3-D face recognition: Features, databases, algorithms and challenges

3 authors:

Hemprasad Patil
VIT University
14 PUBLICATIONS   **108** CITATIONS

SEE PROFILE

Ashwin Kothari
Visvesvaraya National Institute of Technology
68 PUBLICATIONS   **406** CITATIONS

SEE PROFILE

K. M. Bhurchandi
Visvesvaraya National Institute of Technology
73 PUBLICATIONS   **408** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Rough set Based UANN based pattern classifier View project

MIMO antenna for WLAN View project

3-D face recognition: features, databases, algorithms and challenges

Hemprasad Patil, Ashwin Kothari & Kishor Bhurchandi

Volume

ONLINE
FIRST

A
I
R

Artificial
Intelligence
Review

⌂ Springer

⌂ Springer

Springer

CrossMark

3-D face recognition: features, databases, algorithms and challenges

Hemprasad Patil[1] · **Ashwin Kothari**[1] ·
Kishor Bhurchandi[1]

Abstract Face recognition is being widely accepted as a biometric technique because of its non-intrusive nature. Despite extensive research on 2-D face recognition, it suffers from poor recognition rate due to pose, illumination, expression, ageing, makeup variations and occlusions. In recent years, the research focus has shifted toward face recognition using 3-D facial surface and shape which represent more discriminating features by the virtue of increased dimensionality. This paper presents an extensive survey of recent 3-D face recognition techniques in terms of feature detection, classifiers as well as published algorithms that address expression and occlusion variation challenges followed by our critical comments on the published work. It also summarizes remarkable 3-D face databases and their features used for performance evaluation. Finally we suggest vital steps of a robust 3-D face recognition system based on the surveyed work and identify a few possible directions for research in this area.

1 Introduction

Face recognition is becoming a commonly used biometric technique with widespread applications in public records, authentication, intelligence, safety, security, and many other vigilance systems. Unlike other biometric traits such as iris, fingerprint, palm-print, signature, vein-

✉ Hemprasad Patil
hemprasadpatil@gmail.com

Ashwin Kothari
ashwinkothari@ece.vnit.ac.in

Kishor Bhurchandi
bhurchandikm@ece.vnit.ac.in

1   Visvesvaraya National Institute of Technology, Nagpur, India

 Springer

pattern etc., which require a user to be attentive and intentional while capturing input, face recognition requires least user attention. An ideal face recognition system is expected to detect and recognize a human face when the person is unaware of being biometrically examined. This reduces the harm of altering fiducial identity features by vicious persons. Also, tampering of basic facial features is difficult and can be easily noticed. This may make it the most reliable and preferred candidate amongst the other biometric techniques in the days to come (Li and Jain 2011).

Face recognition is defined as either online or offline recognition of one or more people using available face database (Zhao et al. 2003). Face recognition is classified into two domains namely, identification and verification. For an identification/recognition system, input face is matched against all faces in an available database, which in turn provides the determined identity (Zhao et al. 2003). A verification/authentication system compares an input face with a similar claimed face from a database. It either validates or rejects the claim based on the matching score. The input to a face recognition system is usually referred to as a 'query' or 'probe', whereas collection of faces in dataset is denoted as a 'gallery' or a 'database' (Bowyer et al. 2006).

After conduction of Face Recognition Vendor Tests FRVT-2000 (Blackburn et al. 2001), FRVT-2002 (Phillips et al. 2003), researchers realized the importance of 3-D human facial scans. A 3-D facial scan conveys much more discriminating information, as it provides the exact shape information including the depth profile associated with a human face. After the availability of 3-D scanners and subsequent 3-D face databases, many researchers focused their energies toward 3-D face recognition, expecting to utilize the more precise information associated with a 3-D facial shape (Bowyer et al. 2006; Xi et al. 2011; Ballihi et al. 2012; Mohammadzade and Hatzinakos 2013). The acquired 3-D facial data is invariant to illumination (Al-Osaimi et al. 2012; Smeets et al. 2010) and pose changes (Ocegueda et al. 2013). In fact all other poses and occlusion invariant features can be computationally estimated using a complete 3-D frontal scan, even in presence of occlusions, improper illumination or expression variations. Research efforts on 3-D face analysis are focused on improvement of recognition accuracy and to address many challenges faced by these systems such as expression variations and occlusions. The 3-D face recognition systems are more immune to spoofing and deception (Määttä et al. 2012). Various agencies organized the Face Recognition Grand Challenge (FRGC) (Phillips et al. 2005) and FRVT-2006 (Phillips et al. 2010) evaluation tests which involved assessment of 3-D face recognition algorithms on databases with large gallery size. This indicates the growing interest of the research community in 3-D face recognition. Bowyer et al. (2006) have summarized the research trends in 3-D face recognition up to 2006. Abate et al. (2007) have reviewed the associated literature up to year 2007. Smeets et al. (2012) have recently studied various algorithms for expression invariant 3-D face recognition and evaluated the complexity of existing 3-D face databases. Some of the earlier developments in face recognition are surveyed by Chellappa et al. (1995), Jaiswal et al. (2011), Zhao et al. (2003) and Tan et al. (2006).

This paper surveys published work on 3-D face recognition from a general viewpoint, and also reviews it with special focus on some challenges such as expression and occlusion variations. An extensive review of prominent 3-D face databases that will help forthcoming researchers to choose a database for their research problem is also presented. There are some multimodal approaches based on both shape (3-D) and texture (2-D) data. The joint use of these shape and textural features accommodates more information and facilitates extraction of more discriminant features. We have reviewed the literature associated only with 3-D shape, and also approaches that address 3-D face recognition and related challenges. The

survey of pure 2-D and multimodal (2-D and 3-D) face recognition techniques is outside the scope of this paper.

The rest of this paper is organized as follows: in Sect. 2 a brief overview of 3-D face recognition systems, various 3-D data acquisition and representation techniques is presented along with their advantages and limitations. It also includes a summary of 3-D landmark detection, face segmentation, face registration, feature extraction and matching techniques. Section 3 summarizes prominent 3-D face databases. In this section, complexity of each database is examined for challenges such as pose, expression and occlusion variations. The surveyed state-of-the-art methods for expression invariant 3-D face recognition are discussed in Sect. 4 along with their advantages and limitations. Section 5 addresses the review of occlusion invariant 3-D face recognition approaches. Section 6 presents a discussion on various challenges faced by the published research so far and also summarizes various algorithmic steps in 3-D face recognition. Section 7 presents the conclusion and future scope of research in 3-D face recognition approaches.

2 3-D face recognition system

3-D face recognition is defined as a biometric technique which uses individual 3-D facial shape to recognize human faces using 3-D models of both probe and gallery faces. A general 3-D face recognition system is depicted in Fig. 1.

A 3-D human face captured during face acquisition may contain unwanted body parts or areas like hair, ears, neck, shoulders and accessories like glasses and ornaments that need to be effectively eliminated. Major landmarks facilitate the segmentation process which extracts face shape from the entire scan. Facial shape needs to be aligned before actual matching. 3-D registration techniques are used for face shape alignment. The discriminant features are extracted and stored using the surveyed region-based and holistic approaches for all faces in the gallery. Feature vector for a probe (query) scan is extracted and matched against the
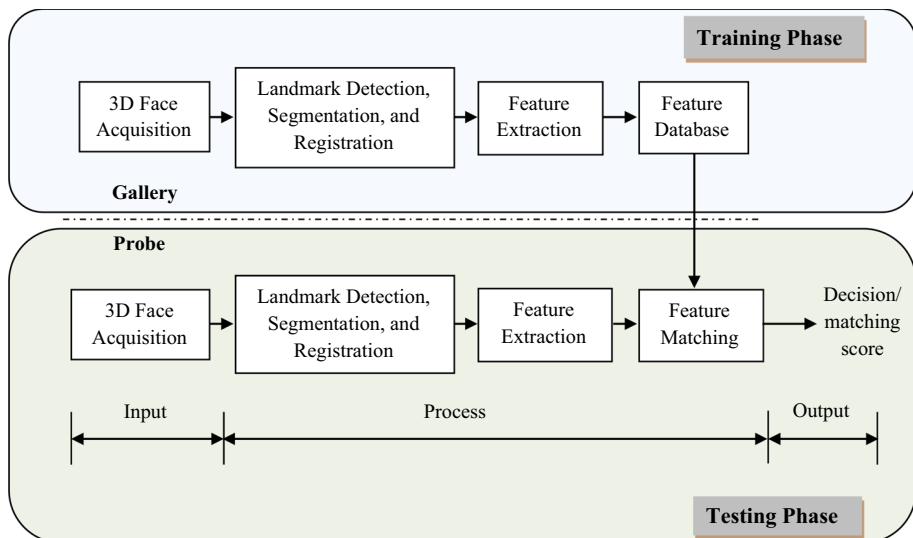

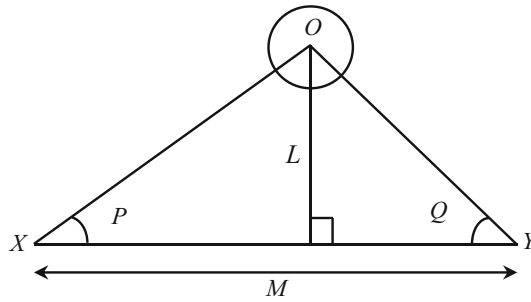**Fig. 1** A general 3-D face recognition system

Fig. 2 Triangulation principle

gallery feature vectors one by one. The gallery scan that has the closest matching distance with the probe scan below a predefined distance threshold is considered as the recognized gallery scan, and the class of the corresponding feature vector is declared as the result.

2.1 3-D facial data acquisition techniques

The methods to acquire 3-D information about a facial surface adopt optical triangulation principle for depth estimation (Amor et al. 2013) and are broadly classified into stereo acquisition, laser beam scanning and fringe pattern based techniques.

Triangulation principle Triangulation principle (Haasbroek 1968) states that the depth (L) and coordinates of a point (O) on an object in 3-D space can be estimated if length (M) of one side of a triangle OXY and angles P and Q between line XY and other sides of the triangle i.e. lines OX and OY at points X and Y could be measured depicted in Fig. 2. The points X and Y represent an illumination source and a sensor or both can be sensors like cameras.

$$M = \frac{L}{\tan P} + \frac{L}{\tan Q}$$

$$M = L\left(\frac{\cos P}{\sin P} + \frac{\cos Q}{\sin Q}\right)$$

$$M = L\left(\frac{\sin(P + Q)}{\sin P \cdot \sin Q}\right)$$

$$L = M\left(\frac{\sin P \cdot \sin Q}{\sin(P + Q)}\right) \tag{1}$$

Thus, depth L of each point O on a 3-D shape can be estimated using (1).

2.1.1 Stereo acquisition

Stereo acquisition based systems capture the 3-D facial data with two calibrated 2-D cameras. The concurrently captured 2-D images are registered and integrated to form a 3-D model using visual features and landmarks in both (Winkler and Min 2013). This is a passive method of 3-D face acquisition (Uchida et al. 2005) and involves complex computations since a 3-D face model is reconstructed from multiple 2-D images. The accuracy of 3-D face model heavily depends upon the surrounding illumination conditions. Illumination intensity should be consistent during the acquisition of 2-D calibrated images (Gökberk et al. 2009) to avoid erratic registration. Figure 3 presents a schematic for stereo 3-D image acquisition.
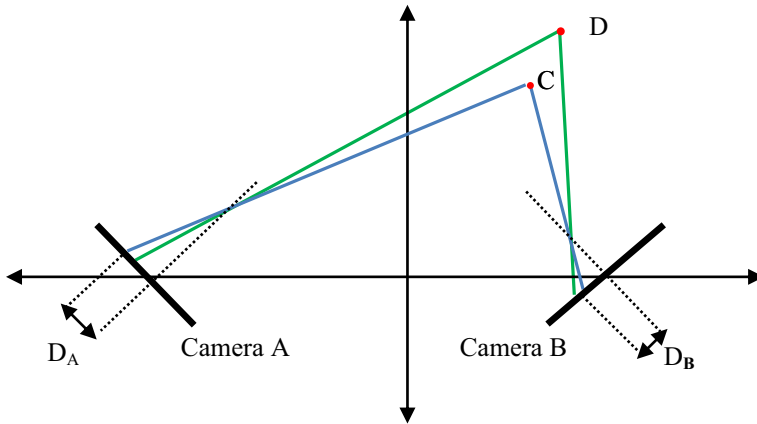
Fig. 3 Depth estimation using Stereo acquisition technique

Locations of calibrated cameras and an object form a triangle and hence the depth can be estimated with the triangulation principle. The manifold images captured using a set of cameras share some common points in 3-D space. Such shifted common points are known as disparity (Koschan et al. 2007). If the value of disparity (d) is small, the estimated depth is large and vice-versa. Figure 3 shows stereo depth acquisition of point C using camera A and camera B. The distances D_A and D_B are measured from the centre of the camera lens. Subsequently, the disparity is computed using (2) (Koschan et al. 2007).

$$d = |D_A - D_B| \tag{2}$$

It can be interpreted that if the point D is relatively farther from the camera planes, its disparity will be lesser (Koschan et al. 2007). In this way, the depth of various points in 3-D space is computed using stereo acquisition.

This method depends upon point-pair correspondence and feature extraction in addition to the conventional triangulation (Sansoni et al. 2009). The stereo acquisition technique is advantageous due to its high sensing speed, ease of availability of acquisition devices, no specific wavelength illumination source requirement and low cost. This technique suffers from errors in exact localization of point-pairs in distinct images which affect the overall reconstruction process. The process of conversion of the captured 2-D images into a 3-D model is computationally heavy. The resultant reconstruction accuracy is low, restricted to well illuminated scenes and depends upon resolution quality of captured 2-D images (Sansoni et al. 2009).

2.1.2 Laser beam scanning

Laser beam based 3-D scanner projects a laser beam on the subject to be scanned, i.e. a human face. The beam gradually scans the whole object just like a 2-D scan. A charge-coupled device camera is used to acquire the reflected light from the subject along with a 2-D intensity image. The 3-D data is generated using triangulation process, which yields the depth information. Laser beam based scanners offer the highest capturing range (∼8 feet) as compared to other acquisition techniques (∼1.6 to 5 feet) (Gökberk et al. 2009). Figure 4 presents the basic principle of a laser beam scanning system for 3-D objects.
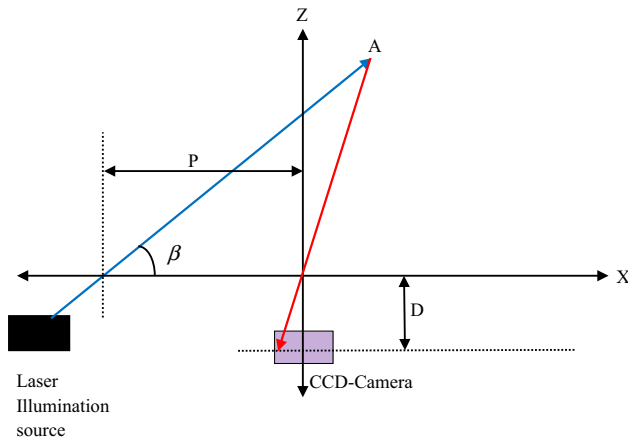
Fig. 4 Single spot laser beam depth acquisition

The depth estimation is contingent upon bending of the emitted laser light. The laser source emits single wavelength laser light to point A in 3-D space at an angle β with the baseline. The intersection of laser source and baseline is P units away from the origin. The focal distance of CCD camera is denoted by D (Koch et al. 2012; Koschan et al. 2007). The 3-D position i.e. (X, Y, Z) coordinates of point A are given by (3).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{P}{D \cot \beta - x} \begin{pmatrix} x \\ y \\ D \end{pmatrix} \tag{3}$$

This is an active triangulation technique which involves laser source as an illumination device and CCD camera as a sensor. The CCD camera is tuned to the wavelength of the transmitted laser light. The advantages of laser beam scanning method are high accuracy due to tight focus of laser source, illumination robustness due to utilization of a single standard wavelength and less power requirements (Scopigno et al. 2002). This method is slow in terms of acquisition speed due to line-wise scanning (Daoudi et al. 2013) of the complete 3-D surface. These devices are costlier than other 3-D acquisition devices.

2.1.3 Fringe pattern acquisition using structured light

This technique employs special system cameras to capture reflections from facial surfaces illuminated with fringe patterns. This is an active triangulation technique where the source projects encoded structured light patterns on the object to be scanned (Sansoni et al. 2009). The fringe patterns are predefined for specific angles and coded using structured light. Generally halogen lights project the structured light pattern on a facial surface (Malassiotis and Strintzis 2005; Breitenstein et al. 2008; Gökberk et al. 2009; Je et al. 2013). These systems are expensive compared to the stereo acquisition devices and offer more legitimate 3-D facial surface representation. They are less expensive and also less accurate compared to laser beam scanner devices. Analogous to stereo acquisition methods, the captured data also depends upon illumination conditions (Gökberk et al. 2009). Basic principle of this method is depicted in Fig. 5.
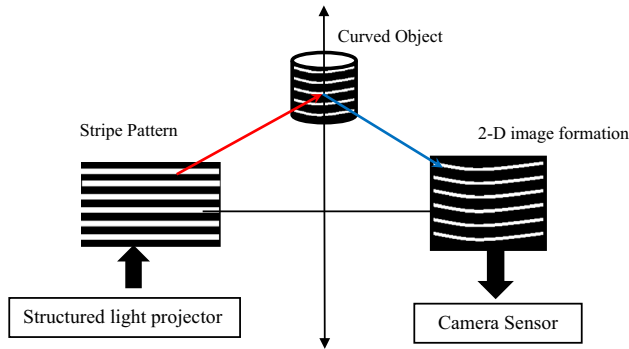
Fig. 5 Depth estimation using structured light projection

Various types of structured light pattern consist of manifold slits, multi-color (RGB), pixel-grid and dotted patterns (Sansoni et al. 2009). This technique is advantageous due to its wide Field-Of-View (FOV), high speed of acquisition, low cost, less computations for conversion from 2-D to 3-D, computerized control and implementation simplicity. This technique suffers from haziness due to use of single multi-stripe encoded pattern (Scopigno et al. 2002; Zhang and Lu 2013).

2.2 3-D face representations

2.2.1 Point cloud representation

Each point on the 3-D point cloud is denoted by $\{x, y, z\}$ coordinates. Most of the scanners use this representation (Beumier and Acheroy 2000; Chenghua et al. 2004) in order to store the captured 3-D facial information. Sometimes texture attributes are also concatenated to the shape information. In this case, the representation simply becomes $\{x, y, z, p, q\}$ (Amor et al. 2013), where p and q are spatial coordinates. The disadvantage of this representation is that the neighborhood information is not available as each point is simply expressed as three/five attribute feature vector (Gokberk et al. 2008). Figure 6a indicates point cloud representation for a sample face from the GavabDB 3-D face database (Moreno and Sanchez 2004). Point cloud is a collection of unstructured coordinates in the 3-D space $\Re^3$ (Koch et al. 2012). Often the point cloud data is fitted to a smooth surface to avoid drastic variations due to noise.

2.2.2 3-D Mesh representation

The 3-D Mesh representation uses pre-computed and indexed local information about the 3-D surface (Amor et al. 2013). It requires more memory and storage space than point cloud representation, but it is more preferred as it is flexible and more suitable for 3-D geometric transformations such as translation, rotations and scaling. Each 3-D polygonal mesh is expressed as a collection of mesh elements: vertices (points), edges (connectors between vertices) and polygons (shapes formed by edges and vertices) (Amor et al. 2013). 3-D mesh representation for an exemplary face belonging to the GavabDB database (Moreno and Sanchez 2004) is shown in Fig. 6b. Most of the 3-D mesh data is represented in terms of triangular mesh elements derived from the 2-D representation in $\Re^3$ (Koch et al. 2012). The 2-D texture coordinates are also embedded in the vertex information which is helpful for construction of accurate 3-D models.
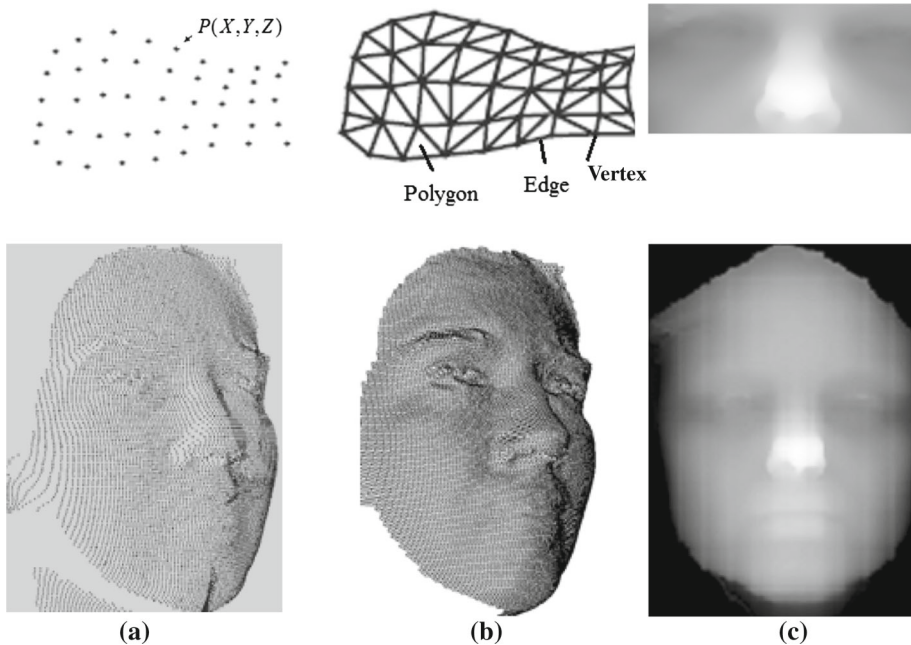
Fig. 6 3-D face representations. **a** Point cloud, **b** 3-D mesh and **c** depth image representation

2.2.3 Depth image representation

Depth images are also called 2.5D or range image representations (Bowyer et al. 2006). The Z-axis intensities of facial information are mapped onto a 2-D plane and the model appears like a smooth 3-D surface represented in terms of intensity. It gives realization of continuous depth information corresponding to various points on the facial shape. Since it is a 2-D representation, many existing 2-D image processing approaches can readily be applied to this representation. Figure 6c indicates a representative 3-D depth image from the Texas 3-D face database (Gupta et al. 2010a). Range image is an orthogonal projection of Z-axis coordinates on the 2-D image surface where uniform sampling is achieved (Koch et al. 2012). Range representation needs fewer memory resources as variable sampling is transformed into uniform sampling. Range images can be readily visualized with gray scale image viewers. Range images can be converted to 3-D mesh format as per the application requirements using the triangulation principle. Figure 6c depicts range image representation of a facial surface where the nose is closer to the camera. Thus, the nose has highest intensity and the intensity diminishes with increase in distance from the camera.

As discussed above, the 3-D facial information is usually represented in the three forms namely; point cloud data, mesh data and depth image representation as shown in Fig. 6a–c. The difference between the isolated points in point cloud, continuous edges and vertices in mesh and smooth surface representation in range data followed by their respective computer screen display shots can be clearly observed.

2.3 Landmark detection and face segmentation

Most of the published 3-D face recognition approaches assume manually localized and registered faces in a common coordinate system (Colombo et al. 2006). However, while designing

Table 1 Various landmark detection approaches

Sr. no.	Author and year	Landmarks detected	Method	Initial pose: frontal?	Initial pose: non-frontal?
1	Koudelka et al. (2005)	NT, IE	Radial symmetry maps	Yes	No
2	Malassiotis and Strintzis (2005)	NT	Principal curvature	Yes	Yes
3	Colombo et al. (2006)	NT, IE	Mean and Gaussian curvature maps	Yes	Yes
4	Conde et al. (2006a)	NT, IE	Mean curvature maps	Yes	Yes
5	Xu et al. (2006b)	NT	Effective energy	Yes	Yes
6	D'Hose et al. (2007)	NT, OE	Gabor wavelets	Yes	No
7	Segundo et al. (2007)	NT, IE	Mean and Gaussian curvature maps	Yes	Yes
8	Breitenstein et al. (2008)	NT	Single signature	Yes	Yes
9	Dibeklioglu et al. (2008)	NT	Principal curvature	Yes	Yes
10	Yi and Lijun (2008)	NT, IE	2-means clustering	Yes	Yes
11	Romero and Pears (2009)	NT, IE	Cylindrically Sampled RBR histograms (CSR)	Yes	Yes
12	Chew et al. (2009)	NT	Effective energy	Yes	Yes
13	Nair and Cavallaro (2009)	NT, IE, OE	Shape index and curvedness index	Yes	Yes
14	Szeptycki et al. (2009)	NT, IE, OE	Mean and Gaussian curvature maps	Yes	No
15	Peng et al. (2011)	NT	2-D profiles based	Yes	Yes
16	Mohammadzade and Hatzinakos (2013)	NT	Eigen-nose space	Yes	Yes

NT nose tip, *IE* inner eye-corners, *OE* outer eye-corners

a complete face recognition system, automated detection and registration of human faces is essential. Facial landmarks detection plays a vital role in automatic face localization and registration process. Facial landmarks consist of a set of fiducial points defined by the anthropometric studies (Farkas 1994), which can be used for face detection, face recognition, facial expression recognition and pose estimation applications (Xi et al. 2011). Many approaches utilize landmarks (such as nose tip and eye corners) to segment a 3-D facial surface from background and other body parts. The nose tip is invariant to most expression changes and it is noticeable for a broad range of non-frontal poses (Malassiotis and Strintzis 2005). Various methods for landmark(s) detection on 3-D facial surfaces are surveyed for landmarks namely the nose-tip (NT), inner eye-corners (IE) and outer eye-corners (OE) which are also known as PRN (pronasal), EN (endocanthion) and EX (exocanthion) points by anthropometric experts (Vezzetti and Marcolin 2012). In Table 1, we enlist the prominent landmark localization approaches along with the detected landmarks and pose flexibility. This subsection is further dedicated to the discussions and our comments on the work summarized in Table 1. A frontal 3-D face registration can be accomplished with only one landmark by means of the iterative closest point (ICP) algorithm, whereas, for non-frontal faces, location information of minimum three landmarks is essential (Colbry et al. 2005; Vezzetti and Marcolin 2012).

The 3-D surface derivatives play a vital role in computation of the mean (H) and Gaussian (K) curvature maps. A facial surface is partitioned into concave, convex and saddle regions by employing signs of the HK maps for segmentation. The nose tip and eye corners are searched within the segmented regions by applying various thresholding algorithms. Colombo et al. (2006) detected the eyes in the concave segmented HK maps (H_{min} and K_{min}) and a nose tip around peak values of the mean curvature map (H_{max}). However, the experimentation is performed on a small dataset. Detection rate of the reported approach is affected by self-occlusions and camouflage. Another method to detect inner eye corners and nose tip was proposed by Conde et al. (2006a). The regions with high values of mean curvature (H) were segmented into three different sets with clustering. Each cluster contained a feature point corresponding to nose tip and inner eye corners. This approach further employed SVM classifier for feature point localization. An advantage of this approach is that it can be directly applied on a 3-D mesh, acquired using laser beam scanning. To locate the nose tip and inner eye corners in a range image (Segundo et al. 2007) employed HK classification and median filtering. Principal component analysis (PCA) was utilized to compute surface normal vectors. Eye-corners are searched in concave regions and nose tip is searched in convex surface. Performance of this method is adversely affected when input range image has pose variations more than $\pm 15°$. In Szeptycki et al. (2009), a coarse-to-fine rotation invariant method was proposed to detect nose tip and inner eye corners. A generic facial statistical model was built and fitted to locate outer eye corners and other fiducial landmarks. Though these landmark detection methods use computationally heavy techniques, in effect they are computationally moderate due to their selective or guided search techniques.

Principal curvatures based nose tip detection method was proposed in Malassiotis and Strintzis (2005). An elliptical region of interest was marked on the range image to locate a nose tip. Nose tip value was detected by considering the regions with large values of the principal curvatures (k_1 and k_2). Principal curvatures k_1 and k_2 are the solutions of the quadratic equation $\lambda^2 + 2H\lambda + K = 0$ and built upon from the mean (H) and Gaussian (K) maps of the facial surface. This method is robust for large variations in pose and expressions. In Dibeklioglu et al. (2008), the surface curvature maps were computed as a part of pre-processing of a range image. Then, a difference map between Gaussian and mean curvatures was transformed to logarithmic scale for nose tip localization. The eye corners clustering based approach was proposed in Yi and Lijun (2008) for localization of nose tip and inner eye corners. The possible inner eye candidate's vertices were segmented from the rest of the 3-D mesh using concave regions of principal curvature based maps. The authors have employed the 2-means clustering method for separation of left and right eye corners. A flat surface was fitted to the candidate clusters using least summation distance method. A point on the positive side of the flat surface with maximum distance from clusters was marked as nose tip. Performance of the surface curvature based approaches is sensitive to the noise present in facial range images. Statistical shape based point distribution model (PDM) was presented for the extraction of nose tip and eye corners in Nair and Cavallaro (2009). Principal curvature based shape index and curvedness index maps were computed and candidate landmarks were located using regional segmentation. Furthermore, they fitted the PDM model for accurate registration of 3-D meshes. Performance of this method suffers when the mesh contains self-occlusions. Due to curvature based approaches, local segmentation and the simple 2-means classification technique, these approaches are computationally efficient. However, the performance is sensitive to noise and mesh occlusions.

In Koudelka et al. (2005), initially, the positively and negatively affected pixels are located using gradient operator on pixels of a range image. Subsequently, a radial symmetry plane is computed to obtain the radial symmetric features. A nose tip was declared within large

positive values and eye corners were localized around dark intensities of the radial symmetry map. Effective energy based hierarchical filtering method was proposed in Xu et al. (2006b) to locate a nose tip. The values of neighboring effective energy, mean and variance of neighboring pixels were fed to the support vector machine (SVM) classifier. This method is suitable for frontal, viewpoint invariant and rotated facial range images. It requires SVM based classification which makes it computationally heavy (Chew et al. 2009). To address this issue, Chew et al. (2009) proposed a direct approach based on the effective energy and avoided the use of SVM classifier. The effective energy for a pixel P with its neighborhood pixels P_n for an angle between normal vector for P and $(P_n - P)$ was computed using $E = \|P_n - P\| \cdot \cos(\theta)$. The surface pixels P with negative effective energy (E) were marked as potential candidates for nose tip location. This approach does not require initial training unlike (Xu et al. 2006b) and is computationally efficient.

In D'Hose et al. (2007), the authors pre-processed range images to cluster characteristic features (including the nose tip and eyes) using the Gabor wavelets. The segmented areas were refined using iterative closest point (ICP) algorithm for fine alignment and detection. The pre-processing stage is computationally efficient since only horizontal and vertical directions for Gabor filters were employed. Though this algorithm employs the computationally heavy ICP algorithm, it is applied only on small tentative feature regions in wavelet domain resulting in minimum computational requirement.

A 3-D shape signature based approach was proposed in Breitenstein et al. (2008). The 3-D location of a pixel along its orientation was examined and segmented to check if it has the maximum value compared to its neighbors. These local directional maximas were combined to form a sparse aggregated signature map. The cell located at the center of the signature map was denoted as nose tip. This approach is robust to partial occlusion, expression variations and vast range of poses such as $\pm 90°$ yaw, $\pm 30°$ roll and $\pm 45°$ pitch, though it is computationally heavy due to localization of a pixel in 3-D space in presence of the said pose variations.

In Romero and Pears (2009), the authors proposed a graph matching based method to detect eminent landmarks. The regional distribution of neighborhood pixels was conceptualized as a pattern of sample points around concentric circles. These patterns were encoded as cylindrically sampled radial basis function (CSR) histograms. The graph edges were effectively encrypted into these descriptors and utilized for detection of eye corners and nose tip. Like most graph-based techniques, this approach is computationally light.

A face profile based approach for detection of nose tip in range images or uniformly sampled point clouds was proposed in Peng et al. (2011). Initially, 61 left-most (min group, $[-90°, -3°]$) and right-most (max group, $[0°, 90°]$) 2-D profiles were computed from the input facial scan. The possible nose tip candidates were located using the proposed perimeter difference parameter. This method does not involve any computationally complex operations and is independent of any statistical model or training and hence demands fewer computations.

In Mohammadzade and Hatzinakos (2013), the authors constructed an Eigen-nose space using a training set of nose regions. The pixel in the input facial scan is marked as a nose tip if the mean square error between a candidate feature vector and an Eigen nose-space projection is less than a predefined threshold. Additionally, to reduce computational complexity, the authors have proposed a low-resolution wide-nose Eigen space, which speeds up the nose tip detection process. This reduces the otherwise heavy computational requirement of the original algorithm.

2.4 Registration techniques

Face registration means transforming a probe (query) face such that it is aligned with other faces in the gallery (database). Face registration is an essential step prior to feature extraction and the matching of two facial surfaces. Various algorithms to perform registration between two facial representations are briefed in the following subsections.

2.4.1 Iterative closest point (ICP) algorithm

Most of the 3-D face recognition systems introduced so far use the iterative closest point (ICP) algorithm (Besl and McKay 1992) for registration and matching of 3-D facial surfaces. The ICP algorithm transforms a 3-D point cloud shape '*M*' for alignment with another 3-D point cloud shape '*N*'. A series of incremental transformations are performed repeatedly. During each iteration step, the mean square error (MSE) between both point clouds is computed. The algorithm converges when the value of the minimum MSE is achieved based on the registered 3-D shape information. The ICP algorithm does not depend on extraction of any local features. Thus, ICP essentially is a semi-blind 3-D shape matching algorithm. ICP based registration is performed either between a probe and set of gallery faces or between a probe and average face model. A few methods that use ICP algorithm or it's variants for face registration are presented in Yue et al. (2010), Cai and Da (2012), Störmer and Rigoll (2008), Alyuz et al. (2010, 2008, 2012) Al-Osaimi et al. (2009), Ben Amor et al. (2008), Cook et al. (2004), Xueqiao et al. (2010a, b), Abate et al. (2006), Colombo et al. (2008), Kakadiaris et al. (2007), Amberg et al. (2008), Xiaoguang et al. (2006), Queirolo et al. (2010), Faltemier et al. (2008), Li and Da (2012), Ballihi et al. (2012) and Russ et al. (2005). A disadvantage of ICP-based registration is that it is prone to converge at a local minimum point and requires sufficient initialization for facial surface alignment (Yi and Lijun 2008). Thus, predefined localized landmarks enhance the accuracy of initial correspondence between the two point clouds (D'Hose et al. 2007; Nair and Cavallaro 2009). ICP based templates require large amount of storage space and it is comparatively slow in terms of registration time (Spreeuwers 2011). Nonetheless, it has been used by many researchers due to its robust registration performance.

2.4.2 Spin images

Spin images-based global registration method maps every point on 3-D facial surface to an oriented point p with surface normal n forming a 2-D basis (p, n) (Johnson and Hebert 1998, 1999). Point p acts as origin resulting in a dimensionality reduction $(x, y, z) \rightarrow (\alpha, \beta)$ using a 3-D to 2-D transformation (Johnson and Hebert 1999). The cylindrical coordinates (α, β) stand for the perpendicular span through a surface normal and elevation respectively (Johnson and Hebert 1999). A spin image is analogous to a distance histogram and represents neighborhood around a specific oriented point with associated directions. Spin image representation is used for face registration in Conde et al. (2006a), Kakadiaris et al. (2007) and Conde and Serrano (2005). The spin map $S_0(x)$ is calculated using (4). The spin image based registration is robust to partially occluded or noisy 3-D surfaces (Johnson and Hebert 1999). Compared to conventional ICP, this approach is computationally lighter due to dimensionality reduction in the form of the said transformation from 3-D to 2-D space.

$$S_0(x) \rightarrow (\alpha, \beta) = \left(\sqrt{\|x - p\|^2 - (n \cdot (x - p))^2}, n \cdot (x - p) \right) \tag{4}$$

2.4.3 Simulated annealing

Simulated annealing is an iterative stochastic approach which looks for better neighbor solution while matching two range images. Parameters for translation and rotation are treated as a transformation vector. Initially, centers of mass for two facial images are aligned. The *M-estimator* sample consensus (MSAC) measure (Torr and Zisserman 2000) and MSE are minimized for coarse alignment. The fine alignment is attained based on surface interpenetration measure (SIM) distance (Queirolo et al. 2010). This approach is employed for face registration in Kakadiaris et al. (2007), Queirolo et al. (2008) and Bellon et al. (2006). The SIM is preferred over MSE for quality evaluation of two registered range images (Queirolo et al. 2010). Due to the statistically heavy stochastic search algorithms, this technique is computationally inefficient. However, it has minimum tendency to get trapped in local maxima or minima. This ensures a correct and robust estimation of facial features.

2.4.4 Intrinsic coordinate system

All 3-D point cloud surfaces are transformed to a common reference coordinate system using a set of localized landmark structures. This system is referred to as the 'Intrinsic Coordinate System' (Spreeuwers 2011). The (x, y, z) coordinates are mapped to intrinsic coordinates (u, v, w) using vertical symmetry plane, nose tip, and slope of the nose bridge. This approach is effective in terms of time and computational complexity as all faces in the gallery and the probe image share a common coordinate system instead of extrinsic alignment with another 3-D facial shape or an average 3-D facial model. The efficiency of registration depends upon accurate localization of 3-D facial landmark structures.

2.5 Feature extraction and matching

After facial shape registration, most approaches extract significant features and match them using metrics such as the Euclidean distance (Heseltine et al. 2004; Mpiperis et al. 2008; Berretti et al. 2010; Gordon 1992; Gupta et al. 2010b), Hausdorff distance (Koudelka et al. 2005; Russ et al. 2005; Yeung-hak and Jae-chang 2004; Achermann and Bunke 2000; Gang et al. 2003), and angular Mahalanobis distance (Amberg et al. 2008) against the stored set of feature vectors for gallery faces. Major feature extraction algorithms, dimensionality reduction techniques and classifiers are discussed below.

2.5.1 Gabor wavelet filters

The Gabor kernel functions $\psi_j(\vec{x})$ are formulated as the Gaussian modulated complex exponentials. Convolution of a range image $I(\vec{x})$ with the family of Gabor functions results in directional Gabor wavelet subbands (Moorthy et al. 2010). The family of linear bandpass Gabor functions is obtained using variations in parameters of mother Gabor kernel $\psi_j(\vec{x})$. The human brain's primary visual cortex is responsible for extraction and computation of perceptual data (Tai Sing 1996). It is advantageous to employ the Gabor wavelet filters since they model the receptive fields of visual cortex (Tai Sing 1996; Moorthy et al. 2010; Daugman 1985). The Gabor wavelet functions typically have oriented localization properties which are essential for robust feature extraction. The resultant Gabor coefficients obtained after filtering a range image with oriented Gabor wavelets are known as Gabor features. The Gabor wavelet filters based feature extraction techniques on range images are used for precise 3-D feature extraction (Moorthy et al. 2010; Jahanbin et al. 2008; Xueqiao et al. 2010b). Gabor wavelets

yield precise features in transform domain. The computational complexity depends upon the level of Gabor wavelet transformations used in the specific application.

2.5.2 Local binary pattern (LBP)

Local binary pattern (LBP) approach is introduced by Ojala et al. (2002) for feature extraction from spatial 2-D texture images. A set of binary values is obtained upon comparison of each pixel value with intensities of its 3×3 neighboring pixels. These binary values are subsequently encoded to a decimal number which represents the LBP value for a particular pixel. A histogram of such LBP values acts as a discriminant feature vector and is widely used by face recognition researchers. A histogram of Multi Scale Local Binary Pattern (MS-LBP) based on LBP is introduced by Di et al. (2010) where, feature information is extracted using the MS-LBP histogram and employed for 3-D face recognition. Huang et al. (2006) realized that range images of two different persons may have same depth variations and the 2D-LBP is not suitable to distinguish them. Based on the inference that the exact values of differences may be different, they have proposed the extended 3D-LBP (Di et al. 2012) for feature detection in 3-D scans. The LBP representation is in integer numbers and hence LBP is a computationally effective technique compared to other feature detection techniques demanding heavy floating point computations.

2.5.3 Markov random field model

Ocegueda et al. (2011) proposed the Markov random field (MRF) model-based approach for extraction of discriminant features from 3-D faces. The MRF model is based on combination of probability and graph theory and is widely used for solving image processing and computer vision related complex problems such as face recognition, image denoising, image segmentation, image reconstruction and 3-D vision. The MRF model considers mutual dependencies amongst label field of range image pixels. The MRF model is based on the concept of near-neighbor dependence which is commonly observed in most of the range images. The discriminant features are selected based on Posterior marginal probabilities of the MRF model. This approach is computationally efficient due to involvement of graph theory and computation of Markov random probabilities based on a small number of prior event probabilities.

2.5.4 Scale invariant feature transform

Scale invariant feature transform (SIFT) detects landmarks using local extrema in the scale-space difference of Gaussian (DOG) function (Lowe 2004). SIFT features are invariant to orientation and scale. Di et al. (2010) applied the SIFT transform on shape index (SI) maps and multi-scale local binary pattern (MS-LBP) depth maps. It is also utilized by Berretti et al. (2013) for feature extraction on multi-scale extended local binary pattern (MS-eLBP) maps. SIFT is extended for feature extraction on 3-D mesh representation using MeshSIFT algorithm (Maes et al. 2010) and employed in Berretti et al. (2013), Smeets et al. (2011, 2013). Inherently, SIFT offers scale and rotation invariant features at moderate computational costs. However, application of SIFT on 3-D data demands considerable computational resources in terms of memory and computational speed.

2.5.5 Local shape patterns

The local shape pattern (LSP) computes differential structure factor (similar to LBP) as well as orientation factor (similar to SIFT) on a 3×3 neighborhood for a range image (Huang et al. 2011). The LBP is an established technique for efficient feature vector generation from facial images. The top-left first four pixel values are subtracted from their opposite pixels on the same diameter. The binary values are obtained in a way similar to the LBP. The 4-bit code is then converted to a decimal code to form an orientation pattern. The LSP combines the LBP with orientation information, which results in obtaining a more discriminant feature vector for 3-D face recognition. This approach is computationally heavy compared to LBP but it is lighter compared to transformation techniques like Gabor wavelets or Markov random fields.

2.5.6 Haar-like features

Viola and Jones (2001) developed this Haar-like feature detection algorithm. This technique deals with neighboring regions, adds up intensities in each region and computes difference between sums from each region. For a range image, Haar-like features calculate discontinuities in depth values. Mian (2011) computed Haar-like features on range images and its gradients. The Haar-like features are employed for description of characteristics of the signed shape difference maps (SSDM) (Yue et al. 2010). Due to the use of computationally simple Haar-like features, this algorithm is computationally very fast. This is also one of the most popular online available algorithms for real time face detection and tracking.

Major dimensionality reduction and feature classification techniques used by 3-D face recognition approaches are enlisted in Table 2.

Principal component analysis (PCA) (Turk and Pentland 1991) is a popular technique for dimensionality reduction of feature vectors. The procedure finds a set of orthonormal basis vectors which optimally represents the feature vector data. It transforms the original feature data vectors into a lower dimensional subspace. The new coordinate system forms the axes in such a way that they are a linear combination of input data vectors. The first component with maximum variance is denoted as the principal component. The second component is perpendicular to primary axis of principal component and indicates the direction of second highest statistical variance from the original data and so forth. PCA is computationally complex due to computation of multiple orthogonal Eigen-space vectors. Linear discriminant analysis (LDA) (Belhumeur et al. 1997) projects high dimensional feature vectors into a low dimensional subspace and computes an optimal linear combination of components that maximize intra-class distance and minimize inter-class distance simultaneously. The true reconstruction error rate of LDA is a parametric function of Gaussian distribution of classes. LDA is based on PCA and hence requires more computations than PCA but yields a smaller but more robust feature vector compared to PCA. Support vector machine (SVM) classifier (Cortes and Vapnik 1995) is a supervised machine learning method which is popular for addressing binary classification problems. SVM attempts to place a decision boundary which highly separates the feature vectors of genuine class from imposter class. The performance of SVM classifier is notably robust to sparse and noisy input feature vectors. During classification phase, they isolate a training data using a hyper-plane which is maximally separated from binary labeled classes. If the linear separation between the classes is not feasible, SVMs utilize kernel based methods to isolate the classes with non-linear mapping in the feature vector space. The SVM is extended to multi-class domain for face recognition, object classification and fingerprint recognition related problems. The SVM classifier for multiclass domain is

Table 2 Major dimensionality reduction and classification methods

Sr. no.	Method	Purpose	Used by
1	Principal component analysis (PCA) (Turk and Pentland 1991)	Dimensionality reduction	Heseltine et al. (2004), Russ et al. (2006), Chang et al. (2006), Störmer and Rigoll (2008) and Spreeuwers (2011)
2	Linear discriminant analysis (LDA) (Belhumeur et al. 1997)	Dimensionality reduction	Gupta et al. (2007), Kin-Chung et al. (2007), Alyuz et al. (2010), Sala Llonch et al. (2010), Heseltine et al. (2008), Lei et al. (2013) and Mohammadzade and Hatzinakos (2013)
3	Support vector machine (SVM) (Cortes and Vapnik 1995)	Classification	Conde and Serrano (2005), Hyoungchul and Kwanghoon (2006), Srivastava et al. (2006), Kin-Chung et al. (2007), Mousavi et al. (2008), Zhang et al. (2011), Ballihi et al. (2012) and Lei et al. (2013)
4	Nearest neighbor (NN) (Cover and Hart 2006)	Classification	Hesher et al. (2003), Samir et al. (2006), Jahanbin et al. (2007), Xueqiao et al. (2010b), Alyuz et al. (2010) and Yue et al. (2010)

computationally very heavy. The nearest neighbor (NN) classifiers are subdivided into 1-NN and k-NN classifiers (Cover and Hart 2006). The 1-NN classifiers allocate the class label of its nearest neighbor to a queried test feature point. The k-NN classifiers denote a class label to queried test feature point based on majority voting between the classes of k-nearest neighbors. The nearest neighbor classifiers are computationally very light. However, their performance depends upon the used distance measure.

2.6 Performance measures

The performance of a 3-D face verification/identification system is examined in terms of the percentage of false decisions it gives during verification/identification process. While comparing a probe face (query) with a claimed or possible identity (a face from gallery data), a system is observed for two types of errors. They are known as false reject errors and false accept errors. False reject error appears when a genuine face is compared with a possible or claimed identity and similarity score (Z) is lower than predefined threshold (λ), and the genuine person is not recognized by the system. False accept error arises when fake/imposter probe is compared with a possible or claimed gallery face and similarity score (Z) resulting from the match is higher than predefined threshold (λ) and the wrong person is authorized as established identity (Phillips et al. 2010). Figure 7 portrays the errors that occur during face verification process.
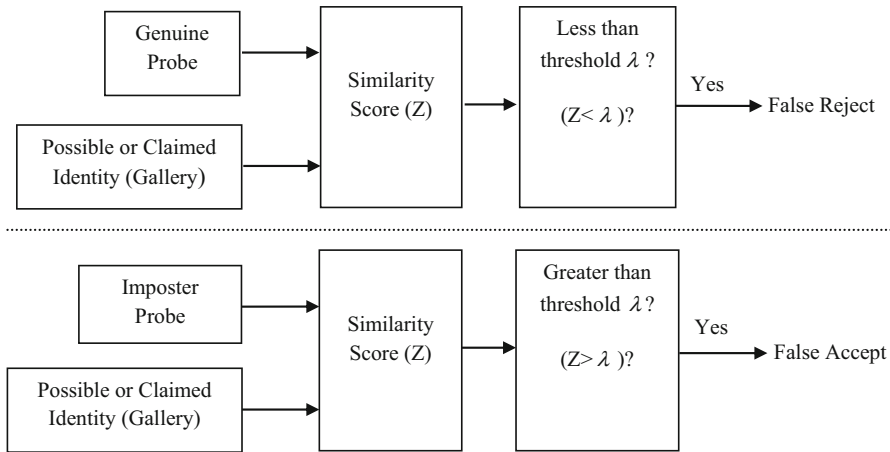
Fig. 7 False reject and false accept errors during face verification process

False accept rate (FAR) is the percentage of fake (imposter) probes which have similarity score greater than predefined threshold. False reject rate (FRR) is the percentage of genuine probes which have similarity scores less than predefined threshold. When FRR versus FAR is plotted on logarithmic scale or its variants, emanating graph is called receiver operating characteristic (ROC) curve. A point on ROC graph, where FRR and FAR are equal, provides the equal error rate (EER) point which is a single-valued error metric (Jain et al. 2008) for the rejection rate. For an accurate face recognition algorithm, the FRR, FAR and EER should be as minimum as possible. The performance of identification system is evaluated through the cumulative match characteristics (CMC) curve (Jain et al. 2008). The CMC curve is a plot between ranks and recognition rates which contains Rank-1 recognition rate (R_1-RR) that implies the first best matching rate when a probe is compared with all gallery faces (Smeets et al. 2012). These performance measures offer robustness to the performance benchmarking of the face recognition algorithms.

3 3-D face databases

For the development of a high quality robust face recognition algorithm, its testing on multiple large datasets which have inherent complexities similar to real life situations is a must. After the availability of 3-D face scanners and powerful digital computers capable of processing huge data at a time, many 3-D face databases have become available for research. Various publicly available 3-D face databases are enlisted in Table 3.

The 3DRMA database is captured at Royal Military Academy, Belgium using a structured light based 3-D face scanner. Initially, the camera and structured light projectors are aligned and calibrated in order to reduce the relative errors. They have acquired facial information of 120 different subjects. The 3-D facial surface is extracted through labeled stripe detection method (Beumier and Acheroy 2000). The FSU 3-D database is developed at Florida State University, FL and contains facial images of 37 distinct persons in a mesh format. The database is acquired using the Minolta Vivid 700 laser scanner (Hesher et al. 2003). The GavabDB database consists of mesh surfaces of 61 individuals. The age variation of the subjects is in between 18 and 40. They have acquired both intrinsic (pose, expression, occlusion) as well

Table 3 Prominent 3-D face databases

Sr. no.	Database and ref.	Number of acquisitions	Data acquisition method	2D intensity image size
1	3DRMA (Beumier and Acheroy 2000)	720	Structured light	768 × 576
2	FSU (Hesher et al. 2003)	222	Laser scanner	N.A.
3	GavabDB (Moreno and Sanchez 2004)	427	Laser scanner	N.A.
4	FRGC v 2.0 (Phillips et al. 2005)	4007	Structured-light	640 × 480
5	BU3D-FE (Lijun et al. 2006)	2500	Structured-light	512 × 512
6	CASIA (Xu et al. 2006a)	4059	Laser scanner	640 × 480
7	FRAV3D (Conde et al. 2006b)	1696	Laser scanner	N.A.
8	ND 2006 (Faltemier et al. 2007)	13,450	Laser scanner	N.A.
9	MSU (Xiaoguang and Jain 2006)	533	Laser scanner	N.A.
10	ZJU-3DFED (Wang et al. 2006)	360	Structured-light	N.A.
11	IV^2 multimodal (Petrovska-Delacretaz et al. 2008)	2880	Laser scanner	780 × 576
12	Bosphorus (Savran et al. 2008)	4666	Structured-light	1600 × 1200
13	University of York (Heseltine et al. 2008)	5250	Stereo acquisition	N.A.
14	BJUT-3D (Bao-Cai et al. 2009)	500	Laser scanner	489 × 478
15	Texas 3-D (Gupta et al. 2010a)	1149	Stereo acquisition	751 × 501
16	UMB-DB (Colombo et al. 2011)	1473	Laser scanner	640 × 480
17	3D TEC (Colombo et al. 2011)	428	Laser Scanner	480 × 640
18	EURECOM Kinect (Huynh et al. 2013)	∼900	Structured-light	256 × 256

as extrinsic variations (varying background, scale, lighting effects) (Moreno and Sanchez 2004). The Face Recognition Grand Challenge V 2.0 (FRGC V 2.0) database is a part of a protocol specially designed to access the increase in recognition performance by different state-of-the-art algorithms (Phillips et al. 2005). The 3-D data is captured under controlled illumination conditions. The 3-D shape and the corresponding texture image of a specific person were acquired with Minolta Vivid 900/910 scanner. The training partition of the dataset consists of 943 scans whereas the validation partition contains 4007 3-D scans for 466 individuals. Experiment 3 of FRGC challenge focuses on 3-D face recognition where, both training and query faces are 3-D facial surfaces. The BU3D-FE (Binghamton University 3D Facial Expression) database specifically incorporates the expression variation challenges for 3-D face recognition (Lijun et al. 2006). The various expressive faces of 100 subjects were

acquired with the 3DMD digitizer which stores a range image as well as mesh structure of the person in front of a projection beam. CASIA 3-D face database is collected at the Chinese Academy of Sciences Institute of Automation (Xu et al. 2006a). It contains facial images of 123 subjects, captured with laser based Minolta Vivid 910 scanner. The advantage of this database is that it not only contains expression and pose variations but also combinations of poses under different expressions. It also involves scans with illumination variations. Face Recognition and Artificial Vision group acquired the FRAV3D database which consists of facial scans of 105 subjects (Conde et al. 2006b). Minolta Vivid-700 laser scanner was used for capturing 3-D mesh and texture information. ND 2006 database is a superset of FRGC v 2.0 database (Phillips et al. 2005), acquired at the University of Notre Dame (Faltemier et al. 2007). The facial scans of 888 individuals were captured using Minolta Vivid 910 scanner (Konica 2013). The output of this laser based scanner consists of color texture along with 3-D facial range images. Michigan state university (MSU) has developed a database which contains simultaneous expression and large pose variations. They have acquired the range scans of 90 subjects. Since expression and pose variations are imposed simultaneously, face recognition using this database is challenging and complex (Xiaoguang and Jain 2006). The ZJU-3DFED database consists of 3-D facial scans of 40 individuals captured using structured light based Inspeck 3-D Mega capturor sensor at the Zhejiang University (Wang et al. 2006). Lack of pose variations is the main limitation of this database. The 'Identification par l'Iris et le Visage via la Video' (IV^2) multimodal database (Colineau et al. 2008) involves 3-D faces as well as 2-D videos, audios and iris data of 365 subjects. 3-D facial surfaces are captured using Konica Minolta Vivid 7000 sensor. Bosphorus 3-D face database is acquired using Inspeck Mega Capturor-II scanner at the Bogazici University (Savran et al. 2008). The shape and texture data were captured within less than a second per subject. The raw point cloud data of 105 subjects was processed for noise removal using median and Gaussian filters. It also offers manual marking of 24 facial landmarks like chin middle, eye corners for each facial scan. Researchers from the University of York (Heseltine et al. 2008) acquired 3-D facial data with help of two cameras using stereo based modeling technique. In addition, another camera mapped the captured 2-D texture information onto the 3-D mesh model. Beijing University of Technology gathered BJUT-3D database which contains 3-D mesh and color texture data of 500 subjects. The database is acquired with the Cyberware 3030 RGB/PS 3-D scanner, which stores texture along with shape data for each facial scan. In order to avoid occlusion due to hair, all subjects were asked to wear a swimming cap. Smoothing has been performed for the sake of removing noise from the captured data. Hole filling filters were operated on the facial scan data in order to deal with missing parts. Finally, after preprocessing, the facial region from the redundant parts has been separated and stored as a part of BJUT-3D database (Bao-Cai et al. 2009). The range images of 118 subjects along with corresponding 2-D color textures were acquired at the University of Texas at Austin using stereo based scanning system (Gupta et al. 2010a). Locations of 25 anthropometric fiducial landmarks are included along with the database (Gupta et al. 2010b) which is a unique feature of this database. UMB-DB 3-D face database consists of frontal facial scans of 143 subjects acquired using Minolta Vivid 900 laser scanner at university of Milano Bicocca (Colombo et al. 2011). This database is useful for the evaluation of algorithm performance under unconstrained scenarios since they have captured naturally occurring occlusions. The 3D-TEC dataset is a subset of the twins' days dataset captured with Minolta Vivid 910 scanner during the twin's days festival in Twinsburg, Ohio. It contains 3-D facial scans of 107 twins acquired with neutral and smiling expressions (Vijayan et al. 2011). The RGB color and depth data of 52 subjects is acquired using Microsoft's Kinect camera (Shotton et al. 2013) at the EURECOM institute. The data was captured during two sessions separated

Table 4 Expression specific 3-D face databases

Sr. no.	Name	Expressions
1	FSU	Neutral, smile, scared, angry, squint, frown
2	GavabDB	Neutral, smile, accentuated laugh, random gesture
3	FRGC v 2.0	Neutral, surprise, happy, puffy cheeks, anger, frown
4	BU3D-FE	Neutral, angry, fear, sadness, disgust, happiness, surprise
5	CASIA	Neutral, smile, eyes closed, anger, laugh, surprise
6	FRAV3D	Neutral, smile, open mouth, and gesture
7	ND2006	Neutral, surprise, sadness, disgust, happiness, undetermined
8	ZJU-3DFED	Neutral, smile, surprise, sad
9	IV^2 multimodal	Neutral, surprise, happy, eyes closed, disgust
10	Bosphorus	Neutral, happy, anger, disgust, fear, sadness, surprise
11	University of York	Neutral, eyes closed, eyebrows raised, happy, anger
12	Texas-3D	Neutral, smile/talk with open/closed eyes and/or open/closed mouth
13	UMB-3D	Neutral, smile, angry, bored
14	3D-TEC	Neutral, smile
15	EURECOM Kinect	Neutral, smile, open mouth

with half a month's difference. The manual annotation information is also available for six fiducial landmarks. The 3D expression variations are captured by various databases enlisted in Table 4. These databases can also be used for 3D expression recognition.

From Table 4, it is evident that the FSU, FRGC v 2.0, BU3D-FE, ND 2006 and Bosphorus are the databases which contain large variations amongst facial expressions and can be utilized for benchmarking of identification/verification performance of a novel algorithm.

As indicated in Table 5, the CASIA and Bosphorus database captured stepwise pose variations. These databases are most challenging for the automatic landmark localization and normalization algorithms. The FRAV3D database contains severe Z-axis pose variations which also severely affect the performance of real time 3-D face recognition systems.

Table 6 enlists various occlusions available in publicly accessible 3-D face databases. The most common intrinsic occlusions that affect the performance of face recognition systems are due to hair, randomly posed hands that cover facial region and eyeglasses. The occlusion invariant face recognition system should address the extrinsic occlusions like missing facial data due to hat, scarf, paper and eyeglasses. It should be noted that the Bosphorus dataset appears with all the three constraints. Thus it should be imperative for the validation and benchmarking of novel 3-D face recognition algorithms.

4 Expression invariant 3-D face recognition

3-D face recognition is comparatively robust to illumination and pose variations than 2-D image-based face recognition. However, owing to its non-rigid nature, the captured 3-D faces deform with expression variations which is a prominent challenge for recognition from the

Table 5 Pose specific 3-D face databases

Sr. no.	Name	Pose variations
1	3D RMA	Frontal, up/down, limited left/right
2	GavabDB	Frontal, left profile, right profile, looking up, looking down
3	CASIA	Frontal, tilt left and right $20°-30°$, up and down $20°-30°$, left and right $20°-30°$, left and right $50°-60°$, left and right $80°-90°$
4	FRAV3D	Frontal looking up and down in X-axis direction, 25° Y-axis right turn, 5° Y-axis left turn, small and severe Z-axis right turn
5	IV^2 multimodal	Frontal, left and right profile
6	Bosphorus	Frontal, right-downwards, right-upwards, upwards, downwards, slight upwards and slight downwards, or as represented by exact numerical angles $+10°, +20°, +30°, +45°, +90°, -45°, -90°$
7	University of York	Frontal, up, down
8	EURECOM Kinect	Frontal, left and right profile

Table 6 Occlusion specific 3-D face databases

Sr. No.	Name	Occlusion due to
1	Bosphorus	Hair, mouth, eye, eyeglasses
2	UMB-DB	Scarf, hat, hands in random positions, eyeglasses, hair, miscellaneous
3	EURECOM Kinect	Hand, paper

3-D shape (Samir et al. 2006; Kin-Chung et al. 2007). An obvious approach to take on this challenge is to identify expression invariant 3-D human face features and use them for matching. Expression invariant face recognition approaches are broadly classified as holistic and region based methods. Figure 8 presents the classification of different approaches for expression invariant 3-D face recognition.

4.1 Local processing based expression invariant approaches

In recent years, there has been a growing interest in using local processing based approaches for addressing the expression variation problem in 3-D face recognition. The facial surface is divided into different local sub-regions. Instead of matching entire facial surfaces, face recognition algorithms are applied on smaller facial sub-regions. Generally, those sub-regions which have lesser effect of expression variations are selected for matching process. Figure 9 indicates the general steps used in local processing based approaches which address expression variation problem in 3-D face recognition.

After acquisition of a 3-D probe face, it is segmented to avoid outliers and registered with the gallery faces. Usually, local patches invariant to expression variations such as the area surrounding nose, forehead and eyes are extracted and matched with the corresponding local
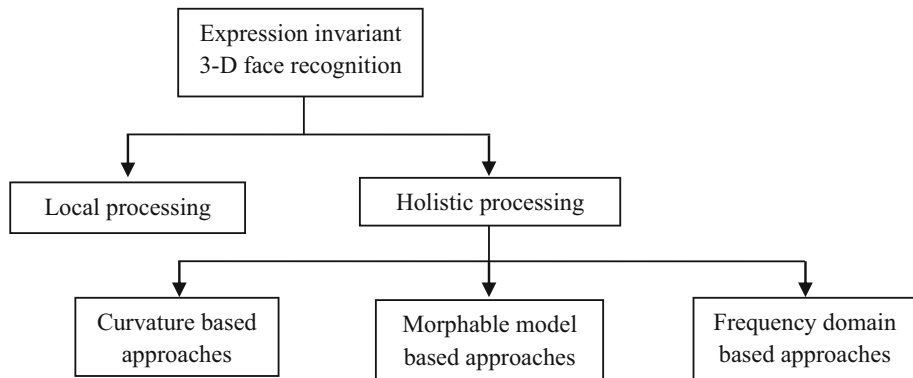
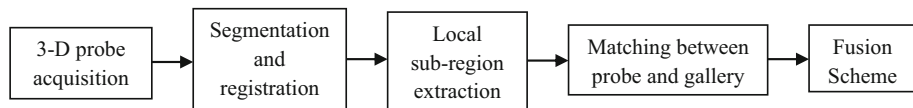**Fig. 8** Broad classification of expression invariant approaches

Fig. 9 General steps involved in local processing based approaches

patches from gallery faces to find the closest match. Individual matching results are typically fused to achieve a better rank-1 recognition rate. For performance improvement, various fusion schemes are proposed in the literature. Major fusion schemes are briefly discussed below. More details on fusion schemes can be found in Spreeuwers (2011).

4.1.1 Fusion schemes

Sensor level fusion Sensor level fusion combines the raw sensor data during initial stages of 3-D face recognition system, before feature extraction and matching. The raw sensor data includes the biometric information of the same person acquired with different sensors. Sensor level fusion is widely referred in multi-biometrics literature as the image or pixel level fusion. This fusion scheme belongs to the 'fusion before matching' class (Ross et al. 2006; Li and Jain 2011). For example, multiple calibrated 2-D cameras enable pose invariant face recognition by forming a 3-D model that can be rotated and translated. The calibration and compatibility of the 2-D cameras is a must for reconstruction of a reliable 3-D model. As this scheme achieves fusion at sensor level, it requires minimum computational overhead.

Feature level fusion Feature level fusion involves combination of features extracted from different facial representations of a single subject. Using the results of multiple feature extractors, the ultimate feature vector is derived. If the feature extraction algorithms have compatible outputs, template improvement or template update schemes are employed to obtain the resultant feature vector (Ross et al. 2006; Li and Jain 2011). Template improvement attains the removal of duplicate data in feature sets. On the other hand, template update modifies the resultant feature vector based on certain statistical operations such as mean, variance or standard deviation. This fusion scheme is among the 'fusion before matching' class. If the feature vectors participating in fusion have incompatible ranges and distributions, feature normalization is employed to modify the scale and location of feature vectors. Transforma-

tion techniques such as min-max normalization and median normalization are popularly used in the reported literature for feature normalization (Ross et al. 2006). This requires considerable computational overhead as registration and feature detection is to be done individually for all images.

Rank level fusion During identification experiment, ranks are assigned to gallery images according to decreasing sequence of confidence. The rank outputs of several subsystems are fused to construct a consensus rank for each individual probe image. The consensus of ranks provides eminent information to the decision-making module. The highest-rank, Borda-count and logistic regression are popular statistical methods to attain the rank level fusion. Rank level fusion is independent of normalization techniques (Ross et al. 2006; Li and Jain 2011) and belongs to the 'fusion after matching' class. This is computationally complex as rank of each matching gallery image is to be computed.

Decision level fusion The decisions of individual classifiers are combined in decision level fusion. Exemplary techniques are behavior knowledge space, AND and OR rules, Bayesian decision fusion, majority voting, Dempster–Shafer evidence-based method and weighted majority voting schemes (Ross et al. 2006; Li and Jain 2011). The behavior knowledge space-based technique computes a storage map during training that links the decision of individual classifiers to a single decision. The AND rule declares a match only when all classifiers agree with the specified template, whereas, OR rule affirms a match when at least one classifier agrees with the particular template. The Bayesian decision fusion selects a label with highest probabilistic discriminant function value. Majority voting based technique performs decision level fusion based on majority of classifiers agreeing with the specific template. If the face recognition system contains say, M classifiers, then at least $g(g = (M/2) + 1, (M \bmod 2) = 0$ and $g = (M + 1)/2, (M \bmod 2) = 1)$ classifiers must agree with the specified template to declare it as a match. Dempster–Shafer evidence based method adopts belief theory to obtain more flexible and robust results than the probability based Bayesian decision level fusion. This method is preferred when the complete knowledge base related to the problem is not available. Weighted majority voting technique is utilized when classifier efficiency in terms of accuracy is not similar with other classifiers. The classifier 'C' may give more accurate matching rates and hence assigned more weight in the majority voting scheme. This scheme belongs to 'fusion after matching' class.

Score level fusion It is the combination of the match scores of individual classifiers based on weighing schemes. Sum, product, max, min and median rules are utilized for performing score level fusion. The combination of matching scores from various classifiers provides a new feature vector. The efficiency of score level fusion depends upon similarity metrics, probability distributions and numerical ranges of individual classifiers. It is the most popular fusion technique for face recognition due to readily available match scores and reduced processing overheads. Score level fusion scheme belongs to the 'fusion after matching' class (Ross et al. 2006; Li and Jain 2011). Score level fusion is computationally heaviest as matching scores given by multiple classifiers are to be computed for all the gallery images and fused for the final decision.

4.1.2 Approaches

The first systematic study of 3-D face recognition based on local processing was carried out by Chang et al. (2006). After segmentation and pose normalization, the circular regions in

the vicinity of the nose tip are selected and registered with a generic 3-D face model. Three regions namely: (i) entire nose region, (ii) interior nose region and (iii) circular region around the nose tip are extracted as small probe surfaces and matched with each region in gallery using the ICP algorithm. Score level fusion is performed based on sum, product and minimum rules. The neutral probe to neutral gallery identification rate is 97.1 % and expressive probe to neutral gallery identification rate is 87.1 %. Although they have achieved 97.1 % identification rate, the probe images in the real life scenario may comprise of expressive faces. It is stated that the performance may increase with extra local patches in the fusion process at the cost of increased computations.

Kin-Chung et al. (2007) proposed linear discriminant analysis (LDA) and linear support vector machine (LSVM) fusion based 3-D face recognition system by capturing local characteristics from multiple regions as summation invariants. From a single frontal facial image, ten sub-regions and subsequent feature vectors are extracted. Further, summation invariants are computed with the moving frames (Fels and Olver 1998) technique. LDA and LSVM based linearly optimum fusion rules provide improved performance. The performance of reported approach decreases with increasing severity of expressions.

Faltemier et al. (2008) used 28 best-performing sub-regions from a facial surface for 3-D face recognition. For a probe image, total 38 sub-regions are extracted; a few of them overlap with each other. Each sub-region from a probe image is matched with a gallery image region by employing the ICP algorithm. Highest rank-one recognition rate achieved by individual region matching is 90.2 % which promoted the use of the fusion strategy. The modified Borda-count fusion method attains overall 97.2 % rank-one recognition rate. This algorithm performed well despite incomplete facial information in some of the range images from the FRGC v 2.0 database.

Boehnen et al. (2009) proposed an alternative method to the ICP based matching. Instead of comparing whole 3-D point clouds with each other which is time consuming, each face is represented as a 3-D signature template. The comparison time for such templates is much less than ICP based approaches. Eight reference regions are described for signature matching. A novel match sum threshold fusion scheme is proposed which performs better than the modified Borda-count technique. The authors have mentioned that the performance can be improved upon addition of regions with variability of shapes and density. However, identification and separation of such patches, their processing and fusion for final decision requires considerable computational power.

Queirolo et al. (2010) developed a new technique for range image registration based on Simulated Annealing. Subsequently, the elliptical and circular area in the vicinity of the nose, upper head and the entire facial region are extracted from the facial surface. To match the range images, surface interpenetration measure (SIM) is utilized as a similarity metric with sum rule for fusion. If two range images present good interpenetration, it means that, they are registered correctly and have high SIM values. This approach is quite faster than the techniques which utilize the ICP based registration with MSE as a similarity metric.

Spreeuwers (2011) proposed an intrinsic coordinate system defined by the nose tip, perpendicular symmetry plane passing through the nose and nose bridge slope for face registration. Thirty overlapping sub-regions are defined for 3-D face recognition. After matching probe and gallery images, the majority voting scheme is opted for performing decision level fusion. This approach performs fastest computation of similarity matrix on FRGC v 2.0 database using 'all versus all' protocol for the verification experiments. Each local classifier has different parameters in order to further improve the performance of the proposed system. In addition, the fusion is performed using techniques based on majority voting approach. More advanced fusion and matching techniques may yield enhanced performance.

Li and Da (2012) used local processing based approach for matching 3-D faces. Initially, the facial surface is split into six regions which are forehead, left mouth, right mouth, nose, left cheek and right cheek. Deformation mapping based adaptive region selection scheme is proposed for automatic region selection. The partial central profile is extracted and compared against partial central profiles of the other gallery faces to form an initial rejection classifier. Sub-regions are matched using the ICP algorithm. Subsequently, weighted sum fusion rule method is employed for overall recognition performance improvement. The approach can fail when the nose tip localization is incorrect which causes improper normalization leading to inaccurate matching.

Lei et al. (2013) divided 3-D facial surface into nose, eyes-forehead and mouth regions. By assuming the nose region as rigid, eyes-forehead region as semi-rigid and mouth as non-rigid regions, only rigid along with semi-rigid regions are considered to eradicate the effect of facial expressions. Four fundamental geometric features namely: A: Angle between the two predefined random points and the nose tip, C: Radius of a circle between the nose tip and two predefined random points, N: Angle of a line joining predefined points and Z-axis, D: Length of two predefined points on a facial surface are extracted and subsequently transformed into a region based histogram descriptor. A machine learning based SVM technique is employed for classification. The results indicate that feature level fusion performs better than score level fusion for this experimental scenario. Authors also recommend the fusion of low level features of the rigid and semi-rigid facial parts. The local region segmentation is performed with a fixed size binary mask which may not be the best fit for all facial scans. This may affect the final matching scores and recognition rates.

Tang et al. (2013) proposed a local binary pattern (LBP) based 3-D face division scheme. The facial surface is divided into 29 blocks which is denoted as a sparse division and another 59 blocks are denoted by a dense division. They have utilized the nearest neighbor based classifier for 3-D face recognition.

Li et al. (2014) computed the normal vectors to a range image using local plane fitting technique. Subsequently, the normal component images were extracted for each x, y and z axis. Each normal component image was subdivided into local patches of fixed size. The local patches were subjected to multi-scale uniform LBP to obtain multi-scale multi component local normal pattern (MSMC-LNP) histograms. Specific weights have been assigned to different patches depending upon their position in a face to obtain an expression robust feature vector. The method was tested using three scales and respective normal components. Weighted sparse representation (W-SRC) classifier has been employed for final classification. The range image registration is performed with the ICP algorithm which is computationally heavy. The overall algorithm is computationally heavy.

Regional Bounding SpheRe descriptors (RBSR) were presented for efficient feature extraction on 3-D facial surfaces (Ming 2015). Initially, the facial surface is divided into local patches by using shape band algorithm. The local patches involve the forehead, mouth, left cheek, right cheek, nose, left eye and right eye regions. The RBSR descriptors extracted from the local patches are transformed with prior weightage into the spherical domain. The regional and global regression mapping (RGRM) captured the manifold configuration of pre-weighted RBSR feature vectors. This method is heavily dependent upon robust, consistent and accurate detection of the local patches.

Different region based approaches reported so far are summarized in Table 7.

As indicated in Table 7, Spreeuwers (2011) has achieved the highest rank-one recognition rate on FRGC v 2.0 database. This approach requires significantly less time than the other state-of-the-art methods for registration and matching.

Table 7 Local processing based 3-D face recognition approaches

Sr. no.	Author and year	Dataset	Local sub-regions	Identification rate (%)	Verification rate (%)	Classifier/feature detection	Fusion scheme
1	Chang et al. (2006)	FRGC v 2.0 Superset	Interior nose, center face, entire nose	87.10	N.A.	PCA	Score level
2	Kin-Chung et al. (2007)	FRGC v 2.0	10 Sub-regions from nose, eyes, cheeks and eyebrows	N.A.	90.45	LDA, LSVM	Score level
3	Faltemier et al. (2008)	FRGC v 2.0	38 Sub-regions from nose, eyes, forehead and chin	97.20	93.2	N.A.	Rank level
4	Boehnen et al. (2009)	FRGC v 2.0	Eight sub-regions around eyes and nose	95.50	N.A.	Nearest neighbor	Rank level
5	Queirolo et al. (2010)	FRGC v 2.0	Sub-regions from nose and forehead	98.40	96.50	N.A.	Score level
6	Spreeuwers (2011)	FRGC v 2.0	30 Overlapping sub-regions from all over the face	99.00	94.60	PCA-LDA	Decision level
7	Li and Da (2012)	FRGC v 2.0	Six sub-regions: forehead, left mouth, right mouth, nose, left cheek and right cheek	97.80	96.00	PCA	Score level
8	Lei et al. (2013)	FRGC v 2.0	Sub-regions from nose, eyes and forehead area	95.60	97.60	Modified LDA, polynomial-linear-RBF SVM	Feature level
		BU-3DFE		97.70	98.20		Feature level
9	Tang et al. (2013)	FRGC v 2.0	Special feature based sparse division (29 blocks) and dense division (59 blocks)	94.89	N.A.	Nearest neighbor	Score level
		BJUT-3D		95.10	N.A.		Score level
10	Li et al. (2014)	FRGC v 2.0	Multi-scale multi resolution patches (MSMC-LNP)	96.30	N.A.	W-SRC	Score level
		BU-3DFE		92.21	N.A.		
		Bosphorus		95.40	N.A.		

Table 7 continued

Sr. no.	Author and year	Dataset	Local sub-regions	Identification rate (%)	Verification rate (%)	Classifier/feature detection	Fusion scheme
11	Ming (2015)	FRGC v 2.0	Seven sub-regions: forehead, mouth, left cheek, right cheek, nose, left eye and right eye regions	N.A.	95.03	N.A.	Feature level
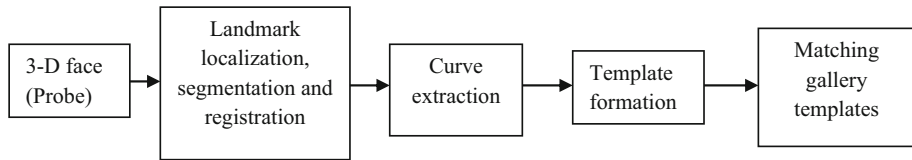
Fig. 10 General steps in curvature based expression invariant 3-D face recognition

4.2 Holistic approaches

Unlike local processing based techniques, the whole facial surface is considered simultaneously for face recognition in holistic approaches. The various methods using holistic approaches are broadly classified as (i) curvature based, (ii) morphable model based and (iii) frequency domain approaches.

4.2.1 Curvature based approaches

A few holistic approaches extract curvature information associated with the facial shape. General steps involved in curvature based approach are presented in Fig. 10.

The curvature information is represented as level or radial curves. Level curves are extracted at distance 'X' along the depth direction from some reference point. In most of the approaches, nose tip is the reference point. Figure 11a depicts level curves marked on a range image. Profile information is recorded by rotating a vertical plane which makes angle 'Y' with the reference plane to derive radial curves. The reference plane passes through a landmark such as nose-tip. Figure 11b illustrates radial curves marked on 3-D facial surface. The extracted curves represent the facial surface profile and information in a compact form and are further converted into a feature template for matching templates of gallery faces. Since the facial surface is represented by a very compact representation, these approaches are much appreciated due to less computational requirements during feature computation and matching.

Samir et al. (2006) represented the facial surface as a height map and extracted the level curves. Geodesic lengths of facial curves are used as a similarity metric for matching amongst different curves using the nearest neighbor classifier. This approach is invariant to range image acquisition noise.

Haar and Veltkamp (2009) used profile and contour (level) curves for face matching. Profile curves are collected starting from the nose tip. To compare two profile curves, sample values were examined along the curves. The contour curves are obtained by combining samples in all profile curves with the same depth value. The number of profiles and contours-based distance between the corresponding samples for two faces are used as dissimilarity measure. This technique requires exact nose-tip location for initial pose normalization. Good recognition rate is achieved subject to the accurate localization of the nose-tip.

Ballihi et al. (2012) used the level sets/circular curves and streamlines/radial curves as Euclidean distance functions of the 3-D face. The feature extraction is performed using the machine learning based Adaboost method (Freund et al. 1999). Pair-wise comparison of facial curves is accomplished using differential geometry tools. Each face is characterized using a highly compact 3-D face signature which is effective in terms of storage and retrieval cost. This yields good recognition rates at reasonably less computational cost.
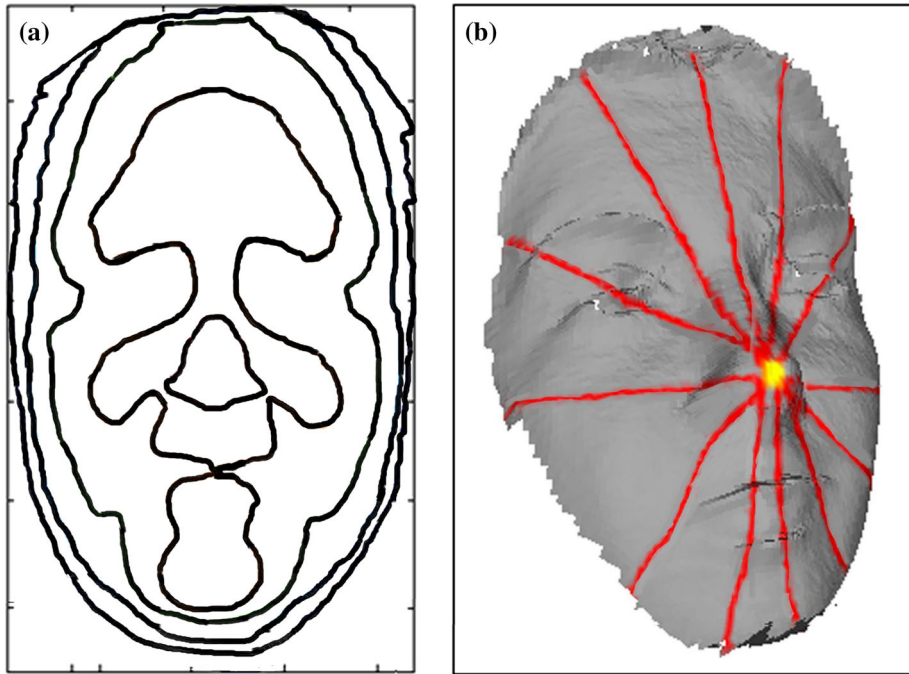
Fig. 11 **a** Level curves marked on depth image and **b** radial curves marked on a 3-D facial surface

Drira et al. (2013) proposed a 3-D face recognition method based on curvature analysis of a 3-D facial surface. The Riemannian framework is used for analyzing 3-D facial surfaces. The radial curves are extracted with nose tip as the center. They have compared radial curves from a probe with gallery faces with the help of elastic shape analysis framework for curves (Srivastava et al. 2011) of 3-D facial surfaces. The authors have manually located the nose tip for non-frontal facial surfaces in the Bosphorus and GavabDB datasets. This technique also achieves good recognition rates subject to accurate localization of nose-tip by a human operator.

Lei et al. (2014) extracted a set of radial curves from upper part of a face due to its semi-rigid nature. A novel feature extraction method called angular radial signatures (ARS) is proposed which represent a 3-D face into a set of mapped 1-D feature vectors. The 1-D feature vectors are found to be very close to each other in the feature vector space. Thus, Kernal principal component analysis (KPCA) is employed to address the linear inseparability issue. The KPCA transforms 1-D ARS features to a high-dimensional feature space. Subsequently, a SVM classifier is trained with a set of KPCA feature vectors. For a probe face, its corresponding KPCA feature vector is compared against the learned SVM model to find the closest match. This technique is computationally complex due to the use of SVM and required training.

Table 8 presents the major curvature based approaches published and discussed so far along with their benchmarking.

As indicated in Table 8, the Riemannian framework based approach (Drira et al. 2013) is found to be a robust curve based 3-D face recognition technique as it achieves consistent performance for identification as well as verification experiments on FRGC v 2.0 dataset, but at a considerable computational cost.

Table 8 Curvature based 3-D face recognition approaches

Sr. no.	Author	Year	Dataset	Persons in dataset	Images in dataset	Identification rate (%)	Verification rate (%)	Classifier/feature detection
1	Samir et al. (2006)	2006	FSU	50	300	92.00	N.A.	Nearest neighbor
			UND	277	953	90.40	N.A.	
2	Haar and Veltkamp (2009)	2009	SHREC'08	61	427	92.50	N.A.	N.A.
			UND	277	953	97.60	N.A.	
3	Ballihi et al. (2012)	2012	FRGC v 2.0	466	4007	98.02	N.A.	Adaboost, neural network, SVM
4	Drira et al. (2013)	2013	FRGC v 2.0	466	4007	97.70	97.14	PCA
			GavabDB	61	427	96.99	N.A.	
			Bosphorus	105	4652	87.06	N.A.	
5	Lei et al. (2014)	2014	FRGC v 2.0	466	4007	N.A.	96.70	KPCA and SVM
			SHREC'08	61	427	N.A.	88.50	
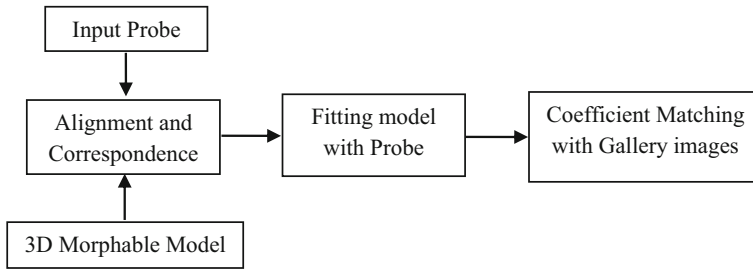
Fig. 12 Morphable model based approaches

4.2.2 Morphable model-based approaches

The morphable models vary their shape in accordance with an unknown facial shape. They are also known as deformable models. A general methodology applicable to the morphable model-based approaches is depicted in Fig. 12.

In order to compare an input probe face with gallery faces, initially a probe is registered with the morphable model to ensure the unified coordinate system. After this alignment, point-to-point feature correspondence is established between a probe and the morphable model. The 3-D morphable model is fitted to the probe. The deformation parameters/coefficients obtained after the completion of fitting procedure are encoded into a feature vector or represented in some suitable form to achieve compact template representation. A morphable model is built by training samples from a train dataset. Once the morphable model is constructed, it can be used for matching probes having any alignment, pose, scale or even occlusion. Major approaches that use the morphable models are enlisted below in Table 9 with their performance on various databases.

Kakadiaris et al. (2007) proposed a fully automatic expression invariant 3-D face recognition system based on the morphable models. An annotated face model (AFM) is built using the average 3-D facial mesh model. After initial alignment, the AFM fitting gives geometric characteristics of probe face based on elastically deformable model scheme by Metaxas and Kakadiaris (2002). A difference between probe scan and the AFM is transformed into a 2-D-geometry image which constitutes encoded information on a 2-D grid sampled from the deformed AFM. Consequently, the geometry image is transformed to a normal map which distributes the information evenly amongst three components. The Haar wavelet transform is applied on a normal map and geometry images. The pyramid transform is employed for the geometry images. The L_1 distance and structural similarity index measure (SSIM) are used as performance metrics. The transformation to a spin image is computationally heavy but it facilitates fast matching in terms of system-time (Spreeuwers 2011).

Xiaoguang and Jain (2008) proposed an algorithm to match range images having expression variations using 3-D face models with neutral expression. A collection of 94 landmarks is utilized to study the 3-D surface deformation due to specific expressions. The mapping is then established between the collection of landmarks having neutral expression and a 3-D neutral model. A comparable deformation is incorporated onto a 3-D neutral model by applying thin plate spline (TPS)-based mapping. Every probe scan associated with a deformable model is matched against gallery faces based on the cost function minimization algorithm. The formation of the user specific deformable model for each person in the gallery is computationally complex.

Table 9 Morphable model based approaches

Sr. no.	Author	Year	Dataset	Persons in dataset	Images in dataset	Identification rate (%)	Classifier/feature detection
1	Kakadiaris et al. (2007)	2007	FRGC v 2.0	466	4007	97.0	N.A.
2	Xiaoguang and Jain (2008)	2008	FRGC v 2.0	100	877	92.0	PCA
3	Mpiperis et al. (2008)	2008	BU-3DFE	50	1250	86.0	Maximum likelihood
4	Amberg et al. (2008)	2008	UND	953	953	100	N.A.
			GavabDB	61	427	99.70	
5	Al-Osaimi et al. (2009)	2009	FRGC v 2.0	466	4007	96.52	PCA
6	Haar and Veltkamp (2010)	2010	UND	277	953	99.0	PCA
			GavabDB	61	427	98.0	
			BU-3DFE	100	2500	100.0	
			FRGC v 2.0	466	4007	97.0	

Mpiperis et al. (2008) established the combination of elastically deformable model framework and Bilinear models for jointly addressing the problems of identity invariant 3-D face expression recognition and expression invariant 3-D face recognition. Based on the elastically deformable models, bilinear model coefficients are computed. During a test phase, elastically deformable model is fitted to both probe and gallery face sets to associate point to point correspondences. Euclidean distance is employed as a metric for comparison of probe and gallery surfaces. In order to build a bilinear model, a large annotated training dataset is needed. The recognition performance using bilinear models hampers with improper localization of facial landmarks.

Amberg et al. (2008) introduced a method for expression invariant face recognition by fitting identity and expression separated 3-D morphable models to a 3-D facial probe shape. A statistical PCA based 3-D morphable model is trained with a combination of 270 individual neutral and 135 expressive face components. A variant of the non-rigid ICP is employed as a fitting algorithm. A 3-D morphable model is fit to the point correspondences using systematic Jacobian and Gauss–Newton Hessian approximations. Angular Mahalanobis distance is used as dissimilarity metric. This approach suffers from a large processing time (40 s per probe).

Al-Osaimi et al. (2009) employed the PCA to learn and model expression deformations. The PCA eigenspace technique is employed because it is computationally effective than other closed form solutions. The generic PCA deformation model is built using non-neutral faces of distinct persons. The expression deformation templates are used to eliminate (morph out) the expressions from non-neutral face scan. The ICP based matching increased the computational time complexity.

Haar and Veltkamp (2010) built a strong multi-resolution PCA model using a limited collection of facial landmarks along with neutral and expression scans. A single morphable identity model and seven isolated morphable expression models per subject are built. Due to separate models for identity and expressions, an expression is neutralized and coefficients of identity model are utilized for face recognition. L_1 distance is used as similarity metric for face matching. This method requires initial manual localization of fiducial landmarks to build morphable models. Obviously performance of this technique is subject to accurate localization of the landmarks.

The morphable model-based approach proposed by Haar and Veltkamp (2010) is found to be the most robust expression invariant approach as they have proved its superiority in terms of identification and verification rates on the UND, GavabDB, BU-3DFE and FRGC v 2.0 databases.

4.2.3 Frequency domain approaches

A few approaches extract frequency domain information from 3-D facial shape and perform 3-D face recognition. Figure 13 depicts the overall procedure for frequency domain 3-D face recognition. The 3-D face is subjected to segmentation and registration algorithms for proper alignment. Further, the facial data is transformed into frequency domain for discriminant feature extraction and formation of template vectors. Templates are matched using established classification algorithms like SVM and nearest neighbor classifier. The frequency domain based approaches are summarized further on.

Cai and Da (2012) proposed shape filtering approach and utilized manifold harmonics to split a facial model into three distinct components relevant to their frequency domain representation. The multi-scale model separates an expressive facial shape into person-specific deformation, expression effects and noise based on its energy in frequency domain. The
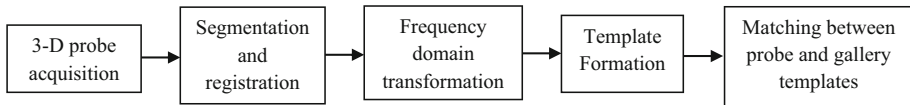
Fig. 13 Frequency domain 3-D face recognition scheme

intra-person variations are captured using low frequency components. The depth distance based method utilized for nose tip localization is not robust for non-frontal poses. Classical frequency domain approaches are computationally complex.

Peijiang et al. (2013) developed the harmonic feature based approach using energies in spherical harmonics at different frequencies. Point cloud data is represented as the spherical depth map (SDM) which acquires both abstract and fine details of a facial surface. The spherical harmonic features (SHF) were extracted from the SDM. It is convenient to handle different poses as pose variation is simply a rigid rotation of the fitted sphere. The recognition performance of this approach degrades when both occlusions and expression variations are considered simultaneously. This is also computationally complex as occlusions and expression variations are handled in frequency domain.

Elaiwat et al. (2014) presented the curvelet transform based 3-D face recognition approach. The reported algorithm initially inspects the curvelet coefficients for fiducial keypoints detection. The keypoint selection is based on the magnitude of the curvelet coefficients. Each keypoint is characterized with direction, scale and spatial as well as positions in the curvelet domain. The resultant rotation invariant feature vector is computed at multiple scales. It is evident from this approach that scale-2 curvelet coefficients contain more discriminant information. Computation of the curvelet coefficients is computationally light compared to other statistical features.

Table 10 presents frequency based approaches and their vital statistics.

The multi-scale frequency domain approach proposed in Cai and Da (2012) achieved a high identification rate on a complex database like FRGC v 2.0. The importance of low and high frequency components both has been established by representing the person specific and discriminating facial information in terms of the frequency domain features. It is clear that, the SDM based SHF frequency domain approach adopted by Peijiang et al. (2013) does not yield very high verification rates though it yields good identification rates on all the listed databases in Table 10. In this technique the SHF features of the same face with different expressions in frequency domain SDM map lie very close to each other and hence the Euclidean distance classifier based matching of an expressive face with its other versions fails considerably. On the other hand, when the SHF features of different faces are mapped into the SDM domain they are wide apart and hence identification rates are high compared to verification rates. The frequency domain approach of detecting key-points using curvelet transform employed by Elaiwat et al. (2014) has yield good verification rate but moderate identification rate in presence of expression variations. Curvelet transform based features depend upon detection of sharp directional facial curves and subsequent feature extraction. The facial curves like face boundary, eye-sockets, nose tip and nose base are not much affected by expression variations represented by large contours and contribute dominantly to the feature matching process. This results in higher verification rate. But matching of only key-points which may not be very robust to expression variations with multiple database image key-points as required in identification is definitely not going to yield high matching scores due to the following two reasons: (i) other faces may have similar key-points and (ii) its own key-point features may be distorted to a large extent due to expression variations.

Table 10 Frequency domain approaches

Sr. no.	Author	Year	Dataset	Persons in dataset	Images in dataset	Identification rate (%)	Verification rate (%)	Classifier/feature detection
1	Cai and Da (2012)	2012	FRGC v 2.0	466	4007	97.50	N.A.	N.A.
2	Peijiang et al. (2013)	2013	FRGC v 2.0	466	4007	96.94	90.00	Relief-F Robnik-Šikonja and Kononenko (2003), mRmR Peng et al. (2005), Random Forests Breiman (2001)
			SHREC2007	64	640	97.86	92.00	
			Bosphorus	105	4666	95.63	81.40	
3	Elaiwat et al. (2014)	2014	FRGC v 2.0	466	4007	94.40	97.83	N.A.

5 Occlusion invariant 3-D face recognition

While capturing non-cooperative subjects under uncontrolled environments, some part of the face may not be acquired as a subject may wear sunglasses, hats, eyeglasses or face may be partially covered by hair. The non-availability of 3-D facial data due to external objects causes external occlusion. For a non-frontal pose of the subject, some parts of the face may not be captured during the scan. This results in missing data and is referred to as internal occlusion. Although many researchers dealt with expression variations, very few have attempted occlusion variation problem for 3-D face recognition. Assumed facial symmetry has an important role in occlusion invariant approaches. A few approaches use the curvature information from a statistical model to obtain relevant curves for missing data. In morphable model based approaches, whole facial model is fitted to the partial probe scan.

Perakis et al. (2009) proposed an algorithm to deal with internal occlusions. The annotated face model (AFM) is fitted into a probe scan on the basis of localization data from landmark detector. The AFM creates geometry images that are invariant to missing data. Thus, the approach addresses the incomplete data problem due to pose variations. Experimentation is performed with FRGC v 2.0 as well as a subset of side pose faces in the UND Ear dataset (F, G collection). Images in this dataset include ear of the subjects in the captured side poses. For 45° side scans, 119 left and 119 right side pose scans from 119 different persons are considered for experimentation along with 88 left and 88 right 60° pose face scans of 88 persons. UND45LR denotes a set of scans with 45° pose where for every person, the left pose scan belongs to gallery and the right pose scan is considered as probe. Similarly UND60LR denotes a collection of side scans at 60° pose where the left pose considered as gallery and the right pose is denoted as probe for each person. Also, UND45F stands for gallery with a single frontal scan from FRGC v 2.0 dataset and the probes are left and right 45° poses of a person. UND60F indicates a gallery set with a single frontal scan from FRGC v 2.0 dataset and 60° left and right poses are denoted as probes.

Drira et al. (2013) proposed an algorithm to address occlusion as well as expression invariance. Facial surface is transformed into radial curves starting from the nose tip and ending at the facial boundary. A curve quality filter is utilized to handle internal occlusions. It checks continuity and length of facial curves. If a particular curve does not pass the quality filter test, it is discarded and restored later by reconstructing the missing data. In order to handle external occlusions, the points belonging to object, which cause occlusion, are detected. Subsequently, the occluding area is removed from the facial surface. The broken curves, caused by the removal of occluding object, are completed with the help of PCA-based statistical model. This model is used to complete the incomplete curves using training data. After the retrieval of missing curves, they have utilized a complete set of curves as an input to the proposed curvature based 3-D face recognition system. They have validated the algorithm with Bosphorus database (Savran et al. 2008) which contains four distinct occlusion variations per subject.

Alyuz et al. (2013) addressed the problem of external occlusions. They have proposed a registration framework in which, a possible non-occluded model is adaptively selected for each probe face, by employing non-occluded facial parts. After the model selection, occluded portions of the face are removed. A masked projection is computed using a set of non-occluded distinct regions. Furthermore, the classification is performed with the 1-nearest neighbor (*1-NN*) classifier. They have experimented with Bosphorus and UMB-DB database. These are the most complex occluded face databases available till-date, with realistic occlusions. This

algorithm also relies on accurate nose tip detection. Since the UMB-DB database has some occlusions covering nose area, the overall system performance is adversely affected.

Smeets et al. (2013) proposed the local feature based MeshSIFT algorithm to deal with missing 3-D data (Maes et al. 2010). They performed the experimentation with SHREC'11 (Veltkamp et al. 2011), UND45LR and UND60LR (Perakis et al. 2009; Passalis et al. 2011) databases. MeshSIFT algorithm automatically extracts salient features from scale space representation of a facial surface. Symmetric MeshSIFT method improves the performance by flipping the histogram vector. Mean distances between the corresponding landmarks were used for face matching. The pose normalization is performed using the RANdom SAmple Consensus (RANSAC) technique which is computationally complex in terms of estimation time (Fischler and Bolles 1981).

Berretti et al. (2013) utilized probe faces acquired under unconstrained situations. They used Scale Invariant Feature Transform (SIFT) key-point detection approach to locate key-points on the depth image along with facial curves that connected those key-points. They have utilized 45° and 60° side-pose scans from the UND database. Since the same organization has collected the UND and FRGC v 2.0 databases, they have found 39 common subjects between 45° side-pose scans in the UND and frontal scans in the FRGC v 2.0. Also, 33 subjects were found as common in 60° side-pose scans in the UND and frontal faces in the FRGC v 2.0. Face matching is achieved using the curvature information across landmarks.

Bellil et al. (2014) introduced the Gappy wavelet transform-based neural network for occlusion invariant 3-D face recognition. They have treated occlusions as regional face deformations. The difference between wavelet sub-bands of a probe face and a generic facial model indicated the regional deformations. Subsequently, the occluded regions are restored using a multi library wavelet transform neural network. A gallery face with minimum distance in the Gappy Wavelet transform neural network (GWNN) based parametric space is declared as the closest match. The authors have manually performed the 3-D face registration. The performance of this technique is subject to the manual 3-D face registration. Apart from the one time neural network training required at the time of new face registration, the algorithm is computationally light.

Performance of various approaches incorporating occlusion invariant 3-D face recognition is presented in the Table 11.

Smeets et al. (2013) has achieved good occlusion invariant recognition rates on small databases like SHREC'11, UND60LR and UND45LR. The SIFT features of the usually symmetric human faces detected from the mesh 3-D representations result in high degree of matching while using the selected databases. Bellil et al. (2014) used comparatively large databases: Bosphorus and UMB-DB and achieved good identification rates. The estimation of intermediate features using Gappy wavelet-based interpolation and neural networks based classification resulted into good identification results.

6 Discussion

In recent years, the 3-D face recognition field saw a sharp growth in terms of number of publically available large size databases, features, algorithms and even practically deployed systems. Despite large research efforts invested so far, a robust 3-D face recognition system capable of operating under unconstrained situations could not be implemented. During the initial years of 3-D face recognition research, the iterative closest point (ICP) was the most preferred algorithm for rigid face registration and matching of 3-D faces. Due to the changes

Table 11 Occlusion invariant 3-D face recognition approaches

Sr. no.	Author	Year	Dataset	Persons in dataset	Images in dataset	Identification rate $(R_1\text{-}RR)$ (%)	Classifier
1	Perakis et al. (2009)	2009	UND45LR	119	119	67.00	N.A.
			UND60LR	88	88	64.00	
			UND45F	466	466	69.00	
			UND60F	466	466	44.00	
2	Alyuz et al. (2013)	2013	Bosphorus	105	381	83.99	Nearest neighbor
			UMB-DB	142	590	65.25	
3	Smeets et al. (2013)	2013	SHREC'11	130	780	98.57	N.A.
			UND60LR	66	66	100.00	
			UND45LR	120	120	98.31	
4	Berretti et al. (2013)	2013	UND45left	39	39	84.60	N.A.
			UND45right	39	39	79.50	
			UND45Avg	78	78	82.05	
			UND60left	33	33	66.70	
			UND60right	33	33	66.70	
			UND60Avg	66	66	66.70	
			UNDleftAvg	72	72	76.40	
			UNDrightAvg	72	72	73.60	
			UNDtotalAvg	114	114	75.00	
5	Bellil et al. (2014)	2014	UMB-DB	951	951	97.12	GWNN
			Bosphorus	105	4652	93.29	

in shape caused by various expressions, facial surface appears to be non-rigid and rather elastic. ICP is costly in terms of time complexity (Boehnen et al. 2009). Therefore, in pursuit of developing an accurate and robust system with small response time, researchers started representing facial surface into other representations such as signed shape difference map (Yue et al. 2010), intrinsic coordinate system (Spreeuwers 2011), spherical depth map (Peijiang et al. 2013) etc. It has been observed that in an effort to reduce the computational complexity one has to sacrifice the robustness of matching or recognition. Although, 3-D face recognition is consolidating its importance in the biometric community, there are a few shortcomings of 3-D face recognition techniques; such as 3-D face scanners require specific ambient illumination (Bowyer et al. 2006). Even after capturing faces in a controlled environment, holes (missing data) and spikes are present in the data during the acquisition process which require subsequent processing. If they are not addressed properly, the recognition efficiency may be adversely affected.

As the field of view (FOV) of a scanner is limited, profile scans of a facial surface suffer from internal occlusion as the other side of the face is not visible. Therefore, internal occlusions due to pose variations is an intrinsic challenge while processing non-frontal 3-D facial scans. However, it is clear that the reconstruction of missing data due to occlusion is computationally heavy. Additionally, since aging changes the human facial shape over time, capturing age variation databases (Smeets et al. 2012) and extracting age invariant facial features is essential. A detailed study of transformation of features with age will also be very important. Practically, capturing of such age variation database and subsequent computation of age invariant features is a tough task. Even very rigid human facial features like jaws, cheek bones, nose shape etc. deform due to ageing. Another issue with all the biometric techniques is that they are prone to spoofing and deception. But the 3-D face recognition techniques are minimum prone to such attacks due to the available details of the facial information, its possible integration with 2D data and other biometric techniques (Määttä et al. 2012).

In our opinion, the process of human face observation, abstracting, memorizing and subsequent recognition by a human brain consists of many psycho-visual, neural, basic scientific and mathematical processes in the background. If the ongoing research is supported by some fundamental research in psycho-visual activities in this area and their repercussions on the facial structure and change in features due to various expressions, it may throw light on a few new expression invariant human face features. Neural activities in human brain after observing a human face are also of very high importance and are supported by human intelligence. As human beings, we can recognize a person even from a 2-D picture of a resolution adequate for human eyes. What goes on in this recognition process is still a mystery. The only thing that can be pointed out at this stage is the study of psycho-visual processes that take place while observing and recognizing a human face may be of great help in designing robust face recognition algorithms and systems. Also it is obvious that the human face recognition process might not simultaneously compute or estimate all the features required for recognizing a face out of human memory. Instead, based on some primary features, it either accepts the face for further matching or rejects it only on the basis of mismatch of a few initial parameters. In other words, the matching process is hierarchical. This is also a very crucial step that may require high level of human intelligence. Thus, the initial rejection step in Li and Da (2012) may be very important in minimizing the recognition time per probe. In computer based face recognition systems, the hierarchical approach can be implemented using well organized or specially structured databases. The effects of laser scans on living tissues have not so far been studied extensively. Such study may add a different dimension to the utility of laser or other wavelength scanners for face scanning. It is evident from discussions that computational systems with very high capabilities are a mandatory part of 3-D face recognition systems

introduced so far. Our opinions and critical comments on advantages and limitations of the specific 3-D face acquisition, registration, feature extraction and matching algorithms have already been presented in the respective sections. The most important 3-D face databases for evaluation and benchmarking of novel face algorithms have also been pointed out for general applications. However, a specific application may demand use of a specific database specially designed or captured under constrained conditions.

Brief discussions on the major approaches discussed in the state of the art selected publications have been presented. Most of them use ICP technique for face registration and PCA and its derivatives for feature extraction followed by subsequent nearest neighbor matching. Both of them are computationally heavy and blind techniques. Recently, researchers have proposed alternate methods for 3-D face registration like simulated annealing (Queirolo et al. 2010), 3-D signatures (Ross et al. 2006), transformation to intrinsic coordinate system (Spreeuwers 2011) etc. This review has already presented significant discussions on expression invariant and occlusion invariant major 3-D face recognition work.

The traditional 2-D face recognition techniques that are essentially face recognition from either gray-scale or color facial images have been researched for more than last 40 years. Many algorithms performed well under constrained conditions and are deployed as limited capacity commercial face recognition systems under such conditions. Face Recognition Vendor Tests FRVT-2000 (Blackburn et al. 2001), FRVT-2002 (Phillips et al. 2003) were organized to evaluate the performance of commercially deployed algorithms under challenging conditions. Conclusions of FRVT-2000 highlight the need for further research on handling facial images with indoor and outdoor illumination variations. FRVT-2002 recorded significant progress on handling indoor face recognition and eventually expressed the need of improvement in outdoor face recognition systems under uncontrolled illumination conditions.

The disadvantages of 2-D face recognition are:

I. Performance of 2-D face recognition algorithms badly suffers due to illumination variations, pose changes, poor image quality, occlusions and facial expression variations (Ocegueda et al. 2013; Tang et al. 2013; Al-Osaimi et al. 2012; Petrovska-Delacretaz et al. 2008).

II. Since 2-D intensity image is a projection of original 3-D facial shape, some geometric information is lost during this mapping (Yue et al. 2010; Smeets et al. 2010). Human beings are intelligent enough to recover the lost 3-D information for face recognition process even from 2-D images.

However, 2-D face recognition has the inherent advantage of lower dimensionality of the feature space, and subsequently simpler registration, feature extraction and low recognition time. It should be noted that general algorithmic steps for 3-D face recognition and 2-D face recognition are similar.

Based on this survey work, we present functional steps of a robust 3-D face recognition system in Fig. 14.

7 Conclusion and future scope

In this paper, a comprehensive survey of 3-D face acquisition, representation, registration, feature detection and matching techniques have been presented along with our opinions and comments. Initially we summarized various available 3-D face databases. Almost all researchers used three basic approaches for 3-D facial data representations namely (1) point cloud (2) mesh (3) range data. Out of these three representations, range data is found to
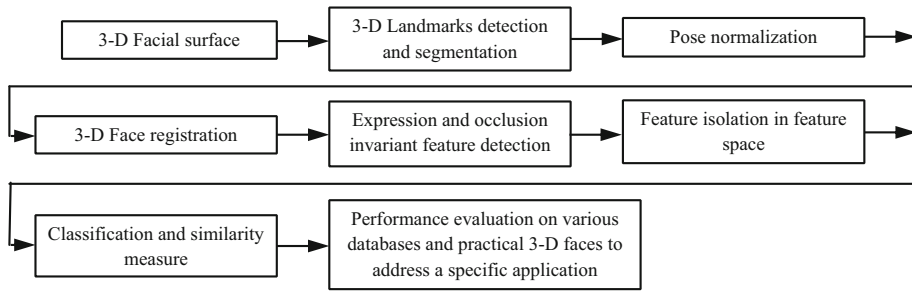
Fig. 14 General flow of a 3-D face recognition system

be simplest for handling and processing due to its similarity with 2-D data representations. Mesh data is found to be most complex due to heavy statistical representation of every mesh element. However, this is preferred due to its suitability in designing pose, expression and occlusion invariant robust face recognition algorithms. Fast, deterministic and accurate robust feature detection, registration and matching techniques are the need of the hour. Though the ICP algorithm requires more disk space for storing probe and gallery samples, and high computational time for registration, it is most frequently used because of its robustness. LDA is the most frequently used technique for dimensionality reduction. It is obvious that 3-D shape contributes more information compared to conventional 2-D intensity images. However, acquisition of very accurate 3-D data in itself is a major challenge. Furthermore, the 3-D data acquisition and processing algorithms are expected to be computationally heavier and demand larger primary and secondary memory sizes leading to increased costs of 3-D face recognition systems.

Spreeuwers (2011) offered the most accurate, robust rank-1 recognition rate on the most complex databases such as FRGC v 2.0 Smeets et al. (2012). LBP and LSP-based approaches offer reasonably robust features at moderate computational overheads. Many algorithms require inputs like nose-tip, eye location from human operators. Performance of such algorithms directly depends on the accuracy of the information provided by the human operator.

After 40 years of rigorous research, the 2-D face recognition still face shortcomings like sensitivity to pose, illumination and occlusion variations. This is where 3-D face recognition finds space for itself. Although major computer-based systems process 3-D information with reasonable computation time, it is challenging to deploy such a system on portable embedded platform till date. Thus, in general, it can be inferred that 3-D face recognition is a wide open area for research; in terms of robustness, accuracy and computational time.

A desired 3-D face recognition system should have the following aspects:

1. Robust to expression, pose, illumination, occlusion, ageing, makeup variations.
2. Ability to identify a person at a reasonably long distance from 3-D acquisition sensor.
3. Have higher Rank-one Recognition Rates (R_1-RR) and verification rates (VR) with lower false acceptance rates (FAR).
4. Fast and accurate 3-D face registration mechanism with a large database (more than 10,000 faces) with minimum training data.
5. Should have provision for easy registration, algorithmic up-gradation, expansion and maintenance of the database via a user friendly graphical user interface (GUI) in minimum time.
6. Should be deployable in the form of portable embedded system (Kisačanin and Nikolić 2010).

7. Should have fast and portable 3-D data acquisition devices with low cost and least acquisition spikes and noise.
8. Should be capable to accurately differentiate and recognize similar faces like that of twins.
9. The system should be available at a reasonable cost.

Many state-of-the-art 3-D face scanners acquire texture (2-D intensity) information along with 3-D shape. Thus, multimodal (2-D + 3-D) approaches can be explored in the coming years to achieve more accuracy. Most 3-D approaches do not concentrate on color information or depth information represented by colors. This may be an interesting area for future work. Detection of 3-D facial landmarks under challenging conditions (such as pose variations, occlusion and expression variations) is also an open research problem. After observing the performance characteristics of occlusion invariant approaches, it is clear that, there is a lot of scope in research areas related to 3-D occlusion invariant face matching. Face tampering is difficult to hide specially in case of 3-D face scanning techniques. Microsoft Kinect 3-D face acquisition technology (Kinect 2013) is becoming popular due to its very low cost, high capture speed (0.033 s) and small size (41.25 inch^3) (Li et al. 2013) compared to other 3-D scanners. The limitation of the data acquired from the Kinect scanners is the low resolution of the output scan and high noise contents. In the coming years, with the use of robust preprocessing techniques, it may be possible to perform real time 3-D face recognition with low cost and light weight sensors in real time. After introduction of the 'Twin-Expression Challenge' (3-D-TEC) (Vijayan et al. 2011) database, twin 3-D face recognition is an approachable research area. In view of age-invariant 3-D face recognition techniques, 3-D facial ageing model (Unsang et al. 2008) may provide further cues for research that may lead to age invariant recognition. Best possible combinations of modules such as preprocessing and registration techniques, feature extraction algorithms and data retrieval methods can be worked out by comparison of existing individual module performance. Robust 3-D feature template data protection techniques are needed to be discovered. As discussed in Sect. 6, the psycho-visual processes and structural changes in human faces due to expression variations and aging provide a lot of scope for research in this area not only for engineers but also for psychologists, physiologists and anthropologists.

References

Abate AF, Nappi M, Riccio D, Sabatino G (2006) 3D face recognition using normal sphere and general Fourier descriptor. In: 18th international conference on pattern recognition, pp 1183–1186. doi:10.1109/ICPR.2006.25

Abate AF, Nappi M, Riccio D, Sabatino G (2007) 2D and 3D face recognition: a survey. Pattern Recogn Lett 28:1885–1906. doi:10.1016/j.patrec.2006.12.018

Achermann B, Bunke H (2000) Classifying range images of human faces with Hausdorff distance. In: 15th international conference on pattern recognition, pp 809–813. doi:10.1109/ICPR.2000.906199

Al-Osaimi F, Bennamoun M, Mian A (2009) An expression deformation approach to non-rigid 3D face recognition. Int J Comput Vis 81:302–316. doi:10.1007/s11263-008-0174-0

Al-Osaimi FR, Bennamoun M, Mian A (2012) Spatially optimized data-level fusion of texture and shape for face recognition. IEEE Trans Image Process 21:859–872. doi:10.1109/TIP.2011.2165218

Alyuz N, Gokberk B, Akarun L (2010) Regional registration for expression resistant 3-D face recognition. IEEE Trans Inf Forensics Secur 5:425–440. doi:10.1109/TIFS.2010.2054081

Alyuz N, Gokberk B, Akarun L (2013) 3-D face recognition under occlusion using masked projection. IEEE Trans Inf Forensics Secur 8:789–802. doi:10.1109/TIFS.2013.2256130

Alyuz N, Gokberk B, Dibeklioglu H, Akarun L (2008) Component-based registration with curvature descriptors for expression insensitive 3D face recognition. In: 8th IEEE international conference on automatic face and gesture recognition, pp 1–6. doi:10.1109/AFGR.2008.4813359

Alyuz N, Gokberk B, Spreeuwers L, Veldhuis R, Akarun L (2012) Robust 3D face recognition in the presence of realistic occlusions. In: 5th IAPR international conference on biometrics, pp 111–118. doi:10.1109/ICB.2012.6199767

Amberg B, Knothe R, Vetter T (2008) Expression invariant 3D face recognition with a Morphable model. In: 8th IEEE international conference on automatic face and gesture recognition, pp 1–6. doi:10.1109/AFGR.2008.4813376

Amor BB, Ardabilian M, Chen L (2013) 3D face modeling. In: 3D face modeling, analysis and recognition. Wiley, Singapore, pp 1–37. doi:10.1002/9781118592656.ch1

Ballihi L, Ben Amor B, Daoudi M, Srivastava A, Aboutajdine D (2012) Boosting 3-D-geometric features for efficient face recognition and gender classification. IEEE Trans Inf Forensics Secur 7:1766–1779. doi:10.1109/TIFS.2012.2209876

Bao-Cai Y, Yan-Feng S, Cheng-Zhang W, Yun G (2009) BJUT-3D large scale 3D face database and information processing. J Comput Res Dev 46:1009–1018

Belhumeur PN, Hespanha JP, Kriegman D (1997) Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19:711–720. doi:10.1109/34.598228

Bellil W, Brahim H, Ben Amar C (2014) Gappy wavelet neural network for 3D occluded faces: detection and recognition. Multimed Tools Appl:1–16 doi:10.1007/s11042-014-2294-6

Bellon O, Silva L, Queirolo C, Drovetto S, Pamplona M (2006) 3D face image registration for face matching guided by the surface interpenetration measure. In: IEEE international conference on image processing, pp 2661–2664. doi:10.1109/ICIP.2006.313057

Ben Amor B, Ardabilian M, Liming C (2008) Toward a region-based 3D face recognition approach. In: IEEE international conference on multimedia and expo, pp 101–104. doi:10.1109/ICME.2008.4607381

Berretti S, Del Bimbo A, Pala P (2010) 3D face recognition using isogeodesic stripes. IEEE Trans Pattern Anal Mach Intell 32:2162–2177. doi:10.1109/TPAMI.2010.43

Berretti S, Del Bimbo A, Pala P (2013) Sparse matching of salient facial curves for recognition of 3-D faces with missing parts. IEEE Trans Inf Forensics Secur 8:374–389. doi:10.1109/TIFS.2012.2235833

Besl PJ, McKay ND (1992) A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell 14:239–256. doi:10.1109/34.121791

Beumier C, Acheroy M (2000) Automatic 3D face authentication. Image Vis Comput 18:315–321. doi:10.1016/S0262-8856(99)00052-9

Blackburn DM, Bone M, Phillips PJ (2001) Face recognition vendor test 2000: evaluation report. DTIC Document, http://www.face-rec.org/vendors/FRVT_2002_Evaluation_Report.pdf. Accessed Dec 20, 2013

Boehnen C, Peters T, Flynn P (2009) 3D Signatures for fast 3D face recognition. In: Tistarelli M, Nixon M (eds) Advances in biometrics, vol 5558. Lecture notes in computer science. Springer, Berlin, pp 12–21. doi:10.1007/978-3-642-01793-3_2

Bowyer KW, Chang K, Flynn P (2006) A survey of approaches and challenges in 3D and multi-modal 3D + 2D face recognition. Comput Vis Image Underst 101:1–15. doi:10.1016/j.cviu.2005.05.005

Breiman L (2001) Random forests. Mach Learn 45:5–32. doi:10.1023/A:1010933404324

Breitenstein MD, Kuettel D, Weise T, Van Gool L, Pfister H (2008) Real-time face pose estimation from single range images. In: IEEE conference on computer vision and pattern recognition, pp 1–8. doi:10.1109/CVPR.2008.4587807

Cai L, Da F (2012) Estimating inter-personal deformation with multi-scale modelling between expression for three-dimensional face recognition. IET Comput Vis 6:468–479. doi:10.1049/iet-cvi.2011.0105

Chang KI, Bowyer W, Flynn PJ (2006) Multiple nose region matching for 3D face recognition under varying facial expression. IEEE Trans Pattern Anal Mach Intell 28:1695–1700. doi:10.1109/TPAMI.2006.210

Chellappa R, Wilson CL, Sirohey S (1995) Human and machine recognition of faces: a survey. Proc IEEE 83:705–741. doi:10.1109/5.381842

Chenghua X, Yunhong W, Tieniu T, Long Q (2004) Automatic 3D face recognition combining global geometric features with local shape variation information. In: Sixth IEEE international conference on automatic face and gesture recognition, pp 308–313. doi:10.1109/AFGR.2004.1301549

Chew WJ, Seng KP, Ang L-M (2009) Nose tip detection on a three-dimensional face range image invariant to head pose. In: Proceedings of the international multiconference of engineers and computer scientists, pp 858–862

Colbry D, Stockman G, Jain A (2005) Detection of anchor points for 3D face Veri.cation. In: IEEE computer society conference on computer vision and pattern recognition—workshops, pp 118–118. doi:10.1109/CVPR.2005.441

Colineau J, D'Hose J, Amor B, Ardabilian M, Chen L, Dorizzi B (2008) 3D face recognition evaluation on expressive faces using the IV2 database. In: Blanc-Talon J, Bourennane S, Philips W, Popescu D, Scheunders P (eds) Advanced concepts for intelligent vision systems, vol 5259. Lecture notes in computer science. Springer, Berlin, pp 1050–1061. doi:10.1007/978-3-540-88458-3_95

Colombo A, Cusano C, Schettini R (2006) 3D face detection using curvature analysis. Pattern Recogn 39:444–455

Colombo A, Cusano C, Schettini R (2008) Recognizing faces In 3D images even in presence of occlusions. In: 2nd IEEE international conference on biometrics: theory, applications and systems, pp 1–6. doi:10.1109/BTAS.2008.4699345

Colombo A, Cusano C, Schettini R (2011) UMB-DB: a database of partially occluded 3D faces. In: IEEE international conference on computer vision workshops, pp 2113–2119. doi:10.1109/ICCVW.2011.6130509

Conde C, Rodríguez-Aragón L, Cabello E (2006a) Automatic 3D face feature points extraction with spin images. In: Campilho A, Kamel M (eds) Image analysis and recognition, vol 4142. Lecture notes in computer science. Springer, Berlin, pp 317–328. doi:10.1007/11867661_29

Conde C, Serrano A (2005) 3D Facial normalization with spin images and influence of range data calculation over face verification. In: IEEE Computer Society conference on computer vision and pattern recognition, pp 115–115. doi:10.1109/CVPR.2005.379

Conde C, Serrano A, Cabello E (2006b) Multimodal 2D, 2.5D & 3D face verification. In: IEEE international conference on image processing, pp 2061–2064. doi:10.1109/ICIP.2006.312863

Cook J, Chandran V, Sridharan S, Fookes C (2004) Face recognition from 3D data using iterative closest point algorithm and gaussian mixture models. In: 2nd international symposium on 3D data processing, visualization and transmission, pp 502–509

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297. doi:10.1023/A:1022627411411

Cover T, Hart P (2006) Nearest neighbor pattern classification. IEEE Trans Inf Theor 13:21–27. doi:10.1109/tit.1967.1053964

Daoudi M, Srivastava A, Veltkamp R (2013) 3D face modeling, analysis and recognition. Wiley, London

Daugman JG (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Opt Soc Am J A Opt Image Sci 2:1160–1169

D'Hose J, Colineau J, Bichon C, Dorizzi B (2007) Precise localization of landmarks on 3D faces using gabor wavelets. In: First IEEE international conference on biometrics: theory, applications, and systems, pp 1–6. doi:10.1109/BTAS.2007.4401927

Di H, Ardabilian M, Yunhong W, Liming C (2012) 3-D face recognition using eLBP-based facial description and local feature hybrid matching. IEEE Trans Inf Forensics Secur 7:1551–1565. doi:10.1109/TIFS.2012.2206807

Dibeklioglu H, Salah AA, Akarun L (2008) 3D facial landmarking under expression, pose, and occlusion variations. In: 2nd IEEE international conference on biometrics: theory, applications and systems, pp 1–6. doi:10.1109/BTAS.2008.4699324

Di H, Guangpeng Z, Ardabilian M, Yunhong W, Liming C (2010) 3D face recognition using distinctiveness enhanced facial representations and local feature hybrid matching. In: Fourth IEEE international conference on biometrics: theory applications and systems, pp 1–7. doi:10.1109/BTAS.2010.5634497

Drira H, Ben Amor B, Srivastava A, Daoudi M, Slama R (2013) 3D face recognition under expressions, occlusions, and pose variations. IEEE Trans Pattern Anal Mach Intell 35:2270–2283. doi:10.1109/TPAMI.2013.48

Elaiwat S, Bennamoun M, Boussaid F, El-Sallam A (2014) 3-D face recognition using curvelet local features. IEEE Signal Process Lett 21:172–175. doi:10.1109/LSP.2013.2295119

Faltemier TC, Bowyer KW, Flynn PJ (2007) Using a multi-instance enrollment representation to improve 3D face recognition. In: First IEEE international conference on biometrics: theory, applications, and systems, pp 1–6. doi:10.1109/BTAS.2007.4401928

Faltemier TC, Bowyer KW, Flynn PJ (2008) A region ensemble for 3-D face recognition. IEEE Trans Inf Forensics Secur 3:62–73. doi:10.1109/TIFS.2007.916287

Farkas L (1994) Anthropometry of the head and face, 2nd edn. Raven Press, New York

Fels M, Olver P (1998) Moving coframes: I. A practical algorithm. Acta Appl Math 51:161–213. doi:10.1023/A:1005878210297

Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. ACM Commun 24:381–395. doi:10.1145/358669.358692

Freund Y, Schapire R, Abe N (1999) A short introduction to boosting. J Jpn Soc For Artif Intell 14:1612

Gang P, Yijun W, Wu Z (2003) Investigating profile extracted from range data for 3D face recognition. In: IEEE international conference on systems, man and cybernetics, pp 1396–1399. doi:10.1109/ICSMC.2003.1244607

Gokberk B, Dutagaci H, Ulas A, Akarun L, Sankur B (2008) Representation plurality and fusion for 3-D face recognition. IEEE Trans Syst Man Cybern Part B Cybern 38:155–173. doi:10.1109/TSMCB.2007.908865

Gökberk B, Salah A, Alyüz N, Akarun L (2009) 3D face recognition: technology and applications. In: Tistarelli M, Li S, Chellappa R (eds) Handbook of remote biometrics. Advances in pattern recognition. Springer, London, pp 217–246. doi:10.1007/978-1-84882-385-3_9

Gordon GG (1992) Face recognition based on depth and curvature features. In: IEEE computer society conference on computer vision and pattern recognition, pp 808–810. doi:10.1109/CVPR.1992.223253

Gupta S, Markey M, Bovik A (2010b) Anthropometric 3D face recognition. Int J Comput Vis 90:331–349. doi:10.1007/s11263-010-0360-8

Gupta S, Castleman KR, Markey MK, Bovik AC (2010a) Texas 3D face recognition database. In: IEEE southwest symposium on image analysis and interpretation, pp 97–100. doi:10.1109/SSIAI.2010.5483908

Gupta S, Markey MK, Aggarwal JK, Bovik AC (2007) Three dimensional face recognition based on geodesic and Euclidean distances. In: Proceedings of SPIE 6499, vision geometry XV, pp 64990D–64990D-6911. doi:10.1117/12.704535

Haasbroek ND (1968) Gemma Frisius. Rijkscommissie voor Geodesie, Delft, W. D. Meinema, Netherlands, Tycho Brahe and Snellius and their triangulations

Heseltine T, Pears N, Austin J (2008) Three-dimensional face recognition using combinations of surface feature map subspace components. Image Vis Comput 26:382–396. doi:10.1016/j.imavis.2006.12.008

Heseltine T, Pears N, Austin J (2004) Three-dimensional face recognition: an eigensurface approach. In: International conference on image processing, pp 1421–1424. doi:10.1109/ICIP.2004.1419769

Hesher C, Srivastava A, Erlebacher G (2003) A novel technique for face recognition using range imaging. In: Seventh international symposium on signal processing and its applications, pp 201–204. doi:10.1109/ISSPA.2003.1224850

Huang D, Ouji K, Ardabilian M, Wang Y, Chen L (2011) 3D Face recognition based on local shape patterns and sparse representation classifier. In: Lee K-T, Tsai W-H, Liao H-Y, Chen T, Hsieh J-W, Tseng C-C (eds) Advances in multimedia modeling, vol 6523. Lecture notes in computer science. Springer, Berlin, pp 206–216. doi:10.1007/978-3-642-17832-0_20

Huang Y, Wang Y, Tan T (2006) Combining statistics of geometrical and correlative features for 3D face recognition. In: British machine vision conference, BMVA Press, pp 879–888. doi:10.5244/C.20.90

Huynh T, Min R, Dugelay J-L (2013) An efficient LBP-based descriptor for facial depth images applied to gender recognition using RGB-D face data. In: Park J-I, Kim J (eds) Computer vision—ACCV 2012 workshops, vol 7728. Lecture notes in computer science. Springer, Berlin, pp 133–145. doi:10.1007/978-3-642-37410-4_12

Hyoungchul S, Kwanghoon S (2006) 3D face recognition with geometrically localized surface shape indexes. In: 9th international conference on control, automation, robotics and vision, pp 1–6. doi:10.1109/ICARCV.2006.345192

Jahanbin S, Bovik AC, Hyohoon C (2008) Automated facial feature detection from portrait and range images. In: IEEE southwest symposium on image analysis and interpretation, pp 25–28. doi:10.1109/SSIAI.2008.4512276

Jahanbin S, Hyohoon C, Bovik AC (2007) Castleman KR three dimensional face recognition using wavelet decomposition of range images. In: IEEE international conference on image processing, pp 145–148. doi:10.1109/ICIP.2007.4378912

Jain AK, Flynn PJ, Ross AA (2008) Handbook of biometrics, 2nd edn. Springer, USA. doi:10.1007/978-0-387-71041-9

Jaiswal S, Bhadauria S, Jadon R, Divakar T (2011) Brief description of image based 3D face recognition methods. 3D Res 1:1–14. doi:10.1007/3DRes.04(2010)02

Je C, Lee KH, Lee SW (2013) Multi-projector color structured-light vision. Sig Process Image Commun 28:1046–1058. doi:10.1016/j.image.2013.05.005

Johnson AE, Hebert M (1998) Surface matching for object recognition in complex three-dimensional scenes. Image Vis Comput 16:635–651. doi:10.1016/S0262-8856(98)00074-2

Johnson AE, Hebert M (1999) Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Trans Pattern Anal Mach Intell 21:433–449. doi:10.1109/34.765655

Kakadiaris IA, Passalis G, Toderici G, Murtuza MN, Yunliang L, Karampatziakis N, Theoharis T (2007) Three-dimensional face recognition in the presence of facial expressions: an annotated deformable model approach. IEEE Trans Pattern Anal Mach Intell 29:640–649. doi:10.1109/TPAMI.2007.1017

Kin-Chung W, Wei-Yang L, Yu Hen H, Boston N, Xueqin Z (2007) Optimal linear combination of facial regions for improving identification performance. IEEE Trans Syst Man Cybern Part B Cybern 37:1138–1148. doi:10.1109/TSMCB.2007.895325

Kinect (2013) Kinect for Windows. http://www.microsoft.com/en-us/kinectforwindows/. Accessed Dec 20, 2013

Kisačanin B, Nikolić Z (2010) Algorithmic and software techniques for embedded vision on programmable processors. Sig Process Image Commun 25:352–362. doi:10.1016/j.image.2010.02.003

Koch R, Pears N, Liu Y (2012) 3D imaging, analysis and applications. Springer, London. doi:10.1007/978-1-4471-4063-4

Konica (2013) Konica minolta color, light and shape measuring instruments. http://sensing.konicaminolta.us/applications/3d-scanners/. Accessed Dec 20, 2013

Koschan A, Pollefeys M, Abidi M (2007) 3D imaging for safety and security (computational imaging and vision). Springer, New York. doi:10.1007/978-1-4020-6182-0

Koudelka ML, Koch MW, Russ TD (2005) A Prescreener for 3D face recognition using radial symmetry and the hausdorff fraction. In: IEEE computer society conference on computer vision and pattern recognition, pp 168–168. doi:10.1109/CVPR.2005.566

Lei Y, Bennamoun M, El-Sallam AA (2013) An efficient 3D face recognition approach based on the fusion of novel local low-level features. Pattern Recogn 46:24–37. doi:10.1016/j.patcog.2012.06.023

Lei Y, Bennamoun M, Hayat M, Guo Y (2014) An efficient 3D face recognition approach using local geometrical signatures. Pattern Recogn 47:509–524. doi:10.1016/j.patcog.2013.07.018

Li BYL, Mian AS, Wanquan L, Krishna A (2013) Using Kinect for face recognition under varying poses, expressions, illumination and disguise. In: IEEE workshop on applications of computer vision, pp 186–192. doi:10.1109/WACV.2013.6475017

Li SZ, Jain AK (2011) Handbook of face recognition, 2nd edn. Springer, London. doi:10.1007/978-0-85729-932-1

Li H, Huang D, Morvan J-M, Chen L, Wang Y (2014) Expression-robust 3D face recognition via weighted sparse representation of multi-scale and multi-component local normal patterns. Neurocomputing 133:179–193. doi:10.1016/j.neucom.2013.11.018

Li X, Da F (2012) Efficient 3D face recognition handling facial expression and hair occlusion. Image Vis Comput 30:668–679. doi:10.1016/j.imavis.2012.07.011

Lijun Y, Xiaozhou W, Yi S, Jun W, Rosato MJ (2006) A 3D facial expression database for facial behavior research. In: 7th international conference on automatic face and gesture recognition, pp 211–216. doi:10.1109/FGR.2006.6

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60:91–110. doi:10.1023/B:VISI.0000029664.99615.94

Määttä J, Hadid A, Pietikäinen M (2012) Face spoofing detection from single images using texture and local shape analysis. IET Biometrics 1:3–10. doi:10.1049/iet-bmt.2011.0009

Maes C, Fabry T, Keustermans J, Smeets D, Suetens P, Vandermeulen D (2010) Feature detection on 3D face surfaces for pose normalisation and recognition. In: Fourth IEEE international conference on biometrics: theory applications and systems, pp 1–6. doi:10.1109/BTAS.2010.5634543

Malassiotis S, Strintzis MG (2005) Robust real-time 3D head pose estimation from range data. Pattern Recogn 38:1153–1165. doi:10.1016/j.patcog.2004.11.020

Metaxas DN, Kakadiaris IA (2002) Elastically adaptive deformable models. IEEE Trans Pattern Anal Mach Intell 24:1310–1321. doi:10.1109/TPAMI.2002.1039203

Mian A (2011) Robust realtime feature detection in raw 3D face images. In: IEEE workshop onapplications of computer vision, pp 220–226. doi:10.1109/WACV.2011.5711506

Ming Y (2015) Robust regional bounding spherical descriptor for 3D face recognition and emotion analysis. Image Vis Comput 35:14–22. doi:10.1016/j.imavis.2014.12.003

Mohammadzade H, Hatzinakos D (2013) Iterative closest normal point for 3D face recognition. IEEE Trans Pattern Anal Mach Intell 35:381–397. doi:10.1109/TPAMI.2012.107

Moorthy AK, Mittal A, Jahanbin S, Grauman K, Bovik AC (2010) 3D facial similarity: automatic assessment versus perceptual judgments. In: Fourth IEEE international conference on biometrics: theory applications and systems, pp 1–7. doi:10.1109/BTAS.2010.5634494

Moreno A, Sanchez A (2004) GavabDB: a 3D face database. In: 2nd COST workshop on biometrics on the internet: fundamentals, advances and applications, pp 77–82

Mousavi MH, Faez K, Asghari (2008) A three dimensional face recognition using SVM classifier. In: Seventh IEEE/ACIS international conference on computer and information science, pp 208–213. doi:10.1109/ICIS.2008.77

Mpiperis I, Malassiotis S, Strintzis MG (2008) Bilinear models for 3-D face and facial expression recognition. IEEE Trans Inf Forensics Secur 3:498–511. doi:10.1109/TIFS.2008.924598

Nair P, Cavallaro A (2009) 3-D face detection, landmark localization, and registration using a point distribution model. IEEE Trans Multimed 11:611–623. doi:10.1109/TMM.2009.2017629

Ocegueda O, Tianhong F, Shah SK, Kakadiaris IA (2013) 3D face discriminant analysis using Gauss–Markov posterior marginals. IEEE Trans Pattern Anal Mach Intell 35:728–739. doi:10.1109/TPAMI.2012.126

Ocegueda O, Shah SK, Kakadiaris IA (2011) Which parts of the face give out your identity? In: IEEE conference on computer vision and pattern recognition, pp 641–648. doi:10.1109/CVPR.2011.5995613

Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987. doi:10.1109/TPAMI.2002.1017623

Passalis G, Perakis P, Theoharis T, Kakadiaris IA (2011) Using facial symmetry to handle pose variations in real-world 3D face recognition. IEEE Trans Pattern Anal Mach Intell 33:1938–1951. doi:10.1109/TPAMI.2011.49

Peijiang L, Yunhong W, Di H, Zhaoxiang Z, Liming C (2013) Learning the spherical harmonic features for 3-D face recognition. IEEE Trans Image Process 22:914–925. doi:10.1109/TIP.2012.2222897

Peng H, Fulmi L, Ding C (2005) Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 27:1226–1238. doi:10.1109/TPAMI.2005.159

Peng X, Bennamoun M, Mian AS (2011) A training-free nose tip detection method from face range images. Pattern Recogn 44:544–558. doi:10.1016/j.patcog.2010.09.015

Perakis P, Passalis G, Theoharis T, Toderici G, Kakadiaris IA (2009) Partial matching of interpose 3D facial data for face recognition. In: IEEE 3rd international conference on biometrics: theory, applications, and systems, pp 1–8. doi:10.1109/BTAS.2009.5339019

Petrovska-Delacretaz D et al. (2008) The IV2 multimodal biometric database (including Iris, 2D, 3D, stereoscopic, and talking face data), and the IV2-2007 evaluation campaign. In: 2nd IEEE international conference on biometrics: theory, applications and systems, pp 1–7. doi:10.1109/BTAS.2008.4699323

Phillips PJ et al. (2005) Overview of the face recognition grand challenge. In: IEEE computer society conference on computer vision and pattern recognition, vol 941, pp 947–954. doi:10.1109/CVPR.2005.268

Phillips PJ, Scruggs WT, O'Toole AJ, Flynn PJ, Bowyer KW, Schott CL, Sharpe M (2010) FRVT 2006 and ICE 2006 large-scale experimental results. IEEE Trans Pattern Anal Mach Intell 32:831–846. doi:10.1109/TPAMI.2009.59

Phillips P, Grother P, Micheals R, Blackburn D, Tabassi E, Bone J (2003) FRVT 2002: evaluation report. http://www.frvt.org/FRVT2002. Accessed Dec 20, 2013

Queirolo CC, Silva L, Bellon ORP, Segundo MP (2008) 3D face recognition using the surface interpenetration measure: a comparative evaluation on the FRGC database. In: 19th international conference on pattern recognition, pp 1–5. doi:10.1109/ICPR.2008.4761696

Queirolo CC, Silva L, Bellon ORP, Pamplona Segundo M (2010) 3D face recognition using simulated annealing and the surface interpenetration measure. IEEE Trans Pattern Anal Mach Intell 32:206–219. doi:10.1109/TPAMI.2009.14

Robnik-Šikonja M, Kononenko I (2003) Theoretical and empirical analysis of ReliefF and RReliefF. Mach Learn 53:23–69. doi:10.1023/A:1025667309714

Romero M, Pears N (2009) Point-pair descriptors for 3D facial landmark localisation. In: IEEE 3rd international conference on biometrics: theory, applications, and systems, pp 1–6. doi:10.1109/BTAS.2009.5339009

Ross AA, Anil K. Jain, and Karthik Nandakumar (2006) Levels of fusion in biometrics. In: Handbook of multibiometrics, vol 6. international series on biometrics. Springer, USA, pp 59–90. doi:10.1007/0-387-33123-9_3

Russ TD, Koch MW, Little CQ (2005) A 2D range hausdorff approach for 3D face recognition. In: IEEE computer society conference on computer vision and pattern recognition—workshops, pp 169–169. doi:10.1109/CVPR.2005.561

Russ T, Boehnen C, Peters T (2006) 3D face recognition using 3D alignment for PCA. In: IEEE computer society conference on computer vision and pattern recognition, pp 1391–1398. doi:10.1109/CVPR.2006.13

Sala Llonch R, Kokiopoulou E, Tošić I, Frossard P (2010) 3D face recognition with sparse spherical representations. Pattern Recogn 43:824–834 doi:10.1016/j.patcog.2009.07.005

Samir C, Srivastava A, Daoudi M (2006) Three-dimensional face recognition using shapes of facial curves. IEEE Trans Pattern Anal Mach Intell 28:1858–1863. doi:10.1109/TPAMI.2006.235

Sansoni G, Trebeschi M, Docchio F (2009) State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation. Sensors 9:568–601. doi:10.3390/s90100568

Savran A, Alyüz N, Dibeklioğlu H, Çeliktutan O, Gökberk B, Sankur B, Akarun L (2008) Bosphorus database for 3D face analysis. In: Schouten B, Juul N, Drygajlo A, Tistarelli M (eds) Biometrics and identity management, vol 5372. Lecture notes in computer science. Springer, Berlin, pp 47–56. doi:10.1007/978-3-540-89991-4_6

Scopigno R, Andujar C, Goesele M, Lensch H (2002) 3D data acquisition. http://www.gris.informatik.tu-darmstadt.de/mgoesele/download/Scopigno-2002-3DA.pdf. Accessed Dec 20, 2013

Segundo MP, Queirolo C, Bellon ORP, Silva L (2007) Automatic 3D facial segmentation and landmark detection. In: 14th international conference on image analysis and processing, pp 431–436. doi:10.1109/ICIAP.2007.4362816

Shotton J et al (2013) Real-time human pose recognition in parts from single depth images. ACM Commun 56:116–124. doi:10.1145/2398356.2398381

Smeets D, Claes P, Vandermeulen D, Clement JG (2010) Objective 3D face recognition: evolution, approaches and challenges. Forensic Sci Int 201:125–132. doi:10.1016/j.forsciint.2010.03.023

Smeets D, Claes P, Hermans J, Vandermeulen D, Suetens P (2012) A comparative study of 3-D face recognition under expression variations. IEEE Trans Syst Man Cybern Part C Appl Rev 42:710–727. doi:10.1109/TSMCC.2011.2174221

Smeets D, Keustermans J, Vandermeulen D, Suetens P (2013) meshSIFT: Local surface features for 3D face recognition under expression variations and partial data. Comput Vis Image Underst 117:158–169. doi:10.1016/j.cviu.2012.10.002

Smeets D, Keustermans J, Hermans J, Claes P, Vandermeulen D, Suetens P (2011) Symmetric surface-feature based 3D face recognition for partial data. In: International joint conference on biometrics, pp 1–6. doi:10.1109/IJCB.2011.6117539

Spreeuwers L (2011) Fast and accurate 3D face recognition. Int J Comput Vis 93:389–414. doi:10.1007/s11263-011-0426-2

Srivastava A, Liu X, Hesher C (2006) Face recognition using optimal linear components of range images. Image Vis Comput 24:291–299. doi:10.1016/j.imavis.2005.07.023

Srivastava A, Klassen E, Joshi SH, Jermyn IH (2011) Shape analysis of elastic curves in Euclidean spaces. IEEE Trans Pattern Anal Mach Intell 33:1415–1428. doi:10.1109/TPAMI.2010.184

Störmer A, Rigoll G (2008) A multi-step alignment scheme for face recognition in range images. In: 15th IEEE international conference on image processing, pp 2748–2751. doi:10.1109/ICIP.2008.4712363

Szeptycki P, Ardabilian M, Liming C (2009) A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking. In: IEEE 3rd international conference on biometrics: theory, applications, and systems, pp 1–6. doi:10.1109/BTAS.2009.5339052

Tai Sing L (1996) Image representation using 2D Gabor wavelets. IEEE Trans Pattern Anal Mach Intell 18:959–971. doi:10.1109/34.541406

Tan X, Chen S, Zhou Z-H, Zhang F (2006) Face recognition from a single image per person: a survey. Pattern Recogn 39:1725–1745. doi:10.1016/j.patcog.2006.03.013

Tang H, Yin B, Sun Y, Hu Y (2013) 3D face recognition using local binary patterns. Signal Process 93:2190–2198. doi:10.1016/j.sigpro.2012.04.002

ter Haar FB, Veltkamp RC (2009) A 3D face matching framework for facial curves. Graph Models 71:77–91. doi:10.1016/j.gmod.2008.12.003

ter Haar FB, Veltkamp RC (2010) Expression modeling for expression-invariant face recognition. Comput Graph 34:231–241. doi:10.1016/j.cag.2010.03.010

Torr PHS, Zisserman A (2000) MLESAC: a new robust estimator with application to estimating image geometry. Comput Vis Image Underst 78:138–156. doi:10.1006/cviu.1999.0832

Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3:71–86

Uchida N, Shibahara T, Aoki T, Nakajima H, Kobayashi K (2005) 3D face recognition using passive stereo vision. In: IEEE international conference on image processing, pp II-950–II-953. doi:10.1109/ICIP.2005.1530214

Unsang P, Yiying T, Jain AK (2008) Face recognition with temporal invariance: A 3D aging model. In: 8th IEEE international conference on automatic face and gesture recognition, pp 1–7. doi:10.1109/AFGR.2008.4813408

Veltkamp RC et al. (2011) SHREC'11 track: 3D face models retrieval. In: 4th Eurographics conference on 3D object retrieval, Llandudno, UK, Eurographics Association, pp 89–95. doi:10.2312/3dor/3dor11/089-095

Vezzetti E, Marcolin F (2012) 3D human face description: landmarks measures and geometrical features. Image Vis Comput 30:698–712. doi:10.1016/j.imavis.2012.02.007

Vijayan V et al. (2011) Twins 3D face recognition challenge. In: International joint conference on biometrics, pp 1–7. doi:10.1109/IJCB.2011.6117491

Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: IEEE computer society conference on computer vision and pattern recognition, pp 511–518. doi:10.1109/CVPR.2001.990517

Wang Y, Pan G, Wu Z, Wang Y (2006) Exploring facial expression effects in 3D face recognition using partial ICP. In: Narayanan PJ, Nayar S, Shum H-Y (eds) Computer vision, vol 3851. Lecture notes in computer science. Springer, Berlin, pp 581–590. doi:10.1007/11612032_59

Winkler S, Min D (2013) Stereo/multiview picture quality: overview and recent advances. Sig Process Image Commun 28:1358–1373. doi:10.1016/j.image.2013.07.008

Xi Z, Dellandrea E, Liming C, Kakadiaris IA (2011) Accurate landmarking of three-dimensional facial data in the presence of facial expressions and occlusions using a three-dimensional statistical facial feature model. IEEE Trans Syst Man Cybern Part B Cybern 41:1417–1428. doi:10.1109/TSMCB.2011.2148711

Xiaoguang L, Jain AK, Colbry D (2006) Matching 2.5D face scans to 3D models. IEEE Trans Pattern Anal Mach Intell 28:31–43. doi:10.1109/TPAMI.2006.15

Xiaoguang L, Jain AK (2006) Deformation modeling for robust 3D face matching. In: IEEE computer society conference on computer vision and pattern recognition, pp 1377–1383. doi:10.1109/CVPR.2006.96

Xiaoguang L, Jain AK (2008) Deformation modeling for robust 3D face matching. IEEE Trans Pattern Anal Mach Intell 30:1346–1357. doi:10.1109/TPAMI.2007.70784

Xu C, Tan T, Wang Y, Quan L (2006) Combining local features for robust nose location in 3D facial data. Pattern Recogn Lett 27:1487–1494. doi:10.1016/j.patrec.2006.02.015

Xueqiao W, Qiuqi R, Yue M (2010a) 3D face recognition using corresponding point direction measure and depth local features. In: IEEE 10th international conference on signal processing, pp 86–89. doi:10.1109/ICOSP.2010.5656654

Xueqiao W, Qiuqi R, Yue M (2010b) A new scheme for 3D face recognition. In: IEEE 10th international conference on signal processing, pp 657–661. doi:10.1109/ICOSP.2010.5656861

Xu C, Tan T, Li S, Wang Y, Zhong C (2006a) Learning effective intrinsic features to boost 3D-based face recognition. In: Leonardis A, Bischof H, Pinz A (eds) Computer vision, vol 3952. Lecture notes in computer science. Springer, Berlin, pp 416–427. doi:10.1007/11744047_32

Yeung-hak L, Jae-chang S (2004) Curvature based human face recognition using depth weighted Hausdorff distance. In: International conference on image processing, pp 1429–1432. doi:10.1109/ICIP.2004.1421331

Yi S, Lijun Y (2008) Automatic pose estimation of 3D facial models. In: 19th international conference on pattern recognition, pp 1–4. doi:10.1109/ICPR.2008.4760973

Yueming W, Jianzhuang L, Xiaoou T (2010) Robust 3D face recognition by local shape difference boosting. IEEE Trans Pattern Anal Mach Intell 32:1858–1870. doi:10.1109/TPAMI.2009.200

Yue M, Qiuqi R, Xiaoli L, Meiru M (2010) Efficient Kernel discriminate spectral regression for 3D face recognition. In: IEEE 10th international conference on signal processing, pp 662–665. doi:10.1109/ICOSP.2010.5655733

Zhang D, Lu G (2013) Biometrics: systems and applications. Springer, New York. doi:10.1007/978-1-4614-7400-5

Zhang H, Zhang Y, Guo Z, Lin Z, Zhang C (2011) 3D face recognition based on principal axes registration and fusing features. Front Electr Electron Eng China 6:347–352 doi:10.1007/s11460-011-0155-x

Zhao W, Chellappa R, Phillips PJ, Rosenfeld A (2003) Face recognition: a literature survey. ACM Comput Surv 35:399–458. doi:10.1145/954339.954342