A 3D ALGORITHM FOR UNSUPERVISED FACE IDENTIFICATION

A. Lagorio, M. Cadoni, E. Grosso, M. Tistarelli

VisionLab - Computer Vision Laboratory
University of Sassari, Italy

ABSTRACT

With the increasing availability of low-cost 3D data acquisition devices, the use of 3D face data for the recognition of individuals is becoming more appealing and computationally feasible. This paper proposes a completely automatic algorithm for face registration and matching. The algorithm is based on the extraction of stable 3D facial features characterizing the face and the subsequent construction of a signature manifold. The facial features are extracted by performing a continuous-to-discrete scale-space analysis. Registration is driven from the matching of triplets of feature points and the registration error is computed as shape matching score. Conversely to most techniques in the literature, a major advantage of the proposed method is that no data pre-processing is required. Therefore all presented results have been obtained exclusively from the raw data available from the 3D acquisition device.

The method has been tested on the Bosphorus 3D face database and the performances compared to the ICP baseline algorithm. Even in presence of noise in the data, the algorithm proved to be very robust and reported identification performances which are aligned to the current state of the art, but without requiring any pre-processing of the raw data.

***Index Terms*—** Face recognition, 3D Face recognition

1. INTRODUCTION

In recent years there has been an increasing interest in the development of 3D face recognition algorithms, leading to great improvements in recognition performance. 3D acquisition systems are becoming affordable, user friendly and easy to install in different environments. For these reasons it is envisaged that, in the near future, 3D face acquisition and matching can be successfully employed in different scenarios, from access control to forensic applications.

The recognition of 3D objects often involves the alignment of their shapes followed by a measure of the shapes similarity. Particularly with deformable objects, such as human faces, shape registration based on 3D or texture data can be very difficult due to ambiguities in the characterization of anchor points.

An accurate registration of the face shapes of two individuals, or of the face areas stable to expression, could readily provides a measure of their similarity, [1, 2, 3, 4]. In fact, one can define an appropriate registration error measure to be used a matching score between the two individuals. The Iterative Closest Point (ICP) algorithm [5] has extensively been used as a benchmark to evaluate the performances of face recognition algorithms. ICP proved to be very effective to accurately register (or match) 3D face scans, but an approximate initial alignment of the two point sets not only is required to bootstrap the algorithm, but it is mandatory to prevent the ICP to converge to a local minimum, i.e. to fail to reach the optimal solution. An accurate and efficient face pre-registration is therefore mandatory to perform face recognition based on scan registration. In this paper, 3D face identification is tackled as a by product of the registration of 3D point sets.

A number of 3D databases are now available to the scientific community, many of them consisting of high resolution scans of many individuals acquired with different poses and expressions (see [6] for an up to date list). This data allows to simulate real scenarios, where we need to take into account a high degree of variability. Therefore, face recognition algorithms should be able to handle densely sampled scans, with pose and expression variations, and from large populations of subjects. In this light, and given the rationale of matching 3D faces, an algorithm devised to register 3D face scans must be:

- Completely automatic, not requiring human intervention which can be prohibitive when dealing with thousands or millions of subjects.
- Robust to changes in poses and facial expressions. In a real scenario, especially for visual surveillance or in forensic applications, it may be difficult or not practical to require a subject to "pose" for a 3D scan to be taken.

In this paper a completely automatic algorithm for face identification is proposed. The algorithm is based on the extraction of facial features characterizing the face and the subsequent construction of a signature manifold. Registration is driven from the matching of triplets of feature points. After registration we propose two different methods: in the first, the matching score is given by the registration error,in the second, the registration obtained with the invariants is refined with the

978-1-4799-8105-2/15/$31.00 © 2015 IEEE

ICP and the matching score is defined as the registration error computed after the last iteration.

The proposed algorithm was tested on the Bosphorus database [7], which has a good variety of high resolution scans relative to different poses and expressions. Previous works on this database have concentrated on landmarks detection robust to occlusions and noise ([8, 9]). In [10], benchmark algorithms have been tested on selected subsets of the database. The algorithm proposed in this paper significantly outperforms the benchmarks algorithms based on automatic features extraction. The paper is organized as follows. In section 2 the feature extraction method is described. In section 3 the registration algorithm is detailed and in section 4 results of the tests run on the Bosphorus database are presented.

2. FEATURES EXTRACTION

2.1. Scale-space theory for 3D face analysis

Recognition of faces from 3D information only can be achieved by registering the data from two individuals and measuring the goodness of fit. This process requires to identify anchor points on the faces which are similar for all faces but also to locate 3D features which may be highly distinguishing.

Starting from the observation "all faces are similar and different at the same time" the aim is to localize points in areas that almost every face share, "common" points such as eyes corners, nose tip etc, and, at the same time, points that are peculiar to a face such as a chin dimple or a prominent cheekbone. The first kind of points present a certain degree of variability amongst faces which is useful to distinguish faces from different individuals. These points should be localized with the highest possible accuracy. Moreover, in order to compute the signature of an individual's face also the 3D surface normals at those points are required. Considering a 3D face scan as a smooth surface, both kind of points are either local maxima or minima of the Gaussian curvature. Our aim is then to find an algorithm to extract local maxima and minima of curvature, with a given approximation.

The scale-space theory [11], originally proposed to describe the gray level variations in 2D intensity images, can be applied to 3D face scan to optimally select all "common" points, namely 3D features, to be extracted from a set of 3D faces. According to this theory, a signal $f : \mathbb{R}^n \to \mathbb{R}$ (the face surface in our case) can be modeled by a scale-space representation: $L : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ where $L(x,t) = G(x,t) \otimes f(x)$, $G(\, ,t)$ is a Gaussian kernel of width t and $\otimes$ is the convolution operation [12]. Given a scale-space representation of the face, we can characterize the face at each scale by means of the Gaussian curvature at each point. Following this finding, we conjecture that the scale at which the scale space curvature reaches its maximum is likely to be a relevant scale to represent that patch of the face. This would imply that, by varying the scale, it is possible to localize all required points (common and peculiar) on a face. Furthermore, the surface normal computed at the same relevant scale is expected to be more robust to noise.

In order to adapt the scale-space theory to a face scan represented by a discrete set of points, a similar scheme as in [13] is adopted. Due to computational time and memory limits, the scale can not be varied continuously, nor can the cloud of points be model with a parameterized surface. This problem can be overcome by extracting, for each 3D scanned point p_i, an approximation of the Gaussian curvature computed on the set of spherical neighbors $N_{p_i}(r_j)$, centered at the point p_i and of increasing radius r_j. The scale step, i.e. the difference between the radii of two consecutive neighbors can be chosen on the basis of the sampling density of the scan. In the performed experiments the scale step was determined by constraining, on average, the difference between two neighbors to be equal to 10 points. Given a 3D point p_i and the 3D neighbors $N_{p_i}(r_j)$, an approximation of the Gaussian curvature can be obtained by computing the Principal Components of $N_{p_i}(r_j)$. The eigenvalues $\lambda_0 \leq \lambda_1 \leq \lambda_2$ and the respective eigevectors v_0, v_1, v_2 corresponding to the principal directions, are computed. The absolute value of the curvature is then defined as $\mathcal{C}(p_i, r_j) = \dfrac{2|(p_i - p_g) \cdot v_0|}{d_m^2}$, where p_g is the center of gravity of the neighbor $N_{p_i}(r_j)$ and d_m^2 is the mean of distances $|p_i - p_j|, \; p_j \in N_{p_i}(r_j)$. The surface normal $\nu(p_i, r_j)$ at the point p_i at scale r_j is computed as the principal direction corresponding to the smallest eigenvalue λ_0.

2.2. Multi-scale feature extraction

The scale-space analysis of the 3D face scans can be summarized as follows:

- Two extreme values for the search radius are set, i.e. a starting radius r_s and an end radius r_e. It is worth noting that these two values are metric parameters which are fixed on the basis of anthropometric facial measures. Therefore they do not depend on the training data nor on the acquisition device. For example, the smallest radius r_s should be small enough to detect the nose of a child, while the largest radius r_e should be large enough to detect an adult nose. In the experimental tests the two radii were set empirically to 6mm and 22mm.
- The scale step σ_s is defined to partition the interval (r_s, r_e) into a set of $n_\sigma = \dfrac{r_s - r_e}{\sigma_s} + 1$ intervals of equal length.
- For each point p_i of a face scan, the curvature $\mathcal{C}(p_i, r_j)$ is computed for $j = s, s + \sigma_s, s + 2\sigma_s, \ldots, e$. The curvature values are then interpolated to produce a func-

tion $\mathcal{C}(p_i) : [s, e] \to \mathbb{R}$. A median filter is applied to smooth the curve, and the scale $\sigma_m(p_i)$ for which the curvature $\mathcal{C}(i) = \mathcal{C}(p_i, \sigma_m(p_i))$ reaches a maximum is computed. Should the maximum correspond to the first scale (r_s), $\sigma_m(p_i)$ is set to be the scale at which the curvature is equal to the median value of all curvatures. This is necessary because often the maximum is a consequence of noise which can highly affect the processing at small radius scales. The normal ν_i at point p_i is determined as $\nu(p_i, \sigma_m(p_i))$. Three sample graphs of the function $\mathcal{C}(p_i) : [s, e] \to \mathbb{R}$ are shown in figure 1. The three curves have been drawn respectively from the tip of nose (blue curve), the inner left eye corner (red curve) and the right nostril (green curve).

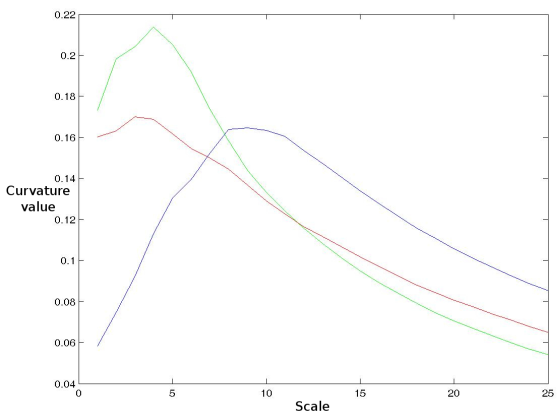


Fig. 1. Curvature computed at selected positions on the face as a function of scale.

As a result, for each point p_i of the face scan an optimal curvature value $\mathcal{C}_i$ and an optimal normal vector $\nu(i)$ are obtained. In order to avoid detecting the face edges as local maxima, all points belonging to the border of the face scan are first detected and marked to be excluded from the successive processing. Given $r = (r_e - r_s)/2$, and for each p_i in the face scan, p_i is defined to be a local maxima or minima of the curvature if $|\mathcal{C}_i|$ is the largest of all $|\mathcal{C}_k|$ for $p_k \neq p_i \in N_{p_i}(r)$. The extracted extrema curvature points are retained as 3D face feature points.

While the number of features is naturally bounded by the radius r, up to 12 points of highest curvature are selected amongst them. This value was experimentally proved to be sufficient to include a sufficient number of common and distinguishing points to characterize and match a 3D face. In figure 2 the projected surface of a sample 3D face scan is shown. The surface color encodes the curvature values computed at the fixed scale $(r_e - r_s)/2$. The marked points on the surface represent the extracted 3D features. As it can be noticed, the extracted features include stable points such as the nose tip, the eye corners as well as distinguishing points for this face such as the chin dimples and the little bump on the nose.

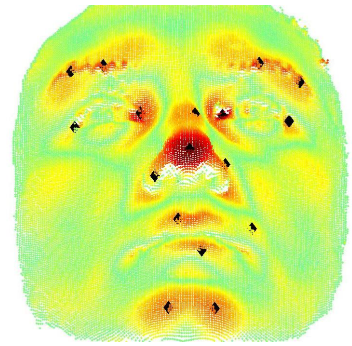


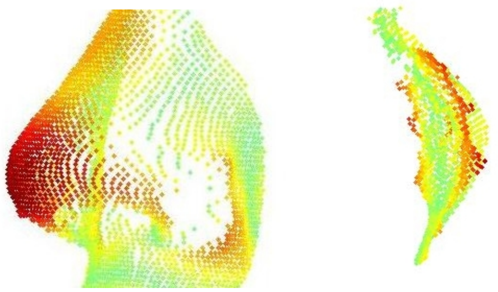
Fig. 2. Feature points extracted from a sample face scan from the Bosphorus database.

Even though from the projection of the sample face scan in figure 2 the point sampling appears to be quite uniform, this is not the case. In fact, this is just a visual artifact due to the coincidence of the direction of view with the direction of the original scanning. Most of the times the face scan is non-uniformly sampled just in the areas corresponding to the curvature variations. This non-uniform sampling may lead to occlusions and thus impair the extraction of feature points. For example, the nostril on the right hand side of the surface in figure 2 is located slightly upwards with respect to the left one. This is not due to an anatomical asymmetry or to an error in the curvature computation, but rather to a missing patch in the nostril area (see figure 3(a)). Another example of errors in the sampled points is shown in figure 3(b). In this case the eye area contains spurious points which are detected as spikes on the surface. Probably due to the specular reflectance of the cornea all facial scans of the Bosphorus database contain noise peaks within the areas including the eyes. Despite the occlusions and noise in the data, preprocessing of the data has been carefully avoided. It is worth stressing that all results presented in the experimental section were obtained without applying any kind of data preprocessing. This allows to better evaluate the performance on face registration and matching as related to the raw data only and not to the quality of any preprocessing step.

3. 3D FACE REGISTRATION

The registration algorithm is based on the Moving Frame Theory [14]. The procedure that leads to the generation of the invariants and the signature are discussed in full detail in [1]. Only the fundamental issues are discussed here.

Given a surface F, the Moving Frame Theory defines a framework (and an algorithm) to calculate a set of invariants, say $\{I_1, \ldots I_n\}$, where each I_i is a real valued function that depends on one or more points of the surface. By construction, this set contains the minimum number of invariants



(a) (b)

Fig. 3. Enlarged view of the details from the sample face scan in figure 2, showing the effects of noise in the distribution of the 3D points. (a) Occlusion produced in the nostril area of the nose. (b) Local spikes in the eye area.

that are necessary and sufficient to parametrize a "signature" $S(I_1, \ldots, I_n)$ that characterizes the surface up to Euclidean motion. The framework offers the possibility of choosing the number of points the invariants depend on, and this determines both the number n of invariants we get and their differential order. The more the points the invariants depend on the lower the differential order. For instance, invariants that are functions of only one point varying on the surface ($I = I(p)$, $p \in F$) have differential order equal to 2. These are the classical Gaussian and Mean curvatures. In order to trade the computational time with robustness to noise the invariants are built depending on three points at one time. The result is a set of nine invariants, three of differential order zero, and six of order one.

3.1. 3-Points invariants

Let p_1, p_2, $p_3 \in F$ and ν_i be the normal vector at p_i. The directional vector v of the line between p_1 and p_2 and the normal vector ν_t to the plane through p_1, p_2, p_3, are defined as:

$$v = \frac{p_2 - p_1}{\|p_2 - p_1\|} \qquad \text{and} \qquad \nu_t = \frac{(p_2 - p_1) \wedge (p_3 - p_1)}{\|(p_2 - p_1) \wedge (p_3 - p_1)\|}.$$

The zero order invariants are the 3 inter-point distances $I_k(p_1, p_2, p_3)$ for $k = 1, 2, 3$:

$$I_1 = \|p_2 - p_1\|, \quad I_2 = \|p_3 - p_2\| \qquad \text{and} \qquad I_3 = \|p_3 - p_1\|$$

whereas the first order invariants are

$$J_k(p_1, p_2, p_3) = \frac{(\nu_t \wedge v) \cdot \nu_k}{\nu_t \cdot \nu_k} \qquad \text{for} \quad k = 1, 2, 3$$

and

$$\tilde{J}_k(p_1, p_2, p_3) = \frac{v \cdot \nu_k}{\nu_t \cdot \nu_k} \qquad \text{for} \quad k = 1, 2, 3.$$

Each triplet (p_1, p_2, p_3) on the surface can now be linked with a point of the signature in 9-dimensional space whose coordinates are given by $(I_1, I_2, I_3, J_1, J_2, J_3, \tilde{J}_1, \tilde{J}_2, \tilde{J}_3)$.

3.2. Registration of face pairs

For each triplet of feature points extracted from a sample face scan F the invariants are computed and stored into a signature S that characterizes F. Given a test scan F', the same procedure is applied to obtain another signature S'. The two face scans can be compared by computing the intersection between the two signatures S and S'. If the intersection between S and S' is not null, then exists a subset of feature points belonging to the two scans holding the same properties, i.e. the same inter-point distances and normal vectors (up to Euclidean motion). The signature points are compared by computing the Euclidean distance: given a threshold ϵ, if $s \in S$, $s' \in S'$ and $|s - s'| \leq \epsilon$, then the triplets that generated the signature points are matched. From the triplets the roto-translation $(\mathbf{R}, \mathbf{t})$ that takes the second into the first can be computed. Given $\{t_1, \ldots, t_m\}$ the set of triplets of the face scan F that are matched to the triplets in S', each matched triplet generates a roto-translation $(\mathbf{R}_i, \mathbf{t}_i)$. To select the best registration parameters among those computed, each $(\mathbf{R}_i, \mathbf{t}_i)$ is applied to F', so that $F'' = \mathbf{R}F' + \mathbf{t}$ and the registration error is computed according to the following procedure.

For each point $q_i \in F'$ the closest point p_i in F is computed together with the corresponding Euclidean distance $d_i = \|q_i - p_i\|$. A set of distances $D = \{d_i\}_{i \in I}$ is obtained where I is the cardinality of F'. The registration error is defined to be the median of $D = \{d_i\}_{i \in I}$. The pair $(\mathbf{R}_m, \mathbf{t}_m)$ corresponding to the minimum registration error d_m is chosen as the best registration between the two faces. It might happen that the scans are so different that the registration step fails (there are no matching points in the signature space and so triplets). In this case the result is accounted as a negative match.

3.3. Classification

The registration error defined above can be used as matching score between two faces F and F'. Given a set of gallery faces G and a probe F, we could simply match face F to the face in G that, after registration, has the minimum registration error. We should keep into consideration though, that the input data might not be reliable enough for the feature points to be calculated accurately, simply because often the same point is not present in two scans of the same subject due to occlusions (see figure 3). Also, big variations in sampling density might lead to a slight displacement of a feature point. This will lead to a coarse registration that, if refined, would lead to a smaller registration error. In light of this, the feature extraction and subsequent registration through invariants can be thought of an automatic coarse registration of faces, to be followed by a refinement. We chose to use ICP to refine the registration. In the first iteration ICP will take as input the two scans aligned through invariants. The registration error after the last iteration will be the matching score.

4. EXPERIMENTAL RESULTS

The proposed algorithm was tested on the Bosphorus database [7]. The database contains scans of 105 individuals, of which 61 male and 44 female subjects. From the total male subjects, 31 male subjects have a beard and mustaches. For each subject there are about 50 scans. Each scan either presents a different facial expression (anger, happiness, disgust), corresponding to a "Face Action Unit", or a head rotation along different axes. Examples of the scans for two subjects are shown in figure 4. The pictures represent, left to right and up down: neutral, slight rotation down, slight rotation up, $10°$ on the right, lip corner puller, happiness.

Fig. 4. Sample images for two subjects in the Bosphorus database.

For each pose, the data points are stored in a file containing the coordinates of about 30,000 3D points, a color 2D image of the face texture and a set of landmark points. The landmarks were manually selected on the 2D images and mapped on the corresponding 3D points. This database has been chosen because it contains a large number of subjects and an excellent variety of poses and expressions. Furthermore, the available landmark points constitute a ground truth which makes it possible to compare the methodology with a baseline algorithm such as ICP.

The database was divided into a gallery set and a probe set. The gallery consists of one neutral face scan for each individual. Different tests were run with probes containing variations in pose and expressions. In particular, all poses and expression showed in figure 4 were considered. In all tests described below, the gallery consisted of the neutral pose whereas the probes were taken in the same order as in figure 4 and are detailed below.

- **Test 1**. 105 faces labeled in the database as PR-SD, which consists of a slight rotation of the face downwards as shown in the second figure of 4, were matched against a gallery of neutral faces of the same individuals.
- **Test 2**. 105 faces labeled as PR-SU, consisting of a slight rotation of the face upwards as shown in the third figure of 4, were match against neutral poses.
- **Test 3**. 105 faces labeled as YR-R10, which consists of a slight rotation of the face on the right were matched against the neutral poses.
- **Test 4**. 105 faces labeled as LFAU-12 which consists of a timid smile (see 5th image in figure 4) were matched against neutral poses.
- **Test 5**. 105 faces labeled as E-HAPPY, consisting of a full smile as shown in the last image in figure 4.

Following experiments 2 and 3, the best registration obtained with the invariants method between each probe and gallery face was refined with ICP. Furthermore, experiment 2 was compared to the followed baseline algorithm: on each probe and gallery face the landmarks corresponding to the two inner eye corners and the nose tip were selected from the available list of 3D landmarks of the database and a coarse registration performed using the three points correspondences. The coarse manual registration was then refined with ICP. In this way the automatic features selection and registration is compared with the manual operation.

The results of the tests are summarized in table 1. *F.R.*, standing for failed registrations, is the number of subjects for which the registration failed (after features extraction no triples were matched in the signature space). These numbers are indicative of the robustness of the method, since if a registration fails there is no later chance of refinement. As it can be seen from table 1, in tests 2 and 3 no failures occurred whereas only one occurred in test 1. The registrations in test 4 and 5 proved to be more difficult because of the great expression variations.

Table 1. Identification Rates

Experiment	F.R.	Inv
Test 1	1	77%
Test 2	0	89%
Test 3	0	72%
Test 4	3	53%
Test 5	4	45%

In the third column of table 1, *Inv* indicates identification rates (number of correct identifications over total number of subjects) obtained using as matching score the registration error that follows from the automatic feature extraction and the registration through invariants. Considering that no preprocessing was applied to the raw data, the performance reported for the tests where only the head rotation is involved (1, 2 and 3) are quite encouraging. As for test 4 and 5, it should be noticed that the matching score is given by the "overall distance" of the two registered scans, which is a measure not at all adequate in case of facial expressions. In fact this observation is reinforced from the result of the 3rd column of 2. In this case the registrations were refined with ICP and the matching score (*Inv+ICP*) is given by the registration error at the end of the last iteration. In the case of test 5, any improvement can be hardly noticed, which confirms that such matching score is weak in case of changes in facial expression. On the contrary, refining the coarse registration in the case of similar facial expressions improves the result (lines corresponding to tests 2 and 3 in table 2).

Table 2. Identification Rates

Experiment	Inv	Inv+ICP	Man+ICP
Test 2	89%	98.1%	99%
Test 3	72%	93%	-

The last column of table 2 shows the result of the manual registration followed by ICP in the case of the faces slightly rotated upwards. This baseline algorithm failed to recognize one of the two individuals that the automatic registration followed by ICP failed. So, manually registering the scans and then applying ICP hardly improved the result.

In figure 5, the Cumulative Matching Characteristic Curves (CMC) are drawn for test 2. The invariant recognition method improves significantly with the rank (at rank 7 is over 96%), whereas the same test followed by ICP does not improved at rank 15. The manual+ICP baseline algorithm reaches 100% accuracy at rank 4.

5. CONCLUSIONS

The identification of individuals on the basis of 3D shape information only has been addressed. This is a very promising and challenging biometric technology at the same time, because of the difficulties in processing three-dimensional data and of the advantages such as the relative insensitivity to illumination changes. In this framework, the main original contribution of this paper is the application of the scale-space theory for the extraction of relevant features from discrete 3D face scans aimed at the construction of geometrical 3D invariants to perform face recognition. Within this context, a single signature for a cloud of 3D facial points is generated using robust joint differential invariants. Even though the Scale-space analysis and the invariant signature are defined for continuous surfaces, they have been successfully adapted to the discrete set of 3D face points.

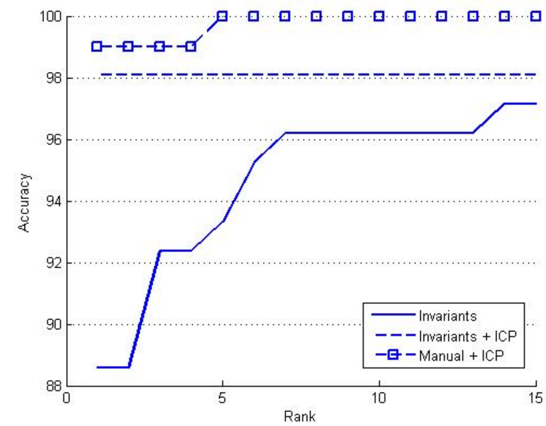


Fig. 5. CMC curve obtained by the tests performed according to the three recognition algorithms described.

A major drawback of 3D face scans is the presence of noise, which generally needs heavy data pre-processing to obtain reliable information for matching. Even though it can not be claimed to be *noise-free*, the developed methodology, based on the scale-space theory, proved to be very robust and can provide very good performances in terms of matching accuracy, without requiring any data pre-processing to either fill-in holes or smooth the face surface to remove spikes within the points cloud. An efficient procedure for the extraction of stable 3D feature points and the generation of an invariant signature in 9-dimensional space was presented, suitable to be employed for registration-based matching. It can be envisioned that even a light data pre-processing, e.g. cropping the central part of the face to remove spikes due to hair or acquisition artifacts and the choice of a more robust surface similarity measure such as the Surface Intepenetration measure ([3]) would positively affect the results further improving the matching accuracy. The method proved to be effective with non neutral input data, particularly with moderate rotations of the faces. This feature can be exploited in scenarios which assume non cooperative subjects as it often happens in forensic applications.

Acknowledgments

This work has been partially supported by the European Commission COST Action IC1106.

6. REFERENCES

[1] Marinella Cadoni, Manuele Bicego, and Enrico Grosso, "3d face recognition using joint differential invariants,"

Advances in biometrics, pp. 279–288, 2009.

[2] Marinella Cadoni, Andrea Lagorio, and Enrico Grosso, "Augmenting sift with 3d joint differential invariants for multimodal, hybrid face recognition," *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*, pp. 1–6, 2013.

[3] Chaua C Queirolo, Luciano Silva, Olga RP Bellon, and Mauricio Pamplona Segundo, "3d face recognition using simulated annealing and the surface interpenetration measure," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 2, pp. 206–219, 2010.

[4] Ajmal S Mian, Mohammed Bennamoun, and Robyn Owens, "An efficient multimodal 2d-3d hybrid approach to automatic face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 11, pp. 1927–1943, 2007.

[5] N.D. McKay P.J. Besl, "A method for registration of 3-d shapes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 14, no. 2, pp. 239–256, 1992.

[6] "http://www.face-rec.org/databases/," .

[7] Arman Savran, Neşe Alyüz, Hamdi Dibeklioğlu, Oya Çeliktutan, Berk Gökberk, Bülent Sankur, and Lale Akarun, "Bosphorus database for 3d face analysis," *Biometrics and Identity Management*, pp. 47–56, 2008.

[8] Oya Celiktutan, HC Akakin, and Bülent Sankur, "Multi-attribute robust facial feature localization," *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*, pp. 1–6, 2008.

[9] Hamdi Dibeklioglu, Albert Ali Salah, and Lale Akarun, "3d facial landmarking under expression, pose, and occlusion variations," *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, pp. 1–6, 2008.

[10] Neşe Alyüz, Berk Gökberk, Hamdi Dibeklioğlu, Arman Savran, Albert Ali Salah, Lale Akarun, and Bülent Sankur, "3d face recognition benchmarks on the bosphorus database with focus on facial expressions," *Biometrics and Identity Management*, pp. 57–66, 2008.

[11] Tony Lindeberg, "Feature detection with automatic scale selection," *International journal of computer vision*, vol. 30, no. 2, pp. 79–116, 1998.

[12] Andrew P. Witkin, "Scale-space filtering: A new approach to multi-scale description," *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84.*, vol. 9, pp. 150–153, 1984.

[13] Mark Pauly, Richard Keiser, and Markus Gross, "Multi-scale feature extraction on point-sampled surfaces," *Computer graphics forum*, vol. 22, no. 3, pp. 281–289, 2003.

[14] Peter J. Olver, "Joint invariant signatures," *Foundations of computational mathematics*, vol. 1, no. 1, pp. 3–68, 2001.