Reference Face Graph for Face Recognition

Mehran Kafai*, *Member, IEEE,* Le An*, *Student Member, IEEE,* and Bir Bhanu, *Fellow, IEEE*

Abstract—Face recognition has been studied extensively; however, real-world face recognition still remains a challenging task. The demand for unconstrained practical face recognition is rising with the explosion of online multimedia such as social networks, and video surveillance footage where face analysis is of significant importance. In this paper, we approach face recognition in the context of graph theory. We recognize an unknown face using an external Reference Face Graph (RFG). A RFG is generated and recognition of a given face is achieved by comparing it to the faces in the constructed RFG. Centrality measures are utilized to identify distinctive faces in the reference face graph. The proposed RFG-based face recognition algorithm is robust to the changes in pose and it is also alignment-free. RFG recognition is used in conjunction with DCT locality sensitive hashing for efficient retrieval to ensure scalability. Experiments are conducted on several publicly available databases and the results show that the proposed approach outperforms the state-of-the-art methods without any preprocessing necessities such as face alignment. Due to the richness in the reference set construction, the proposed method can also handle illumination and expression variation.

Index Terms—Face recognition, graph analysis, centrality measure, alignment-free, pose robust.

I. INTRODUCTION

As the explosion of multimedia has been observed in the past decade, the processing and analysis of the multimedia data, including images and videos, are of broad interest. Among these data, face images and videos take a large fraction. The recognition of the faces in different contexts enables various applications such as surveillance monitoring, social networking, and human-computer interaction. For example, in the recent event of Boston marathon bombings, images of the suspects taken from street surveillance cameras aided the FBI to identify the suspects [1].

Face recognition includes the studies of automatically identifying or verifying a person from an image or video sequences. Face identification refers to determining the ID of the person, given as a probe, from a large pool of candidates in the gallery. Face verification or face authentication is the problem of deciding whether a given pair of images are of the same person or not. Although face recognition has been studied extensively and impressive results have been reported on benchmark databases [2] [3] [4], unconstrained
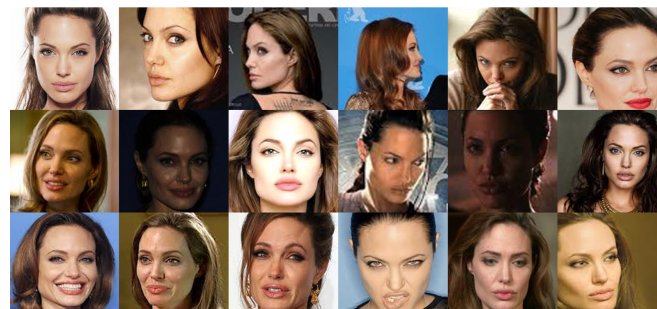
Fig. 1. Sample unconstrained face images of the same person with variations in pose, illumination, expression, and alignment. Such variations make face recognition a challenging task.

face recognition is of more practical use and it is still a challenge due to the following factors:

- Pose variation. The face to be recognized may come under arbitrary poses.
- Misalignment. As the faces are usually detected by an automated face detector [5], the cropped faces are not aligned. However since most feature descriptors require alignment before feature extraction, misalignment degrades the performance of a face recognition system.
- Illumination variation. The illumination on a face is influenced by the lighting conditions and the appearance of the face would vary significantly under different lighting.
- Expression variation. The face images may differ with different expressions.

Figure 1 shows some examples of unconstrained faces. The appearance of the same person varies significantly due to the change in imaging conditions. In most of the previous literature either some or all of the aforementioned factors have to be taken care of before the recognition algorithm is able to perform. For instance, illumination normalization is performed as part of a pre-processing step to remove the variations of the lighting effects in [6]. For commercial face recognition software, some specific constraints have to be imposed. For example, the locations of the eyes have to be determined for the FaceVACS recognition system [7].

In this paper, a *Reference Face Graph framework (RFG)* is presented for face recognition. We focus on modeling the task of face recognition as a graph analysis problem. An early work by Wiskott *et al.* [8] proposed to represent each face by a bunch graph based on a Gabor Wavelet transform. In our approach, the identity of an unknown face is described by its similarity to the reference faces in the constructed graph. This is an indirect measure rather than comparing the unknown face to the gallery faces directly. The purpose of using a reference set of external faces is to represent a given face by its degree of similarity to the images of a set of N reference individuals. Specifically, two faces which are visually similar in one pose,

for example profile, are also, to some extent, similar in other poses, for example frontal. In other words, we assume that visual similarity follows from underlying physical similarity in the real world. We take advantage of this phenomenon in the following way: compare a gallery/probe face image, with all the images of a reference individual from the reference set. The similarity with the best matching image of the reference individual is the degree of the similarity between the gallery/probe subject and the reference individual. By repeating this procedure for each of the reference individuals, we create a basis descriptor for the gallery/probe face image that reflects the degree of similarity between this face image and the reference set individuals. Through the use of reference-based descriptors, we mitigate the pose and other variations by not using the appearance features extracted directly from the original images for face matching. The purpose for building a reference face graph is to obtain the node centralities. By doing this, we determine which face in the reference set is more discriminative and important. The centralities are used as weights for the faces in the reference face graph. The proposed alignment-free approach is robust to pose changes and it is tolerant to illumination and expression change.

II. RELATED WORK AND CONTRIBUTIONS

A. Related Work

1) General approaches towards robust face recognition: Recent work focuses on unconstrained face recognition, which is inherently difficult due to the uncontrolled image acquisition that allows variations in pose, illumination, expression, and misalignment.

In the pursuit of an advanced feature descriptor, Barkan *et al.* [9] form an over-complete face representation using modified multi-scale LBP features for face recognition. Together with Diffusion Maps as a feature reduction technique, competitive results are obtained. Lei *et al.* [10] learn discriminant local features called discriminant face descriptors (DFD) in a data-driven manner instead of a handcrafted way and it is effectiveness on both homogeneous face recognition (*i.e.*, images from the same modality) and heterogeneous face recognition (*i.e.*, images from different modalities). Contrary to tradition in recognition where high-dimensional features are not preferred, Chen *et al.* [11] empirically show that using high-dimensional LBP feature descriptors, the state-of-the-art is achieved. To make the use of high-dimensional features applicable, a sparse projection method is proposed to reduce computation and model storage.

To tackle the pose variations, pose-invariant face recognition is achieved by using Markov random field based image matching [12]. In [13] an expression invariant face recognition algorithm is proposed based on compressive sensing. Recently Liao *et al.* [14] propose an alignment-free approach for partial face recognition using Multi-Keypoint Descriptors (MKDs). However, none of these methods are able to simultaneously handle the faces with unconstrained pose, illumination, expression, and alignment. In [15], a 3D model for each subject in the database is constructed using a single 2D image, then the synthesized face images at different poses are generated

from the 3D model for matching. In contrast to most distance-based methods for face recognition, the probability that two faces have the same underlying identity cause is evaluated in [16]. This approach renders comparable or better results than current methods for face recognition with varying poses. In [17] a probabilistic elastic matching method is proposed to handle pose variation in recognition. In this method, a Gaussian mixture model (GMM) is used to capture the spacial-appearance distribution of the faces in the training set and SVM is used for face verification.

To eliminate the lighting effects that hinder face recognition systems, a face representation called Local Quantized Patterns (LQP) is proposed in [18]. The illumination invariance of this representation leads to improved performance for state-of-the-art methods on the challenging Labeled Faces in the Wild (LFW) [19] database. In [6], a robust illumination normalization technique is proposed using Gamma correction, difference of Gaussian filtering, masking and contrast equalization. The preprocessing step improves the recognition performance on several benchmark databases. In [20], near infrared images are used for face recognition regardless of visible illumination changes in the environment.

For face recognition with expressions, most approaches aim at reproducing the neutral faces for matching. In [13], the images of the same subject with different expressions are viewed as an ensemble of intercorrelated signals and the sparsity accounts for the variation in expressions. Thus, two feature images, the holistic face image and the expression image, are generated for subsequent face recognition. Hsieh *et al.* [21] remove the expression from a given face by using the optical flow computed from the input face with respect to a neutral face.

Face mis-alignment can abruptly degrade the face recognition system performance [27]. However, for the unconstrained face images, accurate alignment is a challenging topic itself. In [14] an arbitrary patch of a face image can be used to recognize a face with an alignment-free face representation. Wang *et al.* [28] propose a misalignment-robust face recognition method by inferring the spatial misalignment parameters in a trained subspace. Cui *et al.* [29] try to solve misalignment in face recognition by extracting sparse codes of position-free patches within each spatial block in the image. A pairwise-constrained multiple metric learning is proposed to integrate the face descriptors from all blocks.

For unconstrained face with multiple variations such as pose, expression, *etc.*, approaches have been developed to take some of these aspects into account simultaneously. For instance, multiple face representations and background statistics are combined in [30] for improved unconstrained face recognition. In [31], a reference set of faces is utilized for identity-preserving alignment and identity classifier learning. A collection of the classifiers is able to discriminate the subjects whose faces are captured in the wild. Müller *et al.* [32] separate the learning of the invariance from learning new instances of individuals. A set of examples called model is used to learn the invariance and new instances are compared by rank list similarity.

TABLE I
RECENT WORK SUMMARY FOR FACE RECOGNITION VIA INDIRECT SIMILARITY AND REFERENCE-BASED DESCRIPTORS

Methodology & Authors	Training?	Comments
Attribute and simile classifiers Kumar *et al.* [22]	Yes	requires trained classifier for each face part of each reference individual, facial key points need to be extracted, incapable of handling large pose variations
AP and LP models Yin *et al.* [23]	Yes	training required for probe images during online processing, person specific classifiers need to be trained
Doppelgänger lists Schroff *et al.* [24]	No	effective for across pose recognition, rank-lists comparison is expensive, slow retrieval
Message passing model Shen *et al.* [25]	Yes	ineffective for recognition across pose, message passing algorithm is used to choose the best reference images
Joint Bayesian formulation Chen *et al.* [26]	Yes	probabilistic reference-based method approach, incapable of handling large pose variations
Reference face graph This paper	Yes	alignment-free, fast retrieval, compatible with efficient indexing structures

2) Recognition via indirect matching: Recognition and retrieval via reference-based descriptors and indirect similarity has been previously explored in the field of computer vision. Shan *et al.* [33] and Guo *et al.* [34] use exemplar-based embedding for vehicle matching. Rasiwasia *et al.* [35] label images with a set of pre-defined visual concepts, and then use a probabilistic method based on the visual features for image retrieval. Liu *et al.* [36] represent human actions by a set of attributes, and perform activity recognition via visual characteristics symbolizing the spatial-temporal evolution of actions in a video.

Kumar *et al.* [22] propose using attribute and simile classifiers for verification, where face-pairs are compared via their similes and attributes rather than a direct comparison. Experiments are performed on PubFig and Labeled Faces in the Wild (LFW) databases [19] in which images with significant pose variation are not present. Also, attribute classifiers require extensive training for recognition across pose.

In the field of face recognition, method based on rank-lists have been investigated. Schroff *et al.* [24] propose describing a face image by an ordered list of similar faces from a face library, *i.e.*, a rank list representation (Doppelgänger list) is generated for each image. Proximity between gallery and probes images is determined via the similarity between ordered rank lists of the corresponding images. The main drawback of this approach is the complexity of comparing the Doppelgänger lists. The authors propose using the similarity measure of Jarvis and Patrick [37] which is computationally expensive. Also, no suitable indexing structure is available for efficient ranked-list indexing.

An associate-predict model is proposed in [23] to handle the intra-personal variations due to pose and viewpoint change. The input face is first associated with similar identities from an additional database and then the associated face is used in the prediction. Cui *et al.* [38] introduce a quadratic programming approach based on reference image sets for video-based face recognition. Images from the gallery and probe video sequences are bridged with a reference set pre-defined and pre-structured to a set of local models for exact alignment. Once the image sets are aligned, the similarity is measured by comparing the local models.

Chen *et al.* [26] represent a face x by the sum of two independent Gaussian random variables: $x = \mu + \epsilon$, where μ represents the identity of the face and ϵ represents face variations within the same identity. The covariance matrices of μ and ϵ, S_μ and S_ϵ, are learned from the training data using an expectation maximization algorithm. Having learned these, the log-likelihood ratio for two input images x_1 and x_2 being of the same individual can be computed. A summary of the most recent work for face recognition via indirect similarity is shown in Table I.

B. Contributions

As compared to the state-of-the-art approaches, the distinguishing characteristics of the proposed method are:

- A novel reference face graph (RFG) based face recognition framework is proposed. The contribution of the approach is shown by performing empirical experiments on various databases.
- DCT locality-sensitive hashing [39] is incorporated, in a novel manner, into the proposed framework for fast similarity computation, efficient retrieval, and scalability.
- The proposed framework can be used in conjunction with any feature descriptor (*e.g.*, LBP [40], LGBP [41]). The results are shown on several publicly available face databases and they are compared with the state-of-the-art techniques.

III. TECHNICAL APPROACH

Our proposed framework consists of two main steps: 1. Preparing the reference face graph; 2. Generating the reference face descriptors. In the following, first we define the terms used in this paper, and then we discuss each of the above steps in detail.

A *reference face graph (RFG)* is a structure of nodes and the dyadic relationships (edges) between the nodes. A *reference face* is a node representing a single individual in the reference face graph. Each reference face has multiple images with various poses, expressions, and illumination. All the images of these reference faces build a set called the *reference basis set*. We use the term *basis set* because we represent each probe face or gallery face as a linear combination of the similarities to the reference faces in the reference basis set. A *basis descriptor* of an image is a vector that describes the image in terms of its similarity to the reference faces. A *reference face descriptor*

(RFD) is defined by incorporating the RFG node centrality metrics into the basis descriptor.

Figure 2 shows the overall system diagram for the RFG framework for face identification. A similar process is used for face verification.

A. Building the Reference Face Graph

1) Initializing the Reference Face Graph: We define a RFG structure and pre-populate it with a set of individuals called the reference faces. The reference faces are not chosen from the gallery or probes. The reference faces build a set called the reference basis set. Figure 3 illustrates a sample reference basis set with N reference faces.

Each image in the reference basis set is partitioned into regions. In total we use eight different partitioning schemes to create 4, 9, 16, 25, 49, 64, 81, and 100 regions. The partitioned regions do not overlap. In each scheme, the regions are translated 1/4 of their size both horizontally and vertically. By translation, we mean the shifting of grid lines (e.g., after translation of one grid line, the resulting partition regions of an image no long have the same size). For each patch in a specific partitioning, a feature vector is extracted. These feature vectors are not concatenated. Instead, these feature vectors of the same person in the reference set construct a reference face R_i and DCT hash based similarity [39] computation is performed as shown in Table II.

The motivation for having such regions is related to the common mis-alignment of face images after face detection, in addition to having a scale-free approach without the need for facial feature point detection (*e.g.*, eyes). Most state-of-the-art face recognition algorithms require aligned faces to operate with high accuracy. By using multiple regions with different scale and translation setting we eliminate the need for face alignment. Our approach is capable of working on nonaligned images with different scale. Figure 4 illustrates a sample partitioning for a reference basis set image.

Each reference face is assigned a node $R_i, i = 1 \ldots N$ in the RFG. Each R_i is connected to all other nodes via a direct edge e, *i.e.*, the RFG is complete. The weight of edge e_{ij} between node i and node j, w_{ij}, represents the similarity between reference faces R_i and R_j, which is defined as

$$w_{ij} = sim(R_i, R_j) = \max_{u,v} sim(R_i^u, R_j^v), \qquad (1)$$

where R_i^u and R_j^v refer to images u and v of reference faces R_i and R_j, respectively. Equation 1 shows that the similarity
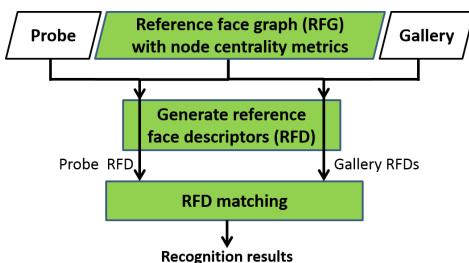

Fig. 3. Reference basis set: a set of images containing multiple individuals (reference faces). Each reference face has multiple images with various poses, expressions, and illumination settings.

between two reference faces is defined as the maximum similarity between images of two reference faces. The Chi-square measure [40] is used to compute the similarity between two reference basis set images. Specifically, given a Chi-square distance x, we convert it into a similarity score s using the Chi-square cumulative distribution function given by

$$s = 1 - \frac{\gamma(\frac{k}{2}, \frac{x}{2})}{\Gamma(\frac{k}{2})}, \qquad (2)$$

where $\Gamma(\frac{k}{2})$ is the gamma function and γ is the lower incomplete gamma function. k refers to the degrees of freedom, which is chosen as 59 in the experiments. As for the features, we utilize various texture-based descriptors (e.g., LBP [40], LGBP [41]). Later in Section IV, we state the type of features used for each experiment.

It's important to note that each node represents a reference face which itself consists of several images of the same person with various pose, expression, and illumination. Figure 5 shows a sample RFG with four reference faces. The smaller circles within each node refer to the multiple images for each reference face in the reference basis set.

2) Node Centrality Measures: We represent the proposed RFG as a static undirected weighted graph $G = (V, E)$ consisting of a set of nodes V and a set of edges E. We utilize both structural and linkage-based analysis to determine the more distinctive reference faces. Our goal is to use the linkage structure to propagate the labels among different nodes. We measure the *centrality* for each reference face to determine

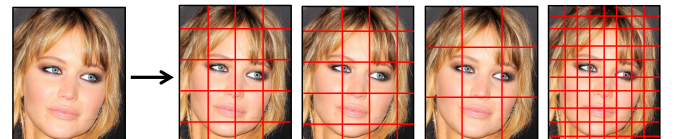Fig. 2. System diagram for reference face graph recognition.

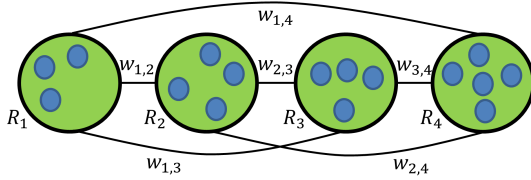Fig. 4. An example of oversampling regions.

Fig. 5. Sample reference face graph. The smaller circles within each node refer to multiple images for each reference face in the reference basis set. The edge weights represent the similarity between reference faces following Equation 1.

how important each reference face in the RFG is. In order to measure the centrality of each reference face, we adopt three measures of node centrality including degree, betweenness, and closeness for weighted graphs [42].

Generally, in weighted graphs, degree of a node is extended to the sum of weights and named as *node strength*. Since G is an undirected weighted graph, we use node strength instead of degree as a measure of centrality. $C_D^w(i)$, the node strength [42] of node i, is defined as the sum of weights of all edges connected to i. That is,

$$C_D^w(i) = \sum_j^N w_{ij}, \qquad (3)$$

where j represents all other nodes, N is the total number of nodes, and w_{ij} denotes the weight of the edge between node i and node j. For the purpose of this paper, we use the average node *weakness* for each node, $C_D^{'w}(i)$. We use the term weakness because in the proposed RFG, a node with high node strength is less distinct compared with other nodes with low node strength. In other words, a node with high node weakness is less similar to its neighboring nodes. $C_D^{'w}(i)$ is defined as:

$$C_D^{'w}(i) = 1 - \frac{C_D^w(i)}{N-1}. \qquad (4)$$

Before discussing betweenness and closeness, the distance between two nodes should be defined. In a weighted graph, the distance between node i and node j is defined as

$$d^w(i,j) = \min\left(\frac{1}{w_{ih_u}} + \ldots + \frac{1}{w_{h_v j}}\right), \qquad (5)$$

where h_u and h_v represent intermediary nodes on the path from i to j.

Betweenness is the second measure for centrality used in this paper. The betweenness for node i, $C_B^w(i)$, is defined as the number of shortest paths from all nodes to all other nodes that pass through node i. In other words,

$$C_B^w(i) = \frac{g_{jk}^w(i)}{g_{jk}^w}, \qquad (6)$$

where g_{jk}^w is the number of shortest paths connecting j and k, and $g_{jk}^w(i)$ is the number of shortest paths connecting j and k and for which node i is a part of. We use the normalized value of betweenness, $C_B^{'w}$, defined as

$$C_B^{'w}(i) = \frac{C_B^w(i)}{0.5 \times (N-1) \times (N-2)}. \qquad (7)$$

The third measure of centrality used in this paper is closeness. *Closeness* is defined as the inverse sum of shortest distances to all other nodes from a focal node [42]. In other words, closeness represents the length of the average shortest path between a node and all other nodes in a graph. More formally,

$$C_C^w(i) = \left[\sum_j^N d^w(i,j)\right]^{-1}. \qquad (8)$$

For the proposed RFG, the normalized value of closeness, $C_C^{'w}(i)$, is used. It is defined as

$$C_C^{'w}(i) = \frac{C_C^w(i)}{N-1}. \qquad (9)$$

Node weakness, betweenness, and closeness are computed for all reference faces in the RFG. For the nodes in weighted graph G (Figure 5),

$$weakness: C_D^{'w} = \{C_D^{'w}(1), C_D^{'w}(2), \ldots, C_D^{'w}(N)\} \qquad (10)$$

represents a vector of node weakness values,

$$betweenness: C_B^{'w} = \{C_B^{'w}(1), C_B^{'w}(2), \ldots, C_B^{'w}(N)\} \qquad (11)$$

denotes a vector of normalized betweenness values, and

$$closeness: C_C^{'w} = \{C_C^{'w}(1), C_C^{'w}(2), \ldots, C_C^{'w}(N)\} \qquad (12)$$

refers to a vector of normalized closeness values. $C_D^{'w}$ represents a vector of node weakness values, $C_B^{'w}$ denotes a vector of normalized betweenness values, and $C_C^{'w}$ refers to a vector of normalized closeness values.

B. Generating the Reference Face Descriptors

In the following, we discuss how the RFDs are generated for the probe and gallery image for face identification. The same methodology holds for verification as well.

A probe or gallery image is described as a function of its region-based cosine similarity to the reference faces. Computing the reference face descriptor G_A for image A consists of two steps. *First*, the basis descriptor F_A is generated. F_A (Equation 13) is an N-dimensional vector representing the similarity between image A and the reference faces in the reference basis set.

$$F_A = [sim(A, R_1), \ldots, sim(A, R_N)] \qquad (13)$$

Second, node centrality measures are incorporated to compute the reference face descriptor G_A from the basis descriptor F_A.

1) Weights for Face Regions: When computing the similarity between two images, for each region p of a face, a weight w_p is assigned, where $w_p \in [0, 1]$. The weights are computed via a genetic optimization algorithm [43]. The goal of the objective function is to maximize the classification accuracy. A set of training data disjoint from the testing data is used to compute the weights for different face regions. The motivation for computing weights for each region is based on the fact that specific regions of the human face (*e.g.*, eyes and mouth) are more discriminative than others for face recognition [40]. Images from the FERET face database [3] are used to learn the weights of face regions. We ensure that the identities for learning region weights are disjoint to the training data for building up the reference face graph and the test data.

2) DCT Locality Sensitive Hashing [39]: Computing $sim(A, R_i)$ between image A and a reference face R_i is computationally expensive because of the large number of oversampled regions. For this purpose, we use DCT locality sensitive hashing [39] to compute $sim(A, R_i)$. The DCT hash function maps input vectors to a set of hashes of size H chosen from a universe $1 \ldots U$, where U is a large integer (we chose U to be 2^{16}). In this way, computing F_A is fast and efficient, and it is performed in constant time. The number of hashes per input vector, H, is 200 and common hash suppression (CHS) is utilized. The algorithm for computing the DCT hash-based similarity between A and R_i is shown in Table II. Further details on DCT hashing can be found in [39].

3) Reference Face Descriptor: Having generated the basis descriptor F_A for image A, the reference face descriptor G_A is defined as

$$G_A = [F_A \odot C_D'^w, F_A \odot C_B'^w, F_A \odot C_C'^w], \qquad (15)$$

where $\odot$ represents element-wise multiplication.

The reference face descriptor G_A represents A in terms of its similarity to the reference faces, incorporating the centrality measures corresponding to each reference face. For face identification the reference face descriptors are generated for all probe and gallery images, and recognition is performed by comparing the reference face descriptors. For face verification the similarity between the image pair's reference face descriptors is utilized.

For efficiency, reference face descriptors are compared via DCT hash based retrieval [39]. The pseudocode for similarity computation between reference face descriptors is provided in Table II. To evaluate the contribution of the centrality measures, we provide a baseline using only reference-based (RB) face recognition without centrality measures (i.e., $G_A = F_A$ as contrast to Equation (15)).

TABLE II
PSEUDOCODE FOR COMPUTING DCT HASH-BASED SIMILARITY BETWEEN IMAGE A AND A REFERENCE FACE R_i

–Input: image A and a reference face R_i
–Output: $sim(A, R_i)$

1. Compute DCT hashes for all oversampled regions for R_i.
2. Insert DCT hashes into a hash table.
3. Generate DCT hashes for image A.
4. Retrieve bucket contents from the hash table from corresponding hashes generated for image A.
5. For each region in A determine the most similar region in R_i. Similarity between two regions is defined as the number of hashes they have in common.
6. The similarity between A and R_i is defined as

$$sim(A, R_i) = \sum_p w_p \max_q \psi(A^p, R_i^q), \qquad (14)$$

where p and q denote the oversampled regions of A and R_i respectively, ψ refers to the number hashes in common, and w_p represents the weight of region p.

IV. EXPERIMENTS

To show the effectiveness of the proposed method, we evaluate RFG for both face verification and identification. which are defined as:

- Verification: given a pair of face images, determine if they belong to the same person or not (pair matching). The verification threshold is obtained by a linear SVM classifier trained on images not used during evaluation.
- Identification: Automatically searching a facial image (probe) in a database (gallery) resulting in a set of facial images ranked by similarity.

The proposed RFG based method is compared with several state-of-the-art algorithms on multiple face databases. RFG face recognition does not require any face alignment. That is, the input to the proposed methods is the output of the face detection algorithm. In our experiments, we adopt the OpenCV implementation of the Viola-Jones face detector [5]. The detected faces have different resolutions depending on the original image size. We normalize all the detections to the size of 100×100 in all the experiments.

A. Data

The images used in the experiments are from the following databases.

- **LFW** (Labeled Faces in the Wild) database [19] : 5749 individuals, 13233 images gathered from the web
- **Multi-PIE** face database [44]: 337 individuals, $755,370$ images, 15 camera angles, 19 illumination settings, various facial expressions
- **FacePix** database [45]: 30 individuals, 16290 images spanning the spectrum of $180°$ with increments of $1°$ with various illumination conditions
- **FEI** face database [2]: 200 individuals, 2800 images, 14 images per individual with various poses and illumination conditions
- **CMU-PIE** database [4]: 68 individuals, 41368 images, 13 poses, 43 illumination conditions, 4 facial expressions

Each aforementioned database has its own specific settings. For example LFW [19] contains images taken under unconstrained settings with limited pose variation and occlusion, whereas FacePix [45] images include a variety of poses and illumination settings. We choose these databases to show how our algorithm is capable of performing under various constrained and unconstrained image settings. Figure 6 presents sample images from the databases.

(a) LFW [19] (b) M-PIE [44] (c) FacePix[45] (d) FEI [2] (e) PIE [4]

Fig. 6. Sample images from public databases

We prepare two reference basis sets. The reference set with unconstrained faces is used for recognition on unconstrained

database (*e.g.*, LFW). The reference set with lab-controlled faces is used for recognition on lab-controlled database (*e.g.*, Multi-PIE). The first set contains unconstrained images gathered from the web of 300 subjects, where each subject has 30 images, and is used for experiments on the LFW database [19]. No individual is common between the first reference basis set and the LFW database. The second reference basis set, used for experiments on the other databases, includes 500 individuals from the FEI [2] and Multi-PIE [44] databases. For each individual from the FEI database, 13 images are selected under various poses from full profile left to full profile right. For each individual from the Multi-PIE database 13 poses are chosen from profile left to profile right with 15 degree increments. In addition, three images are selected for each pose with varying illumination (images #01, #07, and #13).

B. Reference Basis Set Evaluation

The reference face descriptors are obtained by projecting the gallery and probe images into the reference basis set; therefore, the selection of reference faces of the reference basis set plays an important role, and affects the overall system performance. In order to evaluate the reference basis set, we adopt the method in [46] but from a different perspective. In [46], an image alignment method is proposed using sparse and low-rank decomposition. The low-rank decomposition learns the common component of a set of images with variations in illumination, expression, pose, etc. The calculated sparse error component indicates the specificity of that image in the image set. In our application, we examine the diversity of the reference faces. Specifically, for each pose or specific expression in the reference basis set, a matrix D whose columns represent images $I_1, \ldots, I_N$ of all reference faces is constructed. The goal is to determine the diversity of D and how effective it is as the basis function matrix. The optimization follows the formulation in [46]:

$$\min_{A,E,\tau} ||A||_\star + \lambda ||E||_1 \quad \text{s.t.} \quad D \circ \tau = A + E, \qquad (16)$$

where A is the aligned version of D, $||A||_\star$ takes the nuclear norm of A, E is the error matrix, λ is a positive weighting parameter, and $||E||_1$ corresponds to the 1-norm of E. τ is a set of transformations specified by,

$$D \circ \tau = [I_1 \circ \tau_1, \ldots, I_n \circ \tau_n], \qquad (17)$$

where $I_i \circ \tau_i$ represents applying transformation τ_i to image I_i. If the images in D are similar, the sparse error E would be small since the common component will be dominating. On the other hand, if the images in D are very dissimilar, the error E would be larger. In the reference basis set, dissimilar images are preferred to define a more definitive basis set. Figure 7 shows the averaged mean squared error of E over multiple experimental runs as the size of the reference basis set increases. The black line is a polynomial trendline of order 2 fitted to the mean squared values.

The peak of the MSE is observed with 400 reference faces in the reference basis set which contains images from FEI [2] and Multi-PIE [44]. Thus, from the 500 individuals in the second reference basis set only 400 are chosen for the experiments.
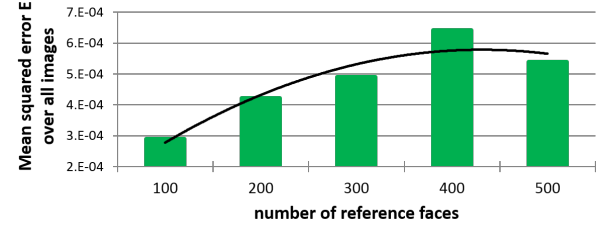

Fig. 7. Mean squared error vs. reference basis set size. The black line is a polynomial trendline of order 2 fitted to the mean squared values.

A similar evaluation is performed for the first reference basis set, and based on the results 250 individuals are chosen.

Figure 8 illustrates how the number of reference faces affects the verification accuracy on the FacePix database [45]. Four Patch LBP (FPLBP) [30] is the feature used in this experiment. As the number of reference faces increases over 400, the plot flattens around 78%. Note that the number of reference faces (N) determines the dimensionality of the reference face descriptors.
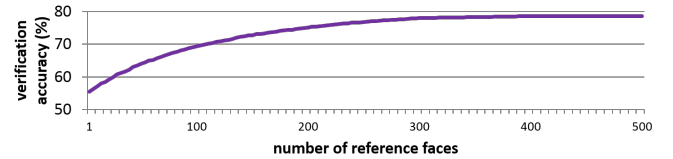

Fig. 8. Verification accuracy on FacePix [45] with increasing number of reference faces.

C. Comparison With Other Methods

We compare the proposed method with several state-of-the-art algorithms on multiple databases.

• **Comparison on LFW database [19]:** We followed the default 10-fold verification protocol on the LFW database to provide a fair comparison. Table III shows the average classification accuracies of the proposed RFG approach and other methods including a commercial recognition system. With simple LBP feature used and without face alignment, our method achieves an average classification accuracy of 0.9284 with a standard deviation of 0.0027, which is very competitive.

To examine the performance in detail, we further compare RFG recognition with associate-predict (AP) method [23], likelihood-predict (LP) model [23], cosine similarity metric

TABLE III
CLASSIFICATION ACCURACY ON LFW [19] USING PROTOCOL OF
"UNRESTRICTED, LABELED OUTSIDE DATA RESULTS"

Method	Accuracy
Attribute classifiers [22]	0.8525 ± 0.0060
Multiple LE + comp [47]	0.8445 ± 0.0046
Associate-Predict [23]	0.9057 ± 0.0056
Combined Joint Bayesian [48]	0.9242 ± 0.0108
POOF-HOG [49]	0.9280 ± 0.0047
Tom-vs-Pete [31]	0.9310 ± 0.0135
face.com [50]	0.9130 ± 0.0030
RFG (this paper)	0.9284 ± 0.0027

learning [51], and message passing model [25] using Receiver Operating Characteristic (ROC) curve for face verification on the LFW database. For these methods same LBP features are used as compared to the features in our approach. Specifically, for all the methods, uniform LBP descriptors with radius 3 and 8 neighboring pixels are used. Note that all the methods in Figure 9 except the proposed RFG perform on the aligned or funneled version of LFW, whereas RFG uses the non-aligned version. The results in Figure 9 clearly demonstrate that RFG outperforms the other methods even when no alignment is performed. The results using RB are also shown in this Figure. As compared to RB which does not utilize centrality measures, RFG achieves better results.
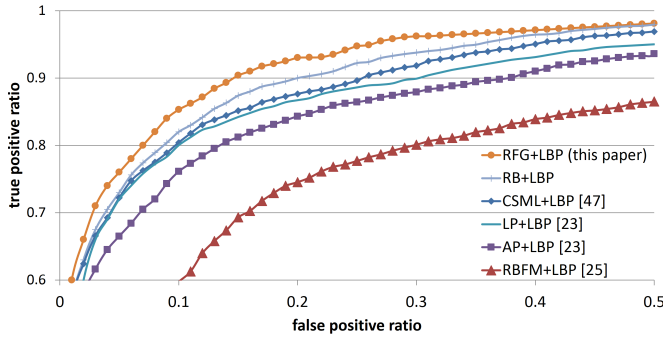

Fig. 9. ROC plot for verification comparison on LFW [19].

Figure 10 demonstrates how the proposed RFG performs on the aligned version of LFW in comparison with the original non-aligned LFW. The results show that RFG performs quite closely for aligned and non-aligned images.
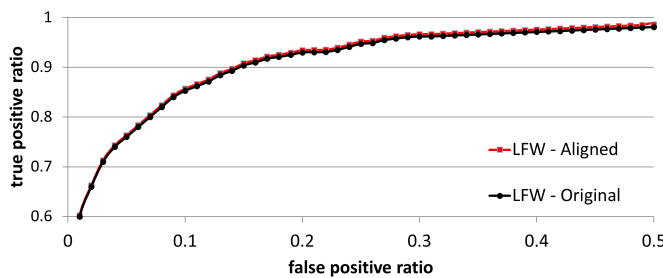

Fig. 10. ROC plot comparison for original vs. aligned LFW.

Note that although there are some recent methods (*e.g.*, [31], [11], [29]) which have reported better performance on the LFW database, but they have limitations which are not present in our method. For example, in [31] 95 parts of the face have to be located to perform alignment and 5000 classifiers have to be trained involving a lot of parameters. In [11] features are extracted at dense facial landmarks and the method heavily relies on accurate facial landmark detection. In [29] the input face has to be aligned and for each face region a Mahalanobis matrix has to be learned. Besides, it is not scalable to larger database. On the contrary, our method does not require *any* face alignment. In addition, the use of DCT hashing guarantees the efficiency and scalability of our method and this framework is compatible with any advanced
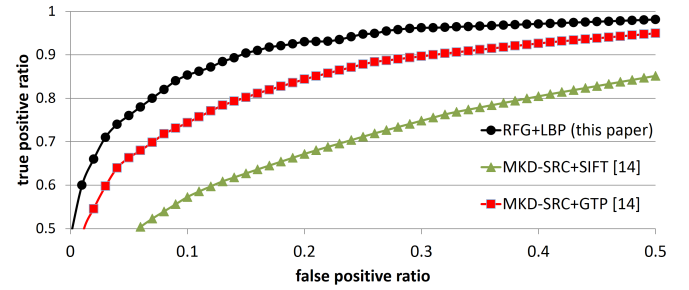

Fig. 11. Comparison with MKD-SRC [14] on LFW [19].

feature descriptors (*e.g.*, Local Quantized Patterns (LQP) [18], Learning-based (LE) descriptor [23]) to further improve the performance.

Figure 11 compares the proposed RFG with the multi-keypoint descriptor sparse representation-based classification method (MKD-SRC) [14]. MKD-SRC is a recent paper that proposes an alignment free approach for practical face recognition. For RFG, LBP features are used with parameters similar to those used for Figure 9. Results for MKD-SRC are reported using two features; Gabor Ternary Patterns (GTP) [14], and SIFT descriptors [53]. RFG performs significantly better than MKD-SRC even though standard LBP features are used.

• **Comparison with 3D pose normalization [52]:** We compare the proposed method with the 3D pose normalization method discussed in [52] for face identification on the FacePix and Multi-PIE [44] databases. Table IV demonstrates how RFG outperforms 3D pose normalization [52] in terms of rank-1 recognition rate on the FacePix database. The experiment setup is as follows. The gallery includes the frontal face image of all 30 subjects from the FacePix database. For the probe images, 180 images per subject are chosen with poses ranging from -90° to +90° in yaw angle. The 3D recognition system from [52] requires exact alignment of the face images, and it is limited to poses ranging from -45° to +45°. For fair comparison, Local Gabor Binary Pattern [41] descriptors are used for all methods in this experiment.

For close to frontal poses where the pose ranges from -30° to 30°, 3D pose normalization has better performance than RFG. This relies on the fact that reconstructing the frontal face image from a close to frontal face image is accurately performed via 3D pose normalization. For poses ranging from -90° to -31° and 31° to 90°, 3D pose normalization is either incapable of recognition or it results in a lower recognition rate than our proposed method. These results show that the proposed reference face graph based recognition algorithms have superior performance in recognition across pose. When the pose displacement is large, we observe that the 3D approach fails to compete with our proposed methods. More importantly, 3D pose normalization requires additional steps such as face boundary extraction and face alignment. The results by RB without centrality measures are also provided, it is observed that for different poses, RFG consistently outperforms RB.

Table V compares RFG with 3D pose normalization on the Multi-PIE database. The gallery and probes are chosen according to the experiments performed in [52]. Images that

TABLE IV
RANK-1 RECOGNITION RATES (%) ON FACEPIX [45]; 3D POSE NORMALIZATION [52] VS. RB AND RFG

Pose →	-90 to -46	-45 to -31	-30 to -16	-15 to -1	1 to 15	16 to 30	31 to 45	46 to 90	Avg
3D [52]	—	71.6	**90.0**	**97.3**	**95.8**	**92.7**	74.8	—	87.0
RB	70.6	78.9	84.7	90.1	89.9	87.2	81.3	73.2	82.0
RFG (this paper)	**75.0**	**82.9**	88.3	93.2	93.4	91.6	**85.5**	**78.0**	**89.2**

TABLE V
RANK-1 RECOGNITION RATES (%) ON MULTI-PIE [44]; 3D POSE NORMALIZATION [52] VS. RFG

Pose →	-45°	-30°	-15°	15°	30°	45°
3D [52]	74.1	91.0	95.7	95.7	89.5	74.8
RFG (this paper)	**86.4**	**91.2**	**96.0**	**96.1**	**90.9**	**85.4**

Fig. 12. FacePix [45] example image pairs. (a) image pairs correctly classified by RFG but incorrectly classified by DLC [24] (b) image pair incorrectly classified by both RFG and DLC [24].

are part of this experiment are excluded from the reference set to avoid training and testing on the same images. In addition, we ensure that the identities in the test set do not overlap with identities in the reference set, which is made up from both Multi-PIE and FEI databases. Thus, a fair comparison with [52] is made. The chosen probes contain six different poses from -45° and +45°. Similar to the previous experiment, LGBP descriptors are used as features. The results in Table V show that RFG outperforms 3D pose normalization for all six poses from the Multi-PIE database [44]. It is to be noted that there are recent methods that report better results on Multi-PIE database (*e.g.*, [54], [55]). However, these methods have to normalize the faces using *manually* labeled facial points, while alignment is not necessary for our method.

• **Comparison with Doppelgänger Lists [24]:** Table VI compares the proposed RFG algorithm with Doppelgänger list comparison introduced in [24] in terms of verification accuracy at equal error rate. FPLBP [30] is utilized as the feature descriptor for all methods. Similar to [24], ten test sets are chosen from the FacePix database [45], each including 500 positive and 500 negative pairs. Selected poses range from -90° to +90°. For direct FPLBP, the probe FPLBP descriptors are compared directly, and verification is solely based on direct comparison of FPLBP features. Note that, our proposed method does not either require alignment of the face images or any canonical coordinates or facial control points (eye, nose, *etc.*), whereas Doppelgänger list comparison [24] requires face image alignment.

The results in Table VI show that even only RB is used, the verification rate is higher than the competing methods. The proposed RFG algorithm outperforms Doppelgänger list comparison [24] by 10.6% without image alignment. Table VI also shows that similar to Doppelgänger list comparison, our proposed method is robust to pose variation. Figure 12 shows examples FacePix image pairs and if they are correctly or incorrectly classified.

• **Comparison with MRF model image matching [12]:** Table VII demonstrates how reference graph based face recognition performs under uneven illumination settings in addition to pose difference. For this reason, we select the probes as a subset of the CMU-PIE database [4] images with three illumination settings and three pose angles (frontal, profile, and 3/4 profile). The gallery consists of frontal face images of all 68 subjects of the CMU-PIE database. From the 21 images with different illumination settings available for each pose, we randomly choose 3 images to perform recognition and obtain the varying illumination results in Table VII. Uniform LBP features with a radius of 3 and 8 neighboring pixels are used as features [40].

The MRF model image matching [12] results in Table VII show that the rank-1 recognition rate decreases by 38.8% (79% to 40.2%) when illumination conditions change for the profile pose. Under the same conditions, we observe a rank-1 recognition rate decrease of only 1.3% (85.2% to 83.9%) for the proposed RFG algorithm. A similar recognition rate drop-off is also observed for the 3/4 profile pose. This shows how the proposed face recognition algorithm performs with little degradation under varying illumination settings.

• **Comparison with stereo matching [56] [57]:** The proposed RFG algorithm is compared with two stereo matching techniques for recognition across pose [56], [57]. Stereo matching provides a robust measure for face similarity which is pose-insensitive. Comparison is performed on the CMU PIE database [4]. For each of the 68 individuals in the PIE database, one image is randomly selected for the gallery and the remaining images as probes. In this experiment, the Stereo Matching Distance [56] and Slant Stereo Matching Distance [57] methods report 82.4% and 85.3% rank-1 recognition rate respectively, whereas the proposed RFG algorithm achieves 92.4% rank-1 recognition rate as shown in Table VIII.

Compared to the second best results by slant stereo matching distance [57], the proposed method reports a performance gain of over 8% in terms of rank-1 recognition rate on the CMU-PIE database. Using the stereo matching cost as a measure of similarity between two images with different pose and illumination settings is expensive, it requires detection of landmark points, and the cost of retrieval is high for large databases. On the contrary, the proposed approach does not require facial feature detection and is fast and efficient due to its integration with DCT hashing-based retrieval.

TABLE VI
VERIFICATION ACCURACY (%) OVER VARIOUS POSE RANGES ON FACEPIX [45]; DOPPELGÄNGER LISTS [24] VS. RB AND RFG

Probe 1	Probe 2	Direct FPLBP	DLC [24]	RB	RFG (this paper)
-30° - 30°	-30° - 30°	72.0 ± 2.0	74.5 ± 2.6	80.3 ± 0.9	**83.7** ± 0.8
-90° - 90°	-90° - 90°	54.7 ± 1.9	68.3 ± 1.0	75.2 ± 0.8	**78.9** ± 0.7
-10° - 10°	> |70°|	51.1 ± 1.7	66.9 ± 1.0	74.4 ± 0.8	**78.7** ± 0.7

TABLE VII
RANK-1 RECOGNITION RATE (%) COMPARISON UNDER NEUTRAL AND VARYING ILLUMINATION ON CMU-PIE [4]; MRF MODEL IMAGE MATCHING [12] VS RFG

Illum.→	Neutral		Varying		
Pose→	3/4 profile	Profile	3/4 profile	Profile	Frontal
MRF [12]	**98**	79	71.5	40.2	95.6
RFG	96.9	**85.2**	**92.3**	**83.9**	**96.1**

TABLE VIII
RANK-1 RECOGNITION RATE COMPARISON ON CMU-PIE [4]; STEREO MATCHING [56], [57] VS RFG

Method	Accuracy
Stereo Matching Distance [56]	82.4%
Slant Stereo Matching Distance [57]	85.3%
RFG (this paper)	**92.4%**

D. Computational Cost

The proposed method incorporates DCT locality sensitive hashing [39] for efficient similarity computation. Experiments were performed on a laptop with Intel Core i7 CPU and 8GB of RAM. The time required to generate the reference face descriptor (online processing) for a given image is 0.01 second on average with non-optimized code. In [39] the results showed that the performance of DCT hashing is very close to linear scan. The time needed to retrieve a probe from a 40k size gallery using DCT hashing is about 10 *ms*, while linear scan takes about 9000 *ms*. In addition, the time for linear scan increases linearly with the size of the gallery. The time for DCT retrieval, however, is essentially constant. Please see [39] for more details and experimental results. The main limitation of the current approach is that batch-training is performed to build the reference face graph, which however needs to be performed offline and only once.

V. CONCLUSIONS AND FUTURE WORK

We proposed a novel reference face graph (RFG) based approach towards face recognition in real-world scenarios. The extensive empirical experiments on several publicly available databases demonstrate that the proposed method outperforms the state-of-the-art methods with similar feature descriptor types. The proposed approach does not require face alignment, and it is robust to changes in pose. Results on real-world data with additional complications such as expression, scale, and illumination suggest that the RFG based approach is also robust in these aspects. The proposed approach is scalable due to the integration of DCT hashing with the descriptive reference face graph that covers a variety of face images with different pose, expression, scale, and illumination.

Future work includes development of other effective methods for reference set selection. Another direction is to study feature transformation or selection methods to further improve the performance of the proposed reference face graph based face recognition.

REFERENCES

[1] "The New York Times, www.nytimes.com/2013/04/19/us/fbi-releases-video-of-boston-bombing-suspects.html."

[2] "FEI face database," http://www.fei.edu.br/~cet/facedatabase.html.

[3] P. Phillips, H. Moon, P. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 1997, pp. 137–143.

[4] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression (PIE) database," in *IEEE International Conference on Automatic Face and Gesture Recognition*, May 2002, pp. 46–51.

[5] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal on Computer Vision*, vol. 57, no. 2, pp. 137–154, May 2004.

[6] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635–1650, 2010.

[7] "Facevacs software developer," KitCognitec Systems GmbH, 2012. [Online]. Available: http://www.cognitec-systems.de

[8] L. Wiskott, J.-M. Fellous, N. Kuiger, 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, Jul 1997.

[9] O. Barkan, J. Weill, L. Wolf, and H. Aronowitz, "Fast high dimensional vector multiplication face recognition," in *IEEE International Conference on Computer Vision (ICCV)*, Dec 2013, pp. 1960–1967.

[10] Z. Lei, M. Pietikainen, and S. Li, "Learning discriminant face descriptor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 289–302, Feb 2014.

[11] D. Chen, X. Cao, F. Wen, and J. Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2013, pp. 3025–3032.

[12] S. Arashloo and J. Kittler, "Energy normalization for pose-invariant face recognition based on MRF model image matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 6, pp. 1274–1280, June 2011.

[13] P. Nagesh and B. Li, "A compressive sensing approach for expression-invariant face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1518–1525.

[14] S. Liao, A. K. Jain, and S. Z. Li, "Partial face recognition: Alignment-free approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 5, pp. 1193–1205, 2013.

[15] U. Prabhu, J. Heo, and M. Savvides, "Unconstrained pose-invariant face recognition using 3D generic elastic models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1952–1961, Oct. 2011.

[16] P. Li, Y. Fu, U. Mohammed, J. H. Elder, and S. J. Prince, "Probabilistic models for inference about identity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 144–157, 2012.

[17] H. Li, G. Hua, Z. Lin, J. Brandt, and J. Yang, "Probabilistic elastic matching for pose variant face verification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2013, pp. 3499–3506.

[18] S. U. Hussain, T. Napolon, and F. Jurie, "Face recognition using local quantized patterns," in *Proceedings of the British Machine Vision Conference*. BMVA Press, 2012, pp. 99.1–99.11.

[19] 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, Oct. 2007.

[20] S. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 627–639, 2007.

[21] C.-K. Hsieh, S.-H. Lai, and Y.-C. Chen, "Expression-invariant face recognition with constrained optical flow warping," *IEEE Transactions on Multimedia*, vol. 11, no. 4, pp. 600–610, 2009.

[22] N. Kumar, A. Berg, P. Belhumeur, and S. Nayar, "Attribute and simile classifiers for face verification," in *IEEE International Conference on Computer Vision*, Oct. 2009, pp. 365–372.

[23] Q. Yin, X. Tang, and J. Sun, "An associate-predict model for face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2011, pp. 497–504.

[24] F. Schroff, T. Treibitz, D. Kriegman, and S. Belongie, "Pose, illumination and expression invariant pairwise face-similarity measure via Doppelgänger list comparison," in *IEEEInternational Conference on Computer Vision (ICCV)*, Nov. 2011, pp. 2494–2501.

[25] W. Shen, B. Wang, Y. Wang, X. Bai, and L. J. Latecki, "Face identification using reference-based features with message passing model," *Neurocomputing*, vol. 99, pp. 339–346, Jan. 2013.

[26] D. Chen, X. Cao, L. Wang, F. Wen, and J. Sun, "Bayesian face revisited: A joint formulation," in *European Conference on Computer Vision*, 2012, vol. 7574, pp. 566–579.

[27] S. Shan, Y. Chang, W. Gao, B. Cao, and P. Yang, "Curse of misalignment in face recognition: problem and a novel mis-alignment learning solution," in *IEEE International Conference on Automatic Face and Gesture Recognition*, 2004, pp. 314–320.

[28] S. Yan, H. Wang, J. Liu, X. Tang, and T. Huang, "Misalignment robust face recognition," *IEEE Transactions on Image Processing*, vol. 19, pp. 1087–1096, 2010.

[29] Z. Cui, W. Li, D. Xu, S. Shan, and X. Chen, "Fusing robust face region descriptors via multiple metric learning for face recognition in the wild," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2013, pp. 3554–3561.

[30] L. Wolf, T. Hassner, and Y. Taigman, "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1978–1990, Oct. 2011.

[31] T. Berg and P. Belhumeur, "Tom-vs-pete classifiers and identity-preserving alignment for face verification," in *Proceedings of the British Machine Vision Conference*. BMVA Press, 2012, pp. 129.1–129.11.

[32] M. K. Müller, M. Tremer, C. Bodenstein, and R. P. Würtz, "Learning invariant face recognition from examples," *Neural Networks*, vol. 41, pp. 137 – 146, 2013.

[33] Y. Shan, H. Sawhney, and R. Kumar, "Vehicle identification between non-overlapping cameras without direct feature matching," in *IEEE International Conference on Computer Vision*, vol. 1, Oct. 2005, pp. 378–385.

[34] Y. Guo, Y. Shan, H. Sawhney, and R. Kumar, "PEET: Prototype embedding and embedding transition for matching vehicles over disparate viewpoints," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2007, pp. 1–8.

[35] N. Rasiwasia, P. Moreno, and N. Vasconcelos, "Bridging the gap: Query by semantic example," *IEEE Transactions on Multimedia*, Aug. 2007.

[36] J. Liu, B. Kuipers, and S. Savarese, "Recognizing human actions by attributes," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2011, pp. 3337–3344.

[37] R. Jarvis and E. Patrick, "Clustering using a similarity measure based on shared near neighbors," *IEEE Transactions on Computers*, vol. C-22, no. 11, pp. 1025–1034, Nov. 1973.

[38] Z. Cui, S. Shan, H. Zhang, S. Lao, and X. Chen, "Image sets alignment for video-based face recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2012, pp. 2626–2633.

[39] M. Kafai, K. Eshghi, and B. Bhanu, "Discrete cosine transform locality-sensitive hashes for face retrieval," *IEEE Transactions on Multimedia*, vol. 16, no. 4, pp. 1090–1103, June 2014.

[40] 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, Dec. 2006.

[41] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition," in *IEEE International Conference on Computer Vision*, vol. 1, Oct. 2005, pp. 786–791.

[42] T. Opsahl, F. Agneessens, and J. Skvoretz, "Node centrality in weighted networks: Generalizing degree and shortest paths," *Social Networks*, vol. 32, no. 3, pp. 245–251, 2010.

[43] A. Popov, *Genetic Algorithms for Optimization*, 2005. [Online]. Available: http://www.automatics.hit.bg

[44] R. Grossa, I. Matthewsa, J. Cohnb, T. Kanadea, and S. Bakerc, "Multi-PIE," *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010.

[45] D. Little, S. Krishna, J. Black, and S. Panchanathan, "A methodology for evaluating robustness of face recognition algorithms with respect to variations in pose angle and illumination angle," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, 2005, pp. 89–92.

[46] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2233–2246, Nov. 2012.

[47] Z. Cao, Q. Yin, X. Tang, and J. Sun, "Face recognition with learning-based descriptor," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2010, pp. 2707–2714.

[48] D. Chen, X. Cao, L. Wang, F. Wen, and J. Sun, "Bayesian face revisited: A joint formulation," in *European Conference on Computer Vision*, vol. 7574, no. 566-579, 2012.

[49] T. Berg and P. Belhumeur, "POOF: Part-based one-vs.-one features for fine-grained categorization, face verification, and attribute estimation," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2013, pp. 955–962.

[50] Y. Taigman and L. Wolf, "Leveraging billions of faces to overcome performance barriers in unconstrained face recognition," *ArXiv e-prints*, 2011.

[51] H. V. Nguyen and L. Bai, "Cosine similarity metric learning for face verification," in *Asian Conference on Computer Vision*, 2010, pp. 709–720.

[52] A. Asthana, T. Marks, M. Jones, K. Tieu, and M. Rohith, "Fully automatic pose-invariant face recognition via 3D pose normalization," in *IEEE International Conference on Computer Vision (ICCV)*, Nov. 2011, pp. 937–944.

[53] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal on Computer Vision*, vol. 60, pp. 99–110, 2004.

[54] X. Cai, C. Wang, B. Xiao, X. Chen, and J. Zhou, "Regularized latent least square regression for cross pose face recognition," in *International Joint Conference on Artificial Intelligence*, 2013, pp. 1247–1253.

[55] A. Li, S. Shan, and W. Gao, "Coupled bias-variance tradeoff for cross-pose face recognition," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 305–315, Jan. 2012.

[56] C. D. Castillo and D. W. Jacobs, "Using stereo matching with general epipolar geometry for 2D face recognition across pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pp. 2298–2304, Dec. 2009.

[57] C. D. Castillo and D. W. Jacobs, "Wide-baseline stereo for face recognition with large pose variation," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2011, pp. 537–544.
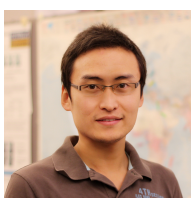
Mehran Kafai (S'11-M'13) is a research scientist at Hewlett Packard Laboratories in Palo Alto, California, USA. He received the M.Sc. degree in computer engineering from Sharif university of technology, Tehran, Iran, in 2005, the M.Sc. degree in computer science from San Francisco state university in 2009, and the PhD degree in computer science from the Center for Research in Intelligent Systems (CRIS), University of California, Riverside in 2013. His recent research has been concerned with secure computation, information retrieval, and big-data analysis.

Le An (S'13) received the B.Eng. degree in telecommunications engineering from Zhejiang University, Hangzhou, China in 2006 and the M.Sc. degree in electrical engineering from Eindhoven University of Technology, Eindhoven, The Netherlands in 2008. He is currently a PhD candidate in the Department of Electrical Engineering and a Graduate Student Researcher at the Center for Research in Intelligent Systems (CRIS), University of California, Riverside. His research interests include image processing, computer vision, pattern recognition, and machine learning. His current research focuses on face recognition, person re-identification, and facial expression recognition. He received the best paper award from the 2013 IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS).

Bir Bhanu (S'72-M'82-SM'87-F'95) received the S.M. and E.E. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA; the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA; and the M.B.A. degree from the University of California, Irvine, CA.

Dr. Bhanu is the Distinguished Professor of electrical and computer engineering and serves as the interim Chair of Bioengineering and the Founding Director of the Interdisciplinary Center for Research in Intelligent Systems at the University of California at Riverside (UCR). He was the founding Professor in the College of Engineering and served as the founding Chair of Electrical Engineering from 1991 to 1994. Since 1991, he has been the Director of Visualization and Intelligent Systems Laboratory, UCR. In addition, currently he serves as the Director of the National Science Foundation (NSF)Interdisciplinary Graduate Education, Research, and Training Program in video bioinformatics with UCR. Since 1991, 2006, and 2008, he has been a Cooperative Professor of computer science and engineering, bioengineering, and mechanical engineering, respectively, with UCR. He was a Senior Honeywell Fellow with Honeywell Inc., Minneapolis, MN, USA. He has been with the Faculty of the Computer Science, University of Utah, Salt Lake City, UT, USA, and with Ford Aerospace and Communications Corporation, Newport Beach, CA; French Institute for Research in Computer Science and Control (INRIA); and IBM San Jose Research Laboratory, San Jose, CA. He has been the Principal Investigator of various programs for the NSF, the Defense Advanced Research Projects Agency (DARPA), NASA, the Air Force Office of Scientific Research, the Office of Naval Research, the Army Research Office, and other agencies and industries in the areas of video networks, video understanding, video bioinformatics, learning and vision, image understanding, pattern recognition, target recognition, biometrics, autonomous navigation, image databases, and machine-vision applications. He is the coauthor of the books *Computational Learning for Adaptive Computer Vision* (to be published), *Human Recognition at a Distance in Video* (Berlin, Germany: Springer-Verlag, 2011), *Human Ear Recognition by Computer* (Berlin, Germany: Springer-Verlag, 2008), *Evolutionary Synthesis of Pattern Recognition Systems* (Berlin, Germany: Springer-Verlag, 2005), *Computational Algorithms for Fingerprint Recognition* (Norwell, MA: Kluwer, 2004), *Genetic Learning for Adaptive Image Segmentation* (Norwell, MA: Kluwer, 1994), and *Qualitative Motion Understanding* (Norwell, MA: Kluwer, 1992). He is the coeditor of *Computer Vision Beyond the Visible Spectrum* (Berlin, Germany: Springer-Verlag, 2004), *Distributed Video Sensor Networks* (Berlin, Germany: Springer-Verlag, 2011), and *Multibiometrics for Human Identification* (Cambridge, U.K.: Cambridge University Press, 2011). He is the holder of 18 (five pending) U.S. and international patents. He has more than 470 reviewed technical publications, including over 125 journal papers and 44 book chapters.

Dr. Bhanu is a Fellow of the American Association for the Advancement of Science, the International Association of Pattern Recognition, and the International Society for Optical Engineering. He has served as the General Chair for the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), the IEEE Conference on Advanced Video and Signal-Based Surveillance, the Association for Computing Machinery/IEEE Conference on Distributed Smart Cameras, the DARPA Image Understanding Workshop, the IEEE Workshops on Applications of Computer Vision (founded in 1992, now Winter Applications of Computer Vision Conference), and the CVPR Workshops on Learning in Computer Vision and Pattern Recognition, Computer Vision Beyond the Visible Spectrum, and Multi-Modal Biometrics. He serves as the General Chair for the IEEE Winter Conference on Applications of Computer Vision (2015). He has been on the Editorial Board of various journals and has edited special issues of several IEEE publications, such as the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS−PART B:CYBERNETICS, IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE SENSORS AND IEEE COMPUTER. He served on the IEEE Fellow Committee from 2010−2012. He was the recipient of the Best Conference Papers and Outstanding Journal Paper Awards, and the Industrial and University Awards for Research Excellence, Outstanding Contributions,Team Efforts and UCR Doctoral/Dissertation Advisor/Mentor Award.