

High-Resolution Face Verification Using Pore-Scale Facial Features

Dong Li, Huiling Zhou, and Kin-Man Lam

Abstract—Face recognition methods, which usually represent face images using holistic or local facial features, rely heavily on alignment. Their performances also suffer a severe degradation under variations in expressions or poses, especially when there is one gallery per subject only. With the easy access to high-resolution (HR) face images nowadays, some HR face databases have recently been developed. However, few studies have tackled the use of HR information for face recognition or verification. In this paper, we propose a pose-invariant face-verification method, which is robust to alignment errors, using the HR information based on pore-scale facial features. A new keypoint descriptor, namely, pore-Principal Component Analysis (PCA)-Scale Invariant Feature Transform (PPCASIFT)—adapted from PCA-SIFT—is devised for the extraction of a compact set of distinctive pore-scale facial features. Having matched the pore-scale features of two-face regions, an effective robust-fitting scheme is proposed for the face-verification task. Experiments show that, with one frontal-view gallery only per subject, our proposed method outperforms a number of standard verification methods, and can achieve excellent accuracy even the faces are under large variations in expression and pose.

Index Terms—Pore-scale facial feature, face verification, face recognition, pose invariance, expression invariance, alignment-error-robust.

I. INTRODUCTION

FACE verification is a one-to-one matching problem, which validates (or not) the claimed identity of a person. The claim is either accepted or rejected based on a certain threshold. Face verification has been widely used in security systems and electronic commercial systems due to the easy access to face acquisition.

Many of the face recognition algorithms are based on holistic facial features, which project the lexicographic ordering of raw pixels onto a certain subspace. Examples include eigenfaces [1], fisherfaces [2], locality preserving projection (LPP) [3], etc. All these algorithms will suffer significant degradation in their performances when the face images

considered are under pose, expression, and/or illumination variations. Local features can be used to achieve better performances under the different variations. The local features are extracted from local regions or parts of the images only, based on some transformations or descriptors. Commonly used local features include Gabor wavelets [4], local binary pattern (LBP) [5], etc. The Gabor representation of faces is similar to that of the human visual system, which is robust against illumination and expression changes. LBP was originally designed for texture classification [5], [6], and was introduced to face recognition in [7]. The representation is invariant to illumination changes and rotation. Recently, the Local Phase Quantization (LPQ) operator [8] was proposed for the recognition of blurred faces based on quantizing the Fourier transform phase in local neighborhoods. Another method, named Patterns of Oriented Edge Magnitudes (POEM) [9], was proposed to extract facial features by applying a self-similarity operator on accumulated edge magnitudes across different directions. Multi-Directional Multi-Level Dual-Cross Patterns (MDML-DCPs) [10] was proposed to encode discriminative information along the directions of facial components. However, feature representations and face recognition algorithms always require the face images to be normalized and aligned to achieve a satisfactory accuracy level. In addition, the pose, expression, and illumination variations will cause non-linear distortions on the 2D face images, due to the fact that the facial features (eyes, nose, mouth, etc.) do not appear on a planar surface. When the 3D faces are projected onto a 2D plane, non-linear distortions will happen, which linear subspace methods cannot cope with. Although non-linear kernel methods [11], [12] can be used, the improvement is still limited and the methods become more computationally intensive.

A lot of research has been conducted to solve non-frontal-view face recognition, including multi-view [13], [14], cross-view [4], [15], [16], and matching-based [17], [18] face recognition. Multiview face recognition requires a number of training images at different poses per subject, which is impractical in real applications. The cross-view approach usually applies 2D or 3D appearance models to synthesize face images in specific views. Alignment is needed to establish correspondences between two faces in different poses. Matching-based methods attempt to establish local correspondences between the gallery and query images, which are usually alignment-free. Our proposed method adopts this advantage, so the accurate alignment of faces is unnecessary.

With the development of multimedia hardware such as HDTV and digital cameras, it has become easy to access high-resolution (HR) images. This enables us to analyze more sophisticated feature, in addition to the traditional facial

Manuscript received October 3, 2013; revised October 6, 2014 and January 2, 2015; accepted February 23, 2015. Date of publication March 12, 2015; date of current version April 15, 2015. This work was supported by The Hong Kong Polytechnic University, Hong Kong, under Project G-YBA2. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. A. N. Rajagopalan.

D. Li was with the Center for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong. He is now with the Faculty of Automation, Guangdong University of Technology, Guangzhou 510006, China (e-mail: leedong111@gmail.com).

H. Zhou and K.-M. Lam are with the Center for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong (e-mail: hl.zhou@connect.polyu.hk; enkmlam@polyu.edu.hk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2015.2412374

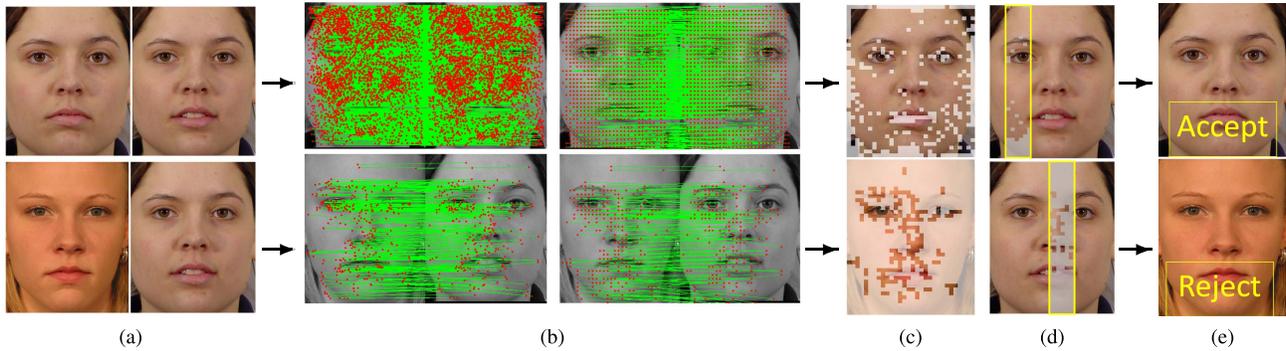


Fig. 1. The face-verification framework based on the PPCASIFT feature: (a) two examples of query pairs for face verification - the top row is a client and the claimed identity, and the bottom row is an impostor and the claimed identity; (b) the pore-scale facial-keypoint-matching results and the initial block-matching results between the query and the claimed-identity images; (c) the refined block-matching results (the color regions represent those aggregated, matched blocks); (d) the local region (in the yellow box) of the gallery image which has the maximum local-matching density; and (e) the verification results.

features like face shape, eyes, nose, and mouth. HR face recognition is a relatively recent topic; it extracts subtle and detailed information, such as mark-scale features (e.g. moles, scars) and pore-scale features (e.g. pores, hair), which contains more distinctive information than low-resolution (LR) face images do.

An analysis from macrocosm to microcosm was proposed [19] to solve HR face recognition. This method uses Gabor filters to extract pore-scale features (namely skin texton) to form a texton histogram (a similar idea to bag-of-words [20]). Then, regularized LDA is applied to preserve intrinsic information and reduce interference. The Gabor-based skin texture extracted achieves a recognition rate of between 38.5% and 57.3%, so the skin texture can be used as an auxiliary feature only. Also, since the method uses LDA, more than one HR face is needed for training. This may not be feasible for some real-world applications.

Another HR face-recognition method was proposed based on facial-marker detection [21]. This method uses LoG blob detection for marker extraction after applying the active appearance model (AAM) [22] to detect and remove facial features such as the eyes, nose, mouth, etc. However, only a very limited quantity of marker-scale features can be extracted from human faces, so the features only complement the traditional methods.

Similarly, in [23], facial marks, which are manually annotated by multiple observers, are used as biometric signatures to distinguish between identical twins. That work paid more attention to particular biometric traits, like facial markers, than to the overall facial appearance. However, there is no guarantee that a face image has a sufficient number of traits (e.g. scars, moles, freckles, etc.) for recognition.

With human biology, it is impossible for two people, even identical twins [24], to have an identical skin appearance. Inspired by this idea, a novel pore-scale facial feature has been proposed in [25]. By adapting the SIFT detector and descriptor to the pore-scale facial-feature framework, and using a candidate-constrained matching scheme, the algorithm [25] can establish a large number of reliable correspondences of keypoints between two face images of the same subject which may have a big difference in pose. Such pore-scale facial

features are dense and distinguishable, which are the desirable aspects for face verification.

In this paper, we propose a face-verification algorithm based on the pore-scale facial features to take advantage of the HR information. One of the major advantages of the proposed approach is that the facial-skin regions under consideration are usually more linear - i.e. approximate to a planar surface - than other facial features, so the recognition performance will be very robust to pose, expression, and illumination variations, etc. Furthermore, only one gallery sample per subject is needed, and an accurate face alignment is not necessary to achieve a good performance. An overview of our proposed framework is shown in Fig. 1. Firstly, the pore-scale facial features are detected and extracted from a testing or a query image. Then, initial keypoint matches between the testing image and that of the claimed identity in the gallery are established; the initial keypoint matches are then converted to block matches, which are further aggregated to eliminate the outliers, as shown in Fig. 1(b) and (c). Finally, the verification result is determined based on the maximum density of the local, aggregated, matched blocks on the face images, as illustrated in Fig. 1(d). In our experiments, we will show the superior performance of our face-verification algorithm compared to the standard face-verification methods.

The contributions of this paper and the novel aspects of the proposed method are listed as follows:

- An alignment-error-insensitive and pose-invariant face-verification approach is proposed. In other words, only the approximate locations of facial features such as the eyes and mouth are necessary. Non-frontal-view face images do not need to be included in the gallery. These make our method suitable for practical and real applications.
- To the best of our knowledge, our method is the first to perform face verification using pore-scale facial features rather than landmark-features (e.g. contours, eyes, nose, mouth) or marker-scale features (e.g. moles, scars).
- A new descriptor is proposed, namely Pore-Principal Component Analysis (PCA)-Scale Invariant Feature Transform (SIFT) (PPCASIFT), which can achieve a similar performance to the Pore-SIFT (PSIFT) [25]

descriptor but which requires only 9% of the PSIFT descriptor’s computation time in the matching stage.

- A fast and robust fitting method is proposed to establish the block matching of two faces based on matched keypoints, which considers the non-rigid structure of faces and which can also remove outliers at the same time.
- A pose-invariant similarity measure, namely the maximized local-matching density, is proposed to provide a normalized similarity measure for pose-invariant face verification. Based on the maximized local-matching density, no prior knowledge of pose information is needed.

The remainder of this paper is organized as follows. Section II briefly reviews the detection of pore-scale facial features. Section III presents two descriptors for the pore-scale facial features, namely PSIFT and PPCASIFT. Section IV introduces the feature-matching and robust fitting method for refining the correspondences of blocks between two faces for face verification. Section V describes a new similarity measure for face verification, namely the maximized matching density, which makes our method robust to pose variations. Section VI presents the experimental results and compares our face-verification method with some standard face-verification methods. In particular, we will evaluate the performances of the different methods when the query faces are under pose and expression variations. Section VII summarizes this paper, and discusses our future work on the pore-scale facial feature.

II. PORE-SCALE FACIAL-FEATURE DETECTION

Pore-scale facial features include pores, fine wrinkles and hair, which commonly appear in the whole face region. Most of the pore-scale facial features are blob-shaped features. Hence, the PSIFT detector [25] employs the Hessian-Laplace detector on the multiscale difference of Gaussians (DoG) for blob detection.

In order to generate a sufficient number of correspondences between two face images, a large number of reliable feature points should first be detected on each of the two faces. Meanwhile, from the biological viewpoint, different people should have a similar quantity of pores in their facial skin. Hence, the PSIFT detector employs a quantity-driven approach via an adaptive threshold. To avoid the extremely dense DoG responses at hairy (e.g. bearded) areas, the PSIFT detector estimates the adaptive threshold based on a cropped skin region in the cheek instead of the whole face image, as shown in Fig. 2. With a consideration of both the robustness and the completeness of matching, the number of detected pore-scale keypoints N_k in the cropped region is specified to be within the range [210, 240] by an adaptive threshold. To determine the value of the adaptive threshold, the binary search method [26] is performed on a threshold list, which is from 0 to 0.0025. Also, the quantity of inliers is used to evaluate the matching performance by experiments at different sampling frequencies and different priors of smoothing, respectively. Thus, the number of DoG octaves is set at 3, and 8 scales are sampled in each DoG octave. In addition, the prior smoothing, which is applied to each image level

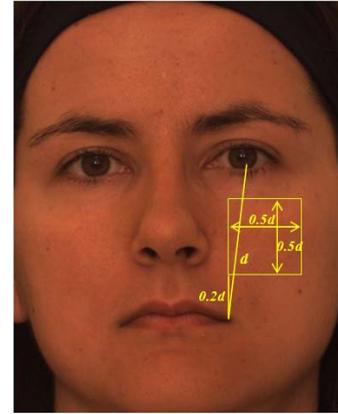


Fig. 2. The size of a cropped skin region, whose pore features are to be extracted.

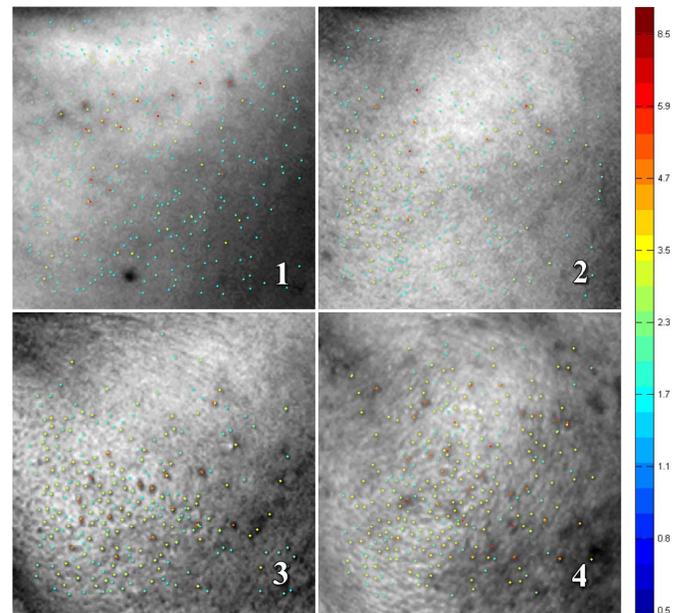


Fig. 3. Visualization of keypoints on the skin images of 4 distinct subjects (different colors for the keypoints indicate their scales as represented by the color bar).

before building the scale-space representation, is set at 1. Usually, about 4,500 keypoints are detected for a whole face region. Fig. 3 illustrates the keypoint-detection results on the cropped region of four distinct subjects, where keypoints with different scales are represented using different colors.

III. PORE-SCALE FACIAL-FEATURE DESCRIPTION

Most pore-scale facial features are similar to each other when they are observed individually, because most of them are blob-shaped, and the surrounding region of each keypoint has almost the same color. However, the spatial distribution of pores on the skin is distinctive. Based on this biological observation, designing a distinctive pore-scale facial-feature descriptor becomes possible.

PSIFT [25], adapted from SIFT [27], was proposed using the gradient information of neighboring keypoints to describe the textures and the spatial distribution of pores. The number

of sub-regions and the support size of each sub-region are expanded in PSIFT so as to extract the relative-position information about the neighboring keypoints. In this way, a PSIFT descriptor is constructed from the gradient orientations within a region containing 8×8 sub-regions with each sub-region represented by a histogram of 8 orientation bins. Therefore, PSIFT is represented as a 512-element feature vector for each keypoint description. In addition, the keypoints are not assigned a main orientation because most of them are blob-shaped and do not have a coherent orientation. Furthermore, as the rotation of a face image is usually not large, generating a rotation-free description of the pore-scale facial feature is not necessary.

However, matching two keypoints using descriptors of 512 dimensions is computationally expensive. To improve the method's efficiency, we propose a more compact keypoint descriptor in this paper, namely Pore-PCA-SIFT (PPCASIFT), which is adapted from PCA-SIFT [28] and which uses PCA to reduce the dimensionality of the descriptor. We extract the description of a keypoint using a patch/sampling size of 41×41 , with the keypoint at the center, at a given window size (48 times the scale determined by the PSIFT detector). The parameters are determined by a Powell method [29] based on the verification rates on a small dataset. The initial setting was chosen so that the PPCASIFT descriptor window has the same size as PSIFT. Unlike PCA-SIFT, the patches of pore-scale facial keypoints do not need to be aligned or assigned their main orientations. However, PCA can still represent these patches. The main reason for this is that the patches contain the relative spatial information about the neighboring keypoints, and the patterns of patches are relatively simple. If a sufficient number of patches is available for learning the principal components, the patches can then be represented efficiently in a much lower subspace. We selected 16 face images from 4 distinct subjects with different skin conditions in order to extract about 90,000 patches (after removing the keypoints near the borders). The horizontal and vertical gradient maps in the 41×41 patch are computed, and are represented by a vector containing $2 \times 39 \times 39 = 3,042$ elements. The vectors are normalized to unit magnitude, and then PCA is applied to these training vectors. The 72 leading eigenvectors are used to form the projection matrix for PPCASIFT, which is of dimension $3,042 \times 72$. To generate the PPCASIFT descriptor for a given keypoint, its normalized gradient vector is computed and is then projected onto the eigenspace formed by the 72 eigenvectors. Compared to the PSIFT descriptor which has a dimension of 512, the dimension of the PPCASIFT descriptor is 72 only. In other words, the PPCASIFT feature is much more compact and computationally efficient in the matching stage than PSIFT.

IV. PORE-SCALE FACIAL-FEATURE MATCHING AND ROBUST FITTING

In [25], a double-matching scheme (namely *candidate-constrained matching*) was proposed to narrow the matching of keypoints between two face images and to achieve accurate face matching, based on both intra- and inter-scale

facial information. RANSAC [30], as a robust fitting method, is then applied to the matched keypoints to remove the outliers reliably. To perform face verification efficiently, we modify the candidate-constrained matching from two passes to one pass only. In addition, a new robust fitting scheme, namely *parallel-block aggregation*, is proposed to refine the candidate-constrained matching results. As the keypoint/block matching may result in one-to-many or many-to-many matches, matching from gallery faces may differ from that from testing faces. In our experiments, we only consider the block matching from a testing face image to a gallery face image.

A. Feature Matching

For verifying whether or not two face images are of the same identity, correspondences are established from the query image to the gallery image of the claimed identity. Suppose that the position and the scale of a keypoint in the query image are (x^q, y^q) and σ^q , respectively, while the position and the scale of the i^{th} keypoint in the claimed image are (x_i^c, y_i^c) and σ_i^c , respectively. Assume that the height of the gallery image is H .

First, the spatial information of the face image is considered in feature matching. Considering that the poses of faces are limited to within a certain range, and the y coordinates of the two keypoints from the two face images at different poses are close, then the position of the matched keypoint in the gallery image should satisfy the following constraint:

$$|y_i^c - y^q| < \lambda H, \quad (1)$$

where λ is a factor and is set at 0.2 in our experiments.

Second, the scales of the two keypoints, one from the query and the other from the gallery, should be close to each other. Therefore, the ratio of the scale of the keypoint in the query image and the i^{th} keypoint in the gallery image should be close to 1, and is defined to be within the range as follows:

$$1/\mu \leq |\sigma_i^c / \sigma^q| \leq \mu, \quad (2)$$

where μ is a constant larger than 1. When μ is close to 1, the scales of the two keypoints are similar. In our experiments, μ is set at 2.

Based on these two constraints, the number of keypoint candidates is narrowed to about 30% of all the keypoints in the gallery image. Then, the distances between a keypoint in the query image and the remaining keypoints in the gallery image are computed. The distance between two keypoints is measured using the Euclidean distance between their corresponding pore-scale feature descriptors. The best-matched keypoint in the gallery image is the one with the smallest Euclidean distance. We define the *distance ratio*, which is the ratio between the distances of a keypoint from the query image to its nearest keypoint and to its second-nearest keypoint in the gallery image. We accept the match if the distance ratio is smaller than 0.85, which is set empirically by experiments. The initial matching results using our method are shown in Fig. 4; this shows the effectiveness of using the pore-scale facial features.

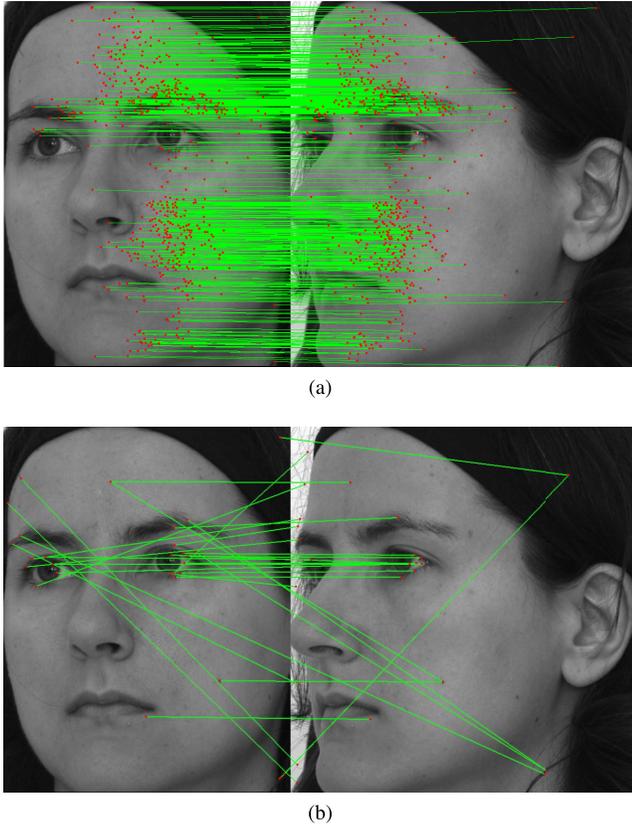


Fig. 4. Initial matching results for face images of pose R10 and pose R45: (a) PPCASIFT, and (b) SIFT. The matching results of PSIFT are much denser and more structurally accurate than those of SIFT.

B. Robust Fitting

After the detection, description, and initial matching of the keypoints, a large number of matched keypoint pairs between a query/testing face and a gallery/claimed face have been established. However, the matching results still include many outlier pairs. Consequently, further refinement is necessary to improve the matching accuracy. In [25], RANSAC [30] is used to refine the matching results by fitting to the epipolar constraint. However, this process is of high complexity and requires a large number of iterations, which is not desirable for real-time face verification. In addition, the number of matches cannot be used directly as a similarity measure for verification, because the number depends on the degree of variation between the two faces to be matched. In this paper, we propose a more efficient and effective scheme to refine the matches and provide a normalized measure of the correspondences for face verification. Our basic idea is to transfer the keypoint correspondences to block-based correspondences. The line connecting two correctly matched blocks in the two face images should be approximately parallel to the other lines of the corresponding neighboring blocks.

First, matched keypoint pairs are transformed into matched block pairs, which can further remove some outliers. All the face images are divided into non-overlapping blocks of size $W' \times H'$. Assume that a query face image, I_q , establishes keypoint correspondences to a gallery face image, I_g .

Algorithm 1 Parallel-Block-Aggregation Algorithm

```

1: Transform the matched keypoint pairs to initially-matched
   block pairs
2: for each block  $B'$  of the  $W' \times H'$  blocks do
3:   for each initially-matched block  $B$  do
4:      $N_B = 0$ 
5:     for each of the  $n_u$  neighbors  $U$  do
6:       Compute the location of  $U'_{\text{perfect}}$  based on  $B$ ,  $B'$ ,
         and  $U$ 
7:       if  $\text{distance}(U'U'_{\text{perfect}}) < R$  then
8:          $n_B = n_B + 1$ 
9:          $U'$  is a parallel-supporting block
10:      end if
11:    end for
12:    if  $n_B \geq n_t$  then
13:       $B'$  is aggregated
14:    end if
15:  end for
16: end for

```

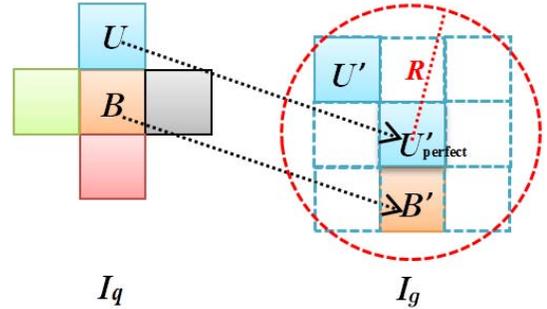


Fig. 5. Use block B and its upper neighboring block U in a query image (I_q) to illustrate the parallel-block-aggregation scheme. B' and U' in a gallery image (I_g) are the corresponding matched blocks of B and U , respectively. U'_{perfect} is the perfect location for U such that UU'_{perfect} is exactly parallel to BB' . However, human faces are non-rigid and may have expressions. Block B' is aggregated if U' is inside the neighborhood of U'_{perfect} , which is represented as a circle with the red, dashed line.

If a keypoint resides in a block, denoted as B , in I_q and is matched to a keypoint in I_g which resides in block B' , the block pair B and B' is considered to be initially matched. As shown in Fig. 1(b), the initial block-matching result of an impostor is much sparser than that of a genuine subject; this characteristic is useful for face verification.

However, some of these initially matched block pairs are still outliers. In order to achieve a robust and accurate face-verification performance, we propose a new, robust fitting scheme, namely parallel-block aggregation. Algorithm 1 shows the pseudocode of the parallel-block-aggregation algorithm, and Fig. 5 illustrates the robust fitting scheme. Block B and one of its n_u neighboring blocks, U , in I_q are initially matched to B' and U' in I_g , respectively. For the perfect matching of an inlier, U should match U'_{perfect} . However, due to the fact that faces are non-rigid and may have local changes caused by facial expressions, the matching from U to any block in the neighborhood of U'_{perfect} within a certain radius, R , is considered valid, i.e. the line joining U and U' ,

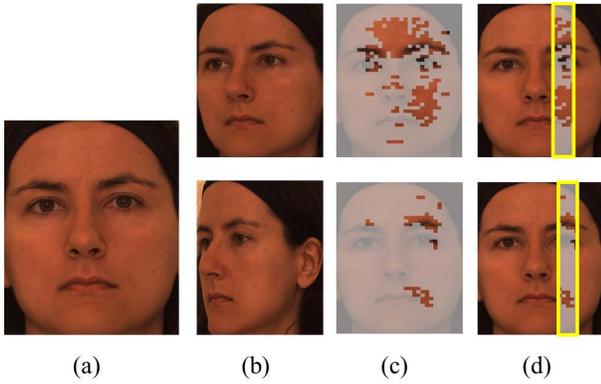


Fig. 6. Selection of a local region with the maximized local-matching density: (a) the frontal-view gallery face image, (b) two query face images at pose “right 10°” (the 1-st row) and pose “right 45°” (the 2-nd row), (c) the corresponding block-aggregation results on the gallery face for the two query images at different poses, and (d) the selected local regions in the two images with maximum local-matching density.

and that joining B and B' , are considered to be parallel. R is the threshold of the distance between U' and U'_{perfect} , which forms an acceptable, circular region for block matching. As illustrated in Fig. 5, the block U' inside the circular region is called a parallel-supporting block of B' . The block B' is aggregated if the number of its parallel-supporting blocks is larger than or equal to n_t . By experiments, we set $n_u = 4$, the distance threshold $R = 1.5 \times \text{block size}$, and $n_t = 1$, which can produce the best performance. Fig. 1(c) and Fig. 6(c) illustrate the parallel-block-aggregation results.

V. SIMILARITY MEASUREMENT

The area of the aggregated blocks (or the number of aggregated blocks) is variant to poses, due to the fact that the areas of a corresponding region in two faces with different poses are not the same. In this section, we propose a new normalized similarity measure, namely the maximized local-matching density, for face verification with pose variations.

A. Matching Density

Denote $\mathbf{R}(\mathbf{x})$ as a face region of a particular size located at \mathbf{x} . The total number of blocks in $\mathbf{R}(\mathbf{x})$ is counted, and is denoted as $N_{\text{total}}(\mathbf{R}(\mathbf{x}))$. After robust fitting, the number of matched blocks in the region $\mathbf{R}(\mathbf{x})$ is denoted as $N_{\text{matched}}(\mathbf{R}(\mathbf{x}))$. Then, the matching density ρ of the region $\mathbf{R}(\mathbf{x})$ is defined as follows:

$$\rho(\mathbf{R}(\mathbf{x})) = N_{\text{matched}}(\mathbf{R}(\mathbf{x})) / N_{\text{total}}(\mathbf{R}(\mathbf{x})), \quad (3)$$

where the value is within the range $[0, 1]$.

B. Local-Region Selection

The well-matched regions of two faces of the same subject are always unoccluded and more planar areas, as illustrated in Fig. 6(c). These regions have significantly fewer non-linear distortions when the faces to be matched have different poses. To ensure that the matching density of a well-matched local

region is robust to poses, the size of such a local region $\mathbf{R}(\mathbf{x})$ should always be smaller than the common area $\mathbf{R}'(\theta)$ of the two faces, where θ is the pose difference between the frontal-view gallery face image and the query image. Thus, the matching density of the local region $\rho(\mathbf{R})$ is more invariant to the pose difference θ . Generally, we define the size of the local region \mathbf{R} as in Fig. 6, by considering the following conditions:

$$\begin{aligned} N_{\text{total}}(\mathbf{R}(\mathbf{x})) &\leq \min_{\theta} N_{\text{total}}(\mathbf{R}'(\theta)) \approx N_{\text{total}}(\mathbf{R}'(45^\circ)) \\ &\approx 20\%W \times H. \end{aligned} \quad (4)$$

Determined by experiments, the size of the local region \mathbf{R} is set at $15\%W \times H$, as shown in Fig. 6.

Then, the location \mathbf{x} , where the local-matching density $\rho(\mathbf{R}(\mathbf{x}))$ is a maximum, is searched as follows:

$$P = \max_{\mathbf{x}} \rho(\mathbf{R}(\mathbf{x})), \quad (5)$$

where P is the maximized local-matching density of the local region \mathbf{R} (represented by the yellow box in Fig. 6(d)), which is insensitive to pose variations and which represents the similarity between the claimed gallery image and the query image.

VI. EXPERIMENTS

The performances of our proposed face-verification methods (based on PSIFT and PPCASIFT features) are evaluated using images under pose variations, expression variations, different capture times, and alignment errors. In all the experiments, only a single frontal-view face of each subject is in the gallery set. To compare the performances of the different methods, we measure the receiver-operating characteristic (ROC) curve by varying a threshold to produce different false-rejection rates (FRR) and false-acceptance rates (FAR). The equal-error rate (EER), where the above two rates are equal, is also measured.

A. Preprocessing

The performance of the pore-scale face-verification algorithm is evaluated on three public databases: the Bosphorus dataset [31], the Multi-PIE dataset [32], and the FRGC v2.0 dataset [33]. All the face images used in the experiments are converted to gray-scale images. For each database, a single neutral, frontal-view facial image of each subject is taken for the gallery.

The Bosphorus dataset [31] contains 4,666 HR face images of 105 subjects. The Multi-PIE database [32] contains 755,370 images of 337 subjects, which were recorded over a span of 6 months. Individual attendance at sessions where the HR images were captured varies from 203 to 249 subjects. Overall, 129 subjects appear in all four sessions, which are used in the experiment. The third dataset used is the FRGC v2.0 database [33]. It contains approximately 50,000 images of over 200 subjects, which were collected about once a week from 2002 (Fall) to 2004 (Spring). In the experiments, we use the landmark information to retrieve high-quality images from the FRGC v2.0 database. A total of 9,844 images, whose number of pixels between the centers of the two eyes is larger

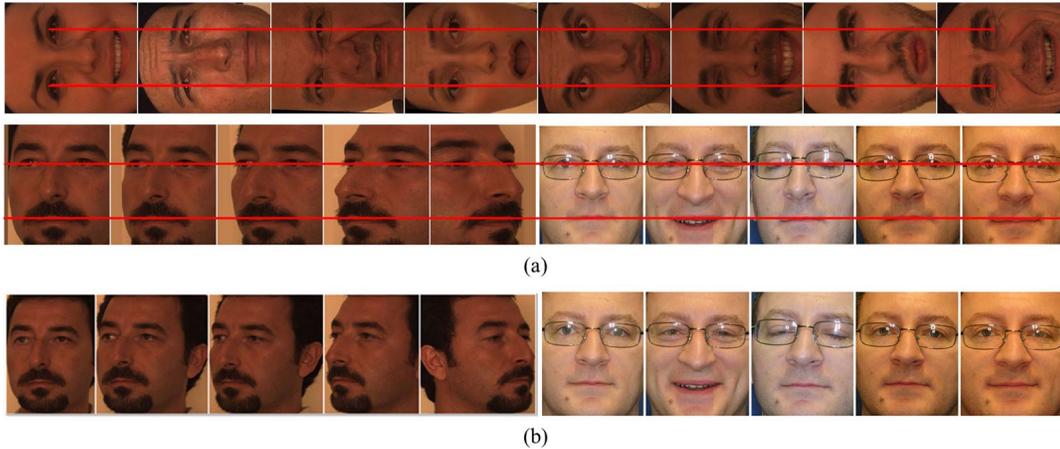


Fig. 7. Examples of preprocessed faces from the different datasets: (a) Facial images with different poses and expressions, which are aligned both horizontally and vertically for the alignment-required methods; (b) Facial images cropped for our proposed face-verification method, which is robust to alignment errors.

than 280, are selected. Then, these high-quality images are divided into five sessions according to the capture time.

Since our proposed method is robust to alignment errors, we simply crop the images to include the faces only. In order to further improve efficiency, we down-sample all the cropped facial images to a resolution of about 560×670 . The impact of resolution on keypoint matching was discussed in [25], which has shown that down-sampling within a certain range has a slight effect on the matching result. For feature-block matching, we partition each face image into 30×45 blocks uniformly, which is experimentally determined.

To compare our method with other standard face-verification methods wherein face alignment must be performed, all the face images are manually aligned according to the centers of the two eyes and the outer corners of the lips. All of the eyes and lips are aligned to a corresponding vertical and a horizontal line, as illustrated in Fig. 7. Then, the aligned faces are normalized to the same size, 560×670 . Some example face images used for verification are shown in Fig. 7. All of these aligned images are then down-sampled by a factor of 0.2 (i.e. to the size of 112×134), which can result in a performance that is better than or similar to using either the resolution 560×670 or down-sampled images corresponding to the factors set at 0.5 and 0.1.

Our proposed method is compared with the Eigenface method (PCA), the Gabor feature with PCA (Gabor+PCA), the LBP method [7] and the LBP feature with PCA (LBP+PCA). For the Gabor+PCA face-verification method, Gabor filters of eight orientations ($0, \pi/8, \dots, 7\pi/8$) and five scales ($\pi/2, \pi/2\sqrt{2}, \dots, \pi/8$) are employed to extract the features, which are concatenated and then normalized to zero mean and unit variance. Since the Gabor features are extremely huge, PCA is applied to reduce the feature dimensionality. To retain as much information as possible, $N - 1$ components are used, where N is the number of training samples. For LBP-based face verification, the $LBP_{(8,2)}^{u2}$ operator is used, and the images are divided into 7×7 non-overlapping windows. All the images are down-sampled to the size of 112×134 , which is similar to the image resolution used in [7]. For the

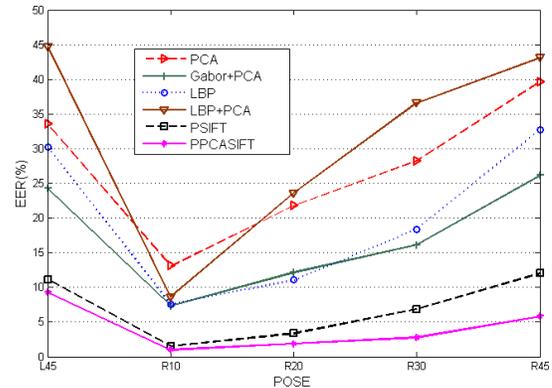


Fig. 8. EER of the different methods for face images under different poses.

LBP+PCA method, the LBP features from non-overlapping windows are concatenated to a long feature vector. Then, PCA is applied to reduce the feature dimensionality to $N - 1$. The similarity metric for the LBP method is the weighted Chi-square distance where the same weight matrix as [7] is used. For PCA, Gabor+PCA and LBP+PCA, similarity metric used is the l_2 distance.

B. Face Verification With Pose Variations

To evaluate the robustness of the different face-verification methods to pose variations, the 105 frontal-view faces from the Bosphorus dataset were selected to form the gallery set, while images of the 5 poses (R10, R20, R30, R45, L45) form 5 testing sets, respectively.

Fig. 8 and Fig. 9 show the EER results for each pose and the ROC curves under different poses, respectively. We can see that the performances of all the other methods degrade significantly when the pose variation becomes larger. Note that the performances of the LBP-based face-verification methods degrade significantly under large pose variations. This may be because the histograms representing the texture information about the frontal and non-frontal faces become more uncorrelated when the pose difference increases.

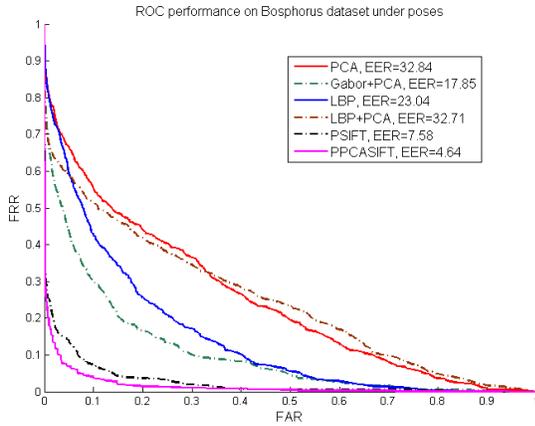


Fig. 9. ROC curves of the different methods for face images under all poses.

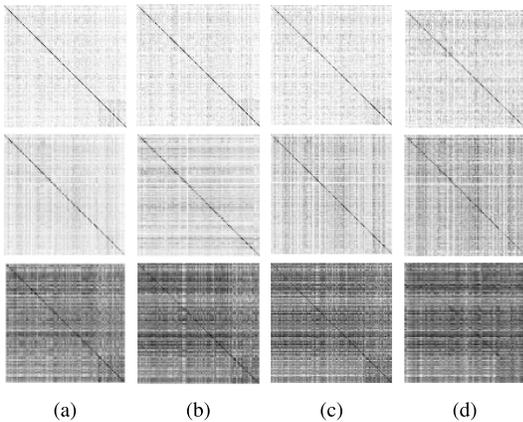


Fig. 10. Distance matrices of the PPCASIFT, PSIFT and Gabor+PCA methods (from top to bottom) under different poses: (a) R10, (b) R20, (c) R30, and (d) R45. PPCASIFT, which has its diagonal line the darkest, performs the best in distinguishing clients from impostors.

The LBP method using a constant weight matrix is not suitable for face images with different poses; i.e. the prior knowledge of pose for adaptive weights is necessary. In contrast, both the PPCASIFT- and PSIFT-based face-verification methods can maintain their performances with a lower EER under a large pose difference, such as 45 degrees. In particular, PPCASIFT achieves a slightly better performance than PSIFT, and has a feature dimension of 72 only. The result shows that PSIFT and PPCASIFT are robust to large pose variations.

We also transform the PSIFT and PPCASIFT similarity metrics into distance metrics by subtracting each similarity score or matching density from one, respectively. Fig. 10 shows the distance matrices of the three face-verification methods with the best performance under pose variations: these are PPCASIFT, PSIFT, and Gabor+PCA, respectively. This can provide a more intuitive way of illustrating their verification performances. Both the PPCASIFT and PSIFT methods can effectively distinguish clients from impostors, and they can achieve a better performance than the Gabor+PCA method, as their results show a much darker diagonal line than the Gabor+PCA method for all four poses. In addition, PPCASIFT performs better than PSIFT, especially under large pose variations such as R30 and R45.

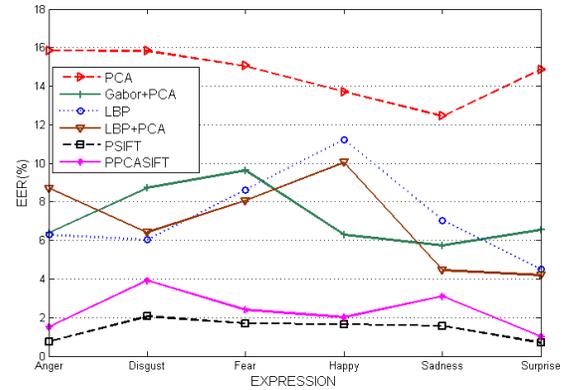


Fig. 11. EER of the different methods under different expressions.

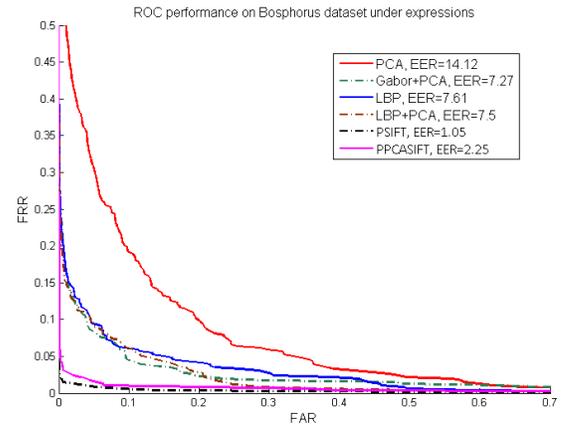


Fig. 12. ROC curves of the different methods under different expressions.

C. Face Verification Under Different Expressions

The verification of faces under facial-expression variations is another hot topic for real applications. We evaluate the robustness of our proposed method using images expressing 6 different emotions (anger, disgust, fear, happiness, sadness, and surprise) from the 105 subjects in the Bosphorus dataset. We compare our proposed method with the other methods in terms of the EER for each expression and the ROC curves under expressions, as shown in Fig. 11 and Fig. 12, respectively.

From the results, all the verification methods can achieve a better EER than the results in the previous section, since all the testing faces are frontal view. The Gabor+PCA, the LBP and the LBP+PCA methods are more effective than PCA. Both the PSIFT and the PPCASIFT methods outperform the other four methods, and achieve lower EERs in all cases. However, in this expression-variation case, the PPCASIFT method with local-matching density falls a little behind PSIFT. One reason for this may be that the PPCASIFT subspace is learned only from images with a neutral expression rather than with large expression variations.

D. Face Verification on Face Images Captured in Different Time Sessions

Our method relies on matching facial-skin regions, which involves the challenge of skin conditions changing with time.

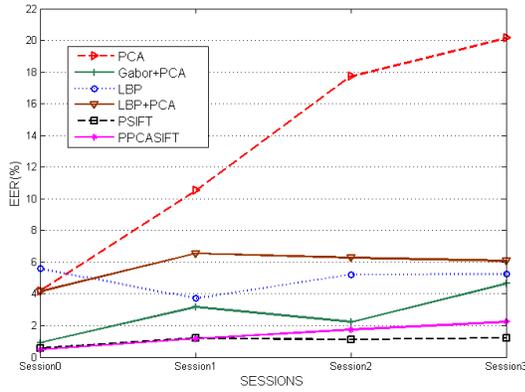


Fig. 13. EER of the different methods through different sessions.

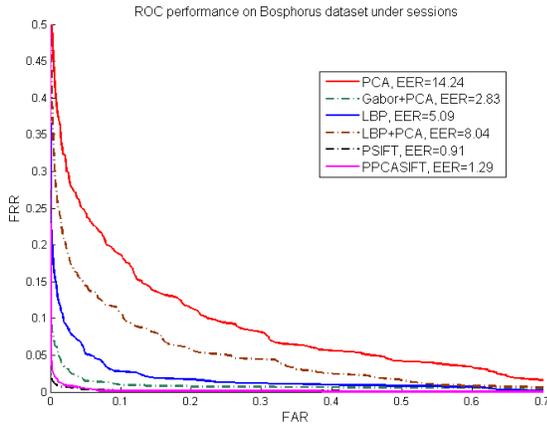


Fig. 14. ROC curves of the different methods through different sessions.

Therefore, in this section, we will evaluate the robustness of our proposed algorithm to face images captured at different times.

In this experiment, faces in the Multi-PIE dataset, which appear in all four sessions, are used. The longest time interval between the photos captured is 6 months. We select faces with a neutral expression in Session 0 to form the gallery set, while the faces captured in Session 0 with expressions, and those captured in the other three sessions, form the testing sets. The EERs and ROC curves of each face-verification method in sessions are shown in Fig. 13 and Fig. 14, respectively. The results show that our method still outperforms the other three methods for face images captured at different sessions. The superior performance of our facial-skin-based method may be due to the fact that, over time, facial appearances can change in ways other than just their skin condition. In addition, the geometric relations between the pores in a skin region should be very stable over time.

E. Face Verification Under Large Time Span and Different Expressions

In order to further examine the robustness of our proposed PPCASIFT and PSIFT methods, we use the FRGC v2.0 database for an extensive face-verification experiment. After eliminating the low-quality images in uncontrolled

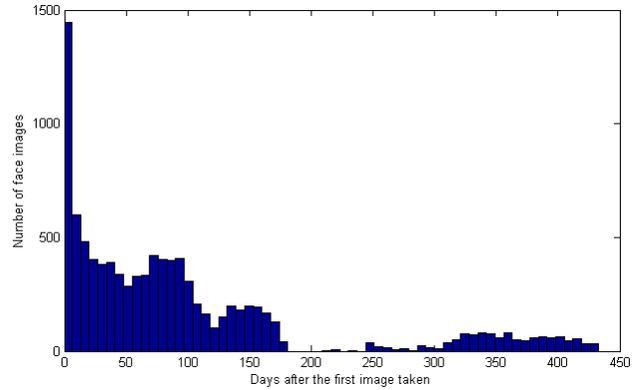


Fig. 15. Time span of the FRGC dataset.

TABLE I
EER(%) OF DIFFERENT FACE-VERIFICATION METHODS FOR ALL THE EXPERIMENTS

Variations	Equal Error Rate(%)					
	PCA	Gabor+PCA	LBP	LBP+PCA	PSIFT	PPCASIFT
P-R10	13.11	7.34	7.60	8.16	1.49	1.00
P-R20	21.79	12.13	11.06	23.6	3.37	1.90
P-R30	28.21	16.14	18.4	6.88	36.61	2.77
P-R45	39.61	26.13	32.76	43.14	12.01	5.82
P-L45	33.54	24.25	30.2	44.75	11.14	9.28
P-All	32.84	17.85	23.04	32.71	7.58	4.64
E-Anger	15.85	6.39	6.3	8.72	0.76	1.51
E-Disgust	15.83	8.74	6.03	6.42	2.06	3.92
E-Fear	15.05	9.63	8.63	8.08	1.7	2.4
E-Happy	13.71	6.3	11.23	10.06	1.65	2.02
E-Sadness	12.46	5.75	7.04	4.46	1.58	3.12
E-Surprise	14.86	6.55	4.5	4.21	0.7	1.01
E-All	14.12	7.27	7.61	7.5	1.05	2.25
T-Session0	4.17	0.91	5.59	5.44	0.55	0.47
T-Session1	10.51	3.18	3.71	8.56	1.2	1.18
T-Session2	17.73	2.22	5.22	8.28	1.11	1.74
T-Session3	20.13	4.64	5.24	8.09	1.21	2.24
T-All	14.24	2.83	5.09	8.04	0.91	1.29
FRGC-0-2W	15.6	4.21	8.28	8.02	2.01	3.3
FRGC-3-10W	27.11	9.67	12.87	12.2	4.25	6.75
FRGC-11-18W	33.63	12.05	16.23	15.54	6.32	8.21
FRGC-19-26W	28.94	9.62	14.2	12.76	5.6	7.7
FRGC-Aft26W	28.75	14.26	18.09	16.14	8.11	10.55
FRGC-All	27.57	10.02	14.12	13.16	5.51	7.36

environments (e.g. images taken in the wild with a resolution of less than 400×400), we use the remaining 9,844 images of 362 subjects for our verification task. According to the capture-time duration of each image, a new experiment protocol is proposed. The total time span is about 400 days (58 weeks), as shown in Fig. 15. We used one randomly selected frontal face of each subject, taken at the beginning of this time span, as the gallery image, and we divided the remaining images into five groups (0-2 weeks, 3-10 weeks, 11-18 weeks, 19-26 weeks, and more than 26 weeks) for testing. We also compared the EERs of the different face-verification methods. The results are tabulated in Table I, together with those of the previous three experiments. In general, these methods using pore-scale features can achieve much better performances when faces are under variations of pose, expression, and capture time. In particular, by extracting the PPCASIFT features, we can achieve greater efficiency while still maintaining a similar performance to PSIFT.

TABLE II
EER(%) OF THE DIFFERENT FACE-VERIFICATION METHODS WITH DIFFERENT ALIGNMENT ERRORS

σ of $(\Delta x, \Delta y)$	Equal Error Rate(%)					
	PCA	Gabor+PCA	LBP	LBP+PCA	PSIFT	PPCASIFT
0	14.12	7.27	7.61	7.50	0.77	1.94
10	15.23	7.35	7.89	7.96	0.73	2.20
20	19.49	9.38	13.88	11.26	0.87	2.12
30	21.86	14.06	21.98	14.38	0.79	2.59
40	24.90	16.97	25.22	17.79	0.63	2.30
50	29.13	22.97	31.06	21.59	0.97	2.60

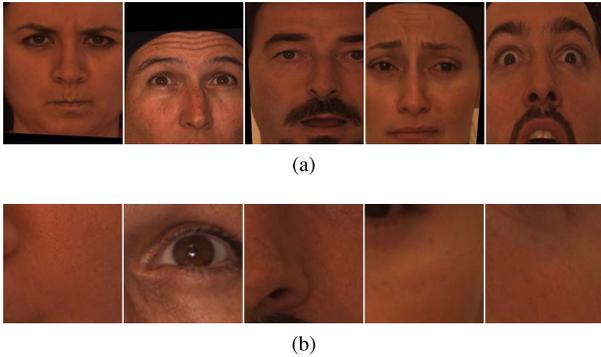


Fig. 16. Face images with a displacement vector added: (a) Samples of testing images when $\sigma = 50$, and (b) the corresponding cropped regions for the PSIFT detector used to estimate the adaptive threshold.

F. Robustness to Alignment Error

In this section, we will evaluate the performances of PSIFT and PPCASIFT in cases of different alignment errors. Those well-aligned frontal face images with neutral expression among the 105 subjects in the Bosphorus dataset are chosen to form the gallery set. The other 453 face images, expressing 6 different emotions (anger, disgust, fear, happiness, sadness, and surprise), form the testing set. All these images are normalized to size 560×670 . A random displacement vector $(\Delta x, \Delta y)$ is added to the location of each face in the testing set, where Δx and Δy are uncorrelated and normally distributed with zero mean and a standard deviation of σ . Fig. 16(a) shows some samples of the testing images when $\sigma = 50$, and Fig. 16(b) shows the corresponding cropped regions for the PSIFT detector used to estimate the adaptive threshold for pore keypoint detection. It is obvious that the estimated thresholds will also be distorted. For PCA, Gabor+PCA, LBP and LBP+PCA, these images are further down-sampled by a factor of 0.2. Six experiments were conducted with the testing images distorted by the displacement vector with six different σ values.

The EERs of PSIFT and PPCASIFT, as well as other face-verification algorithms, with different alignment errors are summarized in Table II. It can be seen that PSIFT and PPCASIFT are robust to different alignment errors, and perform well even when suffering from a large alignment error. Gabor+PCA, LBP and LBP+PCA work well only when there is a small or no alignment error; their performances drop significantly when the alignment error increases. Of these methods, PCA is sensitive to alignment errors even with small σ values.

VII. CONCLUSION AND DISCUSSION

In this paper, we have addressed the problem of HR face verification based on pore-scale facial features. The proposed method is robust to alignment errors and pose variations, while the gallery set requires only a single image per subject. The PSIFT and PPCASIFT features are highly distinctive, and PPCASIFT can efficiently reduce the computational time of the matching process to about 9% of that of PSIFT, while a similar performance level can be maintained. For each query in the feature-matching stage, PSIFT needs 1.45 seconds, while PPCASIFT needs only 0.13 seconds on an Intel i7 3.4GHz CPU with 8 threads and 8GB Ram PC under the MATLAB R2014a programming environment. These runtimes can be further reduced by using GPU parallel computing techniques. The runtime of the robust-fitting stage is less than 0.01 seconds.

Furthermore, the proposed parallel-block-aggregation and matching-density schemes can be applied to other image-analysis tasks such as object recognition, image annotation, since they provide an approach to transforming point matching into similarity measurement. Experimental results have shown that our method can achieve a superior performance under a range of variations, especially under large pose variations. To the best of our knowledge, this is the first work on HR face verification that uses pore-scale facial features and establishes such a large number of correspondences between faces. In addition, our proposed face-verification method can tackle pose, expression, and capture-time variations simultaneously.

In our future work, we will investigate the fusion of pore-scale facial features from HR images with larger-scale facial features from LR images. We will also study how to further improve the efficiency of the proposed method, and apply it to other important areas like face recognition and 3D face reconstruction.

REFERENCES

- [1] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Comput. Soc. Conf. CVPR*, Jun. 1991, pp. 586–591.
- [2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
- [3] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.
- [4] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775–779, Jul. 1997.
- [5] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.

- [6] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
- [7] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
- [8] T. Ahonen, E. Rahtu, V. Ojansivu, and J. Heikkilä, "Recognition of blurred faces using local phase quantization," in *Proc. 19th Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2008, pp. 1–4.
- [9] N.-S. Vu and A. Caplier, "Enhanced patterns of oriented edge magnitudes for face recognition and image matching," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1352–1365, Mar. 2012.
- [10] C. Ding, J. Choi, D. Tao, and L. S. Davis. (2014). "Multi-directional multi-level dual-cross patterns for robust face recognition." [Online]. Available: <http://arxiv.org/abs/1401.5311>
- [11] K. I. Kim, K. Jung, and H. J. Kim, "Face recognition using kernel principal component analysis," *IEEE Signal Process. Lett.*, vol. 9, no. 2, pp. 40–42, Feb. 2002.
- [12] M.-H. Yang, "Kernel eigenfaces vs. kernel Fisherfaces: Face recognition using kernel methods," in *Proc. 10th IEEE Int. Conf. Autom. Face Gesture Recognit.*, May 2002, pp. 210–215.
- [13] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in *Proc. IEEE Comput. Soc. Conf. CVPR*, Jun. 1994, pp. 84–91.
- [14] D. J. Beymer, "Face recognition under varying pose," in *Proc. IEEE Comput. Soc. Conf. CVPR*, Jun. 1994, pp. 756–761.
- [15] V. Blanz, S. Romdhani, and T. Vetter, "Face identification across different poses and illuminations with a 3D morphable model," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recognit.*, May 2002, pp. 192–197.
- [16] X. Chai, S. Shan, X. Chen, and W. Gao, "Locally linear regression for pose-invariant face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1716–1725, Jul. 2007.
- [17] C. D. Castillo and D. W. Jacobs, "Wide-baseline stereo for face recognition with large pose variation," in *Proc. IEEE Conf. CVPR*, Jun. 2011, pp. 537–544.
- [18] S. R. Arashloo and J. Kittler, "Energy normalization for pose-invariant face recognition based on MRF model image matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 6, pp. 1274–1280, Jun. 2011.
- [19] D. Lin and X. Tang, "Recognize high resolution faces: From macrocosm to microcosm," in *Proc. IEEE Comput. Soc. Conf. CVPR*, vol. 2, Jun. 2006, pp. 1355–1362.
- [20] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Proc. ECCV Int. Workshop Statist. Learn. Comput. Vis.*, vol. 1, 2004, pp. 1–22.
- [21] U. Park and A. K. Jain, "Face matching and retrieval using soft biometrics," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 3, pp. 406–415, Sep. 2010.
- [22] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.
- [23] N. Srinivas, G. Aggarwal, P. J. Flynn, and R. W. V. Bruegge, "Facial marks as biometric signatures to distinguish between identical twins," in *Proc. IEEE Comput. Soc. Conf. CVPR*, Jun. 2011, pp. 106–113.
- [24] A. K. Jain, S. Prabhakar, and S. Pankanti, "On the similarity of identical twin fingerprints," *Pattern Recognit.*, vol. 35, no. 11, pp. 2653–2663, 2002.
- [25] D. Li and K.-M. Lam, "Design and learn distinctive features from pore-scale facial keypoints," *Pattern Recognit.*, vol. 48, no. 3, pp. 732–745, 2015.
- [26] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. Cambridge, MA, USA: MIT Press, 2001.
- [27] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [28] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proc. IEEE Comput. Soc. Conf. CVPR*, vol. 2, Jun./Jul. 2004, pp. II-506–II-513.
- [29] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.
- [30] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [31] A. Savran *et al.*, "Bosphorus database for 3D face analysis," in *Biometrics Identity Management*. Berlin, Germany: Springer-Verlag, 2008, pp. 47–56.
- [32] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-PIE," in *Proc. 8th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, Sep. 2008, pp. 1–8.
- [33] P. J. Phillips *et al.*, "Overview of the face recognition grand challenge," in *Proc. IEEE Comput. Soc. Conf. CVPR*, vol. 1, Jun. 2005, pp. 947–954.



Dong Li received the B.Eng. degree in computer science and the M.Eng. degree in computer science from Tianjin University, Tianjin, in 2006 and 2009, respectively, and the Ph.D. degree from The Hong Kong Polytechnic University, Hong Kong, in 2014. He is currently an Assistant Professor with the Guangdong University of Technology. His research interests include topics in the fields of computer vision, pattern recognition, and image processing, such as feature matching, face recognition, and color correction.



Huiling Zhou received the B.Eng. degree in communication engineering from the University of Electronic Science and Technology of China, Chengdu, in 2008. She is currently pursuing the Ph.D. degree with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong. Her research interests include image processing, computer vision, and pattern recognition.



Kin-Man Lam received the Associate (Hons.) degree in electronic engineering from The Hong Kong Polytechnic University, in 1986, the M.Sc. degree in communication engineering from the Department of Electrical Engineering, Imperial College of Science, Technology and Medicine, London, U.K., in 1987, and the Ph.D. degree from the Department of Electrical Engineering, University of Sydney, Sydney, Australia, in 1996.

He was a Lecturer with the Department of Electronic Engineering, The Hong Kong Polytechnic University, from 1990 to 1993. He joined the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, as an Assistant Professor, in 1996, became an Associate Professor in 1999, and has been a Professor since 2010. He has been a member of the Organizing Committee and Program Committee of many international conferences. He was the Chairman of the IEEE Hong Kong Chapter of Signal Processing from 2006 to 2008, and was the Director-Student Services of the IEEE Signal Processing Society from 2012 to 2014. He also served as an Associate Editor of the IEEE TRANSACTION ON IMAGE PROCESSING from 2009 to 2013.

Dr. Lam is currently the VP-Member Relations and Development of the Asia-Pacific Signal and Information Processing Association (APSIPA), and the Director-Membership Services of the IEEE Signal Processing Society. He serves as an Associate Editor of the *APSIPA Transaction on Signal and Information Processing*, *Digital Signal Processing*, and the *EURASIP International Journal on Image and Video Processing*. He is also an Editor of *HKIE Transactions* and an Area Editor of the *IEEE Signal Processing Magazine*. His current research interests include human face recognition, image and video processing, and computer vision.