Matching Composite Sketches to Face Photos: A Component-Based Approach

Hu Han, Brendan F. Klare, *Member, IEEE*, Kathryn Bonnen, and Anil K. Jain, *Fellow, IEEE*

Abstract—The problem of automatically matching composite sketches to facial photographs is addressed in this paper. Previous research on sketch recognition focused on matching sketches drawn by professional artists who either looked directly at the subjects (viewed sketches) or used a verbal description of the subject's appearance as provided by an eyewitness (forensic sketches). Unlike sketches hand drawn by artists, composite sketches are synthesized using one of the several facial composite software systems available to law enforcement agencies. We propose a component-based representation (CBR) approach to measure the similarity between a composite sketch and mugshot photograph. Specifically, we first automatically detect facial landmarks in composite sketches and face photos using an active shape model (ASM). Features are then extracted for each facial component using multiscale local binary patterns (MLBPs), and per component similarity is calculated. Finally, the similarity scores obtained from individual facial components are fused together, yielding a similarity score between a composite sketch and a face photo. Matching performance is further improved by filtering the large gallery of mugshot images using gender information. Experimental results on matching 123 composite sketches against two galleries with 10,123 and 1,316 mugshots show that the proposed method achieves promising performance (rank-100 accuracies of 77.2% and 89.4%, respectively) compared to a leading commercial face recognition system (rank-100 accuracies of 22.8% and 52.0%) and densely sampled MLBP on holistic faces (rank-100 accuracies of 27.6% and 10.6%). We believe our prototype system will be of great value to law enforcement agencies in apprehending suspects in a timely fashion.

Index Terms—Component-based face representation, composite sketch, face recognition, forensic sketch, heterogeneous face recognition, modality gap.

I. INTRODUCTION

AUTOMATED face recognition has made dramatic progress over the past decade [16]. Helping to determine the identity of criminals is an important application of face

H. Han and K. Bonnen are with the Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824 USA (e-mail: hhan@msu.edu; bonnenka@msu.edu).

B. F. Klare is with Noblis, Annandale, VA 22042 USA (e-mail: klarebre@msu.edu).

A. K. Jain is with the Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824 USA, and also with the Department of Brain and Cognitive Engineering, Korea University, Seoul 136-713, Republic of Korea (e-mail: jain@msu.edu).

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

recognition systems; however, in many cases the facial photograph of a suspect is not available. In these circumstances, drawing a sketch following the description provided by an eyewitness or the victim is a commonly used method to assist the police to identify possible suspects. Fig. 1(a), (b), (c) show three cases reported in the media where suspects were successfully identified through forensic sketches drawn by artists. In fact, criminal investigations have leveraged face sketches as far back as the late 19th century [1], [2]. The original, and still commonly used, paradigm for identifying subjects using facial sketches relies on human identification.[1]

Due to budgetary reasons, many law enforcement agencies use facial composite software, which allows the user to create a computer generated facial composite (composite sketch) [2], instead of employing forensic sketch artists. Examples of computer generated composite sketches can be found in Fig. 1(d), (e), (f), which shows three different cases in which suspects were identified based on computer generated composite sketches. However, while these examples illustrate cases in which human identification was successful, it is often the case that no suspect is identified based on computer generated composite sketches.

Much like its counterparts in forensic sketch recognition, the paradigm for identification using computer generated composites can be greatly expanded through automated face recognition technology. However, despite several methods for matching hand drawn sketches to photographs appearing in the literature (notably [3]–[7]), to our knowledge, there has been only limited research on matching computer generated composites to mugshot photographs [8]. This work attempts to develop a system that can successfully match composite sketches to mugshots. This would aid in quickly and accurately identifying suspects involved in criminal activities.

Because a skilled police sketch artist generally needs a large amount of training in both drawing and sculpting, facial composite software kits, which allow even nonartists to synthesize a sketch after only several hours of training, have become a popular alternative in criminal justice and law enforcement agencies. Some of the most widely used facial composite software kits include IdentiKit [38], Photo-Fit [9], FACES [37], Mac-a-Mug [9], and EvoFIT [10]. The first two kits (IdentiKit and Photo-Fit) are examples of early facial composite systems that synthesize a sketch by selecting a collection of facial components, e.g. hair, eyebrow, eyes, nose, mouth, shape, and eyeglasses. The next two (FACES and Mac-a-Mug) systems are modern component based facial composite systems which

---

[1]Human identification of sketches involves disseminating the sketches to the media in the hope that a citizen who recognizes the subject will come forward to provide a clue.
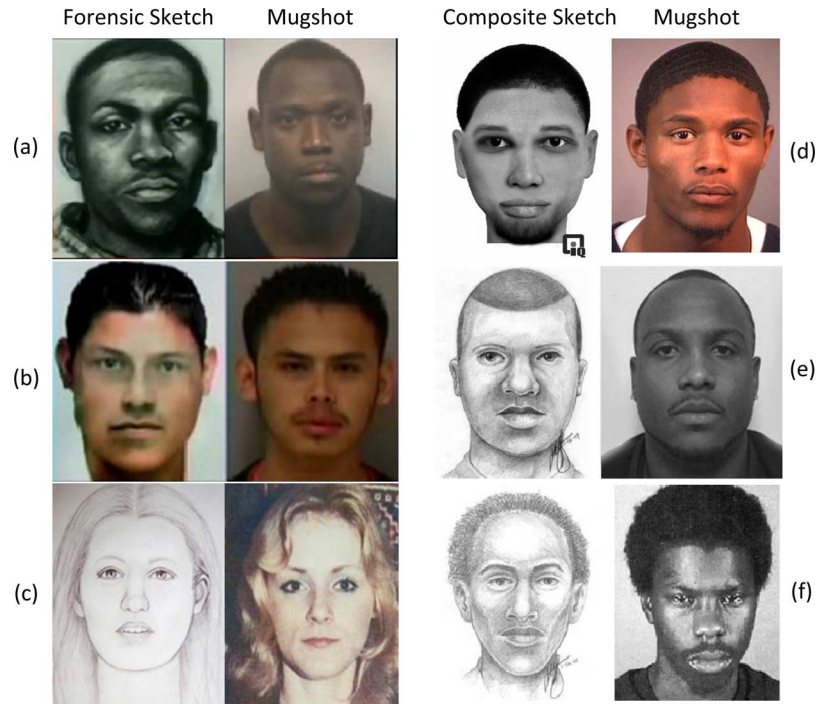
Fig. 1. Several cases recently reported in the media where the suspects were identified by using forensic sketches drawn by artists (a)–(c), and composite sketches created with facial composite software (d), (e). (a) Forensic sketch drawn based on descriptions provided by the female victims led to arrest of suspect in Virginia Highland armed robbery.[1] (b) A police sketch helped identify a man suspected of attacking a woman on a street.[2] (c) Forensic sketch by face reconstruction led to identification of the victim in a possible murder.[3] (d) Composite sketch created using FACES [37] aided in identifying the culprit of several armed robberies.[4] (e) Composite sketch created using SketchCop [44] led to arrest of a suspect that attacked several female drivers.[5] (f) Composite sketch created using SketchCop [44] led to arrest of an escaped suspect.[6]

[1] http://www.askaforensicartist.com/composite-sketch-leads-to-arrest-in-virginia-highland-robbery/

[2] http://www.woodtv.com/dpp/news/local/grand_rapids/Sketch-leads-to-mans-assault-arrest#

[3] http://www.askaforensicartist.com/sketch-by-texas-rangers-forensic-artist-leads-to-id/

[4] http://www.facesid.com/mediacenter_frontline_stories.html

[5] http://captured.sketchcop.com/2011sketchcopcaptured_010.htm

[6] http://captured.sketchcop.com/2011sketchcopcaptured_002.htm

include additional facial components (e.g., smile line, mole, scar, and tattoo), and have more candidate patterns for each facial component. The final system for drawing composite sketch (EvoFIT) is an evolutionary facial composite system, which is different from component based systems in that it creates a holistic likeness to the suspect using a genetic algorithm based on several selected shapes and textures that most resemble a suspect. We illustrate the procedure for creating a composite using IdentiKit and FACES in Fig. 2. In both these facial composite systems, each facial component is selected from a candidate list shown on the right side of the system GUI.

The difference between hand drawn sketches and computer generated composite sketches can be seen in Fig. 3. Compared to face photos, both hand drawn sketches and composite sketches lack detailed texture, especially around the forehead and cheeks. However, artists can depict each facial component with an exact shape, and even shading. Thus, artist-drawn sketches can usually capture the most distinctive characteristics of different faces. By contrast, facial components in a composite sketch must be approximated by the most similar component available in the composite software's database (see Fig. 2). Moreover, the psychological mechanism of an artist guarantees that hand drawn sketches look natural while composite sketches may look synthetic. Therefore, while we

may easily recognize a person from his hand-drawn sketch, it is often more challenging to identify a person from a composite sketch. Similar observations have also been reported in the cognitive psychology community [2], [11].

A survey in [2] showed that 80% of the officers in law enforcement agencies used computer generated composites. Despite this high percentage of law enforcement agencies using computer generated composites, the application of automated face recognition algorithms to computer generated composites has not been adequately studied [8]. By contrast, sketch recognition, both viewed sketches[2] and forensic sketches, has received relatively more attention. In this paper, we present a study to assess the ability of a face recognition system to match computer generated facial composites to facial photographs or mugshots. To facilitate this recognition task, we propose a new face recognition algorithm for matching composite sketches to face photos. The objectives of the proposed approach are to (i) provide a common representation for composite sketches and face photos that can diminish intraperson variations while still maintaining interperson discriminability, (ii) leverage the component-based approach by which computer generated

[2] A viewed sketch is a facial sketch drawn while viewing a photograph of the subject. While this scenario has little real world application because the photograph itself could be queried in the face recognition system, such research has proved quite valuable in improving automatic sketch recognition.
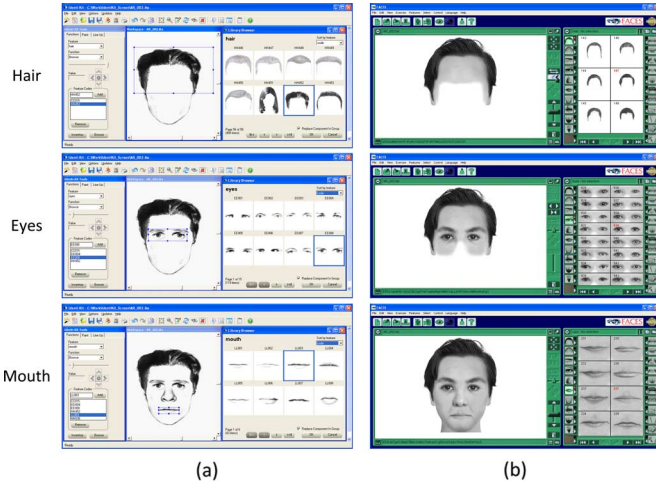
Fig. 2. Procedure for creating a composite sketch using (a) IdentiKit [38] and (b) FACES [37]. The three rows, respectively, show how the hair, eyes, and mouth components are selected from the candidate list.
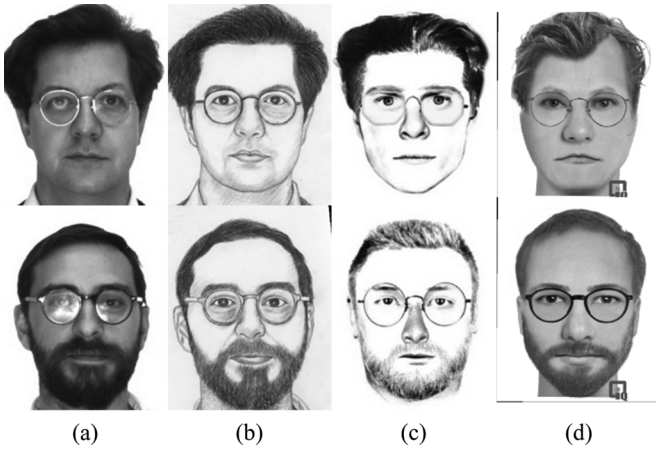

Fig. 3. Difference between hand drawn sketches and composite sketches. (a) Face photographs, (b) the corresponding viewed sketches, (c) the corresponding computer generated composite sketches synthesized using IdentiKit [38], and (d) using FACES [37]. Both the viewed sketches and composite sketches were constructed while viewing the face photographs.

composites are formed, and (iii) effectively match composite sketches against large-scale mugshot gallery databases.

The remainder of this paper is structured as follows. Section II introduces related work on hand drawn sketch recognition. The proposed component based representation (CBR) for composite sketch recognition is presented in Section III. In Section IV, experiments are performed to verify the effectiveness of the proposed CBR in matching composite sketches to large scale mugshot gallery sets. Finally, Section V summarizes our contributions and lists the directions of future work.

II. RELATED WORK

Compared to photo-based face recognition, there is only a limited amount of research on sketch recognition, and most of the published work is focused on hand drawn sketches. The initial research in this field focused on matching sketches that are drawn by an artist while looking at the corresponding photograph of the person or the person himself (called viewed sketches). These studies on viewed sketches can be grouped into

two categories: modal transformation and modal-insensitive feature representation. Approaches in the first category [3]–[6] convert images from one modality (e.g. sketch) into a different modality (e.g. photo). Methods for modal transformation include eigentransformation [3], [12], local linear embedding (LLE) [4], multiscale Markov Random Fields model [6], and embedded hidden Markov model (E-HMM) [15]. The merit of these approaches is that traditional face matching algorithms, designed for the target modality, can be used following the modal transformation. However, the synthesized photo (for example) can only be called a pseudo-photo due to its inferred content. In fact, these synthesis methods are often solving a more difficult problem than the recognition task.

The second approach to sketch recognition attempts to learn or design feature representations that reduce the intraclass difference caused by modality gap while preserving interclass separability. Representative methods in this category include common discriminant space [5], coupled spectral regression (CSR) [17], coupled information-theoretic projection (CITP) [13] and partial least squares (PLS) [14].

Klare *et al.* [7] demonstrated that state of the art face matchers can successfully match (with over 95% accuracy) viewed sketches (drawn by artists while viewing a person or photograph) to face photographs. Therefore, Klare *et al.* focused on the more challenging and realistic problem of matching forensic sketches (drawn by a forensic artist using only on a verbal description) to face photographs. Although both viewed and forensic sketches are drawn by an artist, the difference is that forensic sketches are drawn following the verbal description of an eyewitness or the victim, instead of looking at a person or photograph. During the drawing of a forensic sketch, the witness usually cannot exactly recall the facial appearance of a suspect. Additionally, it is often the case that a disparity exists between the understanding and depiction of facial features between an artist and the eyewitness. Thus, additional challenges are posed when matching forensic sketches against face photographs. To handle such difficulties, Klare *et al.* developed a local feature-based discriminant analysis (LFDA), which learns a discriminative representation from partitioned vectors of SIFT [42] and LBP [33] features using multiple discriminative subspace projections.

Component-based face recognition methods (which are used in this study) were studied in [18]–[20], [22], [23]; however, these algorithms either directly utilized intensity features that are sensitive to changes in facial appearance or employed supervised algorithms for classification whose performance is sensitive to the amount of training data available. Moreover, these algorithms were mainly proposed to resolve the misalignment problem in photo-to-photo face matching, and do not address the heterogeneous modality gap present when matching computer generated composite sketches to facial photographs.

Yuen and Man's [8] study on composite sketch recognition is the only previous work found to address this topic. A sketch matching algorithm using local and global features followed by relevance feedback from the users was proposed for recognizing composite sketches created using AICAMS [24]. Table I differentiates the proposed method and experimental protocol of this paper from those in [8].

TABLE I
FEATURES OF THE PROPOSED APPROACH FOR MATCHING COMPOSITE SKETCHES TO FACE PHOTOS

Comparison		Yuen and Man [8]	Proposed Approach
Facial composite system		AICAMS [24] with around 100 candidate patterns per component.	Identi-Kit [35] with around 100 candidate patterns per component. FACES [34] with around 1,000 candidate patterns per component. Both are commercial software that are publicly available.
Composite database		300 composite-mugshot pairs. Publicly availability is unknown.	123 composite-mugshot pairs for subjects from the public AR database. Composites are publicly available: http://biometrics.cse.msu.edu/pubs/databases .html.
Component localization		Manually locate high-curvature points.	Fully automatic
Face matching algorithm		LDA	Non-learning based method with MLBP; A leading commercial matcher - FaceVACS and holistic MLBP are used as the baseline.
Cross-race bias		Not studied	Studied with Asian and American users
Large scale background images		Not studied	Studied with 10,000 subjects from PCSO database and 1,193 mugshots from MEDS-II database.
Accuracy	Rank-1	~8% (Fig. 6 (c) in [8])	10.6% with PCSO background images 12.2% with MEDS background images 13.8% w/o additional background images
	Rank-50	~41% (Fig. 6 (c) in [8])	52.0% with PCSO background images 74.8% with MEDS background images 82.1% w/o additional background images

According to the relationship between Rank and Ra defined by [8, eq. (7)], Rank-1 and Rank-50, respectively, correspond to Ra = 0% and Ra = 16.4%.

III. COMPONENT BASED REPRESENTATION (CBR)

As discussed in Section I, most facial composite software systems used by law enforcement agencies are essentially component based systems. To facilitate the use of the system, these software kits provide a predefined set of candidate facial components. A composite sketch is constructed by individually selecting each facial component following the description of a witness or victim. The forehead and cheek areas are left blank, unless there are striking irregularities, such as a nevus (mole).

Inspired by the principle of facial composite systems and existing work on component-based methods in photo-to-photo matching, we propose a component based representation (CBR) for matching composite sketches to facial photographs. As illustrated in Fig. 4, the proposed approach consists of the following steps:

1) Face normalization using a geometric transformation and color space conversion.
2) Facial component localization using an active shape model (ASM), followed by component-wise alignment.
3) Per component feature extraction using block-based multiscale local binary patterns (MLBP).
4) Per component similarity measurement, followed by score normalization and fusion.

Each of the above steps is detailed in the following sections. We further improve the performance of CBR by making use of gender information to filter the gallery set.

A. Face Normalization

Although most photographs in the mugshot databases maintained by law enforcement agencies are captured with the cooperation of subjects, there are still some variations in face pose and scale. Most facial composite systems generate frontal sketches, but there are also variations in face scale.[3] To handle these variations in pose and scale, a geometric transformation (rotation and scaling) is first applied to both composite sketches and facial photographs based on the detected eye centers using PittPatt Face Recognition SDK [25]. The geometrically normalized face images are then cropped to the same size, which includes the whole face (including hair, chin and ears).

Color information may be available for gallery facial photographs, and recent research [29], [30] has shown that color information is helpful in photo-to-photo recognition. However, facial composites systems can seldom generate color sketches. Thus, the face normalization step of the proposed CBR method converts all color photographs (in our case, police mugshots) into gray scale images.

B. Facial Component Localization

In photo-to-photo matching tasks, to compensate for large facial pose variations, complicated facial component detectors have been proposed by training component specific SVM classifiers [18], [19], [23]. However, mugshot photographs are usu-

---

[3]A small number of mugshots may include slightly nonuniform illumination. We would like to study this problem in our future work using illumination pre-processing based methods [26]–[28].
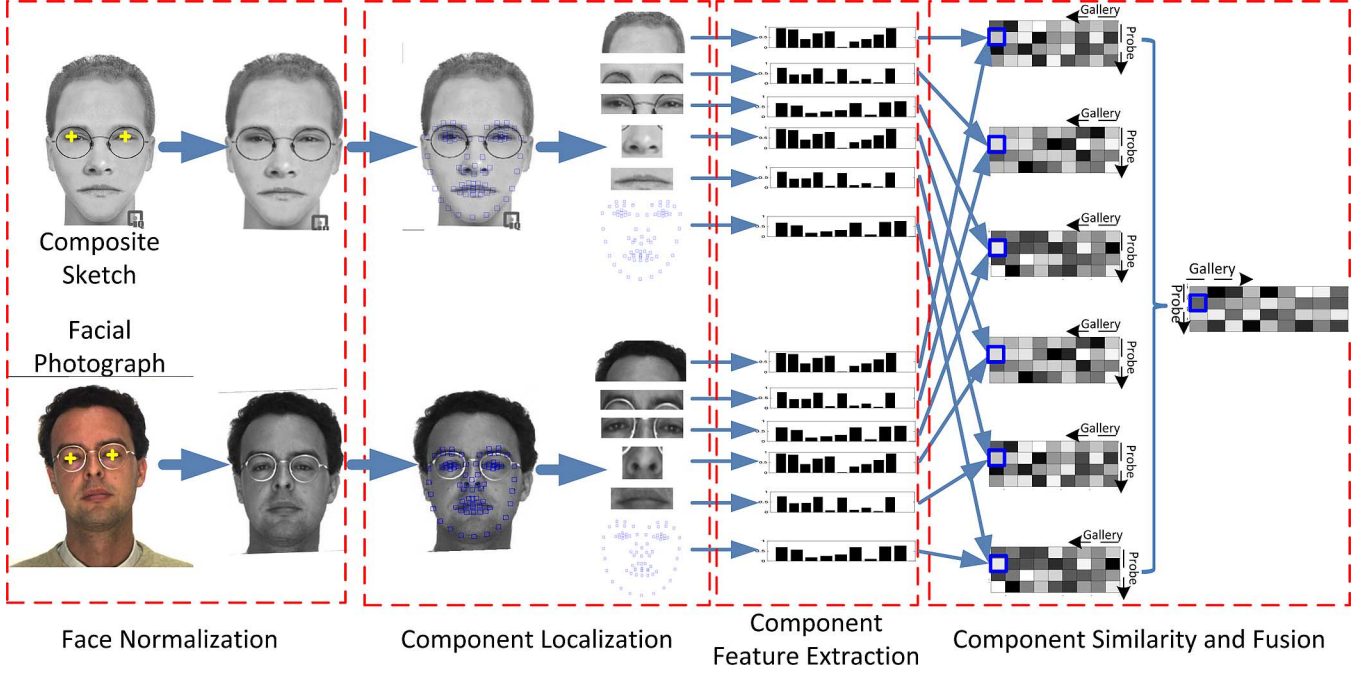
Fig. 4.   Overview of the component-based representation approach for matching composite sketches to face photos.
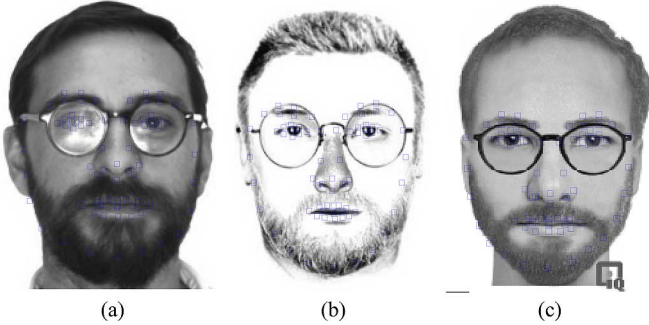

Fig. 5.   Facial landmark detection with ASM. Facial landmarks detected on (a) facial photographs, and the corresponding composite sketches synthesized using (b) IdentiKit and (c) FACES.

ally captured with the cooperation of the subjects, and the composite sketches are typically synthesized in frontal pose. The pose and scale variations between photographs and composites can be effectively eliminated with the above mentioned face normalization step. Therefore, in the proposed method, a more efficient approach is utilized for localizing the facial components from normalized frontal photographs and composites. An Active Shape Model (ASM) [31] is first utilized to detect a set of facial landmarks on the face, and then each component is localized according to the predefined keypoints in the ASM model. Specifically, we use the Stasm [32] open source ASM implementation, which detects 76 predefined facial landmarks. As shown in Fig. 5, even though Stasm was not trained with composite sketches, facial landmarks can still be accurately detected on composite sketches. This reveals that although composite sketches are a different modality than facial photographs, composite sketches still possess similar gradient information.

Once the facial landmarks are detected, different facial components can be localized according to the corresponding

landmarks in ASM model. Typically, a composite sketch is constructed by selecting the hair style, eyebrows, eyes, nose, mouth, and shape. In the proposed CBR, we follow the same face decomposition, and divide a face into five components (See Fig. 6). We also employ the face shape information for face matching by using vectorized coordinates of facial landmarks detected with the ASM. As shown in Fig. 6, each facial component is localized with a rectangle. We determine the horizontal boundaries of each rectangle based on the x coordinates of the left-most and right-most keypoints, and determine its vertical boundaries by maintaining a fixed aspect ratio across all subjects.

Considering that the hair component has a simple structure but is relatively large, we scale down the hair component in order to reduce the computational cost of feature extraction. We increase the size of components with lower height to width ratios (the eyebrows, eyes, and mouth) so that we can apply a block-based feature extraction method to these components. Fig. 7(a) and (b) show the average components calculated from 10,123 facial photographs and 123 composite sketches, respectively. The sharpness of the average components indicates accurate alignment of each facial component. Another benefit of per-component scaling is to counter the diversity in relative component sizes. For example, a facial component in a composite sketch may have the same shape as that in a photograph, but its relative size against the whole face is different from that in a photograph.

C. Component Feature Extraction

After per-component alignment and scaling, each component in the photographs is observed to be similar in structure. However, when looking closely at each component in a photograph and the corresponding composite sketch, we can find that there
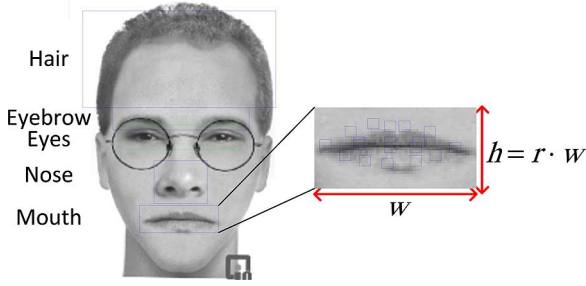
Fig. 6.   Each facial component is localized with a rectangle. The width of the rectangle w is determined by the left-most and right-most keypoints, while the height of the rectangle h is determined by a fixed ratio r.

Fig. 8.   Robustness of different local feature descriptors. (a) The same facial region in a composite sketch (blue rectangle) and the corresponding photograph (red rectangle); (b) feature representation for the regions in (a) with LBP; (c) feature representation for the regions in (a) with MLBP.
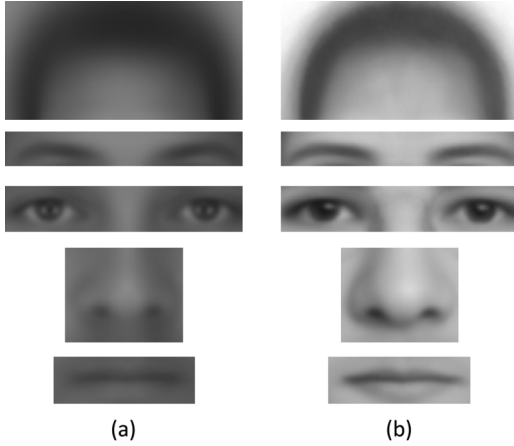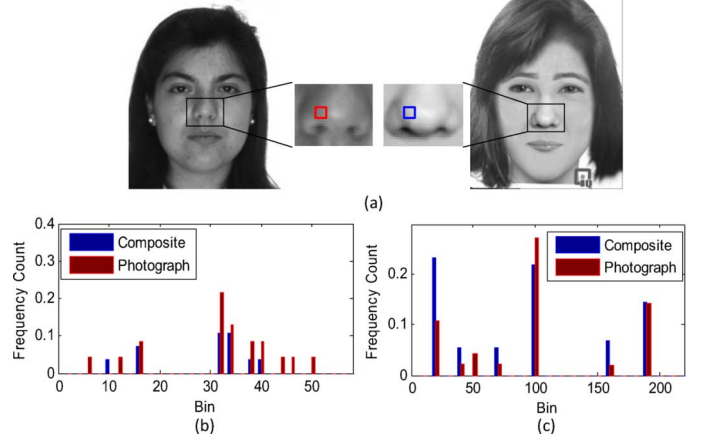

Fig. 7.   Average facial components calculated from (a) 10,123 photographs, and (b) 123 composite sketches synthesized using FACES.

is a discrepancy in component intensity. The intensity variations in facial photographs are due to shading of faces, and thus make the faces in photographs look more like 3-D shapes (see Fig. 3(a)). However, both hand drawn sketches and computer generated composite sketches lack shading information (see Fig. 3(b), (c), and (d)). We use the gray-scale invariant local binary pattern (LBP) [33] descriptor, which allows use to robustly match corresponding facial components under the intensity discrepancy between composite sketches and facial photographs.

LBP is a local texture descriptor that has been successfully applied in a wide range of applications, including texture classification [33], object detection [34], and object recognition [35]. The LBP code of a pixel is calculated by thresholding neighboring pixel intensities with the intensity of the center pixel. Formally, the LBP code for pixel (x, y) in image I can be calculated as

$$LBP_{P,R}(x,y) = \sum_{p=0}^{P-1} 2^p \, sign\left(I(x_p, y_p) - I(x,y)\right) \qquad (1)$$

where (x_p, y_p) is one of the P sampled neighboring pixels of pixel (x, y), the distance from (x_p, y_p) to (x, y) is no larger than R pixels and $sign(\cdot)$ function is defined as

$$sign(d) = \begin{cases} 1 & \text{if } d \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

Many facial structures or components are observable at different scales; therefore, LBP defined at a single scale is generally insufficient to capture all features of different facial components. Thus, multiscale LBP (MLBP) (a concatenation of LBP histograms at different radii $R \in \{1, 3, 5, 7\}$) is introduced in [7] to capture features at various scales in different facial components. Fig. 8 shows an example where LBP and MLBP are applied to extract features for the same facial region in a composite sketch and the corresponding photograph. This figure indicates that MLBP is more robust in matching two corresponding facial regions from a composite sketch and the corresponding photograph.

MLBP can effectively extract local texture features for each facial component; however, if a single MLBP descriptor is extracted from the entire facial component, the spatial structure of the component, which is also important for matching composite sketches to facial photographs, will be lost. Thus, to roughly capture the spatial structure information in each facial component, we further decompose each facial component into overlapped blocks (see Fig. 9).

D. Component Similarity and Fusion

With overlapped block decomposition, a facial component is divided into n patches. Let A_i and B_i be the photo patch and the corresponding composite patch, respectively (see Fig. 9). With MLBP features extracted for A_i and B_i, the similarity S_i between A_i and B_i is calculated based on the normalized histogram intersection. Specifically, S_i is calculated as

$$S_i = \frac{\sum_{j=1}^{D} \min\left(H_{A_i}^j, H_{B_i}^j\right)}{\min\left(\sum_{j=1}^{D} H_{A_i}^j, \sum_{j=1}^{D} H_{B_i}^j\right)} \qquad (3)$$

where H_{A_i} and H_{B_i} are D-dimensional MLBP histograms that are respectively extracted from A_i and B_i. Next, the similarities for all corresponding blocks are summed up to obtain the total component similarity S_{Nose} (for example). The similarities for the other facial components (hair, eyebrow, eyes and mouth) are
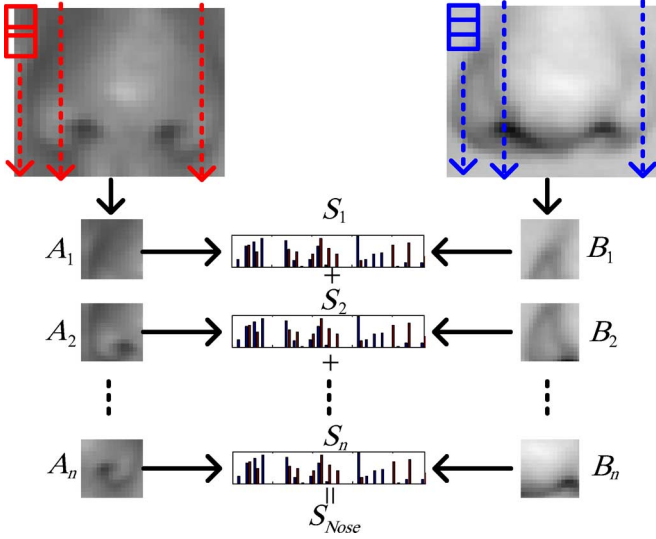
Fig. 9. Similarity calculation for each facial component. Similarities are first calculated for corresponding blocks, and then summed up to get the component similarity.

also calculated in the same manner. We also include the facial shape similarity by calculating the cosine distance between two facial landmark vectors. Finally, all the component similarities are again summed up to obtain the final matching score between a facial photograph and a composite sketch. However, the similarities calculated for different facial components need not be in the same numerical scale (range). Therefore, score normalization is applied to the match scores from each facial component prior to combining them. Tanh and z-score are the two most preferred score normalization methods in biometrics [41]. While both these normalization schemes had similar performance, the Tanh score normalization achieved a slightly higher accuracy over z-score normalization.

E. Gender Information Integration

Mugshot databases managed by law enforcement agencies generally include demographic information such as the gender, race, birthdate, height, weight, etc. From the view point of an eyewitness or the victim, it may be difficult for him/her to accurately specify the race or estimate the age, height and weight for the suspect. However, most eyewitnesses can provide the gender of the suspect. Thus, in practical applications, it is feasible to utilize gender information to improve the matching performance. Specifically, in the proposed method, gender information is used to filter the gallery set so that only composite faces and mugshots of similar gender are matched.

IV. EXPERIMENTAL RESULTS

In this section, we study the performance of the proposed CBR approach for matching composite sketches to facial photographs. A photo-composite database is first built for this experimental study. We analyze the discriminability of the different facial components for composite-to-photo matching. The accuracy of the proposed CBR method is then evaluated and compared against (i) a state-of-the-art, commercial-off-the-shelf (COTS), face recognition system, and

(ii) densely sampled MLBP on holistic faces. Finally, the diversity in composite sketch formed by two different users is analyzed.

A. Composite Sketch Database

We built a composite sketch database for our experimental study. It includes 123 subjects from the AR public domain face database [36]. We chose to use the AR database because it has been used in several previous studies on viewed sketches [6], [7], [14]. We synthesized composite sketches for each subject using two commercial facial composite systems: (i) FACES [37], and (ii) IdentiKit [38]. Thus, we have two composite sketches for each of the 123 subjects in the AR database. For each subject, there is also a photograph taken in frontal pose under ideal lighting conditions with a neutral expression. The composite sketches were created while viewing the facial photographs using the two different facial composite systems. As shown in Fig. 2, each facial component is selected from the list provided by the facial composite software. Examples of the composite sketches used in our experiments can be seen in Fig. 3, 4, and 5.

In our experiments, the composite sketches are used as the probe set. For the gallery set, in addition to the 123 photographs from the AR database, mugshots from two additional databases are used to populate the gallery. Specifically, 10,000 mugshot photographs (one image per subject) from the Pinellas County Sheriff's Office (PCSO) database[4] and 1,193 mugshots with prelabeled landmarks from Multiple Encounter Dataset II (MEDS-II) database [39][5] are, respectively, used to populate the gallery set. Thus, the matching experiments should closely replicate real-world scenarios where law enforcement agencies perform retrieval using a large scale mugshot database to find potential suspects by inputting a composite sketch.

We manually labeled the gender of the 123 subjects in the AR database. The 10,000 subjects from the PCSO database and 1,193 images from MEDS-II were prelabeled with gender information. As shown in Table II, there are 70 male and 53 female subjects in the composite sketches and corresponding photographs of the AR database. For the mugshots from the PCSO database, there are 8,043 males and 1,957 females. For the mugshots in the MEDS-II database, there are 1,082 males and 111 females. Example mugshot images for PCSO and MEDS-II are shown in Fig. 10(a) and (b), respectively.

B. Evaluation Metrics

In the scenario of photo-to-photo matching under controlled environments, face images are usually captured following a strict protocol, e.g. normal lighting, frontal pose, natural expression, etc. For example, border crossing officers need to determine whether a person crossing the border is presenting his own passport by comparing his photograph with that in the passport. Thus, exact and efficient matching is of crucial importance for face recognition systems designed for these scenarios. However, when a law enforcement agency matches a composite sketch to mugshot gallery database, the first ranked

---

[4]PCSO is a mugshot database that is collected in the state of Florida, U.S.A.

[5]There are two data sets in the MEDS database: MEDS-I and MEDS-II; MEDS-I is a subset of MEDS-II. We have used MEDS-II in our experiments.

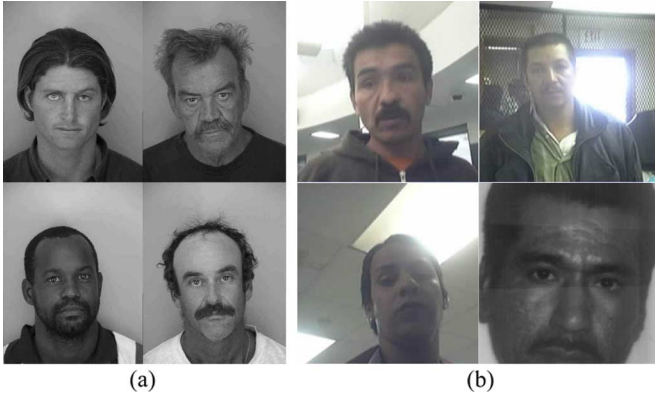| Dataset	Male		Female	
	# Images	Percentage	# Images	Percentage
AR composites and photographs	70	56.9%	53	43.1%
Mugshots in PCSO database	8,043	80.4%	1,957	19.6%
Mugshots in MEDS-II database	1,082	90.7%	111	9.3%

Fig. 10. Example mugshot photographs from (a) PCSO and (b) MEDS-II databases.

mugshot match is often not the suspect that the eyewitness or victim described. This is due to the challenge an eyewitness or victim faces in recalling and describing the facial details of the suspect. Therefore, forensic face recognition scenarios generally involve an examination by the eye witness or detective of the top hundred or so retrieved mugshots [45]. Thus, computer generated composite recognition methods are best evaluated by examining the rank retrieval rates using a Cumulative Match Characteristic (CMC) curve. While high rank-1 accuracy is ideal, success in these forensic recognition scenarios is generally measured around rank-100 accuracy [45].

The accuracy of the proposed CBR method is compared against a state-of-the-art COTS face recognition system, Face-VACS [40] and densely sampled MLBP on holistic faces. FaceVACS is representative of the recognition accuracies law enforcement agencies can expect from their current face recognition systems when queried with composite sketches, and thus serves as a faithful baseline method.

C. Discriminative Ability of Different Facial Components

Our first experiment evaluates the proposed approach by investigating the discriminative ability of different facial components. In this experiment, 10,000 mugshots from the PCSO database are first used to populate the gallery. Both the composite sketches synthesized using FACES and those generated using IdentiKit are used in this experiment. The similarities between 123 composite sketches and 10,123 mugshot photographs are calculated for all the six facial components (hair, eyebrows, eyes, nose, mouth and shape), and the matching rates of individual facial components are shown in Fig. 11. A comparison between the two component sketch kits FACES (Fig. 11(a)) and IdentiKit (Fig. 11(b)) shows that FACES provides a more accurate depiction for the different faces in the AR database. Thus,

in the following experiments, we focus on the recognition of composite sketches generated using only FACES.

From Fig. 11(a), it can be noticed that all the six facial components achieve fairly low rank-1 matching performance. The rank-1 matching rates of different components are below 5%. This reveals the difficulties in exactly depicting a face by using a facial composite system, and the challenges in matching a composite sketch to its true facial photograph. However, there is still a disparity in the discriminability of different components, especially when we compare the matching rates at higher ranks (e.g., rank-100). Among all the six facial components, the eyebrow component appears to be the most discriminative for matching composites to photos. Nose, hair and mouth components show higher discriminability than facial shape and eye components. In photo-to-photo matching, hair is seldom used due to its lack of persistence over time lapse; however, for composite-to-photo matching, any available clue should be utilized to determine the identity of the suspect. It should be understandable that the face shape is less discriminative, as only the coordinates of 76 landmarks detected with ASM are used to represent it. The finding that the eyes are the least discriminative component (rank-200 accuracy of only 6.5%) is contrary to the findings in photo-to-photo matching [20], [21]. As shown in [21], both eyebrows and eyes are among the most important components for photograph based face recognition.

Our finding here is understandable if we look closely at how a composite is synthesized. In the component menu (database) of composite software, it is likely that there are very similar patterns for the simple components, e.g. eyebrows, nose, mouth, and hairstyle. However, the pattern/structure of the eyes is more complicated (iris, sclera, white, pupil, eyelid, eyelash) than the other components. This makes it more difficult to replicate these fine structures in the template database available in the composite software. As a result, the eyes in a composite sketch are generally the least discriminative for composite-to-photo matching. In our experiments, users of the facial composite software also reported that it is quite difficult for them to choose the eye component when they created the composite sketches.

D. Performance Evaluation

We apply a score level fusion to the component-wise similarities to calculate the final matching score between a facial photograph and a composite sketch. As detailed in Section III-D, scores of different facial components are transformed into a common domain with Tanh normalization, and then summed to yield the final similarity score between a composite sketch and photograph. Due to the disparity in the discriminability between different components, we tried a sum of score fusion of the top M (2, 3, 4 and 5) most discriminative components in addition to the fusion of all components. As shown in Fig. 12, fusion of the four (i.e., $M = 4$) most discriminative facial components (eyebrow, nose, hair and mouth) achieves the best matching performance. Using only these four components, the matching rate at rank-200 increased from 37% for the best component (eyebrow) to 73%. Thus, in all subsequent experiments, we use the fusion of these four most discriminative components as our final matching score.
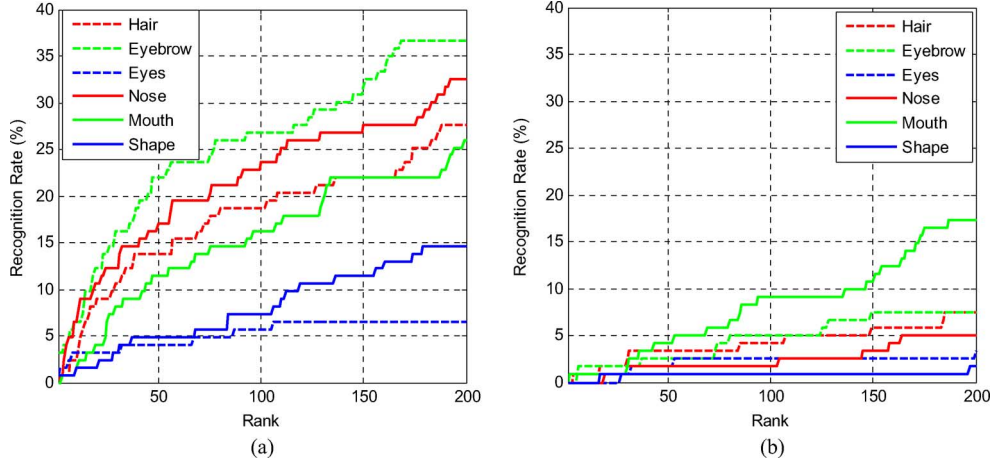
Fig. 11. Discriminability of different facial components for matching composite sketches to facial photographs with 10,000 background mugshots from the PCSO database. Composites sketches are respectively synthesized using commercial facial composite systems: (a) FACES and (b) IdentiKit.
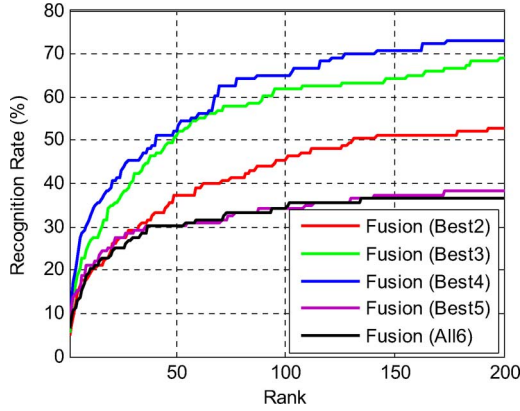

Fig. 12. Score level fusion of component similarities is applied by considering the top M ($M = 2, 3, \ldots, 6$) most discriminative facial components in matching composites to photographs with 10,000 background mugshots from the PCSO database.

The proposed approach is compared with the COTS matcher FaceVACS and densely sampled MLBP on holistic faces (Holistic MLBP) by using enlarged gallery sets with mugshots from the PCSO and MEDS-II databases. FaceVACS has been reported to achieve recognition accuracies on viewed sketches that is comparable to the leading viewed sketch recognition algorithms proposed in the literature [7]. Densely sampled MLBP on holistic faces has also been widely used in photo-to-photo matching [33]. For densely sampled MLBP, we use the same face image size, patch size and radii as the proposed approach. The matching performance of the three methods (proposed, FaceVACS and MLBP on holistic faces) on the computer generated composites is reported in Fig. 13. As shown in Fig. 13(a), the proposed CBR significantly outperforms the commercial matcher FaceVACS and holistic MLBP when the gallery is enlarged by 10,000 mugshots from the PCSO database. We also provide the matching accuracies at several retrieval ranks in Table III. The proposed CBR achieves 8% and 42.2% higher matching rates at rank-1 and rank-100, respectively, than those achieved by FaceVACS. Compared with holistic MLBP, the proposed CBR achieves 6.5% and 37.4% higher matching rates

at rank-1 and rank-100, respectively. When 1,193 mugshots from public MEDS-II database are utilized to populate the gallery set, the proposed approach is again found to be much more effective for matching composite sketching to mugshots (see Fig. 13(b) and Table IV). Both the proposed approach and FaceVACS achieve higher performance, compared to the performance on the larger gallery set populated with mugshots from the PCSO database; however, there is performance degradation for holistic MLBP. We attribute this to the sensitivity of holistic MLBP to the generally lower image quality in the MEDS-II database. As shown in Fig. 10(b), MEDS-II database contains blurred images due to the movement of camera or subjects, low contrast images due to backlighting, and out of image facial regions. Additionally, some mugshot images in MEDS-II database are of low resolutions, e.g. 240 × 240. By contrast, the mugshots in the PCSO database are clear (see Fig. 10(a)), and all face images are with the resolutions of 624 × 486.

In the above experiments, mugshots from two different databases (PCSO and MEDS-II) were used to populate the gallery set. One concern regarding the populated gallery is the different characteristics of face images in AR, PCSO, and MEDS-II databases. As a result, if face images in the AR database contain any significant differences from the mugshots in PCSO or MEDS-II, the populated gallery could not make the identification task more challenging. This issue would be of serious concern, especially if the task addressed is photo-to-photo matching. To appease this concern, we also perform composite sketch recognition without using mugshots to populate the gallery. The performance of the proposed approach, FaceVACS, and holistic MLBP is shown in Fig. 14. From Fig. 14, we can notice that all the three methods achieve much higher performance without using the mugshots from PCSO or MEDS-II databases to enlarge the gallery.[6] This experiment reveals that including the mugshots from PCSO or

---

[6]We should point out that if the mugshots from MEDS-II cannot make the matching task more challenging, the performance of FaceVACS and holistic MLBP should not drop dramatically; however, as shown in Fig. 14(b) and (c), mugshots from MEDS-II greatly reduce the performance of FaceVACS and holistic MLBP.
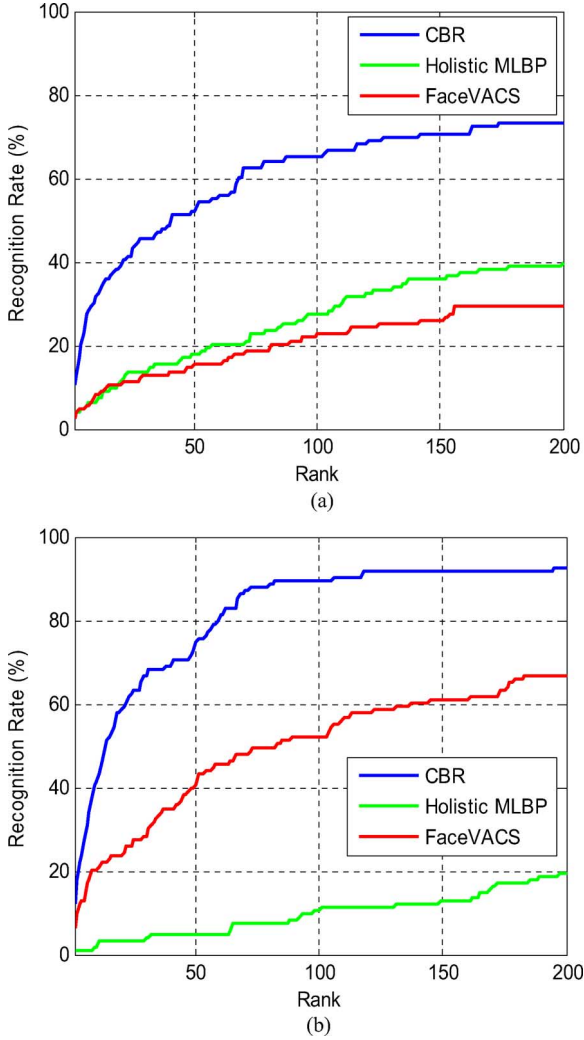
Fig. 13.   Matching performance of the proposed CBR, FaceVACS, and holistic MLBP on enlarged gallery set with mugshots from (a) PCSO and (b) MEDS-II databases.

TABLE III
MATCHING PERFORMANCE OF THE PROPOSED CBR, FACEVACS, AND HOLISTIC MLBP AT SEVERAL SPECIFIC RETRIEVAL RANKS ON ENLARGED GALLERY SET WITH MUGSHOTS FROM THE PCSO DATABASE

Matcher	Matching Rate (%) at Rank-N				
	Rank-1	Rank-50	Rank-100	Rank-150	Rank-200
FaceVACS	2.4	15.4	22.8	26.0	29.3
Holistic MLBP	4.1	17.9	27.6	35.8	39.8
CBR	10.6	52.0	65.0	70.7	73.2

TABLE IV
MATCHING PERFORMANCE OF THE PROPOSED CBR, FACEVACS, AND HOLISTIC MLBP AT SEVERAL SPECIFIC RANKS ON ENLARGED GALLERY SET WITH MUGSHOTS FROM THE MEDS-II DATABASE

Matcher	Matching Rate (%) at Rank-N				
	Rank-1	Rank-50	Rank-100	Rank-150	Rank-200
FaceVACS	6.5	40.7	52.0	61.0	66.7
Holistic MLBP	0.8	4.9	10.6	13.0	19.5
CBR	12.2	74.8	89.4	91.9	92.7

MEDS-II database in the gallery does make it more challenging to match composites to photographs. As shown in Fig. 14(c),

both the large galley size (10,123 images) imposed by PCSO mugshots, and low image quality in MEDS-II database make it challenging for densely sampled MLBP on holistic faces to correctly match composites to photographs. Fig. 14(b) shows that FaceVACS is more robust to the low image quality in MEDS-II, but it is also challenging for FaceVACS to handle a large scale gallery set. By contrast, from Fig. 14(a), we find that the proposed approach is more robust to the challenges posed by the large gallery size and low image quality.

As discussed in Section III-E, gender information can be utilized to further improve the matching performance of the proposed CBR by filtering the mugshot photos in the gallery set. From Table II, it can be noticed that for a male suspect, the size of the enlarged gallery set with PCSO can be reduced from 10,123 to 8,113, and for a female suspect, the gallery size is reduced to 2,010. Similarly, for a male suspect, the size of the enlarged gallery set with MEDS-II can be reduced from 1,316 to 1,152, and for a female suspects, the gallery size is reduced from 1,316 to 164. The performance of the proposed CBR with and without gender information is show in Fig. 15. On the gallery set populated with mugshots from the PCSO database, there is nearly 10% improvement, on average, from rank-1 to rank-200. On the gallery set populated with mugshots from the MEDS-II database, there is about 5% improvement, on average, from rank-1 to rank-200. This demonstrates the importance of utilizing any demographic information the eyewitness or victim is able to provide.

Fig. 16 shows examples of correctly and incorrectly matched composite-photo pairs at rank-1 with the proposed CBR utilizing gender information. The first two columns in Fig. 16(a) and (b), respectively, show the composite sketches and the rank-1 matched photographs. The last column in Fig. 16(b) gives the ground-truth photographs corresponding to the composite sketches in the first column in Fig. 16(b). As to be expected for the proposed CBR, most facial components in the correctly matched composite-photo pairs do attain high similarities for all facial components. However, the incorrectly matched photographs in the second column of Fig. 16(b) also show high visual similarity with the composites sketches. For these composite sketches, it is even difficult for a human to tell which photos in the last two columns of Fig. 16(b) are the ground-truth. This illustrates one of the challenges in using facial composite software, even though users are able to choose the most similar pattern for each facial component, it is still difficult for them to guarantee accurate spatial relations among different facial components.

E. Composites Synthesized by Different Users

The quality of the composite sketches depends not only on the specific facial composite system used but also the skills and background of the users of the software. Therefore, it should be valuable to investigate composite sketches synthesized by different users. In the above experiments, all 123 composite sketches, corresponding to the 123 faces in the AR database, are synthesized using FACES by a Caucasian user (a U.S. graduate student in our lab) with research experience in face recognition. In this experiment, the same 123 faces were synthesized again with the same facial composite system but by an Asian user who
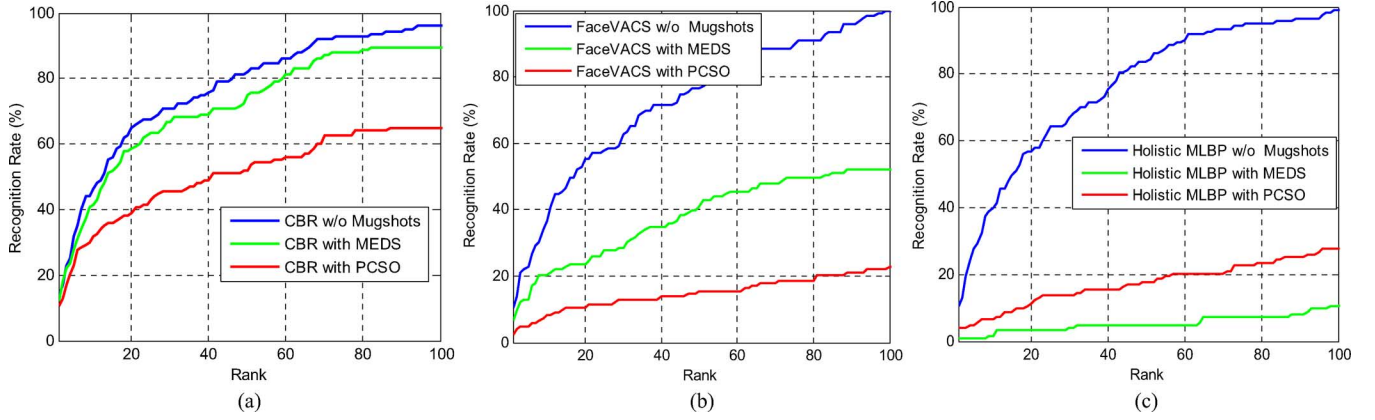
Fig. 14.   Matching performance of (a) the proposed CBR, (b) FaceVACS, and (c) holistic MLBP with and without using mugshots to populate the gallery set.
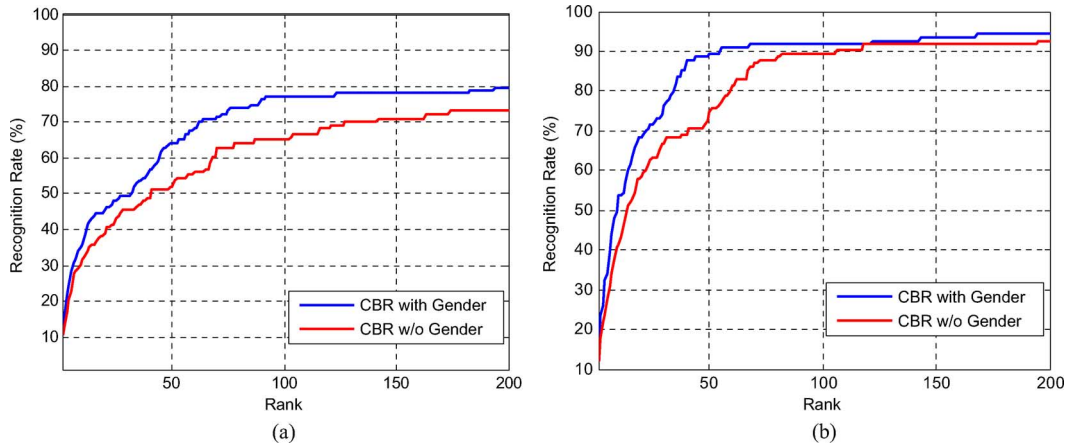

Fig. 15.   Performance of the proposed CBR with and without using gender information to filter the gallery set populated by mugshots from (a) PCSO and (b) MEDS-II databases.
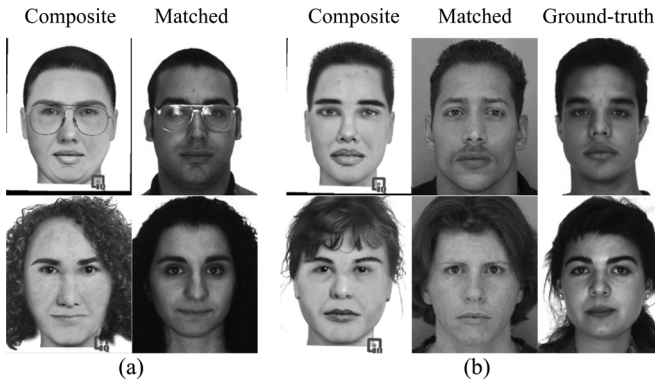

Fig. 16.   Examples of rank-1 retrievals for matching composite-photo pairs with the proposed CBR utilizing gender information. (a) Correctly matched composites-photo pairs; (b) incorrectly matched composites-photo pairs and the ground-truth photographs.
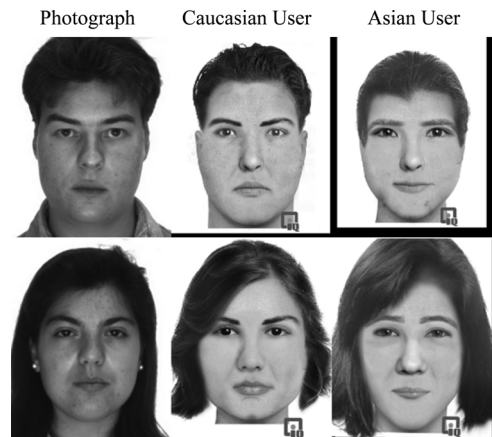

Fig. 17.   Two examples of composite sketches synthesized by two different users (with different cultural background) who used FACES to form the composite sketches.

also has research experience in face recognition but has been in the U.S. for only 6 months. Fig. 17 shows some photographs and the corresponding composite sketches synthesized by these two users. To a certain degree, the composite sketches created by the Asian user look like Asians. This reflects the "other-race effect" [46], which is the tendency for people of one race to have difficulty recognizing and processing faces and facial expressions

of members of a race or ethnic group other than their own [43]. Due to the cross-race bias, an Asian user tends to choose the facial components that he is familiar with, when he creates composite sketches for Caucasian subjects. It is difficult for him to determine the similarity of the facial components that he is not
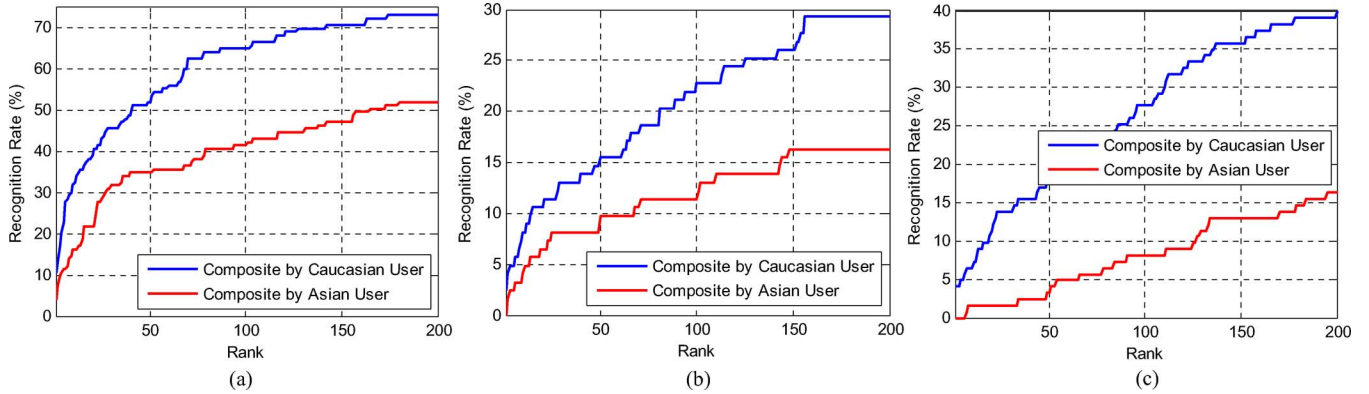
Fig. 18. Variations in composite sketches synthesized by two different users. Matching performance without using gender information of (a) the proposed CBR, (b) FaceVACS, and (c) holistic MLBP on composites that are, respectively, created using FACES by a Caucasian user and an Asian user.

familiar with. However, if an Asian user is asked to create composite sketches for Asian individuals, the Asian user should be expected to produce more favorable composites compared with a non-Asian user. Another issue is that the facial composite software that we have used here have been designed by Caucasian designers for use in non-Asian market. The available choices of facial components in the database of the facial composite system also reflect this bias.

The final experiment compares the accuracy of matching composite sketches synthesized by the two users to the mugshot gallery set, using both the proposed CBR method and Face-VACS. Matching performance of the proposed CBR and FaceVACS on the composite sketches synthesized by two different users is shown in Fig. 18. Our statistics shows that more than 95% of the 123 subjects in the AR database are non-Asian. As such, most composite sketches created for non-Asian faces by the Asian user are less discriminative. As expected, we notice that CBR (Fig. 18(a)), FaceVACS (Fig. 18(b)), and holistic MLBP (Fig. 18(c)) achieve dramatically higher performance on the composites created by the Caucasian user. Our experimental results suggest that the users of composite software in law enforcement agencies should be trained to reduce the cross-race bias in order to create high quality composite sketches for potential suspects of various races.

V. Conclusions and Future Work

This paper investigated a challenging heterogeneous face recognition problem: matching composite sketches to facial photographs. We studied composite sketches generated using two publicly available facial composite software, and proposed a component based representation (CBR) to address the modality gap between composite sketches and facial photographs. Facial components are automatically localized by detecting landmarks with an ASM. A local descriptor-based representation, MLBP, is then utilized to capture local texture and structure at various scales in each facial component. Block based feature extraction is employed to spatially encode the structural information of each facial component. This feature representation used with the proposed CBR is consistent with the process of how a computer generated facial composite sketch is synthesized in facial composite systems.

The proposed method is tested by matching composite sketches against a large scale mugshot gallery set. Our method significantly outperforms a state-of-the-art COTS face matcher. By filtering the gallery set with gender information, the performance of the proposed method is further improved. This suggests that face matcher vendors should also make use of any possible demographic information when their systems are applied for law enforcement. Analysis of the variations in the quality of composite sketches synthesized by two users of different ethnic and cultural background suggests that users of facial composite software should be trained to reduce the cross-race bias in order to create high quality composite sketches for potential suspects of various races.

Currently, the spatial relationship between different facial components in a composite sketch is encoded in the landmarks extracted using ASM. Additional correlations between the choices of different facial components in creating a composite sketch will be investigated in the future work. Also, we would like to further investigate the proposed approach by increasing the number of subjects in the composite sketch database,[7] and by considering additional facial composite systems to create composite sketches. Another challenge for matching composite sketches to mugshots is the age difference between a previously captured mugshot and the current face of a suspect. Our study in [47] shows that a component based representation is effective in achieving age-invariant photo-to-photo matching. In our future work, we also would like to study the aging problem in composite-to-photo matching using the proposed approach.

References

[1] K. T. Taylor, *Forensic Art and Illustration.* Boca Raton, FL: CRC Press, 2000.

[7]Specifically, we are generating sketches from additional databases such as MEDS and FERET to demonstrate the general applicability of our composite to photo matching approach.

[2] D. Mcquiston, L. Topp, and R. Malpass, "Use of facial composite systems in US law enforcement agencies," *Psychology, Crime and Law*, vol. 12, no. 5, pp. 505–517, 2006.

[3] X. Tang and X. Wang, "Face sketch recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 50–57, Jan. 2004.

[4] Q. Liu, X. Tang, H. Jin, H. Lu, and S. Ma, "A nonlinear approach for face sketch synthesis and recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2005, pp. 1005–1010.

[5] D. Lin and X. Tang, "Inter-modality face recognition," in *Proc. Eur. Conf. Computer Vision*, 2006, pp. 13–26.

[6] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 1955–1967, Nov. 2009.

[7] B. Klare, Z. Li, and A. Jain, "Matching forensic sketches to mug shot photos," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 3, pp. 639–646, Mar. 2011.

[8] P. C. Yuen and C. H. Man, "Human face image searching system using sketches," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 37, no. 4, pp. 493–504, Jul. 2007.

[9] G. Wells and L. Hasel, "Facial composite production by eyewitnesses," *Current Directions Psychol. Sci.*, vol. 16, no. 1, pp. 6–10, Feb. 2007.

[10] C. Frowd, P. Hancock, and D. Carson, EvoFIT: A Holistic, "Evolutionary facial imaging technique for creating composites," *ACM Trans. Appl. Psychol.*, vol. 1, no. 1, pp. 19–39, Jul. 2004.

[11] C. Frowd, D. Carson, H. Ness, D. McQuiston, J. Richardson, H. Baldwin, and P. Hancock, "Contemporary composite techniques: The impact of a forensically-relevant target delay," *Legal Criminol. Psychol.*, vol. 10, no. 1, pp. 63–81, Feb. 2005.

[12] X. Wang and X. Tang, "Hallucinating face by eigentransformation," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 35, no. 3, pp. 425–434, Aug. 2005.

[13] W. Zhang, X. Wang, and X. Tang, "Coupled information-theoretic encoding for face photo-sketch recognition," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2011, pp. 513–520.

[14] A. Sharma and D. Jacobs, "Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2011, pp. 593–600.

[15] X. Gao, J. Zhong, J. Li, and C. Tian, "Face sketch synthesis algorithm based on E-HMM and selective ensemble," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 4, pp. 487–496, Apr. 2008.

[16] *Handbook of Face Recognition*, S. Z. Li and A. K. Jain, Eds., 2nd ed. New York: Springer, 2011.

[17] Z. Lei and S. Li, "Coupled spectral regression for matching hegerogeneous faces," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009, pp. 1123–1128.

[18] B. Heisele, P. Ho, J. Wu, and T. Poggio, "Face recognition: Component-based versus global approaches," *Comput. Vis. Image Understand.*, vol. 91, no. 1, pp. 6–21, Jul. 2003.

[19] Y. Ivanov, B. Heisele, and T. Serre, "Using component features for face recognition," in *Proc. Automatic Face and Gesture Recognition*, 2004, pp. 421–426.

[20] D. Lin and X. Tang, "Recognize high resolution faces: From macrocosm to microcosm," in *Proc. IEEE Computer Vision and Pattern Recognition*, 2006, pp. 1355–1362.

[21] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell, "Face recognition by humans: Nineteen results all computer vision researchers should know about," *Proc. IEEE*, vol. 94, no. 11, pp. 1948–1962, Nov. 2006.

[22] B. Klare, A. Paulino, and A. Jain, "Analysis of facial features in identical twins," in *Proc. Int. Joint Conf. Biometrics*, 2011, pp. 1–8.

[23] J. Huang, V. Blanz, and B. Heisele, "Face recognition using component-based SVM classification and morphable models," in *Proc. First Int. Workshop on Pattern Recognition With Support Vector Machines*, 2002, pp. 334–341.

[24] J. W. Brahan, K. P. Lam, H. Chan, and W. Leung, "AICAMS—Artificial intelligence crime analysis and management system," *Knowledge-Based Syst.*, vol. 11, no. 5–6, pp. 355–361, 1998.

[25] PittPatt Face Recognition SDK, Pittsburgh Pattern Recognition [Online]. Available: http://www.pittpatt.com

[26] H. Han, S. Shan, X. Chen, and W. Gao, "Illumination transfer using homomorphic wavelet filtering and its application to light-insensitive face recognition," in *Proc. Automatic Face and Gesture Recognition*, 2008, pp. 1–6.

[27] H. Han, S. Shan, L. Qing, X. Chen, and W. Gao, "Lighting aware preprocessing for face recognition across varying illumination," in *Proc. Eur. Conf. Computer Vision*, 2010, pp. 308–321.

[28] H. Han, S. Shan, X. Chen, S. Lao, and W. Gao, "Separability oriented preprocessing for illumination-invariant face recognition," in *Proc. Eur. Conf. Computer Vision*, 2012, pp. 307–320.

[29] C. Jones and A. Abbott, "Color face recognition by hypercomplex gabor analysis," in *Proc. Automatic Face and Gesture Recognition*, 2006, pp. 126–131.

[30] J. Yang and C. Liu, "A general discriminant model for color face recognition," in *Proc. Int. Conf. Computer Vision*, 2007, pp. 1–6.

[31] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models-their training and application," *Comput. Vis. Image Understand.*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[32] S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," in *Proc. Eur. Conf. Computer Vision*, 2008, pp. 504–513.

[33] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[34] M. Heikkila and M. Pietikainen, "A texture-based method for modeling the background and detecting moving objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 657–662, Apr. 2006.

[35] M. Pietikainen, T. Nurmela, T. Maenpaa, and M. Turtinen, "View-based recognition of real-world textures," *Pattern Recognit.*, vol. 37, no. 2, pp. 313–323, Feb. 2004.

[36] A. M. Martinez and R. Benavente, The AR Face Database CVC Tech. Rep., 1998.

[37] FACES 4.0, IQ Biometrix 2011 [Online]. Available: http://www.iqbiometrix.com

[38] Identi-Kit, Identi-Kit Solutions 2011 [Online]. Available: http://www.identikit.net/

[39] National Institute of Standards and Technology (NIST), NIST Special Database 32—Multiple Encounter Dataset (MEDS) 2011 [Online]. Available: http://www.nist.gov/itl/iad/ig/sd32.cfm

[40] FaceVACS Software Developer Kit, Cognitec Systems GmbH 2010 [Online]. Available: http://www.cognitec-systems.de

[41] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognit.*, vol. 38, no. 12, pp. 2270–2285, Dec. 2005.

[42] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[43] C. Meissner and J. Brigham, "Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review," *Psychology, Public Policy, and Law*, vol. 7, no. 1, pp. 3–35, Mar. 2001.

[44] SketchCop, SketchCop Solutions 2011 [Online]. Available: http://sketchcop.com/

[45] A. K. Jain, B. Klare, and U. Park, "Face matching and retrieval in forensics applications," *IEEE Multimedia*, vol. 19, no. 1, p. 20, Jan. 2012.

[46] R. K. Bothwell, J. C. Brigham, and R. S. Malpass, "Cross-racial identification," *Personality Social Psychol. Bull.*, vol. 15, no. 1, pp. 19–25, 1989.

[47] C. Otto, H. Han, and A. K. Jain, "How does aging affect facial components?," in *Proc. ECCV WIAF Workshop*, Florence, Italy, Oct. 7–13, 2012, pp. 189–198.

Hu Han is a postdoctoral research fellow in the Department of Computer Science and Engineering at Michigan State University, East Lansing. He received the B.S. degree in computer science from the Shandong University, Jinan, China, in 2005, and the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2011. His research interests include computer vision, pattern recognition, and image processing, with applications to biometrics, forensics, law enforcement, and security systems.

Brendan F. Klare (S'09–M'11) is a lead scientist at Noblis. He received the B.S. and M.S. degrees in computer science from the University of South Florida in 2007 and 2008, and the Ph.D. degree in computer science from Michigan State University in 2012. From 2001 to 2005, he served as an airborne ranger infantryman in the 75th Ranger Regiment. He has authored several papers on the topic of face recognition, and he received the Honeywell Best Student Paper Award at the 2010 IEEE Conference on Biometrics: Theory, Applications and Systems (BTAS). His other research interests include pattern recognition, image processing, and computer vision.

Kathryn Bonnen received the B.S. degree in computer science from the Department of Computer Science and Engineering at Michigan State University in 2011. Her research interests include pattern recognition, image processing, and computer vision, with applications to face recognition.

Anil K. Jain (S'70–M'72–SM'86–F'91) is a university distinguished professor in the Department of Computer Science and Engineering at Michigan State University, East Lansing. 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 *Introduction to Biometrics* (2011), *Handbook of Face Recognition* (2nd edition, 2011), *Handbook of Fingerprint Recognition* (2009), *Handbook of Biometrics* (2007), *Handbook of Multibiometrics* (2006), *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. ISI has designated him a highly cited researcher. According to Citeseer, his book *Algorithms for Clustering Data* (Englewood Cliffs, NJ: Prentice-Hall, 1988) is ranked 93 in most cited articles in computer science.