Neurocomputing 275 (2018) 1295–1307

Contents lists available at ScienceDirect

Neurocomputing

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

Expression-insensitive 3D face recognition by the fusion of multiple subject-specific curves

Ye Li*, YingHui Wang*, Jing Liu, Wen Hao

Institute of Computer Science & Engineering, Xi'an University of Technology, No.5 South Jinhua Road, Xi'an 710048, China

ARTICLE INFO

Article history:
Received 21 September 2016
Revised 20 July 2017
Accepted 22 September 2017
Available online 4 October 2017

Communicated by Yue Gao

Keywords:
3D face recognition
Expression-insensitive
Subject-specific curve
Fusion
Feature-level

ABSTRACT

This study proposes a 3D face recognition method using multiple subject-specific curves insensitive to intra-subject distortions caused by expression variations. Considering that most sharp variances in facial convex regions are closely related to the bone structure, the convex crest curves are first extracted as the most vital subject-specific facial curves based on the principal curvature extrema in convex local surfaces. Then, the central profile curve and the horizontal contour curve passing through the nose tip are detected by using the precise localization of the nose tip and symmetry plane. Based on their discriminative power and robustness to expression changes, the three types of curves are fused with appropriate weights at the feature-level and used for matching 3D faces with the iterative closest point algorithm. The combination of multiple expression-insensitive curves is complementary and provides sufficient and stable facial surface features for face recognition. In addition, for each convex crest curve, an expression-irrelevant factor is assigned as the adaptive weight to improve the face matching performance. The results of experiments using two public 3D databases, GavabDB and BU-3DFE, demonstrate the effectiveness of the proposed method, and its recognition rates on both databases reflect an encouraging performance.

1. Introduction

Automated face recognition is one of the most promising areas of computer vision with many applications. Because of its non-intrusive nature compared to other biometric technologies, face recognition has received increasing attention over the past two decades. Although considerable work has been done in 2D face recognition, which delivers satisfactory results under constrained conditions, the recognition accuracy in uncontrolled environments can be significantly hampered by variations in illumination and pose. Because human bodies are in essence three-dimensional, the absence of depth information causes significant degradation of the representation and discriminative capability of the features extracted from 2D images [1]. Fortunately, due to the rapid advances in graphics hardware, computing techniques, and 3D technologies, the applications of 3D models have spread widely and effective 3D object retrieval technologies have become increasingly accessible [2,3]. Facial data obtained from 3D models can directly provide ample geometric information that is invariant to external changes, which cannot be obtained from a single 2D image. The prevention of face recognition caused by illumination and pose variance can be ameliorated by using 3D face representations. Hence, it is commonly believed that 3D face recognition has the potential to provide greater accuracy than 2D face recognition [4]. However, as human faces are non-rigid, expression variations bring about significant internal changes in the geometric shapes of the facial surface. Moreover, the faces of different subjects are similar to each other in terms of the primary structure. This inter-subject similarity causes 3D face recognition to be sensitive to variations in expressions. Thus, the distortions caused by facial expressions have become the main challenge of 3D face recognition [5]. To be useful and perform effectively in real-world applications, a 3D face recognition approach should handle expression variations effectively.

Existing 3D face recognition algorithms can be categorized into two classes: holistic and local feature-based algorithms [5,6]. The holistic feature-based algorithms use the entire facial surface as an input to perform face recognition, while the local feature-based algorithms extract a set of representative local features (e.g., landmarks, curves, regions) from the parts of the facial surface least affected by changes in expressions and then use these local features to compute the similarity between facial representations. Compared to the holistic feature-based algorithms, the local feature-based algorithms are commonly more robust to facial expressions [7,8]. Therefore, there has been a growing interest in using local feature-based algorithms for addressing the expression variation problem in 3D face recognition in recent years [9]. For

* Corresponding authors.
E-mail addresses: liye@xaut.edu.cn (Y. Li), wyh@xaut.edu.cn (Y. Wang).

https://doi.org/10.1016/j.neucom.2017.09.070
0925-2312/

the local feature-based algorithms, a critical requirement is to select relatively stable structures and patches on the facial surface and utilize them to obtain those features that are highly discriminative and invariant to different expressions. Most of the popular approaches have proposed to extract local features from the several segmented rigid facial regions. However, only a few regions on the entire facial surface remain rigid with varying expressions. Because the geometric variations on the facial surface caused by facial expressions are complex, the distortions corresponding to the same expression vary from one person to another, and the distortion caused by different expressions on the 3D facial surfaces belonging to the same subject is also diverse. Thus, the rigid facial regions are specific to each subject and to each particular expression. In addition, the non-rigid facial regions also bear significant discriminative information and should not be ignored [10]. For example, the region around the mouth, which is the portion of the face exhibiting the maximum distortion with changing expressions for almost all person and almost all expressions, may remain relatively stable for some people under some expressions and can thus offer distinctive information about different subjects. The approaches, which extract local features only from the relatively rigid regions, do not use all available facial features but exclude those affected by expressions. This leads to a loss of information that could be discriminative for face recognition. The segmentation of a 3D facial surface into relatively rigid and non-rigid regions is problematic and is not universally applicable to all subjects and all expressions.

To overcome these limitations, this study proposes to address the aforementioned problem of local feature-based 3D face recognition algorithms involving expression variations and presents a novel expression-insensitive 3D face recognition method based on the fusion of multiple subject-specific curves. The reasons that curves, instead of regions or landmarks, were chosen as the facial local features, are manifold. Firstly, as the result of sparse sampling from the facial surface, if selected properly, curves have good potential for representing the salient features of regional shapes of 3D face models (for instance, the central profile has been proved to work well [11,12]). Secondly, the algorithms based on curves, which encode more geometric information, are more discriminative and stable compared to the algorithms based on landmarks. Compared to the algorithms directly using local regions for facial representation, representing the facial surface in terms of curves reduces the amount of face data, and thus makes the procedure of face matching easier and more time-efficient. To achieve good performance under expression variations, the most critical issue in this study was the selection of the optimal set of 3D curves and the best way to match them. Because a single type of curve has a limited discriminative ability, three types of curves were selected from the entire face representation to provide the local feature descriptors. Closely related to the bone structure, these curves are specific to each subject and are insensitive to various changes in expressions. In order to achieve better performance, the three types of curves were fused at the feature-level. The proposed method was evaluated on two public databases, GavabDB and BU-3DFE. In the experiment on the GavabDB, for the entire dataset containing both neutral and expressive 3D faces, the proposed approach achieved a rank-1 recognition rate of 95.5%. The experiment on the expressive faces with different types and intensity levels in the BU-3DFE database also demonstrated the feasibility and effectiveness of the proposed approach and produced a rank-1 recognition rate of 95.25%.

The main contributions of this study can be summarized as follows:

(1) Introduction of a new local feature-based descriptor to handle expression variations in 3D face recognition: It consists of three types of subject-specific curves with distinctive shape information, which are invariant to expression. These are the convex crest curves, the central profile, and the horizontal contour passing through the nose tip. These subject-specific curves are distributed more equably and cover almost the entire facial surface.
(2) Design of a feature-level fusion scheme for multiple facial curves: Based on the discriminative power of each type of subject-specific curve, grouping weighted fusion was adopted on multiple curves at the feature-level to form a single compact representation. The feature-level fusion of these curves affords sufficient and stable features to achieve accurate 3D face recognition under expression variations.
(3) Improvement of the feature extraction mode of the traditional local feature-based face recognition algorithms: Instead of fixedly segmenting a 3D face into a few relatively rigid and non-rigid regions before the procedure of feature extraction to reduce the impact of expression changes, the proposed method can flexibly extract the expression-insensitive, subject-specific curves from the entire facial surface.
(4) The convex crest curves are selected as the most vital subject-specific facial curves based on the principle curvature extrema in convex local surfaces. They are closely related to the facial bone structure of each individual. Therefore, these curves can tolerate the distortion caused by expressions while making differences between individuals discernible. Furthermore, to achieve a good performance in case of severe expressions, each convex crest curve is given an adaptive weight in terms of its stability under expression variations in the face-matching phase.

The remainder of this study is organized as follows. In Section 2, the related work is introduced. Then, the proposed method is described in detail in Section 3. Section 4 shows the experiments conducted using two public 3D face databases with expressions. Finally, the conclusions are provided in Section 5.

2. Related work

The work in [5] provided a comprehensive survey of the 3D face recognition methods under expressions. The studies referred to are mainly restricted to the local feature-based face recognition algorithms, which are closely related to this work.

Compared to global features, local features extracted from the regional facial surface are more stable and less affected by expression changes. Many researchers have focused on expression-invariant 3D face recognition using local features. Local feature-based algorithms can further be divided into landmark-based, curve-based, and region-based algorithms [13]. The majority of expression-invariant approaches employ the relatively rigid regions to calculate the similarity between 3D faces. The most widely used region-based 3D expression-invariant face recognition method is to select well-defined rigid regions based on prior knowledge. Chang et al. [14] proposed a method for 3D face recognition under varying facial expressions by matching multiple regions around the nose. They selected the multiple overlapping rigid regions around the nose, which are less affected by the expression. Then, they matched each one of these regions individually using the iterative closest point (ICP) algorithm [15] and combined the results from all matches to achieve greater accuracy. Lei et al. [8] presented a 3D face recognition approach based on low-level geometric features collected from the eyes-forehead and nose regions. A region-based descriptor is formed on the basis of four types of histograms, which are obtained by computing the spatial relationships between the vertices on the 3D face. A support vector machine (SVM) is used to perform classification at both a feature-level and a

score-level fusion scheme. The experiment results demonstrate that the performance of feature-level fusion is better than that of score-level fusion. Mian et al. [16] used a region-based matching algorithm to recognize non-neutral 3D faces. They segmented each 3D face into three regions: eyes-forehead, nose and cheeks. Then, face recognition was performed using a modified version of the ICP algorithm based only on the eyes-forehead and nose regions in order to minimize the effects of expressions. Finally, the matching results obtained from the two regions were fused at the score-level. Ming et al. [17] presented a regional bounding spherical descriptor (RBSR) to perform 3D face recognition and emotion analysis. Firstly, each 3D face was segmented into a group of regions based on shape index and spherical bands. Then, the corresponding facial patches were projected to regional bounding spheres to obtain the RBSR descriptor. Finally, a regional and global regression mapping technique was employed to the weighted regional descriptor for enhancing the recognition accuracy. Li et al. [10] introduced a weighted multi-scale and multi-component local normal pattern descriptor to represent 3D faces. Three components of normal vectors were first estimated to produce three normal component images. Then, each normal component image was divided into several patches. Their local normal patterns were encoded on different scales and concatenated according to the facial configuration for 3D face recognition.

Considering that the segmentation of a 3D facial surface into relatively rigid and non-rigid regions based on a priori knowledge is inaccurate and problematic, some methods selected the more rigid regions or features dynamically. In these studies, the points with a lower registration error were considered to belong to the more rigid part of the facial surface. Wang et al. [18] developed a partial-ICP method for 3D face recognition. Based on the extent of distortion under different expressions, the method can extract the rigid regions of the facial surface by selecting a part of the nearer point pairs during registration of facial surfaces. The matching between the faces of the probe and the gallery is performed based on only the same relatively rigid parts. Chua et al. [19] presented a 3D point signature-based face recognition algorithm. After registration, the point's distances from two facial surfaces are modeled as a Gaussian distribution, and an adaptive threshold is computed to extract the rigid region. Li et al. [26] proposed a robust 3D face recognition approach based on sparse representation using low-level geometric features. A uniform remeshing scheme is used to establish a consistent sampling pattern across 3D faces. In addition, a feature pooling and ranking scheme is applied to choose expression-insensitive features.

As the overall 3D shape of the face is deformed by expression variations, there are almost no regions on the entire facial surface that can remain absolutely rigid. To overcome the problems occurring due to the sensitivity of the region-based algorithms, some researchers focused on the landmark-based algorithms, which compare 3D faces by detecting a set of landmarks and calculating the relations between them. Mian et al. [20] proposed a landmark-based method to handle expression variations in 3D face recognition. In this method, a novel 3D spherical face representation is used in conjunction with the Scale-Invariant Feature Transform (SIFT) descriptor to form a rejection classifier. The results show that the eyes-forehead and nose regions of a face contain the maximum number of discriminative features critical for robust face recognition. Li et al. [21] presented a SIFT-like algorithm for registration-free 3D face recognition under expressions. First, two principal curvatures were used for 3D keypoint detection. Then, three keypoint descriptors corresponding to three surface differential quantities were fused at the feature-level to describe the local shapes of all keypoints. 3D face recognition was finally achieved using a sparse representation-based fine-grained matching algorithm. Gupta et al. [22] proposed a face recognition algorithm based on anthropometric facial fiducial points. By identifying 10 facial fiducial points that are associated with the discriminative anthropometric features, they developed a completely automatic face recognition algorithm that employs facial 3D Euclidean, evaluates geodesic distances between these 10 facial fiducial points and generates a linear discriminant classifier. Guo et al. [7] presented a landmark-based expression-invariant 3D face recognition method. A set of class-specific keypoints are detected on each 3D face and the local surface around each keypoint is represented with a Rotational Projection Statistics (RoPS) feature. Face recognition is achieved using both RoPS feature matching and face registration. Yu et al. [23] presented a 3D directional vertices ($3D^2V$) approach for efficient 3D face matching. They used a few sparsely distributed structured vertices that carry structural information transferred from their deleted neighboring points to represent 3D faces. The results of experiments demonstrate that their approach can significantly reduce the data storage requirement and computation time.

Exploiting the fact that curves can be parameterized canonically and can thus be compared naturally, some 3D expression-invariant curve-based face recognition approaches have been proposed. Mahoor et al. [24] presented an approach for 3D face recognition using frontal range data based on the ridge lines on the face surface. A face range image is represented as a 3D binary image by using the principal curvature to extract the points lying on ridge lines. The robust Hausdorff distance and ICP algorithms are utilized for matching two face images. Drira et al. [25] proposed a local representation for facial surfaces using radial curves emanating from the nose tips. In order to handle the variations in facial expressions, they extended the elastic shape analysis of curves to 3D facial surfaces and developed a Riemannian framework for face matching. Similar to [25], Samad et al. [27] introduced a Frenet frame-based mathematical framework, which also represents 3D faces by extracting the radial curves originating from the nose tip. Their framework can reduce about 30,000 3D points per 3D face, bringing down the effective feature space required to less than 200 feature points. Ter Haar et al. [12] proposed a 3D face-matching framework that employs curve-based face recognition. Using this framework, they evaluated profile and contour types including those described in the literature and selected subsets of facial curves for effective and efficient face matching. Mpiperis et al. [29] developed an expression-invariant approach using geodesic polar parameterization for 3D face recognition. Each 3D face is represented by a set of geodesic circles based on geodesic polar coordinates. Face matching is performed using surface attributes defined on the geodesic plane. In their study, they considered that the limitations on recognition performance are imposed mainly by the isometric approximation of facial distortions and a face with an open mouth can violate the isometry assumption. Lee et al. [30] designed a method for matching 3D faces by using deformed circular curves, the shortest geodesic distances, and the properties of the Fourier Transform. By measuring the geodesic distance between the reference point (e.g., nose tip) and a point on the curve and by using Fourier Transform, they achieved efficient face matching by reducing the 3D face data to a low dimension vector.

To enhance the ability to represent 3D faces under expressions, a combination of different local features or a hybrid schema of local feature-based and other techniques was employed. Emambakhsh et al. [31] conducted a thorough investigation of the potential of the nasal region for expression-invariant 3D face recognition and introduced a novel surface normal-based recognition algorithm using a set of spherical patches and curves over the nasal region. The algorithm achieves the highest reported recognition ranks based on the nasal region as recorded in the FRGC, Bosphorus, and BU-3DFE databases. Gao et al. [32] proposed a local shape descriptor to explore expression-invariant, discriminative features on both depth and normal maps. After

preprocessing, the landmarks on the nasal surface and the adjoining cheek regions were detected. Then, inspired by Local Binary Patterns (LBPs), local shape differences for 3D points on a set of horizontal curves joining selected landmarks provided a novel representation of the local shape information. Huang et al. [33] introduced an effective method for 3D face recognition using a novel geometric facial representation along with a local feature hybrid-matching scheme. A set of facial depth maps is extracted by utilizing multi-scale extended LBPs to form the facial surface descriptors. A hybrid matching process, which combines a local matching step using the SIFT-based features with a global one under the facial component and configuration constraints, is carried out to enhance performance. Smeets et al. [34] proposed an expression-invariant 3D face recognition method in which an isometric distortion modeling approach and a region-based approach are combined. The isometric distortion modeling approach is based on a spectral decomposition of the geodesic distance matrix. The region-based approach uses the ICP algorithm to match two faces based only on the spherical rigid region around the nose. The results of the two approaches are fused at the score-level and rank-level.

Thanks to the advances in neural networks, deep learning has been proven to outperform traditional methods in various visual recognition tasks. A recent research [35–39] shows that the performance of 2D face recognition has significantly improved with the use of deep learning. The successful application of these deep learning approaches has been possible largely due to the use of an extensive labeled training dataset, created for face representation learning, that is invariant to different factors such as expressions or poses. Compared to 2D face recognition, training discriminative features for 3D face recognition is challenging because of the long time of training networks and the lack of large-scale 3D face datasets. To improve the efficiency of training, Zhang et al. [40] selected 30 feature points, which mainly lie in non-flat areas and on the border of the skull, from the 113 feature points of a Candide-3 face model to characterize an original 3D face. A stacked denoising auto-encoder algorithm based on deep learning was proposed to conduct the unsupervised preliminary training of face depth data and for supervised fine-tuning. The experiment results show that when the training sample set is small (below 5000 scans per subject), the robustness of the algorithm is not satisfactory. Lee et al. [41] proposed a face recognition system from RGB-D images based on deep learning. To achieve more accurate recognition performance, an efficient pipeline was designed to include 3D face shapes into the deep network for learning effective color and depth feature transformation. In consideration of the limited size of available RGB-D data for learning, their deep network was first trained on color face images and was later fine-tuned for face depth images to help machine transfer the knowledge of color images to depth images. Kim et al. [42] presented a 3D face recognition algorithm using a deep convolutional neural network (DCNN) and a 3D augmentation technique. To cope with the limited amount of available 3D data, they utilized the existing network used for 2D face recognition and fine-tuned it for 3D face recognition. To improve the performance of 3D face recognition under expression variations, they augmented their 3D face database by generating a number of synthesized 3D faces with expression changes from a single raw 3D scan.

3. Expression-insensitive 3D face recognition method based on the fusion of multiple curves

3.1. Overview of the proposed approach

In this study, a novel 3D face recognition approach was developed to handle informative expression variations by identifying three types of subject-specific and expression-insensitive curves and fusing them at the feature-level. Unlike most region-based or landmark-based techniques, the proposed curve-based method exploits the entire face for the extraction of discriminative facial features. Considering that the most convex regions (e.g., the nose and forehead regions) on the entire facial surface are closely related to the bone structure and are relatively insensitive to various expression changes, we focus mainly on the features that originate in these regions. In these regions, the points at which the local surface bends sharply are subject-specific and very important for the facial shape analysis of different subjects. These key points were detected on the entire facial surface based on rotation invariant curvature information to form a set of subject-specific features. To reduce the disturbance of noise, curve features were adopted instead of point features. These curves, named convex crest curves, are bound to the facial skeletons. For the sake of the completeness and sufficiency of features, these convex crest curves were combined with two other types of curves, which were selected based on a priori knowledge. The central profile and the horizontal contour passing through the nose tip were selected because they both contain distinctive shape information about the face and are more rigid under expression variance. These multiple curves, which are distributed more equably and basically cover the entire facial surface, were fused into a single compact representation at the feature-level to enhance the discriminative power of the system. To improve performance, each type of curve is assigned an appropriate weight based on its capacity to differentiate, and each convex crest curve is assigned an adaptive weight in terms of its stability under expression changes during the face-matching phase. By selecting a definite proportion of the nearer point pairs on the curves around the mouth during matching two faces with the ICP algorithm, the useful information is implicitly extracted while the effect of distortion caused by expressions is effectively reduced. To the best of the authors' knowledge, this is the first study that uses the feature-level fusion of these types of subject-specific curves to handle expression variations in 3D face recognition.

Fig. 1 shows the three primary procedural steps of the proposed method. In the first procedure, the convex crest curves, central profile, and the horizontal contour passing through the nose tip are extracted. In the second procedure, these multiple curves are fused into a compact representation at the feature-level with a grouping weighted strategy. In the third procedure, the ICP algorithm is used to perform face matching and obtain the recognition result. The specific steps are as follows: After preprocessing, the ridge points are detected from all local convex surfaces on the entire face according to the high-order curvatures, and the convex crest curves are obtained by connecting the corresponding ridge points. Then, the more stable and salient of all the convex crest curves are selected based on their properties of prominence and anti-interference. Next, the nose tip is localized based on a combination of low-order and high-order curvature information, and the nose bridge is detected from all the distinct convex crest curves based on its position and linear features. Then, the asymmetry plane is determined based on the positions of the nose tip and the bridge, and the central profile and the horizontal contour through the nose tip of the 3D face can be obtained. Each type of curve, according to its representational capacity and invariance to facial expressions, is given an appropriate weight, and all three types of curves are fused at the feature-level. Finally, face matching is achieved by using the ICP algorithm.

3.2. Subject-specific curves extraction

3.2.1. Convex crest curves detection

The major characteristic of the facial surface is its noteworthy convex and concave variances. The positions and degrees of

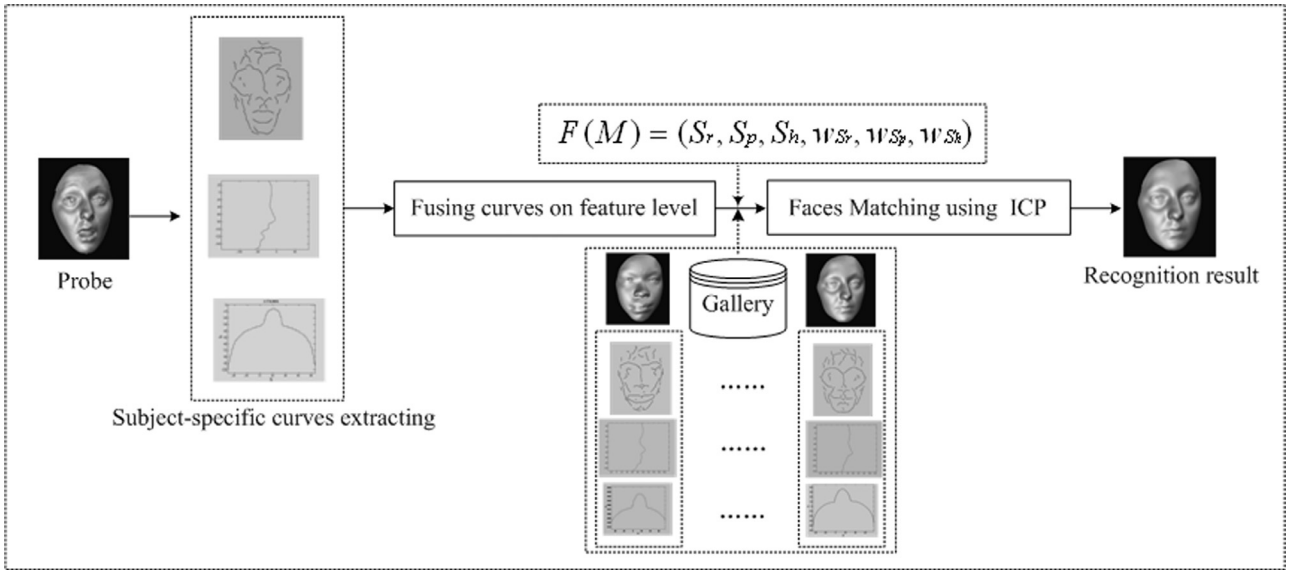


Fig. 1. Schema of the expression-insensitive 3D face recognition method based on fusion of multiple curves.

these variances are specific to different persons and can be used as key information in recognition. In addition, most sharp variances in convex regions are closely related to the facial bone structure. Thus, they have a high discriminative ability and are insensitive to various expressions. On a given smooth generic surface M, a non-umbilic point p is called a ridge point if its maximal principal curvature attains a local positive maximum along the associated principal direction. As the computation of curvatures is sensitive to noise on the surface, in order to improve the stability of the features, convex crest curves formed by connecting the salient ridge points are used instead of directly using the ridge points. The convex crest curves can be characterized by

$$e_{\max} = 0, \quad \partial e_{\max}/\partial t_{\max} < 0, \quad k_{\max} > |k_{\min}| \tag{1}$$

where $k_{\max}$ and $k_{\min}$ are the maximal and minimal principal curvatures, $\boldsymbol{t}_{\max} = (t_{max}{}^1, t_{max}{}^2, t_{max}{}^3)$ is the principal direction corresponding to the maximal principal curvature and $e_{\max}$ is the derivative of the maximal principal curvature along its corresponding principal direction. In Thirion's paper [43], $e_{\max}$ is called an extremality coefficient. For a surface in implicit form $F(x, y, z) = 0$, the extremality coefficient $e_{\max}$ is given by

$$e_{\max} = \frac{\partial k_{\max}}{\partial t_{\max}} \tag{2}$$

$$= \frac{F_{xyz} \cdot t^1_{\max} \cdot t^2_{\max} \cdot t^3_{\max} + 3k_{\max} \cdot Fxy \cdot t^1_{\max} \cdot n^2}{|\nabla F|} \tag{2}$$

where $\boldsymbol{N} = (n^1, n^2, n^3)$ is the normal vector of surface M, and F_{xy} and F_{xyz} are the second and third partial derivatives of $F(x, y, z)$, respectively.

To obtain the subject-specific convex curves that represent 3D faces, the salient ridge points in convex regions are first detected from the entire facial surface. In the proposed method, the detection of ridge points and convex crest curves on a 3D face is accomplished through the following steps: First, for each mesh vertex v with normal vector $\boldsymbol{N}_v$, suppose there are m vertices adjacent to v and let v_i denote the ith adjacent vertex. $\boldsymbol{N}_v$ can be computed from the weighted facet normals around vertex v. Then, a new local coordinate is chosen such that the origin of the coordinate is situated on v, v becomes (0, 0, 0), and $\boldsymbol{N}_v$ lies along the positive z-axis. Each adjacent vertex v_i is transformed to the local coordinates and the adjacent-normal cubic approximation method [44] is used to fit the local surface of each mesh vertex v_i and its k-ring ($k = 1, 2, 3 \ldots$) neighborhood. In these local coordinates, the local cubic surface in the Monge form is:

$$f(x,y) = \frac{1}{2}\left(b0x^2 + 2b1xy + b2y^2\right) + \frac{1}{6}\left(c0x^3 + 3c1x^2y + 3c2xy^2 + c3y^3\right) \tag{3}$$

Next, the principal curvatures are computed by

$$K = \frac{f_{xx}f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)} \tag{4}$$

$$H = \frac{f_{xx} + f_{yy} + f_{xx}fy^2 + f_{yy}f_x^2 - 2f_xf_yf_{xy}}{2(1 + f_x^2 + f_y^2)^{3/2}} \tag{5}$$

$$k_{\max} = H + \sqrt{(H^2 - K)} \tag{6}$$

$$k_{\min} = H - \sqrt{(H^2 - K)} \tag{7}$$

where K and H are the Gaussian curvature and mean curvature at vertex v, and f_x and f_y are the first order and f_{xx}, f_{xy}, and f_{yy} are the second order partial derivatives of $f(x, y)$ at vertex v, respectively. Therefore, the curvature extremality coefficients of each mesh vertex v can be estimated. Finally, the method of Ohtake [45] is used to detect ridge points and convex curves on the entire facial mesh model.

The full set of extracted convex curves in small local surfaces frequently contains many insignificant curves, which are caused by or are sensitive to expression variance. To reduce the interference of these convex curves, an appropriate larger k-ring neighborhood is selected for fitting the local surface during the ridge point detection step and a parameter, called ridge strength (RS), is used to select the convex curves that are more noteworthy and have stronger convex features. In this experiment, the value of k is experimentally determined as 6 and is effective for different 3D face models. RS is obtained as

$$RS = \sum_{i=1}^{k-1} \delta i \frac{k_{\max}(pi) + k_{\max}(p_{i+1})}{2} \|p_i - p_{i+1}\| \tag{8}$$

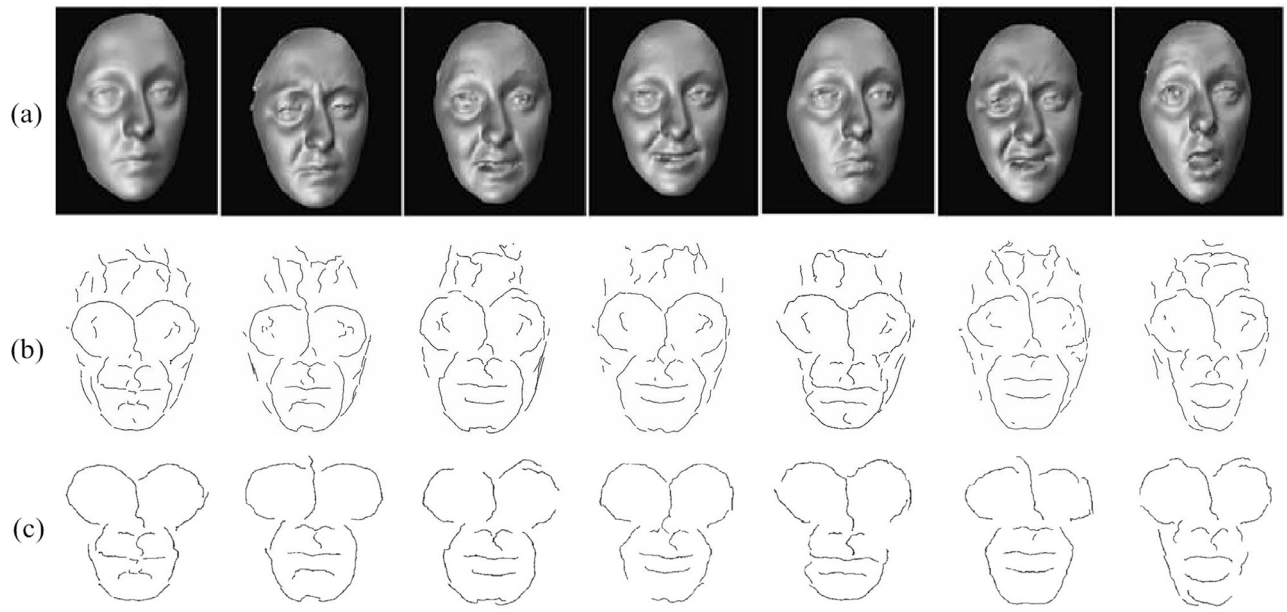


Fig. 2. Convex crest curves on the faces of the same subject under different expressions: (a) 3D faces, (b) convex crest curves, (c) the convex crest curves of which the number of ridge points is greater than 30 and RS is larger than 15.

$$\delta i = \frac{1}{1 + \left| \frac{k_{\max}(pi) + k_{\max}(pi+1)}{2} - \frac{1}{n}\sum_{i=1}^{n} k_{\max}(pi) \right|} \tag{9}$$

where pi is a ridge point on a convex crest curve and n is the number of ridge points on this crest curve, $k_{\max}(pi)$is the maximal principal curvature at the ridge pointpi, and δi is a stability factor at the ridge pointpi. The role of δi is to decrease the disturbance of noise and increase the RS of the stable convex curves. At the ridge point p_i on the mesh edge $[v_1,v_2]$, k_{max} is estimated as

$$k_{\max}(pi) = \frac{|e_{\max}(v_2)|k_{\max}(v_1) + |e_{\max}(v_1)|k_{\max}(v_2)}{|e_{\max}(v_1)| + |e_{\max}(v_2)|} \tag{10}$$

Since the convex crest curves, particularly those that are prominent, are closely related to facial bone structure and can represent the major facial shape information, they are insensitive to expression changes and have a high discriminative ability. The detection results of the extracted convex crest curves on the 3D faces of the same subject under different expressions and on the 3D faces of different subjects under the same expression are shown in Fig. 2 and Fig. 3, respectively. The results of these experiments show that the larger the number of ridge points and the greater the RS of the convex crest curves, the better their stability and retrieval performance. Therefore, an expression-irrelevant factor is used as the adaptive weight for each convex crest curve to increase the prominence of those curves that are more rigid under expressions in face matching. The expression irrelevant factor on the kth convex crest curve $E\lambda k$ is calculated as

$$E\lambda k = \frac{nk \cdot RS_k}{\frac{1}{m}\sum_{i=1}^{m} ni \frac{1}{m}\sum_{i=1}^{m} RS_i} \quad (k = 1, 2, \ldots, m) \tag{11}$$

where m is the number of the convex crest curves, n_k is the number of the ridge points on the kth convex crest curve, and RS_k is the ridge strength of the kth convex crest curve.

3.2.2. Central profile and horizontal contour curves detection

In order to enhance the completeness and sufficiency of the features, the central profile and the horizontal contour curves are extracted and combined with the convex crest curves for subject-specific face representation. It has been found that the central profile curve contains different information of the face that is not present in the frontal view and has the most distinctive geometrical properties among different subjects [11,28]. As the shapes of the nose ala and cheekbones are important for face recognition and are more stable under expression variances, the horizontal contour curve through the nose tip is also selected as the key feature of a 3D face. Because the recognition rate depends on the precision of the extracted curves, it is important to extract them accurately. In the proposed method, the two curves mentioned above are detected based mainly on the localization of the nose tip and the symmetry plane.

Since the region surrounding the nose tip has a higher Gaussian curvature and dome-like shape (approximately spherical), the nose tip is located on the convex curves based on the Gaussian curvature distribution characteristic of a spherical surface. A more detailed description of the nose tip detection process was presented in our previous study [46]. The localization of the symmetry plane is based on the detection of the nose bridge. As the nose bridge constitutes the intersection of the two plane-like nose wings and runs horizontally across the top part of the nose, it is detected from the convex curves based on its position relative to the nose tip and its feature of linear regression. According to its position relative to the nose tip and assuming that some data for the nose bridge may be missing because of self-occlusion, the nose bridge candidates are detected in a spherical neighborhood of the nose tip, wherein the radius of the sphere is experimentally set to 15 mm. The distinct convex curves are selected as the nose bridge candidates if a portion of the curves falls within the spherical neighborhood. Based on its feature of linear regression, the least square method is used to fit each nose bridge candidate with a straight line L, which passes through the nose tip $P_t(x_t, y_t, z_t)$. Since the ridge points at a distance from the nose tip have limited effect on fitting and may sometimes cause loss of precision, the ridge points that are at a distance greater than 30 mm from the nose tip are removed from each nose bridge candidate. The equation of the line L in standard form is

$$\frac{x - xt}{t_1} = \frac{y - y_t}{t_2} = \frac{z - z_t}{t_3} \tag{12}$$

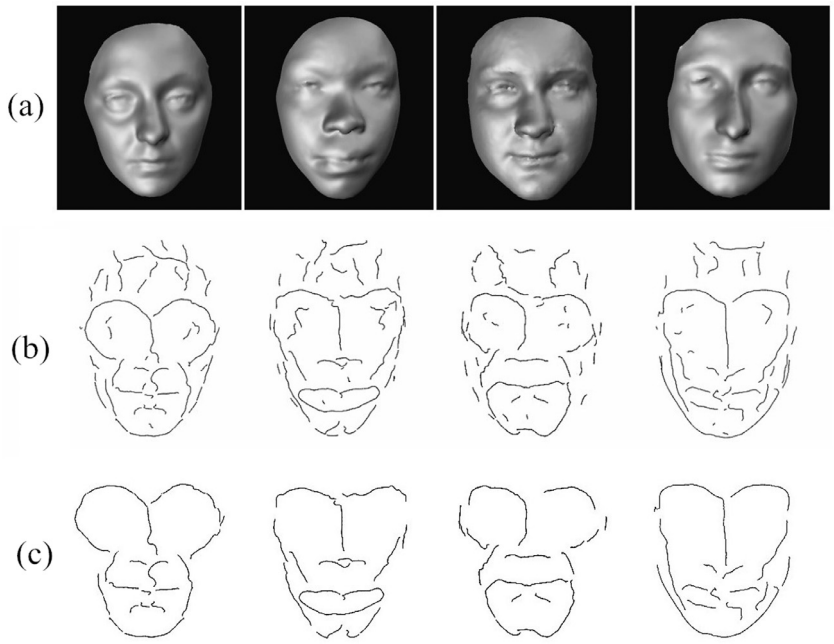


Fig. 3. Convex crest curves on the faces of different subjects under the same expression: (a) 3D faces, (b) convex crest curves, (c) the convex crest curves of which the number of ridge points is greater than 30 and RS is larger than 15.

where t_1, t_2, and t_3 are the three parameters of L to be determined. A parameter, Line Strength (LS), is used to evaluate the approximation degree between a nose bridge candidate and L. LS is calculated as

$$LS = \frac{1}{\sqrt{\frac{1}{m}\sum_{i=1}^{m}[x_{pi} - (az_{pi} + b)]^2 + \frac{1}{m}\sum_{i=1}^{m}[y_{pi} - (cz_{pi} + d)]^2}}$$
$$a = \frac{t_1}{t_3}, b = xt - \frac{t_1}{t_3}z_t, c = \frac{t_2}{t_3}, d = y_t - \frac{t_2}{t_3}z_t \quad (13)$$

where m is the number of ridge points on a nose bridge candidate and p_i is one of the ridge points. The nose bridge candidate with the largest LS is selected as the true nose bridge. As the nose bridge and the normal vectors of the ridge points on it all lie on the symmetry plane, the normal vector of the symmetry plane $\overrightarrow{\mathbf{n}S}$ can be calculated as

$$\overrightarrow{\mathbf{n}S} = \overrightarrow{\mathbf{n}}_L \times \mathrm{mean}(\overrightarrow{\mathbf{n_{pi}}}) \quad (i = 1, \ldots, M_p) \quad (14)$$

where $\overrightarrow{\mathbf{n}L}$ is the normal vector of L, p_i is the ith ridge point on L, $\overrightarrow{\mathbf{npi}}$ is the normal vector of p_i, M_p is the number of ridge points on L and mean $\overrightarrow{\mathbf{n}_p i}$ is the mean of the normal vectors of all ridge points on L. Based on the normal vector of the symmetry plane $\overrightarrow{\mathbf{n}S} = (ns1, ns2, ns3)$ and the position of nose tip P_t (x_t, y_t, z_t), the symmetry plane can be described as

$$n_{s1}(x - x_t) + n_{s2}(y - y_t) + n_{s3}(z - z_t) = 0 \quad (15)$$

By cutting a 3D face model with its symmetry plane, the central profile curve can be obtained.

To obtain the horizontal contour, first, the nasal base point and nasal root point need to be determined on the central profile. To locate the nasal root point p_r, the intersection points of L and the central profile are obtained. Then, the point that has the greatest distance from the nose tip in the set of all intersection points is selected as the nasal root point. Since the nasal base point p_b is the point which lies on the central profile beneath the nose tip with a conspicuously sharp turn, it is sought along the central profile, starting at the nose tip p_t and moving away from the nasal root p_r. The first inflection point detected on the central profile curve is determined as the nasal base point. Then according to the normal vector determined by $p_b(x_b, y_b, z_b)$ and $p_r(x_r, y_r, z_r)$, the plane that is perpendicular to the symmetry plane and passes through the nose tip can be *described as:*

$$(x_b - x_r)(x - x_t) + (y_b - y_r)(y - y_t) + (z_b - z_r)(z - z_t) = 0 \quad (16)$$

By cutting the 3D facial model with the plane, the horizontal contour curve is obtained.

3.3. Fusion of multiple curves and face matching

To achieve effective face retrieval, it is essential to use features obtained by the fusion of multiple curves. Multiple information fusion, which is commonly used in pattern classification, can be performed at three different levels: the data, feature, and decision level. Feature-level fusion can fuse multiple features extracted from different sources directly and effectively, without data redundancy. Since it combines the advantages of fusion on the data-level and decision-level, the fusion at the feature-level usually can achieve a better performance [47]. In our method, the three types of curves are fused at the feature-level and are assigned different weights in terms of their discriminative capacity. Through the grouping weighted fusion of the three types of curves at the feature-level, a 3D face representation with a sextuplet can be expressed as:

$$F(M) = (S_r, S_p, S_h, wS_r, wS_p, wS_h) \quad (17)$$

where M is a 3D facial model, S_r is the set of the ridge points on the convex crest curves, S_p is the set of points on the central profile, S_h is the set of points on the horizontal contour through the nose tip, and wSr, wSp, and wSh are the weights corresponding to the three types of curves mentioned above. The points of S_p and S_h are obtained by sampling along the central profile and the horizontal contour at an interval of 1 mm, respectively.

Based on a single type of feature curve, Euclidean distance d(A,B) is used to measure the dissimilarity between two feature sets A and B, corresponding to a probe and a gallery facial model, respectively. d(A,B) is calculated by

$$d(A, B) = \sum_{k=1}^{m} \left\| p_A^k - p_B^k \right\| \quad (18)$$

where p_A^k is the kth point in set A and p_B^k is the point in set B corresponding to p_A^k, m is the number of feature points in set A. Based on the fusion of multiple feature curves, the dissimilarity $D(Mi, Mj)$ between probe facial model M_i and gallery facial model M_j is measured by the sum of all weighted Euclidean distances corresponding to the three types of feature curve.

$$D(Mi, Mj) = wsr \cdot d(S_r^i, S_r^j) + wsp \cdot d(S_p^i, S_p^j) + wsh \cdot d(S_h^i, S_h^j) \tag{19}$$

$$d(S_r^i, S_r^j) = \sum_{k=1}^{nr} E\lambda k \cdot (S_r^i(k), S_r^j) \tag{20}$$

where $S_r^i = \{S_r^i(1), S_r^i(2), \ldots, S_r^i(nr)\}$ is the feature set of the convex crest curves corresponding to facial model M_i, $S_r^i(k)$ is the set of ridge points on the kth convex crest curve on the probe facial model M_i; S_p^i and S_h^i are the feature sets of the central profile and the horizontal contour through the nose tip, respectively; n_r is the number of the convex crest curve on M_i; and $E\lambda k$ is the adaptive weight assigned to the kth convex crest curves on the probe facial model M_i. The use of the adaptive curve-specific weight $E\lambda k$ can enhance the effect of the convex crest curves providing high distinguishing capacity and stability (see Section 3.2.1).

The ICP algorithm can iteratively align two 3D models by the minimization of the Euclidean distance between the closest points and the ICP error can be used as a dissimilarity measure to select the best matching face for recognition. In the proposed method, facial model matching is achieved based on the multiple feature curve fusion using an improved ICP algorithm [48] after initial registration based on the nose tip position. For each point p_A^k in one type of curve feature set $A = \{p_A^1, p_A^2, p_A^3, \ldots, p_A^m\}$ of the probe facial model M_i, the corresponding point p_B^k, which is the point that is the smallest distance from p_A^k, is searched in the same type of feature curve set $B = \{p_B^1, p_B^2, p_B^3, \ldots\ldots\}$ of the gallery facial model M_j. During the matching process using the improved ICP algorithm, one-to-one correspondence of set A and set B between each point pair is not required, and overlap is allowed. This can reduce the applicable conditions and the error accumulation of the face matching to some extent. Due to severe distortion in the mouth region on expressive faces, 40% (determined according to the best result in the experiments on two databases) of the points were eliminated in S_r and S_p, which are under the nasal base point and at a large distance from their corresponding points in the face-matching phase. Because the initial registration according to the position of the nose tip is performed before face matching, the number of iterations of the ICP algorithm is reduced. Furthermore, for a certain type of curve feature, the searching of point pairs is conducted only within the scope of the same type of feature. The complexity of the proposed algorithm for matching a probe face scan with all face scans in the gallery is $O(Ng \times Nf_p \times Nf_g)$, where Ng is the number of scans in the gallery, Nf_p is the number of feature points on the three types of curves of the probe scan, and Nf_g is the average number of feature points on three types of curves of all gallery scans. By representing a 3D face with fewer features, the number of which is approximately 7%−11% of the original mesh points on the 3D face, the complexity of the proposed face-matching algorithm is lower.

4. Experiment results

In order to evaluate the feasibility and effectiveness of the proposed approach, validation experiments were conducted using two 3D public face databases, GavabDB [28] and BU-3DFE [50].

4.1. Experiments on the GavabDB database

GavabDB is an expression-rich and noise-prone 3D face database. It contains 549 3D facial surfaces. These facial surfaces correspond to 61 different subjects (45 males and 16 females) and each subject has 9 facial models. Each model is presented by a mesh of connected 3D points of the facial surface without texture. The database provides systematic variations of the pose and facial expression of each subject. In particular, the nine models corresponding to each subject include two frontal and four rotated scans with neutral expressions and three frontal scans in which the subject presents a laugh, smile, and a random expression chosen by the user, respectively. The random expressions are frequently accentuated facial expressions, such as bulging of the cheeks, pouting, sticking out tongue, and making faces. All 3D face models were captured by a Minolta VI-700 digitizer, a range laser sensor based on the active triangulation principle. Samples of the facial models with expressions in GavabDB are shown in Fig. 4. In our experiments, all of the five frontal facial scans were selected for each subject. One neutral scan of each subject was used to construct the gallery set (a total of 61 scans), and the remaining frontal scans (a total of 244 scans, one neutral and three expressive scans for each subject) were used as probe faces.

The detection results of the three types of curves for the facial models of different subjects under the same expression are shown in Fig. 5. Although under the same expression, the same types of curves corresponding to the different subjects are distinctly different. The inter-subject differences of the horizontal contour curves passing through the nose tip are reflected mainly in the lateral shape of the nose and cheek, while the inter-subject differences of the central profiles are embodied in the degree and position of the convex-concave in the vertical direction from the forehead to the chin. For convex crest curves, the inter-subject differences are more significant and are shown in the spatial distribution of the local convex surfaces bending sharply on the entire face. Among the three types of curves, the convex crest curves are quite subject-specific.

To further demonstrate the discriminative power of each type of curve and to achieve the best performance, the proposed method was tested using several feature fusion strategies with different weights of the three types of curves during the face matching procedure. The results are shown in Table 1, where S_r means the convex crest curves, S_p means the central profile, S_h means the horizontal contour through the nose tip, and *wSr*, *wSp*, and *wSh* are the weights assigned to the three types of curves. The results show that the discriminative power of the convex crest curves is highest among the three types of curves, and the fusion of the three types of curves can improve the performance effectively. The best recognition rate of 94% is achieved on the test set of the GavabDB following the "expressive vs. neutral" protocol. In addition, the average computation time of the proposed approach is 5.51 s, running on a PC with Intel Core i5 and 2.60 GHz, 4 GB RAM.

The rank-1 recognition rate of the proposed approach obtained with optimum fusion weights was compared with the state-of-the-art approaches on the GavabDB following the same protocol. As shown in Table 2, the proposed approach achieves the second-best performance with an overall rank-1 recognition rate of 95.5%, which outperforms the majority of other approaches and is close to the highest published results on the GavabDB. The best performance was achieved by Drira et al. [25], with an overall rank-1 recognition rate of 95.9%. A same performance as the proposed method was achieved by Huang et al. [33]. Note that, in the preprocessing step, Drira et al. [25] aligned the probe face to the gallery face rigidly using the ICP algorithm, and Huang et al. [33] utilized a coarse alignment based on three landmarks (i.e. the two inner corners of the eyes and the nose tip) to rotate and

Fig. 4. Samples of the facial models with expression in GavabDB.

Table 1
Recognition rates of the proposed method on the GavabDB (Rank-1).

	S_r	S_p	S_h	$S_r + S_p$	$S_r + S_h$	$S_r + S_p + S_h$
wSr	1	0	0	0.60	0.70	0.55
wSp	0	1	0	0.40	0	0.30
wSh	0	0	1	0	0.30	0.15
Neutral (61 scans)	96.7%	88.5%	65.6%	100%	98.4%	100%
Expressive (183 scans)	82.5%	69.9%	46.5%	86.9%	85.2%	94%
Neutral + expressive (244 scans)	86.1%	74.6%	48.8%	90.2%	88.5%	95.5%

Table 2
Recognition rate comparison of the proposed method with other methods on the GavabDB (Rank-1).

Approaches	Neutral (%)	Expressive (%)	Neutral + expressive (%)
Mahoor et al. [24]	95	72	78
Li et al. [26]	96.7	93.3	94.7
Huang et al. [33]	100	94	95.5
Drira et al. [25]	100	94.5	95.9
Tabia et al. [49]	100	93.3	94.9
Yu et al. [23]	95.6	79.8	87.7
Our approach	100	94	95.5

translate the probe face, while the face alignment process is not required in the proposed approach. Since the time-consuming preprocessing in Drira et al. [25], the reported computation time of their approach is 7.45 s, and 6.18 s is used for preprocessing. Thus, the proposed approach achieves a very close performance with Drira et al. [25] at a 26% reduction of computation time.

4.2. Experiments on the BU-3DFE database

In order to illustrate the effectiveness of the proposed method for a larger 3D face database with richer expression types, the BU-3DFE database was selected for our experiments. The BU-3DFE database was built at Binghamton University. It is an important dataset for face recognition under expression variations. It contains 100 subjects (56% female, 44% male), ranging in age from 18 years to 70 years, with a variety of ethnicities or races, including Caucasian, African, East Asian, Middle-East Asian, Indian, and Latino. Each subject has 25 scans with texture, including six universal expressions of anger, fear, disgust, happiness, sadness, and surprise at four different intensities and one neutral expression. All the 3D facial models were captured by a 3DMD digitizer and were preprocessed appropriately. Each model was saved as a polygonal mesh with a resolution ranging from 20,000 to 35,000 polygons. An example of 3D facial models with six basic expressions in four intensities of the same subject in BU-3DFE is shown in Fig. 6. In the experiments, the gallery set is composed of one neutral scan per subject, the remaining 24 gesture scans per subject, a total of 2400 scans, constituted the probe set. All the scans were used without texture. The whole probe set, named set-whole, was divided into six subsets. The six subsets are named according to the corresponding expression type, set-an (angry), set-fe (fear), set-ha (happy), set-sa (sad), set-di (disgust), and set-su (surprise), respectively. The detection results of the three types of curves on 3D faces of the same subject with six different basic expressions in low intensity are shown in Fig. 7. The detection results of the three types of curves on 3D faces of the same subject with the same expression (exemplified as happy) in four different intensities are shown in Fig. 8.

As shown in Figs. 7 and 8, the three types of curves extracted in the proposed method are closely related to the expression-independent structure and shape of facial bones (such as brow bone, nasal bone, and cheekbones), and the distribution of all the curves takes into account both directionality and spreadability. Therefore, these curves are insensitive to the variations in expression type and intensity. They can be used to effectively represent the vital and stable characteristics of 3D faces and have the ability to distinguish intra-class from inter-class changes. Since the noise is less, the convex crest curves extracted from the faces in BU-3DFE are clearer and more specific. The recognition results on the six test subsets of different expression types are shown in Table 3. The results indicate that the proposed method achieves promising recognition results for all six types of expressive faces in both the low-intensity test sets and the entire four levels test sets. Although the recognition rates for all expression types

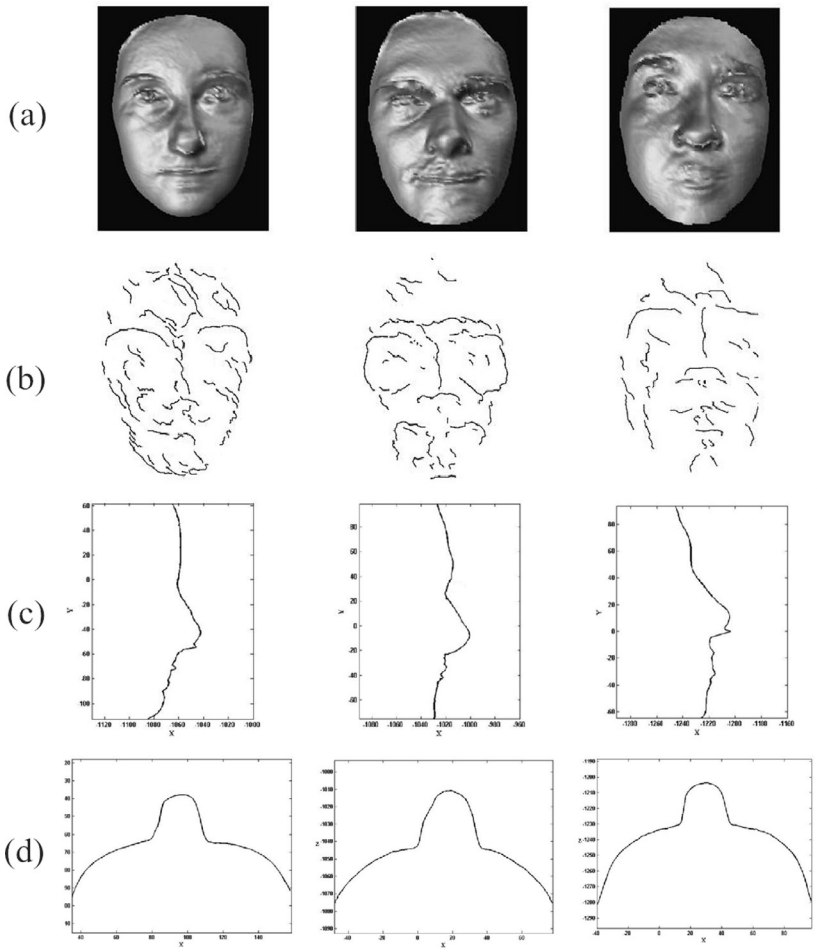


Fig. 5. Three types of feature curve on the facial models of different subjects under the same expression in GavabDB: (a) Facial models, (b) convex crest curves, (c) central profiles, (d) horizontal contours.

Fig. 6. 3D expressive facial models (total 24, with texture) of a sample subject in BU-3DFE.

Table 3
Recognition rates on six 3D expression-face subsets of BU-3DFE (Rank-1).

	set-an	set-fe	set-ha	set-sa	set-di	set-su	set-whole
Low intensity (level $\leq$ 2)	97.5	98	99	98.5	98.5	97.5	98.17
All four levels	93.25	92.75	98.0	97.25	96.5	93.75	95.25

decline when the expression intensity increases, the extent of the decline is modest. The decline is sharper for the expressions of anger, fear, and surprise, a drop of approximately 4%, which can be explained by the fact that the distortions around the brow bone and cheekbones caused by these expressions are more severe than those caused by other expression types. As compared to the results on the GavabDB, the recognition rate of the proposed method on the BU-3DFE remains approximately constant, while the number of subjects increases from 61 to 100 and the number of probe scans increases from 244 to 2400. Despite the fact that the faces in the BU-3DFE are less noisy and are well preprocessed, it is verified that our method can be applied effectively to larger 3D face databases with rich expression types. The subject-specific curves adopted in this method retain sufficient inter-subject facial differences while reducing the interference caused by expression changes.

The results obtained using the proposed approach were compared with those of other approaches that have been published and validated on the BU-3DFE following the same "expressive vs. neutral" protocol. The comparative results of the recognition rate in rank-1 (R1RR) and the equal error rate (EER) are shown in Table 4. The table shows that the proposed approach achieves the best face recognition performance on all 2400 expressive scans in the

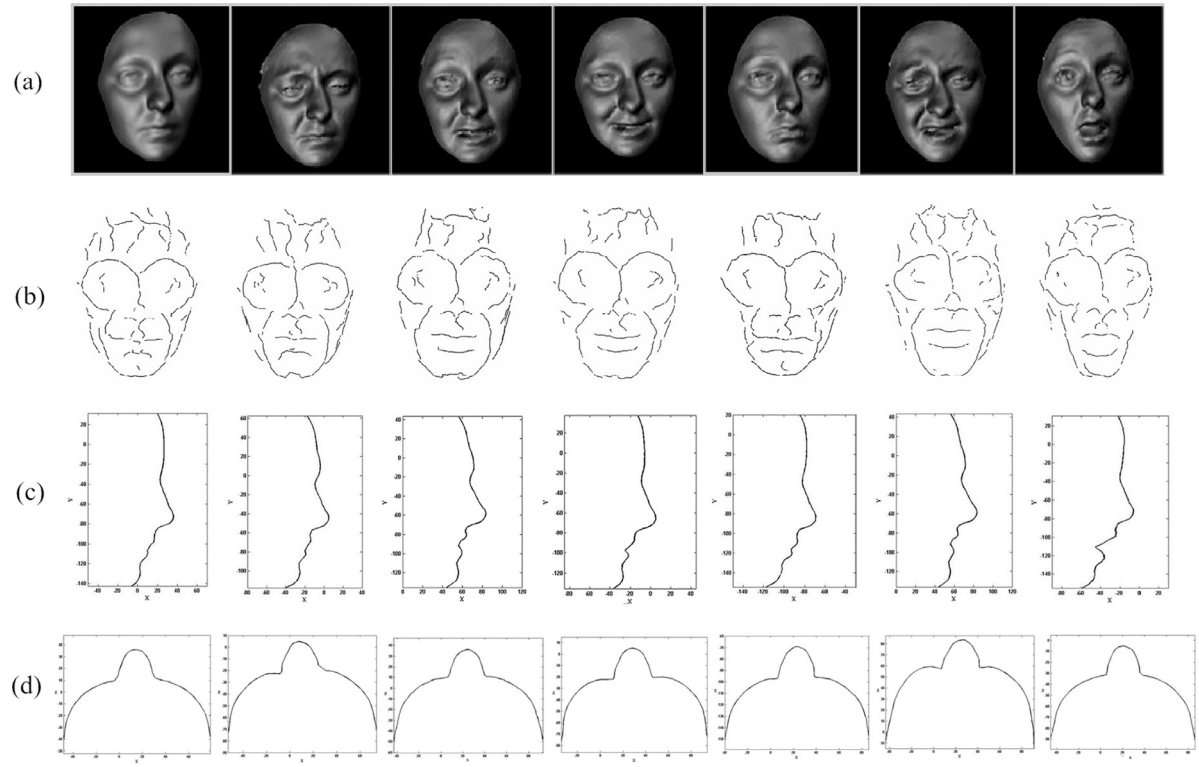


Fig. 7. Three types of curve extractions on the facial models of the same subject in BU-3DFE under six basic expressions (a) Facial models, (b) convex crest curves, (c) central profiles, (d) horizontal contours.

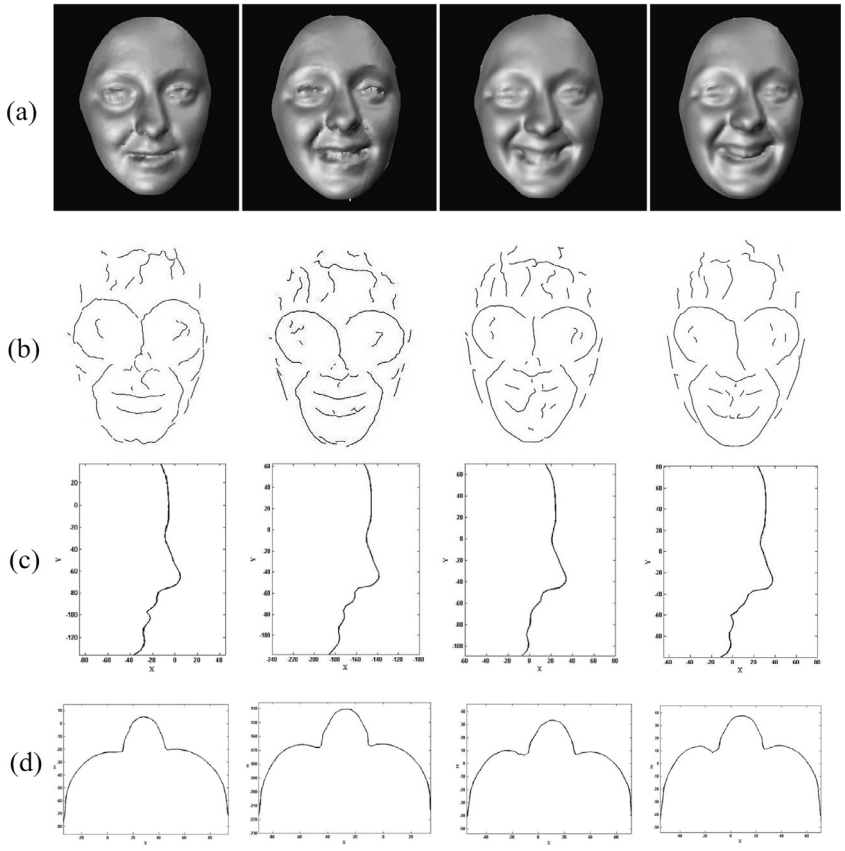


Fig. 8. Three types of curve extractions on the facial models of the same subject in BU-3DFE under the same expression for four different intensities: (a) Facial models, (b) convex crest curves, (c) central profiles, (d) horizontal contours.

Table 4
Comparison with other approaches on 3D expression-face sets of BU-3DFE.

Expressive(Probe) vs. Neutral(Gallery)	No. of subjects	No. of scans	R1RR(%)	EER(%)
Mpiperis et al. [29](without texture)	100	1500	84.4	12
Mpiperis et al. [29](with texture)	100	1500	80.3	9.8
Smeets *et al*[34].	100	900	94.5	5.8
Han et al. [51]	100	2400	91.1	–
Li et al. [10]	100	2400	92.2	–
Lei et al. [6]	100	2400	94	–
Emambakhsh et al. [31]	100	2400	88.9	–
Our approach	100	2400	95.25	5.3

BU-3DFE, with a rank-1 recognition rate of 95.25% and an EER of 5.3%.

5. Conclusions

The problem of 3D face recognition under expression variations is addressed in this study and a novel expression-invariant 3D face recognition method based on the combination of multiple subject-specific curves is presented. In the proposed method, instead of fixedly dividing each 3D face into a few rigid and non-rigid regions to reduce the impact of expression changes, the subject-specific curves are flexibly extracted with more rigidity from the entire facial surface. The subject-specific curves are composed of convex crest curves, the central profile, and the horizontal contour passing through the nose tip. Grouping weighted fusion is adopted on the three types of curves at the feature-level and an adaptive weight is assigned to each convex crest curve according to its discriminative power and insensitivity to expressions. The feature-level fusion of these curves affords sufficient and stable distinctive shape information to achieve robust 3D face recognition under expression variations. Comprehensive experiments were conducted on the GavabDB and BU-3DFE databases, and an encouraging and competitive recognition performance was achieved. In future work, we plan to study the application of the proposed method for non-frontal and partially occluded 3D faces.

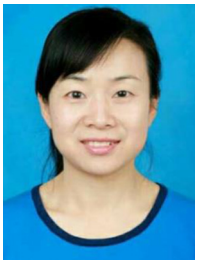
Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grants No. 61472319, No. 61602373, and in part by Scientific Research Project of Shaanxi Provincial Education Department under Grants no. 17JK0535, and in part by Shaanxi Natural Science Foundation under Grants no. 2015JZ015, no. 2017JQ6023, and in part by China Postdoctoral Science Foundation under No. 2014M552469, and in part by Dr. Start-up fund of Xi'an University of Technology.

References

[1] Z. Gao, H. Zhang, P. Xu G, et al., Multi-perspective and multi-modality joint representation and recognition model for 3D action recognition, Neurocomputing 151 (2) (2015) 554–564.
[2] A. Liu, W. Nie, Y. Gao, et al., Multi-modal clique-graph matching for view-based 3D model retrieval, IEEE Trans. Image Process. 25 (5) (2016) 2103–2116.
[3] Y. Gao, M. Wang, R. Ji, et al., 3-D Object retrieval with Hausdorff distance learning, IEEE Trans. Ind. Electron. 61 (4) (2014) 2088–2098.
[4] W. Bowyer K, K. Chang, P. Flynn, A survey of approaches to three-dimensional face recognition, in: Proceedings of the 17th International Conference on Pattern Recognition, 2004, pp. 358–361.
[5] D. Smeets, P. Claes, J. Hermans, et al., A comparative study of 3-D face recognition under expression variations, IEEE Trans. Syst. Man Cybern. Part C 42 (5) (2012) 710–727.
[6] Y. Lei, Y. Guo, M. Hayat, et al., A two-phase weighted collaborative representation for 3D partial face recognition with single sample, Pattern Recogn. 52 (3) (2016) 218–237.
[7] Y. Guo, Y. Lei, L. Liu, et al., EI3D: expression-invariant 3D face recognition based on feature and shape matching, Pattern Recogn. Lett. 83 (3) (2016) 403–412.
[8] Y. Lei, M. Bennamoun, A.A. El-Sallam, An efficient 3D face recognition approach based on the fusion of novel local low-level features, Pattern Recogn. 46 (1) (2013) 24–37.
[9] H. Patil, A. Kothari, K. Bhurchandi, 3-D face recognition: features, databases, algorithms and challenges, Artif. Intell. Rev. 44 (3) (2015) 1–49.
[10] H. Li, D. Huang, M. Morvan J, et al., Expression-robust 3D face recognition via weighted sparse representation of multi-scale and multi-component local normal patterns, Neurocomputing 133 (14) (2014) 179–193.
[11] T. Nagamine, T. Uemura, I. Masuda, 3D facial image analysis for human identification, in: Proceedings of International Conference on Pattern Recognition (ICPR), 1992, pp. 324–327.
[12] B. Ter Haar F, C. Veltkamp R, A 3D face matching framework for facial curves, Graph. Models 71 (2) (2009) 77–91.
[13] M. Bennamoun, F. Sohel, Y. Guo, Feature selection for 2D and 3D face recognition, in: G. John (Ed.), Encyclopedia of Electrical and Electronics Engineering Webste, Wiley, 2015, pp. 1–54.
[14] I. Chang K, W. Bowyer K, J. Flynn P, Multiple nose region matching for 3D face recognition under varying facial expression, IEEE Trans. Pattern Anal. Mach. Intell. 28 (10) (2006) 1695–1700.
[15] J. Besl P, D. Mckay N, Method for registration of 3-D shapes, IEEE Trans. Pattern Anal. Mach. Intell. 14 (2) (1992) 239–256.
[16] S. Mian A, M. Bennamoun, R.A. Owens, Region-based matching for robust 3D face recognition, in: Proceedings of the British Machine Vision Conference, 2005, pp. 207–217.
[17] Y. Ming, Robust regional bounding spherical descriptor for 3D face recognition and emotion analysis, Image Vis. Comput. 35 (2015) 14–22.
[18] M. Wang Y, G. Pan, H. Wu Z, et al., Exploring facial expression effects in 3D face recognition using partial ICP, Lect. Notes Comput. Sci. 3851 (2006) 581–590.
[19] S. Chua C, F. Han, K. Ho Y, 3D human face recognition using point signature, in: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, 2000, pp. 233–238.
[20] A. Mian, M. Bennamoun, Owens RAn efficient multimodal 2D-3D hybrid approach to automatic face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 29 (11) (2007) 1927–1943.
[21] H. Li, D. Huang, J.M. Morvan, et al., Towards 3D face recognition in the real: a registration-free approach using fine-grained matching of 3D keypoint descriptors, Int. J. Comput. Vis. 113 (2) (2015) 128–142.
[22] S. Gupta, M.K. Markey, A.C. Bovik, Anthropometric 3D face recognition, Int J Comput Vis 90 (3) (2010) 331–349.
[23] X. Yu, Y. Gao, J. Zhou, Sparse 3D directional vertices vs continuous 3D curves: efficient 3D surface matching and its application for single model face recognition, Pattern Recogn. 65 (1) (2017) 296–306.
[24] M.H. Mahoor, M. Abdel-Mottaleb, Face recognition based on 3D ridge images obtained from range data, Pattern Recogn. 42 (3) (2009) 445–451.
[25] H. Drira, B. Ben Amor, A. Srivastava, et al., 3D face recognition under expressions, occlusions, and pose variations, IEEE Trans. Pattern Anal. Mach. Intell. 35 (9) (2013) 2270–2283.
[26] X. Li, T. Jia, H. Zhang, Expression-insensitive 3D face recognition using sparse representation, in: Proceedings of IEEE Conference on Computer Vision & Pattern Recognition, 2009, pp. 2575–2582.
[27] M.D. Samad, K.MFrenet Iftekharuddin, Frame-based generalized space curve representation for pose-invariant classification and recognition of 3-D face, IEEE Trans. Human-Mach. Syst. 46 (4) (2016) 522–533.
[28] B. Moreno A, A. Sánchez, GavabDB: a 3D face database, in: Proceedings of International Conference on Workshop Biometrics, 2004, pp. 77–85.
[29] I. Mpiperis, S. Malassiotis, M.G. Strintzis, 3-D face recognition with the geodesic polar representation, IEEE Trans. Inf. Forensics Secur. 2 (3) (2007) 537–547.
[30] D. Lee, H. Krim, 3D face recognition in the Fourier domain using deformed circular curves, Multidimens. Syst. Signal Process. 28 (1) (2017) 105–127.
[31] M. Emambakhsh, A. Evans, Nasal patches and curves for an expression-robust 3D face recognition, IEEE Trans. Pattern Anal. Mach. Intell. (2017) 1–14, doi:10.1109/TPAMI.2016.2565473.
[32] J. Gao, A.N. Evans, Expression robust 3D face recognition by matching multi–component local shape descriptors on the nasal and adjoining cheek regions, in: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, 2015, pp. 1–8.

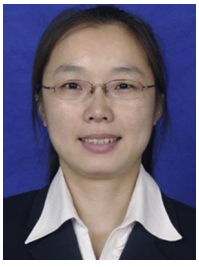
[33] D. Huang, M. Ardabilian, Y. Wang, et al., 3-D face recognition using eLBP-based facial description and local feature hybrid matching, IEEE Trans. Inf. Forensics Secur. 7 (5) (2012) 1551–1565.
[34] D. Smeets, T. Fabry, J. Hermans, et al., Fusion of an isometric deformation modeling approach using spectral decomposition and a region-based approach using ICP for expression-invariant 3D face recognition, in: Proceedings of IEEE International Conference on Pattern Recognition (ICPR), 2010, pp. 1172–1175.
[35] O.M. Parkhi, A. Vedaldi, A. Zisserman, Deep face recognition, in: Proceedings of the British Machine Vision Conference, 2015 pp. 41.1-41.12.
[36] Y. Taigman, M. Yang, M. Ranzato, et al., DeepFace: closing the gap to human-level performance in face verification, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1701–1708.
[37] Y. Sun, X. Wang, X. Tang, Deep learning face representation from predicting 10,000 classes, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1891–1898.
[38] Y. Sun, X. Wang, X. Tang, Deeply learned face representations are sparse, selective, and robust, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2892–2900.
[39] Y. Sun, D. Liang, X. Wang, et al., DeepID3: face recognition with very deep neural networks, Comput. Sci. (2015). arXiv:1502.00873[cs.CV].
[40] J. Zhang, Z. Hou, Z. Wu, et al., Research of 3D face recognition algorithm based on deep learning stacked denoising autoencoder theory, in: Proceedings of IEEE International Conference on Communication Software and Networks, 2016.
[41] Y. Lee, J. Chen, C. Tseng, S. Lai, Accurate and robust face recognition from RGB-D images with a deep learning approach, in: Proceedings of the British Machine Vision Conference, 2016.
[42] D. Kim, M. Hernandez, J. Choi, et al., Deep 3D face identification, Comput. Vis. Pattern Recogn. (2017) arXiv preprint arXiv:1703.10714.
[43] J.P. Thirion, The extremal mesh and the understanding of 3D surfaces, Int. J. Comput. Vis. 19 (2) (1996) 115–128.
[44] J. Goldfeather, V. Interrante, A novel cubic-order algorithm for approximating principal direction vectors, ACM Trans. Graph. 23 (1) (2004) 45–63.
[45] Y. Ohtake, A. Belyaev, H.P. Seidel, Ridge-valley lines on meshes via implicit surface fitting, ACM Trans. Graph. 23 (3) (2004) 609–612.
[46] Y. Li, H. Wang Y, B.B. Wang, et al., Nose tip detection on three-dimensional faces using pose-invariant differential surface features, IET Comput. Vis. 9 (1) (2015) 75–84.
[47] Y. Lei, M. Bennamoun, A.A. El-Sallam, An efficient 3D face recognition approach based on the fusion of novel local low-level features, Pattern Recogn. 46 (1) (2013) 24–37.
[48] S. Rusinkiewicz, Levoy MEfficient variants of the ICP algorithm, in: Proceedings of IEEE Third International Conference on 3-D Digital Imaging and Modeling, 2001, pp. 145–152.
[49] H. Tabia, H. Laga, D. Picard, et al., Covariance descriptors for 3D shape matching and retrieval, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 4185–4192.
[50] L. Yin, X. Wei, Y. Sun, et al., A 3D facial expression database for facial behavior research, in: Proceedings of IEEE 7th International Conference on Automatic Face and Gesture Recognition, 2006, pp. 211–216.
[51] X. Han, M.H. Yap, I. Palmer, Face recognition in the presence of expressions, J. Softw. Eng. Appl. 5 (5) (2012) 321–329.



Ye Li is an associate professor at the Institute of Computer Science and Engineering in Xi'an University of Technology. She received her M.S. and Ph.D. degrees from Xi'an University of Technology, Xi'an, China, in 2003 and 2016, respectively. Her research interests include computer vision, image processing, and pattern recognition.

Yinghui Wang received his Ph.D. degree from North-west University, Xi'an, China, in 2002. From 2003 to 2005, he was a post-doctoral fellow at Peking University, Beijing, China. Now he is a professor at the Institute of Computer Science and Engineering, Xi'an University of Technology, China. His research interests include image analysis and pattern recognition.

Jing Liu is an associate professor at the Institute of Computer Science and Engineering in Xi'an University of Technology. She received her M.S. degree in computer science from Xi'an University of Technology, china, in 2001. Her current research interests include digital watermarking, image processing, and 3D model representation.



Wen Hao received her MS degrees from the Department of Computer Science at Shann'xi Normal University, Xi'an, China, in 2011 and received her Ph.D. degree from Xi'an University of Technology, Xi'an, China, in 2015. She is currently working at the Institute of Computer Science and Engineering, Xi'an University of Technology, Xi'an, China. Her research interests include pattern recognition, point cloud processing.