

An investigation of Face Matching and Retrieval in Cyber Forensics Applications

Dr. I. Lakshmi

Assistant Professor, Department Of Computer Science, Stella Maris College, Chennai, Bengal, India

ABSTRACT

A latent fingerprint on the counter. A drop of blood on the floor. Law enforcement has successfully used these forensic clues to catch criminals for decades. But consider a face image captured by a surveillance camera that needs to be matched against millions of mug shots across the country. With the rapid increase in the number of surveillance cameras and mobile devices with built-in cameras, the forensics world is changing, and the progress in face recognition is helping to lead the way. In fact, in 2009, an estimated 30 million surveillance cameras were deployed in the US, shooting 4 billion hours of footage a week.[1] However, although recent research advances have helped lay the foundations for realizing face-matching scenarios for utilizing this data, face recognition in the forensics arena still poses a number of challenges. This article highlights the challenges in applying face-recognition technology to forensics applications. We explain why forensic face recognition differs from typical portrait face recognition and why a human examiner is often needed to carefully interpret and verify the matching results. Furthermore, we address three specific research problems that pose challenges to commercial-off-the-shelf (COTS) face recognition systems (FRSs):

- ✓ robustness to facial aging,
- ✓ retrieval using facial scars and marks, and
- ✓ matching forensic (composite) sketches to
- ✓ face photograph databases.

Solutions to these three problems are necessary to accurately remove duplicates in various government face databases, including mug shot, passport, and driver license photos (aging-invariant FRS); to search a large face database when only partial or low-quality face images are available (scar and mark matching); and to apprehend criminals when no photo of the suspect is available (sketch to- photo matching). Additionally, we discuss methods that can augment existing COTS face-recognition systems by improving the quality of a face image prior to submission.

Keywords : Photo Matching, COTS Face-Recognition Systems-FRS

I. INTRODUCTION

Face Recognition Overview

Face recognition is the task of recognizing a person using digital face images. A FRS is typically designed to output a measure of similarity between two face images. Automated FRSs typically involve finding key facial landmarks (such as the center of the eyes) for alignment, normalizing the face's appearance,

choosing a suitable feature representation, learning discriminative feature combinations, and developing accurate and scalable matching schemes.[2] Figure 1 illustrates the major steps in automatic face recognition. Two decades of vigorous research has yielded face-recognition systems that are highly accurate in constrained environments (see Figure 2). However, the face-recognition community has recognized four key factors that significantly

compromise recognition accuracy: pose, illumination, expression, and aging (see Figure 3). Figure 4 shows the impact of facial aging on face-recognition performance. Thus, deployments of fully automated FRSs are mostly limited to scenarios in which we can largely constrain these factors. Face images in government-issued identification documents (such as driver's licenses and passports) and mug shots are two scenarios that offer such constraints, which has led to success in the de-duplication (that is, a 1:N matching process to detect ID cards enrolled under different names but belonging to the same subject) of identification cards and prevention of false prisoner releases.

Paradigm for Forensic Face Recognition

In forensic identification, investigators must use any available information to facilitate multimedia in Forensics, Security, and Intelligence subject identification. Typically, the sources of face images are surveillance cameras, mobile device cameras, forensic sketches, and images from social media sites. These face images are difficult to match because they are often captured under non-ideal conditions (see Table 1). Non-forensic, fully automated scenarios are not severely impacted by these performance degrading factors. As a result, forensic face recognition often requires a pre-processing stage of image enhancement or a specialized matcher to perform recognition.

Another important aspect in face recognition in forensics is the continuously increasing size of face databases or galleries. For example, the mug shot database at the Pinellas County Sheriff's Office in Florida contains more than 7.5 million face images.

Most Departments of Motor Vehicles (DMV) in the US (34 states) utilize FRSs[4]. The US Department of State hosts one of the largest face databases in the world, with a gallery of approximately 100 million images, which are being used for de-duplication of passport and visa applicants. We can state the problem of forensic face recognition as follows. A low-quality query (or probe) image of an unidentified subject is available from a source such as a surveillance camera or a forensic sketch. An expansive database (or gallery) of high-quality face images (such as mug shots) exists that might contain the subject. To boost the recognition accuracy in this

difficult matching scenario, a modified matching paradigm with a human in the loop is necessary (see Figure 5). Although mainstream face-recognition research does not often consider this semi automated face recognition, it is necessary to include a human in the recognition loop to boost the accuracy and confidence in forensic scenarios.[5] The role of man and machine can vary in this scenario, with two expectations: the machine is used to return a similarity score from some probe image for each image in the gallery, and the human examines the top-K matches (as opposed to only returning the closest match). This article discusses additional scenarios in which a man or machine can be used to improve the prospects of a successful face identification. We separate such methods into two main categories. The first approach uses pre-processing methods to improve the quality of a face image prior to submission to a COTS FRS. These methods do not require any changes to existing systems, but they are limited in that they can only modify the input face image itself and not the features to be extracted from the face images. Researchers have developed pre-processing methods with the ability to improve the facial pose[6,7] and image illumination, [7] alter the subject's age,[8] and improve the image resolution.[9] The second approach is to design special purpose face-recognition systems for a specific matching problem. These methods let system designers appropriately modify any of the modules of the face-recognition process (such as feature representation). We discuss special-purpose face-recognition systems for problems where pre-processing is infeasible, such as forensic sketch recognition [10] and matching and

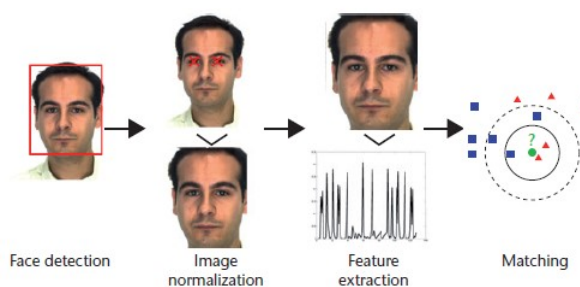


Figure 1. Major steps in automatic face recognition. Automated face recognition systems (FRSs) involve finding key facial landmarks for alignment, normalizing the appearance of the face, choosing a suitable feature representation, learning

discriminative feature combinations, and developing accurate and scalable matching schemes.

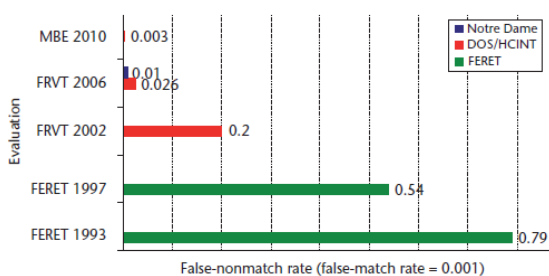


Figure 2. The reduction in error rate for state-of-the-art face-recognition algorithms as documented through the FERET, FRVT 2002, FRVT 2006, and MBE 2010 face evaluations conducted by NIST. Performance is shown separately for the FERET, DOS/HCINT, and the Notre Dame FRVT 2006 datasets.[3]

retrieval using soft biometrics such as facial marks, scars, and demographics.[11] We also discuss a face-recognition system for facial aging that augments the aging pre-processing approach.[12] The four specific problems that we address in this article—sketch recognition; aging, facial scars, and marks; pose and illumination correction; and matching low-resolution images—reflect a body of research that has been directly motivated by recent research progress and that has been the most difficult for law enforcement agencies to overcome.

Pre-processing Approach to Forensic FR

Pre-processing methods in forensic face recognition receive a face image as input, and they output an enhanced-quality face image. The key benefit to such methods is they are compatible with the COTS FRS already in use by law enforcement agencies. Volker Blanz and Thomas Vetter proposed a 3D morphable model that lets a user correct a face image that is off pose and compensate for adverse illumination conditions.[6] Their 3D morphable model uses a training set of face images to learn the distribution of 3D facial shape and texture in a parameterized feature space. Variations in the feature space alter the facial texture, shape, pose, and the 3D illumination

model. Given a 2D face image, gradient descent is performed on the difference between the 2D face image and the texture of the 3D morphable model. At convergence, the 2D face image is parameterized and thus controlled in a 3D graphics environment.

This lets users adjust the pose of the 3D face to frontal and set the illumination to ideal ambient conditions. The accuracy of morphable models is generally improved by manually adjusting the initial alignment between the input 2D image and the 3D morphable model. Often, many off-pose and low-quality face images are available from surveillance video frames. Users can manually inspect these multiple image frames to find the highest quality frame, which they can then feed into the COTS FRS. Alternatively, Unsang Park and Anil Jain used a structure from motion algorithm that utilizes facial landmarks obtained from video sequences to infer the 3D face shape.[7] The factorization method factorizes a set of 2D landmark points, W , to initially estimate the rotation matrix, M' , and 3D shape, S' . These initial estimates of M' and S' are adjusted by a correction matrix, A , obtained from the orthogonal constraint on the true rotation matrix $M = M' \times A$. The true 3D shape is obtained as $S = A^{-1} \times S'$. The reconstructed 3D model can be used to generate the corrected frontal face images for improved identification accuracy[7] (see Figure 6).

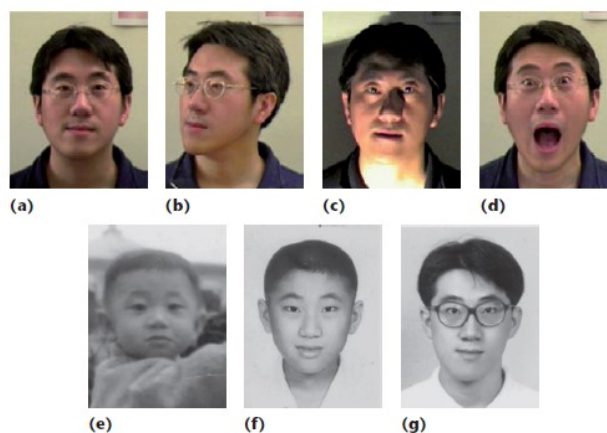


Figure 3. Four key factors compromising face-recognition accuracy are pose, illumination, expression, and aging variations. Face images of one subject are shown under (a) an ideal capture condition and with (b) pose, (c) illumination, and (d)

expression. We also show aging variations for the subject in (a) for (e) 32, (f) 21, and (g) 15 years younger. Often, these factors do not occur in isolation, further compounding the face-recognition problem.

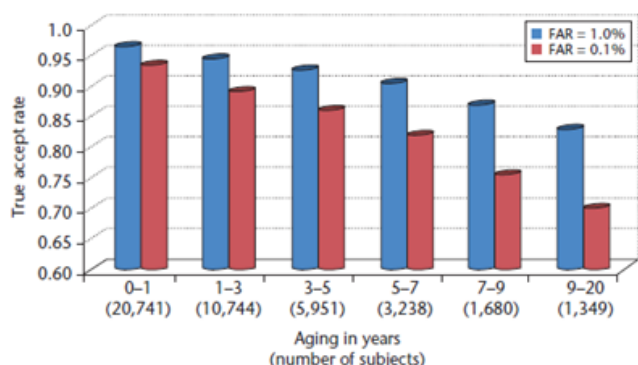


Figure 4. Accuracy of a leading COTS FRS as a function of the time lapse between the probe and gallery images across a mug shot database containing 94,631 face images of 28,031 subjects. The degradation in recognition accuracy suggests that face-recognition systems might need to be specially designed to match across large age variations.

In addition to correcting pose and illumination variations, face-pre-processing algorithms can simulate the facial aging process. Unsang Park, Yiyang Tong, and Anil Jain proposed a generative 3D aging modelling method where the input image is projected into the parametric 3D aging pattern space and simulated face images are generated at target ages (see Figure 6c). The ability to take face images and simulate aging (or de-aging) is useful in cold case investigations, missing children cases, and de-duplication of government-issued documents.

Another problem often encountered in forensic face recognition involves low-resolution face images. A common source of such images in law enforcement units is face images that have been faxed, printed, or heavily compressed (see Figure 7). Thirimachos Bourlai, Arun Ross, and Anil Jain proposed a method that applies image filtering, linear denoising, and thresholding-based nonlinear denoising methods to enhance the quality of the low resolution images for improved matching accuracy.[9] Manual or interactive enhancement techniques can also be used

for pre-processing. For example, Figure 8 shows the result of forensic artists inferring accurate facial appearances from low-quality videos in a recent case in Los Angeles.[13] Two of the suspects shown were manually identified after the sketches were posted in public places. Using specialized sketch recognition further improves the ability to perform identification from the original low quality images.

Special-Purpose FRSs

In certain face-recognition scenarios, image enhancement alone is insufficient for automated identification. Instead, specially designed face recognition algorithms are needed in place of the COTS FRSs. These systems let users tailor any aspect of the face-recognition process (such as the feature representation and statistical learning) to the given scenario.

Table 1. Challenges in forensic face recognition for various purposes.

Sources	Pose	Illumination	Expression	Aging	Heterogeneous
Nonforensic					
Access control		X			
De-duplication				X	
Border control		X			
Forensic					
Missing person	X		X	X	
Child trafficking	X	X	X	X	
Surveillance	X	X	X		
Forensic sketch					X

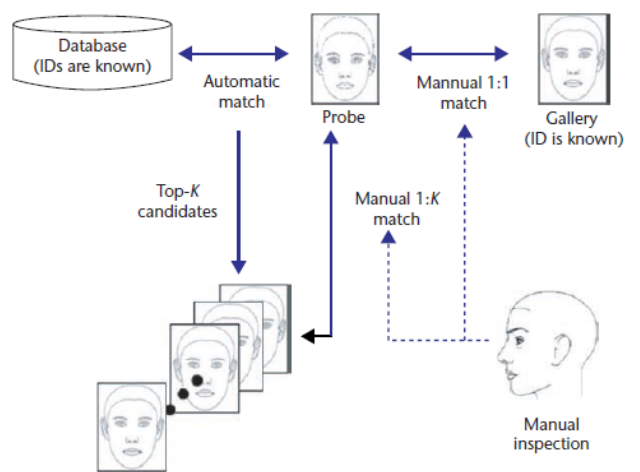


Figure 5. Forensic face recognition. (a) In many forensic scenarios, face recognition is not yet fully automatic and requires manual intervention during

pre-processing and an examination of top-K retrieved faces from a large gallery. (b) Typical face images encountered in forensic face recognition could include, from left to right, a mug shot, deceased subject, sketch, video frame, and near infrared (NIR) images.

Facial Aging

In addition to simulating the facial-aging process to improve a generic COTS FRS,⁸ Zhifeng Li, Unsang Park, and Anil Jain developed a specialized FRS to compensate for aging.^[12] Their approach uses a discriminative aging model to learn a robust face representation. The discriminative model is trained on a set of age separated image pairs using scale-invariant feature transformation (SIFT) and multiscale local binary pattern (MLBP) descriptors as well as random-sampling linear discriminate analysis (LDA) subspace analysis. Combining the discriminative-aging modelling method with the generative pre-processing method we discussed earlier offers significant improvements in identification accuracy over a leading COTS FRS (see Figure 9). Rank-1 indicates the top match out of all matching candidates.

Forensic-Sketch Recognition

Automated identification of a subject based on a composite sketch query expands face recognition capabilities to situations where a suspect's face image is not available from the crime scene. In such situations, only a verbal description of a subject, provided by witnesses or victims, is available for use by a forensic sketch artist or a composite software tool to generate a depiction of the subject's facial appearance. Forensic sketches have been successfully used for more than a century in criminal identification; however, the paradigm for using a forensic sketch has been limited to dissemination of the sketch to the media and law enforcement agencies with the hopes that citizens will provide tips to enable the suspect's arrest. Despite the strengths of COTS FRSs in matching photographs, their ability to match forensic sketches to face photographs is severely limited. To fill this void, Brendan Klare, Zhifeng Li, and Anil Jain designed a

FRS for this task called local feature-based discriminate analysis (LFDA).¹⁰ LFDA operates by representing both forensic sketches and photographs using SIFT and MLBP feature descriptors. A column-wise concatenation of these descriptors are used to learn discriminate subspace projections that attempt to maximize the Fisher criterion, where the within-class feature spaces consist of both a sketch and photo from the same subject.

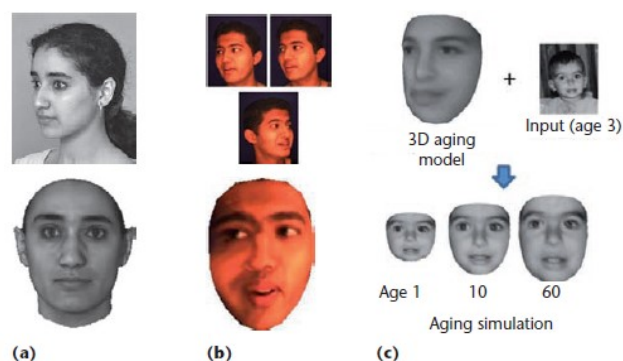


Figure 6. Pre-processing methods. (a) A 3D morphable model allows a single off pose face image to be fit to a 3D model, correcting the pose to a frontal view for improved face-recognition performance.⁶ (b) Given multiple off-pose face images of a subject, a structure from motion algorithm can infer the 3D face shape to generate frontal pose face image.^[7] Tools such as these offer forensic investigators methods for normalizing a face image prior to submitting for a match against large face databases. (c) Aging simulation methods let investigators use the age-progressed (or regressed) output from the aging model to improve face-recognition accuracies across existing COTS FRS configurations.



Figure 7. Degraded and low-resolution probe face images: (a) original image, (b) JPEG compressed (medium quality), (c) resized to 10 percent and up-scaled to the original spatial resolution, (d) half-toning, and (e) scanned passport.

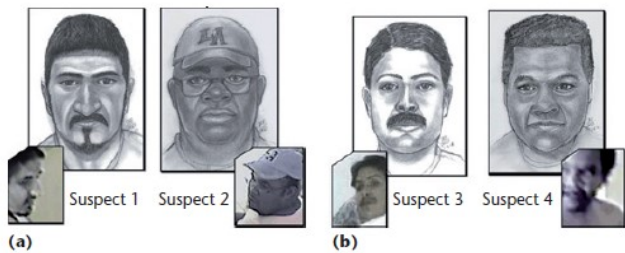


Figure 8. The detailed sketches of four suspects drawn by Sandra Enslow from low-resolution surveillance video in a high-profile Los Angeles Sheriff's Department case.¹³ An example of the highest quality frame available for each subject is shown below each sketch. This manual enhancement technique is useful for both human-based identification and recognition using a specially designed sketch-recognition system.

Forensic-sketch recognition is an example of a heterogeneous face-recognition (HFR) problem, where the probe and gallery images are from different imaging modalities. As the number of imaging devices increases (such as near infrared, thermal infrared, and LIDAR [light detection and ranging] sensors), a lack of sensor interoperability can impact the face recognition performance. To fill this void, Brendan Klare and Anil Jain proposed a generic HFR framework using kernel prototype similarities.^[14] In addition to the aforementioned LDA based method [10,12] Haibin Ling and his colleagues proposed a gradient orientation pyramid (GOP)-based method for age-invariant face recognition.^[15] To handle the sketch recognition problem, Xiaogang Wang and Xiaoou Tang proposed a Markov random field approach to synthesize a photograph from a facial sketch.^[16]

Facial-Mark-Based Matching and Retrieval

Facial marks consist of temporarily persistent skin irregularities, such as scars, moles, and freckles. Various scenarios exist where a specially designed system can explicitly leverage this information. The first such use is to supplement the similarity score from a COTS FRS, which Park and Jain demonstrated

to improve recognition accuracy.^[11] Another scenario is a face-retrieval system, where facial mark information in the probe image, along with demographic information, can filter a large gallery database. The use of facial marks also enables verbal-description-based retrieval—for example, “Find all faces with a large dark scar on right cheek.” Facial marks can also help to individualize identical twins. In fact, in surveillance videos, where the face image often is of low resolution and not frontal, facial marks are the only strong source of information to identify the suspect. Face images identified based on marks serve as valuable evidence in legal testimony.

Park and Jain developed an automatic facial mark detection method based on the Laplacian of Gaussian operator, $\nabla^2 G(x, y, \sigma)$, in scale space.^[11] A scale-space representation, $D(x, y, \sigma)$, is obtained by convolving the normalized Laplacian of Gaussian operator, $\sigma^2 \nabla^2 G(x, y, \sigma)$, with the input image, $I(x, y)$: $D(x, y, \sigma) = \sigma^2 \nabla^2 G(x, y, \sigma) * I(x, y)$. From a multiscale space representation with scales or standard deviations $(\sigma_1, \dots, \sigma_k)$, local maxima and minima are extracted depending on whether $D(x, y, \sigma)$ is greater or smaller than pixels in a neighbourhood. Figure 10 shows examples of automatic mark-detection results. The size of each circle, ellipse, and rectangles represent the scale (s) at which a mark is detected, which is proportional to each mark's size. Forensic experts often use facial marks to verify a suspect against a candidate face image. Park and Jain's face-image matching and retrieval system provides tools, such as manual and automatic mark labelling; image retrieval using facial marks and demographic information; and an interactive interface for the analyze, compare, evaluate, and verify (ACE-V) forensics operation.^[17]

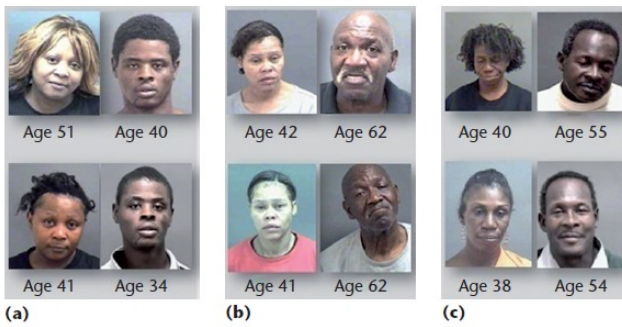


Figure 9. Some examples of successful face retrieval at rank-1. The first row shows the probe images and the second row shows the successfully retrieved or matched images from the gallery. (a) The first set shows two examples where COTS FRS and the generative method fail, while the discriminative method succeeds. (b) The second set shows two examples where the discriminative method fails, while the COTS FRS and the generative method succeed. (c) The third set shows two examples where all three methods (COTS FRS, generative model, and discriminative model) fail, but the score-level fusion of the generative model and the discriminative models succeed. Both the probe and gallery contained 10,000 (different) images of 10,000 different subjects.

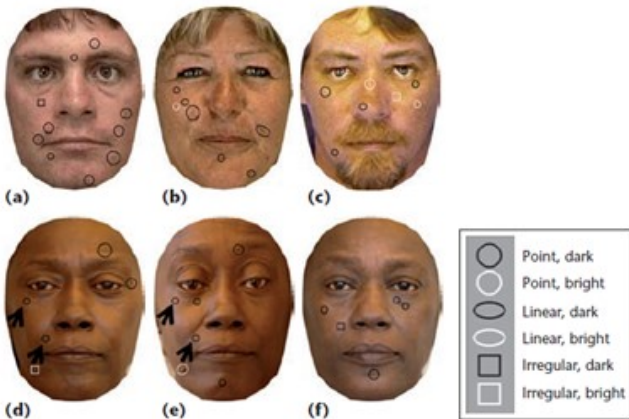


Figure 10. Automatic mark-detection results on face images of three subjects, (a) through (c), and one pair of identical twins, (d) and (f). All significant marks were successfully detected and have been marked using circles, ellipse, and rectangles. For the identical twin pair, the first two images, (d) and (e), show one individual, followed by (f) the image of the twin sister. Using COTS FRS only, the system incorrectly matched the twin sister, but this match was

successfully excluded using COTS FRS and facial marks. The black arrows show the two marks that confirmed the correct match.

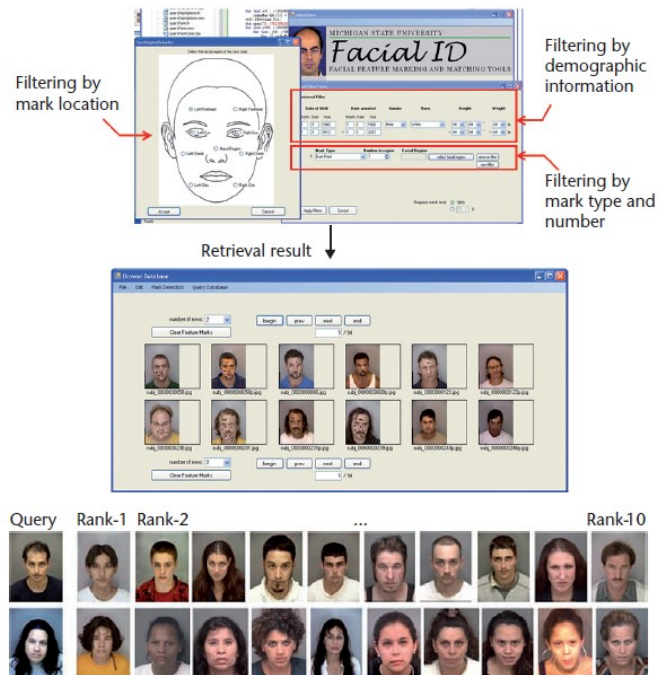


Figure 11. Face-image retrieval system that can filter a large database based on facial-mark type, number, and locations as well as demographic information. (a) System GUI and retrieval results. (b) Two example queries where the correct mates were not retrieved at rank-1 using COTS FRS only. Correct mates were successfully retrieved at rank-1 using COTS FRS and facial marks (first row) and COTS FRS, facial marks, and gender and ethnicity (second row).

Figure 11 shows the retrieval system and two example retrieval results from 100,000 gallery images, where correct mates were found at rank-1 by using facial marks and demographic information in addition to a COTS FRS. In Figure 11b, the use of gender and ethnicity successfully filtered 99.7 percent of the database, which improved the matching accuracy with reduced computation time. The rank-1 matching accuracy using COTS FRS only; COTS FRS and facial marks; and COTS FRS, facial marks, and demographic information (such as gender and ethnicity) are 56.3, 57.1, and 57.7 percent, respectively, using 1,000 probe images and 100,000 gallery images. A facial-mark retrieval system is a

prime example of how forensic face recognition differs from other “lights out” (fully automatic) face recognition applications. Human operators must first verify and then utilize the retrieval system’s results. FRSs are not intended to replace forensic experts but should augment their capabilities, particularly for difficult probe face images.

II. CONCLUSION

The need for the systems we present here was realized through close collaboration with the law enforcement community. The users of such systems (such as forensic investigators) are generally the best source for determining future avenues of forensic face-recognition research that will have the most impact on identifying suspects and victims. Thus, fostering close collaboration between the pattern recognition community and law enforcement is critical to further advance the state of the art in forensic face recognition. One of the most important aspects of forensic face-recognition systems is that they are not always successful in matching faces. This leads to the most critical area of human interaction in the forensic face-recognition process, namely, interpreting the results. The top (rank-1) match returned should not always be given precedence over (say) the rank-10 match. Instead, all the top-K retrieved results should be carefully considered. Forensic experts indicate that they routinely use a K value in the 100 to 200 range, examining up to top 200 returned matches for culpability. Often, most of these results can be quickly discarded using information such as incarnation status or demographics. Another reason why face-recognition results in a forensic setting must be used with caution is that, despite being designed using elements of statistical decision theory, face matchers cannot generate a probability of false match. This deficiency occurs because researchers have yet to develop face individuality models. Klare and Jain proposed organizing facial features into three distinct levels that follow the convention used in fingerprint recognition.¹⁸ This feature organization is meant to

serve as a precursor to face individuality studies. In the meantime, the use of face-recognition results in legal proceedings is severely limited. Forensic face recognition can only point to a suspect; from there, additional evidence must be found for a conviction.

III. REFERENCES

- [1]. J. Vlahos, 'Surveillance Society: New High-Tech Cameras Are Watching You,' *Popular Mechanics*, 1 Oct. 2009; www.popularmechanics.com/technology/military/4236865.
- [2]. S.Z. Li and A.K. Jain, eds., *Handbook of Face Recognition*, 2nd ed., Springer, 2011.
- [3]. P. Grother, G. Quinn, and P.J. Phillips, 'Report on the Evaluation of 2D Still-Image Face Recognition Algorithms,' Nat'l Inst. of Standards and Technology interagency/internal report (NISTIR) 7709, 2010; www.nist.gov/customcf/get_pdf.cfm?pub_id=905968.
- [4]. R. Charette, 'Here's Looking at You, and You, and You. . .,' *blog*, *IEEE Spectrum*, 25 July 2011; <http://spectrum.ieee.org/riskfactor/computing/int/heres-looking-at-you-and-you-and-you->
- [5]. L. Ding et al., 'Computers Do Better than Experts Matching Faces in a Large Population,' *Proc. IEEE Int'l Conf. Cognitive Informatics*, IEEE Press, 2010, pp. 280-284.
- [6]. V. Blanz and T. Vetter, 'Face Recognition Based on Fitting a 3D Morphable Model,' *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, 2003, pp. 1063-1074.
- [7]. U. Park and A.K. Jain, '3D Model-Based Face Recognition in Video,' *Proc. 2nd Int'l Conf. Biometrics (ICB)*, LNCS 4642, Springer, 2007, pp. 1085-1094.
- [8]. U. Park, Y. Tong, and A.K. Jain, 'Age Invariant Face Recognition,' *IEEE Trans. Pattern*

- Analysis and Machine Intelligence, vol. 32, no. 5, 2010, pp. 947-954.
- [9]. T. Bourlai, A. Ross, and A.K. Jain, 'Restoring Degraded Face Images: A Case Study in Matching Faxed, Printed and Scanned Photos,' IEEE Trans. Information Forensics and Security, vol. 6, no. 2, 2011, pp. 371-384.
- [10]. B. Klare, Z. Li, and A.K. Jain, 'Matching Forensic Sketches to Mugshot Photos,' IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 33, no. 3, 2011, pp. 639-646.
- [11]. U. Park and A.K. Jain, 'Face Matching and Retrieval Using Soft Biometrics,' IEEE Trans. Information Forensics and Security, vol. 5, no. 3, 2010, pp. 406-415.
- [12]. Z. Li, U. Park, and A.K. Jain, 'A Discriminative Model for Age Invariant Face Recognition,' IEEE Trans. Information Forensics and Security, vol. 6, no. 3, 2011, pp. 1028-1037.
- [13]. 'Los Angeles Officials Identify Video Assault Suspects,' New York Times, 7 Jan. 2011; www.nytimes.com/2011/01/08/us/08disabled.html.
- [14]. B. Klare and A. K. Jain, 'Heterogeneous Face Recognition using Kernel Prototype Similarities,' tech. report MSU-CSE-11-18, Michigan State Univ., Nov. 2011.
- [15]. H. Ling et al., 'Face Verification Across Age Progression Using Discriminative Methods,' IEEE Trans. Information Forensic and Security, vol. 5, no. 1, 2010, pp. 82-91.
- [16]. X. Wang and X. Tang, 'Face Photo-Sketch Synthesis and Recognition,' IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 31, no. 11, 2009, pp. 1955-1967.
- [17]. H. Tuthill and G. George, Individualization: Principles and Procedures in Criminalistics, Lightning Powder Company, 2002.
- [18]. B. Klare and A.K. Jain, 'On a Taxonomy of Facial Features,' Proc. 4th IEEE Int'l Conf. Biometrics: Theory, Applications, and Systems (BTAS), IEEE Press, 2010, pp. 1-8.