Advances and Challenges in 3D and 2D+3D Human Face Recognition

Shalini Gupta[1], Mia. K. Markey[2], Alan C. Bovik[1]

[1]Department of Electrical and Computer Engineering,
The University of Texas at Austin, TX 78712, USA
[2] The University of Texas Department of Biomedical Engineering,
Austin, TX 78712, USA

1 ABSTRACT

Automated human face recognition is required in numerous applications. While considerable progress has been made in color/two dimensional (2D) face recognition, three dimensional (3D) face recognition technology is much less developed. 3D face recognition approaches based on the appearance of range images and geometric properties of the facial surface have been proposed. Methods that combine 2D and 3D modalities also exist. These innovations have advanced the field and have created novel areas of investigation. The purpose of this chapter is to provide a summary and critical analysis of the progress in 3D and 2D+3D face recognition. The chapter also identifies open problems and directions for future work in the area.

2 INTRODUCTION

Robust and accurate identification of humans is required for numerous tasks. Over the years, a number of scientific approaches have been investigated to identify individuals. A widely explored approach is biometrics, the measurement of anatomical, physiological, and behavioral characteristics believed to be unique to individuals. The measurement of anatomical attributes of human beings is also called anthropometry. Phillips *et al.* outlined the characteristics of an ideal biometric based identification system: a) all members of the population should posses the biometric; b) the biometric signature of each individual should differ from that of other individuals in a controlled population; c) the biometric signature of each individual should not vary under the conditions in which it is recorded; and d) the system should resist countermeasures [87].

The first truly scientific approach to establish an individual's identity based on anthropometry was the Bertillon system introduced in France in the nineteenth century [95]. In this system, anthropometric measurements, *e.g.*, the skull width, foot length, cubit, trunk, along with hair color, eye color, and front and side view photographs were recorded. Bertillonage, as it was called, remained in vogue in police departments around the world for a number of years during the latter part of the nineteenth century. It was eventually abandoned due to questions regarding its infallibility as an accurate identification system, and was replaced by more reliable systems based on fingerprints. Fingerprint matching was considered the most reliable method for person identification for a large part of the twentieth century, and was later augmented with DNA matching.

With the availability of computers, a natural step forward has been to automate the task of human identification. In the recent years, due to the availability of computers with greater speeds and memory, automated biometric systems have become a topic of considerable interest. They have numerous applications including secured access to ATM machines and buildings, automatic surveillance, forensic analysis, retrieval of images from mugshot databases in police departments, automatic identification of patients in hospitals, checking for fraud or identity theft, and human computer interaction. Automated biometric systems expedite processing of large volumes of data, and in some cases can work in realtime. Currently, personal identification numbers, access codes/cards, bar codes, and radio frequency ID tags are popularly employed for identification. However, these are susceptible to loss or theft. Identification numbers and access codes/cards also require substantial user involvement. Hence, they are of limited utility for identifying very young children and seriously ill persons.

Many biometric techniques have been explored for automated human identification including face, iris, retina, fingerprint, palmprint, hand, gait, voice, and handwriting recognition. Of these, iris and fingerprint systems are reported to be highly accurate [93], but they require substantial subject cooperation. They are difficult to deploy in realtime screening and surveillance applications, where minimal user cooperation is desired, or where the system is to be operated covertly. Face recognition as a biometric modality requires less subject cooperation, is amenable to surveillance applications, and can be developed using relatively low cost components.

Although human beings are highly skilled at recognizing faces, there are also deficiencies in the face recognition abilities of humans. For example, a study of DNA exonerations reported that 84% of wrongful convictions were due in part to false recognition by eyewitnesses or victims [99]. Researchers believe that the face cognition abilities of humans are influenced by cross-racial effects [68] and other biases. Furthermore, many psychological aspects of human face processing and cognition are not well understood.

As another example in the criminal justice domain, consider the construction of a lineup, in which eyewitnesses or victims are presented with six face images to inspect. One of these is of the suspect. It is recommended that besides the suspect's facial image, the other images be of individuals close in appearance to the suspect [116]. In order to automatically construct effective lineups, it is necessary to define quantitative measures of similarity of facial appearance. Such measures can also be useful for content-based image retrieval from facial databases. Researchers in the cognitive sciences emphasize the need to define objective, quantitative measures of similarity of human faces [92]. They believe that insights gained from the field of computer vision could aid in enhancing the understanding of the mechanisms driving human face processing and cognition.

Considerable research attention has been directed, over the past two decades, towards developing reliable automatic face recognition systems. For the most part, research efforts have concentrated primarily on intensity/color/two dimensional (2D) images, and commercial systems are now available for this task [115].

Two dimensional face recognition systems are easy to construct with relatively cheap off-the-shelf components, but they are inadequate for robust face recognition. The Face Recognition Vendors Test was conducted in the year 2002 (FRVT 2002) to establish performance metrics for fully automatic 2D face recognition algorithms [91]. It was reported that the performance of the three best algorithms dropped nearly in half for facial images captured under varying ambient illumination conditions, or varying facial poses. Synthetic 2D frontal face images generated by employing three dimensional (3D) morphable models [13], greatly improved recognition results for faces with large pose variations. Hence at the FRVT 2002, using 3D face models for pose correction was identified as a potential solution to the pose problem in face recognition.

Three dimensional face recognition technology has emerged in the recent years, in part, due to the availability of improved 3D image acquisition devices and processing algorithms. For 3D face recognition algorithms, 3D facial models are employed either by themselves (3D algorithms) or in conjunction with 2D facial images (2D+3D algorithms).

In this chapter, we present a comprehensive, up-to-date literature review of the existing 3D and 3D+2D face recognition techniques. We focus primarily on recognition techniques developed for facial models captured in realtime using 3D acquisition devices. While, techniques for 3D and 2D+3D face recognition have also been reviewed previously [3, 15, 100], we present a more comprehensive and up-to-date literature survey.

We first outline the broad categories of tasks that are performed by automatic recognition systems and quantitative measures to assess their performance. We then discuss the main 2D face recognition algorithms that have influenced the development of analogous techniques for 3D face recognition, and have been employed in numerous 2D+3D approaches. A detailed discussion of existing approaches for 3D and 2D+3D face recognition follows. We conclude by enumerating open problems in area and identifying potential directions for future work.

3 FACE RECOGNITION TASKS

The two main tasks performed by any automatic human recognition system are verification/authentication and identification [88]. These are discussed in the following sections.

3.0.1 Verification

Verification/authentication is a one-to-one matching task wherein a person claims to be a specific entity known to the system (Figure 1). The database of people known to the sys-

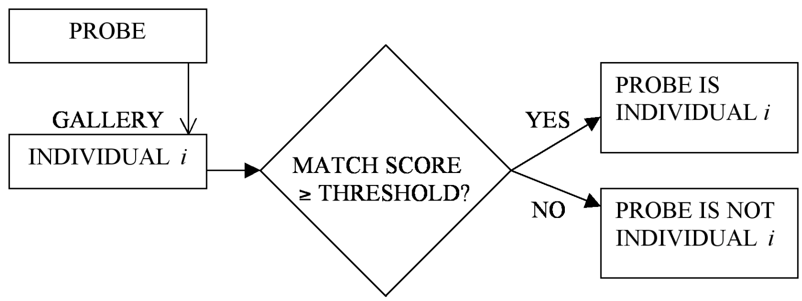


Figure 1: A schematic diagram of an automatic biometric verification system

tem is referred to as the 'gallery'. The individual whose identity is verified/authenticated by the system is referred to as the 'probe'. The facial representation of the probe is compared against the gallery representation of the claimed entity. If the similarity score between the two is greater than a predefined threshold, the individual is identified as the claimed entity; otherwise, he or she is rejected as an imposter. An example of a verification scenario is automated secure access to a building.

The performance of verification systems is reported in terms of a receiver operating characteristic (ROC) curve [94]. A ROC curve is a plot of the relationship between the false acceptance rate (FAR) and the false rejection rate (FRR). The verification scenario really can be thought of as a two class problem where match scores are either classified as intra-intity scores or inter-entity scores. FAR is defined as the proportion of comparisons between two different individuals that are falsely accepted by the system [101]. FRR is the proportion of comparisons between two instances of the same individual that are falsely rejected by the system [101]. Both FAR and FRR vary as the system's decision threshold is varied [33]. A single performance metric typically reported for verification systems is the equal error rate (EER), where FAR = FRR. For an ideal system EER=0%. The area under the ROC curve is also sometimes reported as a measure of performance. This area ranges from zero for an ideal system to 0.5 for a system with chance performance. Figure 2 presents a typical ROC curve. Note that different, but equivalent, formulations of ROC curves are used in evaluating classifications systems in other disciplines (*e.g.*, in medical imaging [71]).

ROC curves are also closely related to precision-recall curves frequently employed for evaluating the performance of general information retrieval systems [31]. Recall is equal to the true positive fraction, which measures the fraction of intra-entity scores that are correctly labeled by the system. Precision on the other hand measures the fraction of match scores that are labeled by the system as intra-entity scores and are truly so. Precision-recall curves are generally regarded superior than ROC curves for evaluating the performance of two-class decision systems with highly skewed data sets. Data sets employed to evaluate face verification systems are also highly skewed, as the number of intra-entity match scores

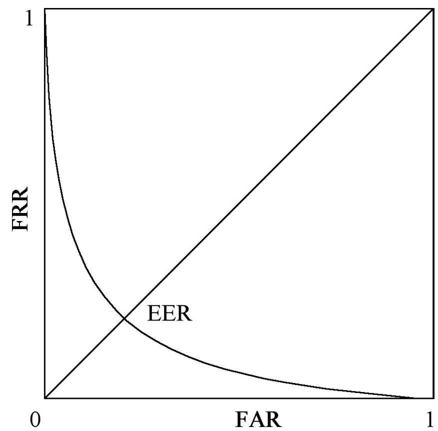


Figure 2: A Receiver Operating Characteristic Curve

is usually much less than the number of inter-entity match scores. Yet, currently only ROC curves are employed to evaluate their performance. Hence, investigations into the utility of precision-recall curves for the task are warranted.

3.0.2 Identification

Identification is a one-to-many matching task wherein an unknown individual's identity is established by comparing his/her biometric against a database of known individuals. The closest matches in the gallery are found. For example a person may be identified by comparing against a database of mugshots in a police department. A generalization of the identification task is the watchlist task. In a watchlist task, each probe is compared against signatures of all entities known to the system and entities resulting in the highest n similarity scores that are also above a predefined threshold value are considered matches. Setting a predefined threshold value seems to be is an arbitrary protection against selecting gallery individuals among the top n matches that are not close in appearance to the probe.

The performance of an identification system can be evaluated in terms of a cumulative match characteristic (CMC) curve [88, 89]. This formulation assumes a 'closed universe' model where all individuals that query the system are present in the gallery. The CMC curve is a plot of the recognition rate (RR) versus the top n database matches considered. It is the ratio of the number of probes for which the correct gallery match is present among the top n matches to the total number of probes that query the system. If the closed universe assumption is false, *i.e.*, a probe is not present in the gallery, then the maximum RR achieved is less than 100%.

Verification, identification, and watchlist tasks present different design challenges. For example, imposters attempting to fool a verification access control system would disguise themselves as someone known to the system. Thus, it is important that a high security ac-

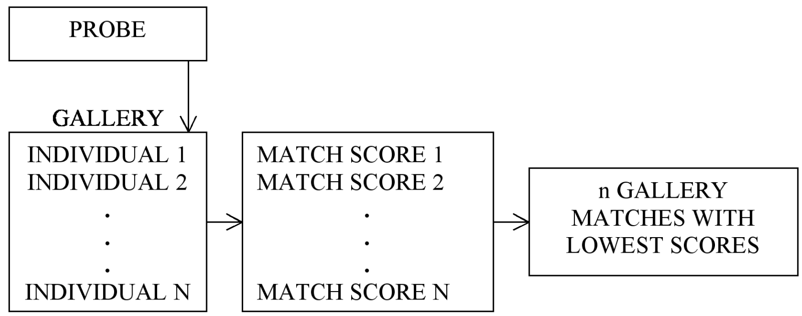


Figure 3: A schematic diagram of an automatic biometric identification system

cess control system have a low FAR so as not to falsely allow access to an imposter. By comparison, persons attempting to evade identification merely need to disguise themselves as anyone other than themselves. Hence, the FRR of, *e.g.*, airport screening systems should be low, so that individuals with disguises are correctly identified. Hence, while designing biometric systems it is necessary to keep in mind the final intended application. A system optimized for one application may not perform acceptably for another.

4 2D FACE RECOGNITION ALGORITHMS

2D face recognition technology has developed considerably over the past two decades. Broadly speaking, the two main approaches employed for 2D face recognition are: (a) based on the appearance of the whole face or (b) based on local facial features/geometric templates. Comprehensive survey papers detailing the progress of 2D face recognition algorithms have been written [26, 124].

Over the last decade, a number of independent tests have also been administered to compare the performance of 2D face recognition algorithms. These include a series of Facial Recognition Technology (FERET) tests [89, 90], and the Face Recognition Vendors Tests [118]. In such evaluations, a few 2D face recognition algorithms have consistently demonstrated superior performance. These include algorithms based on principal component analysis (PCA), linear discriminant analysis (LDA), local feature analysis (LFA), and elastic bunch graph matching (EBGM).

Commercials systems based on mature 2D face recognition algorithms include those by Cognitec Systems GmbH, Dresden, Germany; Identix®, Minnetonka, MN; and Eyematic Interfaces Inc., Inglewood, California now called Neven Vision, Santa Monica, California. These systems were among the top three performers at the FRVT 2002. The system by Identix® was based on LFA, and that by Eyematic Interfaces Inc./Neven Vision was based on EBGM. All these systems performed well for frontal or nearly frontal 2D face images captured indoors, but their performance decreased with variable illumination conditions and head rotations.

These top ranking 2D face recognition algorithms have also inspired similar algorithms for 3D face recognition. Furthermore, many of these 2D techniques have also been combined with 3D techniques to form 2D+3D face recognition algorithms. While, the focus of this paper is not 2D face recognition, the significant ones are discussed here since they form the basis for many 3D and 2D+3D strategies.

4.1 Holistic Appearance Based Techniques

Among the holistic 2D face recognition techniques, based on the appearance of facial grayscale/color images, are a number of subspace projection methods. These include algorithms that employ principal component analysis (PCA), and linear discriminant analysis (LDA). In these methods, a facial image is regarded as an instance in N dimensional feature space, where N is the number of pixels in the image. All human face images are modeled to lie on a linear/non-linear subspace of the N dimensional feature space. Statistical techniques are employed to learn the subspace using an ensemble of facial images. All facial images are projected onto the subspace before classification.

Holistic appearance based methods generally use information from the entire image. They do not employ high level knowledge about human faces to segment parts of the face. Hence, they tend to be less robust to outliers, cluttered backgrounds, and occlusions. Furthermore in order to reliably learn the facial subspace, they require a large number of training images of many subjects under diverse imaging conditions. Two of the important holistic 2D face recognition techniques, PCA and LDA, are discussed here.

4.1.1 Principal Component Analysis

PCA or eigenfaces was one of the first successful techniques developed for 2D face recognition [51, 110]. In this technique, a set of orthogonal basis vectors, that maximize the variance of facial image data, are obtained by eigen decomposition of the scatter matrix of facial images [32]. The eigenvectors are referred to as eigenfaces. PCA results in compact representations of high dimensional data, which are optimal in the mean squared sense. That is, PCA minimizes the mean squared error between the original image, and the corresponding image reconstructed from the eigen directions. If an image is reconstructed from all the eigen directions, the mean squared error between the original image and its reconstructed version is zero.

For face recognition, the top M $(M \leq N)$ eigen vectors, which account for most of the variation of the data, are retained. All gallery and probe facial images are projected onto the M directions. Faces in the transformed space are compared by means of a suitable distance metric. PCA is advantageous in that it possesses a closed form solution. It has resulted in effective 2D face recognition algorithms and is regarded as a benchmark, against which many others are compared [89]. Despite its success with 2D face recognition, it is not intuitively obvious as to what discriminatory information about human faces is encoded in the different eigen directions. Furthermore, its performance degrades with facial variations including expression, pose and illumination changes.

4.1.2 Linear Discriminant Analysis

Another successful approach for 2D face recognition is based on Fisher's linear discriminant analysis [5, 123]. This technique is also called fisherfaces. In this technique, high dimensional data are linearly projected onto $C - 1$ LDA directions, where C is the number of classes, such that the ratio of the between class scatter to the within class scatter is maximized. LDA can successfully discriminate between linearly separable classes [34]. For face recognition, the dimensionality of the data is first reduced via PCA or some other technique, before applying LDA. This is done to ensure that the within class scatter matrix, which is involved in the LDA calculations, is non-singular. Better results for 2D face recognition have been reported with LDA than PCA [5]. This could be explained by the fact while LDA projects data onto novel directions, such that the classes are most separated, PCA is not specifically tailored towards classification problems.

4.2 Local Feature Based Techniques

Two dimensional face recognition techniques based on local facial features/geometric templates, employ characteristics of localized regions of the face as features for recognition. Such techniques require an additional step of automatically locating specific parts of the face using flexible geometric templates or intensity/texture characteristics of specific facial features. The performance of feature based approaches depends on the accurate localization of facial landmarks. Segmentation of facial features is a non-trivial task. Segmentation techniques need to be specific enough to locate only the desired facial features, yet general enough to do so for a diverse variety of facial images. These are contrasting goals and thus it is difficult to achieve both.

If facial features can be reliably segmented, effective face recognition techniques based on local facial features can be developed. They are likely to be more robust to changes in facial pose, expressions and illumination, holes, occlusions, and the presence of noise than holistic appearance based approaches. This is because some facial features derived from localized facial regions would remain unchanged on varying other conditions. Furthermore, the information encoded in face recognition techniques based on local facial features is easier to interpret. For example, the discrimination ability of various sub-parts of the face can be evaluated independently of others.

In this section we discuss two successful 2D face recognition techniques that are based on local facial properties, namely LFA and EBGM.

4.2.1 Local Feature Analysis

It is argued that PCA does not exploit the inherent correlations and redundancies between neighboring pixels of an image. Eigen directions are also not topographical in that it is not understood how the various eigenfaces relate to each other. In order to overcome some of these deficiencies of PCA, the local feature analysis approach to face recognition, was developed [84]. Being a proprietary software of Identix®, Minnetonka, MN, the exact details of how this technique is applied to face recognition are not available. It is known however, that LFA is local in the sense that LFA kernels capture variations in sub-regions of the face. However, like PCA and LDA, it is applied directly to whole face regions without first segmenting different parts of the face. LFA kernels are statistically derived local and topographical sparse representations of the face. Such representations can provide information about the discrimination ability of the different parts of the face. LFA is reported to perform well with 2D frontal faces captured with constant illumination [91].

4.2.2 Elastic Bunch Graph Matching

Wiskott *et al.* developed another successful technique for 2D face recognition called elastic bunch graph matching [117]. EBGM is based on Gabor filter coefficients [14] extracted from specific facial fiducial points. In this technique, fiducial points are manually located on gallery images. From each point, a 'jet', comprising of forty coefficients for Gabor filters at 5 spatial frequencies and 8 orientations, is extracted. Jets from all gallery images are concatenated to form a data structure called a 'bunch'. A flexible 'face graph' is also constructed by connecting fiducial points by straight lines. The face graph and the bunch together form a data structure called an 'elastic bunch graph' (EBG). EBGs are employed for both automatic localization of fiducial points on probe faces, as well as for facial recognition. It is reported that EBGM works well for varying facial expressions and illumination conditions. However, it performs poorly for faces with large pose variations.

5 3D FACIAL MODELS

In this section we describe the techniques used to acquire 3D facial models in realtime and also data structures employed to represent them for face 3D face recognition algorithms.

5.1 3D Facial Model Acquisition

3D facial models are acquired using both active and passive techniques. Besl provides a summary of the different 3D imaging techniques [7]. The most widely employed active 3D acquisition technique is based on laser range finders [23,37,44,59,60,65,72,73,85,96,102]. A range finder projects light from a laser source onto a scene and records its reflection. The depth of the surface closest to the camera is determined by triangulation. Laser range finders produce dense and accurate 3D models, but require longer acquisition times than passive techniques. Furthermore, they require that the human subject be perfectly still during image capture [69]. Thus, laser range finders are unsuitable for high throughput screening applications. Another concern is the intrusive nature of laser light for human eyes.

Passive techniques employed for 3D facial image acquisition include stereo imaging [41–43, 55, 55, 70, 81, 122], and approaches based on structured light [2, 6, 10, 16, 36, 67, 78, 107, 119]. In stereo imaging systems, multiple cameras simultaneously capture a face from different view points. Depth information is resolved using camera calibration parameters and disparity information which is obtained from the different view points. To register points in the different images, a random light pattern is sometimes projected onto the scene [81].

In the structured light approach, a standard light pattern, *e.g.*, a light stripe pattern, is projected onto the scene [10]. Deformation of the light pattern and camera calibration parameters are employed to resolve the depth at each point in the scene by a process of triangulation. While passive techniques are faster, safer, and cheaper than laser range finders, they are typically less accurate and contain more missing data.

Attempts have also been made to recover the 3D shape of faces from one or more 2D images by morphing generic 3D facial models [13, 53, 57, 74, 75, 121]. Many such techniques involve variants of the 'shape from X' algorithms, but the recovery of 3D shape from a single texture image is an ill-posed problem. Such techniques can be useful for generating 3D facial models of subjects whose 2D images are available but their 3D images cannot be captured in realtime.

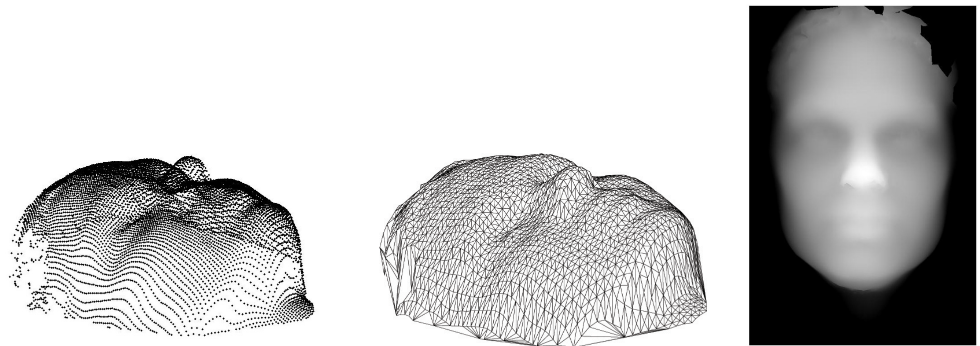
5.2 Facial Surface Representation

Point clouds, triangulated surface meshes, or range images are employed to represent facial surfaces. The point cloud representation contains the (x, y, z) coordinates of a set of points on the facial surface (Figure 4(a)). These points can be connected to their nearest neighbors via straight lines resulting in a triangulated mesh representation (Figure 4(b)). Compact point cloud or surface mesh representations can be obtained by sampling surfaces densely in regions containing detailed information, and sparsely in relatively smooth regions. The 3D points in these representations are usually unstructured and thus they require relatively involved algorithms for processing.

A range image, also referred to as a 2.5D surface or depth map, consists of (x, y) points on a regular rectangular grid. Each (x, y) point is associated with a z value of the point on the surface closest to the acquisition device (Figure 4(c)). Range images can be produced by orthographic projection of surface meshes or 3D point clouds. Three dimensional acquisition devices that capture range images directly are also available. As points in a range image are placed along a regular rectangular grid, they can be processed via relatively straightforward image processing algorithms.

6 3D FACE RECOGNITION ALGORITHMS

In recent years, numerous 3D face recognition techniques have been developed. In some respects, 3D face recognition techniques are advantageous relative to 2D techniques. The pose of 3D face models can be relatively easily corrected by rigid rotation and translation. This substantially alleviates the pose problem. Three dimensional models are normally in real-world dimensions and hence do not need to be rescaled prior to processing. Theoretically speaking, the shape of a face is independent of external factors such as ambient



(a) Point Cloud (b) Triangular Mesh (c) Range Image

Figure 4: This figure shows the different 3D representations of a facial surface. In Fig 4(a), the face is represented as a 3D point cloud. In Fig. 4(b) the 3D points are joined to form a triangular mesh representation. In Fig 4(c) the 3D face is represented as a range image where each value in the matrix corresponds to the depth at that point.

lighting during image acquisition. Knowledge of the shape of a face, and the direction and intensity of ambient illumination during capture, enables the calculation of facial surface reflectance properties. Surface reflectance is an intrinsic property of the facial surface. It can be employed to synthetically generate images of the face under different illumination conditions, which also helps to alleviate the facial illumination problem.

Concerns have been raised with regards to the illumination invariant nature of 3D face recognition algorithms [15]. It is argued that since intensity images are employed to construct 3D models in passive 3D image acquisition techniques, variations in illumination can alter the shape of the model constructed. However, in a recent study it was reported that the performance of 3D face recognition algorithms did not change significantly when 3D models were acquired under varying illumination conditions using a stereo imaging device [54]. It has also been observed that while the performance of 2D face recognition algorithms improves significantly when images are compensated for illumination variations, no significant improvement in performance of 2D+3D algorithms is observed [67, 107]. This evidence suggests that 3D face recognition algorithms may be effected less by changes in illumination conditions than 2D algorithms.

The two main categories of 3D face recognition techniques are: (a) based on the appearance of facial range images ('appearance based'), and (b) based on the geometric properties of 'free form' facial surfaces ("free form' based'). Campbell *et al.* provide an excellent overview of computer vision techniques for 3D object representation and recognition in general [18]. In the following sections, appearance (statistical learning) based and 'free form' based 3D face recognition approaches are discussed.

Although a number of 3D and 2D+3D face recognition algorithms have been proposed, most have been evaluated on small private data sets not accessible to other researchers. This makes it impossible to directly compare their performances. A few studies report the performance of multiple 3D face recognition algorithms on a common data set [6, 10, 12, 36, 41, 43, 45, 49, 50, 80, 83, 102, 119]. However, for most such studies, hypothesis testing procedures have not been conducted to confirm the statistical significance of the observed differences.

Three dimensional and 2D+3D face recognition studies also vary considerably with regards to certain aspects of their experimental design and evaluation protocol. For example, different research groups have employed different sized data sets. Hence, their results should be interpreted, bearing in mind that the performance of face recognition algorithms decreases on increasing the number of subjects in the database [91]. Furthermore, in different studies, different number of images of each person have been included in the gallery. It should be noted that increasing the number of images of each person in the database generally improves the performance of recognition algorithms [23].

There are also inconsistencies in the manner in which results have been reported by different researchers. For example, for the recognition performance, many research groups report only the rank n RR value instead of the rank 1 RR value. Similarly, for the verification performance, instead of reporting both the FAR and the FRR for a particular operating point, in many studies only one out of the two values has been reported. Hence, in order to assess and compare the performance of sate-of-the-art 3D and 2D+3D face recognition algorithms, it is necessary to test them on a large common data set using a fixed evaluation protocol. Results of statistical comparisons between the different algorithms should also be reported.

The Face Recognition Grand Challenge organized in the year 2005 (FRGC 2005) was a move in this direction [85]. As a part of the challenge, the performance of a few 3D and 2D+3D face recognition algorithms was evaluated using a standardized protocol on the two large data sets called FRGC v0.1 and FRGC v0.2 (Table 10). The performance of the 3D PCA algorithms was considered as the baseline [85, 86]. All other algorithms were compared against it. Some limitations of the FRGC data sets have also been noted. These include motion artifacts in range images, inconsistent expressions in the 2D and 3D images because of time delay between their capture, and occlusion problems due to facial hair [69].

6.1 3D Appearance Based Techniques

Three dimensional face recognition techniques based on the appearance of facial range images are similar to 2D holistic appearance based techniques. The only difference being that they employ range images instead of intensity images. For the most part, they are straight forward extensions of techniques that have been successful with 2D facial images.

A number of preprocessing and normalization steps are usually required in these algorithms. Their purpose is to localize and segment the human head; remove spike noise and holes (regions of missing data); align heads to a canonical position; and to generate range images in that position. Three or more points on the face are manually or automatically located to determine the head pose. For most algorithms, the canonical position is the frontal pose with the tip of the nose located at the center of the image.

The appearance based methods that have been investigated for 3D face recognition include PCA, LDA, LFA, independent component analysis (ICA), hidden Markov models (HMM), and optimal component analysis (OCA) (Tables 1, 2, and 3). These are discussed in detail in the following sections.

6.1.1 Principal Component Analysis

PCA is the most widely explored appearance based technique for 3D face recognition (Tables 1, 2, and 10). The PCA algorithm is regarded as a baseline for evaluating the performance of other 3D face recognition algorithms [86]. It has also been applied to 3D facial detection to distinguish between face and non-face regions in range images [30]. For face recognition via 3D PCA, eigensurfaces, similar to 2D eigenfaces, are computed. The optimal spatial resolutions for 3D PCA are reported to be up to 0.98 mm per pixel along the x and y directions, and resolutions ranging from 0.5 to 3 mm per pixel along the z direction [20]. This implies that higher resolutions may not be required for 3D PCA.

Overall, rank one recognition rates ranging from 100% to 68% have been reported for 3D PCA for databases of varying sizes and complexity (Tables 1 and 2). Although earlier experiments with smaller databases resulted in high performance for 3D PCA, more realistic performance estimates can be had from recent studies with larger and more complex databases.

Numerous algorithms have been proposed in which PCA is applied to depth values of range images [2, 20, 21, 23, 24, 35, 44, 45, 108], to (x, y, z) values of 3D facial point clouds [13], to geometry preserving isometric sphere representations of facial surfaces [77], to horizontal gradients, vertical gradients and depth curvature representations [42]. It is not clear as to which of these representations is most effective for use with PCA. Heseltine *et al.* reported that PCA on horizontal gradients of range images was more effective than PCA applied to depth or curvature representations [42]. However, the authors did not report whether these differences were statistically significant. Pan *et al.* also reported improved results for the case when PCA was applied to isometric spherical representations of facial surfaces instead of applying it directly to range images [77]. They evaluated both algorithms on the FRGC 2005 v0.1 data set.

Furthermore, 3D PCA has been applied either to entire range images [2, 108] or only to regions of range images that contain the main facial features [13, 20, 21, 23, 24, 35, 44]. It is likely that eliminating regions of the range image other than the main facial features may be advantageous as it would remove undesirable noise from hairstyles, clothing, and cluttered backgrounds.

The performances of 2D grayscale PCA and 3D PCA have been compared in a number of studies [20, 21, 23, 24, 35, 108] (Table 2). However, it has not been conclusively established as to which of the two modalities results in superior performance. In [21] 2D PCA and 3D PCA were not observed to be significantly different. Three dimensional PCA was observed to be significantly superior to 2D PCA in [20]. However, employing color images instead of grayscale images for 2D PCA has been reported to perform better than 3D PCA [35, 108].

In all these studies it was observed that combinations of the 2D and 3D modalities resulted in significantly superior performance relative to either of them. Chang *et al.* argue that this increase may be merely due to the availability of multiple images of a subject in the gallery for multi-modality approaches [23]. They did, however, observe that the performance for a multi-modality single image 2D+3D PCA approach was superior to a single modality multiple images 2D+2D PCA approach.

6.1.2 Independent Component Analysis

One study has been reported where independent component analysis was applied to facial range images for recognition [45]. ICA considers not only the linear relationships between pixels in a facial image, but also higher order relationships. It projects data linearly onto a set of new basis vectors that are as statistically independent as possible [48]. Analogous to the results reported for 2D face recognition [4], the performance of 3D ICA has been reported to be superior to 3D PCA [45]. Hence, it is worthwhile to further investigate the potential of ICA for 3D face recognition.

6.1.3 Linear Discriminant Analysis

Linear discriminant analysis techniques have also been investigated for 3D face recognition (Tables 1 and 3). LDA has been applied to various 3D representations including range images, and horizontal and vertical gradients of range images. In all studies, the gradient representations are reported to yield the best recognition performance [6, 41, 43]. However, it is unclear as to which of horizontal or vertical gradient representations, is better for 3D LDA. In one analysis the horizontal gradient representation was found to be optimal [41, 43], while in another the vertical gradient representation was reported as the superior of the two [6].

Consistent with results reported for 2D face recognition [5], 3D LDA has also been reported to be better than 3D PCA [36,41,43]. Results also suggest 3D LDA as being superior to 2D LDA, and 2D+3D LDA as being superior to either of the individual modalities [6].

6.1.4 Local Feature Analysis

A single study has been reported that investigates the utility of the LFA technique for 3D face recognition [6]. In this study, the authors employed the 2D LFA FaceIt® package available from Identix®, Minnetonka, MN for 3D face recognition. They observed that the vertical gradient representation was optimal for LFA. They further observed that the 2D LFA algorithm was significantly superior to 3D LFA, and that the 2D+3D LFA algorithm was superior to either of them.

6.1.5 Optimal Component Analysis

Srivastava *et al.* explored a technique called optimal component analysis for 3D face recognition [102]. They derived optimal linear projections of data in a lower dimensional space such that the performance of the nearest neighbor classifier was maximized. The technique involved a computationally expensive iterative procedure to learn the basis vectors. The authors reported that the method was superior to PCA, LDA, or ICA for 3D face recognition and was capable of handling variable facial expressions. However, they also did not report results of hypothesis testing for comparing the performances of these algorithms.

6.1.6 Hidden Markov Models

Two dimensional face recognition techniques based on hidden Markov models (HMM) [98] have also been applied to range images for 3D face recognition [2]. These techniques exploit the fact that facial features naturally occur in a fixed order from top to bottom and from left to right, irrespective of changes in illumination, pose, and facial expression. Different facial components are modeled as states in Markov models that are learnt from an ensemble of facial images. Achermann *et al.* observed the 3D HMM technique to be superior to 3D PCA [2], but statistical analyses were not conducted to confirm whether this difference was significant. Since a small database was employed in the study (Table 1), the difference is less likely to be statistically significant.

Methods based on embedded hidden Markov models (EHMM) [76] have also been studied for face recognition using color, grayscale intensity, and range images, and for combinations of these modalities [66, 67, 105–107] (Table 3). The 3D EHMM technique was reported to perform poorly relative to the grayscale EHMM and color EHMM techniques [105–107]. Like other appearance based techniques, the combined grayscale+3D EHMM and color+3D EHMM approaches are reported to perform better than the individual modalities.

Results of the studies based on EHMM also suggest that while changes in facial pose significantly alter the performance of both the 2D and 3D modalities, 3D algorithms may be effected less than 2D algorithms, by changes in illumination conditions during image acquisition. This can be concluded from the fact that when 2D+3D EHMM techniques were compensated for varying facial pose, by either augmenting the gallery with synthetic facial images with varying facial poses, [105–107], or by transforming all images to a frontal canonical pose before recognition [66, 67], significant improvement in their performances was observed. On the other hand, when pose compensated images were also compensated for varying facial illumination by either augmenting the database with images lit synthetically from various directions [105–107], or by re-lighting all images to frontal illumination [66, 67], the performance of 2D+3D EHMM algorithms did not change substantially.

The pose and illumination compensation technique where all faces were rotated to frontal pose and re-illuminated to frontal lighting [66, 67], was observed to be better than the database enrichment technique [105–107]. For the database enrichment technique it is difficult to ascertain *a priori* the optimal number and types of images of each individual that should be added to the gallery.

6.2 'Free form' 3D Face Recognition Algorithms

A 3D object that cannot be recognized as either a planar or a naturally quadric surface is referred to as a 'free form' object [18]. The human face is an example of a 'free form' object. Numerous 3D face recognition techniques based on 'free form' object descriptions have been proposed. These descriptions may be associated with individual points on a surface, one dimensional surface curves, or two dimensional surface patches. The 'free form' 3D object recognition techniques that have been investigated for human face recognition are discussed in the following sections.

6.2.1 Facial Surface Matching

A number of techniques for 3D face recognition have been developed where the shapes of two facial surfaces are compared directly (Tables 3, 4, 8, 9, and 10). The general philosophy of these approaches is to register two facial surfaces as closely as possible in 3D space and to compare them by means of a suitable metric. Different versions of this basic technique have been investigated by varying the registration procedure employed and/or the metric employed for calculating the distance between the two surfaces.

Coarse normalization, and coarse normalization followed by fine normalization are the two main approaches that have been employed for registering facial surfaces. In course normalization [1, 37, 55, 59, 60], the gross pose and position of each facial surface in 3D space is computed and they are transformed to canonical frontal positions. The two surfaces, in this orientation, are compared by means of either the mean squared error (MSE) metric [37,55,59], or the Hausdorff distance (HD) metric [1,60] proposed by Huttenlocher *et al.* [47]. This normalization technique however is not adequate as the accuracy of these distance metrics is dependent on precise alignment of the two surfaces.

For most 3D face recognition algorithms that employ matching of facial surfaces, coarse normalization is followed by fine normalization [22, 49, 52, 61, 62, 64, 65, 69, 70, 79–81, 96, 97]. The most frequently used algorithm for fine normalization is the Iterative Closest Point (ICP) algorithm [9]. In ICP, one 3D surface is iteratively transformed rigidly in 3D space to place it as close as possible to the other surface. Coarse normalization, in this case, helps the fine normalization iterative optimization procedure to converge correctly and prevents it from being trapped in a local minima.

Mean squared error, between pairs of closest points on the two surfaces, has been often employed as the objective criterion to be minimized in the iterative procedure [22,49,69,70]. In other studies the HD or the partial HD metric [62, 79, 80, 96], a combination of the MSE and HD metrics [52, 97], or a combination of the MSE between the pairs of closest 3D points and the MSE between closest points and surfaces [62, 64, 65] , as proposed by Chen and Medioni [27], have also been employed as the objective function. The partial HD metric is known to be more robust than the MSE metric to small mis-alignmemts between surfaces, and to the presence of holes, noise, and occlusions. It is also known that employing MSE based on point-to-surface distances [27] instead of MSE based on point-to-point distances [9] in the ICP algorithm, makes it less likely to be trapped in a local minimum. Russ *et al.* incorporated a multi-resolution approach with the ICP surface matching algorithm and reported high recognition rates with images from the FRGC v0.1 data set [97]. Slightly lower performance was reported on the same set of images when only a random set of points from the nasal region were employed for facial surface matching [52].

The ICP algorithm and its variants have yielded fairly successful algorithms for 3D and 2D+3D face recognition (Tables 3, 4, 8, 9, and 10). Face recognition algorithms based on ICP have been reported to be robust to changes in facial poses [65], and varying illumination conditions during 3D image acquisition [54]. They have also been reported to perform better than 3D PCA [49, 80], color LFA [70], and to 2D LDA [65] approaches.

It is not clear whether the addition of 2D grayscale/color information to the 3D ICP algorithm results in an increase in its performance. Lu *et al.* reported better performance for a combined 2D LDA+3D ICP algorithm relative to either of the individual modalities [65]. Similarly, Maurer *et al.* noted that the performance of the 3D ICP algorithm improved when it was combined with the 2D EBGM algorithm [69]. On the other hand, significant improvement in performance was not reported for the case when grayscale information and (x, y, z) co-ordinate values were employed for 4D ICP relative to when only the spatial co-oridnates were employed [81].

One major limitation of iterative surface matching based 3D face recognition algorithms is that they are computationally very expensive. For example, it has been reported that a single comparison between a pair of 3D facial surfaces required 20 seconds on an average Pentium 42.8 Ghz CPU [65]. This expense can become prohibitive for searching large databases. This computational expense stems from the two factors. First, the calculation of MSE between pairs of closest points on the two surfaces, or the HD metric, requires computations of the order $O(MN)$, where M is the number of 3D points on one surface and the N is the number of points on the other surface. Second, each matching procedure additionally requires a slow iterative optimization procedure to align a pair of facial surfaces. The iteration can also sometimes undesirably terminate in a local minimum. Recently 3D face recognition algorithms based on facial surface matching have been developed that employ the complex-wavelet structural similarity metric (CW-SSIM) [40]. This metric is known to be robust to small translational and rotational mis-alignments [114]. Matching facial surfaces using the CW-SSIM metric was shown to be more accurate, robust and efficient than employing the MSE or the HD metric.

Another limitation of 3D face recognition algorithms that employ ICP is that they are not robust to changes in facial expression [64, 65, 69, 81]. Employing only regions of the face that are relatively rigid and are not altered significantly by changes in facial expression, has been investigated as a solution to this problem [22].

6.2.2 Surface Normal Orientation Statistics

Attempts have also been made to match facial surfaces using the statistics of the orientation of facial surface normals (Table 5). Specifically, the extended Gaussian image approach [56, 103, 104] and the phase Fourier transforms [25, 82] have been investigated. The underlying idea of these techniques is to obtain a unique signature of each facial surface in terms of the distribution of the orientation of the surface normals. Facial surfaces are then matched by comparing their corresponding signatures.

Such techniques for facial surface matching are advantageous relative to the iterative surface matching techniques in some respects. They do not require an iterative procedure for aligning surfaces. Furthermore they are independent of the scale of 3D models [25], and have been shown to be robust to small rotations [25]. Yet, such techniques have not been explored extensively or rigorously. The few studies that exist, have employed small data sets, and none of them report face recognition performance estimates including the RR and EER values (Table 5). Hence, further investigation into such techniques is warranted in order to establish their true utility for 3D face recognition.

6.2.3 Profile Matching

One dimensional profile curves are can be easily obtained from 3D facial models by intersecting them with planes along various orientations, relative to 2D facial images. A number of studies have been reported where the shapes of the different facial profile curves have been compared in order to recognize faces (Tables 3, 6, and 8).

Interestingly, many studies have reported the central vertical facial profile curve to be the most discriminatory of all profile curves for 3D face recognition. Nagamine *et al.* [73] found vertical profiles located in the central region of the face to be the most discriminatory, followed by annular shaped curves centered at a point 40 mm above the tip of the nose. They also found horizontal facial profiles to be the least discriminatory of the three. Beumier *et al.* reported similar results. They observed a marginal decrease in performance of face recognition algorithms when 15 vertical profiles that spanned the entire facial surface were replaced by only 3 central profiles [10, 12]. Zhang *et al.* found that between the central vertical profile and two horizontal facial profiles (one passing through the forehead and the other through cheeks), the central vertical profile was the most discriminatory [122].

Consequently, a number of techniques have also been investigated to automatically locate the central vertical profile or the natural axis of bilateral symmetry on the human facial surface. Cartoux *et al.* automatically detected the axis of bilateral symmetry by iteratively minimizing the difference between principal curvatures on opposite sides of the facial surface [19]. Others have proposed a method of locating the axis of bilateral symmetry by iteratively aligning a facial surface to its mirror image [78–80, 122].

Consistent with other face recognition algorithms, Beumier *et al.* found that fusion of grayscale and shape information of the central and lateral vertical profiles results in improved performance, relative to either of the two individual modalities [11].

Three dimensional face recognition algorithms based on facial profiles are advantageous in that, they can provide information about the discriminatory information contained in the different sub-regions of the facial surface. However, they require an additional step of reliably locating the profiles. Furthermore, the overall accuracy of the recognition algorithms depends on accurate localization of profiles. The performance of profile based approaches for 3D face recognition lowers with the presence of variable facial expressions [122].

6.2.4 3D Local Geometric Features

Studies that have considered local geometric features of 3D facial surfaces for recognition are fewer in number, relative to holistic techniques that do not require automatic segmentation of facial landmarks (Tables 7, 8, 9, and 10).

A number of methods for automatically locating facial landmarks that are purely based on facial surface characteristics have been proposed. The most widely used method is to locate facial landmarks is to exploit either their unique curvature characteristics [22, 28–30, 37, 38, 46, 65] using the H-K segmentation algorithm [8], or to employ the curvature properties of facial profile curves [58, 122]. Another approach is to align 3D faces rigidly [119] or non-rigidly [49] to generic face templates for which the locations of fiducial points/facial features are known *a priori*. The tip of the nose has been detected as the most prominent point of 3D facial models in a canonical frontal position [59, 73]. However, this heuristic method is not successful for arbitrary facial poses. Yacoob and Davis labeled components of the facial surface based on contextual reasoning [120]. When color/texture information is also available along with the 3D shape information, it has also been employed to automatically locate facial landmarks [46, 55, 109, 111–113].

The local geometric features of facial surfaces that have been employed as features for face recognition include position, surface area and curvatures properties of facial landmarks/fiducial points, and 3D Euclidean distances, ratios of distances, geodesic distances and angles between them. The local shape of facial landmarks/regions about fiducial points have been quantified by means of Gaussian curvature values [38, 72], Guassian-Hermite moments [119], 'point signatures' [28, 113], and by 2D and 3D Gabor filter coefficients of facial range images [46, 111]. In order to quantify the relationships between the facial landmarks, 3D Euclidean distances [38, 46], angles and 3D Euclidean distances [58, 72], or geodesic and 3D Euclidean distances have been employed [39].

The method based on 'point signatures' [28] has also been reported to be superior to methods based on PCA applied to the 3D point clouds, profile matching, and PCA applied to range images [49]. Its performance was equivalent to an approach where the positions of only 10 fiducial points were compared after iteratively aligning a pair of facial models. Another approach based on the shape characteristics of local facial features, and surface matching has also been reported to perform better than profile matching [119].

A technique, where the facial sub-regions of the eyes, nose and mouth by fitting with subdivisional surfaces, and are matched separately, has been evaluated on the FRGC v0.1 and the FRGC v0.2 data sets [50, 83]. It performed better than the baseline 3D PCA algorithm. Another 2D+3D face recognition technique based on EGBM applied to range and grayscale images performed on par with this technique on the FRGC v0.2 data set [52]. In this study, the 2D EBGM algorithm performed better than the 3D EBGM algorithm. This could be explained by the fact that Gabor filter coefficients that are employed to quantify texture in grayscale images, might not encode discriminatory information for facial range images that do not contain as much texture variations (Figure 4(c)).

Face recognition techniques based on local facial features present several advantages relative to holistic techniques. First, instead of an ad hoc set of local facial features, their selection can be based on domain knowledge about the structural diversity of human faces. Such an approach has been demonstrated to result in effective 3D face recognition algorithms [39]. Second, techniques based on local facial features are less affected by global changes in the appearance of facial images including changes facial expressions [28, 39, 113].

Nonetheless, techniques for 3D face recognition based on local facial features have received little attention relative to holistic techniques. This may be due to the fact that such techniques require an additional step of reliably locating facial landmarks, which may effect their overall performance. Nonetheless, if facial landmarks can be reliably located, evidence from the literature on both 2D and 3D face recognition suggests that powerful techniques for 3D face recognition can be developed. Hence, there is a need to further explore the potential of 3D face recognition algorithms that employ features from facial sub-regions, and to find ways to combine them with techniques for robust facial feature detection.

6.3 Ensemble Approaches

A number of modular and ensemble methods consisting of combinations of multiple ‘free form’ 3D face recognition techniques have also been investigated (Table 8). For a majority of such algorithms, it has been observed that the combined approach results in superior performance, relative to the individual constituent approaches. Gökberk *et al.* observed this for the case when scores of 3D face recognition algorithms based on LDA, surface normals, and profiles were combined in parallel using a nonlinear rank-sum method [36]. Significant improvement in performance was also reported for hierarchical combinations of a surface matching classifier with 3D LDA [36], or 2D LDA [65]. A combination of features from the whole face for surface matching, and features from the mouth, nose and orbital regions also resulted in superior performance, relative to the individual sets of features [119]. Pan *et al.* observed superior performance when they combined outputs of a facial surface matching algorithm and a central profile matching classifier using the MAX rule [78, 79].

7 EXPRESSION INVARIANT APPROACHES

Achieving invariance to changes in facial expressions is one of the major open problems of automatic 2D and 3D facial recognition. While the pose of 3D facial models can be easily corrected, changes in facial expression, which are non-rigid transformations of the facial surface are not as trivial to eliminate. Numerous studied have demonstrated that the performance of 2D, 3D and 2D+3D techniques is considerably reduced when facial images with arbitrary expressions are employed [50, 65, 69, 81, 83, 111, 122]. Gallery enrichment has been the most common approach employed in 2D face recognition to introducing some degree of invariance to facial expressions. The idea, is to include multiple images of an individual in the gallery with different facial expressions. There are two problems with this technique. First, it increases computational costs. Second, there is no reliable way to decide on the number and type of facial expressions to add to the gallery, which would, in principle, encompass all facial expressions of an individual.

With the introduction of 3D face recognition techniques, a few other methods for achieving expression invariance have also been explored (Table 9). For 3D expression invariant face recognition, one approach has been to obtain an invariant representation of the face. The other has been to employ regions of the face that are known to remain relatively rigid under facial expression changes. These techniques are discussed in detail in this section.

A number of expression invariant representations of the human facial surface have been proposed. Bronstein *et al.* modeled changes in facial expressions as isometric deformations of the facial surface [16, 17]. For such deformations, the intrinsic properties of surfaces including geodesic distances between all pairs of points remain unchanged. The authors' approach was to isometrically embed facial surfaces into expression invariant canonical forms before recognition using multidimensional scaling. Similarly, invariance to facial expression changes was also observed for a algorithm that employed geodesic and 3D Euclidean distances between anthropometric facial fiducial points as features [39].

Wang and Chua employed 2D/3D Gabor coefficients calculated from range images at specific facial fiducial points to generate expression invariant representations of human face [111]. Their technique met with moderate success. Lu and Jain proposed a technique where 3D models were matched rigidly using ICP then non-rigidly deformed using thin-plate spline deformation to generate a displacement vector image [63]. Each point in the displacement vector image contained a vector connecting a point on the original face before deformation to its location after deformation. These vector images were classified as intra-personal and inter-personal deformations.

The other approach of employing regions of the face that are relatively rigid to changes in facial expressions has not been investigated as extensively. Chang *et al.* investigated matching local regions of the facial surfaces including the nose and the nose bridge using ICP [22]. In another study 'point signatures' were employed to quantify the local surface curvature of fiducial points located at relatively rigid regions of the face [28,113]. However, it is not clear whether this technique achieved expression invariance due to favorable properties of the 'point signature' representation or due the fact that facial features were derived from relatively rigid regions of the face.

8 CONCLUSION

The field of 3D face recognition has grown rapidly since FRVT 2002. Both appearance based and 'free form' based 3D face recognition algorithms have been proposed. Numerous approaches to combine 3D and 2D modalities have also been investigated. While it still remains a matter of debate as to which of 2D or 3D modalities is superior, it has been conclusively established that combing the two modalities holds promise for improving face recognition performance. The availability of 3D data has considerably alleviated the pose problem, but achieving expression invariance still remains to be solved. Debate also exists about the 'illumination invariant' nature of 3D data, but cues from existing studies point towards the fact that 3D techniques may be less sensitive to changes in illumination during image capture than 2D techniques.

Considerable attention is now also being directed towards robust testing and evaluation of 3D face recognition algorithms on a common database. This will help to objectively evaluate the current sate-of-the-art in the area. A number of areas, however, still remain unexplored. Research efforts need to be directed towards developing 3D feature and profile detection algorithms. A better understanding of facial image statistics and surface geometric properties is required to understand the workings of the holistic appearance based statistical approaches for 3D face recognition. Assessing the relationship between multiple modalities and improving methods of combining them are also likely to advance the field further.

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Achermann [2]	PCA, HMM	24	5/subject	5/subject	RR = 100%
Heseltine [42]	PCA	100	40	290	EER = 12.7%
Hesher [44]	PCA	37	5/subject	1/subject	RR = 94%
Hesher [45]	ICA	37	5/subject	1/subject	RR = 97%
Heseltine [41, 43]	LDA	230	–	–	EER = 9.3%
Srivastava [102]	OCA	67	3/subject	3/subject	RR = 99%

Table 1: The table summarizes the studies that have investigated appearance based 3D face recognition algorithms. For some studies, the authors have not specified the number of images/subject they included in the gallery and probe data set, but instead provide only the total number of images employed (*e.g.* Heseltine [42]). Missing information is indicated by '–'. The decrease in performance of the 3D PCA algorithm with the increase in the number of subjects in the database is evident from this table.

Author	Method		Data Set			Performance
	2D	3D	# Subjects	# Gallery Images	# Probe Images	
Tsutsumi [109]	PCA	PCA	24	9/subject	35/subject	$\mathrm{FAR}_{\mathrm{2D+3D}} = 4.5\%$ $\mathrm{FRR}_{\mathrm{2D+3D}} = 3.7\%$
Chang [21]	PCA	PCA	166	1/subject	1/subject	$\mathrm{RR}_{\mathrm{2D+3D}} = 92.8\%$ $\mathrm{RR}_{\mathrm{2D}} = 83.1\%$ $\mathrm{RR}_{\mathrm{3D}} = 83.7\%$
Chang [20]	PCA	PCA	200	1/subject	1/subject	$\mathrm{RR}_{\mathrm{2D+3D}} = 98.5\%$ $\mathrm{RR}_{\mathrm{2D}} = 89\%$ $\mathrm{RR}_{\mathrm{3D}} = 94.5\%$
Chang [24]	PCA	PCA	127	1/subject	297	$\mathrm{RR}_{\mathrm{2D+3D+IR}} = 100\%$ $\mathrm{RR}_{\mathrm{2D}} = 90.6\%$ $\mathrm{RR}_{\mathrm{3D}} = 91.9\%$ $\mathrm{RR}_{\mathrm{IR}} = 71.0\%$
Chang [23]	PCA	PCA	198	1/subject	670	$\mathrm{RR}_{\mathrm{2D+3D}} = 97.5\%$
Tsalakanidou [108]	color PCA	PCA	40	2/subject 3 poses	2/subject	$\mathrm{RR}_{\mathrm{color+3D}} = 98.8\%$ $\mathrm{RR}_{\mathrm{color}} = 93.8$ $\mathrm{RR}_{\mathrm{3D}} = 85$
Godil [35]	color PCA	PCA	200	1/subject	1/subject	$\mathrm{FAR}_{\mathrm{color+3D}} = 20\%$ $\mathrm{FRR}_{\mathrm{color+3D}} = 1\%$ $\mathrm{RR}_{\mathrm{color+3D}} = 82\%$ $\mathrm{RR}_{\mathrm{color}} = 78\%$ $\mathrm{RR}_{\mathrm{3D}} = 68\%$

Table 2: This table summarizes the studies that have combined PCA classifiers applied to different imaging modalities, including grayscale, color, infra red and 3D range images. It is evident from the table that the combination of modalities consistently results in superior performance relative to the individual modalities.

Author	Method		Data Set			Performance
	2D	3D	# Subjects	# Gallery Images	# Probe Images	
BenAbdelkader [6]	LFA,	LFA,	185	2/subject	2/subject	$\mathrm{EER}_{\mathrm{2D+3D}} = 0\%$
Tsalakanidou [106, 107]	gray EHMM	EHMM	50	5/subject	55/subject	$\mathrm{EER}_{\mathrm{2D+3D}} = 5.5\%$ $\mathrm{EER}_{\mathrm{2D}} = 9.3\%$ $\mathrm{EER}_{\mathrm{3D}} = 20\%$
Tsalakanidou [105]	color/gray EHMM	EHMM	20	5/subject	1402	$\mathrm{EER}_{\mathrm{2D+3D}} = 7.5\%$ $\mathrm{EER}_{\mathrm{color}} = 15.1\%$ $\mathrm{EER}_{\mathrm{2D}} = 13.8\%$
						$\mathrm{EER}_{\mathrm{3D}} = 16.7\%$
Tsalakanidou [107]	color/gray EHMM	EHMM	20	2-5/subject	2124	$\mathrm{EER}_{\mathrm{2D+3D}} = 7.7\%$ $\mathrm{EER}_{\mathrm{2D}} = 7.9\%$ $\mathrm{EER}_{\mathrm{3D}} = 14.8\%$ $\mathrm{RR}_{\mathrm{2D+3D}} = 77.3\%$
Malassiotis [66, 67]	color EHMM	EHMM	20	2-5/subject	–	$\mathrm{RR}_{\mathrm{color_3D}} = 99.1\%$ $\mathrm{RR}_{\mathrm{color}} = 90.7\%$ $\mathrm{RR}_{\mathrm{3D}} = 96.4\%$
Lu [62]	color	ICP, shape	18	1/subject	63	$\mathrm{RR}_{\mathrm{2D+3D}} = 84\%$
Lu [64, 65]	LDA	ICP	200	1/subject	598	$\mathrm{RR}_{\mathrm{2D+3D}} = 90\%$ $\mathrm{RR}_{\mathrm{2D}} = 77\%$ $\mathrm{RR}_{\mathrm{3D}} = 86\%$
Papatheodorou [81]	4D ICP	3D ICP	62	1/subject	5/subject	$\mathrm{FRR}_{\mathrm{2D+3D}} = 0\%$ $\mathrm{FAR}_{\mathrm{2D+3D}} = 0\%$
Beumier [11]	Profiles	Profiles	100	9/subject	–	$\mathrm{EER}_{\mathrm{2D+3D}} = 1.2\%$

Table 3: The table summarizes numerous 2D+3D face recognition algorithms. Here again, it can be observed that combining 2D and 3D modalities results in better performance, relative to the individual modalities.

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Gordon [37]	depth MSE	8	2/subject	1/subject	$\mathrm{RR} = 100\%$
Lao [55]	depth MSE	10	3/subject	1/subject	$\mathrm{RR} = 96\%$
Lee [59]	depth μ,σ MSE	35	1/subject	1/subject	Rank 10 $\mathrm{RR} = 100\%$
Medioni [70]	ICP	100	6/subject	1/subject	$\mathrm{EER} < 2\%$
Lu [61]	ICP and shape feature	18	1/subject	113	$\mathrm{RR} = 96.5\%$
Irfanöglu [49]	ICP	30	2/subject	1/subject	$\mathrm{RR} = 96.7\%$
Achermann [1]	Partial HD	24	5/subject	5/subject	$\mathrm{RR} = 99.2\%$
Lee [60]	depth weighted HD	42	1/subject	1/subject	Rank 5 $\mathrm{RR} = 98\%$
Pan [80]	HD-ICP	30	1/subject	2/subject	$\mathrm{EER} = 3.24\%$
Russ [96, 97]	HD-ICP	200	1/subject	1/subject	$\mathrm{FRR} = 2\%$ $\mathrm{FAR} = 0\%$
Gupta [40]	CW-SSIM	12	1/subject	29/subject	$\mathrm{EER} = 9.13\%$ Rank 1 $\mathrm{RR} = 98.6\%$

Table 4: The table presents the existing surface matching based 3D face recognition algorithms. For a majority of them, the mean squared error and the Hausdorff distance metrics have been employed along with the ICP algorithm. This approach has been fairly successful even with large data set of 200 subjects, but is limited by its high computational cost.

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Lee [56]	EGI	6 images	–	–	–
Tanaka [103, 104]	EGI	37 images	–	–	–
Paquet [82]	PFT	24 images	–	–	–
Chang [25]	PFT	15 images	–	–	–

Table 5: The table presents a summary of the 3D face recognition techniques based on the statistics of the orientation of surface normals. Clearly these techniques have not been greatly explored and warrant further investigations.

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Cartoux [19]	Profiles	5	1-3/subject	1/subject	RR = 100%
Pan [78]	HD Profile matching	30	1/subject	2/subject	EER = 2.22%
Nagamine [73]	Profiles	16	9/subject	1/subject	RR = 100%
Beumier [10, 12]	Profiles	120	1/subject	5/subject	EER = 4.75%
Zhang [122]	Profiles	166	166	32	RR = 96.9%

Table 6: This table summarizes the 3D face recognition algorithms that involve the matching of various facial profile curves. For all such techniques, the vertical central profile curve has resulted in the best performance.

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Gordon [38]	Curvature, Distances	8	2/subject	1/subject	RR = 100%
Moreno [72]	Curvature, Angles, Distances	60	1/subject	6/subject	RR = 78%
Lee [58]	Angles and Distances	100	–	–	RR = 96%
Gupta [39]	Euclidean and Geodesic Distances	105	1/subject	663	Rank 1 RR = 98.64% EER = 1.4%

Table 7: This table summarizes the studies based on geometric facial features, including curvatures, angles, Euclidean and geodesic distances between facial landmarks, for 3D face recognition.

Author	Method	Data Set			Performance
		# Subjects	# Gallery Images	# Probe Images	
Pan [79]	Profiles and surface matching	30	1/subject	2/subject	EER = 1.11%
Xu [119]	z values and local shape	30	5/subject	1/subject	RR = 96.1%
Gökberk [36]	LDA and Normals and Profiles	106	3/subject	2-3/subject	RR = 99.1%

Table 8: This table presents the ensemble approaches to 3D face recognition, where multiple 3D classifiers/representaions/fetaures are combined. For all analyses, the ensemble approach results in superior performance, relative to the individual 3D classifiers.

Author	Method		Data Set			Performance
	2D	3D	# Subjects	# Gallery Images	# Probe Images	
Chua [28]	–	Point Sign	6	4/subject	1/subject	RR = 100%
Wang [113]	2D Gabor	Point Sign	50	3/subject	3/subject	RR = 90%
Wang [111]	2D Gabor	3D Gabor	30	1/subject	3/subject	RR = 79.58%
Bronstein [17]	Isometric embedding	Isometric embedding	157	–	–	–
Bronstein [16]	–	Isometric embedding	30	65	155	EER = 1.9% RR = 100%
Chang [22]	–	ICP of rigid regions	449 FRGC v0.2	1/subject	1-9/subject	RR = 88.6%
Lu [63]	–	Spline deformation	100	1/subject	1-2/subject	FRR = 34% FAR = 1% RR = 91%

Table 9: The table summarizes 3D face recognition algorithms that have been reported to be robust to changes in facial expression to some degree. The two main ideologies are to to either transform the facial surface into an expression invariant form or to employ facial shape characteristics that are not significantly altered by changes in facial expression.

Author	Method		Data Set			Performance
	2D	3D	# Subjects	# Gallery Images	# Probe Images	
Pan [77]	–	Isometric flattening & PCA	276 FRGC v0.1	–	–	EER = 2.83%
Russ [97]	–	MSE and HD-ICP	198 FRGC v0.1	1/subject	–	FRR = 6.5% FAR = 0.1% RR = 98.5%
Koudelka [52]	–	HD-ICP	198 FRGC v0.1	1/subject	1/subject	RR = 94%
Kakadiaris [50]	–	AFM	275 FRGC v0.1	152	608	RR = 99.3%
Passalis [83]	–	AFM	466 FRGC v0.2	ROC III	ROC III	FRR = 3.6% FAR = 0.1%
Hüsken [46]	EBGM	EBGM	466 FRGC v0.2	ROC III	ROC III	FRR = 2.7% FAR = 0.1%
Maurer [69]	EBGM	ICP	466 FRGC v0.2	4007	4007	FRR = 6.5% FAR = 0.1%

Table 10: This table summarizes the 3D face recognition algorithms that have been tested on the Face Recognition Grand Challenge data sets. The three algorithms tested on the FRGC v0.2 data set, which contains images of 466 individual, can be regarded as the current state-of-the-art in 2D+3D face recognition. Notably, the algorithm based on 2D+3D EBGM performed the best of the three.

References

[1] B. Achermann and H. Bunke. Classifying range images of human faces with hausdorff distance. In *Proceedings 15th International Conference on Pattern Recognition*, volume vol.2, pages 809–813. IEEE Comput. Soc, Dept. of Comput. Sci., Bern Univ., Switzerland, 2000.

[2] B. Achermann, X. Jiang, and H. Bunke. Face recognition using range images. In *Proceedings. International Conference on Virtual Systems and MultiMedia*, pages 129–136. Int. Soc. Virtual Syst. & MultiMedia (VSMM), Inst. of Comput. Sci. & Appl. Math., Bern Univ., Switzerland, 1997.

[3] L. Akarun, B. Gokberk, and A. A. Salah. 3d face recognition for biometric applications. In *13th European Signal Processing Conference (EUSIPCO)*, Antalya, Turkey, Spetember 2005.

[4] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski. Face recognition by independent component analysis. *Neural Networks, IEEE Transactions on*, 13(6):1450–1464, 2002.

[5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):711–720, 1997.

[6] C. BenAbdelkader and P. A. Griffin. Comparing and combining depth and texture cues for face recognition. *Image and Vision Computing*, 23(3):339–352, 2005.

[7] P. J. Besl. Active, optical range imaging sensors. *Machine Vision and Applications*, 1(2):127–152, 1988.

[8] P. J. Besl and R. C. Jain. Segmentation through variable-order surface fitting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(2):167–192, 1988.

[9] P. J. Besl and H. D. McKay. A method for registration of 3-d shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(2):239–256, 1992.

[10] C. Beumier and M. Acheroy. Automatic 3d face authentication. *Image and Vision Computing*, 18(4):315–321, 2000.

[11] C. Beumier and M. Acheroy. Face verification from 3d and grey level clues. *Pattern Recognition Letters*, 22(12):1321–1329, 2001.

[12] C. Beumier, M. Acheroy, P. H. Lewis, and M. S. Nixon. Automatic face authentication from 3d surface. In P. H. Lewis and M. S. Nixon, editors, *BMVC 98. Proceedings of the Ninth British Machine Vision Conference*, volume vol.2, pages 449–458. Univ. Southampton, Signal & Image Centre, R. Mil. Acad., Brussels, Belgium, 1998.

[13] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(9):1063–1074, 2003.

[14] A. C. Bovik, M. Clark, and W. S. Geisler. Multichannel texture analysis using localized spatial filters. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(1):55–73, 1990.

[15] K. W. Bowyer, K. Chang, and P. J. Flynn. A survey of approaches and challenges in 3d and multi-modal 3d+2d face recognition. *Computer Vision and Image Understanding*, 101:1–15, 2006.

[16] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Three-dimensional face recognition. *International Journal of Computer Vision*, 64(1):5–30, 2005.

[17] A. M. Bronstein, M. M. Bronstein, R. Kimmel, J. Kittler, and M. S. Nixon. Expression-invariant 3d face recognition. In J. Kittler and M. S. Nixon, editors, *Audio- and Video-Based Biometric Person Authentication. 4th International Conference, AVBPA 2003. Proceedings (Lecture Notes in Computer Science Vol.2688)*, pages 62–69. Springer-Verlag, Dept. of Electr. Eng., Israel Inst. of Technol., Haifa, Israel, 2003.

[18] R. J. Campbell and P. J. Flynn. A survey of free-form object representation and recognition techniques. *Computer Vision and Image Understanding*, 81(2):166–210, 2001.

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

[20] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Face recognition using 2d and 3d facial data. In *Multimodal User Authentication, Workshop in*, pages 25–32, 2003.

[21] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Multimodal 2d and 3d biometrics for face recognition. In *Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on*, pages 187–194, 2003.

[22] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Adaptive rigid multi-region selection for handling expression variation in 3d face recognition. In *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, volume 3, pages 157–157, 2005.

[23] K. I. Chang, K. W. Bowyer, and P. J. Flynn. An evaluation of multimodal 2d+3d face biometrics. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(4):619–624, 2005.

[24] K. I. Chang, K. W. Bowyer, P. J. Flynn, and X. Chen. Multi-biometrics using facial appearance, shape and temperature. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pages 43–48, 2004.

[25] S. Chang, M. Rioux, and C. P. Grover. Range face recognition based on the phase fourier transform. *Optics Communications*, 222(1):143–153, 2003.

[26] R. Chellappa, C. L. Wilson, and S. Sirohey. Human and machine recognition of faces: a survey. *Proceedings of the IEEE*, 83(5):705–741, 1995.

[27] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10:145–155, 1992.

[28] C.-S Chua, F. Han, and Y.-K. Ho. 3d human face recognition using point signature. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 233–238, 2000.

[29] A. Colombo, C. Cusano, and R. Schettini. Tri-dimensional face detection and localization. In S. Santini, R. Schettini, and T. Gevers, editors, *SPIE Internet imaging VI, in Proceedings of*, volume 5670, pages 68–75, 2005.

[30] A. Colombo, C. Cusano, and R. Schettini. 3d face detection using curvature analysis. *Pattern Recognition*, 39(3):444–455, 2006.

[31] J. Davis and M. Goadrich. The relationship between precision-recall and roc curves. In *ICML '06: Proceedings of the 23rd international conference on Machine learning*, pages 233–240, New York, NY, USA, 2006. ACM Press.

[32] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, New York, 2nd edition, 2001.

[33] J. P. Egan. *Signal Detection Theory and ROC Analysis*. Academic Press, New York, 1975.

[34] R. A. Fisher. The use of multiple meaures in taxonomic problems. *Annals Eugenics*, 7:179–188, 1936.

[35] A. Godil, S. Ressler, and P. Grother. Face recognition using 3d facial shape and color map information: comparison and combination. In *Proceedings of the SPIE - The International Society for Optical Engineering*, volume 5404, pages 351–361, Nat. Inst. of Stand. & Technol., Gaithersburg, MD, USA, 2004. SPIE-Int. Soc. Opt. Eng.

[36] B. Gökberk, A. A. Salah, and L. Akarun. Rank-based decision fusion fro 3d shape-based face recognition. In *LNCS, International Conference on Audio- and Video-based Biometric Person Authentication*, volume 3546, pages 1019–1028, 2005.

[37] G. G. Gordon. Face recognition based on depth maps and surface curvature. In *SPIE Geometric methods in Computer Vision*, volume 1570, pages 234–247, 1991.

[38] G. G. Gordon. Face recognition based on depth and curvature features. In *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92., 1992 IEEE Computer Society Conference on*, pages 808–810, 1992.

[39] S. Gupta, M. K. Markey, J. K. Aggarwal, and A. C. Bovik. Three dimensional face recognition based on geodesic and euclidean distances. In *Electronic Imaging, Vision Geometry XV*, volume 6499 of *Proc. of SPIE*, San Jose, California, USA, Jan 28 - Feb 1 2007.

[40] S. Gupta, M. P. Sampat, Z. Wang, M. K. Markey, and A. C. Bovik. 3d face recognition using the complex-wavelet structural similarity metric. In *IEEE Workshop on Applications of Computer Vision*, Austin, TX, 2007.

[41] T. Heseltine, N. Pears, and J. Austin. Three-dimensional face recognition: A fisher-surface approach. In A. Campilho and M. Kamel, editors, *International Conference on Image Analysis and Recognition, In Proc. of the*, volume LNCS 3212 of *ICIAR 2004*, pages 684–691. Springer-Verlag Berlin Heidelberg, 2004.

[42] T. Heseltine, N. Pears, and J. Austin. Three-dimensional face recognition: an eigen-surface approach. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 2, pages 1421–1424 Vol.2, 2004.

[43] T. Heseltine, N. Pears, and J. Austin. Three-dimensional face recognition using surface space combinations. In *British Machine Vision Conference*, 2004.

[44] C. Hesher, A. Srivastava, and G. Erlebacher. Principal component analysis of range images for facial recognition. In *International Conference on Imaging Science, Systems, and Technology (CISST), in Proceedings of*, 2002.

[45] C. Hesher, A. Srivastava, and G. Erlebacher. A novel technique for face recognition using range imaging. In *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, volume 2, pages 201–204 vol.2, 2003.

[46] M. Hüsken, M. Brauckmann, S. Gehlen, and C. Von der Malsburg. Strategies and benefits of fusion of 2d and 3d face recognition. In *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, volume 3, pages 174–174, 2005.

[47] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the hausdorff distance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(9):850–863, 1993.

[48] A. Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley and Sons, New York, 2001.

[49] M. O. Irfanöglu, B. Gökberk, and L. Akarun. 3d shape-based face recognition using automatically registered facial surfaces. In *Proceedings of the 17th International Coneference on Pattern Recognition*, volume Vol.4, pages 183–186. IEEE Computer Society, Dept. of Comput. Eng., Bogazici Univ., Turkey, 2004.

[50] I. A. Kakadiaris, G. Passalis, T. Theoharis, G. Toderici, I. Konstantinidis, and N. Murtuza. Multimodal face recognition: combination of geometry with physiological information. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 1022–1029 vol. 2, 2005.

[51] M. Kirby and L. Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(1):103–108, 1990.

[52] M. L. Koudelka, M. W. Koch, and T. D. Russ. A prescreener for 3d face recognition using radial symmetry and the hausdorff fraction. In *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, volume 3, pages 168–168, 2005.

[53] A. Z. Kouzani, F. He, and K. Sammut. Towards invariant face recognition. *Information Sciences*, 123(1):75–101, 2000.

[54] E. P. Kukula, S. J. Elliott, R. Waupotitsch, and B. Pesenti. Effects of illumination changes on the performance of geometrix facevision/spl reg/ 3d frs. In *Security Technology, 2004. 38th Annual 2004 International Carnahan Conference on*, pages 331–337, 2004.

[55] S. Lao, Y. Sumi, M. Kawade, and F. Tomita. 3d template matching for pose invariant face recognition using 3d facial model built with isoluminance line based stereo vision. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 2, pages 911–916 vol.2, 2000.

[56] J. C. Lee and E. Milios. Matching range images of human faces. In *Computer Vision, 1990. Proceedings, Third International Conference on*, pages 722–726, 1990.

[57] M. W. Lee and S. Ranganath. Pose-invariant face recognition using a 3d deformable model. *Pattern Recognition*, 36(8):1835–1846, 2003.

[58] Y. Lee, H. Song, U. Yang, H. Shin, and K. Sohn. Local feature based 3d face recognition. In *Audio- and Video-based Biometric Person Authentication, 2005 International Conference on, LNCS*, volume 3546, pages 909–918, 2005.

[59] Y. Lee and T. Yi. 3d face recognition using multiple features for local depth information. In M. Grgic and S. Grgic, editors, *Proceedings EC-VIP-MC 2003. 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications (IEEE Cat. No.03EX667)*, volume vol.1, pages 429–434. Faculty of Electrical Eng. & Comput, Zagreb, 2003.

[60] Y.-H. Lee and J.-C. Shim. Curvature based human face recognition using depth weighted hausdorff distance. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 3, pages 1429–1432 Vol. 3, 2004.

[61] X. Lu, D. Colbry, and A. K. Jain. Three-dimensional model based face recognition. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 1, pages 362–366 Vol.1, 2004.

[62] X. Lu, D. Colbry, A. K. Jain, and D. Zhang. Matching 2.5d scans for face recognition. In D. Zhang and A.K. Jain, editors, *Biometric Authentication. First International Conference, ICBA 2004. Proceedings (Lecture Notes in Comput. Sci. Vol.3072)*, pages 30–36. Croucher Found., Dept. of Comput. Sci. & Eng., Michigan State Univ., East Lasning, MI, USA, 2004.

[63] X. Lu and A. K. Jain. Deformation analysis for 3d face matching. In *Applications of Computer Vision, WACV '05 7th IEEE Workshop on*, pages 99–104, 2005.

[64] X. Lu and A. K. Jain. Integrating range and texture information for 3d face recognition. In *Applications of Computer Vision, 2005. (WACV 2005). Proceedings. Sixth IEEE Workshop on*, 2005.

[65] X. Lu, A. K. Jain, and D. Colbry. Matching 2.5d face scans to 3d models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(1):31–43, 2006.

[66] S. Malassiotis and M. G. Strintzis. Pose and illumination compensation for 3d face recognition. In *Image Processing, 2004. ICIP '04. 2004 International Conference on*, volume 1, pages 91–94 Vol. 1, 2004.

[67] S. Malassiotis and M. G. Strintzis. Robust face recognition using 2d and 3d data: Pose and illumination compensation. *Pattern Recognition*, 38(12):2537–2548, 2005.

[68] R. S. Malpass and J. Kravitz. Recognition for faces of own and other race. *Journal of Personality and Social psychology*, 13:330–334, 1969.

[69] T. Maurer, D. Guigonis, I. Maslov, B. Pesenti, A. Tsaregorodtsev, D. West, and G. Medioni. Performance of geometrix activeid^{TM} 3d face recognition engine on the frgc data. In *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, volume 3, pages 154–154, 2005.

[70] G. Medioni and R. Waupotitsch. Face modeling and recognition in 3-d. In *Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on*, pages 232–233, 2003.

[71] C. E. Metz. Basic principles of roc analysis. *Seminars in Nuclear Medicine*, 8(4):283–298, 1978.

[72] A. B. Moreno, A. Sanchez, J. Fco, V. Fco, and J. Diaz. Face recognition using 3d surface-extracted descriptors. In *Irish Machine Vision and Image Processing Conference (IMVIP 2003)*, Sepetember 2003.

[73] T. Nagamine, T. Uemura, and I. Masuda. 3d facial image analysis for human identification. In *Pattern Recognition, 1992 . Vol.1. Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on*, pages 324–327, 1992.

[74] D. Nandy and J. Ben-Arie. Shape from recognition: a novel approach for 3-d face shape recovery. *Image Processing, IEEE Transactions on*, 10(2):206–217, 2001.

[75] C. Nastar and A. Pentland. Matching and recognition using deformable intensity surfaces. In *Computer Vision, 1995. Proceedings., International Symposium on*, pages 223–228, 1995.

[76] A. V. Nefian and III Hayes, M. H. An embedded hmm-based approach for face detection and recognition. In *Acoustics, Speech, and Signal Processing, 1999. ICASSP '99. Proceedings., 1999 IEEE International Conference on*, volume 6, pages 3553–3556 vol.6, 1999.

[77] G. Pan, S. Han, Z. Wu, and Y. Wang. 3d face recognition using mapped depth images. In *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, volume 3, pages 175–175, 2005.

[78] G. Pan, Y. Wu, and Z. Wu. Investigating profile extracted from range data for 3d face recognition. In *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, volume 2, pages 1396–1399 vol.2, 2003.

[79] G. Pan, Y. Wu, Z. Wu, and W. Liu. 3d face recognition by profile and surface matching. In *Neural Networks, 2003. Proceedings of the International Joint Conference on*, volume 3, pages 2169–2174 vol.3, 2003.

[80] G. Pan, Z. Wu, and Y. Pan. Automatic 3d face verification from range data. In *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on*, volume 3, pages III–133–6 vol.3, 2003.

[81] T. Papatheodorou and D. Rueckert. Evaluation of automatic 4d face recognition using surface and texture registration. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pages 321–326, 2004.

[82] E. Paquet, H. H. Arsenault, and M. Rioux. Recognition of faces from range images by means of the phase fourier transform. *Pure and Applied Optics*, 4(6):709–721, 1995.

[83] G. Passalis, I. A. Kakadiaris, T. Theoharis, G. Toderici, and N. Murtuza. Evaluation of 3d face recognition in the presence of facial expressions: an annotated deformable model approach. In *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, volume 3, pages 171–171, 2005.

[84] P. S. Penev and J. J. Atick. Local feature analysis: a general statistical theory for object representation. *Network: Computation in Neural Systems*, 7:477–500, 1996.

[85] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 947–954 vol. 1, 2005.

[86] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, and W. Worek. Preliminary face recognition grand challenge results. In *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, pages 15–24, 2006.

[87] P. J. Phillips, A. Martin, C. L. Wilson, and M. Przybocki. An introduction to evaluating biometric systems. *Computer*, 33(2):56–63, 2000.

[88] P. J. Phillips, H. Moon, P. Rauss, and S. A. Rizvi. The feret evaluation methodology for face-recognition algorithms. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 137–143, 1997.

[89] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face-recognition algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10):1090–1104, 2000.

[90] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss. The feret database and evaluation procedure for face-recognition. *Image and Vision Computing*, 16(5):295–306, 1998.

[91] P.J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone. Frvt 2002: Overview and summary. available at www.frvt.org, March 2003.

[92] S. S. Rakover and B. Cahlon. *Face Recognition: Cognitive and Computational Processes*. John Benjamins Publishing Company, Amsterdam, 2001.

[93] N. K. Ratha, A. Senior, R. M. Boile, S. Singh, N. Murshed, and W. Kropatsch. Automated biometrics. In S. Singh, N. Murshed, and W. Kropatsch, editors, *Advances in Pattern Recognition - ICAPR 2001. Second International Conference. Proceedings (Lecture Notes in Computer Science Vol.2013)*, pages 445–453. Springer-Verlag, IBM Thomas J. Watson Res. Center, Yorktown Heights, NY, USA, 2001.

[94] S. A. Rizvi, H. Moon, and P. J. Phillips. The feret verification testing protocol for face recognition algorithms. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pages 48–53, 1998.

[95] S. L. Rogers. *The Personal Identification of Living Individuals*. Charles C Thomas, Springfield, Illinois, 1986.

[96] T. D. Russ, M. W. Koch, and C. Q. Little. 3d facial recognition: a quantitative analysis. In *Security Technology, 2004. 38th Annual 2004 International Carnahan Conference on*, pages 338–344, 2004.

[97] T. D. Russ, M. W. Koch, and C. Q. Little. A 2d range hausdorff approach for 3d face recognition. In *Computer Vision and Pattern Recognition, 2005 IEEE Computer Society Conference on*, volume 3, pages 169–169, 2005.

[98] F. Samaria and S. Young. Hmm-based architecture for face identification. *Image and Vision Computing*, 12(8):537–543, 1994.

[99] B. Scheck, P. Neufeld, and J. Dwyer. *Actual Innocence*. Doubleday, New York, 2000.

[100] A. Scheenstra, A. Ruifrok, and R. C. Veltkarnp. A survey of 3d face recognition methods. In T. Kanade, A. K. Jain, and N. K. Ratha, editors, *Audio- and Video-Based Biometric Person Authentication. 5th International Conference, AVBPA 2005. Proceedings (Lecture Notes in Computer Science Vol.3546)*, pages 891–899. Springer-Verlag, Inst. of Inf. & Comput. Sci., Utrecht Univ., Netherlands, 2005.

[101] W. Shen, M. Surette, and R. Khanna. Evaluation of automated biometrics-based identification and verification systems. *Proceedings of the IEEE*, 85(9):1464–1478, 1997.

[102] A. Srivastava, X. Liu, and C. Heher. Face recognition using optimal linear components of range images. Accepted for publishing in Image and Vision Computing, 2003.

[103] H. T. Tanaka and M. Ikeda. Curvature-based face surface recognition using spherical correlation-principal directions for curved object recognition. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, volume 3, pages 638–642 vol.3, 1996.

[104] H. T. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical correlation. principal directions for curved object recognition. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pages 372–377, 1998.

[105] F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis. Exploitation of 3d images for face authentication under pose and illumination variations. In *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, pages 50–57, 2004.

[106] F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis. Integration of 2d and 3d images for enhanced face authentication. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pages 266–271, 2004.

[107] F. Tsalakanidou, S. Malassiotis, and M. G. Strintzis. Face localization and authentication using color and depth images. *Image Processing, IEEE Transactions on*, 14(2):152–168, 2005.

[108] F. Tsalakanidou, D. Tzovaras, and M. G. Strintzis. Use of depth and colour eigenfaces for face recognition. *Pattern Recognition Letters*, 24(9):1427–1435, 2003.

[109] S. Tsutsumi, S. Kikuchi, and M. Nakajima. Face identification using a 3d gray-scale image-a method for lessening restrictions on facial directions. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pages 306–311, 1998.

[110] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3:7186, 1991.

[111] Y. Wang and C. Chua. Face recognition from 2d and 3d images using 3d gabor filters. *Image and Vision Computing*, 23(11):1018–1028, 2005.

[112] Y. Wang and C. Chua. Robust face recognition from 2d and 3d images using structural hausdorff distanec. *Image and Vision Computing*, 24:176–185, 2006.

[113] Y. Wang, C. Chua, and Y. Ho. Facial feature detection and face recognition from 2d and 3d images. *Pattern Recognition Letters*, 23(10):1191–1202, 2002.

[114] Z. Wang and E. P. Simoncelli. Translation insensitive image similarity in complex wavelet domain. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, volume 2, pages 573–576, 2005.

[115] H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie, , and T. S. Huang, editors. *Face Recognition: From Theory to Applications*. Springer-Verlag, Berlin, 1998.

[116] G. L. Wells and E. P. Seelau. Eyewitness identification: Psychological research and legal policy on lineups. *Psychology, Public Policy, and Law*, 1:765–791, 1995.

[117] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):775–779, 1997.

[118] www.frvt.org, 2002.

[119] C. Xu, Y. Wang, T. Tan, and L. Quan. Automatic 3d face recognition combining global geometric features with local shape variation information. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pages 308–313, 2004.

[120] Y. Yacoob and L. Davis. Labeling of human face components from range data. In *Computer Vision and Pattern Recognition, 1993. Proceedings CVPR '93., 1993 IEEE Computer Society Conference on*, pages 592–593, 1993.

[121] C. Zhang and F. S. Cohen. 3-d face structure extraction and recognition from images using 3-d morphing and distance mapping. *Image Processing, IEEE Transactions on*, 11(11):1249–1259, 2002.

[122] L. Zhang, A. Razdan, G. Farin, J. Femiani, M. Bae, and C. Lockwood. 3d face authentication and recognition based on bilateral symmetry analysis. *Visual Computer*, 22(1):43–55, 2006.

[123] W. Zhao, R. Chellappa, and P. Phillips. Subspace linear discriminant analysis for face recognition. Technical Report CAR-TR-914, Center for Automation Research, University of Maryland, College Park, 1999.

[124] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: a literature survey. *ACM Computing Surveys*, 35(4):399–459, 2003.