Partial Face Recognition: Alignment-Free Approach

Shengcai Liao, Anil K. Jain, *Fellow, IEEE* and Stan Z. Li, *Fellow, IEEE*

Abstract—Numerous methods have been developed for holistic face recognition with impressive performance. However, few studies have tackled how to recognize an arbitrary patch of a face image. Partial faces frequently appear in unconstrained scenarios, with images captured by surveillance cameras or handheld devices (e.g. mobile phones) in particular. In this paper, we propose a general partial face recognition approach that does not require face alignment by eye coordinates or any other fiducial points. We develop an alignment-free face representation method based on Multi-Keypoint Descriptors (MKD), where the descriptor size of a face is determined by the actual content of the image. In this way, any probe face image, holistic or partial, can be sparsely represented by a large dictionary of gallery descriptors. A new keypoint descriptor called Gabor Ternary Pattern (GTP) is also developed for robust and discriminative face recognition. Experimental results are reported on four public domain face databases (FRGCv2.0, AR, LFW, and PubFig) under both the open-set identification and verification scenarios. Comparisons with two leading commercial face recognition SDKs (PittPatt and FaceVACS) and two baseline algorithms (PCA+LDA and LBP) show that the proposed method, overall, is superior in recognizing both holistic and partial faces without requiring alignment.

Index Terms—Partial Face Recognition, Alignment Free, Keypoint Descriptor, Sparse Representation, Open-Set Identification.

✦

1 INTRODUCTION

FACE recognition (FR) is the problem of verifying or identifying a face from its image. It has received substantial attention over the last three decades due to its value both in understanding how FR process works in humans as well as in addressing many challenging real-world applications, including deduplication of identity documents (e.g. passport, driver license), access control and video surveillance. The performance of automatic FR systems has advanced significantly. While face recognition in controlled conditions (frontal face of cooperative users and controlled indoor illumination) has already achieved impressive performance over large-scale galleries, as indicated in a recent IEEE T-PAMI special issue on real-word face recognition [1], there still exist many challenges for face recognition in uncontrolled environments, such as partial occlusions, large pose variations, and extreme ambient illumination.
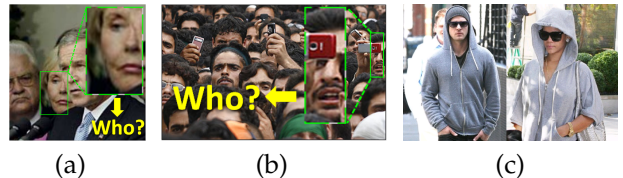
- *Anil K. Jain is with the Dept. of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824 USA. He is also with the Dept. of Brain & Cognitive Engineering, Korea University, Anamdong, Seongbukgu, Seoul 136-713, Republic of Korea.*
  E-mail: jain@cse.msu.edu
- *Shengcai Liao was with the Dept. of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824 USA. He is now with the National Laboratory of Pattern Recognition and the Center for Biometrics and Security Research, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.*
  E-mail: scliao@cbsr.ia.ac.cn
- *Stan Z. Li is with the National Laboratory of Pattern Recognition and the Center for Biometrics and Security Research, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.*
  E-mail: szli@cbsr.ia.ac.cn

Fig. 1. Partial face examples. (a) Partial faces in the LFW database [2]. (b) Partial faces in a crowd [1]. (c) Occluded faces by hooded sweatshirt and sunglasses [2].

Typical applications of face recognition in uncontrolled environments include recognition of individuals in video surveillance frames and images captured by handheld devices (e.g. mobile phones), where a face may be captured in arbitrary pose without user cooperation and knowledge. In such scenarios, it is quite likely that the captured image contains only a partial face. Table 1 lists a categorization of partial face images, and some further illustrations are given in Fig. 1. We call the resulting problem a Partial Face Recognition (PFR) problem, so as to differentiate it from the holistic face recognition problem.

Commercial off-the-shelf (COTS) face recognition systems are not able to handle the general PFR problem since they need to align faces by facial landmarks that may be occluded. For example, FaceVACS [3] requires localization of the two eyes, and PittPatt [4] detects several predefined landmarks for face

1. http://www.textually.org/picturephoning/archives/2008/09/021247.htm

2. http://www.howtovanish.com/2010/01/avoid-nosy-surveillance-cameras/

TABLE 1
A categorization of partial face images

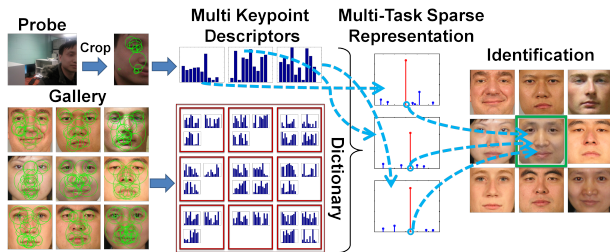| Scenario	External occlusion	Self occlusion	Facial accessories	Limited field of view (FOV)	Extreme illumination	Sensor saturation
Examples	occlusion by other objects	non-frontal pose	hat, sunglasses, scarf, mask	partially out of camera's FOV	gloomy or highlighted facial area	underexposure or overexposure
Image						

Fig. 2. Proposed partial face recognition approach.

alignment. Therefore, research in PFR is important to advance the state of the art in face recognition and enlarge the application domain.

Law enforcement agencies are also in urgent need of a system capable of recognizing partial faces. First, a PFR system will enable them to identify a suspect in a crowd (e.g. Fig. 1b) by matching a partial face captured by, say, a mobile phone to a watch list through a wireless link in real time. Second, given a photo of a certain unlawful event, PFR is needed to recognize the identity of a suspect based on a partial face. As an example, while automatic face recognition resulted in many arrests in the 2011 London riots [5], many suspects in partial face images were not recognized by COTS FR systems [6].

1.1 Proposed Method

In this paper, we present a general formulation of the partial face recognition problem. We require neither face alignment nor the presence of the eyes or any other facial component in the image. Further, we do not know a priori whether the input face is holistic or partial. We provide a general matching solution to accommodate all types of partial faces listed in Table 1. Our approach is based on a Multi-Keypoint Descriptor (MKD) representation for both the gallery dictionary and the probe image. Multi-task sparse representation is learned for each probe face, and the Sparse Representation based Classification (SRC) approach [7] is applied for face recognition. We call the proposed method MKD-SRC. The flowchart of the proposed method is shown in Fig. 2.

The novelty of the proposed approach includes: (i) a

general partial face recognition approach without requiring face alignment; (ii) the MKD-SRC framework that works for both holistic faces and partial faces, and outperforms SRC [7] in addressing the one-sample-per-class problem; (iii) a new keypoint descriptor, called Gabor Ternary Pattern (GTP) which outperforms the Scale Invariant Feature Transform (SIFT) [8] descriptor; and (iv) a fast atom filtering strategy for MKD-SRC to address large-scale face recognition (with 10,000 gallery images).

This paper is built upon our preliminary work reported in [9]. The main differences are summarized as follows. (i) While the method in [9] used the SIFT descriptor, we now propose a new keypoint descriptor (GTP) which outperforms SIFT. (ii) We address the one-sample-per-class problem in large-scale open-set identification setting for PFR, and show that the proposed MKD-SRC method performs better than two leading commercial face matchers, FaceVACS and PittPatt, on the FRGCv2.0, AR, and PubFig databases. (iii) We extend the MKD-SRC method for partial face verification, which is effective on the LFW database.

1.2 Literature Review

The most popular approach for face alignment is to first detect the two eyes and then normalize the face geometrically. Other popular face alignment methods include Active Shape Model (ASM) [10] and Active Appearance Model (AAM) [11], which depend on localizing a certain fixed number (typically 68) of landmarks on holistic face. In [12], a sparse representation based alignment method was proposed in controlled scenarios. However, all these alignment methods would fail for face images with unknown missing portions of the face.

Occluded and non-frontal faces are the most frequently encountered partial faces in practice. While several approaches have dealt with FR under occlusion [7], [12], [13], [14], [15], [16], [17], they all require face alignment. Ekenel and Stiefelhagen showed that under occlusion, face alignment is critical to recognition performance [18].

Non-frontal face recognition has also attracted significant attention, including multi-view [19], [20] and

cross-view [21], [22], [23], [24], [25], [26] face recognition. Multi-view face recognition requires that the gallery contains a large number of poses for each subject, which is difficult to satisfy in practice. For cross-view FR, most approaches apply 2D or 3D appearance models to synthesize face images in specific views. A critical step in these approaches is to localize a certain fixed number of representative facial landmarks and establish correspondences between two images in different views. As a result, the images are expected to have visible anchor points irrespective of the view.

For partial faces resulting from a limited field of view, Yang et al. [27] and Yi et al. [28] proposed an automatic partial face alignment and matching approach. However, their approach requires high resolution images (with an inter-pupil distance of more than 128 pixels) with good skin texture, and it is not applicable to pose variations.

Some FR approaches only require face sub-images as input, such as eye [29], nose [29], one half (left or right portion) of the face [30], or the periocular region [31]. Again, these methods require the presence of certain facial components and pre-alignment.

Instead of holistic representation, some face recognition approaches have adopted parts-based representations to deal with occlusion and pose variations. A simple way is to divide the aligned face image into several sub-regions [32], [33], [34], [35], match each sub-region and then fuse the matching results. Alternatively, one could detect several predefined components (such as eye, nose, and mouth), and then recognize the face by fusing the matching results for the components [20], [26], [36]. Sanderson and Paliwal [37] proposed a parts-based face representation based on block 2D DCT feature extraction and GMM modeling, and obtained better performance than holistic methods like PCA. Cardinaux et al. [38] further showed that, part-based GMM representation is robust to misalignment, which is able to perform face verification on automatically detected faces without perfect alignment. Following these two studies, Lucey and Chen [39] applied relevance adaptation to improve the GMM parts based face representation, and Cardinaux et al. [40] applied HMM instead of GMM, with improved performance. However, both sub-region and component based approaches may fail when a partial face cannot be detected or some facial components are occluded.

Several existing approaches are closely related to the proposed MKD-SRC algorithm. Wright et al. introduced the well-known SRC approach for face recognition [7], achieving robust performance in the presence of illumination variations and occlusions. However, the SRC approach requires well aligned face images. Wagner et al. [12] further improved the SRC method by assuming that face registration might have errors. LBP-SRC [41] is another approach that applies SRC in conjunction with Local Binary Pattern (LBP) [33] features after alignment for face recognition. However, as indicated in [12], SRC is targeted for the access control scenario, where many controlled face images can be captured to construct the gallery.

The proposed MKD-SRC approach uses a different feature representation than the above SRC based approaches. Since both SRC and LBP-SRC require aligned faces and use a single fixed-size feature vector (e.g. concatenated image pixels or LBP histograms) to represent a face, each column of their corresponding dictionary is related to one gallery image. However, in such a scheme a partial face might be difficult to align and represent due to unknown missing facial regions. In contrast, MKD-SRC uses a variable-size description, so each face is represented by a set of descriptors. The MKD dictionary is composed of a large number of gallery descriptors, making it possible to sparsely represent descriptors from a probe image, irrespective of whether it represents a holistic or partial face. Furthermore, SRC requires a sufficient number of training samples covering all possible illumination variations for each subject, which limits its applicability. In contrast, MKD-SRC performs satisfactorily in scenarios where only one training sample per class is available.

The bag-of-words (BoW) representation in the field of visual object categorization [42] is another representation scheme that has been applied to face recognition [43], [44]. However, the underlying assumption in the BoW representation is that the object image should not be significantly occluded, otherwise the descriptor histogram of a partial view will be quite different from that of the whole view. For this reason, the BoW representation is not suitable for PFR.

A number of papers have been published on SIFT-based face recognition [45], [46], [47], [48]. However, all of them were applied on pre-aligned face images. While SIFT matching is fast, it treats each image pair independently and therefore does not utilize collaborative representation [49] of different subjects in the gallery set. Since many local facial patches may look similar, it is possible that the SIFT matching would find more matches for an impostor pair than a genuine pair [9]. To alleviate this, the proposed MKD-SRC approach performs keypoint matching via sparse representation of all gallery images to select the best match automatically.

Characteristics of MKD-SRC in comparison with various existing approaches are summarized in Table 2. The remainder of this paper is organized as follows. In Section 2 we describe the proposed alignment-free partial face representation method. In Section 3 we introduce the MKD-SRC algorithm. Extensive experiments are demonstrated in Section 4, and finally we conclude this work in Section 5.

TABLE 2
Comparison of available partial face recognition approaches with the proposed approach

Approach	Scenario	Image requirement	Database used	#Subjects
Subspace [13], [14], [15]; SRC [7], [12]; SVM [16], [17]; Part-based fusion [26], [32], [33], [34], [35], [36]; Single component [29], [30], [31]	Occlusion	Aligned & cropped frontal faces	AT&T, ORL, AR, Yale B, FERET, FRGC, Multi PIE	$\leq 1,196$
Multi-view [19], [20]; Cross-view [21], [22], [23], [24], [25], [26]	Arbitrary pose	Alignment via facial landmarks	FERET, PIE	≤ 250
Skin texture [50]	Limited FOV	Frontal face	MBGC	114
Proposed method (MKD-SRC)	Occlusion, arbitrary pose, limited FOV	Alignment free	FRGC, AR, LFW, PubFig	$> 20,000$

2 ALIGNMENT FREE PARTIAL FACE REPRESENTATION

While face alignment is often based on detection of facial landmarks, in the scenario of partial faces, there is no guarantee that commonly used landmarks are visible in the image. Therefore, when considering partial faces a canonical frame is not always available for feature extraction. Furthermore, most available approaches, either holistic (e.g. PCA and LDA), or local (e.g. Gabor and LBP), use a fixed-length representation for each face. A fixed-length representation assumes that the face image is aligned and cropped to a predefined size, followed by either concatenating the pixel values or extracting local feature vectors of predetermined dimensionality at fixed locations. However, for partial faces, it is not possible to extract a fixed-length descriptor due to missing facial portions and the difficulty of alignment.

Given the characteristics of partial faces, we establish two principles for partial face representation: (i) alignment free representation, and (ii) variable-length description. The length or size of representation should be determined by the specific content of the input (partial) face. A holistic face generally will have a larger descriptor size than a partial face.

To satisfy these two principles, we adopt a Multi-Keypoint Descriptor (MKD) based representation. A number of salient facial keypoints, along with a descriptor for each keypoint are automatically extracted without pre-alignment. For a face image, holistic or partial, the set of all descriptors associated with the extracted keypoints are used to represent the face. A desirable property of keypoint descriptors is their robust repeatability [8], [51], which means that descriptors from a partial face should be similar to the descriptor set of the corresponding holistic face. Note that the proposed MKD representation is different from the BoW representation. We neither learn visual words (clusters) nor do we compute histograms of cluster centers. In the following we describe the proposed alignment free representation.

2.1 Affine Invariant Keypoint Detection

The SIFT detector proposed by Lowe [8] is one of the most popular keypoint detector. The SIFT keypoints
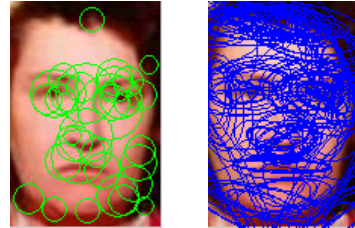

Fig. 3. Comparison of keypoint detection by SIFT and CanAff. Left: 37 keypoints are detected by SIFT. Right: 571 keypoints (only the first 150 keypoints are shown) are detected by CanAff.

have a robust repeatability property against image translation, rotation, and scaling. This property is important for PFR because the faces we are matching are not aligned. However, SIFT keypoint regions are not invariant under affine transformation. Another drawback of SIFT for face image is that it provides a limited number of keypoints because SIFT only seeks blob-like structures. Since most faces look generally similar to each other, a limited number of keypoints may not be sufficient to discriminate between them. This limits its applicability to face recognition.

In this paper, we consider a scale invariant Canny edge [52] based interest point detector proposed in [53], and an affine invariant shape adaptation technique proposed in [54]. We denote the resulting detector as CanAff[3]. The edge based detector finds many more keypoints than the SIFT detector for face images, since there are more edges than blobs on a face (See Fig. 3 for an example). The affine invariant shape adaptation makes image matching more robust to viewpoint changes, which is desired in face recognition with pose variations.

The CanAff detector first extracts edges with a multi-scale Canny edge detector [52], followed by a scale invariant local neighborhood for each edge point. The characteristic size of the local neighborhood is detected by seeking extremum in the local responses

---

3. We have also tried some other affine invariant keypoint detection methods, like the Harris Affine and Hessian Affine detectors proposed in [55] and the MSER detector proposed in [56], and found that the CanAff detector is the best one for our partial face recognition task, because all other detectors provide a limited number of keypoints in a face image.

to the scale normalized LoG operator [51]. Next, the detected local region is adapted to be an anisotropic affine-invariant shape as in [54], where the keypoint neighborhood is iteratively adapted using the second moment matrix information. At convergence, each detected affine invariant region is enclosed with an ellipse $\mathbf{x}^T \mathbf{M} \mathbf{x} = 1$, where the ellipse parameters in $\mathbf{M} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ are estimated from the affine-invariant neighborhood.

All detected ellipses are geometrically normalized to circles by an affine transformation $\mathbf{x}' = \mathbf{M}^{1/2} \mathbf{x}$, and the resulting regions are further cropped and scaled to 40×40 pixels (see Fig. 4). We apply the CanAff detector using the software in [57], with the "-sedgelap -noangle" option and default parameters in all our experiments. We do not apply rotation adaption because most face images are up-right [4], and sometimes a falsely detected dominant orientation [8] will, on the contrary, degrade matching accuracy. To compensate for illumination variations, we normalize the pixel values within each region to be in [0,1] by clipped Z-Score normalization. The idea is, to linearly stretch the pixel values into [0,1] by mapping $(\mu - 3\sigma)$ to 0, and $(\mu + 3\sigma)$ to 1; all values outside [0,1] are clipped.

2.2 Gabor Ternary Pattern Descriptor

Once the detected regions are normalized to a fixed size, a local descriptor is constructed within each region as follows. We first apply the Gabor filter [23], [58] to each image patch. We use Gabor filters because they provide good perception of local image structures and they are robust to illumination variations [23], [58]. The Gabor kernels are defined as

$$\psi_{\mu,\nu}(x,y) = \frac{\|\mathbf{k}_{\mu,\nu}\|^2}{\sigma^2} \exp\left(-\frac{\|\mathbf{k}_{\mu,\nu}\|^2 \|\mathbf{z}\|^2}{2\sigma^2}\right) \times \left[\exp(i\mathbf{k}_{\mu,\nu}^T \mathbf{z}) - \exp(-\frac{\sigma^2}{2})\right], \quad (1)$$

where μ and ν define the orientation and scale of the Gabor kernels, respectively, $\mathbf{z} = (x,y)^T$, and the wave vector $\mathbf{k}_{\mu,\nu}$ is defined as

$$\mathbf{k}_{\mu,\nu} = (k_\nu \cos\phi_\mu, k_\nu \sin\phi_\mu)^T, \quad (2)$$

with $k_\nu = k_{max}/f^\nu, k_{max} = \pi/2, f = \sqrt{2}$, and $\phi_\mu = \pi\mu/8$ [23], [58]. Due to the relatively small region size (40×40 pixels) we process, Gabor kernels at a single scale ($\nu = 0$) and four orientations ($\mu \in \{0, 2, 4, 6\}$, corresponding to 0°, 45°, 90°, and 135°) with $\sigma = 1$ are used. Furthermore, we only use the odd Gabor kernels (imaginary part), which are sensitive to edges and their locations. These four Gabor kernels are able

4. By "Up-right" we mean that while the face can have many possible rotations, it will not appear in the image with a mouth on the top. This is usually the case in most face recognition applications, including surveillance scenarios.
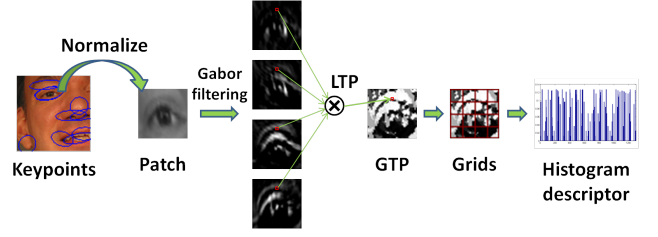

Fig. 4. Major components of the GTP descriptor.

to discriminate local details in the face image. Fig. 4 shows four Gabor filtered images for a local patch. These four response images emphasize edges in four different orientations (0°, 45°, 90°, and 135°).

For each pixel (x,y) in the normalized keypoint region, there are four Gabor filter responses as follows

$$f_i(x,y) = G_i(x,y) * I(x,y), \quad i = 0,1,2,3, \quad (3)$$

where $G_i = imag(\psi_{2i,0})$ is the ith odd Gabor kernel, and $*$ is the convolution operator. The responses of the four filters are combined as a ternary pattern [59]

$$GTP_t(x,y) = \sum_{i=0}^{3} 3^i \left[(f_i(x,y) < -t) + 2(f_i(x,y) > t)\right], \quad (4)$$

where t is a small positive threshold (a value of 0.03 is used in our experiments). We call this local descriptor the Gabor Ternary Pattern (GTP). It encodes local structures from the responses of odd Gabor filters in four different orientations. The local ternary pattern provides a discriminative encoding of the four Gabor filters. Also, these encodings are insensitive to image noise and illumination variation, because the quantization range is tolerant to these corruptions. There are a total of $3^4 = 81$ different GTP patterns. Fig. 4 demonstrates that four corresponding pixels at the same location in the four Gabor response images form a GTP pattern. Next, the 40×40 region is divided into $4 \times 4 = 16$ sub grid cells, each of size 10×10 pixels. A histogram of GTPs is calculated in each grid cell, and all histograms are concatenated to form a 1296-dimensional ($4 \times 4 \times 81$) feature vector. To reduce the influence of outliers in the histogram, the feature vector is first normalized to unit length, followed by a sigmoid function, $\tanh(ax)$, where a is a constant (a=20 in this paper), to suppress extreme values, and a renormalization step. Finally, we apply PCA to reduce the feature dimensionality to M ($M = 128$ in this paper). The major components in computing the GTP descriptor are shown in Fig. 4. Compared to SIFT, GTP performs better for face recognition in uncontrolled scenarios, as will be seen in the experiment section.

3 MKD-SRC

Wright et al. [7] showed that representing a probe image by a sparse linear combination of gallery images is very effective for classification. They call the resulting

algorithm SRC. In this paper, we propose to apply SRC on a large dictionary of keypoint descriptors instead of applying it directly to raw face image pixels; this is the key to the proposed alignment-free partial face recognition approach.

3.1 Gallery Dictionary Construction

The gallery dictionary is constructed as follows. First, an MKD representation (as introduced above) is constructed for each image. Suppose k_c keypoints, say, $\mathbf{p}_{c_1}, \mathbf{p}_{c_2}, \cdots, \mathbf{p}_{c_{k_c}}$, are detected for class (subject) c in the gallery. If class c has multiple face images in the gallery, we simply pool the keypoints extracted from all of them. The corresponding k_c GTP descriptors are denoted by $\mathbf{d}_{c_1}, \mathbf{d}_{c_2}, \cdots, \mathbf{d}_{c_{k_c}}$, where each descriptor is an M-dimensional vector (in our case, $M = 128$). Let

$$\mathbf{D}_c = (\mathbf{d}_{c_1}, \mathbf{d}_{c_2}, \cdots, \mathbf{d}_{c_{k_c}}). \qquad (5)$$

This way the descriptors from the same class form a sub-dictionary of size $M \times k_c$ representing class c. A gallery dictionary for all the C classes is built as

$$\mathbf{D} = (\mathbf{D}_1, \mathbf{D}_2, \cdots, \mathbf{D}_C). \qquad (6)$$

Note that $\mathbf{D}$ has a total of $K = \sum_{c=1}^{C} k_c$ descriptors, resulting in an $M \times K$ dictionary.

Typically, K is very large (e.g. over 1 million), which makes $\mathbf{D}$ an overcomplete description space of the C classes. Therefore, any descriptor from the C classes can be linearly represented in terms of the dictionary $\mathbf{D}$. According to the theory of compressed sensing (CS), a sparse solution is possible for an overcomplete dictionary [60]; therefore, we can express any descriptor from a probe image by a sparse linear combination of the dictionary $\mathbf{D}$.

3.2 Multi-Task Sparse Representation

Given a probe face image with n descriptors

$$\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n), \qquad (7)$$

the sparse representation problem is formulated as

$$\hat{\mathbf{X}} = arg \min_{\mathbf{X}} \sum_{i=1}^{n} \|\mathbf{x}_i\|_0, \ s.t. \ \mathbf{Y} = \mathbf{DX}, \qquad (8)$$

where $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) \in R^{K \times n}$ is the sparse coefficient matrix, and $\| \cdot \|_0$ denotes the ℓ_0 norm of a vector, i.e., $\|\mathbf{x}\|_0 = \sum_i I(x_i \neq 0)$, where $I(s) = 1$ if the statement s is true; otherwise 0. However, the solution to this problem is NP-hard. Based on the results from compressed sensing [61], sparse signals can be recovered with a high probability via the ℓ_1-minimization. Therefore, we solve the following ℓ_1-minimization problem[5] instead of Eq. (8)

$$\hat{\mathbf{X}} = arg \min_{\mathbf{X}} \sum_{i=1}^{n} \|\mathbf{x}_i\|_1, \ s.t. \ \mathbf{Y} = \mathbf{DX}, \qquad (9)$$

5. Some recent studies have also shown advantages of ℓ_2-norm constraint instead of ℓ_1, including the computational efficiency and stability [49], [62].

where $\| \cdot \|_1$ denotes the ℓ_1 norm defined as $\|x\|_1 = \sum_i |x_i|$. This is a multi-task problem since both $\mathbf{X}$ and $\mathbf{Y}$ have multiple columns. Equivalently, we can solve the following set of n ℓ_1-minimization problems, one for each probe descriptor y_i

$$\hat{\mathbf{x}}_i = arg \min_{\mathbf{x}_i} \|\mathbf{x}_i\|_1, \ s.t. \ \mathbf{y}_i = \mathbf{Dx}_i, \ i = 1, 2, \cdots, n. \qquad (10)$$

A number of efficient fast ℓ_1 minimization algorithms can be used to solve Eq. (10), including the ℓ_1 Homotopy method [63]. Since the n ℓ_1-minimization problems in Eq. (10) are independent, it is straightforward to accelerate the algorithm via parallel computation.

Inspired by [7], we adopt the following multi-task SRC to determine the identity of the probe image.

$$\min_c r_c(\mathbf{Y}) = \frac{1}{n} \sum_{i=1}^{n} \|\mathbf{y}_i - \mathbf{D}_c \delta_c(\hat{\mathbf{x}}_i)\|_2^2, \qquad (11)$$

where $\delta_c(\cdot)$ is a function which selects only the coefficients corresponding to class c. Eq. (11) applies a sum fusion among reconstruction residuals of the n descriptors with respect to each class, and determines the identity based on the least residual. Therefore, an unknown partial face in the probe can be recognized by computing Eqs. (10) and (11). The resulting algorithm is what we call MKD-SRC, which does not need face alignment. The flowchart of the MKD-SRC algorithm is shown in Fig. 2.

3.3 Fast Filtering

In practice, the size (K) of the dictionary $\mathbf{D}$ can be of the order of millions, making it difficult to solve Eq. (10). Therefore, we adopt a fast approximate solution. For each probe descriptor $\mathbf{y}_i$, we first compute the following linear correlation coefficients between $\mathbf{y}_i$ and all the descriptors in the dictionary $\mathbf{D}$

$$\mathbf{c}_i = \mathbf{D}^T \mathbf{y}_i, \ i = 1, 2, \cdots, n. \qquad (12)$$

Then for each $\mathbf{y}_i$, we keep only L $(L \ll K)$ descriptors according to the top L largest values of $\mathbf{c}_i$, resulting in a small sub-dictionary $\mathbf{D}_{M \times L}^{(i)}$. Next, $\mathbf{D}$ is replaced by $\mathbf{D}^{(i)}$ in Eq. (10), and Eq. (11) is adjusted accordingly.

We set $L = 100$ in our algorithm. According to our previous finding [9], this approximate solution speeds up the computation, with no significant degradation in recognition performance. According to Eq. (12) and the selection of top L elements (note that this can be done in $O(K)$ by the Introselect algorithm [64]), the computation time of the filtering step scales linearly with respect to K (the number of gallery keypoints). Thus, the algorithm scales almost linearly with respect to the gallery size for each probe image (considering an average number of keypoints per image). The overall MKD-SRC algorithm is outlined in Algorithm 1. The parameter values used in this paper are summarized in Table 3. They were fixed for all the experiments reported in the paper.

Algorithm 1: The MKD-SRC Algorithm

Input: Gallery images of C classes; probe image I; parameter L.

Output: Identity c of the probe image I.

1 **Enrollment:** Extract multi-keypoint descriptors (GTP) from each gallery image and build the Dictionary $\mathbf{D} = (\mathbf{D}_1, \mathbf{D}_2, \cdots, \mathbf{D}_C) \in R^{M \times K}$.

2 **Recognition:**

3 Extract MKDs from the probe image: $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n) \in R^{M \times n}$;

4 **for** $i = 1$ **to** n **do**

5   Compute top L descriptors from Eq. (12), resulting in a sub-dictionary $\mathbf{D}_{M \times L}^{(i)}$;

6   Solve Eq. (10) with $\mathbf{D}_{M \times L}^{(i)}$;

7 **end**

8 Solve Eq. (11) to determine the identity c;

TABLE 3
Parameter Values

Parameter	t	a	M	L
Value	0.03	20	128	100

3.4 MKD-SRC for Partial Face Verification

Given a gallery set, the residual defined in Eq. (11) can be used as the dissimilarity score for face identification (including open-set identification). However, while the SRC algorithm was originally proposed for face identification, little work has been done for SRC based face verification. In this paper, we propose a simple extension of the MKD-SRC algorithm for face verification. The face verification task is to judge whether a given pair of face images, say, I and J, belong to the same subject or not. For this task, we use a set of background face images together with image I as a virtual gallery set, and the other input face image J as the probe. Note that the set of background face images does not contain the same subject as either of the two input face images. Then the MKD-SRC algorithm is applied, and the verification score is defined as $1 - r_c$, where r_c is defined in Eq. (11), and c is the class for image I. To make the verification score a symmetric function of I and J, we also put J in the gallery set and use I as the probe, and the average score is computed as the final score.

To handle pose variations, mirrored face images have been used to improve FR performance in different views [65]. Motivated by this, the input face image, I, is horizontally mirrored as I', then both I and I' are put in the virtual gallery set. In this way, there would be more chance for a left profile face image to match with the corresponding right one.

4 EXPERIMENTS

The performance of the alignment-free MKD-SRC partial face matching algorithm has been evaluated on four public domain databases: FRGCv2.0 [66], AR [67], LFW [2], and PubFig [68], which are summarized in Table 4. With these four databases we focus on three different scenarios of partial face recognition: (i) arbitrary patch extracted from the holistic face (FRGCv2.0), (ii) occluded face (AR), and (iii) real face images with arbitrary occlusion and pose variations (LFW & PubFig), respectively. We conducted three large-scale open-set identification experiments for each of the three scenarios on FRGCv2.0, AR, and PubFig, respectively, with only one image per subject enrolled in the gallery. For the LFW database, we followed the standard LFW verification benchmark test protocol. Previously, in [9] we showed that MKD-SRC performs much better than the SIFT matching approach [8] and the SRC algorithm [7]. In this paper, we focus on performance comparisons with two leading commercial face recognition SDKs, FaceVACS [3] and PittPatt [4]. Besides these two SDKs, two additional baseline algorithms were also compared. The first one is a subspace method (PCA+LDA) implemented in the CSU Face Identification Evaluation System (csuFaceIdEval) V5.1 [69]. The other one is a local feature based approach using LBP [33]. We followed the same procedure as in [33], which divides the whole face image into non-overlapping local blocks, builds an LBP histogram descriptor for each block, matches each block by the χ^2 distance measure, and finally fuses all block-based results by the sum-rule. For the LBP approach we set the patch size to 16×16 pixels, and used the LBP_{81}^{u2} operator. All the above methods used in our comparative evaluation require face alignment. We compared two versions of the proposed alignment free MKD-SRC approach, one with the SIFT descriptor [8] and the other with the newly proposed GTP descriptor, denoted as MKD-SRC-SIFT and MKD-SRC-GTP, respectively. MKD-SRC-SIFT was proposed in our preliminary work [9].

As indicated in [72], the open-set identification task is a more general scenario, with both closed-set identification and verification being its special cases. The watch-list task is an important application that requires open-set identification. Two sub-tasks, detection and identification, are involved in the open-set identification process. In the detection sub-task, the system has to decide if the probe image is in the gallery or not. In the identification task, the system has to report the identity of the accepted probe. The performance evaluation of the open-set identification task involves three sets of images. The first set is the gallery set $\mathscr{G}$, which contains all images that are known to the system. The other two are probe sets $\mathscr{P}_{\mathscr{G}}$ and $\mathscr{P}_{\mathscr{N}}$. While $\mathscr{P}_{\mathscr{G}}$ consists of subjects in the gallery set $\mathscr{G}$ but with different images, $\mathscr{P}_{\mathscr{N}}$ includes subjects that are not present in $\mathscr{G}$. Two performance measures, the detection and identification rate (DIR) P_{DI}, and the false alarm rate (FAR) P_{FA}, are calculated for

TABLE 4
Databases and algorithms used in our experiments

Database	FRGCv2.0+	AR+	PubFig+	LFW
Scenario	Partial patch	Occlusion	Pose & occlusion	Pose & occlusion
#Subjects	20,466	20,135	5,140	5,749
#Gallery	10,466	10,135	5,083	6,000
#Probe	25,562	11,530	8,027	6,000
Methods	MKD-SRC	MKD-SRC, PittPatt, FaceVACS, PCA+LDA, LBP		MKD-SRC, [70], PittPat, [71], FaceVACS, PCA+LDA, LBP

Notes: (1) Open-set identification was conducted on each database with 1 image/subject in the gallery, with the exception that verification was done on LFW. (2) "#Subjects" denotes the number of subjects in both the gallery and probe sets, while "#Gallery" and "#Probe" denotes the number of images in each set. (3) FRGC database has been randomly cropped to generate partial face images as probe, for which the alignment-based algorithms cannot be applied. (4) Some background images from databases other than these specific databases were added to enlarge the gallery set. Please see the text for more details.

evaluation [72]. They are formulated as

$$P_{DI}(\tau) = \frac{|\{p|p \in \mathscr{P}_{\mathscr{G}}, \ s(g^*, p) \geq \tau, \ \text{and} \ id(g^*, p) = 1\}|}{|\mathscr{P}_{\mathscr{G}}|},$$

$$P_{FA}(\tau) = \frac{|\{p|p \in \mathscr{P}_{\mathscr{N}}, \ \text{and} \ s(g^*, p) \geq \tau\}|}{|\mathscr{P}_{\mathscr{N}}|},$$

(13)

where $s(\cdot, \cdot)$ is the similarity score function, τ is the decision threshold, $g^* = arg \max_{g \in \mathscr{G}} s(g, p)$, and $id(g, p)$ is an indicator whether g and p belong to the same identity. By changing the decision threshold τ, a Receiver Operating Characteristic (ROC) curve can be drawn by plotting DIR versus FAR. Note that when FAR=100%, the corresponding DIR is the traditional rank-1 accuracy of the closed-set identification task with the gallery set $\mathscr{G}$ and the probe set $\mathscr{P}_{\mathscr{G}}$.

4.1 Partial Face Recognition for Arbitrary Patch

We synthetically generated a large database of partial faces from 16,028 frontal face images of 466 subjects from the Face Recognition Grand Challenge Ver2.0 (FRGCv2.0) database [66]. We conducted a large-scale open-set identification experiment on this database. The gallery set $\mathscr{G}$ contained 466 full face images of the 466 subjects, with just one image per subject. To make the identification problem more challenging, we included an additional 10,000 full frontal face images (1 image per subject) from a private database to enlarge the gallery set to 10,466 subjects in total. The probe set $\mathscr{P}_{\mathscr{G}}$ consisted of 15,562 partial face images of the 466 subjects from the FRGC database. The other probe set $\mathscr{P}_{\mathscr{N}}$ included 10,000 partial face images of 10,000 subjects from the private database, where none of these subjects were present in the gallery set. The extended database is denoted as FRGCv2.0+ to avoid confusion with the original FRGCv2.0 database.

(a) gallery (b) probe

Fig. 5. Example face images for experiments with partial face patches: (a) gallery images from the FRGCv2.0 database (upper row) and the extended gallery set (bottom row); (b) partial face images cropped from the FRGCv2.0 database (upper row) for $\mathscr{P}_{\mathscr{G}}$ and the private database (bottom row) for $\mathscr{P}_{\mathscr{N}}$.
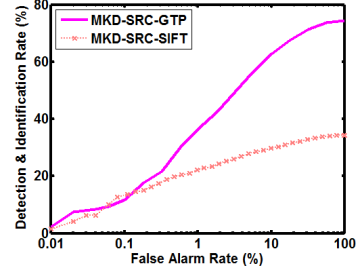

Fig. 6. ROC curves for open-set identification on face patches extracted from the FRGCv2.0+ database. Note that other algorithms (FaceVACS, PittPat, P-CA+LDA and LBP) used in our comparative study could not be applied for arbitrary partial face patches since they all require alignment.

All gallery face images were cropped to 128×128 pixels based on the two eye coordinates. Fig. 5(a) shows examples of cropped face images in the gallery. To generate partial face patches as probe, a full face image was first randomly rotated according to a Gaussian distribution with mean 0 and standard deviation $10°$. Next, a face patch at random position of a random size was extracted to represent a partial face. Finally, the cropped patch was scaled to $h \times w$, where both h and w are uniformly distributed in [64, 256]. Fig. 5(b) shows some instances of partial face patches. Notice that these randomly rotated and scaled face patches are clearly not aligned with the gallery images.

The two commercial face matchers (PittPatt and FaceVACS) and the two baseline algorithms (P-CA+LDA and LBP) could not be evaluated in this experiment. Because they all require face alignment which is not possible for the arbitrary partial face patches in the probe set. Fig. 6 shows the ROC curves of the proposed MKD-SRC-GTP and the MKD-SRC-SIFT methods for open-set identification, where MKD-SRC-GTP performs better than MKD-SRC-SIFT in this experiment. These results show that while the proposed algorithm can operate on these partial face images, it is still a challenging problem.

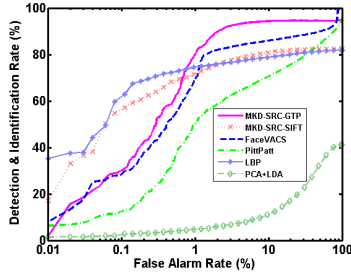To compare MKD-SRC with other methods, we

Fig. 7. ROC curves for open-set identification on full face images from the FRGCv2.0$^+$ database.

Fig. 8. Sample images from the AR database. Top row: gallery images. Bottom row: probe images.

Fig. 9. ROC curves for open-set identification on the AR$^+$ database.

conducted another open-set identification experiment with full face images. This time the probe face images were all aligned and fully cropped as in Fig. 5 (a), so that the two commercial SDKs and the two baseline algorithms could be applied. All other settings were the same as in the previous experiment. The ROC curves of all the compared algorithms are depicted in Fig. 7. Note that in this open-set identification experiment with the full face images in FRGCv2.0$^+$, the proposed MKD-SRC-GTP method performs much better than PittPatt, and slightly better than FaceVACS. Fig. 7 also shows that both MKD-SRC-SIFT and LBP perform better than MKD-SRC-GTP when FAR<1%. Finally, PCA+LDA does not perform well in this large-scale open-set identification experiment.

4.2 Holistic Face Recognition with Occlusion

The AR database contains 135 subjects, including face images wearing sunglasses or a scarf. We selected 135 non-occluded face images (1 image/subject) with neutral expression from the AR database as the gallery set $\mathscr{G}$. Again, an additional 10,000 full frontal faces (1 image/subject) were included, making the gallery size 10,135. For the probe set $\mathscr{P}_{\mathscr{G}}$, 1,530 images from the AR database were selected, all with sunglasses or scarf. Each probe image may have left or right side illumination. For the probe set $\mathscr{P}_{\mathscr{N}}$, 10,000 subjects (not present in the gallery set) with full frontal faces (1 image/subject) were selected from the private database. The extended database is denoted as AR$^+$ to avoid confusion with the original AR database. For the proposed alignment-free methods, all images were cropped to 128×128 pixels after face detection. Fig. 8 shows some cropped face images from the AR database, which are not very well aligned. For PittPatt and FaceVACS, face alignment was done by their respective SDKs (as in all other experiments). For the PCA+LDA and the LBP methods, alignment was done using manually labeled eye locations. Faces were cropped to 128×128 pixels for the LBP method after alignment.

Fig. 9 shows that MKD-SRC-SIFT and MKD-SRC-GTP have comparable performances, and both outperform other face recognition algorithms with a notable margin. At an FAR of 1%, MKD-SRC-GTP can reject 99% of impostors, while correctly detecting and identifying more than 55% of genuine probes. In contrast, both FaceVACS and PittPatt perform worse in this experiment. The second best algorithm, FaceVACS, has almost half the recognition rate of MKD-SRC. Note that this experiment is quite challenging because (i) all gallery face images in $\mathscr{G}$ are holistic without occlusion, but all probe images in $\mathscr{P}_{\mathscr{G}}$ have faces with sunglasses or scarf and also contain illumination variations; (ii) all probe images in $\mathscr{P}_{\mathscr{N}}$ contain holistic faces without occlusion; and (iii) only one sample per class is available in the gallery set. Therefore, all the impostor match pairs by comparing $\mathscr{P}_{\mathscr{N}}$ against $\mathscr{G}$ consist of two well-aligned holistic face images without occlusion, which appear to be more similar than all the genuine match pairs comparing $\mathscr{P}_{\mathscr{G}}$ against $\mathscr{G}$, which consist of one occluded face image and one non-occluded face image. This is why the two baseline algorithms PCA+LDA and LBP perform very poor, as observed in Fig. 9.
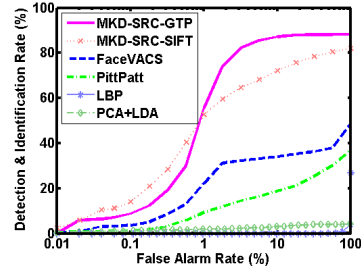
4.3 Face Recognition on LFW

The Labeled Faces in the Wild (LFW) database [2] consists of realistic and naturally occurring face images

Fig. 10. Sample images from the LFW database.

captured in uncontrolled environments and downloaded from the internet. The LFW database includes 13,233 images of 5,749 subjects. Face images in LFW contain large variations in pose, illumination, and expression, and may be arbitrarily occluded. Fig. 10 shows some example images from this database. Unlike many other approaches (e.g. [71]) which used aligned face images (either the "LFW-a" set by commercial software or the "funneled" set) provided at the LFW website [2], we applied our algorithm on the original unaligned LFW images directly. Each face image in LFW was detected and centered using the OpenCV implementation of the Viola-Jones face detector, and the cropping region returned by the detector was enlarged by a factor of 2.2 and then scaled to 250×250 pixels. Since the enlarged images contain many background pixels (see Fig. 10), we cropped the center 150×100 pixels of each image for matching, as shown in Fig. 12. For PCA+LDA, we used PittPatt to detect the two eye locations required by the csuFaceIdEval system. For all the 13,233 images, 13,124 (99.18%) pairs of eyes were detected. For the LBP approach, we used the funnel aligned LFW images, and cropped the face region with 150×100 pixels.

We follow the LFW benchmark test protocol in View 2, where the dataset is divided into 10 subsets for cross validation, with each subset containing 300 pairs of genuine matches and 300 pairs of impostor matches for verification. The mean values of FAR and GAR (Genuine Accept Rate) with fixed thresholds over all the 10 subsets are plotted in an ROC curve for performance evaluation. For the proposed MKD-SRC approach, the procedure described in Section 3.4 was applied for verification. In particular, for each of the 10 test subsets, 100 images were randomly selected from the other 9 subsets to compose the virtual gallery set for our algorithm. Note that the actual identity information about the selected images was not used in training; therefore, our method follows the LFW image-restricted training protocol as described in [2]. Furthermore, as all the faces can be detected automatically, we applied a spatial constraint that only keypoints within a distance of $0.1h$ in the vertical dimension were filtered out in the filtering stage described in Section 3.3, where $h = 150$ is the height of the cropped face. We did not consider the horizontal direction because there are large pose variations in this direction. We found that this spatial constraint improved the performance by about 2%.

Fig. 11 shows the ROC curves of the various face recognition algorithms. The proposed MKD-SRC-GTP method outperforms the commercial matcher FaceVACS and the best published method (V1-like/MKL) [71] which strictly follows the LFW image-restricted training protocol. But, the proposed method does not perform as well as PittPatt. However, a
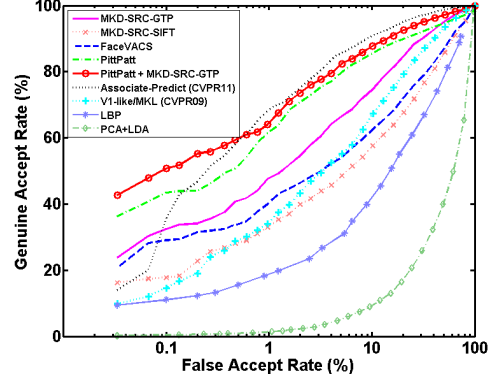

Fig. 11. ROC curves for the LFW database.

sum score fusion[6] of MKD-SRC-GTP with PittPatt improves PittPatt's performance (MKD-SRC-GTP also improves FaceVACS's performance by fusion), showing that our alignment-free partial face recognition method is able to provide complementary information to commercial holistic FR systems. The sum score fusion of MKD-SRC-GTP and PittPatt was applied after min-max score normalization (separately trained in the cross-validation setting for the 10 subsets), and the two matchers were equally weighted. Moreover, the fused result of MKD-SRC-GTP and PittPatt outperforms the best published result (Associate-Predict) [70] on the LFW database when FAR is below 1%. Note that the Associate-Predict method used a large-scale dataset in addition to LFW for training. Furthermore, MKD-SRC-GTP improves the performance of our previous method MKD-SRC-SIFT [9] significantly, suggesting that the new GTP descriptor is more effective than SIFT in uncontrolled scenarios. Compared to the two baseline algorithms PCA+LDA and LBP, our MKD-SRC based methods also perform better. Interestingly, LBP performs much better than PCA+LDA, since the holistic appearance of a face varies drastically in uncontrolled scenarios.

Fig. 12 shows some example face image pairs that can be correctly recognized by MKD-SRC-GTP but not by PittPatt at FAR=1%. Similar to the AR database, MKD-SRC-GTP performs better with faces wearing sunglasses. Besides, MKD-SRC-GTP can also match faces in different views, especially when a mirrored image is included in the virtual gallery set to learn the sparse representation (see Section 3.4).

To further understand the strength of the proposed

6. We found that the fusion of PittPatt and FaceVACS did not perform better than PittPatt, as provided in the supplementary file. This is because PittPatt performs much better than FaceVACS in this experiment, and it appears that the two matchers are not complementary to each other. While we don't know the actual algorithms utilized in these two SDKs, both are alignment based, which is quite different from our alignment-free approach. Therefore, it is expected that the alignment-based and alignment-free approaches would provide sufficient complementary information. That is why the fusion of PittPatt and MKD-SRC-GTP performs better.

Fig. 12. Pairs of face images from the LFW database that can be correctly recognized by MKD-SRC-GTP but not by PittPatt at FAR=1%.

Fig. 13. Genuine pairs of partial face images from a subset of the LFW database.

alignment-free partial face recognition method, we did another experiment with a subset of the LFW database. From the 10 folds of the View 2 experiments, we selected 500 genuine pairs, where either one, or both the face images to be matched have sunglasses, hats, occlusions by hand or other objects, or large pose variations ($> 45°$). All these selected images are partial face images as categorized in Table 1. Some examples of the selected genuine pairs are shown in Fig. 13. We used the same 3,000 impostor pairs as in View 2. The resulting ROC curves are shown in Fig. 14. Comparing with Fig. 11, the performances of all algorithms degrade on this more challenging subset of LFW. Nevertheless, the proposed MKD-SRC-GTP method now performs better than both PittPatt and FaceVACS. Besides, the fusion of PittPatt and MKD-SRC-GTP further improves the recognition performance. This shows that MKD-SRC-GTP is more suitable for general partial face recognition scenarios.

4.4 Face Recognition on PubFig

The PubFig database [68] contains 58,797 face images of 200 famous personalities collected from the internet, where 60 subjects are used for algorithm development and 140 subjects are used as the evaluation set. Due to many broken links to the original images (on various websites like Flickr), we could not download the whole PubFig database for testing. For the evaluation set we downloaded 23,121 images of the 140 subjects. From this available data, we conducted an open-set identification experiment. The gallery set contains 83 full and near frontal face images of 83 subjects from the PubFig database, plus 5,000 full face images of
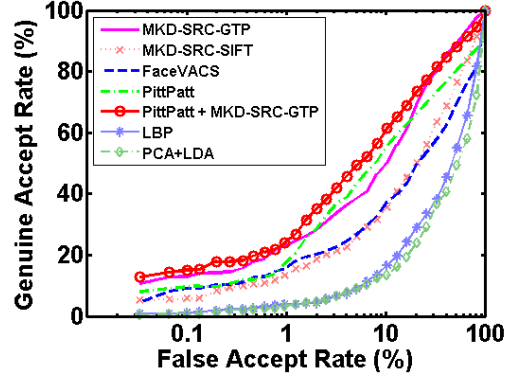

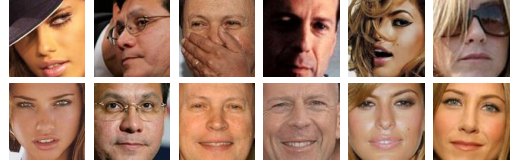Fig. 14. ROC curves for a subset of the LFW database.

Fig. 15. Probe face images (top) and the corresponding gallery images (bottom) from the PubFig database.

5,000 subjects (not overlapping with PubFig) from the LFW database. This extended database is denoted as PubFig$^+$ to avoid confusion with the original PubFig database. The probe set $\mathscr{P}_{\mathscr{G}}$ includes 817 images of the same 83 subjects in the gallery set. All faces in $\mathscr{P}_{\mathscr{G}}$ are partial faces with sunglasses, hats, occluded by hair, hand, or other objects, extreme illumination, or large pose variations ($> 45°$). Fig. 15 shows some examples of genuine pairs of images used in our experiment. The probe set $\mathscr{P}_{\mathscr{N}}$ contains 7,210 holistic or partial face images of 57 subjects from the PubFig database. For each image, a face is annotated by a bounding box provided by a commercial face detector. We cropped the face according to the bounding box, and scaled it to 100×100 pixels for the proposed MKD-SRC methods. For PCA+LDA and LBP, we used PittPatt to detect the two eyes, with a detection rate of 98.33%. Faces were cropped to 128×128 pixels for the LBP approach after alignment.

The challenges associated with this experiment include: (i) large-scale open-set identification; (ii) one sample per class in the gallery; (iii) real-word partial faces, with large intra-class variations in pose, illumination, expression, and occlusions. We compared the proposed MKD-SRC algorithm with PittPatt and FaceVACS, and also the two baseline algorithms PCA+LDA and LBP. The ROC curves are shown in Fig. 16. We can observe that while all algorithms perform poorly in this very challenging experiment, MKD-SRC-GTP performs the best.
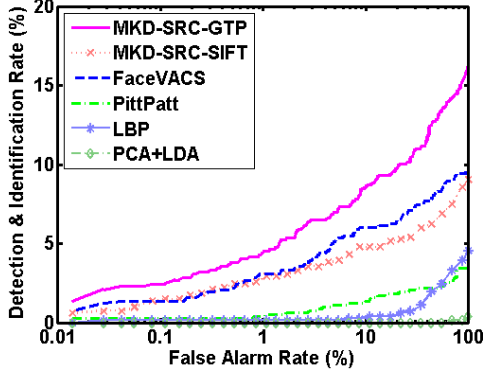
Fig. 16. ROC curves for the PubFig$^+$ database.

4.5 Parameter Values

All the parameter values (t, a, M and L) for the proposed algorithm were fixed for all our experiments on the four databases (see Table 3). In this subsection, we provide a sensitivity analysis of parameter values on our algorithm's performance. We used the development set of the LFW database in View 1 in Image Restricted Configuration. The training set contains 1,100 pairs of genuine samples, and 1,100 pairs of impostor samples. From the training set we random selected 100 images for the virtual gallery set of our algorithm. The testing set contains 500 pairs of genuine samples and 500 pairs of impostor samples for verification experiments. We used the Area Under Curve (AUC) for performance comparison. Each time we evaluated one parameter while fixing the other three parameters. The performance influence of our algorithm in terms of parameter values is demonstrated in Fig. 17. From Fig. 17(a) it can be observed that our choice of the LTP encoding is better than the LBP encoding, and a good range of values for the parameter t is [0.01, 0.07]. Fig. 17(b) shows that applying a sigmoid function to the histogram descriptor is better than without it, and a good range of values for a is [5, 30]. For the number of PCA dimensions, Fig. 17(c) shows that $M \in [128, 256]$ performs well. However, larger values of M increase both the memory and computational costs. Finally, Fig. 17(d) shows that our algorithm is very stable when $L \geq 50$, with lower values of L resulting in faster computation. Therefore, applying the fast filtering technique proposed in Section 3.3 improves the efficiency of the proposed algorithm without loss of accuracy.

5 SUMMARY AND FUTURE WORK

We have addressed the problem of recognizing a face from its partial image, and proposed an alignment free approach, called MKD-SRC. Our approach represents each face image with a set of keypoint descriptors (GTP and SIFT), and constructs a large dictionary from all the gallery descriptors. In this
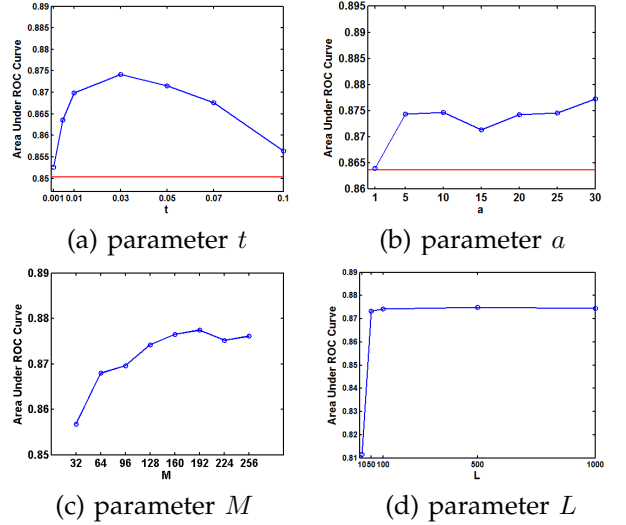

Fig. 17. Sensitivity analysis of parameters. The red line in (a) is the result of LBP encoding instead of LTP. The red line in (b) is the result without a sigmoid function.

way descriptors of a partial probe image can be sparsely represented by the dictionary, and the identity of the probe can be inferred accordingly. The proposed approach shows promising results on synthesized partial faces (from the FRGCv2.0 database), occluded holistic faces (AR database), and occluded or non-frontal faces collected in unconstrained scenarios (LFW and PubFig databases). A comparison with two commercial face matchers, FaceVACS and PittPatt, shows that MKD-SRC, particularly with the proposed GTP descriptor, is well suited for general partial face recognition problem. In case a partial face cannot be detected, our approach can still provide a matching score given a manually cropped face region. Given the general framework of MKD-SRC, it would be useful to apply MKD-SRC to other image classification areas, such as object categorization.

REFERENCES

[1] G. Hua, M.-H. Yang, E. Learned-Miller, Y. Ma, M. Turk, D. J. Kriegman, and T. S. Huang, "Introduction to the special section on real-world face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1921–1924, 2011.

[2] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 07-49, October 2007, http://vis-www.cs.umass.edu/lfw/.

[3] FaceVACS Software Developer Kit, Cognitec Systems GmbH, http://www.cognitec-systems.de.

[4] PittPatt Software Developer Kit, Pittsburgh Pattern Recognition, Inc., http://www.pittpatt.com.

[5] "Police use facial recognition technology to nab rioters," http://www.msnbc.msn.com/id/44110353/ns/technology_and_science-tech_and_gadgets/#.TkR_lnO4KsJ.

[6] "Face recognition technology fails to find UK rioters," http://www.newscientist.com/article/mg21128266.000-face-recognition-technology-fails-to-find-uk-rioters.html.

[7] 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*, vol. 31, pp. 210–227, 2009.

[8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[9] S. Liao and A. K. Jain, "Partial face recognition: an alignment free approach," in *Proceedings of the IAPR/IEEE International Joint Conference on Biometrics (IJCB 2011)*, Oct. 11-13 2011.

[10] T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham, "Active shape models - their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan 1995.

[11] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, June 2001.

[12] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Toward a practical face recognition system: Robust alignment and illumination by sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 372–386, Feb. 2012.

[13] S. Z. Li, X. W. Hou, and H. J. Zhang, "Learning spatially localized, parts-based representation," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001.

[14] J. Kim, J. Choi, J. Yi, and M. Turk, "Effective representation using ICA for face recognition robust to local distortion and partial occlusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1977–1981, 2005.

[15] H. J. Oh, K. M. Lee, and S. U. Lee, "Occlusion invariant face recognition using selective local non-negative matrix factorization basis images," *Image and Vision Computing*, vol. 26, no. 11, pp. 1515–1523, 2008.

[16] K. Hotta, "Robust face recognition under partial occlusion based on support vector machine with local gaussian summation kernel," *Image and Vision Computing*, vol. 26, no. 11, pp. 1490–1498, 2008.

[17] H. Jia and A. Martínez, "Support vector machines in face recognition with occlusions," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 2009.

[18] H. Ekenel and R. Stiefelhagen, "Why is facial occlusion a challenging problem?" in *Proceedings of the 3rd IAPR/IEEE International Conference on Biometrics*, 2009.

[19] D. Beymer, "Face recognition under varying pose," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 1994, pp. 756–761.

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

[21] T. Vetter and T. Poggio, "Linear object classes and image synthesis from a single example image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 733–741, 1997.

[22] D. Graham and N. Allison, "Face recognition from unfamiliar views: subspace methods and pose dependency," in *Proceedings of the 3rd International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 348–353.

[23] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775–779, 1997.

[24] V. Blanz, S. Romdhani, and T. Vetter, "Face identification across different poses and illumination with a 3D morphable model," in *Proceedings of the Fifth International Conference on Face and Gesture Recognition*, 2002, pp. 202–207.

[25] R. Gross, I. Matthews, and S. Baker, "Fisher light-fields for face recognition across pose and illumination," *Pattern Recognition*, pp. 481–489, 2002.

[26] B. Heisele, P. Ho, J. Wu, and T. Poggio, "Face recognition: component-based versus global approaches," *Computer Vision and Image Understanding*, vol. 91, no. 1-2, pp. 6–21, 2003.

[27] J. Yang, S. Liao, and S. Z. Li, "Automatic partial face alignment in NIR video sequences," in *Proceedings of the 3rd IAPR/IEEE International Conference on Biometrics*, 2009.

[28] D. Yi, S. Liao, Z. Lei, J. Sang, and S. Li, "Partial face matching between near infrared and visual images in MBGC portal challenge," in *Proceedings of the 3rd IAPR/IEEE International Conference on Biometrics*, 2009.

[29] K. Sato, S. Shah, and J. Aggarwal, "Partial face recognition using radial basis function networks," in *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 288–293.

[30] S. Gutta, V. Philomin, and M. Trajkovic, "An investigation into the use of partial-faces for face recognition," in *Proceedings of International Conference on Automatic Face and Gesture Recognition*, 2002, pp. 33–38.

[31] U. Park, A. Ross, and A. Jain, "Periocular biometrics in the visible spectrum: A feasibility study," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, 2009.

[32] A. Martínez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 748–763, June 2002.

[33] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[34] K. Pan, S. Liao, Z. Zhang, S. Li, and P. Zhang, "Part-based face recognition using near infrared images," in *Proceedings of IEEE International Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum*, 2007.

[35] R. Min, A. Hadid, and J.-L. Dugelay, "Improving the recognition of faces occluded by facial accessories," in *IEEE Conference on Automatic Face and Gesture Recognition*, 2011.

[36] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042–1052, 1993.

[37] C. Sanderson and K. Paliwal, "Polynomial features for robust face authentication," in *International Conference on Image Processing*, vol. 3, 2002, pp. 997–1000.

[38] F. Cardinaux, C. Sanderson, and S. Marcel, "Comparison of MLP and GMM Classifiers for Face Verification on XM2VTS," in *Audio- and Video-Based Biometric Person Authentication*, 2003.

[39] S. Lucey and T. Chen, "A GMM parts based face representation for improved verification through relevance adaptation," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2004.

[40] F. Cardinaux, C. Sanderson, and S. Bengio, "User authentication via adapted statistical models of face images," *IEEE Trans. on Signal Processing*, vol. 54, no. 1, pp. 361–373, 2006.

[41] C. Chan and J. Kittler, "Sparse representation of (Multiscale) histograms for face recognition robust to registration and illumination problems," in *The 17th IEEE International Conference on Image Processing (ICIP)*, 2010, pp. 2441–2444.

[42] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *ECCV Workshop on statistical learning in computer vision*, 2004.

[43] D. Sun and Z. Qiu, "Bag-of-words vector quantization based face identification," in *International Symposium on Electronic Commerce and Security (ISECS'09)*, vol. 2, 2009, pp. 29–33.

[44] Z. Li, J. Imai, and M. Kaneko, "Robust face recognition using block-based bag of words," in *the 20th International Conference on Pattern Recognition*, Aug. 2010.

[45] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, "On the use of SIFT features for face authentication," in *IEEE Conference on Computer Vision and Pattern Recognition Workshop*, 2006.

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

[47] D. Kisku, A. Rattani, E. Grosso, and M. Tistarelli, "Face identification by SIFT-based complete graph topology," in *IEEE Workshop on Automatic Identification Advanced Technologies*, 2007.

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

[49] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?" in *IEEE International Conference on Computer Vision*, 2011.

[50] D. Yi, R. Liu, R. Chu, Z. Lei, and S. Z. Li, "Face matching from near infrared to visual images," in *Proceedings of the 2nd IAPR/IEEE International Conference on Biometrics*, August 2007.

[51] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.

[52] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[53] K. Mikolajczyk, A. Zisserman, and C. Schmid, "Shape recognition with edge-based features," in *Proceedings of the British Machine Vision Conference*, 2003.

[54] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[55] ——, "An affine invariant interest point detector," in *Proceedings of the European Conference on Computer Vision*, 2002.

[56] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, vol. 22, no. 10, pp. 761–767, 2004.

[57] "Scale & affine invariant feature detectors." [Online]. Available: http://www.robots.ox.ac.uk/˜vgg/research/affine/det_eval_files/extract_features2.tar.gz

[58] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 467–476, 2002.

[59] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, 2007.

[60] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994.

[61] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[62] H. Xu, C. Caramanis, and S. Mannor, "Sparse algorithms are not stable: A no-free-lunch theorem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 187–193, 2012.

[63] D. Donoho and Y. Tsaig, "Fast solution of ℓ_1-norm minimization problems when the solution may be sparse," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 4789–4812, 2008.

[64] D. Musser, "Introspective sorting and selection algorithms," *Software Practice and Experience*, vol. 27, no. 8, pp. 983–993, 1997.

[65] D. Beymer and T. Poggio, "Face recognition from one example view," in *International Conference on Computer Vision*, 1995.

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

[67] A. Martinez and R. Benavente, "The AR face database," CVC Technical Report, Tech. Rep., 1998.

[68] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and simile classifiers for face verification," in *IEEE International Conference on Computer Vision (ICCV)*, Oct. 2009. [Online]. Available: http://www.cs.columbia.edu/CAVE/databases/pubfig/

[69] CSU Face Identification Evaluation System, http://www.cs.colostate.edu/evalfacerec/.

[70] Q. Yin, X. Tang, and J. Sun, "An associate-predict model for face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2011.

[71] N. Pinto, J. DiCarlo, and D. Cox, "How far can you get with a modern face recognition test set using only simple features?" in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[72] P. J. Phillips, P. Grother, and R. Micheals, "Evaluation methods in face recognition," in *Handbook of Face Recognition*, 2nd ed., S. Z. Li and A. K. Jain, Eds. Springer, Sep. 2011, pp. 551–574.

Shengcai Liao received the B.S. degree in mathematics and applied mathematics from the Sun Yat-sen University, Guangzhou, China, in 2005 and the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2010. He was a Post Doctoral Fellow in the Department of Computer Science and Engineering, Michigan State University during 2010-2012. He is currently an Assistant Professor in the Institute of Automation, Chinese Academy of Sciences. His research interests include computer vision, pattern recognition, and machine learning, with a focus on image and video analysis, particularly face recognition, object detection and recognition, and video surveillance. He was awarded the Motorola Best Student Paper award and the 1st Place Best Biometrics Paper award at the International Conference on Biometrics in 2006 and 2007, respectively, for his work on face recognition.

Anil K. Jain is a university distinguished professor in the Department of Computer Science and Engineering at Michigan State University. His research interests include pattern recognition and biometric authentication. He served as the editor-in-chief of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE (1991-1994). The holder of six patents in the area of fingerprints, he is the author of a number of books, including Handbook of Fingerprint Recognition (2009), Handbook of Biometrics (2011), Handbook of Multibiometrics (2006), Handbook of Face Recognition (2005), BIOMETRICS: Personal Identification in Networked Society (1999), and Algorithms for Clustering Data (1988). He served as a member of the Defense Science Board and The National Academies committees on Whither Biometrics and Improvised Explosive Devices. Dr. Jain received the 1996 IEEE TRANSACTIONS ON NEURAL NETWORKS Outstanding Paper Award and the Pattern Recognition Society best paper awards in 1987, 1991, and 2005. He is a fellow of the AAAS, ACM, IAPR, and SPIE. He has received Fulbright, Guggenheim, Alexander von Humboldt, IEEE Computer Society Technical Achievement, IEEE Wallace McDowell, ICDM Research Contributions, and IAPR King-Sun Fu awards.

Stan Z. Li received the B.Eng. degree from Hunan University, Changsha, China, the M.Eng. degree from the National University of Defense Technology, China, and the Ph.D. degree from Surrey University, Surrey, U.K. He is currently a Professor and the Director of Center for Biometrics and Security Research (CBSR), Institute of Automation, Chinese Academy of Sciences (CASIA). He worked at Microsoft Research Asia as a researcher from 2000 to 2004. Prior to that, he was an Associate Professor at Nanyang Technological University, Singapore. His research interest includes pattern recognition and machine learning, image and vision processing, face recognition, biometrics, and intelligent video surveillance. He has published over 200 papers in international journals and conferences, and authored and edited eight books. Dr. Li is currently an Associate Editor of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE and is acting as the Editor-in-Chief for the Encyclopedia of Biometrics. He served as a co-chair for the International Conference on Biometrics 2007 and 2009, and has been involved in organizing other international conferences and workshops in the fields of his research interest.