Computers and Electrical Engineering 70 (2018) 525–537

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

3D face recognition: Multi-scale strategy based on geometric and local descriptors☆

Abdelghafour Abbad[a,*], Khalid Abbad[b], Hamid Tairi[a]

[a] LIIAN, Department of Computer Science, Faculty of Sciences Dhar El Mahraz University Sidi Mohamed Ben Abdelah, Fez, Morocco
[b] LISA, Department of Computer Science, Faculty of Science and Technology University Sidi Mohamed Ben Abdelah, Fez, Morocco

ARTICLE INFO

Article history:
Received 6 November 2016
Revised 18 August 2017
Accepted 21 August 2017
Available online 31 August 2017

Keywords:
3D face recognition
Geometric features
Local features
Facial curves
Expression

ABSTRACT

Most human expression variations cause a non-rigid deformation of face scans, which is a challenge today. In this article, we present a novel framework for 3D face recognition that uses a geometry and local shape descriptor in a matching process to overcome the distortions caused by expressions in faces. This algorithm consists of four major components. First, the 3D face model is presented at different scales. Second, isometric-invariant features on each scale are extracted. Third, the geometric information is obtained on the 3D surface in terms of radial and level facial curves. Fourth, the feature vectors on each scale are concatenated with their corresponding geometric information. We conducted a number of experiments using two well-known and challenging datasets, namely, the GavabDB and Bosphorus datasets, and superior recognition performance has been achieved. The new system displays an overall rank-1 identification rate of 98.9% for all faces with neutral and non-neutral expressions on the GavabDB database.

1. Introduction

Face recognition using 2D images is still a large challenge due to many types of adverse factors, such as changes in the lighting or background, posing changes, and facial expressions. In recent years, many researchers have committed to using three-dimensional information to overcome some of these difficulties. Although 3D face data have more information than 2D images, facial expressions still remain a major challenge for 3D face recognition. They cause non-rigid distortions of the facial surface, increase the intra subject differences and affect the performance of face recognition.

1.1. Related work

In this literature review, the 3D face recognition approaches that address the expression challenge can be categorized into four categories: rigid, non-rigid, geometric form and keypoints detection.

(i) **Rigid matching algorithms**: extract the rigid areas of the face (e.g., eyes, forehead and nose). The partial ICP method [1] 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. [2] developed a binary mask to sub-divide

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Area Editor Dr. E. Cabal-Yepez.

* Corresponding author.
E-mail address: gh.abbad@gmail.com (A. Abbad).

http://dx.doi.org/10.1016/j.compeleceng.2017.08.017
0045-7906/

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. Ming [3] utilized curvature information to separate the rigid areas from the whole face. Then, the corresponding 3D depth image and the orthogonal spectral regression are employed to describe the geometric information.

(ii) **Non-rigid matching algorithms**: eliminate the information related to expression and, at the same time, retain the facial characteristics. Al-Osaimi et al. [4] 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. Amberg et al. [5] use a multilinear model to separate the identity and expression variations. In the fitting process, the offsets from the neutral pose are computed. Then, a PCA model is performed over these offset surfaces, and a variant of the non-rigid ICP algorithm is used to find the closest point pairs, the identity deformation, and the expression deformation.

(iii) **Geometric form matching algorithms**: focus on the geometric shape analysis. Drira et al. [6] 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 [7], 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. [8] 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. Moreno et al. [9] calculated thirty local geometric features that contain the distance, angle, area, curvatures, and centroid to model the 3D facial surfaces. Then, PCA and SVM are deployed for classification. Using similar features, Li et al. [10] characterize each face by a set of registered low-level geometric features (e.g., curvatures or the length of each edge). Then, a feature pooling and ranking algorithm was used to select the features that are robust to the expressions. In [11], Mahoor and Abdel-Mottaleb encode the range data of a 3D face into 3D ridge lines, which describe the locations of the facial regions that contain rich identity information such as the eyes, nose, and mouth. Then, two tools are applied: robust Hausdorff distance (HD) and iterative closest points (ICP), which were utilized for matching the 3D curve points between the ridge images.

(iv) **Keypoints detection matching algorithms**: extract salient point features on the face scans. Mian et al. [12] 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. [13], facial depth images are first represented by multi scale extended Local Binary Patterns (eLBP), and the Scale Invariant Feature Transform (SIFT) framework is then performed on these images to find keypoints, extraction of local features, and matching faces. Recently, Berretti et al. [14] identified first the keypoints from a 3D face based on the meshDOG keypoints 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. Berretti et al. [15] used the meshDOG as a detector to capture the local information of the face surface. After the keypoints detection, different local descriptors are extracted at each keypoint and are used to compare faces during the match. Lei et al. [16] represented a 3D face with a set of local geometrical descriptors called Keypoint based Multiple Triangle Statistics (KMTS) to address the problem of 3D partial face recognition, and then, a Two-Phase Weighted Collaborative Representation Classification (TPWCRC) framework was proposed to address the single training sample problem. In [17], Li et al. detected salient points in 3D faces by maximum and minimum curvatures estimated in the 3D Gaussian scale space. Then, the local region around each salient point is described by three quantities: the histogram of the mesh gradient (HoG), the histogram of the shape index (HoS) and the histogram of the gradient of the shape index (HoGS).

The rigid methods are very simple to implement and provide good results when considering the tradeoff between the accuracy and the time consumption. These types of methods overcome the problem of facial expressions if the rigid part is well extracted, but this is not the case. The face topology does not allow us to have a good extraction of the rigid part. In addition, the rigid part contains some relevant characteristics, but not all of them; these methods completely neglect the other part (non-rigid), which causes a loss of information. Non-rigid methods are among the solutions that are more effective for removing the distortions that are caused by facial expressions. It has the capacity to archive very high rates even in the case of expression vs neutral. However, they also suffer from some problems that are related to the modeling of facial expressions. Most of these techniques require at least two scans for each person in the training phase. In reality, this information is not always available, which limits the applications of these methods. Geometric methods are very fast compared to other methods, and they do not require much calculation because they focus on the facial shape. However, this type of method extracts a large number of forms to overcome the lack of local information. Methods that extract keypoints from the face are the closest methods to real-world applications, and they can surpass almost all problems in recognition,

such as missing parts, occlusions and data corruption. However, the time required by most of these methods to extract keypoints is very high, and the lack of geometric information causes several false matches at the classification level.

1.2. Contributions and organization

To address the limitations reported in the literature review (Section 1.1), this paper proposes a novel and simple yet effective approach for robust 3D face recognition. Our contributions can be summarized as follows:

- An original adaptation of the Wave Kernel Signature (WKS) [18] method to the case of face meshes used as a local descriptor. WKS is used in the field of spectral geometry for different shape classifications, whereas its direct use on the similar objects such as face recognition is not suitable. In the case of different object shapes, WKS is characterized by two important parameters: (1) discriminative and (2) invariant to deformations. It is discriminative for the reason that identical points on the two shapes are assigned with similar signatures. and points not in correspondence have various signatures. This descriptor is based on the spectrum of the Laplace-Beltrami operator (LBO), which makes it invariant to deformations. However, in the case of face scans, the first characteristic is lost because all of the models are similar and symmetric, and as a result, a large number of incorrect matching points exist in the matching stage even if it is the same person.
- Based on empirical mode decomposition (EMD) [19], this work contributes an original solution to complex geometric details of scans. Taking full advantage of the EMD modeling scheme of face geometry, our system benefits from supplementary information that corresponds to the details of face coding at each scale.
- A new matching step designed to measure similarities between the gallery and probe scans. This step works in a hybrid way by exploiting both local and geometric features. This combination is made because the association of the points corresponds to the radial and level curves with their local information in WKS (Section 2.3).

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.

The remainder of the paper is organized as follows: In Section 2, we describe the details of the proposed method. The experimental results and discussion are presented in Section 3. The paper's conclusions are stated in Section 4.

2. Proposed approach

The different stages and components of our method are laid out in Fig. 1. The first stage is data preprocessing. The preprocessing module includes face detection, nose tip detection, face cropping and pose correction. In the second stage, each 3D face is decomposed into different scales using Surfaces Empirical Mode Decomposition (SEMD) [19]. The third stage uses feature extraction based on the Wave Kernel Signature [18] to associate a descriptor vector at each point. In the fourth stage, level and radial curves are extracted from each face and are combined with their corresponding feature vectors, which are extracted in the third stage. Finally, we use the feature vectors associated with the level and radial curves for classification and recognition, with the angle between feature vectors being employed as the similarity measure.

2.1. Preprocessing

The quality of a 3D face is related to the precision with which the information has been taken in the real scene. In general, laser scanner devices produce side effects, such as holes, especially in hairy areas of the face (eyebrows, beard, moustache) and parts of the eyes. Moreover, in the majority of cases, the laser sensor captures, along with the 3D face, some undesired parts, such as clothes, hair, ears and neck. For all of these reasons, our idea starts by preprocessing the input facial images to improve their quality. 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. [20]. The holes produced during the 3D face scans are detected based on the number of neighbors of each vertex. The vertexes with less than 8 neighbors are considered to be holes and are filled by square surfaces. The nose tip is detected by calculating the Mean and Gaussian curvature of the scans. Next, the HK-Classification labels each vertex into basic geometric shape classes (elliptical concave, elliptical convex, hyperbolic concave, hyperbolic convex) based on the curvature values calculated. The nose tip is considered to be the point with the maximum Gaussian curvature in the convex regions. Finally, the face area on each scan is obtained automatically by using a sphere centered at the nose tip with a radius of 100 mm. To correct for 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) algorithm.

2.2. Extraction of scales

As indicated earlier, our goal is to analyze the shapes of the facial surfaces using a local descriptor and facial curves. This analysis is based on the quality of the information provided by the scans in 3D. Using filters is one of the most common

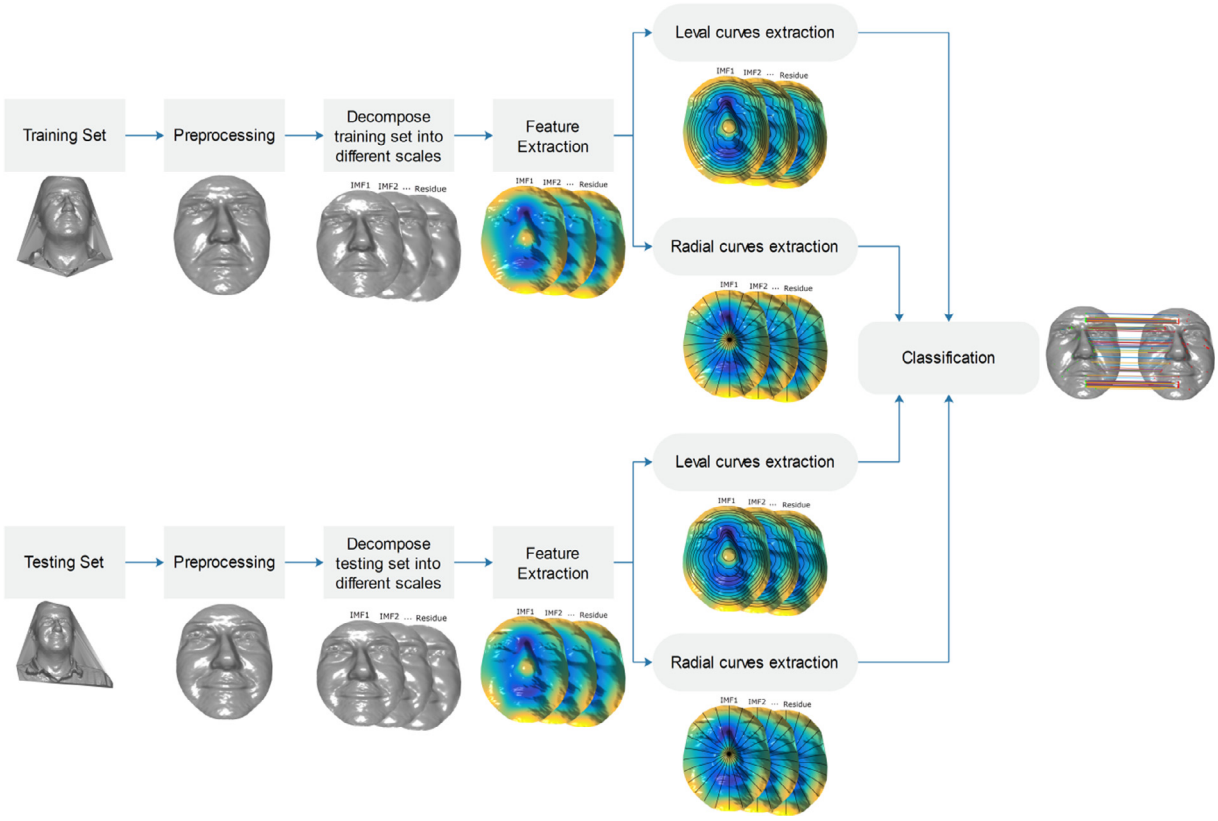

Fig. 1. Overview of the proposed method.

techniques in face recognition to increase the quality of the shapes; however, filters in general do not preserve the small details of the face such as the thickness of the eyebrows, the curvature of the nose, and the curvature of the corners of the eyes. This detail is lost because the noise follows a non-uniform distribution; thus, it is impossible to determine an appropriate filter threshold for the entire 3D face and for all of the faces in the database. As a solution to this problem, we proposed in this paper to decompose the 3D face at different scales using Surfaces Empirical Mode Decomposition (SEMD) [19]. These scales are classified from high frequencies to low frequencies. The new representation of the face information encodes also the noise according to its frequency on these scales.

The Empirical Mode Decomposition (EMD) has received much attention in terms of applications, analysis and data processing in 1D and 2D, but a limited number of studies have been proposed in the context of 3D surface decomposition in recent years. The extension of these methods for surfaces is based on the same stages of the basic EMD algorithm. Wang et al. [19] generalized classical EMD to surfaces, by proposing a new envelope computation method, which is the main difference between 1D EMD and 3D EMD for surfaces. The envelopes from local extrema in 1D EMD are computed by the cubic spline interpolation method, while the interpolation method on the surface such as the face in 3D is computed by minimizing a linearized thin-plate energy.

Unlike other decomposition methods, SEMD is empirical, direct, and adaptive, without pre-determined basis functions. The method decomposes iteratively any surface S into multiple intrinsic mode functions (IMFs) and a residue. The IMF components are obtained from S with an algorithm known as the sifting process. This algorithm subtracts the large-scale features of the signal iteratively until only the fine-scale features remain. The lower-order IMFs capture rapid oscillation modes of S, while higher order IMFs generally represent slow spatial oscillation modes. The residue indicates the lowest frequency of the surface S. The fundamental steps of the SEMD methods are given in Algorithm 1.

The decomposition of 3D face F through SEMD allows us to obtain a smoothed residue R and several Intrinsic Mode Functions (IMFs) IMF_n encoding features at different scales (Fig. 2).

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

where n is the number of IMFs.

Algorithm 1. SEMD algorithm.

Input: S: the surface
Output: IMF_k and R: multiple intrinsic mode functions and a residue, respectively
BEGIN

Step 1: Initialize the residue $R = S$
Step 2: Determine the local maxima and minima of S;
Step 3: Generate the upper and lower surface envelope by connecting those local maxima and minima, respectively, by an interpolation method; The values of each of the vertices in each envelope Q are the solution of the linear system.
$L^2Q = 0$
where L is a Laplace–Beltrami operator discretized as a sparse matrix.
Step 4: Determine the local mean m_1 by averaging the upper and lower signal envelope;
Step 5: Subtract the local mean from S:

$$h_1 = S - m_1$$

Step 6: If h_1 obeys the stopping criteria, then we define $d = h_1$ as an IMF; otherwise, we set $S = h_1$ and repeat the process from Step 1. Then, the empirical mode decomposition of the input surface S can be written as follows:

$$S = \sum_{i=1}^{k} IMF_i + R$$

where k is the number of extracted IMFs, and R is the final residue

END

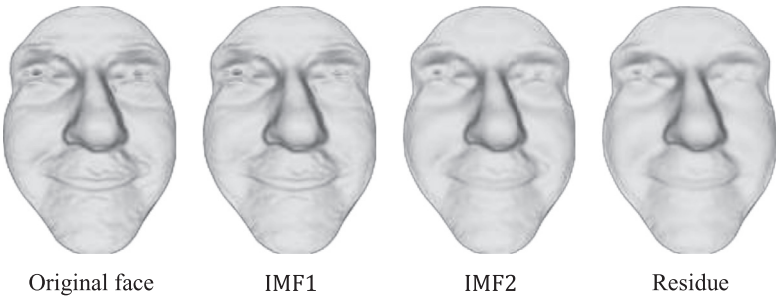

Fig. 2. The decomposition of 3D face scans by SEMD.

2.3. Feature and curve extractions

We considered that the 3D face is divided into two parts: a rigid part such as the forehead and the nose and a non-rigid part such as the mouth and the facial expressions. For the rigid part, the most robust methods are geometric methods, but for the non-rigid parts, the methods based on the extraction of local features are very promising. The exact separation between these parts in the face is almost impossible, it is very close to the problems of fuzzy logic. In this work, we will consider in the first place that the face is not rigid, and then, we will associate with each vertex a vector descriptor. After that, the geometric information is extracted in the form of curves by considering that the face is rigid.

The research literature on matching non-rigid or deformable shapes has grown substantially in recent years, as shown in the survey by Zhouhui et al. [21]. Specifically, the developing field of spectral geometry gives a rich system for the geometric analysis of non-rigid shapes, which depends on the eigensystem of the LBO. A popular and remarkable study shows examples in this direction [22], which comprises heat kernel signature (HKS), global point signature (GPS) and wave kernel signature (WKS). In our work, WKS [18] was chosen for the local features because it is very simple, it is faster than the keypoints detection matching algorithms discussed previously, and it does not need any learned examples as do Optimal Spectral Descriptors (OSD) [23] and ShapeNet [24].

Let S_{IMFn} be a facial surface of the IMF number n. S_{IMFn} is represented by N points. Local features are obtained by associating to each point x of the S_{IMFn} a WKS vector. This vector represents the time-averaged probability of measuring a quantum mechanical particle at a specific location on S_{IMFn}. Presume a quantum particle with an unknown position is on S_{IMFn}. Thus, the wave function of the particle is the Schrödinger equation solution, which can be formulated in the spectral domain as

$$\psi_e(x,t) = \sum_{k\geq 1} e^{i\lambda_k t} \emptyset_k(x) f_e(\lambda_k)$$

where e denotes the energy of the particle at time $t = 0$, and f_e its initial distribution. λ_k and $\emptyset_k(.)$ are, respectively, the eigenvalues and eigenfunctions of the LBO.

The average probability of measuring a particle with the energy distribution f_e is

$$wks_e = \lim_{n\to\infty} \frac{1}{T} \int_0^T |\psi_e(x,t)|^2 = \sum_{k\geq 1} f_e^2(\lambda_k) \emptyset_k^2(x)$$

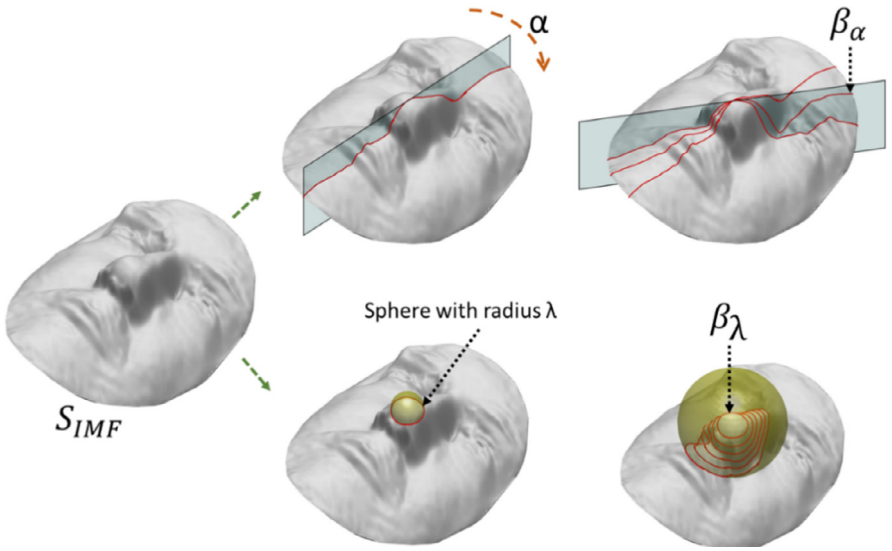

Fig. 3. The Procedure for the extraction of radial and level curves.

Let e_1,e_2,..., e_M be M log-normal energy distributions. Then, each point x on S_{IMFn} is associated with a wave kernel signature, which can be described by an n-dimensional feature vector, as follows:

$$WKS(x) = [wks_{e_1}(x), \ldots, wks_{e_M}(x)]$$

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

Before the classification stage, we associate the geometric information with local information by assigning to each point of the radial and level curves the corresponding wave kernel signature. β_α and β_λ are rewritten as follows: $\beta_{\alpha,wks_\alpha}$ and $\beta_{\lambda,wks_\lambda}$. Thus, the facial features are characterized at different scales by radial curves, level facial curves and their corresponding wave kernel signature. If α and λ are small, then we can approximately reconstruct the surface of the facial F as shown in Eq. (2).

$$F \approx \sum_{k=1}^{n} S_{IMF_k} + S_R \tag{2}$$

where
$S_{IMF_k} \approx \{\bigcup_{k,\alpha} \beta_{\alpha,wks_\alpha}\} \bigcup \{\bigcup_{k,\lambda} \beta_{\lambda,wks_\lambda}\} \approx \{\bigcup_{k,\alpha} \{IMF_k \bigcap P_\alpha\}\} \bigcup \{\bigcup_{k,\lambda} \{IMF_k \bigcap G_\lambda\}\}$
$S_R \approx \{\bigcup_{\alpha} \beta_{\alpha,wks_\alpha}\} \bigcup \{\bigcup_{\lambda} \beta_{\lambda,wks_\lambda}\} \approx \{\bigcup_{\alpha} \{R \bigcap P_\alpha\}\} \bigcup \{\bigcup_{\lambda} \{R \bigcap G_\lambda\}\}$

2.4. Recognition

The classification stage is a critical part in all recognition systems. In this stage, we attempt to match the faces of the probe with the faces of the gallery. Let F_i and F_j denote two faces that belong to the probe and gallery, respectively. F_i and F_j are represented in n scales due to the SEMD. We obtain the characteristics of each scale using radial and level curves associated with the descriptor vector WKS. The methodology used in this paper is based on the symmetrical comparison of the scales, which means that the scale number r of F_i is matched with the scale number r of F_j, and so forth.

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 sets that characterize F_i and F_j at the r scale, respectively. 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 (see Fig. 4). The angle formed between each pair of feature vectors v and u is calculated as follows:

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

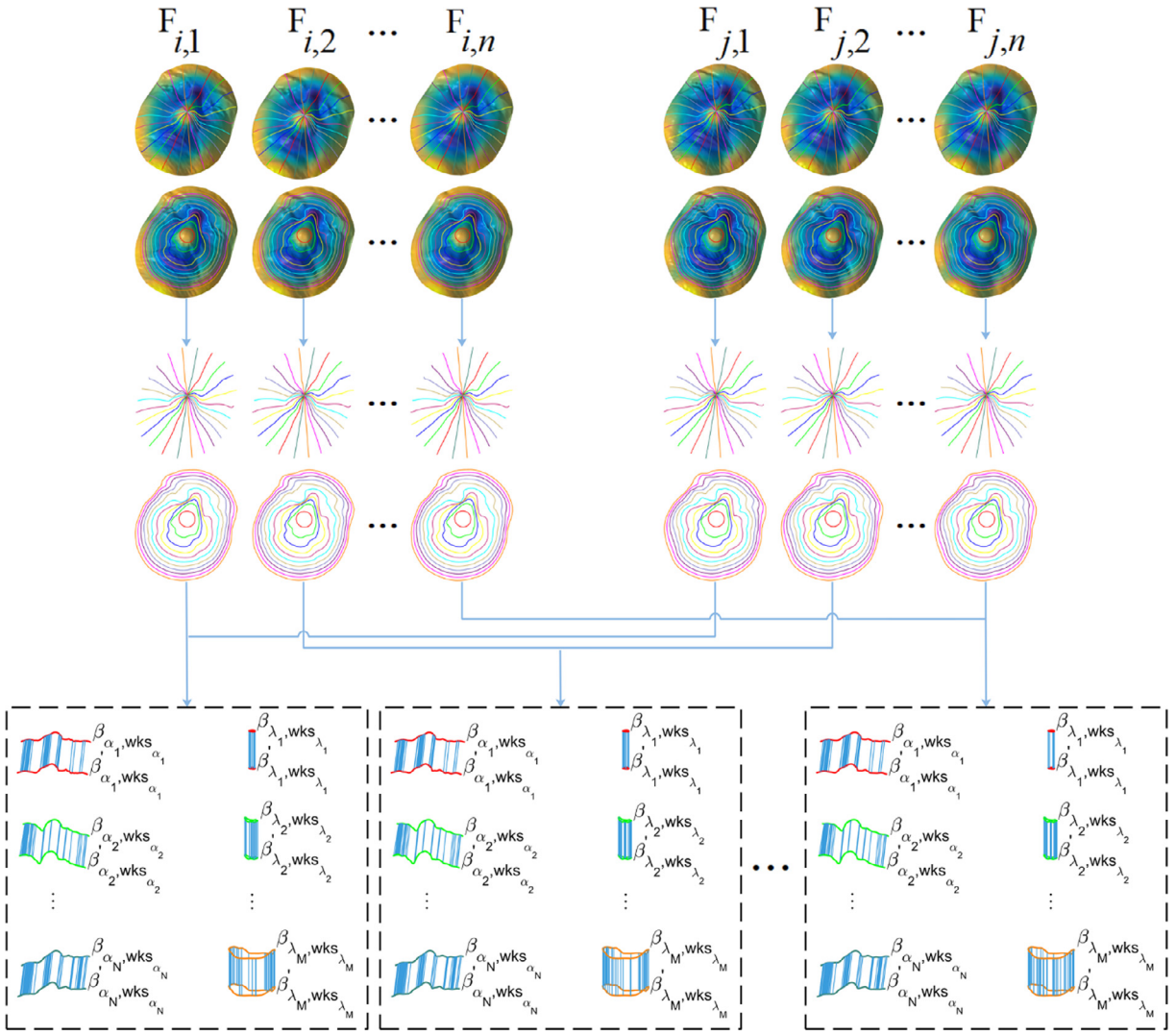

Fig. 4. Illustration of the classification stage.

here, v matches u if the ratio $\frac{\theta_{vu}}{\theta_{vw}}$ is greater than a threshold τ (in our experiments, $\tau = 0.9$). Here, $v \in \beta_{\alpha,wks_\alpha}$ and $(u, w) \in \beta'_{\alpha,wks_\alpha}$, where $\beta_{\alpha,wks_\alpha}$ and $\beta'_{\alpha,wks_\alpha}$ are 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 of the vectors of $\beta'_{\alpha,wks_\alpha}$ with $\theta_{vu} \leq \theta_{vw}$. Fig. 6 shows two examples of matching results of our method on the Bosphorus database. The first example matched the scan labeled 'bs000_N_N_0' with 'bs000_N_N_1' (Fig. 6(a)), and in the second example, the facial scan labeled 'bs000_N_N_0' is matched with 'bs060_N_N_1' (Fig. 6(b)).

3. Experiments and results

To demonstrate the effectiveness of our algorithm for expression-invariant face recognition, we performed our experiments on two datasets, namely, GavabDB [25] and Bosphorus [26]. In the following subsections, we review these two databases and present our experiment results.

In contrast to most of the existing 3D face recognition algorithms, our method does not smooth the faces of the GavabDB database, to keep all of the information that is present in the facial scans. The filtering destroys much of the information that is present in the raw image because the noise is embedded with face information in most cases. However, the faces of the Bosphorus database are characterized by high resolution (The mean number of vertex points per scans after cropping is approximately 25,000 vertex), which allows us to apply the Gaussian filter (scale = 1) to reduce the repetitive noise.

In SEMD, two parameters must be set: the stopping criterion of the IMF extraction and the number of IMFs. In the following experiments, the IMF extraction is stopped if the standard deviation SD goes below a threshold SD_{min}. In this article, we choose the SD_{min} to be 0.1 to overcome the mode mixing [27]. To reduce the time required for the execution of

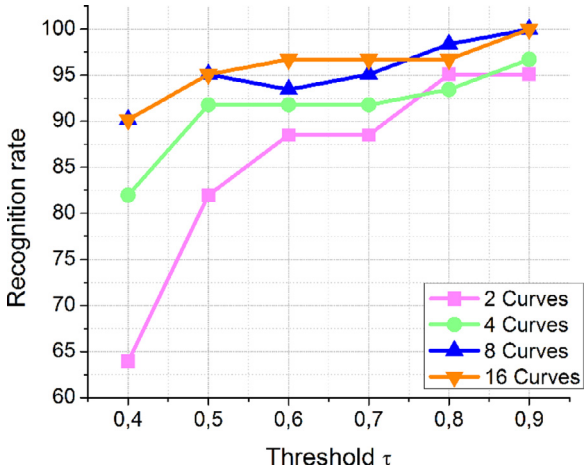

Fig. 5. Effect of changing the number of curves (radial and level) for various thresholds τ.

Table 1
Rank-1 recognition rates on GavabDB.

Approaches	Neutral	Expression	Neutral+Expression
Moreno et al. [9]	90.16	77.90	–
Li et al. [10]	96.67	93.33	94.68
Mahoor et al. [11]	–	72.00	78.00
Huang et al. [13]	**100**	93.99	95.49
Drira et al. [6]	**100**	94.54	95.90
Berretti et al. [15]	**100**	94.00	95.10
Lei, et al. [16]	**100**	95.08	96.31
Our method	**100**	**98.90**	**99.18**

the system, we selected three IMFs, and each IMF represented one of the main variations: high frequency, low frequency and intermediate frequency. We performed many tests with different values for the number of curves and the threshold, to determine the proper value of each, as shown in Fig. 5. These tests are turned on the Gavab3D database using neutral vs. random expressions. It can be seen from this figure that the best recognition rates are obtained with 8 and 16 curves when $\tau = 0.9$. Then, in this work, eight radial and level curves are chosen, and $\tau = 0.9$.

3.1. GavabDB database

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. The gallery includes for each subject the scan called “frontal1”, and the remaining images are used as probe images for recognition.

The performance of our framework against the expression variations on this dataset is demonstrated by some commonly used schemes: the cases of neutral vs. neutral, neutral vs. expression, and neutral vs. neutral+expression. For recognition, we compared our method with recently published results, including Moreno et al. [9], Li et al. [10], Mahoor et al. [11], Drira et al. [6], Huang et al. [13], Berretti et al. [15] and Lei et al. [16]. Table 1 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.

Fig. 7 shows the Cumulative Matching Characteristic (CMC) curves for the three cases. The CMC curve represents the probability that the correct choice is among the N first responses of the system (called rank). A system is said to recognize at rank 1 when it chooses the nearest face as a result of the recognition. We say that a system recognizes at rank 2 when it chooses the two faces that best correspond to the input scan, etc. If we compare the recognition rate for each IMF separately in rank-1, we conclude that the recognition rate of the residue exceeds in all cases the accuracy of IMF1 and IMF2. However, the IMF2 curve is fast because it approaches 100% before the other curves along the ranks lower than 55. However, the

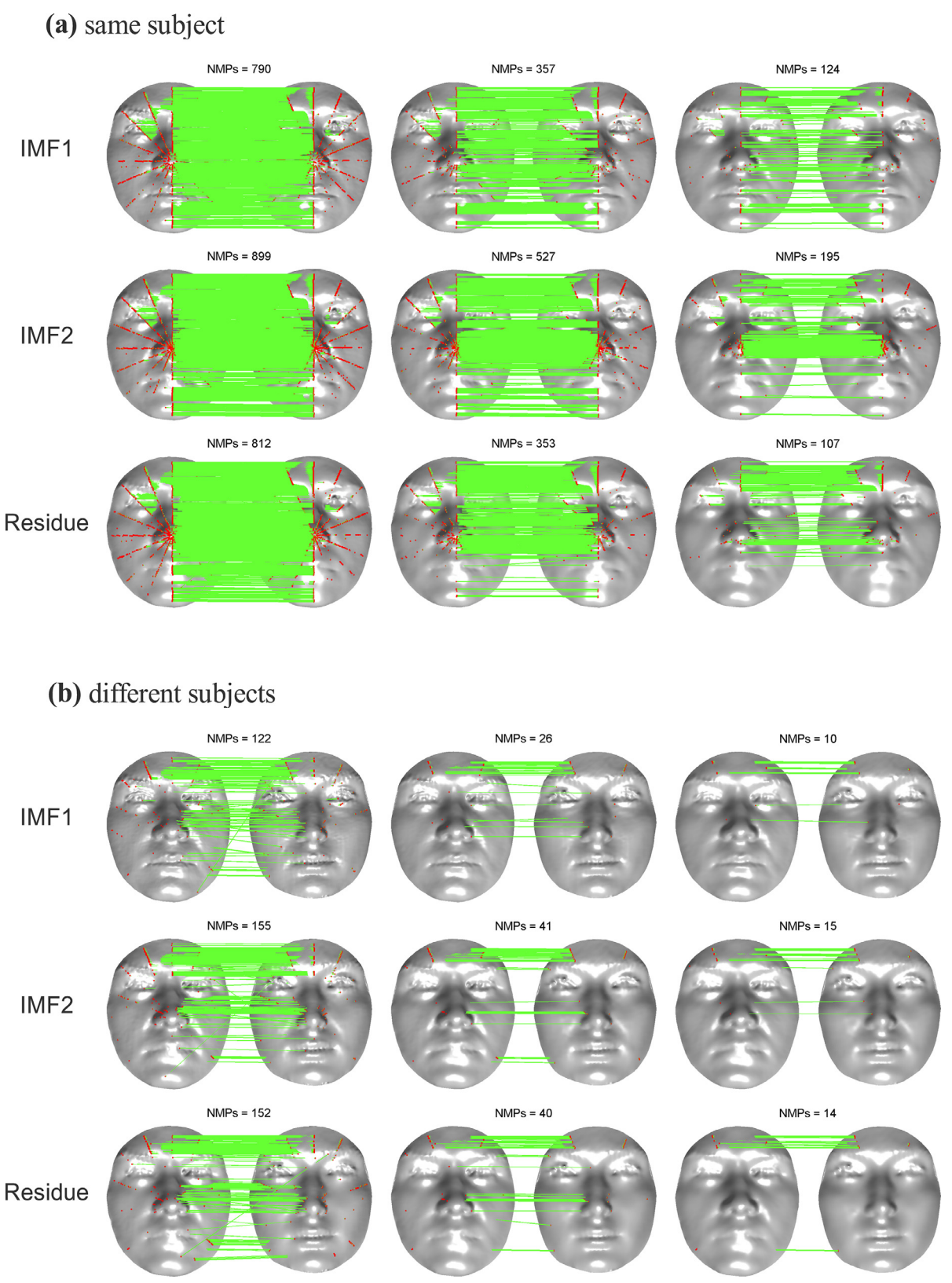

Fig. 6. Comparison between face scans at each scale: (a) same subject; (b) different subjects. Face matching from the left to the right is obtained by the thresholds $\tau = 0.9$, $\tau = 0.7$ and $\tau = 0.5$. NMPs is the number of matched points.

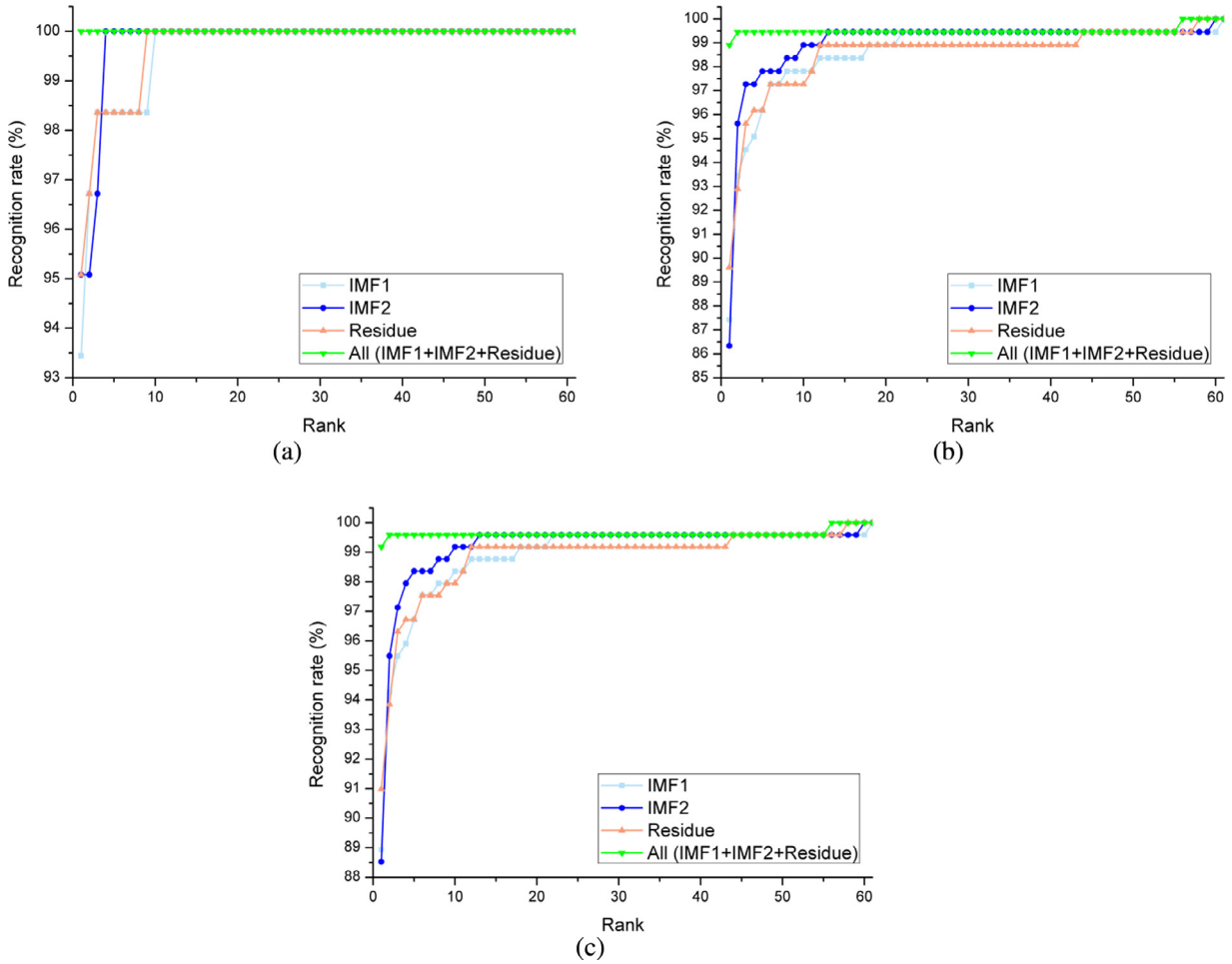

Fig. 7. The CMC curves of the proposed method using the GavabDB database. (a) Neutral (b) Expression and (c) Neutral + Expression.

grouping of different MFIs and the residue (green curve) allow us to have good results in all of the ranks and a very fast curve.

3.2. Bosphorus 3D face database

To include more challenging cases, we tested our algorithm on the Bosphorus dataset. This database includes a total of 4666 facial scans of 105 subjects (60 men and 45 women aged between 25 and 35). Approximately 71 persons have 54 different scans. Each face represents a position, an expression, or both. Overall, the subjects have 34 expressions, 13 poses, 4 occlusions and one or two neutral faces, while there are only 31 scans for the other subjects (34 individuals). These scans are composed of 10 expressions, 13 poses, 4 occlusions and 4 neutrals.

In the same way as for the Bosphorus database, the first neutral scan of each subject constitutes a gallery of 105 scans, and the remaining 2797 scans with frontal pose and without occlusions and pose variations form the probe.

Following the experimental protocol used in [15], we evaluated the performance of our proposed 3D face recognition approach on the Bosphorus database. Table 2 provides a comprehensive comparison between our algorithm and recently published algorithms. The probe scans have been organized in different classes according to the type of expression. The results of the other methods are quoted from their papers. As shown in Table 2, our solution is very close to the state-of-the-art methods in most of the categories. For certain categories, the recognition rate is even better, and more specifically, our algorithm displays the best results in the categories where there is no exaggeration in the opening of the mouth, such as the ANGER and DISGUST categories, which causes an improvement of 7.07% and 7.21%, respectively, compared to the other method. The CMC curve of each category is drawn in Fig. 8.

Table 2
Rank-1 recognition rates on Bosphorus.

Categories	Our method	Yue, 2014 [3] (ROSR[a])	Berretti et al,2013 [15] (meshDOG)	Li et al. 2011 [17] (HoG+ HoS+ HoGS)
Neutral	**100.00**	93.58	97.90	**100.00**
Anger	**95.77**	–	85.90	88.70
Disgust	**88.41**	–	81.20	76.80
Fear	81.43	–	90.00	**92.90**
Happy	88.68	–	92.50	**95.30**
Sadness	**96.97**	–	93.90	95.50
Surprise	92.96	–	91.50	**98.60**
CAU	92.31	–	95.60	**98.80**
UFAU	98.40	–	98.40	**99.10**
LFAU	91.87	–	96.50	**97.20**
Neutral vs. All	**93.24**	93.06	–	–

[a] Rigid-area orthogonal spectral regression.

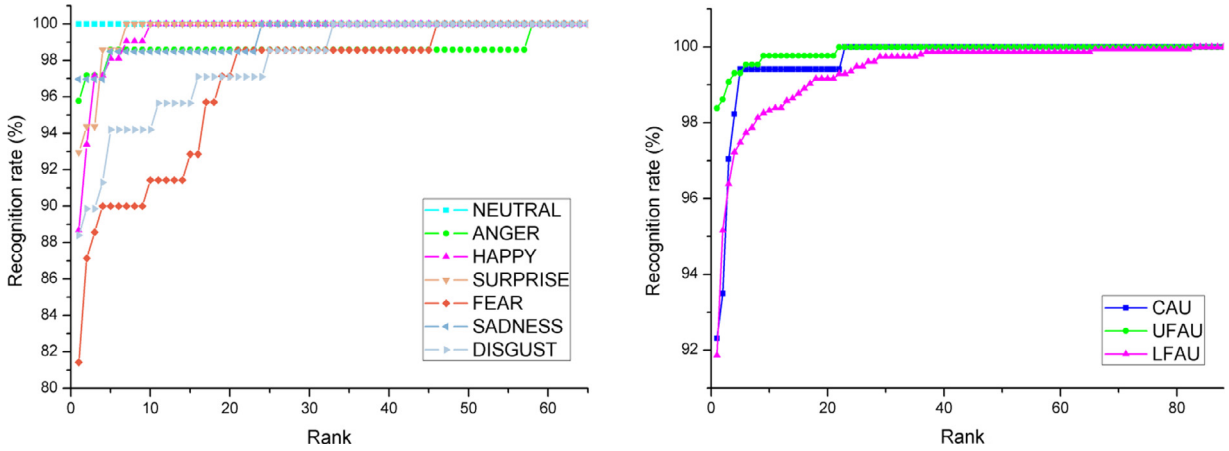

Fig. 8. The CMC curves of the proposed method using the Bosphorus database.

Table 3
Computational cost.

Step	Time consumed (s)
Mesh preprocessing	2.952 (per face)
Mesh decomposition	8.164 (three IMF per face)
Feature extraction	2.23 (per IMF)
Curves extraction	0.175 (per curve)
Curve comparison	0.0246 (two curves)

3.3. Computational performance

The computational cost is another major factor in judging the effectiveness of a 3D face recognition system. Table 3 presents the time consumed for different steps of our algorithm on a PC Intel Core i5-2430M, with a 2.40 GHz processor and 8.0GB of memory. These results were obtained based on the neutral of the Bosphorus dataset. From this table, we can observe that the computational complexity of the proposed matching approach depends on two main cost factors: the face surface decomposition and the feature extraction. The first term resulted in the principal source of the cost, which grows with the number of IMFs extracted from the scans. The elapsed time in this phase of decomposition is consumed almost entirely by the sifting process algorithm. The time of this algorithm is directly related to the value of the standard deviation *SD*, and the higher the *SD* is, the less time is consumed, and vice versa. For example, if the value of $SD = 0.9$, then the time consumed falls to 2 s per face. To avoid the problem of mode mixing [28], we chose in this work a very small value, $SD = 0.1$. The second factor is the feature extraction, and this step consumes time because each IMF consumes 2.23 s, and this time must be multiplied by the number of IMFs (in our case, 3).

If we compare the proposed approach in terms of the computational time to the state-of-the-art, then we can see that our method is located between the computational time of the local method and the geometry method. The proposed approach can perform face identification faster than most of the existing local approaches. The local approaches extract features around the neighboring areas of the detected keypoints by scanning the face vertex by vertex and region by region. In our work, the features are extracted based on the eigenvalues and eigenfunctions of the LBO, which takes as input the

entire face. For example, the time required for the meshSIFT approach [29], on the same PC, to detect the salient keypoints for the first neutral face of the Bosphorus dataset is 941.88 s (≈16 min). However, our system is slightly slower than the geometry approaches because most of the theme is based on low-level geometric features such as the geodesic distance between curves, angle, area, and curvature.

4. Conclusions

In this article, we have presented a new method for 3D facial recognition in the presence of both neutral and non-neutral facial expressions. The proposed algorithm relies on different sides to have more information and reduce the effect of facial distortion when conducting face recognition. We present experimental results on 3D face recognition on the two largest standard 3D face databases: GavabDB and Bosphorus. Compared with the previous approaches, our results outperform previous work on the GavabDB dataset and some cases on the Bosphorus dataset.

In the future, we hope to develop our framework for occlusions and missing data. We can overcome these problems by detecting the occlusions and absence of regions of the face and use the existing data to reconstruct the missing parts.

References

[1] Wang Y, Pan G, Wu Z, Wang Y. Exploring facial expression effects in 3D face recognition using partial ICP. In: Computer vision–ACCV 2006; 2006. p. 581–90.
[2] Lei Y, Bennamoun M, El-Sallam AA. An efficient 3D face recognition approach based on the fusion of novel local low-level features. Pattern Recognit 2013;46:24–37.
[3] Ming Y. Rigid-area orthogonal spectral regression for efficient 3D face recognition. Neurocomputing 2014;129:445–57.
[4] Al-Osaimi F, Bennamoun M, Mian A. An expression deformation approach to non-rigid 3D face recognition. Int J Comput Vision 2009;81:302–16.
[5] Amberg B, Knothe R, Vetter T. Expression invariant 3D face recognition with a morphable model. In: Automatic face & gesture recognition, 2008. FG'08. 8th IEEE international conference on; 2008. p. 1–6.
[6] Drira H, Amor BB, Srivastava A, Daoudi M, Slama R. 3D face recognition under expressions, occlusions, and pose variations. IEEE Trans Pattern Anal Mach Intell 2013;35:2270–83.
[7] Srivastava A, Samir C, Joshi SH, Daoudi M. Elastic shape models for face analysis using curvilinear coordinates. J Math Imaging Vision 2009;33:253–65.
[8] Ballihi L, Amor BB, Daoudi M, Srivastava A, Aboutajdine D. Boosting 3-D-geometric features for efficient face recognition and gender classification. IEEE Trans Inf Forensics Secur 2012;7:1766–79.
[9] Moreno AB, Sanchez A, Velez J, Diaz J. 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; 2005. p. 185–90.
[10] Li X, Jia T, Zhang H. Expression-insensitive 3D face recognition using sparse representation. In: Computer vision and pattern recognition, 2009. CVPR 2009. IEEE Conference on; 2009. p. 2575–82.
[11] Mahoor MH, Abdel-Mottaleb M. Face recognition based on 3D ridge images obtained from range data. Pattern Recognit 2009;42:445–51.
[12] Mian AS, Bennamoun M, Owens R. Keypoint detection and local feature matching for textured 3D face recognition. Int J Comput Vision 2008;79:1–12.
[13] Huang D, Ardabilian M, Wang Y, Chen L. 3-D face recognition using eLBP-based facial description and local feature hybrid matching. IEEE Trans Inf Forensics Secur 2012;7:1551–65.
[14] Berretti S, Werghi N, Del Bimbo A, Pala P. Selecting stable keypoints and local descriptors for person identification using 3D face scans. Visual Comput 2014;30:1275–92.
[15] Berretti S, Werghi N, Del Bimbo A, Pala P. Matching 3D face scans using interest points and local histogram descriptors. Comput Graphics 2013;37:509–25.
[16] Lei Y, Guo Y, Hayat M, Bennamoun M, Zhou X. A two-phase weighted collaborative representation for 3D partial face recognition with single sample. Pattern Recognit 2016;52:218–37.
[17] Li H, Huang D, Lemaire P, Morvan J-M, Chen L. Expression robust 3D face recognition via mesh-based histograms of multiple order surface differential quantities. In: Image processing (ICIP), 2011 18th IEEE international conference on; 2011. p. 3053–6.
[18] Aubry M, Schlickewei U, Cremers D. The wave kernel signature: A quantum mechanical approach to shape analysis. In: Computer vision workshops (ICCV workshops), 2011 IEEE international conference on; 2011. p. 1626–33.
[19] Wang H, Su Z, Cao J, Wang Y, Zhang H. Empirical mode decomposition on surfaces. Graphical Models 2012;74:173–83.
[20] Szeptycki P, Ardabilian M, Chen L. 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; 2009. p. 1–6.
[21] Lian Z, Godil A, Bustos B, Daoudi M, Hermans J, Kawamura S, et al. A comparison of methods for non-rigid 3D shape retrieval. Pattern Recognit 2013;46:449–61.
[22] Li C, Hamza AB. Spatially aggregating spectral descriptors for nonrigid 3D shape retrieval: a comparative survey. Multimedia Syst 2014;20:253–81.
[23] Litman R, Bronstein A, Bronstein M, Castellani U. Supervised learning of bag-of-features shape descriptors using sparse coding. In: Computer graphics forum; 2014. p. 127–36.
[24] J. Masci, D. Boscaini, M. Bronstein, and P. Vandergheynst, Shapenet: convolutional neural networks on non-euclidean manifolds, 2015.
[25] Moreno AB, Sánchez A. GavabDB: a 3D face database. In: Proc. 2nd COST275 workshop on biometrics on the internet Vigo (Spain); 2004. p. 75–80.
[26] Alyuz N, Gokberk B, Akarun L. A 3D face recognition system for expression and occlusion invariance. In: Biometrics: theory, applications and systems, 2008. BTAS 2008. 2nd IEEE international conference on; 2008. p. 1–7.
[27] Abbad A, Douini Y, Abbad K, Tairi H. Post-processing of dimensionality reduction methods for Face Recognition. Pattern Recognit Image Anal 2017;27:265–74.
[28] Huang NE, Wu Z. A review on Hilbert-Huang transform: method and its applications to geophysical studies. Rev Geophys 2008;46.
[29] Smeets D, Keustermans J, Vandermeulen D, Suetens P. meshSIFT: local surface features for 3D face recognition under expression variations and partial data. Comput Vision Image Understanding 2013;117:158–69.

Abdelghafour Abbad received his Master's degree in 2012 from the Faculty of Sciences Dhar El Mehraz Fez Morocco. He is currently working on his Ph.D. in the LIIAN Laboratory at USMBA-Fez University. His research interests include pattern recognition, image processing, and computer vision.

Khalid Abbad received his Master's degree from faculty of sciences Dhar El Mehraz FEZ MOROCCO in 2006 and Ph.D. in computer science from University Sidi Mohammed Ben Abdellah, Fez, Morocco in 2011. He is currently an assistant professor at faculty of science and technologies, Fez, Morocco. His current research activities are in pattern recognition, computer vision and biometric recognition.

Hamid Tairi received his Ph.D. degree in 2001 from the University Sidi Mohamed Ben Abdellah, Morocco. In 2003, he has been an associate professor at the University Sidi Mohamed Ben Abdellah, where he obtained his HDR in 2009. His research interests are in visual tracking for robotic control, in 3-D reconstruction of artificial vision, in medical image, in visual information retrieval and pattern recognition.