3-D Face Recognition Using eLBP-Based Facial Descriptionand Local Feature Hybrid Matching

Di Huang, Mohsen Ardabilian, Yunhong Wang, Liming Chen

3D Face Recognition using eLBP-based Facial Description and Local Feature Hybrid Matching

Di Huang*, *Member, IEEE*, Mohsen Ardabilian,

Yunhong Wang, *Member, IEEE*, and Liming Chen, *Member, IEEE*

Abstract—This paper presents an effective method for 3D face recognition using a novel geometric facial representation along with a local feature hybrid matching scheme. The proposed facial surface description is based on a set of facial depth maps extracted by multi-scale extended Local Binary Patterns (eLBP) and enables an efficient and accurate description of local shape changes; it thus enhances the distinctiveness of smooth and similar facial range images generated by preprocessing steps. The following matching strategy is SIFT-based and performs in a hybrid way that combines local and holistic analysis, robustly associating the keypoints between two facial representations of the same subject. As a result, the proposed approach proves to be robust to facial expression variations, partial occlusions and moderate pose changes, and the last property makes our system registration-free for nearly frontal face models. The proposed method was experimented on three public datasets, *i.e.* FRGC v2.0, Gavab and Bosphorus. It displays a rank-one recognition rate of 97.6% and a verification rate of 98.4% at a 0.001 FAR on the FRGC v2.0 database without any face alignment. Additional experiments on the Bosphorus dataset further highlight the advantages of the proposed method with regard to expression changes and external partial occlusions. The last experiment carried out on the Gavab database demonstrates that the entire system can also deal with faces under large pose variations and even partially occluded ones, when only aided by a coarse alignment process.

Index Terms—3D face recognition and verification, geometric facial description, extended *LBP*, SIFT, hybrid matching

1 Introduction

THE face has its own advantages over other biometrics for people identification and verification-related applications, since it is natural, non-intrusive, contactless *etc.* Unfortunately, all human faces are similar to each other in their configurations and hence offer low distinctiveness, unlike other biometrics, *e.g.* the iris and fingerprint. [1]. Furthermore, intra-class variations, due to factors as diverse as pose and facial expression *etc.* are usually greater than inter-class ones. The past decades have witnessed tremendous efforts firstly focused on 2D face images [2] and more recently on 3D face models or scans [3]. Despite great progress achieved so far within the field [2], face recognition (FR) using 2D facial texture images is still not reliable enough [4], especially in the presence of pose and lighting variations [5]. With the rapid development in 3D imaging systems, 2.5D and 3D facial scans have emerged as a major alternative in dealing with the unsolved issues in 2D face recognition, *i.e.* changes of illumination and pose [3], [6]. Meanwhile, even though 3D facial scans capture exact shape information of facial surfaces, and are thereby theoretically reputed to be robust to variations in illumination, they are likely to be more sensitive to expression changes. Furthermore, they generally require an accurate registration step before 3D shape-based face matching.

1.1 Related Work

Generally, how to describe facial surface is a core topic in 3D face recognition. "Good" features of facial surfaces should have the following properties [7]: first, they can tolerate within-class variations while discriminating different classes well; second, they can easily be extracted from raw facial data to allow fast processing; finally, they should lie in a space with moderate dimensionality to avoid high computational cost in matching. As a result, 3D face recognition techniques can be firstly classified according to the features they use: (1) *original feature-based* techniques make use of the entire face region as input to compute similarity. Several works have explored PCA directly on facial range images [8], [9], [10]; while some have applied the ICP (Iterative Closest Point) algorithm [11] or its modified version on facial point-clouds to match surfaces [12], [13], [14], [15]. The Hausdorff distance has also been investigated for face matching [16], [17]; (2) *region or point feature based* detects representative facial areas or points to construct feature spaces. The eye and nose areas are used in [18]; segmented facial regions and lines are utilized in [19]; anthropometric facial fiducial keypoints are employed in [20]; (3) *curve feature based* extracts discriminative surface curves for facial representation. In [21], three facial curves are found to intersect the facial surface using horizontal and vertical planes as

- *Di Huang and Yunhong Wang are with State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, 100191, Beijing, China (e-mail: dhuang@buaa.edu.cn).*
- *Mohsen Ardabilian and Liming Chen are with Université de Lyon, CNRS, Ecole Centrale de Lyon, LIRIS, UMR5205, 69134, Ecully, France.*

Manuscript received (insert date of submission if desired). Please note that all acknowledgments should be placed at the end of the paper, before the bibliography.

xxxx-xxxx/0x/$xx.00 © 200x IEEE

well as a cylinder; the central profile with maximal protrusion as well as two parallel profiles are searched in [22]; a union of the level curves of a depth function is proposed to represent 3D facial surfaces [23]; (4) *shape feature based* focuses on the attributes of local surfaces, such as curvatures [18], point signature [24], Extended Gaussian Image (EGI) [25], Signed Shape Difference Map (SSDM) [26], *etc.*

3D face recognition techniques can also be categorized according to the nature of their matching strategies, even though it is highly dependent on the facial features used. Zhao et al. [2] have roughly classified 2D face recognition approaches into three main streams: *holistic*, *e.g.* PCA [27] and LDA [28]; *feature-based* such as Elastic Bunch Graph Matching (EBGM) [29]; *hybrid* like Component Eigenfaces [30]. This taxonomy can be extended to 3D face recognition. *The holistic stream* contains ICP-based matching [12], annotated deformable model [31], and isometry invariant description [32] *etc.* This matching scheme is based on holistic facial features and hence generally requires an accurate normalization step with respect to pose and scale changes. Furthermore, it has proved sensitive to expression variations and partial occlusions. The *feature-based* one utilizes local features of 3D facial scans and has been explored in several works in the literature, including point signature [24], and more recently keypoint detection and local feature matching for textured 3D face recognition by Mian et al. [33]. In this last work, an original keypoint detection method on a 3D facial surface was introduced, and a graph-based matching scheme along with a dynamic fusion strategy performed at score level. As also highlighted in [33], feature-based matching has the potential advantage of being robust to facial expression, pose and lighting changes and even to partial occlusions. The downside of this scheme is the difficulty in extracting sufficient informative feature points from similar or smooth 3D facial surfaces. Some studies also exist which present *hybrid matching* combining global facial features with local ones: Region-ICP [34], multiple region-based matching [35] and the component and morphable model-based approach [36]. As this scheme tends to combine facial configuration information with local properties of faces, it is theoretically the most powerful [2]. However, it also risks inheriting both types of shortcomings: sensitivity to pose variations, difficulty in generating sufficient stable descriptive features, *etc.*

1.2 Motivation and Approach Overview

In this paper, our basic assumption, as the one behind feature-based face recognition algorithms, is that, when a variation such as an expression or an occlusion occurs on a probe 3D facial scan, there still exist some small local areas, *e.g.* the nose region during a facial expression, that change slightly or remain invariant as compared to the corresponding 3D facial scan in the gallery set. Once located and characterized accurately, these local regions can be utilized to identify 3D faces, thereby providing robustness to facial expressions and partial occlusions through a proper matching process.

Motivated by this intuition, this paper proposes a novel approach to 3D face recognition, making use of a geometric facial representation along with a local feature hybrid matching strategy. Our main contributions can be summarized as follows:

(1) Because after basic preprocessing, *e.g.* spike removal and hole filling, 3D facial surfaces to be identified are generally smooth and similar, to achieve accurate representations of facial surfaces and enhance their distinctiveness, we propose a 3D shape based geometric facial description, consisting of a set of Multi-Scale extended Local Binary Pattern Depth Faces (*MS-eLBP-DFs*). This method improves the discriminative power of *LBP* for 3D facial surface description by two solutions, *i.e.* encoding exact gray value differences between the central pixel and the neighboring ones as well as embedding a Multi-Scale (MS) scheme.

(2) In order to extract sufficient repeatable local features on smooth facial surfaces, we propose to apply SIFT [46] to these Multi-Scale extended LBP-based facial representations, *i.e. MS-eLBP-DFs*, interpreted as simple gray level images, for the detection of keypoints and the characterization of local geometric properties.

(3) A hybrid matching approach is designed to measure similarities between gallery and probe facial scans once they are represented in terms of the geometric facial description *i.e. MS-eLBP-DF*s. This matching method inherits the principles of local feature matching along with graph based matching as in [33], and it also extends the latter by incorporating a facial component constraint.

(4) Thanks to the local feature-based approach and the hybrid matching scheme, the proposed 3D face recognition algorithm is robust to facial expression changes and partial occlusions whilst remaining tolerant of moderate pose variations. The last property makes our method a registration-free technique for recognizing nearly frontal 3D facial scans as can be the case of most user cooperative biometric applications, *e.g.* access control. This is clearly in contrast to the overwhelming majority of state-of-the-art 3D face recognition algorithms requiring the costly 3D face alignment step before face matching.

As a local feature-based 3D face recognition algorithm, the proposed method and that of Mian et al. in [33] share some similarities, including in particular the overall local-feature oriented framework (though the features used in each of the two approaches are very different), the use of the SIFT method (though SIFT is applied to extended *LBP* based shape representations, *i.e. MS-eLBP-DFs*, instead of texture), the matching strategy combining local feature matching and graph based matching that is further extended to include a facial component constraint-based matching, and finally the weighted fusion scheme at score level with dynamic weight calculation.

The proposed method was evaluated on three public databases, namely FRGC v2.0 [48], Bosphorus [68] and Gavab DB [39]. As experimented on the FRGC v2.0 database for both the tasks of 3D face recognition and verification, our approach achieves a rank-one recognition rate of 97.6% and a 98.4% verification rate with a FAR of 0.1% respectively. Since the 3D facial scans in FRGC are nearly frontal, the costly step of 3D face alignment was not re-

quired. The experiment carried out on the Bosphorus dataset further highlights the ability of the proposed method to identify nearly frontal 3D face models owning expression changes and external occlusions. When it handles large pose variations including left and right profiles which probably lead to self-occlusions, a coarse alignment based on a few landmarks is sufficient as a preprocessing step of this approach. This is demonstrated by the experiments on the Gavab dataset.

The preliminary experimental results of this work appeared in [50] and [64]. The remainder of this paper is organized as follows. The proposed geometric facial description, *MS-eLBP-DFs*, is shown in section II, and section III presents SIFT-based local feature extraction. The hybrid matching step is introduced in section IV. Experimental results of both face recognition and verification are described and analyzed in section V. Section VI concludes the paper.

2 Multi-Scale Extended LBP Depth Faces

Due to their descriptive power of micro-texture structures and computational simplicity, Local Binary Patterns (*LBP*) are among the most successful descriptors for 2D texture based face analysis [65]. In literature, *LBP* has been also investigated for 3D face recognition [37], [38]; however, *LBP* is not as discriminative as we expected for 3D facial representation since it cannot correctly distinguish similar local surfaces because of its thresholding strategy. To address this problem, two solutions are considered. First, extended *LBP* (*eLBP*), generalized from the task on *3DLBP* [38] and capable of handling different numbers of sampling points and various scales, is used. It not only extracts the relative gray value differences from the central pixel and its neighbors provided by *LBP*, but also focuses on their absolute differences that prove critical to describe range faces as well. Secondly, a multi-scale strategy is introduced to represent local surfaces to different extents which are then combined for a comprehensive description. Additionally, previous works simply repeated the histogram based manner as did in 2D facial analysis that firstly divides the face into a number of sub-regions, where *LBP* based histograms are extracted; then concatenates all these local histograms into a global one to construct a final facial feature. Unlike these tasks, we adopt an image based approach by applying *eLBP* directly to a facial range image to generate a set of Multi-Scale extended *LBP* Depth Faces (*MS-eLBP-DFs*) which retain all 2D spatial information of range faces. Finally, this approach inherits the property of computational simplicity from *LBP* and achieves fast processing.

In this section, we firstly recall the basics of *LBP* and analyze its descriptive ability for local facial surface representation. We then present extended *LBP* (*eLBP*) and the multi-scale scheme to generate a novel 3D geometric facial description, called *MS-eLBP* Depth Faces (*MS-ELBP-DFs*) that comprehensively encodes local shape variations of range faces.

2.1 LBP and Its Descriptive Power for Local Shape Variations

LBP, a non-parametric algorithm [40], was first proposed to describe local texture in 2D images. The most important properties of *LBP* are its tolerance to monotonic illumination variations and computational simplicity, so it has been extensively adopted for 2D face recognition in the past few years [41].

Specifically, the original *LBP* operator labels each pixel of a given 2D image by thresholding in a 3×3 neighborhood. If the values of the neighboring pixels are no lower than that of the central pixel, their corresponding binary bits are assigned to 1; otherwise they are assigned to 0. A binary number is hence formed by concatenating all the eight binary bits, and the resulting decimal value is used for labeling. Figure 1 illustrates the *LBP* operator by a simple example.

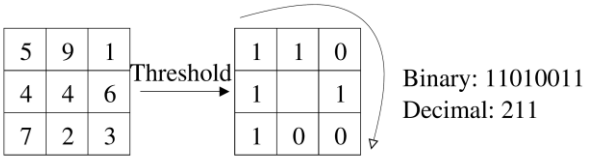


Fig. 1. An example of the original LBP operator.

Formally, given a pixel at (x_c, y_c), the derived *LBP* decimal value is:

$$LBP(x_c, y_c) = \sum_{n=0}^{8} s(i_n - i_c)2^n;\ \ s(x) = \begin{cases} 1 & if\ \ x \geq 0 \\ 0 & if\ \ x < 0 \end{cases} \quad (1)$$

where n covers the eight neighbors of the central pixel, i_c and i_n are the gray level values of the central pixel and its surrounding pixels respectively.

According to equation (1), the *LBP* code is invariant to monotonic gray-scale transformations, preserving their pixel orders in local neighborhoods. When *LBP* operates on the images formed by light reflection, it can be used as a texture descriptor. Each of the 256 (2^8) *LBP* codes can be regarded as a micro-texton. Local primitives codified by the bins include different types of curved edges, spots, flat areas *etc.* Fig. 2 shows some examples. Similarly, as *LBP* works on range images which are based on depth information, it can also describe local shape structures, such as flat, concave, convex *etc.*, as shown in Fig. 3.

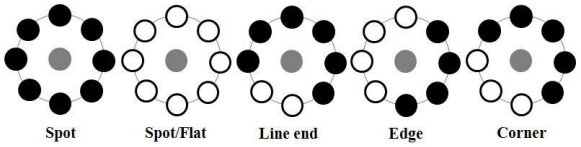


Fig. 2. An example of texture patterns which can be encoded by LBP (white circles represent ones and black circles zeros).

Unfortunately, the direct application of *LBP* to depict the shapes of 3D surfaces also leads to unexpected confusion when distinguishing similar yet different local shapes. Fig. 4 lists two similar shapes with different Shape Index (*SI*) values [42] while indeed sharing the same *LBP* code: shape (A) is a spherical cap; shape (B) is a dome. This lack of descriptive power is problematic when

one needs to derive a facial description to enhance distinctiveness for the task of face recognition.

In order to address such a problem, we propose to adopt two complementary solutions. The first solution aims to improve the discriminative ability of LBP with the eLBP coding method, and the other one focuses on providing a more comprehensive geometric description of a given neighborhood by applying a multi-scale strategy. Each solution is discussed in the following two subsections respectively.

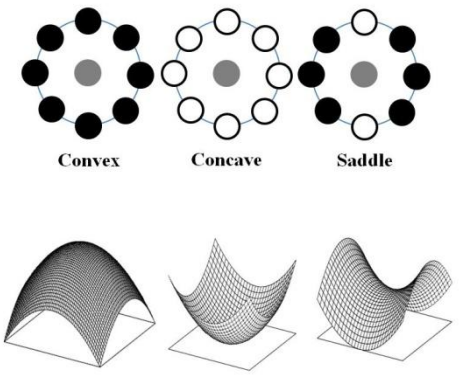


Fig. 3. Examples of local shape patterns encoded by the basic LBP operator (white circles represent ones and black circles zeros).

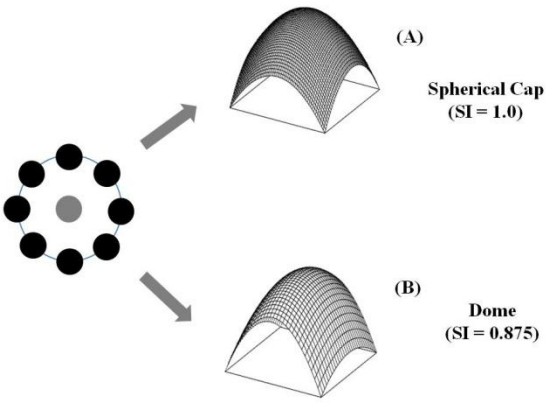


Fig. 4. A confusion case of LBP when it encodes similar but different local geometric shapes.

2.2 Extended Local Binary Patterns

The fact that *LBP* is not competent to distinguish similar local shapes is due to its operation mode. It only encodes relative differences between a central pixel and its neighboring ones. In this section, we introduce *eLBP* in order to better describe local surface properties. Instead of the original *LBP*, *eLBP* not only extracts the relative gray value difference between the central pixel and its neighbors provided by *LBP*, but also focuses on their absolute differences which are also critical to describe local shapes. *eLBP* is a generalized version of the *3DLBP* [38] originally proposed for histogram-based 3D face recognition.

Specifically, the *eLBP* code consists of several *LBP* codes in multiple layers that encode the exact gray value differences (*GD*) between the central pixel and its neighbors. The first layer of *eLBP* is actually the original *LBP* code encoding the *GD* sign. The following layers of *eLBP* then encode the absolute value of *GD*. Basically, each absolute *GD* value is firstly encoded in its binary representation and then all the binary values at a given layer result in an additional local binary pattern. The example of Fig. 1 can be expressed by *eLBP* as shown in Fig. 5. The first layer of *eLBP* code is simply the original *LBP* code that encodes the sign of *GD*, thus yielding a decimal number of 211 from its binary form $(11010011)_2$. The absolute values of *GD*, *i.e.* 1, 5, 3, 2, 1, 2, 3, 0, are first encoded in their binary numbers: $(001)_2$, $(101)_2$, $(011)_2$, $(010)_2$, …, *etc.* Using the same weighting scheme of *LBP* on all the binary bits, we generate the *eLBP* code of its corresponding layer, *e.g.*, L_2 is composed of $(01000000)_2$ and its decimal value is 64; L_3 is composed of $(00110110)_2$ and its decimal value is 54; finally L_4 is composed of $(11101010)_2$ and its decimal value is 234. As a result, when describing two similar local shapes, although the first layer *LBP* is not discriminative enough (both marked by the decimal value of 211), the information encoded in their additional layers can be used to distinguish them so long as the values of the two shapes in all corresponding layers are not exactly the same.

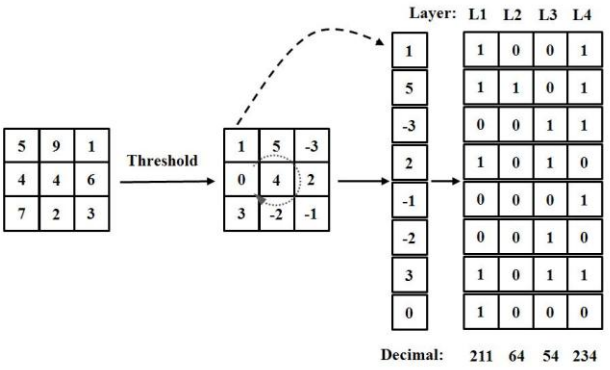


Fig. 5. An example of the eLBP operator.

Theoretically, in one image, the maximum value of *GD*s is 255 (between 0 and 255), which means that 8 additional binary units are required to code *GD*s (2^8-1=255), and thus 7 additional layers should be produced. Nevertheless, we do not need so many layers in e*LBP*. Preprocessed range faces are indeed very smooth; the *GD*s in a local surface generally do not vary dramatically. Some preliminary statistical work reveals that more than 80% *GD*s are smaller than 7 between points within eight pixels. Therefore, the number of additional binary units, k, is determined by *GD*. Meanwhile k can also be exploited to control the trade-off between the description expressiveness of local shapes and the computational simplicity of *eLBP*. All the *GD*s which are larger than 2^k-1 can be assigned to 2^k-1 to decrease computational cost. In this study, three additional layers are extracted and analyzed to illustrate their contributions to the final accuracy.

2.3 Multi-Scale Strategy

The original *LBP* operator was extended later with different sizes of local neighborhood to deal with various scales [40]. The local neighborhood of the *LBP* operator is defined as a set of sampling points evenly spaced on a circle centered on the pixel to be labeled. These sampling points which do not fall exactly on the pixels are expressed us-

ing bilinear interpolation, thus allowing any radius value and any number of points in the neighborhood. Figure 6 shows different *LBP* neighborhoods. The notation (*P*, *R*) denotes the neighborhood of *P* sampling points on a circle of radius *R*. By adopting the same protocol, the e*LBP* operator can handle different sampling points and scales as well.

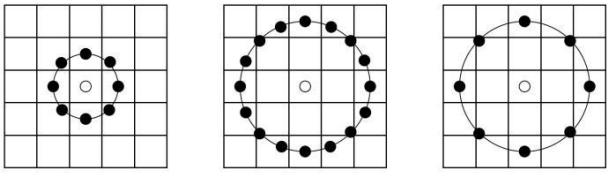


Fig. 6. LBP operator examples: circular (8, 1), (16, 2), and (8, 2).

Some *LBP* histogram-based tasks change the neighborhood of the *LBP* operator for improved performance. By varying the value of radius *R*, the *LBP* of different resolutions is thus obtained. The multi-scale strategy was first used for texture classification [40], and it was also introduced to 2D face recognition [43] [44]. In [45], Shan and Gritti studied *MS-LBP* for facial expression recognition by firstly extracting *MS-LBP* histogram- based facial features and then using AdaBoost to learn the most discriminative bins. They reported that the boosted classifiers of *MS-LBP* consistently outperform those based on single-scale *LBP*, and the selected *LBP* bins distribute at all scales. *MS-LBP* can hence be regarded as an efficient method for facial representation. When considering it in 3D face analysis, this multi-scale technique can be applied to enhance the descriptive power of *LBP*.

2.4 Multi-Scale Extended LBP Depth Faces (MS-eLBP-DFs)

LBP facial representation can be achieved in two ways: one is by *LBP* histogram; the other is by *LBP* face. The general idea of the former is that a human face can be regarded as a composition of micro-patterns described by *LBP*. The images are divided into a certain number of local regions, from which *LBP* histograms are extracted. These histograms are concatenated and thus contain both local and global information about faces. The second method is to generate *LBP* based maps. It regards the decimal number of the *LBP* code as the pixel values of an *LBP* map, and thus produces the corresponding *LBP* face. Due to its own strategy, an *LBP* histogram loses some 2D spatial information for representing faces. In this study, the second, *eLBP* face, is investigated.

For a facial range image, we generate a set of *MS-eLBP-DF*s for facial representation, *i.e.* an original LBP map (describing relative gray value differences between the central pixel and its neighbors) as well as its additional maps (representing exact gray value differences between the central pixel and its neighbors). These *MS-eLBP-DF*s can be achieved by varying the neighborhood size of the *eLBP* operator, or by first down-sampling range faces and then adopting an *eLBP* operator with a fixed radius. Some face samples are shown in Fig.7. In that figure, the number of sampling points is 8, and the radius value varies from 1 pixel to 8 pixels. As we can see, the preprocessed range face is very smooth, whilst the resulting *MS-eLBP-DF*s contain much more detail of local shape variations.

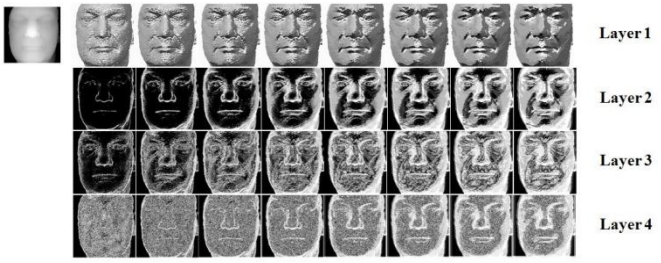


Fig. 7. MS-eLBP-DFs of a range face image with different radii from 1 to 8 (from left to right).

3 Local Feature Extraction

Once the *MS-eLBP-DF*s have been produced, the widely-used SIFT features [46] are extracted from them and exploited to calculate a similarity score between two 3D facial scans in the subsequent matching process.

SIFT applies the Difference-of-Gaussian (DOG) scale-space to detect keypoints in 2D images. The raw images are repeatedly convolved with Gaussians of different scales separated by a constant factor *k* to produce an octave in scale space. As for an input image, $I(x, y)$, its scale space is defined as a function, $L(x, y, \sigma)$, produced by convolution of a variable scale Gaussian $G(x, y, \sigma)$ with the input image I, and the DOG function $D(x, y, \sigma)$ can be computed from the difference of two nearby scales:

$$\begin{aligned} D(x,y,\sigma) &= (G(x,y,k\sigma) - G(x,y,\sigma)) * I(x,y) \\ &= L(x,y,k\sigma) - L(x,y,\sigma) \end{aligned} \tag{2}$$

Then, extremes of $D(x, y, \sigma)$ are detected by comparing each pixel with its 26 neighbors in 3×3 regions at current and adjacent scales (see Fig. 8). At each scale, gradient magnitude, $m(x, y)$, and orientation, $\theta(x, y)$, are computed using pixel differences in (3) and (4).

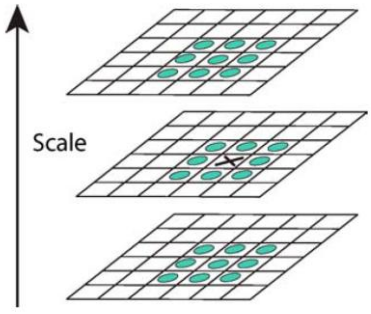


Fig. 8. Extremes (maxima or minima) of the difference-of-Gaussian images are detected by comparing a pixel (marked with "X") to its 26 neighbors in 3×3 regions at current and adjacent scales (marked with circles) [46].

$$\begin{aligned} m^2(x,y) &= (L(x+1,y) - L(x-1,y))^2 \\ &+ (L(x,y+1) - L(x,y-1))^2 \end{aligned} \tag{3}$$

$$\theta(x,y)=\tan^{-1}\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)} \quad (4)$$

For each detected keypoint, a feature vector is extracted as a descriptor from the gradients of the sampling points within its neighborhood. See Fig. 9 for more details. To achieve orientation invariance, coordinates and gradient orientations of sampling points in the neighborhood are rotated relative to keypoint orientation. Then a Gaussian function is used to assign a weight to the gradient magnitude of each point. The points close to the keypoint are given more emphasis than the ones far from it (see [46] for SIFT parameter settings). The orientation histograms of 4×4 sampling regions are calculated, each with eight orientation bins. Hence a feature vector with a dimension of 128 (4×4×8) is produced.

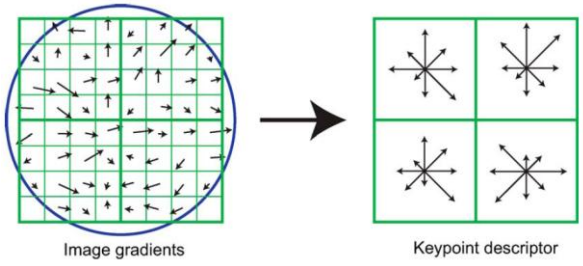


Fig. 9. Computation illustration of the keypoint descriptor [46].

The SIFT operator works on each *MS-eLBP-DF* separately. Because each *MS-eLBP-DF* highlights the local shape changes of an original smooth facial range image by encoding local binary patterns at different scales and thus providing various details, many more SIFT-based keypoints can be detected for the following matching step than in the original smoothed range face. Some statistical work was done using all the 3D facial scans in the FRGC v2.0 database. The average number of descriptors extracted from each of *MS-eLBP-DF*s is 553, while that of each original facial range image is limited to 41 and the detected keypoints are often located on the edge of the face. As a result, the number of detected keypoints is increased by 10 times more on *MS-eLBP-DFs* as compared to the original face range image; and the number of keypoints only covers 0.3% of pixels in each *MS-eLBP-DF* that has an average resolution of 400x400. Figure 10 shows the SIFT-based keypoints extracted from one facial range image and its four associated *eLBP-DFs* respectively.

Meanwhile, we also studied the repeatability of these keypoints detected on *MS-eLBP-DF*s across different scans of the same subject based on three manually labeled facial landmarks, *i.e.* the nose tip and two inner corners of eyes. 25 subjects that have more than 4 facial scans were randomly selected in FRGC v2.0 for this experiment. To each keypoint *K* detected on an *MS-eLBP-DF*, we can compute a vector of three distances, (a, b, c), each of which is the one between the keypoint and one of these three landmarks. These three distances are then used as the coordinates of *K*. Now a keypoint *X* detected by SIFT on an *MS-eLBP-DF* with the coordinates (a_x, b_x, c_x) is considered as in correspondence or matched with the keypoint *Y* detected from another *MS-eLBP-DF* possessing the coordinates (a_y, b_y, c_y), as long as the difference between each pair of coordinates, *i.e.* a_x-a_y, b_x-b_y, c_x-c_y, falls within 4 *mm* in absolute value. This definition is similar to the one as in [33] which considers that two keypoints are matched when their distance falls within 4 mm on two registered facial surfaces. Based on such a definition, the repeatability reaches 84.2% for neutral facial scans. For expressive faces, this repeatability at 4*mm* drops to 79.4% due to the 3D face shape changes. Note that these figures are comparable to that in [33] which proposed an original keypoint detection method based on local shape changes as well. Their repeatability is 86% for neutral facial scans and 75.6% for non-neutral facial scans. Our following experiments in section 5 show that this level of repeatability is quite sufficient to distinguish intra-class variations from inter-class ones for 3D facial scans with moderate pose changes, facial expressions and even partial occlusions when the hybrid matching scheme developed in the following section is applied.

4 The Hybrid Matching Process

Once local features have been extracted from *MS-eLBP-DF*s, a hybrid matching process is carried out, which combines a local matching step using the SIFT-based features with a global one under the facial component and configuration constraints.

4.1 Local Feature-Based Matching

Given local facial features extracted from each *MS-eLBP-DF* pair of the gallery and the probe face scan respectively, two facial keypoint sets can be matched. Matching one keypoint to another is accepted only if the matching distance is less than a predefined threshold, t times the distance to the second closest match. Here, $N_{Li(P,R)}$ denotes the number of matched keypoints in the i_{th} layer of an *eLBP-DF* pair, generated by *eLBP* from range face images with a parameter setting of (P, R).

4.2 Holistic Facial Matching

Unlike the samples used in the domain of object detection, all human faces have the same physical components and share a similar global configuration. Holistic matching is thus carried out to constrain the matched local features with respect to the facial components and configuration.

1) Facial Component Constraint: we propose to divide the entire facial range image into non-overlapped sub-regions, each of which contains roughly one component of nearly frontal faces. Different from the division scheme used for histogram statistics, the one in our method is to restrict the matched keypoints of gallery and probe face scans only to those with similar physical meaning. That means the matched keypoints from the same facial region should be more important. Instead of the costly clustering process [47] to automatically construct sub-regions based on keypoint locations from training samples, we simply use facial component position, and divide the face area into 3×3 rectangle blocks of the same size. The similarity

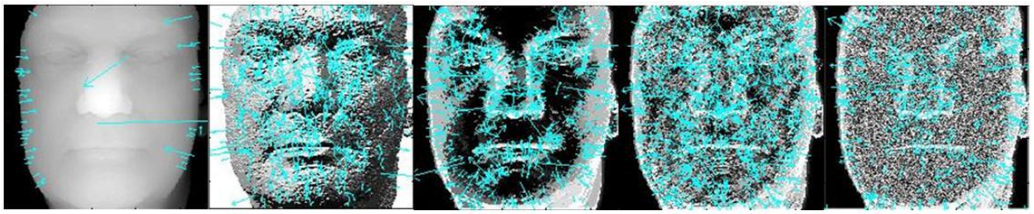


Fig. 10. The SIFT-based keypoints of an original facial range image and its four associated eLBP-DFs.

measurement of the facial component constraint is defined from this facial composition scheme. An *MS-eLBP-DF*, *I*, is represented as (m_1, m_2, ..., m_k); k is 9 in our case and m_i is the number of detected SIFT keypoints that fall within the i_{th} component. The local SIFT-based descriptors in all the k components can be denoted by:

$$I = (f_1^1,...,f_1^{m_1}, f_2^1,...,f_2^{m_2},...,f_k^1,...,f_k^{m_k}) \quad (5)$$

where f_i^j means the j_{th} descriptor in the i_{th} facial component. Then the similarity between a gallery face I_g and a probe face I_p is computed by:

$$C(I_p, I_g) = \frac{1}{k}\sum_{i=1}^{k}(\max(\frac{< f_{p_i}^x, f_{g_i}^y >}{\left\| f_{p_i}^x \right\| \cdot \left\| f_{g_i}^y \right\|})) \quad (6)$$

where $x \in [1, ..., m_{pi}]$; $y \in [1, ..., m_{gi}]$. $<\cdot>$ denotes the inner product of two vectors, and $\| \cdot \|$ denotes the norm of one vector. A bigger C indicates the most similar attributes of the two faces represented by *MS-eLBP-DF*s. We thus obtain similarity values, $C_{Li(P,\ R)}$ for each *MS-eLBP-DF*.

2) Facial Configuration Constraint: the former constraint on facial component emphasizes the importance of the matching score between local features of the same facial component-based area in the gallery and probe set, and we further improve the holistic constraint by facial configuration inheriting the local feature based graph matching implemented in [33].

All facial range images are normalized to a certain size to build a public 2D coordinate system. For each 3D face scan, the *MS-eLBP-DF*s are extracted from the range image. Therefore, all the keypoints of the proposed facial surface representations share the same XY-plane with the range face image, and the pixel values of the corresponding facial range image can be regarded as the Z-axis values of these keypoints. Hence, each keypoint has its position in 3D space. After local feature-based matching, a 3D graph is formed for each *MS-eLBP-DF* of a probe F_p, by simply linking every two keypoints which have a matching relationship with keypoints detected on the corresponding *MS-eLBP-DF* of a gallery face F_g. The matched keypoints of F_g also construct a corresponding graph of F_p. Since all the facial range images are of the same scale, intuitively, if faces F_g and F_p are from the same subject, their corresponding graphs should have similar shapes in 3D space.

The similarity measure between the two graphs can be defined as:

$$D = \frac{1}{n_e}\sum_{i=1}^{n_e}\left| d_{p_i} - d_{g_i} \right| \quad (7)$$

where d_{pi} and d_{gi} are the lengths of corresponding edges in the probe and gallery graphs respectively. The value n_e is the total number of edges. If the number of matched keypoints is n_n, n_e will be n_n*(n_n -1)/2. Equation 7 is an efficient way to measure the spatial error between the matched keypoint pairs of probe and gallery features. As in the facial component constraint, here, $D_{Li(P,\ R)}$ denotes the similarity score of each *MS-eLBP-DF*.

4.3 Similarity Fusion

In summary, the matching process of gallery and probe facial range images contains three types of similarities: the number of matched keypoint pairs N, similarity C of the facial component constraint and similarity D of the facial configuration constraint. Except for D, all the other similarity measures have a positive polarity (a bigger value means a better matching relationship). A range face of the probe set is matched with every face in the gallery, resulting in three vectors S_N, S_C and S_D. The n_{th} element of each score vector corresponds to the similarity score between the probe and the n_{th} gallery face. Each vector is normalized to the interval of [0, 1] using the min-max rule. Elements of S_D are subtracted from 1 to reverse its polarity. The final similarity of the probe face with the ones in the gallery is calculated using a basic weighted sum rule:

$$S = w_N * S_N + w_C * S_C + w_D * (1 - S_D) \quad (8)$$

We further make use of the original weight calculation as proposed by Mian et al. in [33] to dynamically determine their corresponding weights: w_N, w_C, and w_D during the online step:

$$w_{S_i} = \frac{mean(S_i) - \min_1(S_i)}{mean(S_i) - \min_2(S_i)} \quad (9)$$

where i corresponds to the three similarities: N, C, and D, and operators $max_1(S_i)$ and $max_2(S_i)$ produce the first and second maximum value of vector S_i. The gallery face which has the maximum value in vector S is declared as the identity of the probe face image when the decision is to be made on each *MS-eLBP-DF* independently.

5 Experimental Results

To demonstrate the effectiveness of the proposed method, we utilized three public datasets for experiments, *i.e.* FRGC v2.0 [48], Bosphorus [68] and Gavab DB [39]. The first is for evaluating performance with large number of subjects that have slight pose variations and various facial expression changes; the second is to further observe its robustness to expression variations as well as external occlusions; while the last one is to analyze its accuracy on

3D face samples with extreme pose changes (left and right profiles).

5.1 Experiments on the FRGC v2.0 Dataset

5.1.1 Database and protocol description

Some experiments were evaluated on FRGC v2.0 [48], one of the most comprehensive and popular datasets, containing 4007 3D face scans of 466 different subjects. One facial range image was extracted from each 3D face model. A preprocessing step was applied to remove noises with a median filter and fill holes using cubic interpolation, and the technique details can be found in [57]. Thanks to the relative tolerance to moderate pose changes of hybrid matching, we did not perform any registration on these 3D face models, in contrast to most works such as [31], [49], [56] *etc.* The facial range images were automatically cropped using a basic bounding box, which was roughly located according to the vertical and horizontal integral projection of the mask provided by a 3D scanner indicating if the point was valid or not in that position. Cropped faces have moderate pose, expression changes, as well as partial occlusions caused by hair. All the faces are normalized to 150×150 pixels. Fig. 11 shows some preprocessed samples for the following recognition step.

Fig. 11. Examples of preprocessed 3D facial scans of the same subject from the FRGC dataset.

The proposed method was evaluated by face recognition and verification tasks. In order to compare our method with the state of the art, we followed the same experimental protocol that the first 3D face scan with a neutral expression from each subject makes up a gallery of 466 samples and the remaining face scans (4007-466=3541) are treated as probes. The probe face scans were divided into two subsets according to their facial expression labels to evaluate its robustness to facial expression variations. The first subset is made up of facial scans with a neutral expression; whilst the other one with facial scans possessing non-neutral facial expressions. Besides the experiment of Neutral vs. All, two additional experiments: Neutral vs. Neutral and Neutral vs. Non-Neutral were also included. In the Neutral vs. Neutral and Neutral vs. Non-Neutral experiments, only the neutral and non-neutral subsets were used, respectively.

Based on its default setting in SIFT matching [46], we compared the values of t from 0.5 to 0.9 with an increasing step of 0.05, and in the range of [0.6, 0.8], the performance basically keeps stable and outperforms the others. We thus set t at 0.6 in the following experiments.

5.1.2 The influences of eLBP parameters (Experiment A)

The four sub-tables in Table I list the results based on depth faces of each *eLBP* layer with different parameters. Recall that P is the number of sampling points and R is the radius value.

In these sub-tables, all *eLBP* accuracies at different scales with different numbers of sampling points display recognition rates better than 90%, greatly outperforming the ones based on original *LBP* operators (*eLBP* L_1). The accuracies displayed in the last row labeled as the '*eLBP*' performance are fusion results according to the weighted sum rule in (8) combining the scores provided by the first three layers (L_1, L_2 and L_3) using the same parameter setting; similarity scores at L_4 are omitted because of their low performance. As we can see from Table I, using 8 sampling points achieves better results on L_1, and L_2 for almost all radius values (except R = 2, 3), respectively; whilst the setting with 16 sampling points results in better performance on L_3 (except R=2) and L_4, respectively.

TABLE I
RANK-ONE RECOGNITION RATES BASED ON EACH *eLBP* LAYER OF THE DEPTH FACES WITH DIFFERENT PARAMETERS IN FRGC V2.0.

(A)

P=4	*R=2*	*R=3*	*R=4*	*R=5*	*R=6*	*R=7*	*R=8*
eLBP L1	81.6%	84.8%	86.9%	87.7%	87.6%	86.2%	85.9%
eLBP L2	75.2%	83.3%	85.7%	87.1%	87.6%	87.3%	87.0%
eLBP L3	76.9%	74.7%	71.6%	68.8%	67.4%	63.7%	61.9%
eLBP L4	4.5%	8.0%	12.7%	16.0%	25.9%	33.2%	40.6%
eLBP	90.0%	90.9%	92.0%	92.6%	92.4%	92.3%	92.3%

(B)

P=8	*R=2*	*R=3*	*R=4*	*R=5*	*R=6*	*R=7*	*R=8*
eLBP L1	**86.1%**	**87.8%**	**88.5%**	**88.3%**	**87.7%**	**86.6%**	**86.0%**
eLBP L2	73.6%	84.6%	**88.6%**	**89.2%**	**89.2%**	**89.3%**	**89.9%**
eLBP L3	80.1%	78.3%	76.4%	76.3%	75.6%	76.6%	76.3%
eLBP L4	6.6%	11.1%	17.8%	29.8%	40.3%	50.8%	55.6%
eLBP	**91.3%**	**92.5%**	**93.5%**	**93.4%**	**93.0%**	**92.7%**	**92.6%**

(C)

P=12	*R=2*	*R=3*	*R=4*	*R=5*	*R=6*	*R=7*	*R=8*
eLBP L1	85.3%	86.1%	86.2%	87.2%	85.8%	85.4%	84.6%
eLBP L2	71.7%	84.4%	87.3%	88.6%	89.3%	88.9%	88.4%
eLBP L3	**81.9%**	78.7%	78.1%	76.6%	78.5%	78.9%	79.6%
eLBP L4	6.2%	12.3%	22.1%	35.6%	48.7%	57.4%	63.2%
eLBP	90.9%	92.1%	92.9%	93.3%	92.3%	92.3%	91.5%

(D)

P=16	*R=2*	*R=3*	*R=4*	*R=5*	*R=6*	*R=7*	*R=8*
eLBP L1	82.1%	82.9%	85.3%	84.2%	84.3%	83.5%	82.7%
eLBP L2	**73.7%**	**86.1%**	87.9%	88.6%	88.2%	87.5%	87.7%
eLBP L3	81.6%	**80.0%**	**78.7%**	**78.4%**	**79.4%**	**79.1%**	**79.7%**
eLBP L4	**7.2%**	**11.8%**	**27.7%**	**42.3%**	**52.3%**	**60.0%**	**66.1%**
eLBP	90.6%	91.9%	92.4%	92.4%	91.8%	91.6%	91.6%

5.1.3 Face identification performance (Experiment B)

Using the weighted sum rule described in (8), we then fused the similarity measurements of *eLBP* with 8 sampling points and different values of radius from 2 to 8, and compared the rank-one face recognition rate achieved with the state of the art in Table II. Except for ICP, all the results are cited from the original papers.

In order to test the discriminative power of *LBP* and *eLBP* to characterize local geometric shapes, Shape Index (*SI*) faces are also produced and associated with the proposed hybrid matching for comparison in 3D face recognition. Recall that an SI face is generated by computing

the shape index value on each pixel location from a facial range image and quantizing that values to the range of gray level images, *i.e.* [0, 255]. Further technique details can be found in [50]. With a rank-one recognition rate of 91.8% as indicated in table II, *SI* faces outperform any of the single scale *LBP-DFs* (*i.e.* the layer of *eLBP L1*). Meanwhile, most of the results based on single scale *eLBP-DF* surpass that of *SI* faces; furthermore, when fusing the matching scores of the *eLBP-DF*s at different scales to achieve *MS-eLBP-DF*s, the rank-one recognition rate is increased by more than 5 points, from 91.8% for *SI* face to 97.6% for *MS-eLBP-DF*s. These results clearly indicate how well Multi-Scale *eLBP* describes geometric shape variations. On the other hand, from the comparison with the state of the art, we can see that our result is comparable to the best results reported in literature.

TABLE II
RANK-ONE RECOGNITION RATES USING THE PROTOCOL OF NEUTRAL VS. ALL ON THE FRGC V2.0 DATASET.

	Rank-one RR
(1) ICP	72.2%
(2) *SI* Faces	91.8%
(3) *MS-LBP-DF*s	94.1%
(4) Chang et al. 2005 [52]	91.9%
(5) Cook et al. 2006 [55]	94.6%
(6) Mian et al. 2007 [49]	96.2%
(7) Wang et al. 2007 [51]	87.7%
(8) Mian et al. 2008 [33]	93.5%
(9) Kakadiaris et al. 2007 [31]	97.0%
(10) Faltemier et al. 2008 [56]	98.1%
(11) Huang et al. 2010 [50]	96.1%
(12) Alyüz et al. 2010 [67]	97.5%
(13) Queirolo et al. 2010 [66]	98.4%
(14) Wang et al. 2010 [26]	98.4%
(15) *MS-eLBP-DF*s	**97.6%**

5.1.4 Robustness to facial expression changes (Experiment C)

Using the same experimental protocol, we also compared the performance of the proposed approach with those in the literature for robustness analysis on facial expression changes (see Table III). The results of our approach are 99.2% and 95.1% for the Neutral vs. Neutral and Neutral vs. Non-Neutral experiment, respectively. The recognition rates on Subset I and subset II are comparable to the best ones of the state-of-the-art (Subset I in [33] and Subset II in [67]). Moreover, Table III also indicates that the *MS-eLBP-DF*s outperform the *SI* face in both the additional experiments on Subset I and II, and the performance degradation of *MS-ELBP-DF*s is much lower than that of *SI* face. These accuracies hence suggest that our approach tends to be insensitive to facial expression changes.

5.1.5 Face verification performance (Experiment D)

The proposed approach was evaluated for face verification as well using the three protocols, *i.e.* Neutral vs. All, Neutral vs. Neutral, Neutral vs. Non-Neutral, and the results are displayed in Table IV. From the first column (VR I), we can see that our accuracy is also among the best ones of the state of the art, while the results in the third column (VR III) illustrate once again that the proposed method performs quite well when recognizing expressive faces. Both the facts match the phenomenon in face recognition. Fig. 12 indicates the verification rates by the ROC curves in the three experiments in Table IV.

TABLE III
RANK-ONE PERFORMANCE USING THE EXPRESSION PROTOCOL ON THE FRGC V2.0 DATASET.

	SUBSET I	SUBSET II	Degradation
***SI* Faces**	97.2%	84.1%	13.1%
***MS-LBP-DF*s**	97.7%	88.9%	8.8%
***MS-eLBP-DF*s**	99.2%	**95.1%**	**4.1%**
Huang et al. 2010 [50]	99.1%	92.5%	6.6%
Alyüz et al. 2010 [67]	98.4%	96.4%	2.0%
Mian et al. 2008 [33]	99.0%	86.7%	12.3%
Mian et al. 2008 [33]	99.4%	92.1%	7.3%
Kakadiaris et al. 2007 [31]	99.0%	95.6%	3.4%

SUBSET I: NEUTRAL VS. NEUTRAL
SUBSET II: NEUTRAL VS. NON-NEUTRAL

TABLE IV
COMPARISON OF VERIFICATION RATES AT 0.001 FAR USING THE EXPRESSION PROTOCOL ON THE FRGC V2.0 DATASET.

	VR I	VR II	VR III
***SI* Faces**	94.4%	98.9%	87.5%
***MS-LBP-DF*s**	96.1%	99.1%	91.9%
***MS-eLBP-DF*s**	**98.4%**	99.6%	**97.2%**
Maurer et al. 2005 [15]	92.0%	97.8%	NA
Passalis et al. 2005 [53]	85.1%	94.9%	79.4%
Husken et al. 2005 [54]	89.5%	NA	NA
Cook et al. 2006 [55]	95.8%	NA	NA
Mian et al. 2008 [33]	97.4%	99.9%	92.7%
Mian et al. 2007 [49]	98.5%	NA	NA
Wang et al. 2010 [26]	98.6%	NA	NA

VR I: NEUTRAL VS. ALL
VR II: NEUTRAL VS. NEUTRAL
VR III: NEUTRAL VS. NON-NEUTRAL

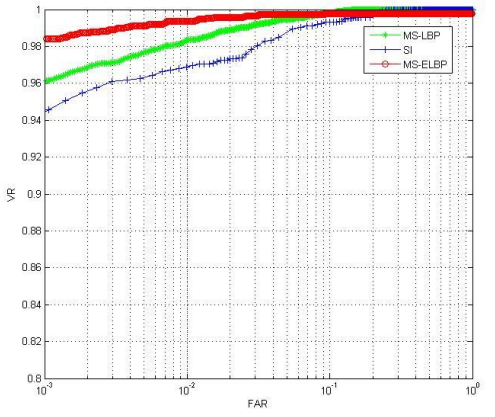
TABLE V
COMPARISONS OF VERIFICATION RATES AT 0.001 FAR USING ROC I, ROC II, ROC III AND ALL VS. ALL PROTOCOL ON THE FRGC V2.0 DATASET.

	ROC I	ROC II	ROC III	All vs. All
Maurer et al. [15]	NA	NA	92.0%	87.0%
Cook et al. [55]	93.7%	92.9%	92.0%	92.3%
Husken et al. [54]	NA	NA	89.5%	NA
Faltemier et al. [56]	NA	NA	94.8%	93.2%
Kakadiaris et al. [31]	97.3%	97.2%	97.0%	NA
Mian et al. [49]	NA	NA	NA	86.6%
Alyüz et al. [67]	85.4%	85.6%	85.6%	NA
Queirolo et al. [66]	NA	NA	96.6%	96.5%
Wang et al. [26]	98.0%	98.0%	98.0%	98.1%
***MS-eLBP-DF*s**	**95.1%**	**95.1%**	**95.0%**	**94.2%**

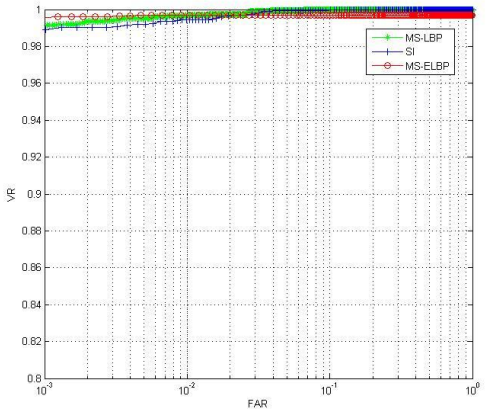
5.1.6 Evaluation by aging factors (Experiment E)

Further experiments were carried out on ROC I, ROC II, and ROC III, and these three ROC curves are based on the three masks provided by the FRGC database. They are

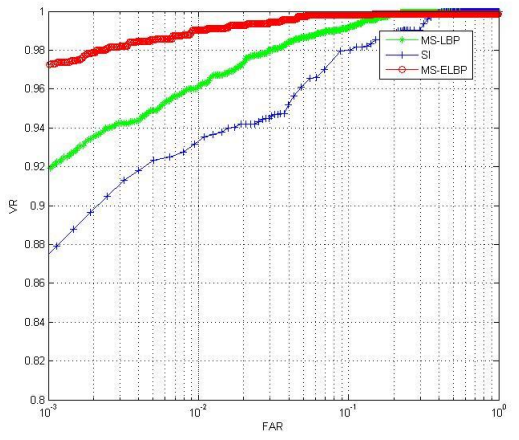
defined over the square similarity matrix with a dimensionality of 4007×4007, and they are of increasing difficulty reflecting the time elapsed between the probe and gallery acquisition sessions. The comparisons are shown in Table V.



(A)



(B)



(C)

Fig. 12. ROC curves using SI Faces, MS-LBP-DFs, and MS-eLBP-DFs respectively in the experiments with neutral faces enrolled. (A) Neutral vs. All. (B) Neutral vs. Neutral. (C) Neutral vs. Non-neutral.

As we can see from Table V, the performance of the proposed method on ROC I, ROC II and ROC III is slightly lower but still close to the best of ones in the literature. Meanwhile, it is noteworthy that our method does not require any registration for nearly frontal face scans such as those in the FRGC dataset. This is clearly in contrast to works [26], [31] and [66]. In [31], Kakadiaris et al. used ICP-based alignment with a coarse to fine strategy for preprocessing, and two kinds of features, *i.e.* Haar and Pyramid were extracted from both normal and geometry maps. In [26], a self-dependent registration step was employed, and a large training database with neutral and non-neutral facial scans is required to learn inter-class and intra-class changes, and fused Haar-, Gabor- and LBP-based facial features for the final decision. In [66], a Simulated Annealing based approach was adopted for range image registration and similarity calculation for corresponding regions of different facial surfaces which were segmented in the previous step.

5.2 Experiments on the Bosphorus Dataset

5.2.1 Database and protocol description

To further confirm the effectiveness of the proposed approach to identify nearly frontal faces without any registration and prove its robustness to expression and pose variations as well as partial occlusions, the Bosphorus database is considered as well, consisting of a large number of 3D face models with extreme pose changes, expression variations (both emotions and action units), and typical occlusions that may occur in real life. The database includes totally 4666 scans collected from 105 subjects, each of whom possesses around 34 expressions, 13 poses, and 4 occlusions. In our experiments, a subset of the dataset containing only nearly frontal face model is collected, and it thus has a total of 3301 scans with roughly 34 different expressions and 4 external occlusions per subject. Since each subject from Bosphorus only has one or two neutral samples while the others are the scans with facial expressions or partial occlusions as we can see in Fig. 13, this subset is potentially much more challenging than FRGC v2.0 to the proposed approach. With rare exceptions like [67] and [69], so far very few works in 3D face recognition have tested their approach on this dataset.

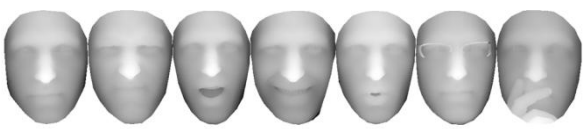


Fig. 13. Some example 3D facial scans of the same subject from the Bosphorus dataset.

Similar to the preprocessing step that was carried out on FRGC v2.0, we only conducted spike removal and hole filling. As the experimental protocol, we constructed a gallery set containing the first neutral facial scan for each subject, and the remaining ones made up of the probe set. Hence, the sizes of gallery and probe set are 105 and 3196, respectively. As in the experiments on the FRGC dataset, there was no registration of 3D face scans.

5.2.2 Experiments on nearly frontal faces with expression changes and occlusions (Experiment F)

In the identification scenario, we computed the rank-one recognition rate. Table VI lists the performance of differ-

ent approaches for comparison. It can be seen from Table VI that the proposed method achieved a recognition rate up to 97.0%, further illustrating its effectiveness to recognize nearly frontal faces without registration and highlighting its robustness to facial expression variations and external occlusions. In comparison with the state of the art by using a similar experimental protocol, the recognition rate is comparable to the one in [69], both of which are slightly inferior to that reported by [67]. The difference between our experimental setup and the one in [67] lies in that the probe set used in [67] excluded the faces with occlusions. Moreover, as the method in [67] also made use of a training set formed by scans from 20 subjects in order to learn LDA subspaces, their probe set was formed by the 3D facial scans of 85 subjects.

TABLE VI
RANK-ONE RECOGNITION RATES ON THE BOSPHORUS DATASET.

Approaches	Rank-one RR
(1) *SI* Faces	89.7%
(2) *MS-LBP-DF*s	93.4%
(3) *MS-eLBP-DF*s	**97.0%**
(4) Maes et al. 2010 [69]	97.7%
(5) Alyüz et al. 2010 [67]	99.3%*

* The experimental protocol in [67] has some differences from that in this study and [69].

5.3 Experiments on the Gavab Dataset

5.3.1 Database and protocol description

To analyze the performance on severe pose changes and even partially occluded 3D facial scans, we also tested our method on the Gavab database. To the best of our knowledge, Gavab is the most noise-prone dataset currently available to the public. This database consists of Minolta Vi-700 laser facial range scans of 61 different subjects. The subjects, of whom 45 are male and 16 are female, are all Caucasian. Each subject was scanned 9 times for different poses and facial expressions. The scans with pose variations contain one facial scan while looking up (+35 degree), one while looking down (-35 degree), one for the right profile (+90 degree), one for the left profile (-90 degree) as well as one with random poses The scans without pose changes include two different frontal facial scans, one with a smile, and one with an accentuated laugh. Figure 14 shows some examples of faces in this dataset.

Fig. 14. Examples of all the 3D facial scans of the same subject from the Gavab dataset [58].

A similar preprocessing step was utilized as in FRGC to remove spikes and fill holes. Since the Gavab contains many severe pose changes, we performed a coarse alignment based on three landmarks for all facial scans. When the two inner corners of the eyes and the nose tip of one scan are available at the same time (all face scans excluding the extreme poses such as the right and left profiles), we used our previous method [57] to find the three landmark points automatically and computed rotation and translation parameters; while for each of these left or right profiles, we manually landmarked four points, *i.e.*, the inner and outer corner of one eye, nose tip, and the corner of the nose, which are visible in that profile. After coarse registration, one range image is extracted from each facial scan; these range images hence only contain partial faces due to the self-occlusion caused by pose variations, and all facial range images are further resized to 150×150 pixels.

In our experiments, the first frontal facial scan of each subject was used as the gallery; while the others were treated as probes. We calculated rank-one face recognition rates, and Table VII shows matching accuracies for different categories of probe faces: (A) displays the results without pose variations; while (B) lists those only with the facial scans with pose changes. In (A), the neutral subset contains one frontal facial scan of each subject, and the expressive subset includes a smile, accentuated laugh and random gesture (random facial expression), three scans of each subject. To the best of our knowledge, work [58] is the only one that carried out experiments on the entire Gavab dataset before this work. Therefore, we compared our results with theirs on the subset of four severe pose variations as well as the overall performance. It is worth noting that the difference between their work and ours is that Drira et al. manually landmarked nose tips on all the face scans in the dataset for an ICP-based fine registration, while we only manually landmarked facial scans of right and left profiles ((c) and (d) in Table VII (B)), and for all the faces, only a coarse alignment is utilized to rotate and translate them.

5.3.2 Evaluation on the faces with large pose variations (Experiment F)

From Table VII (A), we can see that for frontal neutral probes, the rank-one recognition rate is 100% as in [58]; while regarding expressive faces, our approach surpasses all the others. Moreover, when evaluating the robustness to severe pose variations (Table VII (B)), we achieved an overall accuracy of 91.4% on these four subsets; whilst that reported by [58] is 88.9%.

To sum up, the experimental results on the Gavab dataset clearly prove that only aided by a coarse alignment, our method can deal with large pose changes and even partial occlusions.

TABLE VI
COMPARISONS OF RANK-ONE RECOGNITION RATES ON THE GAVAB DATASET: (A) WITHOUT POSE VARIATIONS; (B) ONLY WITH POSE VARIATIONS

(A)

	I. Neutral	II. Expressive	I + II
Li et al. [59]	96.67%	93.33%	94.68%
Moreno et al. [60]	90.16%	77.90%	NA
Mahoor et al. [61]	95.00%	72.00%	78.00%
Berretti et al. [62]	94.00%	81.00%	84.25%
Mousavi et al. [63]	NA	NA	91.00%
Drira et al. [58]	100.00%	NA	94.67%
***MS-eLBP-DF*s**	**100.00%**	**93.99%**	**95.49%**

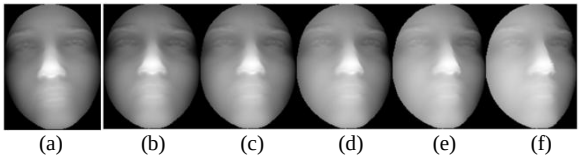
(B)

	(a)	(b)	(c)	(d)	(e)
Mahoor et al. [61]	85.30%	88.60%	NA	NA	NA
Berretti et al. [62]	80.00%	79.00%	NA	NA	NA
Drira et al. [58]	100.00%	98.36%	70.49%	86.89%	88.94%
***MS-eLBP-DF*s**	**96.72%**	**96.72%**	**78.69%**	**93.44%**	**91.39%**

(a): Looking down
(b): Looking up
(c): Right Profile
(d): Left profile
(e): Overall

5.3.3 Analysis of the impact by pose changes on face recognition accuracy (Experiment H)

The previous experiments tend to suggest that our method can cope with moderate 3D face pose changes, in particular for nearly frontal face models as this can be the case for user cooperative biometric applications (*e.g.*, access control), without any prior registration. When aided by a coarse alignment step using very few landmarks, the proposed method can even deal with 3D face models with large pose variations such as the ones in the Gavab dataset. However, the question remains of determining to what extent the proposed approach is still robust in pose changes. To give some insight into this question, we designed the following experiment. Two nearly frontal facial scans of the same subject were randomly chosen from the Gavab database: one used as the gallery face and the other as the probe. We rotate the probe one around the yaw axis by 5, 10, 15 and 20 degrees, respectively, thus leading to range images (shown in Fig. 15) subsequently employed for the analysis on the impact of rotated faces to the matching result.



(a) (b) (c) (d) (e) (f)

Fig. 15. Face samples for matching: (a) gallery face (nearly frontal); (b) probe face (nearly frontal); (c) probe face rotated by 5 degrees; (d) probe face rotated by 10 degrees; (e) probe face rotated by 15 degrees; (f) probe face rotated by 20 degrees.

We generated the *eLBP* based facial representations for a gallery face (a) as well as each of the probe faces (b-f). We then computed the number of keypoints detected on each probe face (c)-(f) and found that the numbers almost remain stable when compared with the one of the original probe face (b). In the matching phase, when the probe face Fig.15 (b) is rotated successively from Fig.15 (c) to (f) each time by 5 degrees, the number of matched keypoints with the gallery face (a) keeps decreasing, as illustrated in Figure 16, from roughly 300 pairs in Fig.16 (a) down to only around 10 pairs in Fig.16 (e) when the probe face is rotated by 20 degrees from the frontal one. Meanwhile, the main quality of a similarity measurement for the purpose of face recognition is its ability to distinguish intra-class changes from inter-class ones. From this perspective, the number of matched keypoints on each *eLBP-DF* pair used as face similarity measurement between the gallery face and the probe rotated by 15 degrees can still clearly discriminate the facial scans of the same subject from the ones of different subjects. Indeed, the average number of matched keypoints in the former case is above 20 even as the face is rotated by 15 degrees while this number over the whole Gavab dataset excluding the left and right profiles is close to 10 in the latter case. However, when probe facial scans are rotated by 20 degrees, it becomes difficult to discriminate intra-class variations from inter-class ones in terms of the matched keypoint pairs. We conducted the experiment 10 times, the phenomena are similar.

5.4 Experiment Summary

In the sub-section 5.1, 5.2 and 5.3, eight experiments (Experiment A-H) are carried out on three public datasets, *i.e.* FRGC v2.0, Bosphours and Gavab. The results achieved in Experiment B, D and E demonstrate that our method, without the registration step for nearly frontal faces displays comparable to the best ones so far presented in the literature. Moreover, Experiments C and F show that the proposed approach is robust to facial expression variations whereas experiment F further highlights its tolerance to partial external occlusions. Additionally, Experiment G indicates that supported by a coarse registration step applied on only a few landmarks, our approach can deal with the faces with large pose variations and even self-occlusions.

At the same time, we observe the computational expense of the proposed approach along with the experiments conducted on the FRGC v2.0 database. Currently, an unoptimized implementation of the proposed method with MATLAB (R2010a) can perform one match between one pair of *eLBP-DF*s of the gallery and probe faces in about 0.32s using a machine where Intel(R) Core(TM) i5 CPU (2.60 GHz) and 4 GB RAM are equipped. Since the similarity scores based on different *eLBP-DF*s can be calculated independently, if implemented in a parallel computing device, our method is quite promising in providing decisions in real time.

6 Conclusion

We have presented an effective method to 3D face recognition using a novel geometric facial representation and local feature hybrid matching. The proposed facial representation is based on *MS-eLBP* and allows for accurate and fast description of local shape variations, thus enhancing the distinctiveness of range faces. SIFT-based hybrid matching that combines local and holistic analysis further robustly associates keypoints between two faces of the same subject. The proposed method was evaluated in 3D face recognition and verification, achieving a recognition rate of 97.6% and a 98.4% verification rate with a 0.001 FAR respectively on the FRGC v2.0 database which consists of nearly frontal 3D facial scans with rich facial expression changes. Additional experiments on the Bosphorus database further confirm the advantages of the proposed approach with regard to facial expression variations and external partial occlusions. The results achieved on the Gavab DB dataset containing severe pose changes

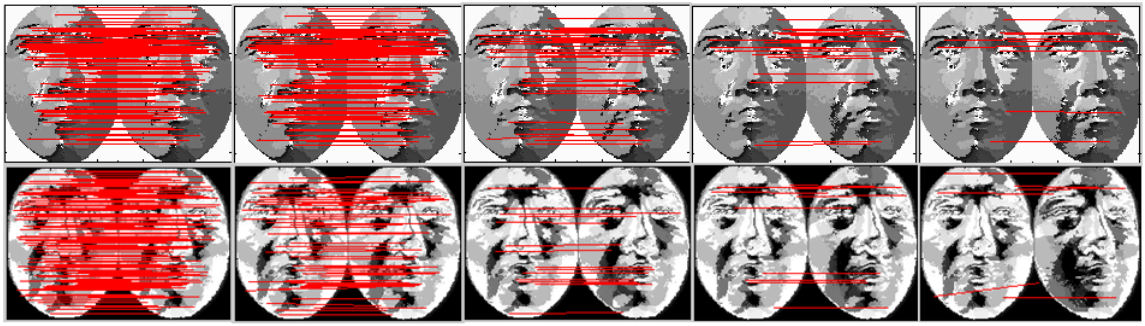


Fig. 16. A face matching demonstration. The upper maps are generated by eLBP(8,3)-L1; while the bottom ones with eLBP(8,3)-L2. (a) matching with nearly frontal probe face; (b) matching with probe rotated by 5 degrees; (c) matching with probe rotated by 10 degrees; (d) matching with probe rotated by 15 degrees; (e) matching with probe rotated by 20 degrees.

clearly illustrate that the entire system also provides a promising solution to recognizing partially occluded faces. Moreover, generally costly registration was not needed thanks to the relative tolerance of the proposed hybrid matching strategy to nearly frontal faces like the ones in the FRGC v2.0 and the subset of Bosphorus. When dealing with extreme poses, *e.g.* left or right profiles, a coarse alignment step based on a few manually landmarked points was sufficient in preprocessing as indicated by the experiments on the Gavab database.

ACKNOWLEDGMENT

This work was in part supported by the National Natural Science Foundation of China under Grant 61061130560, the French research agency, Agence Nationale de Recherche (ANR), within the ANR FAR 3D project (grant ANR-07-SESU-004-03), the Sino-French project 3D Face Analyzer (grant ANR 2010 INTB 0301 01), and the LIA 2MCSI laboratory between the group of Ecoles Centrales, France and Beihang University, China, through the 3D face interpreter project.

REFERENCES

[1] A. K. Jain, A. Ross, and S. Prabhakar. "An introduction to biometric recognition," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4–20, 2004.

[2] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Survey*, vol. 35, no. 4, pp. 399–458, 2003.

[3] K. W. Bowyer, K. Chang, and P. J. Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition," *Computer Vision and Image Understanding*, vol. 101, no. 1, pp. 1-15, 2006.

[4] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: a survey," *Pattern Recognition Letters*, vol. 28, no. 14, pp. 1885-1906, 2007.

[5] P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090-1104, 2000.

[6] A. Scheenstra, A. Ruifrok, and R. C. Veltkamp, "A survey of 3D face recognition methods," in *Proc. International Conference on Audio- and Video-Based Biometric Person Authentication*, 2005.

[7] A. Hadid, M. Pietikäinen, and T. Ahonen, "A discriminative feature space for detecting and recognizing faces," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 797–804, 2004.

[8] B. Achermann, X. Jiang, and H. Bunke, "Face recognition using range images," in *Proc. International Conference on Virtual Systems and MultiMedia*, pp.129-136, 1997.

[9] C. Hesher, A. Srivastava, and G. Erlebacher, "A novel technique for face recognition using range imaging," in *Proc. International Symposium on Signal Processing and Its Applications*, pp. 201-204, 2003.

[10] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Expression invariant 3D face recognition," in *Proc. International Conference on Audio- and Video-Based Biometric Person Authentication*, pp. 62-70, 2003.

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

[12] X. Lu, A. K. Jain and D. Colbry, "Matching 2.5D face scans to 3D models," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 31-43, 2006.

[13] X. Lu and A. K. Jain, "Deformation modeling for robust 3D face matching," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1377-1383, 2006.

[14] K. Chang, K. W. Bowyer, and P. Flynn, "Effects on facial expression in 3D face recognition," in *Proc. SPIE Conference Biometric Technology for Human Identification*, vol. 5779, pp. 132-143, 2005.

[15] T. Maurer, D. Guigonis, I. Maslov, B. Pesenti, A. Tsaregorodtsev, D. West, and G. Medioni, "Performance of geometrix active ID™ 3D face recognition engine on the FRGC data," *IEEE Workshop on FRGC Experiments*, 2005.

[16] Y. Lee and J. Shim, "Curvature-based human face recognition using depth-weighted Hausdorff distance," in *Proc. International Conference on Image Processing*, 2004, pp. 1429-1432.

[17] T. D. Russ, M. W. Koch, C. Q. Little, "A 2D range Hausdorff approach for 3D face recognition," *IEEE Workshop on FRGC Experiments*, 2005.

[18] G. G. Gordon, "Face recognition based on depth and curvature features," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1992, pp. 808-810.

[19] A. B. Moreno, A. Sanchez, J. F. Velez, and F.J. Diaz, "Face recognition using 3D surface-extracted descriptors," in *Proc. Irish Machine Vision and Image Processing Conference*, 2003.

[20] S. Gupta, J. K. Aggatwal, M. K. Markey, and A. C. Bovik, "3D face recognition founded on the structural diversity of human faces," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[21] T. Nagamine, T. Uemura, and I. Masuda, "3D facial image analysis for human identification," in *Proc. International Conference on Pattern Recognition*, 1992, pp. 324-327.

[22] C. Beumier and M. Acheroy, "Automatic 3D face authentication," *Image Vision Computing*, vol. 18, no. 4, pp. 315-321, 2000.

[23] C. Samir, A. Srivastava, and M. Daoudi, "Three-dimensional face recognition using shapes of facial curves," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1858-1863, 2006.

[24] C. Chua, F. Han, and Y. Ho, "3D human face recognition using point signature," in *Proc. International Conference on Automatic Face and Gesture Recognition*, pp. 233-238, 2000.

[25] H. T. Tanaka, M. Ikeda, and H. Chiaki, "Curvature-based face surface recognition using spherical correlation — principal directions for curved object recognition," in *Proc. International Conference on Automatic Face and Gesture Recognition*, pp. 372-377, 1998.

[26] Y. Wang, J. Liu, and X. Tang, "Robust 3D face recognition by local shape difference boosting," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 10, pp. 1858-1870, 2010.

[27] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuro-Science*, vol. 3, no. 1, pp. 71-86, 1991.

[28] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.

[29] L. Wiskott, J. M. Fellous, N. Kruger, and C. v. d. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775–779, 1997.

[30] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 84-91, 1994.

[31] I. A. Kakadiaris, G. Passalis, G. Toderici, M. N. Murtuza, Y. Lu, N. Karampatziakis, and T. Theoharis, "Three-dimensional face recognition in the presence of facial expressions: an annotated deformable model approach," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 640-649, 2007.

[32] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Three dimensional face recognition," *International Journal of Computer Vision*, vol. 64, no. 1, pp. 5-30, 2005.

[33] A. S. Mian, M. Bennamoun, and R. Owens, "Keypoint detection and local feature matching for textured 3D face recognition," *International Journal of Computer Vision*, vol. 79, no. 1, pp. 1-12, 2008.

[34] K. Ouji, B. B. Amor, M. Ardabilian, L. Chen, and F. Ghorbel, "3D face recognition using R-ICP and geodesic coupled approach," in *Proc. International conference on on MultiMedia Modeling*, 2009.

[35] A. S. Mian, M. Bennamoun, and R. Owens, "Region-based matching for robust 3D face recognition," in *Proc. British Machine Vision Conference*, 2005.

[36] J. Huang, B. Heisele, and V. Blanz, "Component-based face recognition with 3D morphable models," in *Proc. International Conference on Audio- and Video-Based Biometric Person Authentication*, pp. 27-34, 2003.

[37] S. Z. Li, C. Zhao, M. Ao, and Z. Lei, "Learning to fuse 3D+2D based face recognition at both feature and decision levels," in *Proc. workshop on Analysis and Modeling of Faces and Gestures*, 2005.

[38] Y. Huang, Y. Wang, T. Tan, "Combining statistics of geometrical and correlative features for 3D face recognition," in *Proc. British Machine Vision Conference*, 2006.

[39] A. B. Moreno and A.Sanchez, "GavabDB: a 3D face database," in *Proc. COST Workshop on Biometrics on the Internet: Fundamentals, Advances and Applications*, pp. 77-82, 2004.

[40] T. Ojala, M. Pietikäinen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[41] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Proc. European Conference on Computer Vision*, 2004.

[42] J. J. Koenderink and A. J. Doorn, "Surface shape and curvature scales," *Image Vision Computing*, vol. 10, no. 8, pp. 557-565, 1992.

[43] S. Yan, H. Wang, X. Tang, and T. S. Huang, "Exploring feature descriptors for face recognition," in *Proc. International Conference on Acoustics, Speech, and Signal Processing*, 2007.

[44] C. Chan, J. Kittler, and K. Messer, "Multi-scale local binary pattern histograms for face recognition," in *Proc. International Conference on Biometrics*, pp. 809-818, 2007.

[45] C. Shan and T. Gritti, "Learning discriminative LBP-histogram bins for facial expression recognition," in *Proc. British Machine Vision Conference*, 2008.

[46] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 4, pp. 91-110, 2004.

[47] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade, and B. Lu, "Person-specific SIFT features for face recognition," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2007.

[48] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, K. I. Chang, K. Hoffman, J. Marques, J. Min, W. Worek, "Overview of the face recognition grand challenge," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, vol. I, pp. 947-954, 2005.

[49] A. S. Mian, M. Bennamoun, and R. Owens, "An efficient multimodal 2D-3D hybrid approach to automatic face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 1927-1943, 2007.

[50] D. Huang, G. Zhang, M. Ardabilian, Y. Wang, and L. Chen, "3D face recognition using distinctiveness enhanced facial representations and local feature hybrid matching," in *Proc. IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2010.

[51] Y. Wang, G. Pan, Z. Wu, "3D face recognition in the presence of expression: guidance-based constraint deformation approach," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[52] K. I. Chang, K. W. Bowyer, P. J. Flynn, "Adaptive rigid multi-region selection for handling expression variation in 3D face recognition," *IEEE workshop on FRGC Experiments*, 2005.

[53] G. Passalis, I. Kakadiaris, T. Theoharis, G. Tederici, and N. Murtaza, "Evaluation of 3D face recognition in the presence of facial expressions: an annotated deformable model approach," *IEEE workshop on FRGC Experiments*, 2005.

[54] M. Husken, M. Brauckmann, S. Gehlen, and C. v. d. Malsburg, "Strategies and benefits of fusion of 2D and 3D face recognition," *IEEE workshop on FRGC Experiments*, 2005.

[55] J. Cook, V. Chandran, and C. Fookes, "3D face recognition using Log-Gabor templates," in *Proc. British Machine Vision Conference*, 2006.

[56] T. Faltemier, K. W. Bowyer and P. Flynn, "A region ensemble for 3D face recognition," *IEEE Trans. on Information Forensics and Security*, vol. 3, no. 1, pp. 62-73, 2008.

[57] P. Szeptycki, M. Ardabilian, L. Chen, "A coarse-to-fine curvature analysis-based rotation invariant 3D face landmarking," in *Proc. IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2009.

[58] H. Drira, B. B. Amor, M. Daoudi, and A. Srivastava, "Pose and expression-invariant 3D face recognition using elastic radial curves," in *Proc. British Machine Vision Conference*, 2010.

[59] X. Li, T. Jia, and H. Zhang, "Expression-insensitive 3d face recognition using sparse representation," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[60] A. B. Moreno, A. Sanchez, J. F. Velez, and F. J. Díaz, "Face recognition using 3d local geometrical features: pca vs.svm," in *Proc. International Symposium on Image and Signal Processing and Analysis*, 2005.

[61] M. H. Mahoor and M. Abdel-Mottaleb, "Face recognition based on 3d ridge images obtained from range data," *Pattern Recognition*, vol. 42, no. 3, pp. 445–451, 2009.

[62] S. Berretti, A. Del Bimbo, and P. Pala, "3d face recognition by modeling the arrangement of concave and convex regions," in *Proc. International Workshop on Adaptive Multimedia Retrieval: Adaptive Multimedia Retrieval: User, Context, and Feedback*, pp. 108–118, 2006.

[63] M. H. Mousavi, K. Faez, and A. Asgharim, "Three dimensional face recognition using svm classifier," in *Proc. International Conference on Computer and Information Science*, pp. 208–213, 2008.

[64] D. Huang, M. Ardabilian, Y. Wang, and L. Chen, "A novel geometric facial representation based on multi-scale extended local binary patterns," in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, 2011.

[65] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Re*views, vol. 41, no. 4, pp. 1-17, 2011.

[66] C. C. Queirolo, L. Silva, O. R. P. Bellon, and M. P. Segundo, "3D face recognition using simulated annealing and the surface interpenetration measure," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 2, pp. 206–219, 2010.

[67] N. Alyuz, B. Gokberk, and L. Akarun, "Regional registration for expression resistant 3-D face recognition," *IEEE Trans. on Information Forensics and Security*, vol. 5, no. 3, pp. 425-440, 2010.

[68] A. Savran, N. Alyüz, H. Dibeklioğlu, O. Çeliktutan, B. Gökberk, B. Sankur, and L. Akarun, "Bosphorus Database for 3D Face Analysis," *The First COST 2101 Workshop on Biometrics and Identity Management*, 2008.

[69] C. Maes, T. Fabry, J. Keustermans, D. Smeets, P. Suetens and D. Vandermeulen, "Feature detection on 3d face surfaces for pose normalisation and recognition," in *Proc. IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2010.

Di Huang (S'10-M'11) received the B.S. and M.S. degrees in computer science from Beihang University, Beijing, China, and the Ph.D. degree in computer science from Ecole Centrale de Lyon, Lyon, France, in 2005, 2008, and 2011 respectively. In December 2011, he joined the Laboratory of Intelligent Recognition and Image Processing, Beijing Key Laboratory of Digital Media, School of Computer Science and Engineering, Beihang University, as a Faculty Member.

His current research interests include biometrics, especially on 2D/3D face analysis, image/video processing, and pattern recognition.

Mohsen Ardabilian received the master and the Ph.D. degree in computer science at Université Technologie de Compiègne, France, in 1996 and 2001, respectively. In 2001, he founded Avisias Company, specialized in media asset management with several other confirmed industrial players of Thomson, Canal+ Technologies and Philips, where he served as Scientific Expert from 2001 to 2003. In 2003, he joined Ecole Centrale de Lyon as an associate professor.

His current research interests include computer vision and multimedia analysis, particularly 3D acquisition and modeling, 3D face analysis and recognition.

Yunhong Wang (M'98) received the B.S. degree in electronic engineering from Northwestern Polytechnical University, Xi'an, China, in 1989, the M.S. degree and the Ph.D. degree in electronic engineering from the Nanjing University of Science and Technology, Nanjing, China, in 1995 and 1998, respectively. She worked at the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China, from 1998 to 2004. Since 2004, she has been a Professor with the School of Computer Science and Engineering, Beihang University, Beijing, China, where she is also the Director of Laboratory of Intelligent Recognition and Image Processing, Beijing Key Laboratory of Digital Media.

Her research interests include biometrics, pattern recognition, computer vision, data fusion, and image processing.

Liming Chen (M'05) received the B.Sc. degree in mathematics and computer science from Université de Nantes, Nantes, France, in 1984 and the M.S. and Ph.D. degrees, both in computer science from the University of Paris 6, Paris, France, in 1986 and 1989, respectively.

He was an Associate Professor with the Université de Technologies de Compiègne, Compiègne, France, and then joined Ecole Centrale de Lyon, Lyon, France, in 1998 as a Professor, where he has been leading an advanced research team on multimedia computing and pattern recognition. From 2001 to 2003, he was the Chief Scientific Officer with the Paris-based company Avivias, where he specializes in media asset management. During 2005, he was a Scientific Expert in multimedia at France Telecom R&D, China. Since 2007, he has been Chairman with the Department of Mathematics and Computer Science, Laboratoire d'InfoRmatique en Image et Systèmes d'information, Ecole Centrale de Lyon. Since 1995, he has been the author of three patents and more than 100 publications in international journals and conferences. He has directed more than 15 Ph.D. dissertations. His research interests include face analysis and recognition in 3-D and 2-D, image and video categorization, and affect computing on multimedia data.

Dr. Chen has been the Chairman and a Program Committee Member for various international conferences and journals. He has been a Reviewer for many conferences and journals, e.g. the *IEEE SIGNAL PROCESSING LETTERS*, *Computer Vision and Image Understanding*, the *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS*, the Proceedings of the *International Conference on Image Processing*, the *IEEE TRANSACTIONS ON IMAGE PROCESSING*, and *Pattern Recognition Letters*. He was a Guest Editor for the special issue on Automatic Audio Classification of the *European Association for Signal Processing (EURASIP) Journal on Audio, Speech, and Music Processing*.