A quantitative assessment of 3D facial key point localization fitting 2D shape models to curvature information

Federico M. Sukno*†, Tarik A. Chowdhury*, John L. Waddington† and Paul F. Whelan*

*Centre for Image Processing & Analysis, Dublin City University, Dublin 9, Ireland
†Molecular & Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin 2, Ireland

Abstract—This work addresses the localization of 11 prominent facial landmarks in 3D by fitting state of the art shape models to 2D data. Quantitative results are provided for 34 scans at high resolution (texture maps of 10 M-pixels) in terms of accuracy (with respect to manual measurements) and precision (repeatability on different images from the same individual). We obtain an average accuracy of approximately 3 mm, and median repeatability of inter-landmark distances typically below 2 mm, which are values comparable to current algorithms on automatic localization of facial landmarks. We also show that, in our experiments, the replacement of texture information by curvature features produced little change in performance, which is an important finding as it suggests the applicability of the method to any type of 3D data.

I. INTRODUCTION

We address the problem of accurate localization of distinctive facial points (*landmarks*) in three dimensions (3D). This is an important aspect for a majority of 3D facial analysis algorithms, as it often constitutes their initial step or a prerequisite (that is sometimes addressed by manual interaction).

Most automatic algorithms for the extraction of facial points are framed within identity recognition. Due to the evolution and higher availability of 3D imaging devices, this field has experienced a considerable growth in the last decade [2], [7], [12], [14], [30]. In this context, although some methods require just a rough spatial normalization of the face to identify, an accurate and precise (repeatable) localization of facial landmarks is accepted to benefit performance. Best results currently obtained indicate average errors between 3 mm and 6 mm for the most distinctive landmarks, with the exception of the nose tip for which average errors slightly above 2 mm have been reported [12], [19], [22], [29], [31].

On the other hand, facial landmarks are also of interest for a number of clinical applications, like facial surgery [23] or craniofacial dysmorphology [9]. While for identity recognition landmark detection can be considered optional, in the applications just listed their accurate localization is crucial, as they constitute the basis for the analysis, often aimed at detecting small shape differences. Depending on the author, localization and repeatability errors are considered clinically relevant when they exceed 1 mm [18] or 2 mm [1]. Recent studies on facial alginate impressions (rigid reconstructions of human heads) suggest that modern 3D scanners allow for sub-millimeter accuracy [11], although results on manual

identification of key-points in real faces indicate that average localization errors are typically between 1 mm and 2 mm [1], [3], [6], [8], [18], [21], [26], [27]. Unfortunately, these studies tend to involve a limited number of patients (usually $n < 20$), and some attempts to use automatic algorithms for this task have reported insufficient accuracy [21]. The evaluation on larger high quality datasets and a direct comparison between manual and automatic localization is thus required.

A. Related work

There is a considerable body of work on the detection of 3D facial landmarks. We concentrate on state of the art works providing quantitative evaluation on accuracy and/or precision, as this allows for quantitative comparison.

Most methods for detecting landmarks in 3D are based on the computation of curvature features (e.g. mean, Gaussian, shape-index, principal curvatures) on the range data (either in 2.5D or 3D) [5], [7], [22], [25]. As the most widespread feature, curvature has been shown to provide state of the art accuracy.

Approaches not based on curvature but still using exclusively 3D geometry as input data, include the response of range data when convolved with a set of primitive filters [29] or Gabor wavelets [4] and combinations of features like spin images, distance to local plane or RBF Shape Histograms [16], [20]. Nonetheless, they do not seem to outperform curvature-based approaches. Other approaches also employ the profile contours of 2D projections to detect a very limited set of key points like the nose tip and eye corners [14], [22].

Recently, some authors have suggested the use of texture information, which is often provided aligned with the 3D data. For example, Zhao et al. [31] fit an Active Shape Model (ASM) to combined data from 2.5D scans and texture. Both depth (z-coordinate) and texture information are sampled and three independent PCA models are built for shape, texture, and depth information. The model is fit to a new instance based on a simplex minimization using normalized correlation and statistical constraints on the PCA-coefficients. Comparison to Szeptycki et al. [25], based on curvature, is favorable for all landmarks except the nose tip and inner-eye corners.

Zhang and Wang [30] also propose to extract key-points independently from texture (using SIFT) and range data (detecting scale space extrema on shape index), although in

this case the algorithm is focused on *salient* points chosen automatically and, in general, not necessarily coincident with landmark definitions (the method is aimed at recognition).

An important advantage of the strategies proposed in [31] and [30] is that they allow for using well established algorithms from the 2D domain. On the other hand, the need for texture information constitutes a limitation, especially for certain clinical applications based on laser scanners, where only geometry is available.

In this work, we analyze the performance of a 2D ASM in the localization of 3D landmarks and provide a quantitative comparison of both the accuracy and precision that can be obtained when using texture or just curvature information mapped in 2D. The latter is computed on the 3D mesh and aligned back into the texture images, allowing for a direct comparison of both input features. As we aim at highly accurate localization, the Invariant Optimal Features (IOF) variant of ASM [24] is used and our data is acquired from a high resolution scanner, which provides texture maps at 10 Mega-pixels, considerably higher than those generally used in the evaluation of automatic methods, and comparable to data reported in clinical studies.

Our results, based on a dataset of 34 facial scans, suggest that an IOF-ASM working on 2D data can be used to localize most salient facial landmarks at an average accuracy of approximately 3 mm, and its precision, measured as the repeatability of inter-landmark distances, is typically below 2 mm (median values). We also show that, in our experiments, the replacement of texture information by curvature features produced little change in performance, which is an important finding as it suggests the applicability of the method to any type of 3D data.

II. METHODOLOGY

This section provides a brief overview of the Active Shape Model with Invariant Optimal Features (IOF-ASM). The reader is referred to [24] for further details.

An IOF-ASM is composed of one shape model or PDM (Point Distribution Model) and as many appearance models as the number of landmarks composing the shape. Both the shape and the appearance models must be trained from a set of annotated images where a set of landmarks (or key points) define the contours of interest.

The shape model is trained by Principal Component Analysis (PCA) on the landmark coordinates from the training set, which results in a set of global constraints about the (statistical) plausibility for any given shape. On the other hand, each appearance model works as a local classifier based on image data. The inputs to the classifier are features based on (scale-space) image derivatives computed in the neighborhood of the landmark. Those image derivatives are appropriately combined to generate differential invariants to rigid transformations.

When the IOF-ASM is used for segmentation, only two inputs are required: an image containing a face and a starting guess of the face position (e.g. provided by a face detector). The process begins by placing an average shape at the initial position on the image. Subsequently, at each iteration and for each landmark, the corresponding appearance model determines the locally best position to place the landmark. Then, the landmarks are constrained by the PDM so that they generate a plausible shape, which is used as the starting point for the following iteration. A predefined number of iterations are executed after which the model is assumed to be fitted.

A. Curvature Mapping

As stated previously, in this paper we apply a 2D IOF-ASM to obtain the 3D position of facial landmarks. This is possible based on the information provided by the 3D scanner for each facial scan, which includes:

- The 3D reconstruction, as a triangulated surface, with coordinates expressed in millimeters.
- Two 2D color images at 10 M-pixels spatial resolution, one from each side (usually indicated as *nearly frontal*).
- The range map relating every pixel from the 2D images to the triangulated surface.

Thus, the results from IOF-ASM on the 2D color images can be directly mapped into 3D through the range map. Additionally, we also generated 2D images with curvature information aligned to the 2D color images. The curvature information was computed on the 3D surface mesh for every vertex based on the method developed by Meyer et al. [13]. This method offers a theoretical justification and is parameter-free, as opposed to other works requiring the definition of neighborhoods or smoothing coefficients. Additionally, the curvature sign was computed based on the relative direction between the curvature vectors obtained as indicated by Meyer et al. and the normals enforced to *point outwards* from the object.

Once curvature was computed, we mapped it into the 2D images by using their corresponding range maps. The value assigned to pixel $\mathbf{p}$, whose 3D position indicated by the range map is $\mathcal{R}(\mathbf{p})$ is:

$$K(\mathbf{p}) = \frac{\sum_{\mathcal{B}_p} \kappa(\mathbf{v}) e^{-\frac{\|\mathbf{v} - \mathcal{R}(\mathbf{p})\|^2}{\sigma_K^2}}}{\sum_{\mathcal{B}_p} e^{-\frac{\|\mathbf{v} - \mathcal{R}(\mathbf{p})\|^2}{\sigma_K^2}}}$$

where $\mathcal{B}_p$ is the set of all mesh vertices $\mathbf{v}$ within $3\sigma_K$ from $\mathcal{R}(\mathbf{p})$ and $\kappa(\mathbf{v})$ is some curvature measure at vertex $\mathbf{v}$. The numerator simulates an *aperture* with a Gaussian weighting and the denominator provides the normalization factor independently of the number of neighbors that fall within $\mathcal{B}_p$. We chose $\sigma_K = 0.33$ mm, a value that is similar to the resolution of our scanner and which resulted in $\mathcal{B}_p$ typically including only one or two vertices. The choice was aimed at avoiding the loss of detail due to smoothing, at the expense of accepting considerable noise on the input data. An example of the resulting *curvature-images* is shown in Fig. 1.

Fig. 1. Example of mean curvature mapped into the 2D view of a subject using Gaussian windows with standard deviation 0.33 mm.

III. RESULTS

Tests were performed on 34 facial scans acquired with a photometric stereo scanner[1], corresponding to 25 different persons. The dataset includes spontaneous expression (61% neutral, 24% happiness, 15% other) and some important pose variation in nodding (12% strongly *facing up*). On the other hand, there are only residual head rotations to either side.

Three different methods were used to localize the facial landmarks:

- Manual annotation in the 3D surface, with the texture mapped into the mesh triangles[2], to be used as ground truth.
- IOF-ASM segmentation of the 2D input images used for the stereo reconstruction. This method is referred in the plots as *Texture-based*.
- IOF-ASM segmentation of the curvature mapped into the images used for the stereo reconstruction. This method is referred in the plots as *Curvature-based*.

The results that are presented in this work were obtained using only the mean curvature. Together with Gaussian curvature, mean curvature would provide a complete representation from which we could derive any other related property (e.g. principal curvatures, shape index). However, in our experiments, adding Gaussian curvature did not significantly change performance, but it rather showed a tendency to increase outliers. We believe this might be due to the higher sensitivity of Gaussian curvature to noise, when compared to mean curvature.

Experiments based on IOF-ASM segmentation need a suitable initialization consisting on a rough localization of the face. We have not addressed the automatization of this point and initialization was provided manually. Nonetheless, considering that the segmentation is performed on 2D images of a single person, state of the art methods for face detection are known to provide excellent performance for this task [15].

---

[1]Di3D FTP001 Scanner (Dimensional Imaging Ltd., Scotland).
[2]We used Di3DView software (Dimensional Imaging Ltd., Scotland).

A. Model-to-image adaptation

Fig. 2 shows the average accuracy of the model to image adaptation, measured as the point-to-curve distance between the landmarks automatically found by the model and the manual annotations. Error values are color coded per landmark, which were arranged in a 205-point template based on the one from PUT database [10] plus 10 additional landmarks to cover the upper part of the nose.

The results correspond to the average from the 34 shapes obtained for the right-profile images. The overall mean (averaged over all landmarks) and standard error were $2.23\% \pm 0.16\%$ for the texture-based model and $2.64\% \pm 0.27\%$ for the curvature-based model (measured as percentage of the inter-ocular distance, which would correspond to 100%). In both cases, the overall initialization error was $14.7\% \pm 7.2\%$, which seems a reasonably rough starting point for the models.

It can be seen that the accuracy from texture and curvature models is similar. However, the latter shows higher errors toward the face side that is furthest from the camera (left in this case). This occurs because, even though the curvature can be easily obtained for all parts of the mesh (it is computed in 3D), there is an increased difficulty to estimate the 2D-3D mapping to that area, including possible occlusions. However, this is not a concern as the other camera of the stereo system provides a complementary (mirrored) behavior.
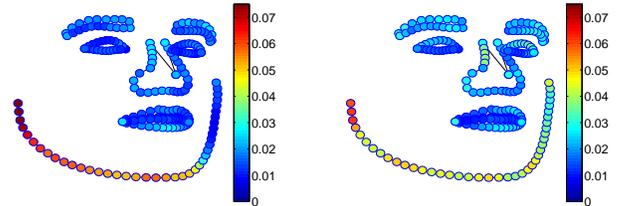

Fig. 2. Point-to-curve error of the model-to-image adaptation in 2D, normalized with respect to the inter-ocular distance (which would correspond to 1.0), for the texture-based (left) and curvature-based (right) models.

B. 3D Accuracy

Fig. 3 shows the accuracy of both texture-based and curvature-based results. Distances were measured with respect to the manual annotations for a total of 34 scans, in a 3-fold cross validation. Since there are repeated scans for the same person, it was made sure that no subject was included in both test and training sets at any time.

It can be seen that the median localization accuracy for eyes and mouth landmarks is between 2 mm and 3 mm, while nose points exhibit errors that can exceed 4 mm. It is worth pointing that, in 2D images (which are the data used to train the models), *pronasale* and both *noistril-base points* are more ambiguous for manual annotation than it is *subnasale*, which correlates with the higher accuracy achieved for the latter.

Regarding the comparison between texture and curvature, differences are negligible for most of the points. Exceptions are *labiale superius*, more accurate with texture, and

TABLE I
LANDMARK DEFINITIONS AND ABBREVIATIONS

Name	Abbr	Description
Cheilion	ch	Mouth corner(s) (labial commissure)
Exocanthion	ex	Outer eye-corner(s)
Endocanthion	en	Inner eye-corner(s)
Labiale superius	ls	Upper-lip midpoint on the vermilion line
Subnasale	sn	Midpoint at which the nasal septum merges with the upper lip (midsagittal plane)
Nostril base	nb	Inferior terminal point of each nostril axis
Pronasale	pm	Most anterior midpoint of the nasal tip

Fig. 4. Repeatability of the measurements for 8 inter-landmark distances from the eyes and nose regions. Landmark names are indicated following the abbreviations from Table I.
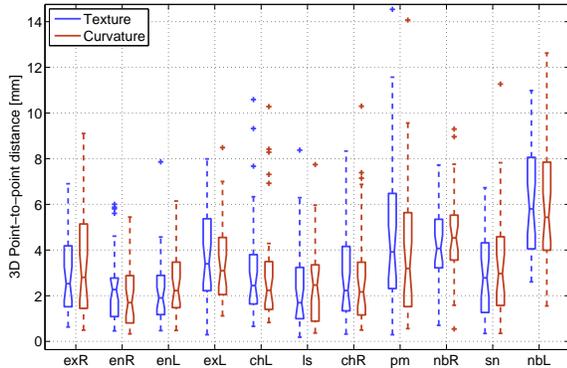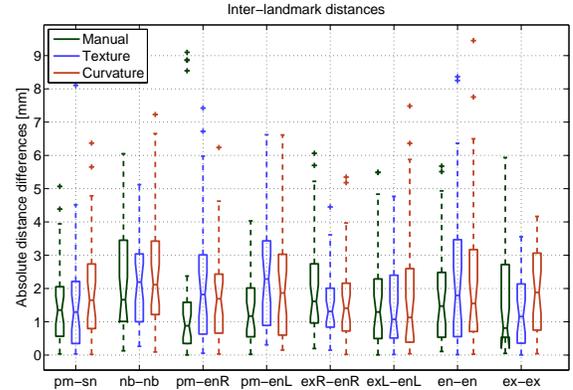

Fig. 3. Accuracy of the automatic methods with respect to ground truth annotations for 11 landmarks, following abbreviations from Table I. For symmetric landmarks, Left or Right are additionally indicated.

pronasale, more accurate with curvature, the latter fact coinciding with results previously reported by Zhao et al. [31].

C. Precision

Fig. 4 provides a measure of precision for all three methods. Taking advantage of the fact that our dataset contains repeated scans for some of the participants, we measured the difference between inter-landmark distances computed on pairs of scans belonging to the same individual. We identified 32 scan pairs that can be compared. As the scans presented considerable variations in facial expressions we restricted ourselves to 8 measurements involving eyes and nose points (as their distances are acknowledged to be less sensitive to facial expression).

It can be noticed that, in most cases, the repeatability of the two automatic methods is comparable to the manual annotations and the medians rarely exceed 2 mm.

IV. DISCUSSION

In this section we discuss the results presented above in the context of other works that have reported quantitative evaluations. We focus on state of the art papers providing average localization errors, and do not include results based on percentages of accurate detection (i.e. error below a certain threshold) as they are more difficult to compare.

A. Accuracy

Table II provides a list of reported results on landmark localization measured as the average distances to manual annotations, which are used as ground truth. Errors are specified on a per-landmark basis, according to the abbreviations provided in Table I, with the exception of *alare* points (al), which are defined as the most lateral points at each alar contour of the nose. Although *alare* and *nostril base* are distinct points, some works provide results on more vaguely defined points, indicated as *nose corners*, so we have decided to group them under the same column. We have also merged (in this case by averaging) the results for symmetric (i.e. left and right) landmarks, as there should not be a fundamental difference in their localization accuracy.

Most papers in the list have been tested on subsets of FRGC database [17]. The 3D scans in this database are of considerably lower resolution than our scans: the average distance between eye centers in FRGC is 160 pixels, while in our case it is above 500 pixels. This is an important consideration when analyzing the results from Table II. However, as showed in the next section, we did not observe a significant improvement in the precision of manual localization as a consequence of higher resolution. To some extent, this is also observed in the results provided by Segundo et al. [22], who obtained comparable accuracy on FRGC (texture maps of 640×480 pixels) and BU-3DFE [28] (texture maps of approx. 1300×900 pixels).

As the title of Table II suggests, when a paper reported results on more than one dataset or method (which is the case of [12], [16], [22]), we chose to include the most accurate ones. An additional list of works analogous (but complementary) to Table II, can be found in [16], although their performance was always below the one reported by Perakis et al. themselves.

With the above observations in mind, it is clear that the comparison provided by Table II must be taken carefully. Nonetheless, it suggests that our results are comparable to the best methods currently reported for all of the analyzed points but *pronasale*. Unfortunately, the size of our dataset is

TABLE II
BEST REPORTED AVERAGES ON LANDMARK LOCALIZATION ERRORS [MM]

Method	ch	en	ex	ls	sn	nb / al	pm
IOF-ASM (texture)	3.2	2.3	3.3	2.4	2.9	5.1	4.9
IOF-ASM (curvature)	3.0	2.3	3.5	2.5	3.5	5.3	4.7
D'Hose et al. [4]	-	-	-	-	-	-	3.17
Lu and Jain [12]	6.1	8.05	9.9	-	-	-	6.1
Perakis et al. [16]	6.03	5.31	5.76	-	-	-	4.88
Segundo et al. [22]	-	3.52	-	-	-	5.34	1.87
Szeptycki et al. [25]	8.56	3.85	2.82	-	-	6.18	2.27
Yu et al. [29]	-	5.17	-	-	-	-	2.14
Zhao et al. [31]	3.93	3.21	4.27	2.72	-	4.47	2.68

considerably smaller than the other works listed in the table, ranging from one to several hundreds of facial scans.

The reason for the lack of accuracy in the localization of *pronasale* could be found in the training data for IOF-ASM, which was obtained by manual annotation of the 2D *texture* images, which i) are nearly frontal, making it difficult to accurately determine the nose tip, and ii) provide very weak texture patterns for most of the nose points. In this aspect, the correction of 2D landmarks based on 3D and/or curvature information could be beneficial. Nonetheless, it is a potential limitation of the method and it might be preferable to use an alternative strategy for this particular point, as it is reported to be the most accurately detected by several other methods.

Among works not included in the table, we shall highlight those from Romero and Pears [19], who report median RMS errors around 3 mm and 5 mm for *pronasale* and *endocanthion(s)*, and Gupta et al. [7] who report only the standard deviation of the localization error (for 10 facial landmarks). It is unfortunate that the averages were not included in the latter work, as the reported standard deviation values are similar to those obtained by the most accurate algorithms.

B. Precision

An important question when evaluating the localization accuracy is the repeatability of the measurements, including those obtained manually as this indicates the quality of the ground truth. Several works address this problem, both from the perspective of image-based measurements and *direct* anthropometry (i.e. measurements derived from a caliper, measuring tape and so on).

Table III provides a list of recent papers reporting the average absolute differences for inter-landmark distances, together with the results of our own manual annotations. Analogously to our accuracy comparison, we merged bilateral measurements (by averaging). As explained earlier, due to the variations in facial expression we constrain our comparison to landmarks on the eyes and nose.

The first thing to notice is the large dispersion of the values reported by different researchers. It is surprising, for example, that errors for the intercanthal width (*en-en*) are larger than the biocular width (*ex-ex*) in half of the works collected in the table, while the opposite holds for the other half. A possible explanation can be found in the relatively small population size of most works, not exceeding 20 cases, with the exception of

Heike et al. [8] ($n = 40$) and ourselves. Indeed, the results by Heike et al. are the closest ones to ours.

It should also be noticed that the numbers reported on Table III are not directly comparable in all cases. The majority of them report the difference between direct anthropometric measurements and image-based measurements performed in 3D, usually with the help of visible marks on the facial surface to facilitate the localization of some of (or all) the landmarks. In the case of Heike et al. the authors reported inter- and intra-observer variabilities, as well as inter-method (direct vs image) differences. We believe the inter-observer variability provides the most fair comparison to our results: it is based on the independent identification of landmarks in the same image by two different observers, while in our case the same observer identified the landmarks in two different images (from the same individual). Additionally, inter-method values reported by Heike et al. were provided as averaged (signed) differences, not fairly comparable with the absolute averages reported on the table (the same applies for the results reported by Schimmel et al. [21]). Nonetheless, we shall point out that some of these inter-method differences showed considerably larger errors than the inter-observer ones (for example, for the biocular distance the average errors were close to 5 mm).

Apart from inter-landmark distances, the repeatability of the landmark coordinates has also been widely reported. Again, population sizes tend to be small and results from different authors do not completely agree. While Plooij et al. [18] report inter-observer averages below 0.5 mm for 80% of the landmarks (from a total of 49 points in 20 patients), Toma et al. [26] report averages between 0.5 mm and 1.5 mm for all landmarks (a total of 21 in 30 patients), with the only exception of *labiale superius* which was located with an average error of 0.39 mm.

On the non-clinical side, Zhao et al. [31] provide the localization error for 15 landmarks on 10 facial scans from FRGC, averaging the manual annotations from 11 observers. Their results show manual errors between 2 mm and 3 mm for landmarks in the eyes and nose, except for vaguely defined points such as *nose corners*, which were found less accurate. Although a comparison between these results and errors of inter-landmark distances is not necessarily direct, it can be seen that the magnitude of the errors are similar to the ones from our manual annotations, in spite of the big differences in spatial resolution.

TABLE III

REPORTED PRECISION ON MANUAL MEASUREMENTS OF INTER-LANDMARK DISTANCES [MM]

Method	en-en	en-ex	ex-ex	en-pm	nb-nb	al-al	pm-sn
Section III-C (different images)	3.0	1.8	2.0	2.2	2.8	-	1.8
Ainechi et al. (direct vs image) [1]	0.35	0.09	0.54	-	-	0.35	0.80
De Menezes et al. (direct vs image) [3]	-	-	0.62	-	-	-	0.28
Ghoddous et al. (direct vs image) [6]	5.0	-	0.6	-	-	-	2.6
Heike et al. (inter-observer) [8]	0.85	1.89	2.09	-	-	0.88	0.88
Wong et al. (direct vs image) [27]	1.0	-	0.5	-	-	0.8	0.7

REFERENCES

[1] N. Aynechi, B.E. Larson, V. Leon-Salazar, and S. Beiraghi. Accuracy and precision of a 3D anthropometric facial analysis with and without landmark labeling before image acquisition. *Angle Orthodontist*, 81(2):245–252, 2011.

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

[3] M. de Manezes, R. Rosati, C. Allievi, and C. Sforza. A photographic system for the three-dimensional study of facial morphology. *Angle Orthodontist*, 79(6):1070–1077, 2009.

[4] J. D'Hose, J. Colineau, C. Bichon, and B. Dorizzi. Precise localization of landmarks on 3D faces using Gabor wavelets. In *Proc. 1st IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, Washington, DC, USA*, pages 1–6, 2007.

[5] H. Dibeklioglu, A.A. Salah, and L. Akarun. 3D faceial landmarking under expression, pose, and occlusion variations. In *Proc. 2nd IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, Arlington, VA, USA*, pages 1–6, 2008.

[6] H. Ghoddousi, R. Edler, P. Haers, D. Wertheim, and D. Greenhill. Comparison of three methods of facial measurement. *International Journal of Oral and Maxillofacial Surgery*, 36:250–258, 2007.

[7] S. Gupta, M.K. Markey, and A.C. Bovik. Antopometric 3D face recognition. *International Journal of Computer Vision*, 90(3):331–349, 2010.

[8] C.L. Heike, M.L. Cunningham, A.V. Hing, E. Stuhaug, and J. Starr. Picture perfect? reliability of craniofacial anthropometry using three-dimensional digital stereophotogrammetry. *Plastic & Reconstructive Surgery*, 124(4):1261–1272, 2009.

[9] R.H. Hennessy, P.B. Baldwin, D.J. Browne, A. Kinsella, and J.L. Waddington. Frontonasal dysmorphology in bipolar disorder by 3D laser surface imnaging and geometric morphometrics: Comparison with schizophrenia. *Schizophrenia Research*, 122(1-3):63–71, 2010.

[10] A. Kasiński, A. Florek, and A. Schmidt. The PUT face database. *Image Processing and Communications*, 13(3):59–64, 2008.

[11] B. Khambay, N. Nairn, A. Bell, J. Miller, A. Bowman, and A.F. Ayoub. Validation and reproducibility of a high-resolution three-dimensional facial imaging system. *British Journal of Oral and Maxilofacial Surgery*, 66:27–32, 2008.

[12] X. Lu and A.K. Jain. Automatic feature extraction for multiview 3D face recognition. In *Proc. 7th IEEE Int. Conf. on Automatic Face and Gesture Recognition, Southampton, UK*, pages 585–590, 2006.

[13] M. Meyer, M. Desbrun, P. Schröder, and A.H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds. In *Int. Workshop on Visualization and Mathematics, Berlin-Dahlem, Germany*, pages 1–26, 2002.

[14] A. Mian, M. Bennamoun, and R. Owens. An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1927–1943, 2007.

[15] S.-K. Pavani, D. Delgado Gomez, and A.F. Frangi. Haar-like features with optimally weighted rectangles for rapid object detection. *Pattern Recognition*, 43(1):160–172, 2010.

[16] P. Perakis. 3D facial landark detection and face registration. Technical report, University of Athens, Greece, 2010.

[17] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowver, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recogmition grand challenge. In *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, San Diego, CA, USA*, volume 1, pages 947–954, 2005.

[18] J.M. Plooij, G.R.J. Swennen, F.A. Rangel, T.J.J. Maal, F.A.C. Schutyser, E.M. Bronkhorst, A.M. Kuijpers-Jagtman, and S.J. Bergé. Evaluation of reproducibility and reliability of 3D soft tissue analysis using 3D stereophotogrammetry. *International Journal of Oral and Maxillofacial Surgery*, 38(3):267–273, 2009.

[19] M. Romero-Huertas and N. Pears. 3D facial landmark localisation by matching simple descriptors. In *Proc. 2nd IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, Arlington, VA, USA*, pages 1–6, 2008.

[20] M. Romero-Huertas and N. Pears. Landmark localisation in 3d face data. In *Proc. 6th IEEE Int. Conf. on Advanced Video and Signal Based Surveillance, Washington, DC, USA*, pages 73–78, 2009.

[21] M. Schimmel, P. Chirstou, O. Houstis, F.R. Hermann, S. Kiliaridis, and F. Müller. Distances between facial landmarks can be measured accurately with a new digital 3-dimensional video system. *American Journal of Orthodontics and Dentofacial Orthepedics*, 137(5):580e1–e10, 2010.

[22] M.P. Segundo, L. Silva, O.R. Pereira Bellon, and C.C. Queirolo. Automatic face segmentation and facial landmark detection in range images. *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, 40(5):1319–1330, 2010.

[23] A. Sharifi, R. Jones, A. Ayoub, K. Moos, F. Walker, B. Khambay, and S. McHough. How accurate is model planning for orthognathic surgery. *International Journal of Oral and Maxillofacial Surgery*, 37(12):1089–1093, 2008.

[24] F.M. Sukno, S. Ordas, C. Butakoff, S. Cruz, and A.F. Frangi. Active shape models with invariant optimal features: Application to facial analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(7):1105–1117, 2007.

[25] P. Szeptycki, M. Ardabilian, and L. Chen. A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking. In *Proc. 3rd IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, Arlington, VA, USA*, pages 1–6, 2009.

[26] A.M. Toma, A. Zhurov, R. Playle, E. Ong, and S. Richmond. Reproducibility of facial soft tissue landmarks on 3D lasser-scanned facial images. *Orthodontics & Craniofacial Research*, 12(1):33–42, 2009.

[27] J.Y. Wong, A.K. Oh, E. Ohta, A.T. Hunt, G.F. Rogers, J.B. Mulliken, and C.K. Deutsch. Validity and reliability of craniofacial anthropometric measurement of 3D digital photogrammetric images. *The Cleft Palate-Craniofacial Journal*, 45(3):232–239, 2008.

[28] L. Yin, X.Wei, Y. Sun, J.Wang, and M.J. Rosato. A 3D facial expression database for facial behavior research. In *Proc. 7th IEEE Int. Conf. on Automatic Face and Gesture Recognition, Southampton, UK*, volume 1, pages 211–216, 2006.

[29] T.H. Yu and Y.S. Moon. A novel genetic algorithm for 3D facial landmark localization. In *Proc. 2nd IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, Arlington, VA, USA*, pages 1–6, 2008.

[30] G. Zhang and Y. Wang. Robust 3D face recognition based on resolution invariant features. *Pattern Recognition Letters*, 32:1009–1019, 2011.

[31] X. Zhao, P. Szeptycki, E. Dellandrea, and L. Chen. Precise 2.5D facial landmarking via an analysis by synthesis approach. In *Proc. Workshop on Applications of Computer Vision, Sonowbird, UT, USA*, pages 1–7, 2009.