression variations. Master's thesis, Bogazici University (2008)
18. Dibeklioglu, H., Salah, A., Akarun, L.: 3D facial landmarking under expression, pose and occlusion variations. In: Proc. 2nd International Conference on Biometrics Theory, Applications and Systems, Arlington, VA, 29 September–1 October 2008
19. Dryden, I., Mardia, K.: Statistical Shape Analysis. Wiley, New York (1998)
20. Efraty, B., Ismailov, E., Shah, S., Kakadiaris, I.: Profile-based 3d-aided face recognition. Pattern Recognit. (2011, in press). Corrected Proof. Available online 19 July 2011
21. Face recognition vendor test 2006 (2006). http://www.frvt.org/FRVT2006/
22. Galton, F.: Numeralised profiles for classification and recognition. Nature **83**, 127–130 (1910)
23. Gao, Y.: Efficiently comparing face images using a modified Hausdorff distance. In: Proc. IEEE Conference on Vision, Image and Signal Processing, pp. 346–350, December 2003
24. Gao, Y., Leung, M.: Human face profile recognition using Attributed String. Pattern Recognit. **35**(2), 353–360 (2002)
25. Gao, Y., Leung, M.: Line segment Hausdorff distance on face matching. Pattern Recognit. **35**(2), 361–371 (2002)
26. Gu, L., Kanade, T.: 3D alignment of face in a single image. In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1305–1312, New York, NY, 17–22 June 2006
27. Harmon, L.D., Khan, M.K., Lasch, R., Ramig, P.: Machine identification of human faces. Pattern Recognit. **2**(13), 97–110 (1981)
28. Husken, M., Brauckmann, M., Gehlen, S., von der Malsburg, C.: Strategies and benefits of fusion of 2D and 3D face recognition. In: Proc. IEEE Workshop on Face Recognition Grand Challenge Experiments, pp. 174–181, San Diego, CA, 20–25 June 2005
29. Kakadiaris, I., Passalis, G., Theoharis, T., Toderici, G., Konstantinidis, I., Murtuza, N.: Multimodal face recognition: Combination of geometry with physiological information. In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1022–1029, San Diego, CA, 20–25 June 2005
30. Kakadiaris, I., Passalis, G., Toderici, G., Murtuza, M., Lu, Y., Karampatziakis, N., Theoharis, T.: Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. IEEE Trans. Pattern Anal. Mach. Intell. **29**(4), 640–649 (2007)
31. Kakadiaris, I., Abdelmunim, H., Yang, W., Theoharis, T.: Profile-based face recognition. In: Proc. 8th IEEE International Conference on Automatic Face and Gesture Recognition, pp. 1–8, Amsterdam, The Netherlands, 17–19 September 2008
32. Kaufman, G., Breeding, K.: The automatic recognition of human faces from profile silhouettes. IEEE Trans. Syst. Man Cybern. **6**, 113–121 (1976)
33. L1 Identity Solutions. L1 faceit SDK
34. Lee, J., Machiraju, R., Pfister, H., Moghaddam, B.: Estimation of 3D faces and illumination from single photographs using a bilinear illumination model. In: Proc. Eurographics Symposium on Rendering, pp. 73–82, Konstanz, Germany, 29 June–1 July 2005
35. Li, C., Barreto, O.: Profile-based 3D Face Registration and Recognition, vol. 3506, pp. 478–488. Springer, Berlin (2005). Chap. 10

36. Lin, T., Shih, W., Chen, W., Ho, W.: 3D face authentication by mutual coupled 3D and 2D feature extraction. In: Proc. 44th ACM Southeast Regional Conference, Melbourne, FL, 10–12 March 2006
37. Liposcak, Z., Loncaric, S.: A scale-space approach to face recognition from profiles. In: Proc. 8th International Conference on Computer Analysis of Images and Patterns, pp. 243–250, London, UK, September 1999
38. Lu, X., Jain, A.: Multimodal facial feature extraction for automatic 3D face recognition. Technical Report MSU-CSE-05-22, Michigan State University, October 2005
39. Lu, X., Jain, A.: Automatic feature extraction for multiview 3D face recognition. In: Proc. 7th International Conference on Automatic Face and Gesture Recognition, Southampton, UK, 10–12 April 2006
40. Lu, X., Jain, A.: Deformation modeling for robust 3D face matching. In: Proc. IEEE Computer Vision and Pattern Recognition, pp. 1377–1383, New York, NY, 17–22 June 2006
41. Lu, X., Jain, A., Colbry, D.: Matching 2.5D face scans to 3D models. IEEE Trans. Pattern Anal. Mach. Intell. **28**(1), 31–43 (2006)
42. Mahoor, M., Abdel-Mottaleb, M.: Facial features extraction in color images using enhanced active shape model. In: Proc. 7th International Conference on Automatic Face and Gesture Recognition, pp. 144–148, Washington, DC, USA, 2–6 April 2006
43. Maurer, T., Guigonis, D., Maslov, I., Pesenti, B., Tsaregorodtsev, A., West, D., Medioni, G.: Performance of Geometrix ActiveIDTM 3D face recognition engine on the FRGC data. In: Proc. IEEE Workshop on Face Recognition Grand Challenge Experiments, San Diego, CA, 20–25 June 2005
44. Mian, A., Bennamoun, M., Owen, R.: An efficient multimodal 2D-3D hybrid approach to automatic face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **29**(11), 1927–1943 (2007)
45. Nair, P., Cavallaro, A.: Matching 3D faces with partial data. In: Proc. British Machine Vision Conference, Leeds, UK, 1–4 September 2008
46. Pan, G., Han, S., Wu, Z., Wang, Y.: 3D face recognition using mapped depth images. In: Proc. IEEE Workshop on Face Recognition Grand Challenge Experiments, pp. 175–181, San Diego, CA, 20–25 June 2005
47. Pan, G., Zheng, L., Wu, Z.: Robust metric and alignment for profile-based face recognition: An experimental comparison. In: Proc. 7th IEEE Workshop on Applications of Computer Vision, vol. 1, pp. 117–122, January 2005
48. Papatheodorou, T., Rueckert, D.: 3D face recognition. In: Face Recognition, pp. 417–446. I-Tech Education and Publishing, July 2007
49. Passalis, G., Kakadiaris, I., Theoharis, T., Toderici, G., Murtuza, N.: Evaluation of 3D face recognition in the presence of facial expressions: An annotated deformable model approach. In: Proc. IEEE Workshop on Face Recognition Grand Challenge Experiments, vol. 3, p. 171, San Diego, CA, 20–25 June 2005
50. Perakis, P., Passalis, G., Theoharis, T., Toderici, G., Kakadiaris, I.: Partial matching of interpose 3D facial data for face recognition. In: Proc. 3rd IEEE International Conference on Biometrics: Theory, Applications and Systems, Arlington, VA, 28–30 September 2009
51. Perakis, P., Theoharis, T., Passalis, G., Kakadiaris, I.: Automatic 3D facial region retrieval from multi-pose facial datasets. In: Proc. Eurographics Workshop on 3D Object Retrieval, pp. 37–44, Munich, Germany, 30 March–3 April 2009
52. Phillips, P., Flynn, P., Scruggs, T., Bowyer, K., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 947–954, San Diego, CA (2005)
53. Phong, B.: Illumination for computer generated pictures. Commun. ACM **18**(6), 311–317 (1975)
54. Pittsburgh Pattern Recognition. PittPatt face tracking & recognition software development kit (2009)
55. Riccio, D., Dugelay, J.-L.: Geometric invariants for 2D/3D face recognition. Pattern Recognit. Lett. **28**(14), 1907–1914 (2007)

56. Russ, T., Koch, K., Little, C.: 3D facial recognition: A quantitative analysis. In: Proc. 45th Annual Meeting of the Institute of Nuclear Materials Management, pp. 338–344, July 2004
57. Scheenstra, A., Ruifrok, A., Veltkamp, R.C.: A survey of 3d face recognition methods. In: Proc. in Lecture Notes in Computer Science, pp. 891–899 (2005)
58. Smith, W., Hancock, E.: Estimating the albedo map of the face from a single image. In: Proc. IEEE International Conference on Image Processing, vol. 3, pp. 780–783, Genoa, Italy, 11–14 September 2005
59. Stegmann, M.B., Gomez, D.D.: A brief introduction to statistical shape analysis. Technical report, Technical University of Denmark, March 2002
60. Toderici, G., Passalis, G., Theoharis, T., Kakadiaris, I.: An automated method for human face modeling and relighting with application to face recognition. In: Proc. Workshop on Photometric Analysis For Computer Vision, Rio de Janeiro, Brazil, 14–21 October 2007
61. Toderici, G., Passalis, G., Zafeiriou, S., Tzimiropoulos, G., Petrou, M., Theoharis, T., Kakadiaris, I.: Bidirectional relighting for 3D-aided 2D face recognition. In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2010)
62. Tsalakanidou, F., Malassiotis, S., Strintzis, M.: A 2D + 3D face identification system for surveillance applications. In: Proc. IEEE International Conference on Advanced Video and Signal based Surveillance, pp. 194–199, London, UK, 5–7 September 2007
63. U. of Notre Dame. University of Notre Dame Biometrics Database (2008). http://www.nd.edu/@cvrl/UNDBiometricsDatabase.html
64. UH Computational Biomedicine Lab. UHDB11 face database (2009). http://cbl.uh.edu/URxD/datasets/
65. UH Computational Biomedicine Lab. UHDB12 face database (2009). http://cbl.uh.edu/URxD/datasets/
66. URxD-PV. UHDB22: CBL database for biometrics research. Available at http://cbl.uh.edu/URxD/datasets/
67. Wang, Y., Zhang, L., Liu, Z., Hua, G., Wen, Z., Zhang, Z., Samaras, D.: Face relighting from a single image under arbitrary unknown lighting conditions. IEEE Trans. Pattern Anal. Mach. Intell. **31**(11), 1968–1984 (2009)
68. Wu, C., Huang, J.: Human face profile recognition by computer. Pattern Recognit. **23**, 255–259 (1990)
69. Yin, L., Yourst, M.: 3D face recognition based on high-resolution 3D face modeling from frontal and profile views. In: Proc. ACM SIGMM Workshop on Biometrics Methods and Applications, pp. 1–8, New York, NY, 8 November 2003
70. Zhou, X., Bhanu, B.: Human recognition based on face profiles in video. In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, June 2005