Int J Comput Vis
DOI 10.1007/s11263-010-0360-8

Anthropometric 3D Face Recognition

Shalini Gupta · Mia K. Markey · Alan C. Bovik

Received: 3 July 2009 / Accepted: 20 May 2010

Abstract We present a novel anthropometric three dimensional (Anthroface 3D) face recognition algorithm, which is based on a systematically selected set of discriminatory structural characteristics of the human face derived from the existing scientific literature on facial anthropometry. We propose a novel technique for automatically detecting 10 anthropometric facial fiducial points that are associated with these discriminatory anthropometric features. We isolate and employ unique textural and/or structural characteristics of these fiducial points, along with the established anthropometric facial proportions of the human face for detecting them. Lastly, we develop a completely automatic face recognition algorithm that employs facial 3D Euclidean and geodesic distances between these 10 automatically located anthropometric facial fiducial points and a linear discriminant classifier. On a database of 1149 facial images of 118 subjects, we show that the standard deviation of the Euclidean distance of each automatically detected fiducial point from its manually identified position is less than 2.54 mm. We further show that the proposed Anthroface 3D recognition algorithm performs well (equal error rate of 1.98% and a rank 1 recognition rate of 96.8%), out performs three of the existing benchmark 3D face recognition algorithms, and is robust to the observed fiducial point localization errors.

S. Gupta (✉) · A.C. Bovik
The Department of Electrical and Computer Engineering,
The University of Texas at Austin, Austin, TX 78712, USA
e-mail: shalini.gupta@ieee.org

A.C. Bovik
e-mail: bovik@ece.utexas.edu

M.K. Markey
Department of Biomedical Engineering, The University of Texas,
Austin, TX 78712, USA
e-mail: mia.markey@mail.utexas.edu

Keywords 3D face recognition · Anthropometric · Local facial features · Geodesic distances

1 Introduction

Automated human face recognition is a non-trivial computer vision problem of considerable practical significance. It has numerous applications including automated secured access, automatic surveillance, forensic analysis, fast retrieval of records from databases in police departments, automatic identification of patients in hospitals, checking for fraud or identity theft, and human-computer interaction. Besides the need for automation, interest in computerized algorithms for face recognition is inspired by the need to develop objective measures of facial similarity.

Considerable research attention has been directed, over the past few decades, towards developing reliable automatic face recognition systems that use two dimensional (2D) facial images (Zhao et al. 2003). Three dimensional (3D) face recognition technology is now emerging, in part, due to the availability of improved 3D imaging devices and processing algorithms. For such techniques, 3D images of facial surfaces are acquired using 3D acquisition devices. For automated human face recognition, 3D facial images have some advantages over 2D facial images. Their pose can be easily corrected by rigid rotations in 3D space. They also provide structural information about the face (*e.g.*, surface curvature and geodesic distances), which cannot be obtained from a single 2D image. Lastly, 3D face recognition algorithms have been shown to be robust to variations in illumination conditions during image acquisition (Kukula et al. 2004).

The existing 3D face recognition algorithms can be broadly classified into 'holistic' and 'local feature based' techniques (Gupta et al. 2007c). The holistic techniques employ

Published online: 05 June 2010

information from the whole face or from large regions of the 3D face. These algorithms include those that are based on the *appearance* of facial range images (hereafter 'appearance based techniques'), *e.g.*, 'eigenfaces' and 'fisherfaces' applied to facial range images. Then, there are the holistic techniques, wherein, 3D facial surfaces are rigidly or non-rigidly aligned and compared (hereafter 'surface matching based'). The other class of 3D face recognition algorithms includes those that employ structural properties of local facial features (hereafter 'local feature based').

Numerous studies in the past have indicated the potential of face recognition algorithms that employ local facial features. For example, at the Face Recognition Vendor's Test (FRVT) in 2002 (Phillips et al. 2003), two of the top three 2D face recognition algorithms, namely local feature analysis (Penev and Atick 1996), and elastic bunch graph matching (EBGM) (Wiskott et al. 1997) were based on local facial features. In EBGM, a face is represented as an 'elastic bunch graph', comprised of Gabor wavelet coefficients computed at specific facial fiducial points, and 2D Euclidean distances between these points. Hüsken et al. (2005) developed a successful 2D+3D face recognition technique called 'hierarchical graph matching', which combined scores of 2D EBGM and 3D EBGM (EBGM applied to facial range images). Their technique was also one of the top performers at the Face Recognition Grand Challenge (FRGC) (Phillips et al. 2005), which was conducted in the year 2005 to evaluate the performance of state-of-the-art 3D face recognition algorithms at the time.

Despite the existence of this evidence in favor of local feature based face recognition techniques, such techniques for 3D face recognition have been investigated less deeply than the holistic techniques for two reasons. First and foremost, there have not been attempts to systematically identify the local discriminatory structural characteristics of the human face for automatic face recognition purposes. For the few reported 3D face recognition techniques that employ local facial features, the choice of facial landmarks has either been *ad hoc* (Gordon 1992; Wang et al. 2002; Moreno et al. 2003; Lee et al. 2005; Cadoni et al. 2009), or has been a straightforward extension of local 2D techniques to range images (Hüsken et al. 2005; Zhang and Wang 2009). Hüsken et al. observed that 2D EBGM outperformed 3D EBGM, and acknowledged that merely extending local 2D techniques to range images may not be optimal for identifying the discriminatory 3D structural characteristics of the human face.

Second, completely automatic local feature based 3D face recognition algorithms also require robust and accurate techniques for automatically detecting local facial features/fiducial points. However, techniques for 3D facial fiducial point detection are currently poorly developed. Progress in this research area is further restricted by the fact that the 'ground truth' locations of the facial fiducial points are not available for many publicly available 3D face databases. It is no surprise then that many studies of 3D or 2D+3D facial fiducial point detection (Gordon 1992; Wang et al. 2002; Hüsken et al. 2005; Lu et al. 2006) report results of *visual inspection only* and do not report error statistics for comparison against any form of 'ground truth' data.

In this paper, we address both these challenging open problems within the area of 3D face recognition. We present a novel, completely automatic Anthropometric 3D Face Recognition algorithm (Anthroface 3D) that employs discriminatory anthropometric local structural characteristics for the human face. Firstly, we employ the information about the structural diversity of human faces from the related scientific discipline of facial anthropometry (Farkas 1987, 1994) to systematically isolate discriminatory structural characteristics of the 3D face. Secondly, we develop novel effective 2D+3D algorithms to automatically and accurately detect 10 facial anthropometric fiducial points that are associated with the identified discriminatory facial characteristics. We develop a successful 3D face recognition algorithm that employs Euclidean and geodesic facial anthropometric distance features and a linear discriminant analysis (LDA) classifier. Lastly, we demonstrate the effectiveness of both the proposed automatic fiducial points detection and the 3D face recognition algorithms on a large database of 1149 3D faces. We compare the locations of the automatically detected points to those of manually detected points. We also demonstrate the significantly superior face recognition performance of the proposed Anthroface 3D algorithm relative to three existing state-of-the-art 3D face recognition algorithms.

This paper is organized as follows. In Sect. 2, we review the existing 3D face recognition algorithms. This is followed by a detailed description of the proposed Anthroface 3D algorithm in Sect. 3. The methodology and the data that were employed to evaluate the performance of all 3D face recognition algorithms in this paper are presented in Sect. 4. The significant results of this paper are outlined and discussed in Sect. 5. The paper concludes with Sect. 6, wherein, the main conclusions of this paper are summarized.

2 Background

The existing 3D face recognition algorithms can be broadly classified into three groups (Gupta et al. 2007c). First, there are the 3D appearance based techniques that are straightforward extensions of successful 2D appearance based techniques (Zhao et al. 2003) to facial range images. Statistical subspace projection methods including principal component analysis (PCA) or 'eigensurfaces' (Chang et al. 2005), independent component analysis (ICA) (Hesher et al. 2003),

and LDA or 'fishersurfaces' (Heseltine and Austin 2004; BenAbdelkader and Griffin 2005) are among the prominent existing appearance based 3D face recognition techniques. The underlying philosophy of these techniques is to regard a 3D facial image as an instance in an N dimensional feature space, where N is the number of pixels/points in the image. All human faces are modeled to lie on a linear subspace of this feature space (Duda et al. 2001). A statistical learning technique is employed to learn the linear subspace from an ensemble of facial range images. Facial images are projected onto the learned subspace and are compared by means of a suitable distance metric in that subspace. Besides the subspace projection techniques, other appearance based 2D face recognition techniques based on hidden Markov models have also been extended to facial range images (Malassiotis and Strintzis 2005; Tsalakanidou et al. 2005).

Of all the existing techniques for 3D face recognition, PCA has been explored most extensively. For FRGC 2005, the performance of 3D PCA was regarded as the baseline (Phillips et al. 2005). PCA is reported to perform well with small 3D facial databases of less than 100 subjects, but poorly for larger data sets. Analogous to 2D face recognition, 3D techniques based on ICA, and LDA have been reported to perform better than 3D PCA. With subspace projection algorithms, in general, gradient maps derived from facial range images have been reported to produce better recognition accuracies than facial range images. The performance of holistic appearance based techniques for 3D face recognition is generally greatly affected by the presence of outlier cases, cluttered backgrounds, occlusions, noise, and variations in facial expression and pose. Interpreting the discriminatory facial structural information that the appearance based techniques encode also remains an open problem (Phillips et al. 2005).

The second class of 3D face recognition techniques is that, wherein, pairs of 3D facial surfaces are rigidly or non-rigidly registered and compared. Among these, techniques based on the Iterative Closest Point (ICP) algorithm (Besl and McKay 1992), wherein, one 3D model is rotated and translated iteratively in space until its distance from the other model converges to a minimum, has been explored most extensively and is reported to be the most successful (Russ et al. 2005; Lu et al. 2006; Koudelka et al. 2005; Maurer et al. 2005). Metrics including the mean squared error, point-to-closest-point mean squared error, point-to-closest-surface mean squared error, and the partial Hausdorff distance (pH) have been employed to iteratively align 3D faces and to compute the final structural dissimilarity between them.

The ICP based 3D face recognition algorithms are reported to be robust to variable facial poses (Lu et al. 2006) and illumination conditions during image capture (Kukula et al. 2004). They are also reported to perform better than 3D PCA (Irfanöglu et al. 2004). On the downside, however, the ICP based registration procedure is not guaranteed to converge to a global minimum. It is computationally expensive and is also affected by the presence of variable facial expressions, which are non-rigid deformations of the facial surface (Lu et al. 2006; Maurer et al. 2005).

To deal with variable facial expressions, Bronstein et al. have proposed matching intrinsic representations of facial surfaces that are computed using multi-dimensional scaling (Bronstein et al. 2005). Recently, techniques to non-rigidly warp one surface into another, using non-elastic deformations (Samir et al. 2009) and elastic deformations (Srivastava et al. 2009) of facial geodesic curves, have also been proposed. However, these techniques have not been tested on a large data set to assess their 3D face recognition performance. In other approaches, the rigid comparison of only the nasal and eye socket regions of the face have also been investigated for expression invariant 3D face recognition (Mian et al. 2007). In yet another approach, the facial expression deformations are modeled using PCA and are removed from probe faces before they are compared to gallery faces (Al-Osaimi et al. 2009).

Finally, there are the local feature based 3D face recognition techniques that employ structural properties of local regions of the 3D face. For a sub-class of these techniques, various facial profile curves have been rigidly aligned and compared (Beumier and Acheroy 2001; Zhang et al. 2006). In numerous such studies, the central vertical facial profile has been noted to be effective at uniquely identifying individuals. Hence, techniques have also been investigated to automatically locate this natural axis of bilateral symmetry of the human face. In general, these algorithms that employ facial profiles for matching 3D faces are adversely affected by the presence of variable facial expressions. Besides facial profiles, local geometric characteristics of facial subregions including their positional co-ordinates, surface areas and curvatures, and 3D Euclidean distances, ratios of distances, joint differential invariants, and angles between the local facial regions, have been employed previously for 3D face recognition (Gordon 1992; Hüsken et al. 2005; Moreno et al. 2003; Lee et al. 2005; Zhang and Wang 2009; Cadoni et al. 2009). The shapes of facial landmarks have been quantified by Gaussian and mean curvature values (Gordon 1992; Moreno et al. 2003), Guassian-Hermite moments (Xu et al. 2004), 'point signatures' (Wang et al. 2002), 2D and 3D Gabor filter coefficients (Wang and Chua 2005; Hüsken et al. 2005), and Scale Invariant Feature Transform (SIFT) descriptors (Zhang and Wang 2009). The successful 2D face recognition technique of local feature analysis (Penev and Atick 1996), has also been applied to 3D facial images (BenAbdelkader and Griffin 2005).

Automatic 3D facial landmark detection algorithms are currently poorly developed. A few attempts have been made

to automatically locate facial landmarks on 3D models using surface (Gordon 1992; Hüsken et al. 2005; Lu et al. 2006; Cadoni et al. 2009) or profile (Zhang et al. 2006) curvature, or by aligning 3D models to generic facial templates with known landmarks (Irfanöglu et al. 2004). The tip of the nose has been detected as the most prominent point for 3D facial models in canonical frontal poses. However, this heuristic fails for faces in arbitrary poses. Moreover, none of these existing studies have reported the accuracy of their facial feature localization techniques against any form of 'ground truth' data.

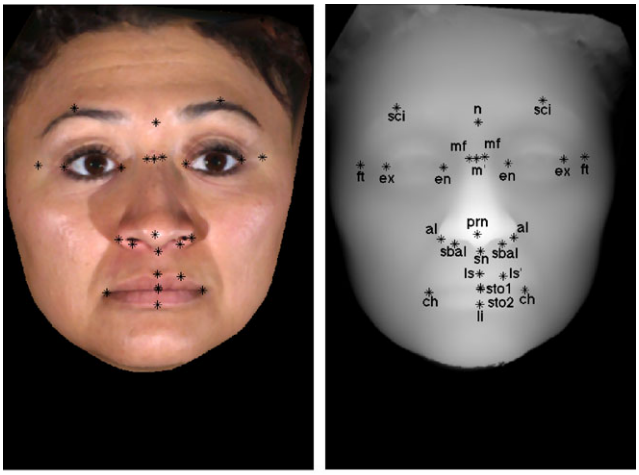
A number of techniques based on local facial features have been reported to perform better than 3D PCA (Irfanöglu et al. 2004). Some have also been reported to perform better than profile matching techniques (Irfanöglu et al. 2004), and ICP based techniques (Cadoni et al. 2009). They have been reported to be less affected by global changes in the appearance of facial range images including variable facial expressions, poses, and the presence of noise and occlusions than the holistic techniques (Mian et al. 2007). Nevertheless, 3D face recognition techniques based on local facial features have been explored much less than the holistic techniques.

3 Anthropometric 3D Face Recognition (Anthroface 3D)

3.1 Anthropometric Cranio-Facial Proportions

Anthropometric cranio-facial proportions (Farkas 1987) are ratios of pairs of straight-line and/or along-the-surface distances between specific cranial and facial fiducial points (*e.g.*, Fig. 1). For example, the most commonly used nasal index $N1$ is the ratio of the horizontal nose width to the vertical nose height ($N1 = (al - al)/(n - sn)$ from Fig. 1(b)). Note that for the anthropometric facial fiducial points, this notation of Farkas' (Fig. 1 and Table 1) is used throughout the paper.

The scientific discipline of cranio-facial anthropometry has existed for nearly three centuries. Over the years numerous anthropometric facial proportions have been proposed, and researchers have collected, recorded and analyzed their values on various human populations. Cranio-facial proportions have been and continue to be used widely employed in art and sculpture as neoclassical canons to aid in the creation of well-proportioned faces; in anthropology for analyzing prehistoric human remains (Comas 1960); for quantifying facial attractiveness (Farkas et al. 1985); for analyzing facial disproportionality in anomalies or after facial injury as an aid to planning facial cosmetic and reconstructive surgery (Farkas 1987; Rogers 1974); and recently for creating parametric models of human faces in computer graphics (DeCarlo et al. 1998). As far back as 1939, Hrdlička (1939)



(a) (b)

Fig. 1 The 25 facial fiducial points associated with highly variable anthropometric facial proportions on (**a**) a color image, and (**b**) a range image

Table 1 The 23 most variable anthropometric facial proportions for adult humans along with their standard deviation values (Farkas 1987). The corresponding fiducial points are presented in Fig. 1(b). N denotes nasal proportions, O denotes orbital proportions, L denotes proportions related to the mouth region, and F denotes facial proportions

S. No	Anthropometric Proportion	σ
1.	$O3 = (ex - en, l)/(en - en)$	7.75
2.	$O10 = (en - en)/(al - al)$	8.29
3.	$O12 = (en - en)(ch - ch)$	6.02
4.	$F32 = (n - sto1)/(ex - ex)$	5.30
5.	$N1 = (al - al)/(n - sn)$	5.81
6.	$N2 = (mf - mf)/(al - al)$	7.08
7.	$N4 = (sbal - sn, l + r)/(al - al)$	8.80
8.	$N6 = (ex - m'_{sag}, l)/(mf - mf)$	14.6
9.	$N7 = (sn - prn)/(al - al)$	6.28
10.	$N8 = (sn - prn)/(sbal - sn, l + r)$	12.8
11.	$N15 = (en - m'_{sag}, l)/(sn - prn)$	11.2
12.	$N16 = (en - m'_{sag}, l)/(en - m, l)$	7.26
13.	$N30 = (mf - mf)/(en - en)$	6.06
14.	$N31 = (ex - m'_{sag}, l)/(en - en)$	7.01
15.	$N32 = (al - al)/ch - ch$	5.04
16.	$N33 = (sn - prn)/(sn - sto1)$	13.8
17.	$L1 = (sn - sto1)/(ch - ch)$	5.40
18.	$L4 = (sn - ls)/(sbal - ls', l)$	10.2
19.	$L5 = (sn - ls)/(sn - sto1)$	5.97
20.	$L6 = (ls - sto1)/(sn - sto1)$	7.10
21.	$L7 = (ls - sto1)/(sn - ls)$	13.3
22.	$L9 = (ls - sto1)/(sto2 - li)$	16.9
23.	$L14 = (sn - sto1)/(n - sn)$	5.10

emphasized the importance of anthropometric facial proportions for comparing groups of people or populations. However, they have not been employed previously to aid in the design of 3D face recognition algorithms.

Farkas and Munro consolidated a list of 155 cranio-facial anthropometric proportions that are used for planning facial reparative and cosmetic surgery (Farkas 1987). By means of physical measurements, they also computed the mean (μ) and standard deviation (σ) values of these proportions for a population of 2564 healthy young and adult human subjects belonging to diverse ethnic, gender, and age groups (Farkas 1987, 1981). From amongst these 155 proportions, we isolated 70 anthropometric proportions that are associated with the facial region, and that can be computed automatically from the fontal 3D facial models normally acquired by 3D imaging devices. We identified a third (23) of these 70 facial proportions with the highest standard deviation values (Table 1) for adult human populations as being representative of discriminatory facial structural characteristics. It is reasonable to hypothesize that characteristics that display wide variation between individuals are likely to be most useful for distinguishing them. Associated with these 23 most variable anthropometric proportions are 25 anthropometric face fiducial points (Fig. 1). This information about the structural diversity of human faces forms the basis of our proposed Anthroface 3D face recognition algorithm (Gupta et al. 2007a, 2007b).

3.2 Manually Detected Fiducial Points

We manually located the 25 anthropometric facial fiducial points associated with the identified diverse anthropometric measurements on all color images (*e.g.*, Fig. 1(a)) of the Texas 3D Face Recognition Database (Gupta et al. 2010) that we employed. These fiducial points were located manually on the facial color images by clicking at appropriate locations with a mouse and a computer based graphical user interface. Since the images in this database were acquired using a stereo imaging system, the range and the color images for a particular acquisition are perfectly aligned. Hence, the locations of the fiducial points for the range images (Fig. 1(b)) are the same as their locations for the color images (Fig. 1(a)).

A previous statistical analysis, on a large database of 994 facial images, has established the reliability and repeatability of manually locating facial anthropometric fiducial points on 2D facial images (Shi et al. 2006). In their analysis, the authors observed that both the variability in the locations of facial fiducial points due to different human subjects, and the variability due to repeated observations by the same subject, were not statistically significant. Hence, manual identification of the fiducial points in our analysis served two purposes. First, it provided the 'ground truth' data for assessing the performance of the algorithms that we developed for automatically detecting facial fiducial points. Second, in a sense it helped to establish an upper bound on the expected performance of 3D face recognition algorithms for reliably detected manual anthropometric facial fiducial points.

Furthermore, in our analyses we first developed the Anthroface 3D face recognition algorithm using manually located facial fiducial points. We reasoned that only upon establishing the potential of our proposed algorithm for well detected manual fiducial points, would it be worthwhile to investigate approaches to automatically locate the fiducial points. This analysis with manually located facial fiducial points, also helped to identify the most useful subset of the 25 anthropometric facial fiducial points (Fig. 1) for the purposes of 3D face recognition.

3.3 Recognition Algorithm

As features for our proposed Anthroface 3D algorithm, we employed 300 3D Euclidean distances and 300 geodesic distances between all of the possible pairs ($\binom{25}{2} = 300$) of the 25 anthropometric facial fiducial points (Fig. 1(b)) that we identified in the previous analysis of facial anthropometric proportions (Gupta et al. 2007b). We computed geodesic distances along the facial surface using Dijkstra's shortest path algorithm (Dijkstra 1959; Tenenbaum et al. 2000). Besides 3D Euclidean distances, the motivation for employing geodesic distances was that previous studies have shown that geodesic distances are better at representing 'free-form' 3D objects than 3D Euclidean distances (Hamza and Krim 2006). Furthermore, a recent study suggested that changes in facial expressions (except for when the mouth is open) may be modeled as isometric deformations of the facial surface (Bronstein et al. 2005). When a surface is deformed isometrically, intrinsic properties of the surface, including Gaussian and mean curvature and geodesic distances, are preserved (Do Carmo 1976). Hence, algorithms based on geodesic distances are likely to be robust to changes in facial expressions.

From among the 300 Euclidean and 300 geodesic distances, we selected subsets of the most discriminatory distance features, using the stepwise linear discriminant analysis (Sharma 1996) procedure ('stepdisc', SAS Institute Inc., NC, USA). Briefly, the stepwise linear discriminant analysis procedure selects a subset of the most discriminatory features, which maximize a chosen statistical discrimination criterion. For our analyses we employed the Wilks' Λ criterion, which is the ratio of the within group sum of squares to the total sum of squares. The cut-off value of statistical significance for both the addition and removal of a feature was set at 0.05. The stepwise linear discriminant analysis procedure begins with no selected features, and at each step a feature is either added or removed. A feature that is already selected is removed if it does not statistically decrease

the discrimination power of the set of selected features, as measured by the selection criterion. If at a step no feature is removed, then a feature that is not already selected, but which adds statistically significantly to the discrimination power of the selected features, is added. This procedure is repeated until no feature is added or removed.

Using this procedure we identified the 106 and 117 most discriminatory Euclidean and geodesic distance features from among the 300 Euclidean and 300 geodesic distances, respectively. We pooled these 106 Euclidean and 117 geodesic anthropometric distances together, and using a second stage stepwise linear discriminant analysis procedure, we identified the final combined set of 123 most discriminatory anthropometric facial distance features. We trained a Fisher's linear discriminant analysis classifier (Duda et al. 2001), which linearly projected these 123 anthropometric distance features onto 11 dimensions (11D). For the stepwise LDA feature selection and for training the LDA classifier we used a training data set, which had no overlap with the test data set employed to test the final algorithm (Table 2). The training data set contained 360 randomly selected images of 12 subjects (30 images per subject) with neutral or arbitrary facial expressions.

The goal of Fisher's LDA is to find a linear projection matrix $\mathbf{W}$ for projecting the input feature vector $\mathbf{x}$ onto a linear subspace $\mathbf{y} = \mathbf{W}^T\mathbf{x}$, such that the discrimination between the different classes in the projected space is maximized (Duda et al. 2001). The discrimination between the different classes is measured by the ratio of their projected between class scatter matrix ($\mathbf{S}_B$) and the sum of their projected within class scatter matrix ($\mathbf{S}_W$) as

$$J(\mathbf{W}) = \frac{|\tilde{\mathbf{S}_B}|}{|\tilde{\mathbf{S}_W}|} = \frac{|\mathbf{W}^T\mathbf{S}_B\mathbf{W}|}{|\mathbf{W}^T\mathbf{S}_W\mathbf{W}|}, \tag{1}$$

where,

$$\mathbf{S}_B = \sum_{j=1}^{c} n_j(\mathbf{m}_j - \mathbf{m})(\mathbf{m}_j - \mathbf{m})^T, \tag{2}$$

$$\mathbf{S}_W = \sum_{j=1}^{c}\sum_{i=0}^{n_j}(\mathbf{x}_i - \mathbf{m}_j)(\mathbf{x}_i - \mathbf{m}_j)^T, \tag{3}$$

where c is the total number of classes in the training data (12 in our case), $\mathbf{m}_j$ is the vector mean of class j, n_j is the number of elements in class j (30 in our case), and $\mathbf{m}$ is the mean of all cases in the training data. Using Lagrange multipliers and the Rayleigh quotient, a closed form solution for $\mathbf{W}$ can be obtained by solving the eigen value problem

$$\mathbf{S}_W^{-1}\mathbf{S}_B\mathbf{w} = \lambda\mathbf{w} \tag{4}$$

and by computing the $(c-1)$ (in our case 11) eigen vectors ($\mathbf{w}$) and eigen values (λ) of ($\mathbf{S}_W^{-1}\mathbf{S}_B$).

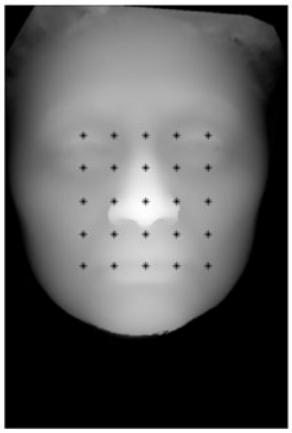


Fig. 2 The set of 25 arbitrarily located facial points

For all faces in the test data set, the 123 most discriminatory anthropometric Euclidean and geodesic distance features $\mathbf{x}$ were first computed. They were projected onto the 11D LDA space as $\mathbf{y} = \mathbf{W}\mathbf{x}$ that was learned using the training data set. The final metric for comparing two faces A and B was the Euclidean distance in the 11D LDA space computed as

$$d = \sqrt{\sum_{i=0}^{10}(y_i^A - y_i^B)^2} \tag{5}$$

3.4 Effect of Choice of Facial Points

We also investigated the effect of the choice of facial fiducial points on the performance of the proposed Anthroface 3D algorithm (Gupta et al. 2007a). We repeated the steps of the Anthroface 3D recognition algorithm (Sect. 3.3), with Euclidean and geodesic distances between 25 arbitrary facial points (Fig. 2) instead of the 25 anthropometric fiducial points (Fig. 1). These points were located in the form of a 5×5 rectangular grid positioned over the primary facial features of each face (Fig. 2). We chose these particular facial points as they measure distances between the significant facial landmarks, including the eyes, nose and the mouth regions, without requiring localization of specific fiducial points. A similar set of facial points was also employed in a previous 3D face recognition algorithm for aligning 3D facial surfaces using the ICP algorithm (Lu et al. 2006).

3.5 Subset of Anthropometric Fiducial Points

Next, we determined if a subset of the 25 manually located anthropometric facial fiducial points could be employed for the Anthroface 3D recognition algorithm, without a significant loss of its performance. This was an important step towards completely automating the proposed Anthroface 3D algorithm. Clearly, the task of automatically detecting all

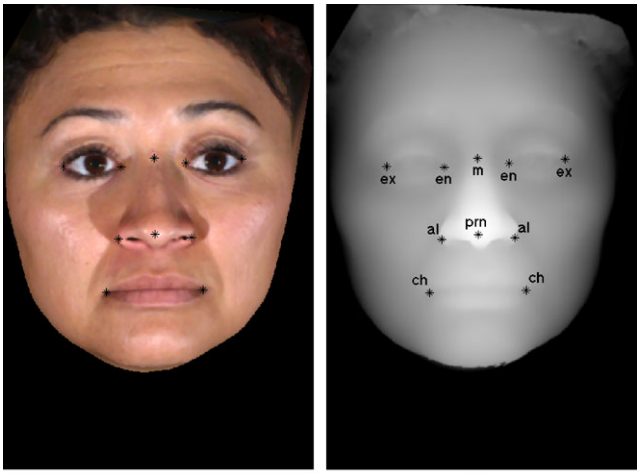


Fig. 3 The subset of 10 anthropometric facial fiducial points that were employed for the final automatic Anthroface 3D algorithm depicted on a (**a**) color, and (**b**) range facial image

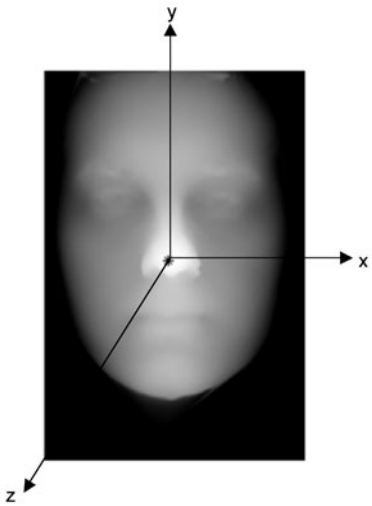


Fig. 4 The 3D template face in canonical frontal upright position which was used to automatically locate the nose tips of all 3D faces, and its manually located nose tip

of the 25 anthropometric fiducial points with a high accuracy is expensive, non-trivial and may even be redundant. To isolate this subset of points, we first removed individual points (*e.g. prn* in Fig. 1(b)), or pairs of symmetric points (*e.g. al-al* in Fig. 1(b)) from the overall Anthroface 3D algorithm and re-evaluated its performance. Interestingly, the removal of none of the individual points/pairs resulted in a statistically significant loss in the performance of the Anthroface 3D recognition algorithm. This indicated that some points in the set of 25 anthropometric points were clearly redundant. We then proceeded to remove larger groups of fiducial points associated with the orbital, nasal, and mouth regions and re-evaluated the overall performance of the Anthroface 3D algorithm. Finally, we isolated a subset of 10 anthropometric facial fiducial points (Fig. 3) that resulted in statistically equivalent recognition performance to that of the algorithm that employed 25 fiducial points. Hence, our final proposed Anthroface 3D algorithm employed only these 10 anthropometric facial fiducial points (Fig. 3), instead of the 25 points that we initially identified from the literature on anthropometric facial proportions (Sect. 3.1). Note also that for this Anthroface 3D algorithm, which employed $\binom{10}{2} = 45$ 3D Euclidean and 45 geodesic distances, nearly all (78) final features were selected using the stepwise linear discriminant procedure as being the most discriminatory. The order of the number of discriminatory features selected for this algorithm was similar to the order of the number of features selected (123) for the Anthroface 3D algorithm with 25 anthropometric points.

3.6 Automatic Detection of Anthropometric Fiducial Points

To completely automate the Anthroface 3D algorithm, we developed algorithms to automatically detect the 10 anthropometric facial fiducial points (Fig. 3) that we isolated in the previous analysis. Out algorithm automatically detects 3 of these points (*prn*, and *al-al* in Fig. 3) using only the 3D shape information of the face. To locate the remaining seven points it employs both the 2D and the 3D information of the face. Furthermore, to locate all the points except for the tip of the nose (*prn*), the algorithm assumes that each face is in a frontal upright position, with its natural axis of bilateral symmetry roughly along the vertical dimension (*e.g.*, Fig. 4).

The overall algorithm for detecting the 10 anthropometric facial fiducial points proceeds in a sequence of cascaded steps. Each stage in the sequence utilizes fiducial point locations found in the previous stages to assist in locating the current fiducial points. The logical sequence used begins with the most reliable and easy to detect feature, proceeding to features that are less reliable and harder to detect. However, as it turns out, all facial fiducial points were quite reliable.

The sequence of stages for automatically detecting the fiducial points in the Anthroface 3D algorithm begins with the detection of the tip of the nose (*prn*). The location of this point is then employed to detect the nose width points *al-al*. These three points are then employed to detect the inner corners of the eyes (points *en-en*) and the center of the nose root (point m'). The tip of the nose (point *prn*) and nose width points (*al-al*) are also employed to detect the corners of the mouth (points *ch-ch*). Lastly, the locations of the in-

ner corners of the eyes (points *en-en*) are used to locate the outer corners of the eyes (points *ex-ex*).

A key underlying concept motivated the design of our proposed Anthroface 3D algorithm for detecting these anthropometric fiducial points. By their definition of being *fiducial* points or *landmarks*, these points have unique structural and/or textural properties that differentiate them from their surrounding regions. When human beings search for these facial fiducial points, either by visual or tactile inspection, it is these characteristics that inherently guide their search. We reasoned that the key to accurately locating these anthropometric fiducial points was to isolate their unique structural and/or textural characteristics and to search for them in an appropriately constrained region of the face. In the following sections, we describe in detail the steps that the Anthroface 3D algorithm employs to automatically detect each of the 10 anthropometric fiducial points. Note that all dimensions in the following discussion are described in millimeters (mm). The actual ratio of pixel/mm of any 3D image is related to the operational conditions of the acquisition system and can be employed to determine the corresponding dimensions in pixel units.

3.6.1 Nose Tip (prn)

In order to locate the tip of the nose (point *prn* in Fig. 3), the algorithm employs a 3D template face (Fig. 4) with a manually located nose tip. We selected this template face from the training partition (Table 2) of our database. It was a relatively holes free and symmetric face with no facial or surrounding hair, and a neutral facial expression. The algorithm registers the entire surface of every 3D face in the database to the surface of the template face using the ICP algorithm (Besl and McKay 1992). After aligning an arbitrary face to the template face, the point on its surface closest to the tip of the nose of the template face is found. This is the initial estimate (ICP estimate) of the tip of the nose.

Although the ICP estimate of the tip of the nose is not very accurate (the standard deviations of errors from the manually located nose tips for the entire database of images were $\sigma_x = 6.271$ and $\sigma_y = 9.415$ pixels), the ICP procedure served two purposes. First, it helped to transform 3D models in arbitrary poses to a frontal upright canonical pose, which was required for detecting all the other points. Second, since the ICP estimate was in the central region of all faces, it helped to limit the search for the tip of the nose in the next stage to a window of 96 mm $\times$ 96 mm about the ICP estimate.

For facial range images of the form $(x, y, z(x, y))$, the Gaussian surface curvature (K), the mean surface curvature (H), and two principal curvatures (κ_1, κ_2) can be computed from their first and second partial derivatives as (Do Carmo 1976)

$$K = \frac{z_{xx} z_{yy} - z_{xy}^2}{(1 + z_x^2 + z_y^2)^2}, \tag{6}$$

$$H = \frac{z_{xx}(1 + z_y^2) + z_{yy}(1 + x_x^2) - 2 z_x z_y z_{xy}}{(1 + z_x^2 + z_y^2)^{3/2}}, \tag{7}$$

$$\kappa_1, \kappa_2 = H \pm \sqrt{H^2 - K}, \tag{8}$$

where z_x and z_y are the first partial derivatives of $z(x, y)$ w.r.t. x and y, respectively, and z_{xx}, z_{yy} and z_{xy} are the second first partial derivatives of $z(x, y)$ w.r.t. x and y. Furthermore, the Gaussian curvature $K = \kappa_1 \kappa_2$, and the mean curvature $H = (\kappa_1 + \kappa_2)/2$. As a part of the Anthroface 3D algorithm, we computed these partial derivatives, and the Gaussian and mean curvature values for the facial range images, using a method developed by Besl (1988). The signs of the Gaussian and the mean curvature values help to identify differently shaped regions of a surface. The regions with $K > 0$ are 'elliptic', those with $K < 0$ are 'hyperbolic', and those with $K = 0$ are either 'planar' or are 'cylindrical'. For the right-handed 3D co-ordinate system defined in Fig. 4, regions of the surface with $H > 0$ are 'concave', while those with $H < 0$ are 'convex'.

Researchers in the past have noted that the sub-parts of the human face have distinct surface curvature properties (Gordon 1992; Moreno et al. 2003). We further observed that in fact, of all the regions on the facial surface, the region surrounding the tip of the nose has the highest elliptic Gaussian curvature (Fig. 5(a)), and more specifically the highest convex elliptic Gaussian curvature (Fig. 5(b)). Hence, the Anthroface 3D algorithm employs a very simple procedure, which reliably and accurately detects the tip of the nose of 3D faces. The algorithm searches for the point with the maximum elliptic Gaussian curvature within a 96 mm $\times$ 96 mm central region of each face (Fig. 5(c)), which surrounds the initial estimate of the nose tip obtained using the ICP procedure. This is the location of the final automatically located tip of the nose. For apparent reasons (Fig. 5(b)), we coined the term 'The Pinocchio Feature' for this reliable facial fiducial point.

3.6.2 Nose Width Points (al-al)

To detect the anthropometric facial fiducial points (*al-al* in Fig. 3(b)), which define the anthropometric measurement of 'nose width', the Anthroface 3D algorithm restricts its search to sub-regions of the range image of size 42 mm $\times$ 50 mm centered about the location of the automatically detected nose tip. We determined the size of this search region using the reported average and standard deviation values of the nose heights and nose widths of adult human males (Farkas 1994). On average, human males are reported

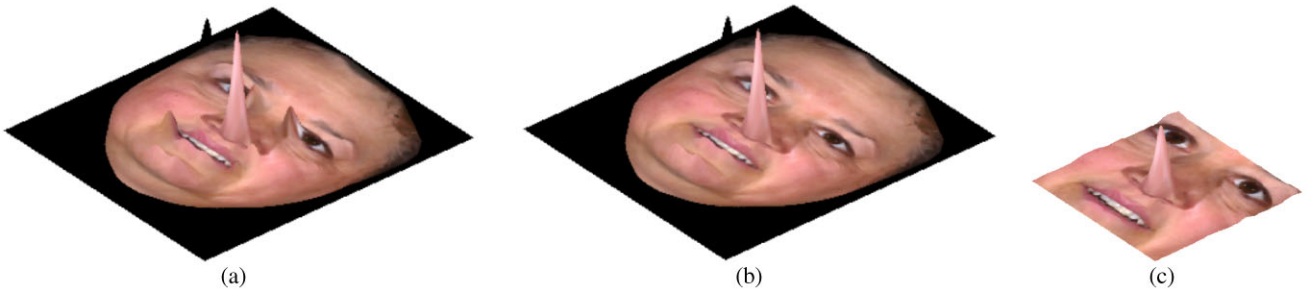


Fig. 5 The magnitudes of the Gaussian curvatures of the (**a**) elliptic, (**b**) convex elliptic regions of a facial range image, and (**c**) of its central 96 mm × 96 mm region. In each image, the Gaussian curvature has been plotted as a 3D surface with the facial texture warped onto it

Fig. 6 (**a**) The edges detected with the LOG edge detector in sub-regions of the facial range images centered about the detected nose tip, and (**b**) critical points along the nose boundary (+) and the tip of the nose (○)

to have wider (μ = 35 mm, and σ = 2.5 mm) and taller (μ = 53 mm with σ = 3.4 mm) noses than females. Hence, to account for variations in the human population, we fixed the width of the search region for points *al-al* about the tip of the nose at the $\mu + 6\sigma$ value of the nose width of human males. Similarly, we fixed the height of the search region at the $0.6 \times (\mu + 6\sigma)$ value of the height of noses of human males.

Within this search region about the tip of the nose, the Anthroface 3D algorithm detects edges on the facial range images using a Laplacian of Gaussian edge detector, with $\sigma = 7$ pixels. Since the human nose is a distinct protrusion in the facial surface, we observed that the left and right boundaries of the nose were always clearly delineated in the edge maps of all faces (Fig. 6(a)). From this edge map, the algorithm isolates the left and right boundaries of the nose by traversing outwards horizontally in both directions from the tip of the nose, and by retaining the first curves encountered.

The algorithm then detects all the 'critical' points (points of high curvature) with negative curvature values (Rodriguez and Aggarwal 1990), which are present along the nasal boundary contours. It traverses the nasal boundary curves in a clockwise direction. For *all* faces in our database, the points *al-al* were among these critical points (shown in Fig. 6(b)). From among the critical points, the algorithm isolates the points *al-al*, by searching for the leftmost and rightmost critical points, that are closest to the tip of the nose along the vertical direction.

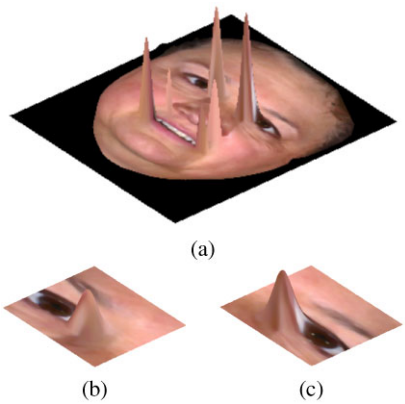


Fig. 7 (**a**) The magnitude of the Gaussian curvature of the concave elliptic regions of a facial range image, (**b**) that of a smaller region defined to search for the right eye's inner corner, and (**c**) that of a smaller region defined to search for the left eye's inner corner. In each image the surface curvature is plotted as a 3D surface with the facial texture warped onto it

3.6.3 Inner Eye Corners (en-en) and Center of Nose Root (m′)

To automatically locate the inner corners of the eyes (points *en-en*) we observed that for *all* faces, these points were located in regions of the face that were distinctly concave elliptic (Fig. 7(a)). The Anthroface 3D algorithm locates the peaks of Gaussian curvature of these two regions, as initial estimates (curvature estimates) of the locations of the inner corners of the eyes. In order to define the search regions for these peaks, we employed the locations of automatically detected points *prn*, *al-al*, the location of the highest vertical point of each 3D model (v), and knowledge of the established horizontal and vertical proportions of a normal adult human face (Farkas 1987). For an average adult, the vertical distance between the inner corners of the eyes and the tip of the nose is ~0.3803 times the vertical distance between the top of the head and the tip of the nose (Farkas 1987). To account for variations in human populations, we fixed the upper limits of the search regions to $(prn_y + 0.3803 \times 1.5 \times |prn_y - v_y|)$, and the lower limits to

$(prn_y + 0.3803 \times 0.33 \times |prn_y - v_y|)$, where prn_y is the vertical co-ordinate of the tip of the nose and v_y is the vertical co-ordinate of the highest point of the 3D model.

To determine the horizontal limits of the two search regions, we employed the fact that for an average face, the ratio of the horizontal distance between the inner corners of the eyes to nose width (the distance between points *al-al*) is unity (Farkas 1987). Thus, Anthroface 3D searches between prn_x, the horizontal co-ordinate of the tip of the nose and $(al_{x,left} + 0.5 \times |al_{x,left} - al_{x,right}|)$ for the curvature estimate of the inner corner of the subject's left eye. Similarly, for the curvature estimate of the inner corner of subject's right eye, the algorithm searches between the horizontal limits that are defined on one side by $(al_{x,right} - 0.5 \times |al_{x,left} - al_{x,right}|)$ and prn_x on the other side. These search regions for the two inner eye corners are shown in Fig. 7(b) and (c), respectively. Within these two regions, the algorithm determines the locations of the points with the highest Gaussian curvatures as the curvature estimates for the inner corners of the eyes.

The Anthroface 3D algorithm obtains the final positions of points *en-en* by further searching within a region of size 20 mm × 20 mm about the curvature estimates for these points, using a recently developed 2D+3D EBGM algorithm (Jahanbin et al. 2008). Briefly, as a part of the 2D+3D EBGM algorithm, the algorithm locates the fiducial points of interest manually on a set of 2D and 3D example images. We selected 68 images of 12 subjects from the training partition of our database, and 21 images of 13 subjects from the Remaining partition (Table 2), as example images for the 2D+3D EBGM algorithm. In this set of images, we included faces with various expressions, *e.g.*, open/closed eyes, and neutral/smiling with open/closed mouths. Forty 2D and forty 3D Gabor coefficients (at 5 scales and 8 orientations) are computed for the manually located fiducial points of all example faces. The Gabor filters that we employed corresponded to a carefully designed filter bank (Bovik et al. 1990) that is widely used in the area of 2D face detection and recognition (Wiskott et al. 1997; Wang et al. 2002; Wang and Chua 2005). For a face with unknown fiducial points, the same 2D and 3D Gabor coefficients are computed for every point within the defined search window as were computed for the example faces. The point, within this search window, that has Gabor coefficients most similar to the Gabor coefficients of any example image, is regarded as the final location of the detected fiducial point. Note that the 2D+3D EBGM algorithm could have been applied directly to the larger search windows (Fig. 7(b) and (c)) that were employed to find the initial curvature estimates of the inner corners of the eyes. However, such an approach produces a significantly large number of false positives at the locations of the inner corners of the eyebrows, which have textural characteristics similar to the inner corners of the eyes. The center of the root of the nose (point m' in Fig. 3) was located at the algebraic mean of the positions of the automatically detected inner corners of the eyes (*en-en*).

3.6.4 Outer Eye Corners (ex-ex)

In order to automatically locate the outer corners of the eyes (points *ex-ex* in Fig. 3), Anthroface 3D employs the positions of the automatically detected inner corners of the eyes (*en-en*) as reference points. For an average human adult, the distance between the inner and the outer corner of an eye is approximately equal to the distance between the inner corners of the two eyes (Farkas 1987). Using this anthropometric information, Anthroface 3D computes the initial estimate for the position of the outer corner of a subject's left eye as $(en_{x,left} + |en_{x,left} - en_{x,right}|, (en_{y,left} + en_{y,right})/2)$, and that of the outer corner of the right eye as $(en_{x,right} - |en_{x,left} - en_{x,right}|, (en_{y,left} + en_{y,right})/2)$. The algorithm then uses the 2D EBGM algorithm to search within a rectangular window of size 20 mm × 34 mm about the initial estimates of these points to obtain the final locations of the points *ex-ex*. The set of example images that were employed for this 2D EBGM algorithm were the same as those that were employed for detecting the point *en-en*. Note, that since the outer corners of the eyes do not have distinct surface curvature characteristics, we used 2D EBGM instead of 2D+3D EBGM.

3.6.5 Mouth Corners (ch-ch)

We examined the curvature of facial surface regions located below the nose, and observed that for *all* faces, the outer corners of the mouth were distinct concavities, *i.e.*, regions of high positive mean curvature (H) (Fig. 8(a)). Hence, the peaks of mean curvature (H) in this region served as the initial estimates (curvature estimates) for the locations of the points *ch-ch*. To find these peaks of mean curvature, we defined appropriate search regions as follows.

We observed that for all faces (including those with beards, mustaches, and arbitrary facial expressions) regions of the upper lip and lower lip have elliptic Gaussian curvature (Fig. 8(b)). By detecting these upper and lower lip regions below the tip of the nose, Anthroface 3D determines the vertical limits of the search regions for the curvature estimates of the corners of the mouth. Furthermore, it employs the locations of the automatically detected points *al-al* to horizontally constrain these search regions. For ch_{left}, Anthroface 3D searches to the left of al_{left} between the horizontal positions $al_{x,left}$ and $(al_{x,left} + 0.7 \times |al_{x,left} - al_{x,right}|)$, and for ch_{right} searches to the right of al_{right} between $al_{x,right}$ and $(al_{x,right} - 0.7 \times |al_{x,left} - al_{x,right}|)$.

Finally, the algorithm refines the positions for the corners of the mouth by searching in a window of size 11 mm × 30 mm about the initial curvature estimates of these points using the 2D+3D EBGM algorithm. We employed the same

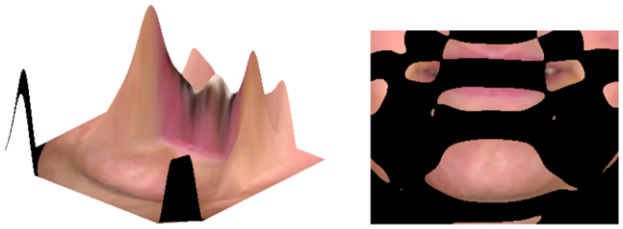


Fig. 8 (**a**) The mean curvature (H) of the mouth region of an example face plotted as a 3D surface, with the facial texture warped onto it. Notable are the distinct peaks at the corners of the mouth. (**b**) *Regions* below the nose with elliptic Gaussian curvature (non-black *regions*) of an example face. Notable are the regions of the upper and the lower lip

set of example images that were employed to detect the points *en-en* and the points *ex-ex*. This second stage, which refines the positions of the corners of the mouth, eliminated a significant number of errors that resulted in the first stage for faces with open mouths.

3.7 Automatic Anthroface 3D Algorithm

Lastly, we developed the completely automatic version of the Anthroface 3D recognition algorithm using 3D Euclidean and 3D geodesic anthropometric facial distances between all pairs of the 10 automatically located facial fiducial points and the classification methodology described earlier (Sect. 3.3).

4 Performance Evaluation

4.1 Database

We employed the publicly available Texas 3D Face Recognition Database (http://live.ece.utexas.edu/research/texas3dfr) to evaluate the performance of all algorithms (Gupta et al. 2010). This database contains 1149 2D and 3D image pairs of 118 adult human subjects. The number of images of each subject varies from 1 per subject to 89 per subject. The subjects' ages range from ~22–75 years. The database contains images of both males and females from the major ethnic groups of Caucasians, Africans, Asians, East Indians and Hispanics. The facial expressions included are neutral and expressive, *e.g.*, smiling or talking faces with open/closed mouths and/or closed eyes (*e.g.* Fig. 9). No subject is wearing a hat or eye-glasses.

The images in the Texas 3D Face Recognition Database were acquired using a stereo imaging system manufactured by 3Q Technologies Ltd. (Atlanta, GA). The final representation of each face in the database is a pair of range and color images in the canonical frontal pose that are perfectly aligned to each other (*e.g.*, Fig. 1). The images are of size

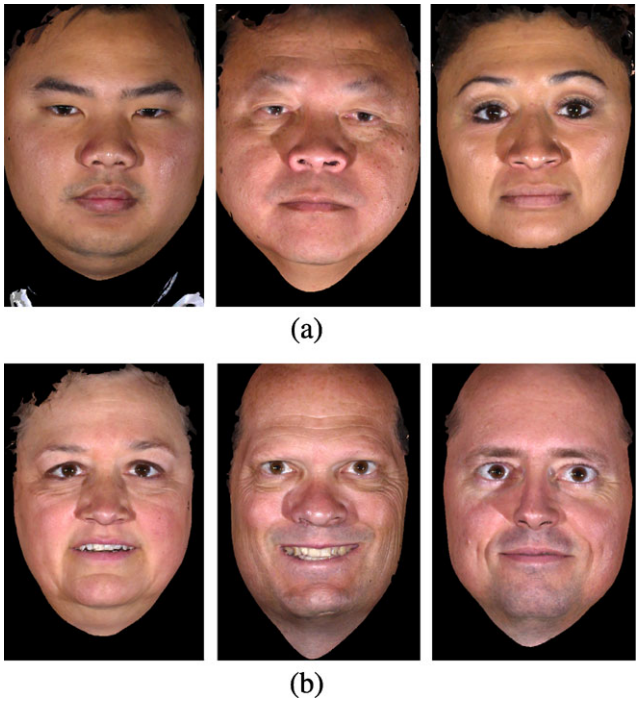


Fig. 9 Color images of faces in the Texas 3D Face Recognition Database in (**a**) neutral, and (**b**) expressive modes

751×501 pixels with a resolution of 0.32 mm along the x, y, and z dimensions. The locations of a large number (25) of manually located anthropometric fiducial points (Fig. 1) that we identified in our analyses are also available for each 2D and 3D face in Texas 3D Face Recognition Database. Details of image acquisition, pose normalization, and preprocessing steps that were applied to generate the final faces in the Texas 3D Face Recognition Database can be found in Gupta et al. (2010).

The Texas 3D Face Recognition Database is closest in terms of variability in facial expression, and facial pose that it contains to the widely used FRGC data set. Both these databases contain mostly frontal faces with neutral or mild facial expressions. It is has also been previously noted that in everyday life most people are likely to exhibit spontaneous emotions in a light (low) intensity without exaggerated appearance and that displays of intense emotions rarely happen (Douglas-Cowie et al. 2000). Hence, it is reasonable to assume that both these databases contain adequate variability to represent a real-world face recognition scenario involving co-operative human subjects in a partially controlled environment. However, the FRGC data set, which although is larger than the Texas 3D Face Recognition database, was unsuitable for our current analyses and could not be employed. This was because no manually annotated facial fiducial points are available for the FRGC data set. Furthermore, the pairs of color and range images in the FRGC database were acquired a few seconds apart, and hence are not perfectly aligned (Phillips et al. 2005). For the same reason,

Table 2 A summary of the data partitions employed for developing 3D face recognition algorithms

Partition		Subjects	No. of Images		
			Neutral	Expressive	Total
Training		12	228	132	360
Test	Gallery	105	105	0	105
	Probes	95	480	183	663
Remaining		13	0	21	21

certain pairs of range and color images in the FRGC 2005 database have inconsistent facial expressions or distortions (Maurer et al. 2005). Using the FRGC database for the current Anthroface 3D analyses would have required manually locating the 25 facial fiducial points on both the ∼5000 color and ∼5000 range images of the database, since the two are not aligned. This is further complicated by the fact that many anthropometric fiducial points (*e.g.* the corners of the eyes), which have distinct textural and properties and not necessarily distinct structural properties, are non-trivial to locate manually on facial range images.

4.2 Data Partitions

For the purposes of developing the Anthroface 3D face recognition algorithm, we partitioned the Texas 3D Face Recognition database into a training data set and a test data set (Table 2). The training data set contained 360 randomly selected images of 12 subjects (30 images per subject) in neutral or expressive modes. For all the 3D face recognition algorithms that we implemented, steps such as automatic facial fiducial point detection, classifier feature selection and classifier optimization were performed using the training data set only. The trained classifier was evaluated on the independent test data set, which did not overlap with the training data set.

The test data set included 768 images of 105 subjects. This test set was further partitioned into a gallery set and a probe set. Consistent with the evaluation protocol of the FRVT 2002 (Phillips et al. 2003) and FRGC 2005 (Phillips et al. 2005), the gallery set contained one range image each of 105 subjects with a neutral facial expression. The probe set contained another 663 images of 95 of the gallery subjects with a neutral or an arbitrary facial expression. In the probe set, the number of images of each subject varied from 1 to 55. Furthermore, in accordance with the widely accepted 'closed universe' model for the evaluation of face recognition algorithms (Phillips et al. 2003), every subject in the probe data set was represented in the gallery data set.

After partitioning the entire database of 1149 images into the training and test data sets, 21 images of 13 subjects remained (hereafter the 'Remaining' set). All these were of faces with an arbitrary facial expression. We employed this Remaining set of images along with 68 images of 12 subjects from the training data set as example images in the 2D/2D+3D EBGM algorithms for detecting facial fiducial points.

4.3 Algorithms Performance

4.3.1 Fiducial Point Detection

In order to evaluate the performance of our proposed automatic facial fiducial point detection algorithm, we regarded the positions of the manually located points as the 'ground truth'. Across 1060 facial images, we computed the standard deviations (σ_x and σ_y) of the positional errors of the automatically detected fiducial points from their manual locations along the x and y dimensions, respectively. This set of 1060 images included all the images in the Texas 3D Face Recognition Database except the 89 images, which were employed as 'example images' for the 2D+3D EBGM fiducial point detection algorithms.

4.3.2 Face Recognition

We evaluated the verification performance of all 3D face recognition algorithms using the Receiver Operating Characteristic (ROC) methodology (Egan 1975), and observed the values of the Equal Error Rates (EER) and the Areas Under the ROC Curves (AUC). The identification performance of the algorithms was evaluated using Cumulative Match Characteristic (CMC) curves, and the rank 1 Recognition Rates (RR) were observed. Statistical 95% confidence intervals for the EER, AUC, and the rank 1 RR values were obtained empirically using bootstrap sampling. All performance statistics were observed separately for neutral faces, for expressive faces, and for all faces in the probe data set.

4.4 Benchmark Algorithms

We also compared the performance of the Anthroface 3D algorithm to three existing state-of-the-art automatic 3D face recognition algorithms by implementing them on the Texas 3D Face Recognition Database. The algorithms included the eigensurfaces algorithm (Chang et al. 2005), the fishersurfaces algorithm (Heseltine and Austin 2004; BenAbdelkader and Griffin 2005), and a 3D face recognition algorithm based on the ICP procedure (Lu et al. 2006; Russ et al. 2005). For the eigensurfaces and the fishersurfaces algorithms, we employed sub-sections of the facial range images between the pixels 147 and 553 along the vertical, and 38 and 478 along the horizontal (*e.g.*, Fig. 10). These limits corresponded to the extrema of the uppermost, bottommost, leftmost, and rightmost co-ordinates, respectively, of the 25 manually located anthropometric fiducial

Fig. 10 The sub-regions of the facial range images that were employed for the eigensurfaces and fishersurfaces 3D face recognition algorithms

points across all faces in the database. In setting these limits to the extrema of all faces, we ensured that the main facial features were not excluded for any face.

Both the eigensurfaces and the fishersurfaces algorithms were trained and tested on exactly the same data sets, which were employed to train and test the Anthroface 3D algorithm. For the eigensurfaces algorithm, we learned 69 eigen directions that accounted for 99% of the variance of the data, and linearly projected all faces in the test data set onto these eigen directions. The final metric for the comparison of 3D facial surfaces in the eigen sub-space was the L_1 norm. For the fishersurfaces algorithm, we first reduced the dimensionality of the range images to 348 using PCA. This was done to ensure that the within-class scatter matrix (4) employed in the LDA computations was non-singular. We then learned 11 LDA directions from the 348 PCA features, projected all faces in the test data set onto these LDA directions, and compared them using the L_2 norm. The 3D face recognition algorithm based on ICP did not require training. Hence, we implemented it only on the test data set and compared all faces in the probe set to all faces in the gallery set, after registering pairs of facial surfaces using the ICP algorithm. The pairs of registered facial surfaces were compared using the partial Hausdorff distance metric (Huttenlocher et al. 1993).

5 Experimental Results and Discussion

5.1 Manual Anthroface 3D

The equal error rates, the areas under the ROC curves, and the rank 1 recognition rates of the Anthroface 3D algorithm, which was based on 25 manually detected anthropometric fiducial points (Fig. 1), and for the algorithm which was based on 25 arbitrary facial points (Fig. 2), are presented in Table 3(a), (b) and (c), respectively. Table 3(a), (b) and (c) also contain the same performance statistics for the benchmark eigensurfaces, fishersurfaces, and ICP algorithms, which were implemented on the Texas 3D Face Recognition Database. The CMC and the ROC curves for all these algorithms are presented in Figs. 11 and 12.

For all faces in the probe data set, the proposed Anthroface 3D algorithm, which was based on 25 manually located

Table 3 The observed (**a**) EER, (**b**) AUC, and (**c**) rank 1 RR values and their 95% confidence intervals for the eigensurfaces, fishersurfaces, ICP, and the Anthroface 3D algorithms based on 25 manually located anthropometric points and 25 arbitrary facial points

EER [Confidence Interval] %

Algorithm	Neutral	Expressive	All
Eigensurfaces	24.0 [21.3 26.3]	23.6 [19.2 26.2]	24.0 [21.8 26.6]
Fishersurfaces	8.11 [6.39 10.9]	3.60 [2.00 6.43]	6.69 [5.27 8.16]
ICP	7.97 [6.85 9.95]	9.92 [6.71 14.7]	9.03 [7.67 10.2]
Anthroface 3D (25 anthro)	0.84 [0.53 1.14]	1.58 [0.64 2.67]	1.00 [0.64 1.45]
Anthroface 3D (25 arbitrary)	8.78 [6.58 10.9]	5.10 [3.37 8.61]	7.65 [6.16 10.1]

(a)

AUC [Confidence Interval] $\times 10^{-2}$

Algorithms	Neutral	Expressive	All
Eigensurfaces	16.7 [14.1 18.8]	14.9 [11.9 18.0]	16.3 [14.8 18.7]
Fishersurfaces	2.88 [2.12 3.71]	1.32 [0.32 2.49]	2.40 [1.79 2.96]
ICP	2.97 [2.16 4.12]	4.39 [2.18 7.80]	3.44 [2.56 4.27]
Anthroface 3D (25 anthro)	0.07 [0.03 0.12]	0.08 [0.04 0.12]	0.08 [0.04 0.12]
Anthroface 3D (25 arbitrary)	3.0 [2.23 3.95]	2.08 [0.68 4.61]	2.70 [2.03 3.65]

(b)

Rank 1 RR [Confidence Interval] %

Algorithms	Neutral	Expressive	All
Eigensurfaces	58.1 [54.0 62.7]	52.5 [45.4 60.1]	56.6 [52.9 60.2]
Fishersurfaces	91.7 [89.4 94.0]	95.1 [91.8 97.8]	92.6 [90.6 94.4]
ICP	88.5 [85.6 91.5]	86.3 [80.9 91.0]	87.9 [85.5 90.2]
Anthroface 3D (25 anthro)	98.8 [97.7 99.6]	95.6 [92.4 98.4]	97.9 [96.8 98.9]
Anthroface 3D (25 arbitrary)	86.0 [82.9 89.0]	91.3 [87.4 95.1]	87.5 [84.9 89.9]

(c)

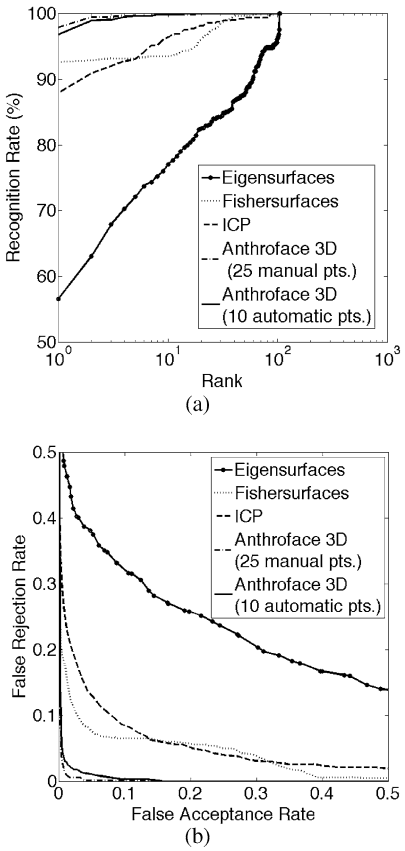


Fig. 11 (**a**) The semi-log CMC curves, and (**b**) the ROC curves for the benchmark 3D face recognition algorithms, the Anthroface 3D algorithm that employed 25 manually located points, and the Anthroface 3D algorithm that employed 10 automatically located points

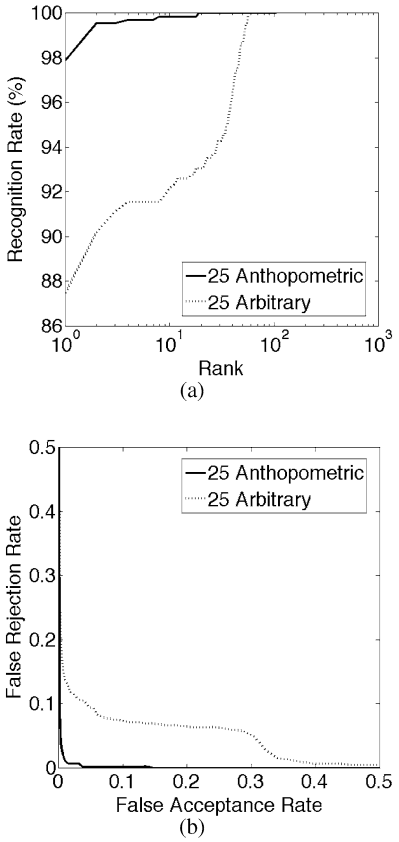


Fig. 12 (**a**) The semi-log CMC curves, and (**b**) ROC curves for the Anthroface 3D algorithm based on 25 manually located facial fiducial points and the algorithm based on 25 arbitrary facial points

anthropometric fiducial points, performed well ($EER = 1\%$, $AUC = 0.0008$, and rank 1 $RR = 97.9\%$). While it may be non-trivial to automatically and accurately locate all the 25 anthropometric facial fiducial points that were employed in this manual Anthroface 3D algorithm, nonetheless, analyzing the performance of this algorithm serves two purposes. Firstly, the performance of this manual version of the Anthroface 3D algorithm can be regarded as an upper bound on the expected performance of any automated version of such an algorithm. Secondly, its high performance is indicative of the potential of 3D face recognition algorithms that incorporate knowledge about the structural diversity and statistical distribution of anthropometric measurements of the human face. This latter conclusion is further supported by the fact that, while the Anthroface 3D algorithm, which employed distances between the carefully selected 25 anthropometric fiducial points, performed well, its counterpart, wherein facial distances between 25 arbitrary facial points were employed, performed significantly poorly ($EER = 7.65\%$, $AUC = 0.027$, and rank 1 $RR = 87.5\%$) in comparison (Table 3 and Fig. 12).

The Anthroface 3D algorithm, which was based on 25 manually located facial fiducial points, also performed significantly better than the three existing holistic benchmark eigensurfaces ($EER = 24.0\%$, $AUC = 0.16$, and rank 1 $RR = 56.6\%$), fishersurfaces ($EER = 6.69\%$, $AUC = 0.024$, and rank 1 $RR = 92.6\%$), and ICP algorithms ($EER = 9.0\%$, $AUC = 0.034$, and rank 1 $RR = 87.9\%$) with nearly an order of magnitude smaller EER and AUC values (Table 3, Fig. 11). These results establish that the manual Anthroface 3D algorithm is more accurate at the task of 3D face recognition than the existing benchmark 3D face recognition algorithms of eigensurfaces, fishersurfaces and ICP. Furthermore, the results also point towards the superiority of local feature based 3D face recognition algorithms relative to holistic techniques. Another interesting observation in this regard is the fact that the verification performance of even the local feature based 3D face recognition algorithm, wherein, facial distances between 25 arbitrary facial points were employed ($AUC = 0.027$ with a confidence interval of [0.020 0.036]), was statistically equivalent to the best performing holistic fishersurfaces 3D face recognition algorithm ($AUC = 0.024$ with a confidence interval of [0.018 0.030]). Interestingly, some studies in the cognitive sciences

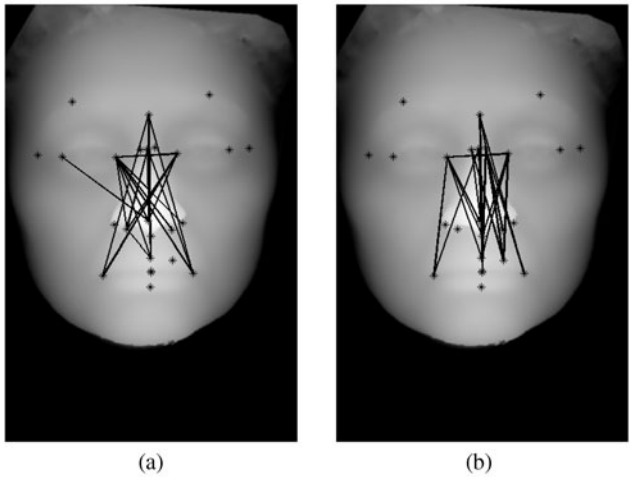


Fig. 13 The 20 most discriminatory facial (**a**) Euclidean, and (**b**) geodesic distance features. The geodesic distances are symbolically depicted by straight lines. In reality, they are along the surface of the face

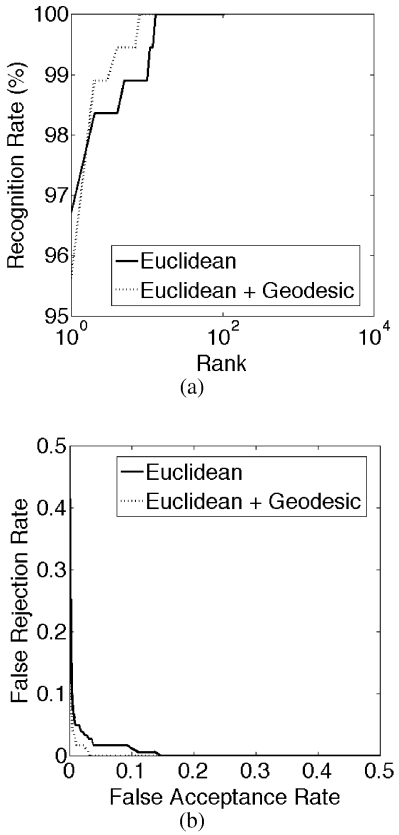


Fig. 14 (**a**) The semi-log CMC curves, and (**b**) ROC curves for the Anthroface 3D algorithm that employed only Euclidean distances, and the Anthroface 3D algorithm that employed both the Euclidean and the geodesic distances for expressive faces only

suggest that similar to Anthroface 3D, which is based on distances between facial fiducial points, humans may also be processing facial information using relational information between parts of the face (Tanaka and Farah 2003). It is further believed that human beings may be acquiring this information by means of sequential eye movements between the different facial features (Henderson et al. 2005).

We attempted to gain further insights into the nature of the discriminatory structural information, which was contained in the variable anthropometric facial proportions that we initially selected (Table 1). We separately ranked the anthropometric Euclidean and geodesic distances between the 25 manually located facial fiducial points in descending order of their individual Fisher's ratio (Duda et al. 2001) values. The 20 most discriminatory facial Euclidean and geodesic distances are presented in Fig. 13. Interestingly, these distances were predominantly associated with the nasal region of the face, as were many of the variable anthropometric facial proportions (Table 1). Furthermore, 17 ($O10$, $O12$, $N1$, $N6$, $N7$, $N8$, $N15$, $N16$, $N30$, $N31$, $N33$, $L1$, $L4$, $L5$, $L6$, $L7$, and $L14$) of the 23 facial proportions that we selected (Sect. 3.1) have also been reported to be significantly different for the two genders by Farkas (1981), and one ($N7$) has been reported to be significantly different for various ethnic groups (Farkas et al. 1985). All these factors very likely contribute to the success of the Anthroface 3D algorithm.

Next, we compared the performance of the proposed Anthroface 3D algorithm, which employed both Euclidean and geodesic distances between the 25 anthropometric facial fiducial points to a similar algorithm that employed *only* Euclidean distances between the same set of fiducial points. We observed that for expressive faces, the verification performance of the algorithm that employed both Euclidean and geodesic distances ($AUC = 0.0008$, $CI = [0.0004\ 0.0012]$), was significantly better than the performance of the algorithm that employed only anthropometric Euclidean distances ($AUC = 0.0015$, $CI = [0.0009\ 0.0023]$) (Fig. 14(b)). Similarly, for expressive faces, the recognition rates of the algorithm that was based on both the Euclidean and geodesic facial anthropometric distances were also generally higher than those of the algorithm that was based on *only* Euclidean distances (Fig. 14(a)) This suggests that facial geodesic distances may be useful for expression invariant 3D face recognition and further strengthens Bronstein et al.'s (2005) proposition that different facial expressions may be modeled as isometric deformations of the facial surface.

The performance of the Anthroface 3D algorithm, which was based on the reduced set of 10 *manually* located anthropometric fiducial points (Fig. 3) is presented in Table 4. For all faces in the probe set, this algorithm resulted in marginally lower verification performance ($EER = 1.68\%$ and $AUC = 0.0014$) than that of the algorithm, which was based on 25 manually located points ($EER = 1.00\%$ and $AUC = 0.0008$). However, the recognition performance of this Anthroface 3D algorithm (rank 1 $RR = 98\%$) was not statistically significantly different from that of the Anthro-

Table 4 The observed (**a**) EER, (**b**) AUC, and (**c**) rank 1 RR values and their 95% confidence intervals for the Anthroface 3D algorithms based on 25 manually located Fiducial points, 10 manually located fiducial points, and 10 automatically located fiducial points

EER [Confidence interval] %			
Algorithms	Neutral	Expressive	All
Anthroface 3D (25 manual)	0.84 [0.53 1.14]	1.58 [0.64 2.67]	1.00 [0.64 1.45]
Anthroface 3D (10 manual)	1.10 [0.65 1.96]	2.34 [1.01 3.10]	1.68 [1.10 2.24]
Anthroface 3D (10 automatic)	1.65 [1.11 2.28]	2.81 [1.27 4.30]	1.98 [1.37 2.88]

(a)

AUC [Confidence interval] $\times 10^{-2}$			
Algorithms	Neutral	Expressive	All
Anthroface 3D (25 manual)	0.07 [0.03 0.12]	0.08 [0.04 0.12]	0.08 [0.04 0.12]
Anthroface 3D (10 manual)	0.12 [0.04 0.30]	0.18 [0.05 0.30]	0.14 [0.07 0.25]
Anthroface 3D (10 automatic)	0.14 [0.08 0.23]	0.25 [0.11 0.42]	0.18 [0.11 0.28]

(b)

Rank 1 RR [Confidence interval] %			
Algorithms	Neutral	Expressive	All
Anthroface 3D (25 manual)	98.8 [97.8 99.6]	95.6 [92.4 98.4]	97.9 [96.8 98.9]
Anthroface 3D (10 manual)	98.8 [97.7 99.6]	96.2 [93.4 98.9]	98.0 [97.0 98.9]
Anthroface 3D (10 automatic)	97.3 [95.8 98.5]	95.6 [92.4 98.4]	96.8 [95.3 98.0]

(c)

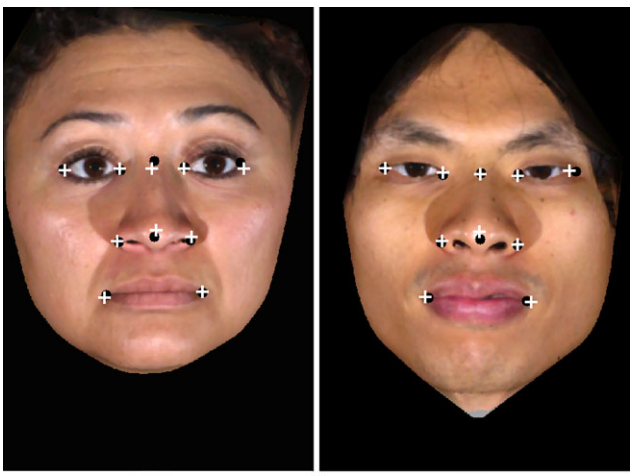
face 3D algorithm, which was based on 25 manually located fiducial points (rank 1 $RR = 97.9\%$) for this database. Note, that since the ability to distinguish the performance of different algorithms depends on the size of the database, a larger database may perhaps be able to distinguish between the performances of these more unequivocally.

5.2 Automatic Anthroface 3D

The standard deviations of the positional errors of the 10 automatically located anthropometric facial fiducial points (Fig. 3) from their respective manual 'ground truth' locations are presented in Table 5. It can be observed that all

Table 5 The standard deviations of the positional errors of the 10 automatically located anthropometric facial fiducial points from their manual locations in mm

Fiducial Point	Error σ_x	Error σ_y	Radial Error σ
prn	1.045	1.680	1.978
al_{left}	0.721	1.655	1.805
al_{right}	0.798	1.646	1.829
en_{left}	1.488	1.245	1.940
en_{right}	1.354	1.344	1.908
m'	1.355	1.811	2.261
ex_{left}	1.795	1.286	2.208
ex_{right}	2.126	1.384	2.537
ch_{left}	1.948	0.933	2.160
ch_{right}	1.976	1.045	2.235



(a) (b)

Fig. 15 Example images showing the automatically detected anthropometric facial fiducial points (+) and their manually located positions (•)

the 10 anthropometric facial fiducial points were detected fairly accurately (e.g., Fig. 15). The radial standard deviation of error for each of the 10 fiducial points was less than 2.54 mm. The average radial standard deviation across all the 10 fiducial points was 2.09 mm.

The nose width points (*al-al*) were located most reliably (within 2 mm of their manual locations), followed by the inner corners of the eyes (*en-en*). The tip of the nose was also located reliably for all faces in the database as the point in the central region of the face with the highest convex elliptic Gaussian curvature. Note that, this characteristic structural property of the fiducial point *prn* also corresponds well with the intuitive definition of the *tip* of the nose. The outer corners of the eyes (*ex-ex*) were detected least reliably (radial errors $\sigma_{ex_{left}} = 2.208$ mm and $\sigma_{ex_{right}} = 2.537$ mm), preceded only by the corners of the mouth (*ch-ch*) with radial errors $\sigma_{ch_{left}} = 2.160$ mm and $\sigma_{ch_{right}} = 2.235$ mm. A majority of

the false positive detections for the corners of the eyes were located at the edges of the irises. This was not surprising since the edges of the irises can have textural properties similar to the corners of the eyes. We also observed that many of the false positive detections for the mouth corners occurred on smiling faces.

Recall that in order to automatically detect the tip of the nose (*prn*), the inner corners of the eyes (*en-en*), the outer corners of the eyes (*ex-ex*), and the corners of the mouth (*ch-ch*) we employed two stages. For each point, we obtained an initial estimate of the point's location using a first detection stage and then searched within a window surrounding this initial estimated location for the final location of the point. We observed that for these points, their respective second detection stages helped to considerably reduce the errors encountered during their first detection stages. It is also instructive to note that for the points *en-en*, *ex-ex* and *ch-ch* a combination of information from the 2D and 3D images resulted in the best overall detection performance, indicating that only the structural properties of these points may not be adequately discriminatory for detecting them accurately.

Lastly, we present the performance of the completely automatic Anthroface 3D algorithm, which employed facial anthropometric distances between these 10 *automatically* detected fiducial points, is presented in Table 4 and Fig. 11. For all faces in the probe data set, the verification performance ($EER = 1.98\%$ and $AUC = 0.0018$) of this algorithm was not statistically different from that of the Anthroface 3D algorithm, which was based on the same 10 *manually* detected fiducial points ($EER = 1.68\%$ and $AUC = 0.0014$). The recognition performance of these two algorithms was also similar (Table 4(c)). Furthermore, the recognition performances of the two Anthroface 3D algorithms, which were based on 10 automatically detected points and on 25 manually detection fiducial points, respectively, were not statistically different (Table 4(c) and Fig. 11(a)). These results suggest that the Anthroface 3D algorithm may be robust to facial fiducial point detection errors with a radial standard deviation (σ) of less than 2.45 mm. Note also that this completely automatic Anthroface 3D algorithm, which was based on 10 automatically detected facial fiducial points, also performed significantly better than the benchmark eigensurfaces, fishersurfaces, and ICP algorithms (Tables 3 and 4).

6 Conclusion

In this paper, we proposed a novel Anthroface 3D recognition algorithm, which presents a number of unique contributions and interesting insights to the field of 3D face recognition and 3D facial processing. The first of which is the introduction of a sound anthropometric perspective for the field of 3D face recognition. We presented a novel way of thinking about the 3D face recognition problem. Rather than employing general purpose pattern recognition algorithms (*e.g.*, PCA, LDA) or an *ad hoc* set of 3D facial features for the task of 3D face recognition, it is for the first time that we have presented a systematic approach of employing domain specific knowledge about the structural diversity of faces for designing successful 3D face recognition algorithms. We have presented a practical method for isolating this information from the scientific discipline of facial anthropometry, and have also developed and demonstrated the successful and robust Anthroface 3D recognition algorithm that employs this information effectively. On a large database of 3D images, we have clearly demonstrated the statistically superior performance of our proposed local feature based algorithm in comparison to a number of the existing benchmark holistic 3D face recognition algorithms.

The second unique novel contribution of this work is the identification of the subset of 10 anthropometric facial points from among the set of 25 points that are most relevant for the task of 3D face recognition. Again, it is for the first time in the field if 3D face recognition that we have also clearly demonstrated by means of rigorous statistical analysis on a large facial database, that the facial distances between these 10 anthropometric facial fiducial points result in statistically identical face recognition performance as the distances between the 25 anthropometric facial fiducial points. This is an important result in the field of 3D face recognition and has the potential to positively influence generations of future local feature based 3D face recognition algorithms. As further research in the largely unexplored area of local feature based 3D face recognition progresses, this result is likely to be significant in providing a sound scientific basis for selecting this optimal set of 10 facial anthropometric fiducial points that we identified for local feature based algorithms, and especially for 3D face recognition algorithms that are based on facial graphs.

The third unique novel contribution of this work is the demonstration that the fusion of carefully selected 2D and 3D information that embodies characteristic discriminatory textural and structural properties of specific facial landmarks, and that the employment of the established anthropometric knowledge about the facial points' positions in relation to the other points, can significantly improve the detection of the fiducial points and, as a consequence, the performance of the recognition system. We envision the highly accurate fiducial point localization algorithms that we have developed to have potential applications not only in the field of 3D face recognition, but also in a number of related areas that require facial anthropometric measurements, including facial surgical planning and computer graphics. Another unique feature of our analyses is that unlike many of the previously reported studies of automatic facial fiducial points

detection, we have systematically assessed and reported the performance of our localization algorithms against manually detected points. Furthermore, we have empirically demonstrated that these localization errors do not adversely affect the recognition performance of the proposed Anthroface 3D algorithm.

In the future, a number of areas of this research could be expanded upon. The fiducial points detection algorithm could be extended to include more than the 10 anthropometric fiducial points that were detected here. Currently the Anthroface 3D recognition algorithm has been evaluated only for nearly frontal faces, with neutral or mild facial expression and/or occlusions due to hair that are presented in the Texas 3D Face Recognition database. In the future, it may be interesting to investigate the performance of the Anthroface 3D on data sets including the Binghamton University 3D Facial Expression Database (Yin et al. 2006) or the Bosphorus databases (Savran et al. 2008), which contain lager pose variations, more extreme facial expressions, and more severe occlusions than those present the Texas 3D Face Recognition database. Investigations with these data sets would require appropriate manual annotations of facial landmarks on the images of these data sets. Investigating the effectiveness of the structural information contained in localized regions about the anthropometric facial fiducial points for the task of 3D face recognition and developing effective techniques for incorporating the learned knowledge into the proposed Anthroface 3D algorithm could be a logical next step. Lastly, numerous studies in the past have demonstrated that a combination of the 2D and 3D imaging modalities for face recognition results in superior performance relative to either of them individually (Mian et al. 2007). Hence, a natural extension of this current work would be to investigate techniques of incorporating discriminatory 2D facial features into the recognition stage of the Anthroface 3D algorithm, which currently only uses 3D facial features.

Acknowledgements We gratefully acknowledge the support of Dr. Kenneth R. Castleman of Advanced Digital Imaging Research, LLC (Friendswood, TX), formerly a subsidiary of Iris International, Inc. (Chatsworh CA) for providing 2D+3D facial data, and technical and financial support for this project. This work was also supported in part by the Advanced Technology Program of the National Institute of Standards and Technology.

References

Al-Osaimi, F., Bennaman, M., & Mian, A. (2009). An expression deformation approach to non-rigid 3d face recognition. *International Journal of Computer Vision*, *81*, 302–316.

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

Besl, P. J. (1988). *Surfaces in range image understanding*. New York: Springer.

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

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

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

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

Cadoni, M., Bicego, M., & Grosso, E. (2009). 3D face recognition using joint differential invariants. In *Advances in biometrics* (pp. 279–288). Berlin/Heidelberg: Springer.

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

Comas, J. (1960). *Manual of physical anthropology*. Charles C. Thomas.

DeCarlo, D., Metaxas, D., & Stone, M. (1998). An anthropometric face model using variational techniques. In *SIGGRAPH* (pp. 67–74).

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*, 269–271.

Do Carmo, M. P. (1976). *Differential geometry of curves and surfaces*. New Jersey: Prentice-Hall.

Douglas-Cowie, E., Cowie, R., & Schroder, M. (2000). A new emotion database: considerations, sources and scope. In *Proceedings of the ISCA ITRW on speech and emotion* (pp. 39–44).

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). New York: Wiley.

Egan, J. P. (1975). *Signal detection theory and ROC analysis*. New York: Academic Press.

Farkas, L. G. (1987). *Anthropometric facial proportions in medicine*. Thomas Books.

Farkas, L. G. (1981). *Anthropometry of the head and face in medicine*. New York: Elsevier.

Farkas, L. G. (Ed.) (1994). *Anthropometry of the head and face*. New York: Raven Press.

Farkas, L. G. et al. (1985). An attempt to define the attractive face: an anthropometric study. In *18th annual meeting of the American society for aesthetic plastic surgery*, Boston.

Gordon, G. G. (1992). Face recognition based on depth and curvature features. In *Computer vision and pattern recognition, 1992. Proceedings CVPR '92, 1992, IEEE Computer Society Conference on* (pp. 808–810).

Gupta, S., Aggarwal, J. K., Markey, M. K., & Bovik, A. C. (2007a). 3d face recognition founded on the structural diversity of human faces. In *Computer vision and pattern recognition, 2007 IEEE computer society conference on*, Minneapolis, MN.

Gupta, S., Markey, M. K., Aggarwal, J. K., & Bovik, A. C. (2007b). 3d face recognition based on Euclidean and geodesic distances. In *Electronic imaging, IS&T/SPIE 19th annual symposium on*, San Jose, CA.

Gupta, S., Markey, M. K., & Bovik, A. C. (2007c). Advances and challenges in 3D and 2D+3D human face recognition. In *Pattern recognition research horizons*. New York: Nova Science.

Gupta, S., Castleman, K. R., Markey, M. K., & Bovik, A. C. (2010). Texas 3d face recognition database. In *Image analysis and interpretation, 2010 IEEE Southwest Symposium on*, Austin, TX.

Hamza, A., & Krim, H. (2006). Geodesic matching of triangulated surfaces. *IEEE Transactions on Image Processing*, *15*(8), 2249–2258.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory and Cognition*, *33*(1), 98–106.

Heseltine, T. N. Pears, & Austin, J. (2004). Three-dimensional face recognition: a fishersurface approach. In A. Campilho, & M. Kamel (Eds.), *LNCS: Vol. 3212. International conference on image analysis and recognition* (pp. 684–691). Berlin/Heidelberg: Springer.

Hesher, C., Srivastava, A., & Erlebacher, G. (2003). A novel technique for face recognition using range imaging. In *Signal processing and its applications, 2003. Proceedings. Seventh international symposium on* (Vol. 2, pp. 201–204).

Hrdlička, A. (1939). *Practical anthropometry*. Wister Institute of Anatomy and Biology, Philadelphia.

Hüsken, M., Brauckmann, M., Gehlen, S., & Von der Malsburg, C. (2005). Strategies and benefits of fusion of 2d and 3d face recognition. In *Computer vision and pattern recognition, 2005 IEEE computer society conference on* (Vol. 3, pp. 174–174).

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

Irfanöglu, M. O., Gökberk, B., & Akarun, L. (2004). 3d shape-based face recognition using automatically registered facial surfaces. In *Proceedings of the 17th international conference on pattern recognition*. IEEE Computer Society, Dept. of Comput. Eng., Bogazici Univ., Turkey (Vol. 4, pp. 183–186).

Jahanbin, S., Bovik, A. C., & Choi, H. (2008). Automated facial feature detection from portrait and range images. In *Image analysis and interpretation, 2008. SSIAI 2008. IEEE southwest symposium on* (pp. 25–28).

Koudelka, M. L., Koch, M. W., & Russ, T. D. (2005). A prescreener for 3d face recognition using radial symmetry and the Hausdorff fraction. In *Computer vision and pattern recognition, 2005 IEEE computer society conference on* (Vol. 3, pp. 168–168).

Kukula, E. P., Elliott, S. J., Waupotitsch, R., & Pesenti, B. (2004). Effects of illumination changes on the performance of geometrix facevision/spl reg/ 3d frs. In *Security technology, 2004. 38th annual 2004 international Carnahan conference on* (pp. 331–337).

Lee, Y., Song, H., Yang, U., Shin, H., & Sohn, K. (2005). Local feature based 3d face recognition. In *LNCS: Vol. 3546. Audio- and video-based biometric person authentication, 2005 international conference on* (pp. 909–918). Springer: Berlin.

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

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

Maurer, T., Guigonis, D., Maslov, I., Pesenti, B., Tsaregorodtsev, A., West, D., & Medioni, G. (2005). Performance of geometrix activeid^{TM} 3d face recognition engine on the frgc data. In *Computer vision and pattern recognition, 2005 IEEE computer society conference on* (Vol. 3, pp. 154–154).

Mian, A., Bennamoun, M., & Owens, R. (2007). An efficient multimodal 2d-3d hybrid approach to automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*, 1927–1943.

Moreno, A. B., Sanchez, A., Fco, J., Fco, V., & Diaz, J. (2003). Face recognition using 3d surface-extracted descriptors. In *Irish machine vision and image processing conference (IMVIP 2003)*.

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

Phillips, P. J., Grother, P., Micheals, R. J., Blackburn, D. M., Tabassi, E., & Bone, M. (2003). *Face recognition vendor test 2002 evaluation report*. Tech. rep., National Institute of Standards and Technology.

Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., Marques, J., Min, J., & Worek, W. (2005). Overview of the face recognition grand challenge. In *Computer vision and pattern recognition, 2005. CVPR 2005. IEEE computer society conference on* (Vol. 1, pp. 947–954).

Rodriguez, J. J., & Aggarwal, J. K. (1990). Matching aerial images to 3-d terrain maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(12), 1138–1149.

Rogers, B. O. (1974). The role of physical anthropometry in plastic surgery today. *Clinical Plastic Surgery*, *1*, 439.

Russ, T. D., Koch, M. W., & Little, C. Q. (2005). A 2d range Hausdorff approach for 3d face recognition. In *Computer vision and pattern recognition, 2005 IEEE computer society conference on* (Vol. 3, pp. 169–176).

Samir, C., Srivastava, A., Daoud, M., & Klassen, E. (2009). An intrinsic framework for analysis of facial surfaces. *International Journal of Computer Vision*, *82*, 80–95.

Savran, A., Alyz, N., Dibekliolu, H., Eliktutan, O., Gkberk, B., Sankur, B., & Akarun, L. (2008). Bosphorus database for 3d face analysis. In *LNCS: Vol. 5372/2008. Biometrics and identity management* (pp. 47–56). Berlin/Heidelberg: Springer.

Sharma, S. (1996). *Applied multivariate techniques*. New York: Wiley.

Shi, J., Samal, A., & Marx, D. (2006). How effective are landmarks and their geometry for face recognition? *Computer Vision and Image Understanding*, *102*(2), 117–133.

Srivastava, A., Samir, C., Joshi, H. S., & Daoudi, M. (2009). Elastic shape models for face analysis using curvilinear coordinates. *Journal of Mathematical Imaging and Vision*, *33*, 253–265.

Tanaka, J., & Farah, M. (2003). The holistic representation of faces. In *Perception of faces, objects and scenes: analytical and holistic processes* (pp. 53–74). New York: Oxford University Press.

Tenenbaum, J. B., de Silva, V., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, *290*(5500), 2319–2323.

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

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

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

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

Xu, C., Wang, Y., Tan, T., & Quan, L. (2004). Automatic 3d face recognition combining global geometric features with local shape variation information. In *Automatic face and gesture recognition, 2004. Proceedings. Sixth IEEE international conference on* (pp. 308–313).

Yin, L., Wei, X., Sun, Y., Wang, J., & Rosato, M. J. (2006). A 3d facial expression database for facial behavior research. In *Automatic face and gesture recognition, 2006. FGR 2006. 7th international conference on* (pp. 211–216).

Zhang, G., & Wang, Y. (2009). Faceprint: fusion of local features for 3D face recognition. In *Advances in biometrics* (pp. 394–403). Berlin/Heidelberg: Springer.

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

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