3D face recognition in the presence of facial expressions based on empirical mode decomposition

Abdelghafour Abbad
Laboratory LIIAN, Faculty of Science
Dhar EL Mahraz
Morocco
gh.abbad@gmail.com

Khalid Abbad
Laboratory ISA, Faculty of Science and Technology
Morocco
abbadkhalid@gmail.com

Hamid Tairi
Laboratory LIIAN, Faculty of Science
Dhar EL Mahraz
Morocco
htairi@yahoo.fr

ABSTRACT

This paper presents an efficient 3D face recognition method to handle facial expression. The proposed method uses the Surfaces Empirical Mode Decomposition (SEMD), facial curves and local shape descriptor in a matching process to overcome the distortions caused by expressions in faces. The basic idea is that, the face is presented at different scales by SEMD. Then the isometric-invariant features on each scale are extracted. After that, the geometric information is obtained on the 3D surface in terms of radial and level facial curves. Finally, the feature vectors on each scale are associated with their corresponding geometric information. The presented method is validated on GavabDB database resulting a rank 1 recognition rate (RR) of 98.9% for all faces with neutral and non-neutral expressions. This result outperforms other 3D expression-invariant face recognition methods on the same database.

CCS CONCEPTS

• **Computing methodologies** → **Biometrics;** • **Computing methodologies** → *Computer vision*

KEYWORDS

3D face recognition, EMD, geometric features, local features, facial curves, expression

ACM Reference format:

Abdelghafour Abbad, Khalid Abbad, and Hamid Tairi. 2018. 3D face recognition in the presence of facial expressions based on empirical mode decomposition. In Proceedings of Mediterranean Conference on Pattern Recognition and Artifcial Intelligence, Rabat, Morocco, March 27–28, 2018 (MedPRAI '18), 6 pages
https://doi.org/10.1145/3177148.3180087

MedPRAI '18, March 27–28, 2018, Rabat, Morocco

ACM ISBN 978-1-4503-5290-1/18/03…$15.00
https://doi.org/10.1145/3177148.3180087

1 INTRODUCTION

Face recognition is extremely desirable in a wide range of applications as well as surveillance, access control and machine-human interaction. Stunningly, recent studies report that automatic face recognition will even surpass the human performance in some specific conditions [26]. However, 2D face recognition systems suffers from many factors, such as pose changes, illumination variations, facial expressions and occlusions[1; 32]. In order to overcome these limitations, 3D scans of the face has been recently proposed. Utilizing 3D data will improve the system performance because of its specific illustration of facial surface. However, the 3D form of the face will be strongly distorted below facial expressions. This distortion can introduce several non-rigid parts which reduces the similarity between faces from the same person. Some efforts have been made to attack this problem.

In literature, four different strategies can be distinguished for modeling the deformations due to expressions: rigid [13; 15; 17; 23], non-rigid [3; 4; 14; 20], geometric form [6; 9] and keypoints detection [8; 12; 22; 27].

(i) Rigid matching algorithms: this category of methods extracts the rigid areas of the face (e.g., eyes, forehead and nose). The partial ICP method [13] overcomes the problem of facial expression variance by calculating dynamically the rigid parts of the facial surfaces at each iteration of the ICP algorithm during similarity analysis. Lei et al.[15] developed a binary mask to sub-divide a facial scan into rigid (nose), semi-rigid (eyes-forehead) and non-rigid (mouth) regions. Rigid and semi-rigid regions are accounted for to eliminate the deformations caused by facial expressions. In each region, they extract four types of local low-level geometrical features. These features are fused into a single descriptor, and then, the Support Vector Machine (SVM) is adopted for classification.

(ii) Non-rigid matching algorithms: eliminate the information related to expression and, at the same time, retain the facial characteristics. Al-Osaimi et al. [3] present a non-rigid approach where the facial expression deformations are modeled from pairs of neutral and non-neutral faces of each individual in the training stage using PCA. These patterns then serve to morph out any expression deformation from the 3D scans before matching and extracting the similarity measures.

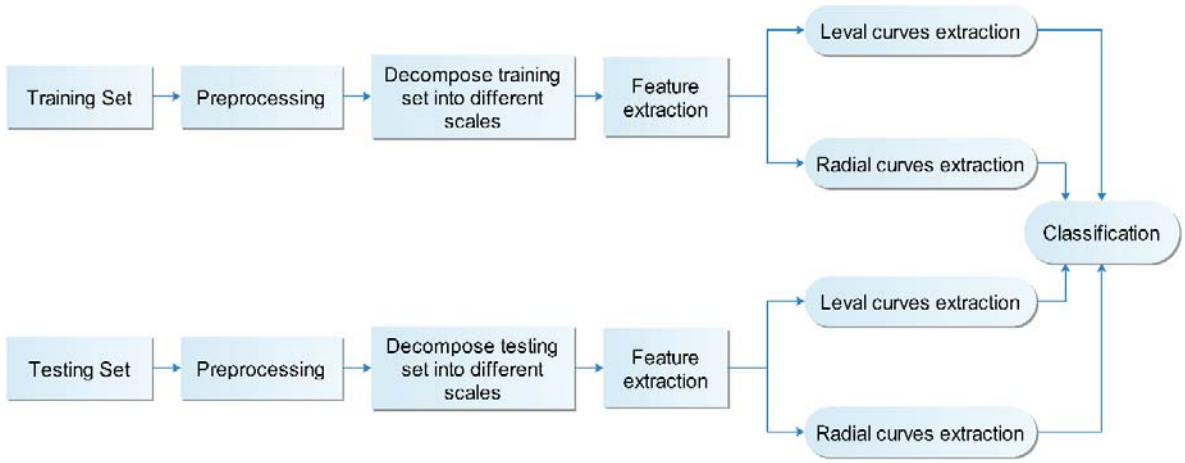

Figure 1: Overview of the proposed method

(iii) Geometric form matching algorithms: focus on the geometric shape analysis. Drira et al. [9] consider curves to be geometric features. The basic idea is to approximate the facial surfaces by a finite and indexed collection of radial curves that emanate from the nose tips. These shapes model the facial deformations that are caused by changes in facial expressions. Faces are compared using Riemannian geometry [28], which defines the geodesic paths between the curves and the distances between them. The final decision is obtained by combining the similarity scores that are produced by each pair of curves. Ballihi et al. [6] also suggested a method that is based on curves to represent the whole facial surface and the Riemannian geometric shape analysis framework to compare and match the facial curves. In this study, they employed both level and radial curves as geometric features. Then, they selected the most stable and most discriminative curves (salient curves) by the Adaboost algorithm [11].

(iv) Keypoints detection matching algorithms: extract salient point features on the face scans. Mian et al. [22] extracted local features individually from the 3D scans and texture of the face. The features are projected to their respective PCA subspaces before they are normalized into a unit. These feature vectors are then concatenated to form a single 2D+3D feature vector. The combined feature vectors are used to calculate the similarity between a probe and a gallery face. According to Huang et al. [12], facial depth images are first represented by multi scale extended Local Binary Patterns (eLBP), and the Scale Invariant Feature Transform (SIFT) [19] framework is then performed on these images to find keypoints, extraction of local features, and matching faces. Recently, Berretti et al. [8] identified first the keypoints from a 3D face based on the meshDOG keypoints [31] detector and the multi-ring geometric histogram descriptor. For face matching, they detected stable keypoints and selected the most effective features from the local descriptors using the RANSAC algorithm [33].

In our work, we propose a new method that is based on geometric and local information of scans in 3D. To have good extraction of facial features, we work on three levels. The first level is the separation of the facial feature information on different scales. The second level is to allocate to each point on the face a descriptive vector to have the local information of each point. In the last level, the geometric information of the face is extracted through the radial and level curves. This work is adapted from our published article [2].

The remainder of the paper is organized as follows. Section 2 first gives the details of the proposed method. The experimental results and discussions are presented in Section 3. We conclude the paper in Section 4.

2 PROPOSED APPROACH

The various sequence of steps followed in our approach is described in Fig. 1.

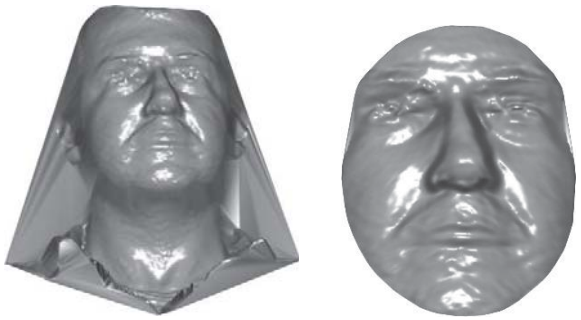

Figure 2: An example of scan before (left) and after preprocessing(right)

2.1 Preprocessing

The preprocessing pipeline contains the following: hole filling, nose tip localization and face cropping. For the nose tip localization and filling of holes, we relied on some work proposed in the literature review, especially the work of Szeptycki et al. [29]. The face area on each scan is obtained automatically by using a sphere centered at the nose tip with a radius of 100 mm (see Fig. 2). To correct the minor pose variations in the scans, the preprocessed 3D facial surfaces are aligned to the corresponding reference model using the iterative closest point (ICP) [10] algorithm.

2.2 Extraction of scales

The following step as illustrated in Fig.1 is to decompose faces into different scales. This decomposition is performed using the Surfaces Empirical Mode Decomposition (SEMD) proposed by Wang et al. [30]. The decomposition of 3D face F allows us to get a smoothed residue R and several Intrinsic Mode Functions (IMFs) IMF_n encoding features at different scales (see Fig. 3).

$$F = \sum_{k=1}^{n} IMF_k + R \quad (1)$$

Where n is the number of IMFs

The decomposition of 3D faces is used for two reasons:

1- The more faces geometric information is separated; the more analysis and feature extraction get better.

2- The analysis of the faces at different scales allows to have more precision in the classification phase.

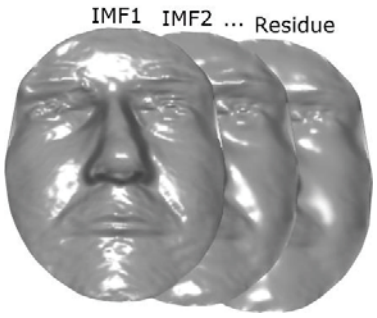

Figure 3: An example of the decomposition of 3D face scan by SEMD

2.3 Feature and curve extractions

Let S_{IMFn} be a facial surface of the IMF number n. S_{IMFn} is represented by N points. As noted previously, our recognition system uses local features and geometric information (see Fig. 4). Local features are obtained by associating to each point x of the IMFs a wave kernel signature vector (WKS) [5]. The WKS represents the time-averaged probability of measuring a quantum mechanical particle at a specific location on the 3D face. The signature reduces the influence of the low frequencies and allows better separation of frequency bands across the descriptor dimensions. This descriptor is based on the spectrum of the Laplace-Beltrami operator which makes it invariant under isometries like facial expressions. The WKS is defined as

$$WKS(x) = \left[wks_{e_1}(x), \dots, wks_{e_M}(x)\right]$$

where:

$$wks_e = \sum_{k\geq 1} f_e^2(\lambda_k)\emptyset_k^2(x) \quad (2)$$

λ_k and $\emptyset_k(.)$ are respectively the eigenvalues and eigenfunctions of the Laplace-Beltrami operator and f_e is a family of log-normal energy distributions.

For geometric information, we extract both radial and level facial curves from the 3D surface. The radial curve β_α on S_{IMFn} which makes an angle α with a reference radial curve is obtained by slicing the facial surface by a vertical plan P_α. The plan passes through the nose tip and makes an angle α with the plan of the reference curve. This step is repeated several times to extract radial curves from the facial surface with a separation angle α. The same methodology is used to extract level curves from IMFs surface. Let β_λ be the level curve obtained on S_{IMFn}. β_λ is the intersection of S_{IMFn} with a sphere G_λ, the curves makes a distance λ (radius of G_λ) from the nose tip chosen as the reference point. In order to extract a collection of level curves, the slicing procedure is repeated many times with different radius.

Before the classification stage, we associate the geometric information with local information by assigning to each point of the radial and level curves it corresponding wave kernel signature. β_α and β_λ are rewritten as follows $\beta_{\alpha,wks_\alpha}$ and $\beta_{\lambda,wks_\lambda}$.

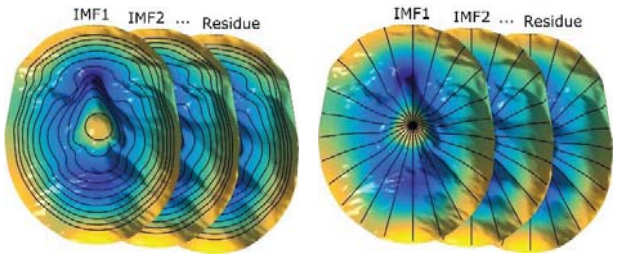

Figure 4: An example of extraction of local features and curves

2.4 Recognition

Let F_i and F_j denote two faces belonging to the probe and gallery respectively. F_i and F_j are represented in n scales thanks to the SEMD. We get the characteristics of each scale using radial and level curves associate with descriptors vectors wks. The

methodology used in this paper is based on the symmetrical comparison of the scales, it means that the scale number r of F_i is matched with the scale number r of F_j and so forth (see Fig. 5). Given $F_{i,r} = \{\cup_{r,\alpha} \beta_{\alpha,wks_\alpha}, \cup_{r,\lambda} \beta_{\lambda,wks_\lambda}\}$ and $F_{j,r} = \{\cup_{r,\alpha} \beta'_{\alpha,wks_\alpha}, \cup_{r,\lambda} \beta'_{\lambda_1,wks_\lambda})\}$ two set that characterizes F_i and F_j at r scale, respectively. In order to find the corresponding characteristics between $F_{i,r}$ and $F_{j,r}$ we perform a symmetrical comparison of curves. Each curve of $F_{i,r}$ is matched with its correspondent in $F_{j,r}$ using the angle as a measure of similarity. The angle formed between each pair of feature vectors v and u is calculated as follows:

$$\theta_{vu} = cos^{-1}\left(\frac{\langle v,u\rangle}{\|v\|\|u\|}\right) \quad (3)$$

v matches u if the ratio $\frac{\theta_{vu}}{\theta_{vw}}$ is greater than a threshold τ (in our experiments $\tau = 0.9$). $v \in \beta_{\alpha,wks_\alpha}$ and $(u, w) \in \beta'_{\alpha,wks_\alpha}$ with $\beta_{\alpha,wks_\alpha}$ and $\beta'_{\alpha,wks_\alpha}$ two curves of $F_{i,r}$ and $F_{j,r}$ respectively. θ_{vu} and θ_{vw} are the two smallest angles formed between v and all vectors of $\beta'_{\alpha,wks_\alpha}$ with $\theta_{vu} \leq \theta_{vw}$.

3 EXPERIMENTS AND RESULTS

To demonstrate the effectiveness of our algorithm for expression-invariant face recognition, we performed our experiments on GavabDB dataset [24].

The GavabDB database is one of the richest expressions and noise-prone 3D face datasets that is currently available to the public. It includes 3D face scans of 61 adult Caucasian individuals (45 males and 16 females). For each person, there are nine different images that differ in the acquisition viewpoint and facial expressions, which results in a total of 549 facial scans. For each individual, there are two frontal face scans with neutral expressions, four scans with large pose variations, e.g., one looking up scan (+35°), one looking down scan (-35°), one left profile scan (+90°) and one right profile scan (-90°) and a neutral facial expression, in which the person laughs, smiles, or shows a random expression.

In our experiments, we used the two neutral frontal images, the frontal images with a smile expression, the frontal images with a laughing expression, and the frontal images with a random gesture (see Fig. 6). The gallery includes for each subject the scan called "frontal1", and the remaining images are used as probe images for recognition.

Rank-1 recognition accuracies are reported in Table 1, separately for the IMFs with neutral vs. neutral, neutral vs. expression, and neutral vs. neutral+expression. From the table, it can be observed that the performance achieved by the Residue is higher than the IMF1 and IMF2, and the combination of all IMFs allows to have a recognition rate very high compared to the rates obtained by the IMFs separately. These results are mainly due to the way of faces are presented. The lower-order IMFs capture rapid oscillation modes of scans, while higher order IMFs generally represent slow spatial oscillation modes. The residue indicates the lowest frequency of the surface.

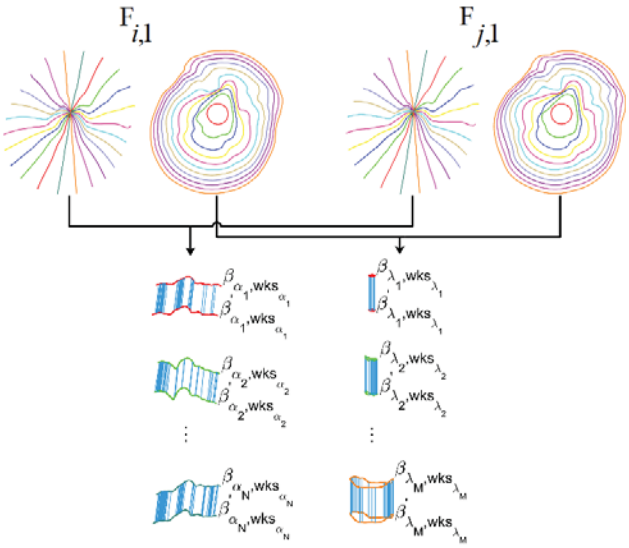

Figure 5: Illustration of classification stage

Figure 6: Examples of facial scans with different expressions in GavabDB

Table 1: Rank-1 recognition rates for different IMFs

	IMF1	IMF2	Residue	All
Neutral	93.44	95.08	95.08	100
Expression	87.43	86.34	89.61	98.90
Neutral+Expression	88.93	88.52	90.98	99.18

Table 2: Rank-1 recognition rates on GavabDB

Approaches	Neutral	Expression	Neutral+Expression
Moreno et al,2005 [25]	**90.16**	**77.90**	**-**
Li et al. 2009 [18]	**96.67**	**93.33**	**94.68**
Mahoor et al, 2009 [21]	**-**	**72.00**	**78.00**
Huang et al.2012 [12]	100	**93.99**	**95.49**
Drira et al.,2013 [9]	100	**94.54**	**95.90**
Berretti, et al, 2013 [7]	100	**94.00**	**95.10**
Lei, et al, 2016 [16]	100	**95.08**	**96.31**
Our method	100	98.90	99.18

We compared our method with recently published results, including Moreno et al. [25], Li et al. [18], Mahoor et al. [21], Drira et al. [9], Huang et al. [12], Berretti et al. [7] and Lei et al. [16]. Table 2 summarizes the evaluation using rank-1 RR. The results of the other methods are quoted from their papers. It can be observed that the proposed approach outperforms the existing approaches; more specifically, our proposed algorithm achieved 98.9% and 99.18%, which have the highest accuracy on rank-1 RR for the gallery scans with expression and neutral+expression, respectively.

4 CONCLUSIONS

In this work, we have proposed an original approach to 3D face recognition in the presence of both neutral and non-neutral facial expressions. Our method is based on the idea of capturing geometric and local information of the face surface at multiple scales. Experimental results compared with previous state-of-the-art methods on GavabDB dataset showed that our method consistently achieved high 3D face recognition performance.

REFERENCES

[1] ABATE, A.F., NAPPI, M., RICCIO, D., and SABATINO, G., 2007. 2D and 3D face recognition: A survey. *Pattern recognition letters 28*, 14, 1885-1906.

[2] ABBAD, A., ABBAD, K., and TAIRI, H., 2017. 3D face recognition: Multi-scale strategy based on geometric and local descriptors. *Computers & Electrical Engineering*.

[3] AL-OSAIMI, F., BENNAMOUN, M., and MIAN, A., 2009. An expression deformation approach to non-rigid 3D face recognition. *International Journal of Computer Vision 81*, 3, 302-316.

[4] AMBERG, B., KNOTHE, R., and VETTER, T., 2008. Expression invariant 3D face recognition with a morphable model. In *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on* IEEE, 1-6.

[5] AUBRY, M., SCHLICKEWEI, U., and CREMERS, D., 2011. The wave kernel signature: A quantum mechanical approach to shape analysis. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on* IEEE, 1626-1633.

[6] BALLIHI, L., AMOR, B.B., DAOUDI, M., SRIVASTAVA, A., and ABOUTAJDINE, D., 2012. Boosting 3-D-geometric features for efficient face recognition and gender classification. *IEEE Transactions on Information Forensics and Security 7*, 6, 1766-1779.

[7] BERRETTI, S., WERGHI, N., DEL BIMBO, A., and PALA, P., 2013. Matching 3D face scans using interest points and local histogram descriptors. *Computers & Graphics 37*, 5, 509-525. DOI= http://dx.doi.org/10.1016/j.cag.2013.04.001.

[8] BERRETTI, S., WERGHI, N., DEL BIMBO, A., and PALA, P., 2014. Selecting stable keypoints and local descriptors for person identification using 3D face scans. *The Visual Computer 30*, 11, 1275-1292. DOI= http://dx.doi.org/10.1007/s00371-014-0932-7.

[9] DRIRA, H., AMOR, B.B., SRIVASTAVA, A., DAOUDI, M., and SLAMA, R., 2013. 3D face recognition under expressions, occlusions, and pose variations. *IEEE Transactions on Pattern Analysis and Machine Intelligence 35*, 9, 2270-2283.

[10] FALTEMIER, T.C., BOWYER, K.W., and FLYNN, P.J., 2008. A region ensemble for 3-D face recognition. *IEEE Transactions on Information Forensics and Security 3*, 1, 62-73.

[11] FREUND, Y. and SCHAPIRE, R.E., 1995. A desicion-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory* Springer, 23-37.

[12] HUANG, D., ARDABILIAN, M., WANG, Y., and CHEN, L., 2012. 3-D face recognition using eLBP-based facial description and local feature hybrid matching. *IEEE Transactions on Information Forensics and Security 7*, 5, 1551-1565.

[13] HUANG, Y., WANG, Y., and TAN, T., 2006. Combining Statistics of Geometrical and Correlative Features for 3D Face Recognition, 90.91-90.10. DOI= http://dx.doi.org/10.5244/c.20.90.

[14] KAKADIARIS, I.A., PASSALIS, G., TODERICI, G., MURTUZA, M.N., LU, Y., KARAMPATZIAKIS, N., and THEOHARIS, T., 2007. Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence 29*, 4.

[15] LEI, Y., BENNAMOUN, M., and EL-SALLAM, A.A., 2013. An efficient 3D face recognition approach based on the fusion of novel local low-level features. *Pattern Recognition 46*, 1, 24-37. DOI= http://dx.doi.org/10.1016/j.patcog.2012.06.023.

[16] LEI, Y., GUO, Y., HAYAT, M., BENNAMOUN, M., and ZHOU, X., 2016. A Two-Phase Weighted Collaborative Representation for 3D partial face recognition with single sample. *Pattern Recognition 52*, 218-237. DOI= http://dx.doi.org/10.1016/j.patcog.2015.09.035.

[17] LI, X. and DA, F., 2012. Efficient 3D face recognition handling facial expression and hair occlusion. *Image and Vision Computing 30*, 9, 668-679. DOI= http://dx.doi.org/10.1016/j.imavis.2012.07.011.

[18] LI, X., JIA, T., and ZHANG, H., 2009. Expression-insensitive 3D face recognition using sparse representation. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* IEEE, 2575-2582.

[19] LOWE, D.G., 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision 60*, 2, 91-110.

[20] LU, X. and JAIN, A., 2008. Deformation modeling for robust 3D face matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence 30*, 8, 1346-1357.

[21] MAHOOR, M.H. and ABDEL-MOTTALEB, M., 2009. Face recognition based on 3D ridge images obtained from range data. *Pattern Recognition 42*, 3, 445-451.

[22] MIAN, A.S., BENNAMOUN, M., and OWENS, R., 2008. Keypoint detection and local feature matching for textured 3D face recognition. *International Journal of Computer Vision 79*, 1, 1-12.

[23] MING, Y., 2014. Rigid-area orthogonal spectral regression for efficient 3D face recognition. *Neurocomputing 129*, 445-457. DOI= http://dx.doi.org/10.1016/j.neucom.2013.09.014.

[24] MORENO, A.B. and SÁNCHEZ, A., 2004. GavabDB: a 3D face database. In *Proc. 2nd COST275 Workshop on Biometrics on the Internet, Vigo (Spain)*, 75-80.

[25] MORENO, A.B., SANCHEZ, A., VELEZ, J., and DIAZ, J., 2005. Face recognition using 3D local geometrical features: PCA vs. SVM. In *Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on* IEEE, 185-190.

[26] PHILLIPS, P.J., SCRUGGS, W.T., O'TOOLE, A.J., FLYNN, P.J., BOWYER, K.W., SCHOTT, C.L., and SHARPE, M., 2007. FRVT 2006 and ICE 2006 large-scale results. *National Institute of Standards and Technology, NISTIR 7408*, 1.

[27] SMEETS, D., KEUSTERMANS, J., VANDERMEULEN, D., and SUETENS, P., 2013. meshSIFT: Local surface features for 3D face recognition under expression variations and partial data. *Computer Vision and Image Understanding 117*, 2, 158-169. DOI= http://dx.doi.org/10.1016/j.cviu.2012.10.002.

[28] SRIVASTAVA, A., SAMIR, C., JOSHI, S.H., and DAOUDI, M., 2009. Elastic shape models for face analysis using curvilinear coordinates. *Journal of Mathematical Imaging and Vision 33*, 2, 253-265.

[29] SZEPTYCKI, P., ARDABILIAN, M., and CHEN, L., 2009. A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking. In *Biometrics: Theory, Applications, and Systems, 2009. BTAS'09. IEEE 3rd International Conference on* IEEE, 1-6.

[30] WANG, H., SU, Z., CAO, J., WANG, Y., and ZHANG, H., 2012. Empirical mode decomposition on surfaces. *Graphical Models 74*, 4, 173-183.

[31] ZAHARESCU, A., BOYER, E., VARANASI, K., and HORAUD, R., 2009. Surface feature detection and description with applications to mesh matching. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* IEEE, 373-380.

[32] ZHAO, W., CHELLAPPA, R., PHILLIPS, P.J., and ROSENFELD, A., 2003. Face recognition: A literature survey. *ACM computing surveys (CSUR) 35*, 4, 399-458.

[33] ZULIANI, M., KENNEY, C.S., and MANJUNATH, B., 2005. The multiransac algorithm and its application to detect planar homographies. In *Image Processing, 2005. ICIP 2005. IEEE International Conference on* IEEE, III-153.