# Face Matching and Retrieval Using Soft Biometrics

Unsang Park, *Member, IEEE*, and Anil K. Jain, *Fellow, IEEE*

***Abstract*—Soft biometric traits embedded in a face (e.g., gender and facial marks) are ancillary information and are not fully distinctive by themselves in face-recognition tasks. However, this information can be explicitly combined with face matching score to improve the overall face-recognition accuracy. Moreover, in certain application domains, e.g., visual surveillance, where a face image is occluded or is captured in off-frontal pose, soft biometric traits can provide even more valuable information for face matching or retrieval. Facial marks can also be useful to differentiate identical twins whose global facial appearances are very similar. The similarities found from soft biometrics can also be useful as a source of evidence in courts of law because they are more descriptive than the numerical matching scores generated by a traditional face matcher. We propose to utilize demographic information (e.g., gender and ethnicity) and facial marks (e.g., scars, moles, and freckles) for improving face image matching and retrieval performance. An automatic facial mark detection method has been developed that uses 1) the active appearance model for locating primary facial features (e.g., eyes, nose, and mouth), 2) the Laplacian-of-Gaussian blob detection, and 3) morphological operators. Experimental results based on the FERET database (426 images of 213 subjects) and two mugshot databases from the forensic domain (1225 images of 671 subjects and 10 000 images of 10 000 subjects, respectively) show that the use of soft biometric traits is able to improve the face-recognition performance of a state-of-the-art commercial matcher.**

***Index Terms*—Demographic information, face marks, face recognition, face retrieval, soft biometrics.**


(a) (b) (c)

Fig. 1. Examples of facial marks. (a) Scar, mole, and freckles (FERET database.). (b) Large birthmark (http://www.wnd.com/index.php?fa=PAGE.view&pageId=63558). (c) Tattoo (http://blog.case.edu/colin.mulholland/2007/09/20/a_way_of_expression).

## I. Introduction

LOCAL facial features have played an important role in forensic applications for matching face images. These features include any salient skin region that appears on the face. Scars, moles, and freckles are representative examples of the local facial features [2]. The use of local features has become more popular due to the development of higher resolution sensors, an increase in face image database size, and improvements in image processing and computer vision algorithms. Local features provide a unique capability to investigate, annotate, and exploit face images in forensic applications by improving both the accuracy and the speed of face-recognition systems. This information is also necessary for forensic experts to give testimony in courts of law where they are expected to conclusively identify suspects [3].

Along with the facial marks, demographic information (i.e., gender and ethnicity) can also be considered as ancillary information that is useful in matching face images. The demographic information and facial marks are collectively referred to as soft biometric traits. Soft biometric traits are defined as characteristics that provide some information about the individual, but lack the distinctiveness and permanence to sufficiently differentiate any two individuals [4]. The use of soft biometric traits is expected to improve the face-recognition performance when appropriately combined with a face matcher. On the other hand, when face images are occluded or partially damaged, soft biometric traits can be considered as alternative means of face matching or retrieval. Gender and ethnicity of a person typically do not change over the lifetime, so they can be used to filter the database to narrow down the candidate face images. While some of the facial marks are not permanent, most of them appear to be temporally invariant, which can be useful for face matching and retrieval. When the face images are occluded or severely off-frontal, as is often the case in surveillance videos, utilizing the soft biometric traits is the only reliable evidence to narrow down the candidate face images. Some of the examples of facial marks are shown in Fig. 1.

Most of the current photo-based identifications in law enforcement units and related security organizations involve a manual verification stage. The identification task is performed by a victim, untrained individual, or a trained forensics officer. This can involve verbal descriptions and hand-drawn sketches [5]. Law enforcement agencies have improved the existing practices and established new techniques for gathering and exploitation of evidence using information technology. A large portion of this evidence is in the form of digital media, especially digital imagery.

Spaun [2], [3] describes the facial examination process carried out in the law enforcement agencies. One of the major examination steps involves identifying "class" and "individual"

Manuscript received December 27, 2009; revised April 27, 2010; accepted April 29, 2010. Date of publication May 24, 2010; date of current version August 13, 2010. This work was supported in part by the NSF Industry University Center on Identification Technology Research (CITeR) at West Virginia University and in part by the National Institute of Justice (NIJ) under Grant 2009-DN-BX-K231. The work of A. K. Jain was supported in part by the World Class University (WCU) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-2008-000-10008-0). An earlier version of this work appeared in the Proceedings of the IEEE International Conference on Image Processing (ICIP), 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Davide Maltoni.

The authors are with the Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48824 USA (e-mail: parkunsa@cse.msu.edu).

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

Digital Object Identifier 10.1109/TIFS.2010.2049842

characteristics. The class characteristics include overall facial shape, hair color, presence of facial hair, shape of the nose, presence of freckles, etc. The individual characteristics include the number and location of freckles, scars, tattoos, chipped teeth, lip creases, number and location of wrinkles, etc., in a face. While these examinations are currently performed manually by forensic experts, an automatic procedure will not only reduce the time-consuming and subjective manual process, but is likely to be more consistent and accurate. This five-step manual examination procedure is referred to as ACE-V [7], which stands for Analyze, Compare, Evaluate, and Verify. However, there have been difficulties in standardizing this manual examination process and the descriptions of local features are limited to only a small number of categories. Therefore, it is expected that the computer-aided automatic feature extraction and representation will help standardize the examination process and make the process more efficient. The automatic extraction of facial marks and their representation in a face centered coordinate system will assist in a quantitative analysis. Inclusion of facial-mark-based matching will also enable a rich description of the face matching evidence from a mere match score to a combination of quantitative and descriptive evidence that can be used by forensic examiners in the courts. Of course, statistical studies on the distribution, frequency, and stability of the facial marks should precede before any attempt is made to fully automate the matching process.

Conventional face-recognition systems typically encode the face images by utilizing either local or global texture features. Local techniques first detect the individual components of the human face (i.e, eyes, nose, mouth, chin, and ears), prior to encoding the textural content of each of these components (e.g., EBGM and LFA) [8], [9] . Global (or holistic) techniques, on the other hand, consider the entire face as a single entity during the encoding process (e.g., PCA, LDA, Laplacianfaces, etc.) [10], [11], [12]. However, these techniques do not explicitly utilize local marks (e.g., scars and moles) and usually expect the input to be a full face image.

The use and significance of soft biometric traits can be summarized into four major categories: 1) supplement existing facial matchers to improve the identification accuracy; 2) enable fast face image retrieval; 3) enable matching or retrieval with partial or off-frontal face images; and 4) provide more descriptive evidence about the similarity or dissimilarity between face images, which can be used in the courts. Since facial marks capture the individual characteristics embedded in a face image that are not explicitly utilized in conventional face-recognition methods, a proper combination of face matcher and mark-based matcher is expected to provide improved recognition accuracy. Fig. 2 shows a pair of images that could not be successfully matched by a state-of-the-art face matcher. The presence of prominent facial marks in these two images strongly support the fact that these two images are of the same subject. The mark-based matcher helps in indexing each face image based on the facial marks (e.g., moles or scars). These indices will enable fast retrieval and also the use of textual or key-word-based query (Fig. 3). Finally, marks can help in characterizing partial, occluded, or off-frontal face images, which will assist in matching or retrieval of tasks based on a partial face images often captured by surveillance cameras. Example face images that are occluded are shown in Fig. 4. These images cannot be automatically processed using commercial face-recognition engines due to the occlusion of the eyes; state-of-the-art matchers require eye locations to align two face images. Fig. 5 shows an example retrieval result based on a facial mark on the right side of the cheek in frontal, partial, and nonfrontal face images. Face images with such distinctive marks can be more efficiently matched or retrieved.

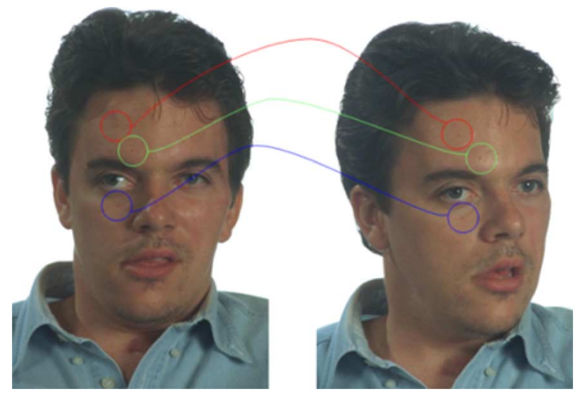


Fig. 2. Two face images of the same person from the FERET database [6]. A commercial face-recognition engine failed to match these images at rank-1. There are a few prominent marks in these two images that can be utilized to make a better decision.

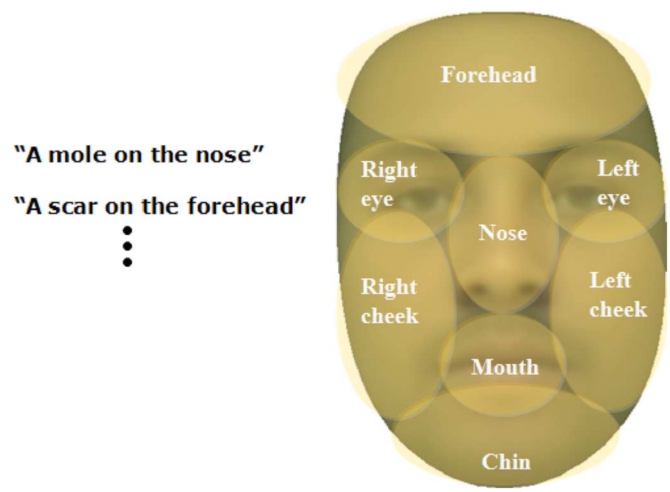



Fig. 3. Examples of textual query and a schematic of face region segmentation.

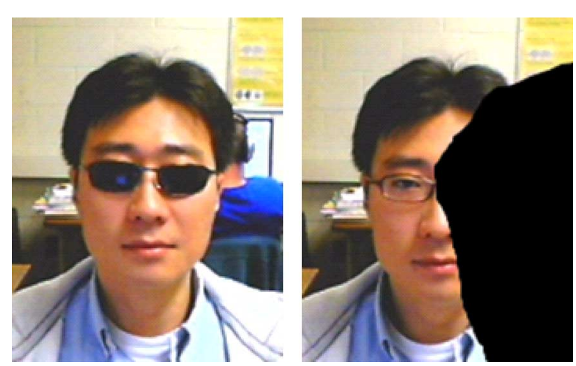


Fig. 4. Example images with facial occlusion.

(a) (b)

Fig. 5. Three different types of example queries and associated retrieval results. (a) Full face, partial face, and off-frontal face (from a video). (b) Retrieval results. The mark that is used in the retrieval is enclosed with a red circle. These images are from the FERET database [6].

Early studies on facial features were conducted by Farkas and Munro [13] and Ekman [14]. Similar studies are also found in [15]. Some of the studies on facial features have been reported in the medical literature [16]. It has been acknowledged that the facial landmarks are largely stable in position and appearance throughout life, since they are associated with the musculo-skeletal system or other body organs.

There have been only a few studies reported in the literature on utilizing facial marks. Lin *et al.* [17] first used the SIFT operator [18] to extract facial irregularities and then fused them with a global face matcher. Facial irregularities and skin texture were used as additional means of distinctiveness to achieve performance improvement. However, the individual types of facial mark were not explicitly defined. Hence, their approach is not suitable for face database indexing. Pierrard *et al.* [19] proposed a method to extract moles using normalized cross correlation method and a morphable model. They claimed that their method is pose and lighting invariant since it uses a 3-D morphable model. However, besides moles, they did not consider other types of facial marks. Pamudurthy *et al.* [20] used distribution of fine-scale skin marks pattern. The correlation coefficients calculated between registered skin patches are aggregated to obtain a matching score. Lee *et al.* [21] introduced "scars, marks, and tattoos (SMT)" in their tattoo image retrieval system. While tattoos can exist on any body part and are more descriptive, we are interested in marks appearing exclusively on the face which typically show simple morphologies. Kumar *et al.* [22] proposed to use a set of descriptive features (e.g., male, chubby, flash, etc.), but facial marks were not defined or considered to be used. There have been a few approaches that use face annotation to efficiently manage digital photographs [23]–[27]. However, these methods or systems utilize either the presence or the identity of faces in photographs; local facial features are not used. Therefore, these are more suitable for multimedia applications rather than the forensic applications of interest here.

We use demographic information (e.g., gender and ethnicity) and facial marks as the soft biometrics traits. We labeled the gender of each face images into three categories (male, female, and unknown) and ethnicity into three categories (Caucasian, African-American, and unknown). We currently label gender and ethnicity manually as practiced in law enforcement. We propose a fully automatic facial mark extraction system using global and local texture analysis methods. We first apply the active appearance model (AAM) to detect and remove primary facial features such as eye brows, eyes, nose, and mouth. These primary facial features are subtracted from the face image. Then, the local irregularities are detected using the Laplacian-of-Gaussian (LoG) operator. The detected facial marks are used to calculate the facial similarity by their morphology and color along with the location. The mark-based matcher can be combined with a commercial face matcher in order to enhance the face matching accuracy or used by itself when the commercial face matcher fails in face matching process due to occlusion or unfavorable viewpoints.

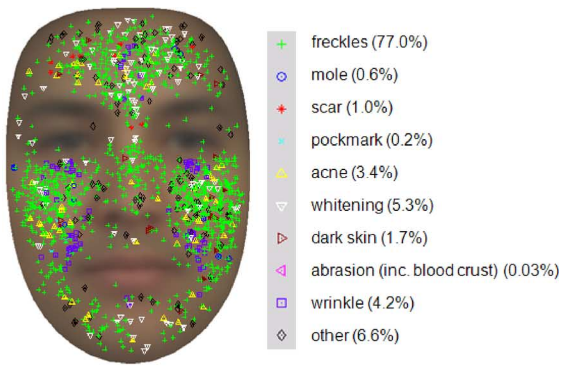
Our method differs significantly from the previous studies in the following aspects: 1) we use a number of soft biometric traits (i.e., gender, ethnicity, facial marks); 2) we extract all types of facial marks that are locally salient; 3) we focus on detecting facial marks and characterize each mark based on its morphology and color; and 4) we evaluate the performance using a state-of-the-art face matcher on a large gallery with 10 213 subjects. The proposed soft biometric matching system will be especially useful to forensics and law enforcement agencies because it will 1) supplement existing facial matchers to improve the identification accuracy, 2) enable fast face image retrieval based on high level semantic query, 3) enable matching or retrieval from partial or off-frontal face images, and 4) help in discriminating identical twins.

The rest of this paper is organized as follows: Section II provides statistics on distributions and frequencies of facial marks in ten different categories, Section III describes our mark detection process, Section IV presents the mark-based matching scheme, and Section V provides experimental results and discussions. Section VI summarizes our contributions and lists some directions for future work.

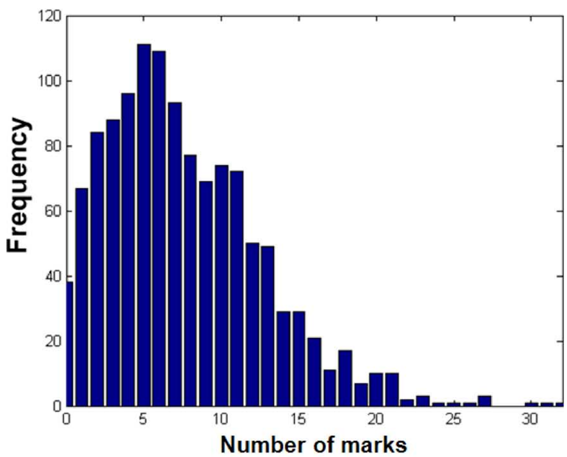
## II. Statistics of Facial Marks

To understand the role of facial marks in face recognition and retrieval, we analyzed the location and characteristics of facial marks in a forensic mugshot database with 1225 images of 671 subjects by manually assigning ground truth labels to these marks. Each facial mark is represented with an enclosing rectangular bounding box. The ground truth labeling process is performed by using the following ten categories provided by a forensics expert in [3]:

1) Freckle: small spots from concentrated melanin;
2) Mole: growth on the skin (brown or black);
3) Scar: marks left from cuts or wounds;
4) Pockmark: crater-shaped scar;
5) Acne: red regions caused by pimple or zit;
6) Whitening: skin region that appears white;
7) Dark skin: skin region that appears dark;
8) Abrasion: wound (includes clots);
9) Wrinkle: fold, ridge, or crease in the skin;




(a)




(b)

Fig. 6. Statistics of facial marks based on a database of 1225 images of 671 different subjects in the Michigan police mugshot database. (a) Distributions of facial mark types on the mean face and the percentage of each mark types. (b) Distribution of the number of marks per subject.

10) Other: all other types of marks.

Freckle is a single or a set of dark spots. When there is a dense set of spots in a small region, we label each of the prominent dark spots rather than labeling the entire set with a single bounding box. Mole is referred to as an extruded region with typically dark skin color. In a 2-D facial image, it is difficult to distinguish between a spot and a mole. A mole typically appears larger in size and darker in color compared with spots. Scar is the discolored region of skin induced from a cut or injury. Pockmark is a crater-shaped scar. Acne is a red region caused by pimples or zits and stays for a few days to several months. Whitening represents a skin region that appears brighter compared with the surrounding region; it is observed more often with dark skinned people. When a larger region of skin is observed as dark, it is labeled as "dark skin." While abrasion is not temporally invariant, it can later be related to the scars that are possibly caused by abrasions. We consider only large wrinkles and ignore small wrinkles especially around the eyes and mouth. We ignore beards and facial hair in constructing the ground truth. All other facial marks that do not belong to the nine groups mentioned above are labeled as "other."

The statistics of mark location and frequency are shown in Fig. 6. Freckles are observed most often, followed by wrinkles and acne. The average number of marks in our database is about 7 per subject. About 97% of the subjects in our database showed at least one mark, which suggests that facial marks can indeed be useful for recognition. Our earlier preliminary work [1] showed that automatically detected facial marks when combined with a commercial matcher improves the face-recognition accuracy. In [1], we only used the location of detected facial marks and evaluated the system on a small database.

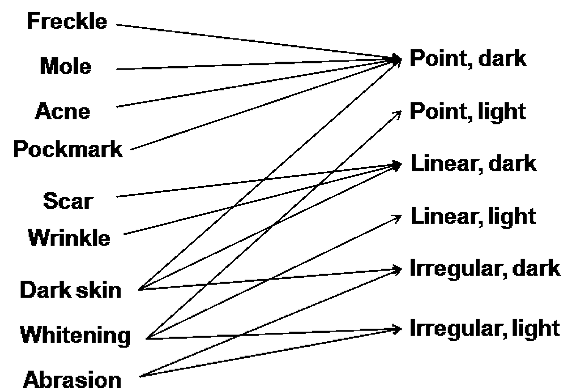



Fig. 7. Schematic of the mapping from the semantic mark categories to the morphology and color-based categories.

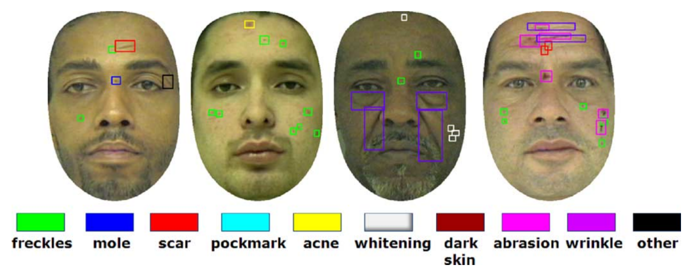



Fig. 8. Example marks in semantic categories.

Useful information about facial marks that should be utilized during matching is the class label of each mark. However, automatic classification of each mark based on the ten categories defined above is extremely difficult because of the ambiguities between those categories. Therefore, to simplify the classification problem we defined a small number of classes that are based on morphology and color of the mark. We use three different morphologies (i.e., point, linear, irregular) and two color characteristics (i.e., dark or bright compared to the surrounding region).[1] Fig. 7 shows how nine of the ten semantic categories can be mapped into the morphology and color-based categories. Example marks in each type of morphology and color are shown in Figs. 8 and 9.

## III. Facial Mark Detection

Face marks appear as salient localized regions on the face image. Therefore, a blob detector based on Difference of Gaussian (DoG) or LoG operator [28] can be used to detect the marks. However, a direct application of a blob detector on a face image will generate a large number of false positives due to the presence of primary facial features (e.g., eyes, eye brows, nose, and mouth). Therefore, first localizing the primary facial features and then extracting facial marks in the rest of the face region is necessary for successful mark detection. The complete

[1]These categories were derived in a meeting with Dr. Richard W. Vorder Bruegge of the Federal Bureau of Investigation (FBI) Academy.

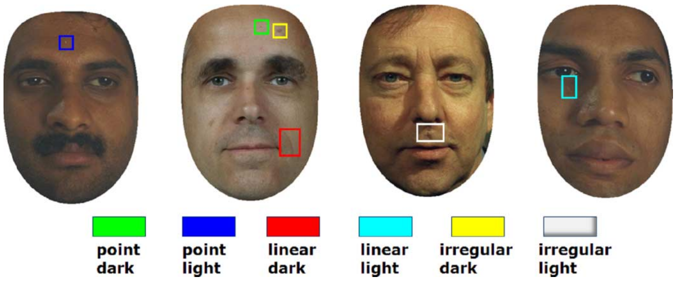



Fig. 9. Example marks with morphology and color-based labels.

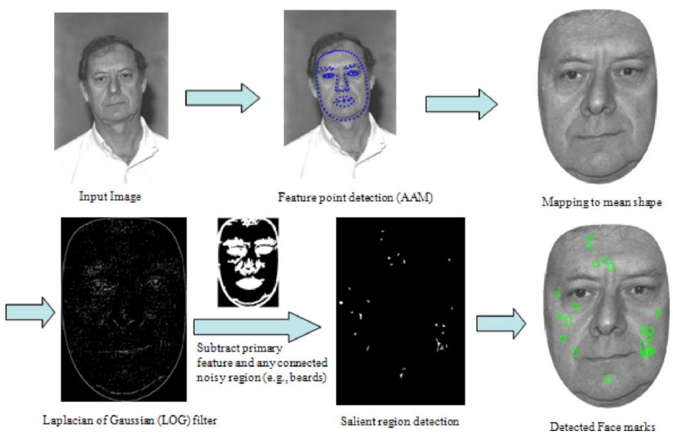



Fig. 10. Schematic of automatic facial mark extraction process. These images are from the FERET database [6].

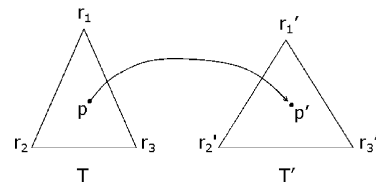



Fig. 11. Schematic of texture mapping process using the triangular Barycentric coordinate system.

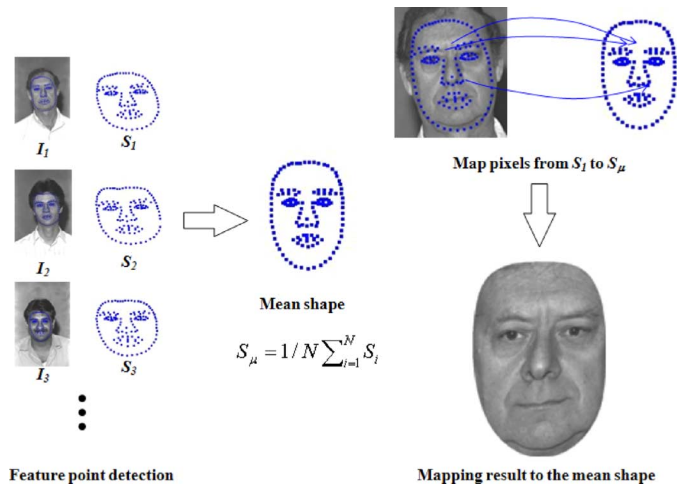



Fig. 12. Schematic of the mean face construction.

facial mark detection procedure is illustrated in Fig. 10. We will describe 1) primary facial feature detection, 2) mapping to the mean shape, and 3) mask construction processes in the following subsections.

### A. Primary Facial Feature Detection

We use AAM [29], [30] to automatically detect 133 landmarks that delineate the primary facial features: eyes, eye brows, nose, mouth, and face boundary (Fig. 10). These primary facial features will be disregarded in the subsequent facial mark detection process. AAM models both the shape and texture of face images using the Principal Component Analysis (PCA). By providing a manually labeled set of landmark points in training data, a set of shapes $X = \{x_1, \ldots, x_N\}$ and corresponding textures $G = \{g_1, \ldots, g_N\}$ are obtained. After applying PCA both on $X$ and $G$, principal components of shape and texture, $P_X$ and $P_G$ are obtained. Then, the shape and texture, $x_{\rm new}$ and $g_{\rm new}$, of a new face can be expressed as $x_{\rm new} = x_\mu + P_X b_X$ and $g_{\rm new} = g_\mu + P_G b_G$, where $x_\mu$ and $g_\mu$ ($b_X$ and $b_G$) are means (weight vectors) of $X$ and $G$, respectively. The 133 landmark points were labeled on a subset of mugshot images for training.

### B. Mapping to Mean Shape

Using the landmarks detected by AAM, we map each face image to the mean shape to simplify the mark detection, matching, and retrieval. Let $S_i$, $i = 1, \ldots, N$ represent the shape of each of the $N$ face images in the database (gallery) based on the 133 landmarks. Then, the mean shape is simply defined as $S_\mu = \sum_{i=1}^{N} S_i$. Each face image $S_i$ is mapped to the mean shape $S_\mu$ by using the Barycentric coordinate-based [31] texture mapping process. First, both $S_i$ and $S_\mu$ are subdivided into a set of triangles. Given a triangle $T$ in $S_i$, its corresponding triangle $T'$ is found in $S_\mu$. Let $r_1$, $r_2$, and $r_3$ ($r'_1$, $r'_2$, and $r'_3$) be the three vertices of $T$ ($T'$). Then, any point $p$ inside $T$ is expressed as $p = \alpha r_1 + \beta r_2 + \gamma r_3$ and the corresponding point $p'$ in $T'$ is similarly expressed as $p' = \alpha r'_1 + \beta r'_2 + \gamma r'_3$, where $\alpha + \beta + \gamma = 1$. This way, the pixel value at $p$ is mapped to $p'$. Fig. 11 shows the schematic of the Barycentric mapping process. By repeating this mapping process for all the points inside all triangles, the texture in $S_i$ is mapped to $S_\mu$.

After this mapping process, all face images are normalized in terms of scale and rotation and it allows us to represent each facial mark in a face-centered common coordinate system. Fig. 12 shows the schematic of mean face construction.

### C. Mask Construction

We construct a mask from the mean shape $S_\mu$ to suppress false positives due to primary facial features in the blob detection process. The blob detection operator is applied to the face image mapped into the mean shape. A mask constructed from $S_\mu$ is used to suppress blob detection on the primary facial features. Let the mask constructed from the mean shape be denoted as $M_g$, namely, a generic mask. However, the generic mask does not cover the user-specific facial features, such as beard or small winkles around eyes or mouth, that are likely to increase the false positives. Therefore, we also build a user-specific mask $M_s$ using the edge image. The edge image is obtained by using the conventional Sobel operator [32]. An example edge image is shown in Fig. 16(b). The user-specific mask $M_s$, constructed

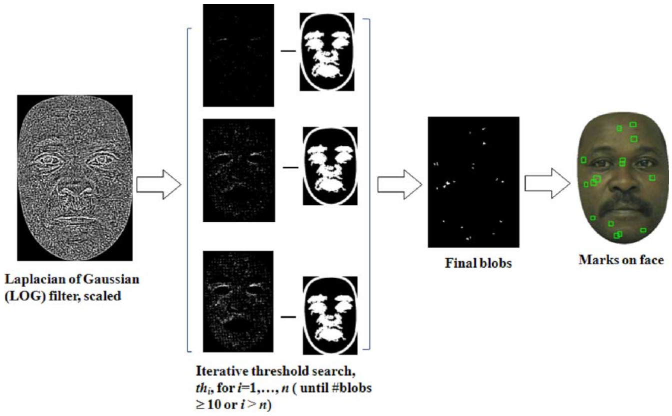



Fig. 13. Schematic of the iterative thresholding process for blob detection.

as a sum of $M_g$ and edges that are connected to $M_g$, helps in removing most of the false positives appearing around the beard or small wrinkles around eyes or mouth.

*D. Blob Detection*

Facial marks mostly appear as isolated blobs. Therefore, we use the well-known blob detector, LoG operator, to detect facial mark candidates. We have used LoG kernels with three different sizes ($= 3 \times 3$, $7 \times 7$, and $9 \times 9$) and five different $\sigma$s ($= \sqrt{2}, \sqrt{2}^2, \sqrt{2}^3, \sqrt{2}^4$, and $\sqrt{2}^5$). We chose the smallest LoG kernel ($3 \times 3$ with $\sigma = \sqrt{2}$) that shows the best identification accuracy in our experiment. The LoG operator is usually applied at multiple scales to detect blobs of different sizes. However, we used a single-scale LoG filter followed by a morphological operator (e.g., closing) to reduce the computation time. The LoG filtered image subtracted with the user-specific mask undergoes a binarization process with a series of threshold values $t_i^b$; $i = 1, \ldots, K$ in a decreasing order. The threshold value $t_i^b$ is successively applied until the resulting number of connected components is larger than a preset value ($t_0^n$) (Fig. 13). A brightness constraint ($t_0^c$) is also applied to each of the connected components to suppress false positives due to weak blob responses. When the user-specific mask does not effectively remove sources of false positives, true marks with lower contrast will be missed in the mark detection process. The overall procedure of facial mark detection is enumerated below:

1) facial landmark detection (AAM);
2) mapping to the mean shape $S_\mu$;
3) construct user-specific mask $M_s$ ($M_g$ is constructed only once and shared by all images);
4) apply LoG operator;
5) using thresholds $t_i^c$, $i = 1, \ldots, K$, binarize, and detect blobs ($m_j$) such that $m_j$ does not overlap with $M_s$ and the average brightness of $m_j \geq t_0^b$; stop if the total number of blobs $\geq t_0^n$;
6) represent each mark with a bounding box.

The overall mark detection process is shown in Fig. 10. Example ground truth and automatically detected marks are shown in Fig. 14. Mark detection accuracy is evaluated in terms of precision and recall curve as shown in Fig. 15. The precision and recall values for the mark detector with a range of brightness contrast thresholds $t_0^c$ varies from (30%, 60%) to (54%, 16%) for the mugshot database. These results represent that the automatic mark detection algorithm shows reasonable performance. In Section V, we will evaluate whether this automatic mark detection method helps to improve the face-recognition accuracy.

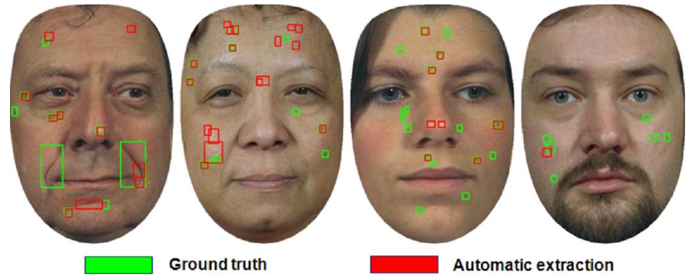



Fig. 14. Ground truth and automatically detected facial marks for four images in the FERET database.

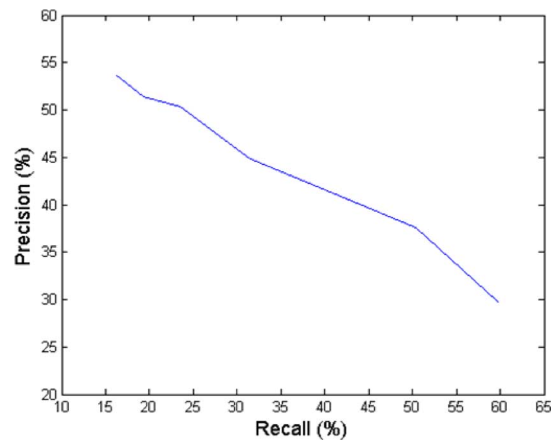



Fig. 15. Precision-recall curve of the automatic mark detection method.

The effect of a generic mask and user-specific mask on mark detection is shown in Fig. 16. The user-specific mask helps in reducing both false positives and false negatives by disregarding small wrinkles and hairs.

*E. Blob Classification*

For each detected blob (i.e., mark), we assign a bounding box tightly enclosing the blob. We then classify a mark in a hierarchical fashion: linear versus all, followed by circular (point) versus irregular. For the linearity classification of a blob, we use two eigen values $\lambda_1$ and $\lambda_2$ obtained from the eigen decomposition on the $x$ and $y$ coordinates of blob pixels. When $\lambda_1$ is significantly larger than $\lambda_2$, the mark is decided as a linear blob. For the circularity detection, we calculate the second moment of the blob pixels $M_2$. A circle $R_{M_2}$ with radius $M_2$ will enclose most of the blob pixels if they are circularly distributed. Therefore, a decision can be made based on the ratio of the number of pixels within and outside of $R_{M_2}$. The color of the blob can be decided based on the ratio of the mean intensity of the pixels inside and outside of the blob ($e_{\text{in}}/e_{\text{out}}$). The classification process can be summarized as below.

i) Morphology classification

if $\lambda_1/\lambda_2 > t_1^l$, then linear shape

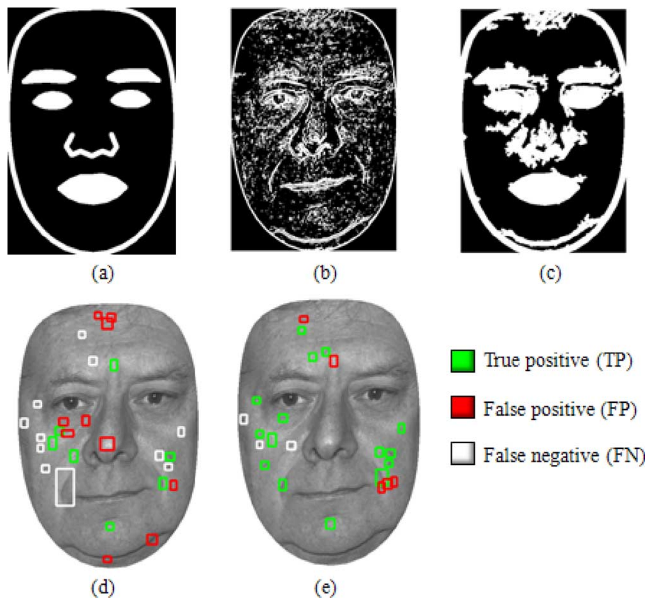



Fig. 16. Effects of generic and user-specific masks on facial mark detection. True positives (TPs) increase but false negatives (FNs) decrease by using a user-specific mask. False positives have moved from small wrinkles around the eyes and beard to the middle of skin region. These are possibly true marks that were missed during the ground truth labeling. (a) Generic mask. (b) Edge map. (c) User-specific mask. (d) Mark detection using generic mask. (e) Mark detection using user-specific mask.

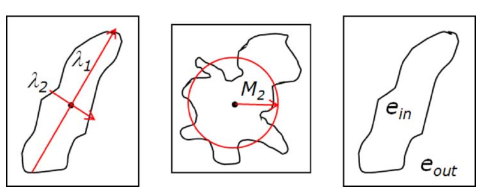



Fig. 17. Schematic of the morphology and color-based mark classification.

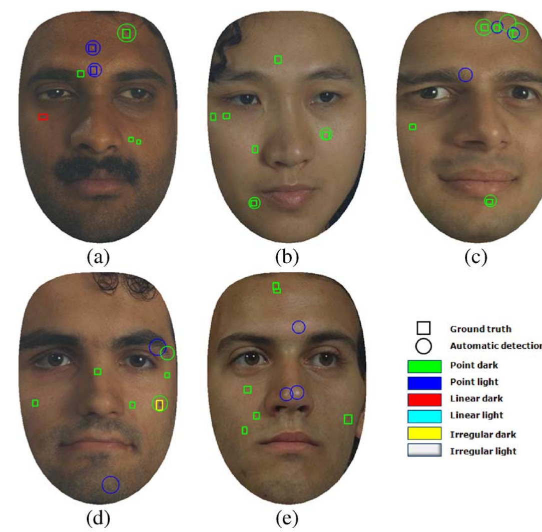



Fig. 18. Example images with ground truth and automatically detected facial marks with their class.

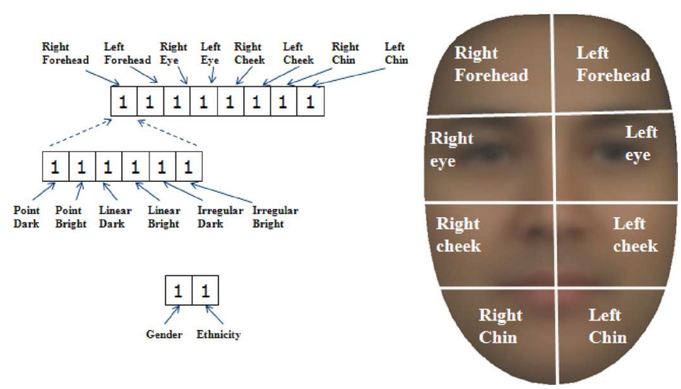



Fig. 19. Schematic of the mark-based indexing scheme.

else

if (#blobs in $R_{M_2}$)/(#blobs out of $R_{M_2}$) $> t_2^l$, then circular

else irregular

ii) Color classification

if $e_{\text{in}}/e_{\text{out}} > t_3^l$, then light

else dark

We have set $t_3^l = 1$ as it is simply comparing the mean intensity of the blob and its surrounding region. We have tried a range of values for $t_1^l$ and $t_2^l$ and selected the set that shows the best identification accuracy, recall, and precision in blob detection and classification process ($t_1^l = 10$ and $t_2^l = 0.6$). The schematic of the blob classification process is shown in Fig. 17. Fig. 18 shows five example images with ground truth and automatically extracted facial marks and their classes.

## IV. Facial-Mark-Based Matching

Given the demographic information and facial marks, we encode them into a 50-bin histogram. The first 48 bins represent the distribution, frequency, morphology, and color of facial marks. The last two bins represent the gender and ethnicity. To encode the facial marks, the face image in the mean shape space is subdivided into eight different regions, as shown in Fig. 19. Each mark is encoded by a six-digit binary value representing its morphology and color (six labels defined in Section II). When there is more than one mark in the same region, a bit by bit summation is performed. The six bin values are concatenated for the eight different regions in the order as shown in Fig. 19 to generate the 48-bin histogram. The gender (ethnicity) has three distinct values, representing male, female, and unknown (Caucasian, African-American, unknown). The histogram intersection method is used to calculate the matching scores.

The soft biometric traits-based matcher can be used to quickly retrieve candidate images from a large database. Since the soft biometric-based matcher also generates a matching score, it can be combined with any face matcher to improve the overall face-recognition accuracy. We used weighted score-sum with a min–max normalization method for the matcher combination [33]. The weights are empirically chosen to obtain the best recognition accuracy.

TABLE I
FACE-RECOGNITION ACCURACY USING FACEVACS MATCHER, PROPOSED FACIAL MARK MATCHER AND THEIR FUSION

| Matcher | No demographic | Gender | Ethnicity | Both |
|---|---|---|---|---|
| FaceVACS only | 90.61% | 91.08% | 91.08% | 91.55% |
| FaceVACS + mark | 91.08% | 91.55% | 91.55% | 92.02% |

## V. EXPERIMENTAL RESULTS

### A. Database

We used the FERET [6] and a mugshot face database to construct two separate databases, DB1 and DB2, to evaluate the proposed mark-based matcher. We have also collected a small video database of five subjects (DB3) and images of five pairs of identical twins (DB4) [34]. DB1 consists of images from FERET and mugshot databases with 213 images, one image per subject, for probe and 10 213 images, one per subject for gallery. DB2 consists of mugshot face images with 554 images from 554 subjects for probe and 671 images from 671 subjects for gallery. The image size varies from $215 \times 323$ to $384 \times 480$ (width $\times$ height) for the mugshot and $512 \times 768$ for FERET databases both at 96 dpi resolution. DB3 and DB4 images are of size $640 \times 480$ at 96 dpi resolution. We manually labeled the ten facial mark types as defined in Section II for DB2 and six facial mark types (e.g., based on morphology and color) for the FERET images in DB1 to create the ground truth. This allows us to evaluate the proposed facial mark extraction method. We used the images from the FERET database in DB1 as the gallery images in the matching experiments with DB3 and DB4.

### B. Matching and Retrieval

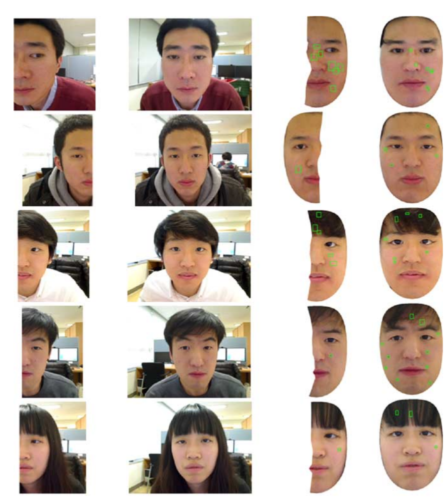
We have primarily used DB1, DB3, and DB4 for the soft biometric matching and retrieval experiments. DB2 was used to provide the statistical analysis of the marks (Fig. 6, Section II). Even though DB2 was used in the matching experiment in our earlier work [1], the images in it are of low resolution and are not suitable for the mark classification-based matching.

Table I shows the matching experiment results on DB1 using facial marks and the proposed soft biometric matcher. The rank-one matching accuracy of the state-of-the-art commercial face matcher, FaceVACS [35] is improved as a result of using gender and ethnicity. The matching accuracy of FaceVACS also improves by combining it with the mark-based matcher. The combined matching accuracy is further improved by incorporating additional soft biometric traits (i.e., gender and ethnicity). Note that FaceVACS is one of the best face matchers available in the market, so improving its accuracy is not an easy task. Fig. 20 shows example images that failed using FaceVACS but succeeded after combining it with soft biometric matcher. Fig. 20 shows three queries that were correctly matched to their mate in the gallery, by using facial marks, gender and ethnicity, respectively. The equal error rates (EERs) are observed as 3.853% and 3.839% from FaceVACS and the fusion of facial marks and FaceVACS, respectively.

We also applied the soft biometric-based matcher to a number of video frames which contain face occlusion. Fig. 21 shows five example query face images with occlusion and their corresponding gallery images. These images failed to be matched using the commercial face matcher at all ranks due to the "enrollment failure." The soft biometric traits-based matcher successfully retrieved the correct images for these five probes at

(a) (b) (c) (d)

Fig. 20. Examples where soft biometrics improves the performance of FaceVACS face image matching results. (a) Query images. (b) Gallery images. (c) Query in mean shape with detected marks. (d) Gallery in mean shape with detected marks. Three different query images and their true mate in the gallery are shown. The first query is correctly matched using face marks, the second query is correctly matched using gender, and the last query is correctly matched using ethnicity. Face matcher with soft biometric [facial mark (first row), gender (second row) and ethnicity (third row)] improves the matching accuracy.



(a) (b) (c) (d)

Fig. 21. Example face image retrieval results using partial faces appearing in video frames with proposed soft biometric matching. (a) Query images. (b) Gallery images. (c) Query in mean shape with detected marks. (d) Gallery in mean shape with detected marks. Commercial matcher failed in matching (a) to (b), but the soft biometric matcher found correct face images for these five probes in the gallery at rank 3, 6, 8, 7, and 4, respectively. Extracted marks used in matching are shown with green bounding boxes.


Fig. 22. Example twin face images. (a) First twin. (b) Second twin. (c) First twin in mean shape with detected marks. (d) Second twin in mean shape with detected marks. All images in (a) are incorrectly matched with those in (b) at rank-one using a commercial matcher, but did not match after combining with mark-based matcher (except the second row). Each distinctive mark is marked by a black arrow.

rank 3, 6, 8, 7, and 4, respectively. The primary facial feature points (133 feature points) were manually labeled for the partial face and automatically labeled for the gallery images. All the marks were detected automatically.

Next, we applied our mark-based matcher to facial images of identical twins [34]. Fig. 22 shows five example pairs of face images of identical twins that were incorrectly matched at rank-one by FaceVACS. However, these images have a few distinctive facial marks that can be used to distinguish the identical twins. When the face matching scores are combined with the mark-based matcher, four of them are no longer matched at rank-one. From the first to the last, each probe image is matched to its identical twin image at rank 5, 1, 45, 22, and 3, respectively, after we combine FaceVACS with the soft biometric matcher. In the case of identical twin images, gender and ethnicity do not help to distinguish them because the identical twins share the same gender and ethnicity. These twin images were captured in an outdoor environment with strong side light causing the AAM-based feature point detection to perform poorly. Therefore, we had to manually adjust some of the detected AAM feature points.

The proposed mark extraction method is currently implemented in Matlab and takes about 15 s per face image. Mark-based matching time is negligible. Further code optimization and migrating into C++ will significantly speed up the process.

## VI. Conclusions and Future Work

Conventional face matching systems generate only numeric matching scores as a similarity between two face images, whereas, the facial mark-based matching provides specific and more meaningful evidence about the similarity of two face images. Thus, automatic extraction and representation of facial marks is becoming important in forensic applications. Facial marks can be used to support other evidence presented in courts of law and may be used as a strong evidence by themselves, when other evidence is either not available or reliable.

We have developed a soft biometric traits-based face matching system. It uses gender and ethnicity information and facial marks. This soft biometric matcher can be combined with any face matcher to improve the recognition accuracy or used by itself when a face matcher fails because of face occlusion. We also show that facial marks can help in discriminating identical twins. With the proposed soft biometric matcher, users can issue semantic queries to retrieve images of interest from a large database. For example, a query could be of the form "Retrieve all face images with a mole on the left side of lip."

The proposed framework of using soft biometric traits in face matching is highly relevant to the goals of the FBI's Next Generation Identification (NGI) R&D effort [36]. The criminal history database maintained by the FBI includes photographic face evidence such as scars, marks, and tattoos as a useful component of the database for the purpose of subject identification. The proposed work is an effort towards establishing the scientific basis and capability of quantitative analysis to exploit the NGI activity [37]. We believe this effort will help avoid the controversies beleaguering the fingerprint matching community in recent years [38].

Our ongoing work includes 1) improving the mark detection accuracy, 2) extending the automatic mark detection to off-frontal face images, and 3) studying the image resolution requirement for reliable mark extraction.

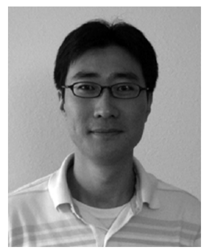
## Acknowledgment

The authors would like to thank Dr. R. W. V. Bruegge of the Federal Bureau of Investigation (FBI) Academy and Inspector G. Michaud of the Michigan State Police Forensic Science Division for their support on mark classification and mugshot database.

## References

[1] A. K. Jain and U. Park, "Facial marks: Soft biometric for face recognition," in *Proc. IEEE ICIP*, 2009, pp. 1–4.
[2] N. A. Spaun, "Forensic biometrics from images and video at the Federal Bureau of Investigation," in *Proc. BTAS*, 2007, pp. 1–3.
[3] N. A. Spaun, "Facial comparisons by subject matter experts: Their role in biometrics and their training," in *Proc. ICB*, 2009, pp. 161–168.
[4] A. K. Jain, S. C. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," in *Proc. ICBA*, 2004, vol. 3072, LNCS, pp. 731–738.
[5] B. Klare and A. K. Jain, "Sketch to photo matching: A feature-based approach," in *Proc. SPIE, Biometric Technology for Human Identification VII*, 2010, pp. 1–10.
[6] J. Phillips, H. Wechsler, J. S. Huang, and P. J. Rauss, "The feret database and evaluation procedure for face recognition algorithms," *Image Vision Comput.*, vol. 16, no. 5, pp. 295–306, 1998.
[7] H. Tuthill and G. George, *Individualization: Principles and Procedures in Criminalistics*. Florida: Lightning Powder Company, Inc., 2002.
[8] L. Wiskott, J.-M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775–779, Jul. 1997.
[9] P. S. Penev and J. J. Atick, "Local feature analysis: A general statistical theory for object representation," *Comput. Neural Syst.*, vol. 7, pp. 477–500, 1996.
[10] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
[11] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.

[12] F. Wang, J. Wang, C. Zhang, and J. T. Kwok, “Face recognition using spectral features,” *Pattern Recognit.*, vol. 40, no. 10, pp. 2786–2797, 2007.
[13] L. G. Farkas and I. R. Munro*, Anthropometric Facial Proportions in Medicine*. New York: Thomas Publishing, 1987.
[14] P. Ekman, W. Friesen, and J. Hager*, The Facial Action Coding System*, 2nd ed. London: Weidenfeld & Nicolson, 2002.
[15] *Craniofacial Anthropometry*, [Online]. Available: http://www.plagiocephaly.info/faqs/anthropometry.htm
[16] Medical Gross Anatomy/Head and Neck/ University of Michigan [Online]. Available: http://anatomy.med.umich.edu/surface
[17] D. Lin and X. Tang, “From macrocosm to microcosm,” in *Proc. CVPR*, 2006, pp. 1355–1362.
[18] D. G. Lowe, “Distinctive image features from scale invariant keypoints,” *Int. J. Comput. Vision (IJCV)*, vol. 60, no. 2, pp. 91–110, 2004.
[19] J. S. Pierrard and T. Vetter, “Skin detail analysis for face recognition,” in *Proc. CVPR*, 2007, pp. 1–8.
[20] S. Pamudurthy, E. Guan, K. Mueller, and M. Rafailovich, “Dynamic approach for face recognition using digital image skin correlation,” in *Proc. Audio- and Video-based Biometric Person Authentication (AVBPA)*, 2005, pp. 1010–1017.
[21] J.-E. Lee, A. K. Jain, and R. Jin, “Scars, marks and tattoos (SMT): Soft biometric for suspect and victim identification,” in *Proc. Biometric Symposium, Biometric Consortium Conf.*, 2008, pp. 1–8.
[22] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, “Attribute and simile classifiers for face verification,” in *IEEE Int. Conf. Computer Vision (ICCV)*, 2009, pp. 1–8.
[23] L. Zhang, Y. Hu, M. Li, W. Ma, and H. Zhang, “Efficient propagation for face annotation in family albums,” in *Proc. ACM Int. Conf. Multimedia*, 2004, pp. 716–723.
[24] J. Y. Choi, S. Yang, Y. M. Ro, and K. N. Plataniotis, “Face annotation for personal photos using context-assisted face recognition,” in *Proc. ACM Int. Conf. Multimedia Information Retrieval*, 2008, pp. 44–51.
[25] S.-W. Chu, M.-C. Yeh, and K.-T. Cheng, “A real-time, embedded face-annotation system,” in *Proc. ACM Int. Conf. Multimedia*, 2008, pp. 989–990.
[26] Picasa [Online]. Available: http://picasa.google.com
[27] iPhoto ’09 [Online]. Available: http://www.apple.com/support/iphoto
[28] T. Lindberg, “Feature detection with automatic scale selection,” *Int. J. Comput. Vision (IJCV)*, vol. 30, no. 2, pp. 79–116, 1998.
[29] T. F. Cootes, G. J. Edwards, and C. J. Taylor, “Active appearance models,” in *Proc. ECCV*, 1998, vol. 2, pp. 484–498.
[30] M. B. Stegmann, “The AAM-API: An open source active appearance model implementation,” in *Proc. MICCAI*, 2003, pp. 951–952.
[31] C. J. Bradley*, The Algebra of Geometry: Cartesian, Areal and Projective Co-ordinates*. Bath: Highperception, 2007.
[32] R. O. Duda and P. E. Hart*, Pattern Classification and Scene Analysis*, 2nd ed. Hoboken, NJ: Wiley, 1995.
[33] A. K. Jain, K. Nandakumar, and A. Ross, “Score normalization in multimodal biometric systems,” *Pattern Recognit.*, vol. 38, no. 12, pp. 2270–2285, 2005.
[34] Z. Sun, A. Paulino, J. Feng, Z. Chai, T. Tan, and A. Jain, “A study of multibiometric traits of identical twins,” in *Proc. SPIE Defense, Security, and Sensing: Biometric Technology for Human Identification*, 2010, pp. 1–12.
[35] FaceVACS Software Developer Kit Cognitec Systems GmbH [Online]. Available: http://www.cognitec-systems.de
[36] Next Generation Identification Overview Page Federal Bureau of Investigation [Online]. Available: http://www.fbi.gov/hq/cjisd/ngi.htm
[37] N. A. of Sciences, Strengthening Forensic Science in the United States: A Path Forward The National Academies Press, 2000.
[38] S. L. Zabell, “Fingerprint evidences,” *J. Law Policy*, vol. 13, pp. 143–170, 2005.



**Unsang Park** (S’08–M’08) received the B.S. and M.S. degrees from the Department of Materials Engineering, Hanyang University, South Korea, in 1998 and 2000, respectively. He received the M.S. and Ph.D. degrees from the Department of Computer Science and Engineering, Michigan State University, in 2004 and 2009, respectively.

From 2009, he was a Postdoctoral Researcher in the Pattern Recognition and Image Processing Laboratory, Michigan State University, East Lansing. His research interests include biometrics, video surveillance, image processing, computer vision, and machine learning.

**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 *Handbook of Fingerprint Recognition* (2009), *Handbook of Biometrics* (2007), *Handbook of Multibiometrics* (2006), *Handbook of Face Recognition* (2005), *BIOMETRICS: Personal Identification in Networked Society* (1999), and *Algorithms for Clustering Data* (1988). He served as a member of the Defense Science Board and The National Academies committees on Whither Biometrics and Improvised Explosive Devices.

Dr. Jain received the 1996 IEEE TRANSACTIONS ON NEURAL NETWORKS Outstanding Paper Award and the Pattern Recognition Society best paper awards in 1987, 1991, and 2005. He is a fellow of the AAAS, ACM, IAPR, and SPIE. He has received Fulbright, Guggenheim, Alexander von Humboldt, IEEE Computer Society Technical Achievement, IEEE Wallace McDowell, ICDM Research Contributions, and IAPR King-Sun Fu awards. 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.