A Group of Facial Normal Descriptors for Recognizing 3D Identical Twins

Huibin Li[1,2], Di Huang[3], Liming Chen[1,2], Yunhong Wang[3], Jean-Marie Morvan[1,4,5]
[1]Université de Lyon, CNRS, [2]Ecole Centrale Lyon, LIRIS UMR5205, F-69134, Lyon, France
[3]IRIP, School of Computer Science and Engineering, Beihang University, Beijing, 100191, China
[4]Université Lyon 1, ICJ, 43 blvd du 11 Novembre 1918, F-69622 Villeurbanne-Cedex, France
[5]KAUST, GMSV Research Center, Bldg 1, Thuwal 23955-6900, Saudi Arabia
huibin.li@ec-lyon.fr, *dhuang@buaa.edu.cn*, *liming.chen@ec-lyon.fr*,*yhwang@buaa.edu.cn*
morvan@math.univ-lyon1.fr

Abstract

In this paper, to characterize and distinguish identical twins, three popular texture descriptors: i.e. local binary patterns (LBPs), gabor filters (GFs) and local gabor binary patterns (LGBPs) are employed to encode the normal components (x, y and z) of the 3D facial surfaces of identical twins respectively. A group of facial normal descriptors are thus achieved, including Normal Local Binary Patterns descriptor (N-LBPs), Normal Gabor Filters descriptor (N-GFs) and Normal Local Gabor Binary Patterns descriptor (N-LGBPs). All these normal encoding based descriptors are further fed into sparse representation classifier (SRC) for identification. Experimental results on the 3D TEC database demonstrate that these proposed normal encoding based descriptors are very discriminative and efficient, achieving comparable performance to the best of state-of-the-art algorithms.

1. Introduction

As the performance of 2D still-image face recognition system in constrained environments continues to increase [6] and the high enough accuracies achieved by existing 3D face recognition algorithms against public datasets like FRGC v2.0, focus is shifting from methods that improve face recognition performance in general, to the ones that handle specic failure cases, caused by interference factors such as illumination conditions, pose, lapse of time, facial expression changes [12]. Recently, the new scenario of distinguishing identical twins becomes a challenging problem of traditional face recognition system since the quite strong similarity of their 2D facial appearance and 3D facial geometric shapes. Some recent studies have evidenced these challenges [17], [16], [12], [20], [19].

To the best of our knowledge, Sun *et al.* [17] were the first to evaluate the face recognition performance on a dataset consisting of multiple biometric traits (fingerprint, face, and iris) of identical twins. They tested face modality using the Cognitec FaceVACS system on 134 subjects (66 pairs of twins and two sets of triplets). They revealed that their face system could distinguish non-twins better than identical twins. Meanwhile, they concluded that the distribution of identical twin impostor was more similar to the genuine distribution than the general impostor distribution.

Phillips *et al.* [16] investigated the performance of three of the top submissions to the Multiple Biometric Evaluation (MBE) 2010 Still Face Track [6] on a dataset of twins acquired at Twins Day [1] in 2009 and 2010. Their experimental dataset is composed of images taken from 126 pairs of identical twins (252 people) collected on the same day and 24 pairs of identical twins (48 people) with images collected one year apart. They revealed that the results were largely dropped by the variations of lighting conditions (studio and outside); expressions (neutral and smiling); gender and age.

Klare *et al.* [12] proposed to classify facial features into three levels, i.e. appearance features (Level 1), local features (Level 2) as well as detailed features (Level 3), and studied the feature distinctiveness of Level 2 (MLBP and SIFT) and Level 3 (facial marks) with respect to distinguish identical twins. They indicate that these features which perform well in twin identification are not always consistent with the ones good at recognizing non-twin faces.

Later, Vijayan *et al.* [20] evaluated the performance of four state-of-the-art 3D face recognition algorithms against the largest dataset of 3D facial scans of twins, namely the 3D Twins Expression Challenge (3D TEC) dataset which contains 107 pairs of identical twins. They found that some algorithms perform very well on the FRGC v2.0 dataset but vastly degrades in performance on the 3D TEC dataset, especially distinguishing the cases combining factors related to facial similarity and the variation of facial expres-

978-1-4673-1228-8/12/$31.00 ©2012 IEEE

sion. Their results show that 3D face recognition of identical twins in the presence of varying facial expression is far from a solved problem, but good performance is possible [20]. In another work of Vijayan *et al.* [19], they also pointed out that distinguishing between identical twins with the variation of facial expression is a very challenging problem.

Recently, Li *et al.* [14] proposed a 3D face descriptor namely Multi-Scale Local Normal Patterns (MS-LNPs) which encodes the three normal components of facial surface using Multi-scale LBP operators. The recognition was carried out by the weighted sparse representation of the encoded facial normal descriptor (MS-LNPs) to enhance the robustness of their system to resist the variation of facial expression. The weights of each facial physical component were learned from training datasets. They achieved a rank-one recognition rate of 96.3% which is comparable to the state-of-the-art tasks. One of the important conclusions they draw in their work lies in that: the encoded normal information is much more discriminative than both the original normal information and the encoded geometric information (i.e., depth images or range images).

On the other side, as we know, except Local Binary Patterns (LBP) [2], Gabor Filters (GFs) [21], and Local Gabor Binary Patterns (LGBPs) [24] are also two popular face descriptors successfully used in the state-of-the-art 2D face biometric systems. Later, LBP [8] and GFs [3], [23] have also been exploited to describe 3D facial shape and applied to 3D face recognition systems. In contrast to what we do in this study, all these 3D face descriptors directly encode the geometric information of 3D facial surfaces.

Inspired by the two facts mentioned above, in this paper, we propose two new ways to encode normal information, including multi-scale and multi-orientation Gabor Filters (GFs) and Local Gabor Binary Patterns (LGBPs). Together with MS-LNPs [14], we achieve a group of facial normal descriptors consisting of Normal Local Binary Patterns descriptor (N-LBPs, i.e. MS-LNPs), Normal Gabor Filters descriptor (N-GFs) and Normal Local Gabor Binary Patterns descriptor (N-LGBPs). The effectiveness of the proposed descriptors are conducted on the 3D TEC dataset, and we find that all these descriptors are very discriminative to distinguish identical twins.

The rest of the paper is organized as follows. The framework overview of the proposed method is presented in section 2. Section 3 introduces the method of facial normal estimation. Facial normal encoding and representations are explained in section 4. Section 5 describes the sparse representation classifier. In section 6, we show the experimental and algorithmic settings as well as the results. Section 7 concludes the paper.

2. Framework Overview

As shown in Fig. 1, our framework consists of four steps: 1) normal estimation; 2) normal encoding; 3) comprehensive representation and 4) classification. Specifically, given a raw 3D facial scan, we first launch the preprocessing step to normalize the range image to an $m \times n \times 3$ matrix (i.e., x, y and z coordinates). Based on the range image, we estimate its three normal components (x, y, and z) by local plane fitting method (see Sec. 3). For each normal component, we propose to use three kinds of encoding and representation methods and thus we achieve a group of normal descriptors: N-LBPs, N-GFs and N-LGBPs. Since all the three descriptors comprehensively encode normal information either by multiple scales or orientations, we make use of two manners to fuse them for classification. Feature-level fusion which concatenates all the features extracted at different scales and orientations; and score-level fusion which creates one feature vector at one scale or orientation for the classifier to calculate an individual similarity score, and all these scores are further fused for the following step. Considering its powerful classification ability, in this study, we apply sparse representation classifier (SRC) for the classification. Due to the high dimension of the feature vector of N-GFs and N-LGBPs, we employ PCA for dimensionality reduction. The final similarity measurement of each descriptor computed by combining the ones of three normal components are used for decision making.

3. Facial Normal Estimation

Given a range image based face model represented by an $m \times n \times 3$ matrix as follows,

$$\mathbf{P} = [p_{ij}(x, y, z)]_{m\times n} = [p_{ijk}]_{m\times n\times\{x,y,z\}}, \quad (1)$$

where $p_{ij}(x, y, z) = (p_{ijx}, p_{ijy}, p_{ijz})^T, (1 \le i \le m, 1 \le j \le n, i, j \in \mathbb{Z})$ represents the 3D coordinates of the point p_{ij}. Let its unit normal vector matrix ($m \times n \times 3$) be

$$\mathbf{N}(\mathbf{P}) = [n(p_{ij}(x, y, z))]_{m\times n} = [n_{ijk}]_{m\times n\times\{x,y,z\}}, \quad (2)$$

where $n(p_{ij}(x, y, z)) = (n_{ijx}, n_{ijy}, n_{ijz})^T, (1 \le i \le m, 1 \le j \le n, i, j \in \mathbb{Z})$ denotes the unit normal vector of p_{ij}. We have $\|n(p_{ij}(x, y, z))\|_2 = 1$. As described in [7], the normal vector **N(P)** of range image **P** can be estimated by using local plane fitting method. That is to say, for each point $p_{ij} \in \mathbf{P}$, its normal vector $n(p_{ij})$ can be estimated as the normal vector of the following local fitted plane:

$$S_{ij} : n_{ijx}q_{ijx} + n_{ijy}q_{ijy} + n_{ijz}q_{ijz} = d, \quad (3)$$

where $(q_{ijx}, q_{ijy}, q_{ijz})^T$ represents any point within the local neighborhood (5×5 window in our paper) of point p_{ij} and $d = n_{ijx}p_{ijx} + n_{ijy}p_{ijy} + n_{ijz}p_{ijz}$.

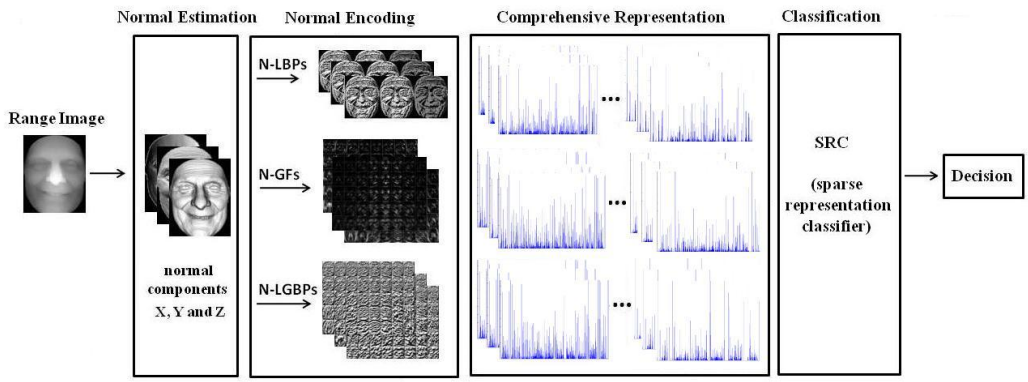

Figure 1. Framework of our proposed method.

Figure 2 shows one example of three normal component matrices (images) estimated from the original range image. we can see that normal components highlight the informative clues of the raw range image, e.g. the details in the forehead region, which corresponds to the conclusion in [14] . To simplify, each normal component in equation (2) can be represented by an $m \times n$ matrix:

$$\mathbf{N}(\mathbf{P}) = \begin{cases} \mathbf{N}(\mathbf{X}) = [n_{ij}^x]_{m\times n}, \\ \mathbf{N}(\mathbf{Y}) = [n_{ij}^y]_{m\times n}, \\ \mathbf{N}(\mathbf{Z}) = [n_{ij}^z]_{m\times n}. \end{cases} \quad (4)$$

where $-1 \leq n_{ij}^k \leq 1, k \in \{x, y, z\}$.

Figure 2. From left to right: the original range image, and its normal images of component x, y and z.

4. Facial Normal Encoding and Representation

Thanks to the matrix form of these normal components in equation (4), we can encode and characterize each of them using the similar way of feature extraction as for 2D texture images. In this paper, we propose to use LBP, Gabor and LGBP descriptors to encode the discriminative information from each of these normal components.

4.1. Normal Local Binary Patterns

Our Normal Local Binary Patterns (N-LBPs) is directly inherited from MS-LNPs proposed in [14], where three scales of LBP, i.e., $Q_{1,8}$, $Q_{2,16}$ and $Q_{3,24}$, ($Q_{n,m}$ denotes a neighborhood of m sampling points on a circle of radius of n) operator were used. Formally, given a point p_{ij}, its normal component noted as $n_{ijk}(0)$, the derived LNPs decimal value is:

$$LNPs(Q_{n,m}(p_{ij})) = \sum_{q=1}^{m-1} t(n_{ijk}(q) - n_{ijk}(0))2^q, \quad (5)$$

where $t(x) = 1$, if $x \geq 0$; $t(x) = 0$, if $x < 0$. $LNPs(Q_{n,m})$ encodes local normal variations of each normal component as decimal value, noted by $e(\{n_{ijk}\}_{m\times n}), k \in \{x, y, z\}$. See Fig. 3 for an example of $LNPs(Q_{1,8})$ representation of three normal components.

To describe a local shape region, the normal component is encoded as the histogram of LNPs as follows:

$$H = \sum_{i,j} I\{e(\{n_{ijk}\}_{m\times n}) = r\}, r = 0, \ldots, R-1, \quad (6)$$

where R is the encoded decimal number, for $Q_{1,8}, R = 2^8 = 256$. $I\{A\} = 1$, if A is ture, else $I\{A\} = 0$. This histogram contains the local micro-patterns of normal component over the whole face model. To utilize spatial information of facial shape, each facial normal component can be further divided into several patches, from which local normal patterns histograms H are extracted; then concatenated by facial configuration to form a global histogram G. Please see [14] for more details.

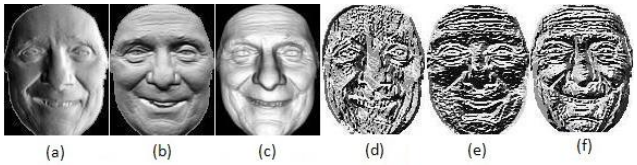

Figure 3. (a) to (c): normal images of component x, y and z; (d) to (e), the corresponding N-LBPs (LNPs) representation $Q_{1,8}$.

4.2. Normal Gabor Filters

Inspired by the motivation that the smooth properties of depth Gabor image cannot describe the facial features in detail [23], our proposed Normal Gabor Filters descriptor

(N-GFs) characterizes facial surfaces by Gabor Filters on its normal components. The mathematical expression of 2D Gabor wavelets can be defined as follows [13]:

$$\psi(z) = \frac{k_{\mu,v}^2}{\sigma^2}\exp(-\frac{k_{\mu,v}^2 z^2}{2\sigma^2})[\exp(ik_{\mu,v}z) - \exp(-\frac{\sigma^2}{2})], \quad (7)$$

where $z = (x, y)$, and μ and v define the orientation and scale of the Gabor wavelets. $k_{\mu,v}$ is define as follows:

$$k_{\mu,v} = k_v e^{i\Phi_\mu}, \quad (8)$$

where $k_v = k_{max}/f^v$ and $\Phi_\mu = \pi\mu/8$. k_{max} is the maximum frequency, and f is the spacing factor between Gabor filters in the frequency domain. In this study, we use Gabor filters with five different scales, $v \in \{0, \ldots, 4\}$, and eight orientations, $\mu \in \{0, \ldots, 7\}$, with the parameters $\sigma = 2\pi, k_{max} = \pi/2$ and $f = \sqrt{2}, asin[23]$.

By the convolution operation of Gabor Filters and normal components, we achieve the Gabor based representation of normal components:

$$O_{\mu,v}^k(x, y) = [n_{ij}^k(x, y)]_{m\times n} * \psi(x, y), k \in \{x, y, z\}, \quad (9)$$

where * denotes the convolution operator.

Given the normal component k, orientation μ and scale v, at each point (x, y), $O_{\mu,v}^k(x, y)$ is a complex number. We compute its magnitude. Thus, totally, for each normal component k, we can achieve 40 (5×8) Gabor magnitude images. See Fig. 4 for an example of normal component z. The comprehensive description of Normal Gabor Filters (N-GFs) can be obtained by feature level fusion or score-level fusion.

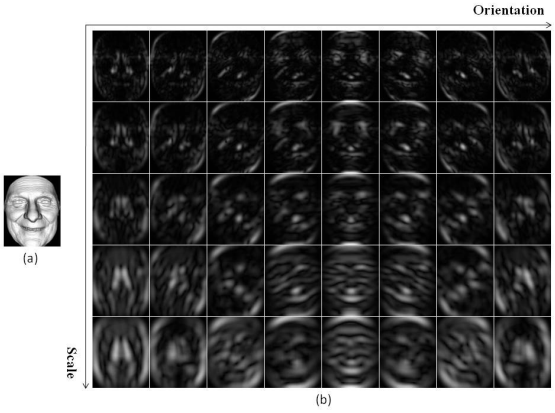

Figure 4. (a). Normal component z. (b) Its Gabor representation.

4.3. Normal Local Gabor Binary Patterns

It have been proved that the combination of Gabor filters and LBP (i.e. Local Gabor Binary Patterns, namely LGBP for short) can improve the performance of LBP for 2D face recognition [24]. Analogously, to characterize 3D facial surfaces, our proposed Normal Local Gabor Binary Patterns descriptor (N-LGBPs) employ LGBP operator to encode the facial normal components. That is to say, we continue to encode the micro-patterns of normal Gabor magnitude images by the LBP operator.

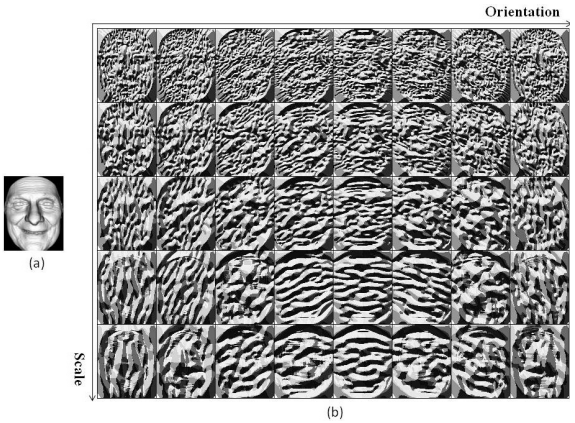

Figure 5. (a). Normal component z. (b) Its LGBP representation.

Formally, given a point p_{ij}^k of normal component k, the derived N-LGBPs decimal value is:

$$N\text{-}LGBPs_{\mu,v}^k(Q_{n,m}(p_{ij}^k)) = \sum_{q=1}^{m-1} t(O_{\mu,v}^k(q) - O_{\mu,v}^k(0))2^q, \quad (10)$$

where $Q_{n,m}$, t, q, 0 have the same meaning as Eq. (5), and $O_{\mu,v}^k$ represents the Gabor representations of normal component k as Eq. (9). See Fig. 5 for an example of $N\text{-}LGBPs_{\mu,v}^z(Q_{1,8})$ based normal representation. Like LBP, to avoid the loss of spatial information of facial normal representation by histograms, each Normal Gabor Magnitude image is further divided into several non-overlapping regions. The global representation can be obtained by concatenating (by facial configuration) all the N-LGBPs histograms extracted from each sub-regions. Similarly, the comprehensive description of Normal Local Gabor Binary Patterns (N-LGBPs) can be achieved using feature level fusion or score-level fusion.

5. Sparse Representation Classifier

Based on the face subspace model: a well-aligned frontal face image under different lighting conditions and various facial expressions lies close to a special low-dimensional linear subspace spanned by sufficient training samples from the same subject, Wright *et al.* [22] modeled the face recognition problem by solving the l_1 minimization sparse representation. They also proposed sparse representation classifier (SRC) for robust 2D face recognition.

Analogously, we assume that a well-aligned frontal test 3D face model represented by a feature vector under dif-

ferent facial expressions approximately lies in a linear subspace spanned by the 3D faces in the training set (represented by the same type of facial features) associated with the same subject.

That is, given n_i training samples of i-th subject, $A_i = [v_{i,1}, v_{i,2}, \ldots, v_{i,n_i}] \in \mathbb{R}^{m \times n_i}$, any test sample $y \in \mathbb{R}^m$ from the same subject can be represented by:

$$y = \alpha_{i,1} v_{i,1} + \alpha_{i,2} v_{i,2} + ... + \alpha_{i,n_i} v_{i,n_i}, \qquad (11)$$

where $\alpha_{i,j} \in \mathbb{R}, j = 1, 2, \ldots, n_i$.

However, the case is not exactly the same, in our experiments, the training set is composed of one face model from each subject (gallery set) ($n_i = 1$). This problem of insufficient training samples introduces a new model error, noted by $\varepsilon \in \mathbb{R}^m$. Model (8) can be modified as:

$$y \approx \alpha_i v_i = \alpha_i v_i + \varepsilon, \qquad (12)$$

where $y \in \mathbb{R}^m$, $v_i \in \mathbb{R}^m$ and $\alpha_i \in \mathbb{R}$ represent a probe face, a gallery face from the same subject and their linear scalar factor respectively.

Considering the whole gallery set with n 3D faces, each of which belongs to one subject, $A \doteq [v_1, v_2, \ldots, v_n] \in \mathbb{R}^{m \times n}$ and any probe $y \in \mathbb{R}^m$, (12) can be rewritten as

$$y = A\mathbf{x} + \tilde{\varepsilon}, \qquad (13)$$

where $\mathbf{x} = [0, \ldots, 0, \alpha_i, 0, \ldots, 0]^T \in \mathbb{R}^n$ is the coefficient vector whose entries are zero except the one associated with the i-subject. Sparse coefficients $\mathbf{x}$ in (13) can be solved as the following l_1 minimization problem:

$$\hat{x}_1 = argmin_x ||x||_1 \text{s.t.} ||A\mathbf{x} - y||_2 \leq ||\tilde{\varepsilon}||_2, \quad (14)$$

We employ OMP [15] algorithm to solve (14) and compute the residuals:

$$r_i(y) = ||y - A\delta_i(\hat{x}_1)||_2, i = 1, 2, \ldots, n. \qquad (15)$$

where δ_i is a characteristic function which selects coefficient associated with the i-th gallery. Finally, the index of minimal $r_i(y)$ corresponding to the identity of y.

6. Experimental and Algorithmic Settings

6.1. Datasets and Preprocessing

To evaluate the performance of our proposed descriptors to distinguish identical twins, as used in [20], all our results were reported on the 3D TEC dataset. 3D TEC is a subset of the Twins Days 2010 dataset acquired at the Twins Days Festival in Twinsburg, Ohio [1]. The whole dataset contains 266 subject sessions, with the 3D scans in the dataset containing two scans: one with a neutral expression and another with a smiling expression. There are 106 sets of identical twins, one set of triplets, and the rest are non-twins. The subset (i.e. 3D TEC) only includes the 3D face scans acquired in the first session and thus consists of of 107 pairs of twins (two of the triplets are included as the 107th set of twins). More instructions of data parameters can be found in [20].

All 3D face models were preprocessed using the tool developed by Szeptycki *et al.* [18], including spike removing, hole filling. Then manually labeled nose tips provided by the dataset were used for face cropping and an ICP fine registration. After normal estimation from each normalized range image, three normal component matrices: $[n_{ij}^x]_{m \times n}, [n_{ij}^y]_{m \times n}$ and $[n_{ij}^z]_{m \times n}$ are resized to the same size of 120×96 for the following encoding step.

6.2. Experimental Settings

We did experiments according to the same experimental protocol as defined in [20]. To be specific, one person in each pair of twins was arbitrarily labeled as Twin A and the other as Twin B. Face identification experiments were performed using four different gallery and probe sets as shown in Table 1.

No.	Gallery	Probe
I	A Smile, B Smile	A Neutral, B Neutral
II	A Neutral, B Neutral	A Smile, B Smile
III	A Smile, B Neutral	A Neutral, B Smile
IV	A Neutral, B Smile	A Smile, B Neutral

Table 1. Gallery and probe sets for cases I, II, III, and IV. "A Smile, B Neutral" means that the set contains all images with Twin A smiling and Twin B neutral [20].

In Case I, all the images in the gallery set possess a smiling expression while all the images in the probe set have a neutral expression. Case II reverses these roles of Case I. Both the two cases model a scenario that the gallery has one expression and the probe has another. As stated in [20], in the identication scenario, theoretically the main challenge would be to distinguish between the probe image and the image of his/her twin in the gallery since they look similar.

In Case III, Twin A smiling and Twin B neutral make up of the gallery set; while Twin A neutral and Twin B smiling as the probe compose the probe set. Case IV reverses these roles of Case III. These two cases model a worst scenario in which the system does not control for the expressions of the subject in a gallery set of twins. As pointed out in [20] as well, in the identication scenario, theoretically the main challenge would be to distinguish between the probe image and the image of his/her twin in the gallery. This is more difficult than Cases I and II since the expression of the probe face is different from his/her image in the gallery but is the same as the image of his/her twin in the gallery.

6.3. Algorithmic Settings

On Normal Local Binary Patterns descriptor (N-LBPs, i.e. MS-LNPs), as introduced in [14], each normal component matrix is empirically divided into local patches of sizes 12×12, 20×16 and 40×32 for the operators LNPs of $Q_{1,8}$, $Q_{2,16}$ and $Q_{3,24}$ respectively. Thus three LBP histograms are achieved with dimensions of 4,720, 8,748 and 4,995. Each of this histogram can be used for score level fusion based comprehensive representation directly. For feature level fusion based comprehensive representation, the three histograms are concatenated to achieve the final feature vector (with a dimension of 18,463).

On Normal Gabor Filters descriptor (N-GFs), for feature level fusion based comprehensive representation, each Gabor magnitude image (size of 120×96) is first down-sampled by a factor of 8 and then reshaped to a column vector with a dimension of 180 (15×12). The final feature vector (with a dimension of 7,200) is achieved by concatenating all the 40 column vectors. For score level fusion based comprehensive representation, each Gabor magnitude image is directly reshaped to a column vector with dimension of 11,520 (120×96).

On Normal Local Gabor Binary Patterns descriptor (N-LGBPs), empirically, we use LBP operator of $Q_{2,16}$ with local patches of 20×16. For feature level fusion based comprehensive representation, all the Normal Gabor Magnitude images are first down-sampled by a factor of 2 before extracting LBP histograms. By concatenating all the 40 LBP based histograms, the final feature thus has a dimension of 87,480 ($40 \times 2,187$). For score level fusion based comprehensive representation, each Normal Gabor Magnitude image is directly encoded by the LBP histogram with a dimension of 8,748.

PCA is employed to reduce the huge dimensions of the N-GFs and N-LGBPs descriptors. For all the cases, including the two descriptors with two kinds of comprehensive representations, according to the number of the gallery samples, we fix the compacted dimension of features at 205. To solve (14), OMP algorithm with sparse coding number of 30 was used.

6.4. Experimental Results

6.4.1 Comparison of comprehensive representations

To evaluate the influences of different comprehensive representations related to different facial normal descriptors, we compared their rank-one performances at both feature level and score level fusion according to Case I. Since the effectiveness of fusion different normal components have been proved by Li *et al.* [14], all the results shown in this study are the final ones obtained by fusing the three normal components. From Tab. 2, we can see that score level fusion is better for N-LBPs; while feature level fusion is better for N-GFs and N-LGBPs. Therefore, we report our following results based on their better schemes, i.e. score level fusion for N-LBP, feature level fusion for N-LGFs and N-LGBPs.

Descriptor	**Rank-1 Recognition Rate**	
	score-level fusion	feature-level fusion
N-LBPs	**94.86%**	92.52%
N-GFs	88.79%	**93.93%**
N-LGBPs	93.46%	**96.73%**

Table 2. Comparison of feature level and score level fusion based comprehensive representation (tested in Case I))

6.4.2 Comparison of identification results

Tab. 3 shows the rank-one recognition rate of our proposed method (noted as Alg. 5) and the algorithms reported in [20]. We can find that N-LGBPs performs best among the group of proposed normal descriptors for all the four cases. N-LBPs and N-GFs achieved comparable results for all the four cases. All the proposed descriptors work better for Cases I and II than Cases III and IV. Compared with other algorithms, N-LGBPs displays comparable results to the best results (given by Alg. 4) for all these four cases. This indicates that our proposed normal descriptors (especially N-LGBPs) prove very high discrimination to recognize identical twins.

Algorithm	**Rank-1 Recognition Rate**			
	I	II	III	IV
Alg.1 (E_{pkn}) [4]	93.5%	93.0%	72.0%	72.4%
Alg.1 (E_{minmax}) [4]	94.4%	93.5%	72.4%	72.9%
Alg.2 (SI) [10]	92.1%	93.0%	83.2%	83.2%
Alg.2 (eLBP) [8]	91.1%	93.5%	77.1%	78.5%
Alg.2 (Range PFI) [9]	91.6%	93.9%	68.7%	71.0%
Alg.2 (Text. PFI) [9]	95.8%	96.3%	91.6%	92.1%
Alg.3 [5]	62.6%	63.6%	54.2%	59.4%
Alg.4 [11]	**98.1%**	**98.1%**	**91.6%**	**93.5%**
Alg.5 (N-LBPs)	94.9%	96.3%	89.3%	88.3%
Alg.5 (N-GFs)	93.9%	94.4%	89.3%	90.2%
Alg.5 (N-LGBPs)	**96.7%**	**96.7%**	**92.5%**	**93.5%**

Table 3. Comparison of rank-one scores of our method (Alg. 5) and the state-of-the-art methods.

7. Conclusion and Perspective

In this paper, we proposed a group of facial normal descriptors consisting of Local Normal Patterns (N-LBPs, i.e. MS-LNPs), Normal Gabor Filters (N-GFs) and Normal Local Gabor Binary Patterns (N-LGBPs). Their discriminations for distinguishing 3D identical twins were evaluated in the face identification scenario. Meanwhile, both feature level fusion and score level fusion based comprehensive representations of each descriptor were compared. We achieved the state-of-the-art results.

In our future work, we will mainly investigate the follow issues: 1) As used in [14], we will study the learning-based weighted sparse representation method for N-GFs and N-LGBPs to resist facial expression variations. 2) Except P-CA, other more efficient techniques for dimensionality reduction and feature selection methods, such as Linear Discriminant Analysis (LDA), AdaBoost, manifold learning, etc. will be discussed for N-GFs and N-LGBPs descriptors.

Acknowledgements

This work is in part jointly supported by the French research agency, Agence Nationale de Recherche (ANR) and Natural Science Foundation of China (NSFC), within the 3D Face Analyzer (grant ANR 2010 INTB 0301 01; grant NSFC 61061130560), and the 3D Face Interpreter project supported by the LIA 2MCSI lab between the group of Ecoles Centrales and Beihang University.

References

[1] Twins days. Available: http://www.twinsdays.org/.

[2] T. Ahonen, A. Hadid, and M. Pietikäinen. Face recognition with local binary patterns. European Conference on Computer Vision (ECCV), 2004, pp. 469–481.

[3] J. Cook, V. Chandran, and C. Fookes. 3d face recognition using log-gabor templates. British Machine Vision Conference (BMVC), Edinburgh, UK, 2006, pp.769–778.

[4] T. C. Faltemier, K. W. Bowyer, and P. J. Flynn. A region ensemble for 3d face recognition. *IEEE Transactions on Information Forensics and Security*, 3(1):62–73, 2008.

[5] B. Gokberk, M. O. Irfanoglu, and L. Akarun. 3d shape-based face representation and feature extraction for face recognition. *Image and Vision Computing*, 24(8):857–869, 2006.

[6] P. J. Grother, G. W. Quinn, and P. J. Phillips. Report on the evaluation of 2d still-image face recognition algorithms. NIST Interagency/Internal Report (NISTIR) - 7709, 2010.

[7] R. Hoffman and A. K. Jain. Segmentation and classification of range images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):608–620, 1987.

[8] D. Huang, M. Ardabilian, Y. Wang, and L. Chen. A novel geometric facial representation based on multi-scale extended local binary patterns. IEEE International Conference on Automatic Face and Gesture Recognition (FG), Santa Barbara, CA, USA, 2011.

[9] D. Huang, W. B. Soltana, M. Ardabilian, Y. Wang, and L. Chen. Textured 3d face recognition using biological vision-based facial representation and optimized weighted sum fusion. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops(CVPRW), Colorado Springs, CO, USA, 2011.

[10] D. Huang, G. Zhang, M. Ardabilian, Y. Wang, and L. Chen. 3d face recognition using distinctiveness enhanced facial representations and local feature hybrid matching. International Conference on Biometrics: Theory, Applications and Systems (BTAS),Washington D.C., USA, 2010.

[11] 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 Transactions on Pattern Analysis and Machine Intelligence*, 29(4):640–649, 2007.

[12] B. Klare, A. A. Paulino, and A. K. Jain. Analysis of facial features in identical twins. International Joint Conference on Biometrics (IJCB), Washington, DC, 2011.

[13] T. S. Lee. Image representation using 2d gabor wavelets. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18:959–971, 1996.

[14] H. Li, D. Huang, J. M. Morvan, and L. Chen. Learning weighted sparse representation of encoded facial normal information for expression-robust 3d face recognition. International Joint Conference on Biometrics (IJCB), Washington, DC, 2011.

[15] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In Proc. 27th Ann. Asilomar Conf. Signals, Systems, and Computers, 1993.

[16] P. J. Phillips, P. J. Flynn, K. W. Bowyer, R. W. V. Bruegge, P. J. Grother, G. W. Quinn, and M. Pruitt. Distinguishing identical twins by face recognition. IEEE International Conference on Automatic Face and Gesture Recognition (FG), 2011.

[17] Z. Sun, A. A. Paulino, J. Feng, Z. Chai, T. Tan, and A. K. Jain. A study of multibiometric traits of identical twins. SPIE Biometric technology for human identification VII, Vol. 7667, Orlando, FL, USA, Apr. 2010, pp. 1–12.

[18] P. Szeptycki, M. Ardabilian, and L. Chen. A coarse-to-fine curvature analysis-based rotation invariant 3d face landmarking. International Conference on Biometrics: Theory, Applications and Systems (ICB), 2009, pp.3206–3211.

[19] V. Vijayan, K. W. Bowyer, and P. J. Flynn. 3d twins and expression challenge. IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, 2011.

[20] V. Vijayan, K. W. Bowyer, P. J. Flynn, D. Huang, L. Chen, M. Hansen, O. Ocegueda, S. K. Shah, and I. A. Kakadiaris. Twins 3d face recognition challenge. International Joint Conference on Biometrics (IJCB), Washington, DC, 2011.

[21] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):775–779, 1997.

[22] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009.

[23] C. Xu, S. Li, T. Tan, and L. Quan. Automatic 3d face recognition from depth and intensity gabor features. *Pattern Recognition*, 42(9):1895–1905, 2009.

[24] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang. Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. IEEE International Conference on Computer Vision (ICCV), 2005, pp. 786–791.