

# Memetically Optimized MCWLD for Matching Sketches with Digital Face Images

Himanshu S. Bhatt, *Student Member, IEEE*, Samarth Bharadwaj, *Student Member, IEEE*, Richa Singh, *Member, IEEE*, Mayank Vatsa, *Member, IEEE*.

**Abstract**—One of the important cues in solving crimes and apprehending criminals is matching sketches with digital face images. This paper presents an automated algorithm that extracts discriminating information from local regions of both sketches and digital face images. Structural information along with minute details present in local facial regions are encoded using multi-scale circular Weber’s local descriptor. Further, an evolutionary memetic optimization is proposed to assign optimal weights to every local facial region to boost the identification performance. Since, forensic sketches or digital face images can be of poor quality, a pre-processing technique is used to enhance the quality of images and improve the identification performance. Comprehensive experimental evaluation on different sketch databases show that the proposed algorithm yields better identification performance compared to existing face recognition algorithms and two commercial face recognition systems.

**Index Terms**—Sketch Recognition, Face Recognition, Memetic Algorithms, Forensic Sketch, Weber’s Local Descriptor.

## I. INTRODUCTION

FACE recognition is a well studied problem in many application domains. However, matching sketches with digital face images is a very important law enforcement application that has received relatively less attention. Forensic sketches are drawn based on the recollection of an eye-witness and the expertise of a sketch artist. As shown in Fig. 1, forensic sketches include several inadequacies because of the incomplete or approximate description provided by the eye-witness. Generally, forensic sketches are manually matched with the database comprising digital face images of known individuals. The state-of-art face recognition algorithms cannot be used directly and require additional processing to address the non-linear variations present in sketches and digital face images. An automatic sketch to digital face image matching system can assist law enforcement agencies and make the recognition process efficient and relatively fast.

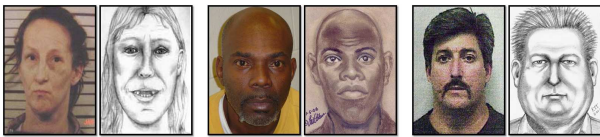


Fig. 1. Examples showing exaggeration of facial features in forensic sketches.

## A. Literature Review

Sketch recognition algorithms can be classified into two categories: *generative* and *discriminative* approaches. Generative approaches model a digital image in terms of sketches and then match it with the query sketch or vice-versa. On the other hand, discriminative approaches perform feature extraction and matching using the given digital image and sketch pair and do not generate the corresponding digital image from sketches or the sketch from digital images.

1) *Generative Approaches*: Wang and Tang [1] proposed Eigen transformation based approach to transform a digital photo into sketch before matching. In another approach, they presented an algorithm to separate shape and texture information and applied Bayesian classifier for recognition [2]. Liu *et al.* [3] proposed non-linear discriminative classifier based approach for synthesizing sketches by preserving face geometry. Li *et al.* [4] converted sketches to photos and used a method similar to Eigen-transform for matching. Gao *et al.* [5] synthesized a series of pseudo sketches using embedded Hidden Markov Models (E-HMM) which are then fused to generate a pseudo sketch. Wang and Tang [6] proposed Markov Random Fields based algorithm to automatically synthesize sketches from digital face images and vice-versa. Zhang *et al.* [7] extended multiscale Markov Random Field (MRF) model to synthesize sketches under varying pose and lighting conditions.

2) *Discriminative Approaches*: Uhl and Lobo [8] proposed photometric standardization of sketches to compare it with digital photos. Sketches and photos were geometrically normalized and matched using Eigen analysis. Yuen and Man [9] used local and global feature measurements to match sketches and mug-shot images. Zhang *et al.* [10] compared the performance of humans and PCA-based algorithm for matching sketch-photo pairs with variations in gender, age, ethnicity, and inter-artist variations. They discussed about the quality of sketches in terms of artist’s skills, experience, exposure time, and distinctiveness of features. Similarly, Nizami *et al.* [11] analyzed the effect of matching sketches drawn by different artists. Klare and Jain [12] proposed a Scale Invariant Feature Transform (SIFT) based local feature approach where sketches and digital face images were matched using the gradient magnitude and orientation within the local region. Bhatt *et al.* [13] extended Uniform Local Binary Patterns to incorporate exact difference of gray level intensities to encode texture features in sketches and digital face images. Klare *et al.* [14] extended their approach using Local Feature Discriminant

Analysis (LFDA) to match forensic sketches. In their recent approach, Klare and Jain [15] proposed a framework for heterogeneous face recognition where both probe and gallery images are represented in terms of non-linear kernel similarities. Zhang *et al.* [16] analyzed the psychological behavior of humans for matching sketches drawn by different sketch artists. Recently, Zhang *et al.* [17] proposed an information theoretic encoding band descriptor to capture discriminative information and random forest based matching to maximize the mutual information between a sketch and a photo.

### B. Research Contributions

After discussing with several sketch artists, it is observed that *generating a sketch is an unknown psychological phenomenon, however, a sketch artist generally focusses on the local facial features and texture which he/she tries to embed in the sketch through a blend of soft and prominent edges.* Therefore, the proposed algorithm is designed based on the following observations:

- information vested in local facial regions can have high discriminating power;
- facial patterns in sketches and digital face images can be efficiently represented by local descriptors.

This research proposes an automatic algorithm for matching sketches with digital face images using the modified Weber's local descriptor (WLD) [18]. WLD is used for representing images at multiple scales with circular encoding. The multi-scale analysis helps in assimilating information from minute features to the most prominent features in a face. Further, memetically optimized  $\chi^2$  distance measure is used for matching sketches with digital face images. The proposed matching algorithm improves the performance by assigning optimal weights to local facial regions. To further improve the performance, a Discrete Wavelet Transform (DWT) [19] fusion based pre-processing technique is presented to enhance the forensic sketch-digital image pairs. Moreover, in this research, three different types of sketches are used for performance evaluation.

- 1) Viewed sketches, drawn by a sketch artist while looking at the digital image of a person.
- 2) Semi-forensic sketches, drawn by a sketch artist based on his recollection from the digital image of a person.
- 3) Forensic sketches, drawn based on the description of an eyewitness from his recollection of the crime scene.

The major contributions of this research can be summarized as follows:

- 1) Previous approaches for matching forensic sketches [14] manually separate *good* and *bad* forensic sketches and generally focus on *good* forensic sketches only. Such a classification is often based on the similarity between the sketch and corresponding digital face image. Since the corresponding digital face image is not available in real-time applications, selecting good and bad forensic sketches is not pragmatic for matching forensic sketches with digital face images. In this research, a pre-processing technique is presented for enhancing

the quality of forensic sketch-digital image pairs. Pre-processing forensic sketches enhances the quality and improves the performance by at least 2 – 3%.

- 2) Multi-scale Circular WLD and memetically optimized  $\chi^2$  based algorithms are proposed for matching sketches with digital face images. The proposed algorithm outperforms existing approaches on different sketch databases.
- 3) To better understand the progression from viewed to forensic sketches, semi-forensic sketches are introduced that bridge the gap between viewed and forensic sketches. In the experiments, it is observed that training sketch recognition algorithms (existing as well as the proposed) on semi-forensic sketches improves the rank-1 identification performance by at least 4% compared to the traditional method of training on viewed sketches.
- 4) Human performance for matching sketches with digital face images is also analyzed. The information collected from individuals corroborate with our initial observation that local regions provide discriminating information.
- 5) The paper also presents a part of the IIIT-Delhi database<sup>1</sup> (Viewed and Semi-forensic Sketch database) and 61 forensic sketch-digital image pairs to the research community to promote the research in this domain.

The paper is organized as follows: Section II describes the pre-processing technique for enhancing forensic sketch-digital image pairs. Section III-A presents Multi-scale Circular WLD (MCWLD) and Section III-B explains memetic optimization for matching sketches with digital face images using weighted  $\chi^2$  distance. Section IV presents the three types of sketch databases used in this research. Sections V to VII present comprehensive experimental results and key observations.

## II. PRE-PROCESSING ALGORITHM

In sketch to digital face image matching, researchers have generally used viewed sketches where the quality of sketch-digital image pair is very good. On these good quality viewed sketches, the state-of-art is about 99% (rank-1) identification accuracy while the state-of-art in forensic sketch recognition is about 16%. One of the reasons for low recognition performance is that forensic sketches may contain distortions and noise introduced due to the excessive use of charcoal pencil, paper quality, and scanning (device noise/errors). Furthermore, digital images in the gallery may also be noisy and of sub-optimal quality because of printing and scanning of images. As shown in Fig. 2, forensic sketch-digital image pairs of lower visual quality may lead to reduced matching performance as compared to good quality sketch-digital image pairs.

In this research, a pre-processing technique is presented that enhances the quality of forensic sketch-digital image pairs. The steps involved in the pre-processing technique are as follows:

- Let  $f$  be the color face image to be enhanced. Let  $f^r$  and  $f^y$  be the red and luma channels<sup>2</sup> respectively. These two channels are processed using the multi-scale

<sup>1</sup><http://research.iiitd.edu.in/groups/iab/sketchDatabase.html>.

<sup>2</sup>In the watermarking literature, it is well established that red and luma channels are relatively less sensitive to the visible noise, therefore, these channels are used for enhancement.

retinex (MSR) algorithm [20] with four iterations. MSR is applied on both red and luma channels to obtain  $f^{rm}$  and  $f^{ym}$ .

- $f^{rm}$  and  $f^{ym}$  are subjected to wavelet based adaptive soft thresholding scheme [21] for image denoising. The algorithm computes generalized Gaussian distribution based soft threshold which is used in wavelet based denoising to obtain  $f^{rm'}$  and  $f^{ym'}$  respectively.
- Noise removal in the previous step may lead to blurring of edges. Experiments show that a symmetric low-pass filter of size  $7 \times 7$  with standard deviation of 0.5 efficiently restores the genuine facial edges. Applying this (Wiener) filter on  $f^{rm'}$  and  $f^{ym'}$  produces  $f^1$  and  $f^2$ .
- After computing the globally enhanced red and luma channels, DWT fusion algorithm is applied on  $f^1$  and  $f^2$  to compute a feature rich and enhanced face image,  $F$ . Single level DWT (with db 9/7 mother wavelet) is applied on  $f^1$  and  $f^2$  to obtain the detailed and approximation bands of these images. Let  $f_{LL}^j$ ,  $f_{LH}^j$ ,  $f_{HL}^j$ , and  $f_{HH}^j$  be the four subbands where  $j = 1, 2$ ,  $LL$  represents the approximation band, and  $LH$ ,  $HL$ , and  $HH$  represent the detailed subbands. To preserve features of both the channels, coefficients from the approximation band of  $f^1$  and  $f^2$  are averaged.

$$f_{LL}^e = \text{mean}(f_{LL}^1, f_{LL}^2) \quad (1)$$

where  $f_{LL}^e$  is the approximation band of the enhanced image. All three detailed subbands are divided into windows of size  $3 \times 3$  and the sum of absolute pixels in each window is calculated. For the  $i^{th}$  window in  $HL$  subband of the two images, the window with maximum absolute value is selected to be used for enhanced subband  $f_{HL}^e$ . Similarly, enhanced subbands  $f_{LH}^e$  and  $f_{HH}^e$  are also obtained. Finally, inverse DWT is applied on the four subbands to generate a high quality face image.

$$F = IDWT(f_{LL}^e, f_{LH}^e, f_{HL}^e, f_{HH}^e) \quad (2)$$

This DWT fusion algorithm is applied on both forensic sketches and digital face images. Fig. 3 shows quality enhanced forensic sketches and digital face images. Note that the pre-processing technique enhances the quality when there are irregularities and noise in the input image, however, it does not alter good quality face images (i.e. sketch-digital image pairs from the viewed sketch database). Sketches are scanned as three channel color images. Further, the forensic images obtained from different sources are also three channel color images. If a gray scale image is obtained, multi-scale retinex and Wiener filtering are applied only on the single channel. Along with quality enhancement, face images are geometrically normalized as well. The eye-coordinates are detected using the OpenCV's boosted cascade of haar-like features. Using the eye-coordinates, rotation is normalized with respect to the horizontal axis and inter-eye distance is fixed to 100 pixels. Finally, the face image is resized to  $192 \times 224$  pixels.

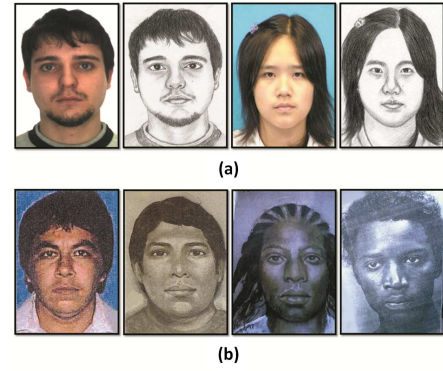


Fig. 2. Paper quality, sensor noise, and old photographs can affect the quality of sketch-digital image pairs and hence reduce the performance of matching algorithms. (a) Good quality sketch-digital image pairs (CUHK database) and (b) poor quality sketch-digital image pairs (Forensic sketch database).

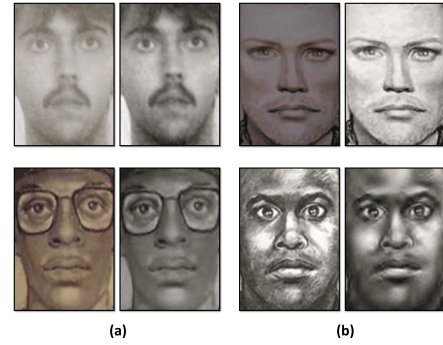


Fig. 3. Quality enhancement using the pre-processing technique. (a) represents digital face image before and after pre-processing and (b) represents forensic sketches before and after pre-processing.

### III. MATCHING SKETCHES WITH DIGITAL FACE IMAGES

Local descriptors have received attention in face recognition due to their robustness to scale, orientation, and speed. Local Binary Patterns (LBP) is one of the widely used descriptors for face recognition [22]. In face recognition literature, several variants of LBP have been proposed. Bhatt *et al.* [13] extended LBP to incorporate exact difference of gray level intensities among pixel neighbors and used it for sketch recognition. Local descriptors such as LBP are generally used as dense descriptors where texture features are computed for every pixel of the input face image. On the other hand, sparse descriptor such as Scale Invariant Feature Transform (SIFT) [23] is based on interest point detection and computing the descriptor in the vicinity of detected interest points. SIFT is computed using the gradient and orientation of neighboring points sampled around every detected key point. As a sparse descriptor, SIFT has been used for face recognition by Geng and Jiang [24]. Klare and Jain [12] applied SIFT in a dense manner (i.e. computing SIFT descriptor at specific pixels) for matching sketches with digital face images. It is our assertion that local descriptors can be used for representing sketches and digital face images because they can efficiently encode the discriminating information present in the local regions.

Recently, Chen *et al.* [18] proposed a new descriptor, Weber's local descriptor, which is based on Weber's law and draws its motivation from both SIFT and LBP. It is similar to SIFT in computing histogram using gradient and orientation,



and analogous to LBP in being computationally efficient and considering small neighborhood regions. However, WLD has some unique features that make it more efficient and robust as compared to SIFT and LBP. WLD computes the salient micro patterns in a relatively small neighborhood region with finer granularity. This allows it to encode more discriminative local micro patterns. In this research, WLD is optimized for matching sketches with digital face images by computing multi-scale descriptor in a circular manner (in contrast to the originally proposed square neighborhood approach). Finally, two multi-scale circular WLD (MCWLD) histograms are matched using memetically optimized weighted  $\chi^2$  distance.

#### A. Feature Extraction using MCWLD

MCWLD has two components: 1) *differential excitation* and 2) *gradient orientation*. MCWLD representation for a given face image is constructed by tessellating the face image and computing a descriptor for each region. As shown in Fig. 4, MCWLD descriptor is computed for different values of parameters  $P$  and  $R$ , where  $P$  is the number of neighboring pixels evenly separated on a circle of radius  $R$  centered at the current pixel. Multi-scale analysis is performed by varying the radius  $R$  and number of neighbors  $P$ . Sketches and digital face images are represented using MCWLD as explained below:

1) *Differential Excitation*: Differential excitation is computed as an arctangent function of the ratio of intensity difference between the central pixel and its neighbors to the intensity of central pixel. The differential excitation of central pixel  $\xi(x_c)$  is computed as:

$$\xi(x_c) = \arctan \left\{ \sum_{i=0}^{P-1} \left( \frac{x_i - x_c}{x_c} \right) \right\} \quad (3)$$

where  $x_c$  is the intensity value of central pixel and  $P$  is the number of neighbors on a circle of radius  $R$ . If  $\xi(x_c)$  is positive, it simulates the case that surroundings are lighter than the current pixel. In contrast, if  $\xi(x_c)$  is negative, it simulates the case that surroundings are darker than the current pixel.

2) *Orientation*: The orientation component of WLD is computed as:

$$\theta(x_c) = \arctan \left\{ \frac{x_{(\frac{P}{2}+R)} - x_{(R)}}{x_{(P-R)} - x_{(\frac{P}{2}-R)}} \right\} \quad (4)$$

The orientation is further quantized into  $T$  dominant orientation bins where  $T$  is experimentally set as eight.

3) *Circular WLD Histogram*: For every pixel, differential excitation ( $\xi$ ) and orientation ( $\theta$ ) are computed using Eqs. 3 and 4 respectively. As shown in Fig. 5, a 2D histogram of circular WLD feature,  $CWLD(\xi_j, \theta_t)$ , is constructed where  $j = 0, 1, \dots, N-1$ ,  $t = 0, 1, \dots, T-1$ , and  $N$  is the dimension of the image. Each column in the 2D histogram corresponds to a dominant orientation,  $\theta_t$ , and each row corresponds to a differential excitation interval. Thus, the intensity of each cell corresponds to the frequency of a certain differential excitation interval in a dominant orientation. Similar to Chen *et al.* [18], a four step approach is followed to compute CWLD descriptor.

Step-1: The 2D histogram  $CWLD(\xi_j, \theta_t)$  is further encoded into 1D histograms. Differential excitations,  $\xi$ , are regrouped into  $T$  orientation sub-histograms,  $H(t)$ , where  $t = 0, 1, \dots, T-1$  corresponds to each dominant orientation.

Step-2: Within each dominant orientation, range of differential excitation is evenly divided into  $M$  intervals and then reorganized into a histogram matrix. Each orientation sub-histogram in  $H(t)$  is thus divided into  $M$  segments,  $H_{m,t}$  where  $m = 0, 1, \dots, M-1$  and  $M = 6$ . For each differential excitation interval  $l_m$ , lower bound is computed as  $\eta_{m,l} = (m/M - 1/2)\pi$  and upper bound  $\eta_{m,u}$  is computed as  $\eta_{m,u} = [(m+1)/M - 1/2]\pi$ .

Each sub-histogram segment  $H_{m,t}$  is further composed of  $S$  bins and is represented as:

$$H_{m,t} = h_{m,t,s} \quad (5)$$

where  $s = 0, 1, \dots, S-1$ ,  $S = 3$  and  $h_{m,t,s}$  is represented as:

$$h_{m,t,s} = \sum_j \delta(S_j == s), \left( S_j = \left\lfloor \frac{\xi_j - \eta_{m,l}}{(\eta_{m,n} - \eta_{m,l})/S} + \frac{1}{2} \right\rfloor \right) \quad (6)$$

Here  $j = 0, 1, \dots, N-1$ ,  $m$  is the interval to which differential excitation  $\xi_j$  belongs i.e.  $\xi_j \in l_m$ ,  $t$  is the index of quantized orientation, and  $\delta(\cdot)$  is defined as follows:

$$\delta(\cdot) = \begin{cases} 1 & \text{if function is true,} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Step-3: Sub-histogram segments,  $H_{m,t}$ , across all dominant orientations are reorganized into  $M$  one dimensional histograms.

Step-4:  $M$  sub-histograms are concatenated into a single histogram thus representing the final  $6 \times 8 \times 3$  ( $M \times T \times S$ ) circular WLD histogram. The range of differential excitation is segmented into separate intervals to account for the variations in a given face image, and assigning optimal weights to these  $H_m$  segments further improves the performance of CWLD descriptor.

4) *Multi-scale Circular WLD*: In Multi-scale analysis, CWLD descriptor is extracted with different values of  $P$  and  $R$  and the histograms obtained at different scales are concatenated. In this research, multi-scale analysis is performed at three different scales with parameters as ( $R = 1, P = 8$ ), ( $R = 2, P = 16$ ) and ( $R = 3, P = 24$ ). A face image is divided into  $6 \times 7$  non-overlapping local facial regions and MCWLD histogram is computed for each region. MCWLD histograms for every region are then concatenated to form the facial representation.

#### B. Memetic Optimization

According to psychological studies in face recognition [25], some facial regions are more discriminating than others and hence, contribute more towards the recognition accuracy. Similarly, MCWLD histograms corresponding to different local facial regions may have varying contribution towards the

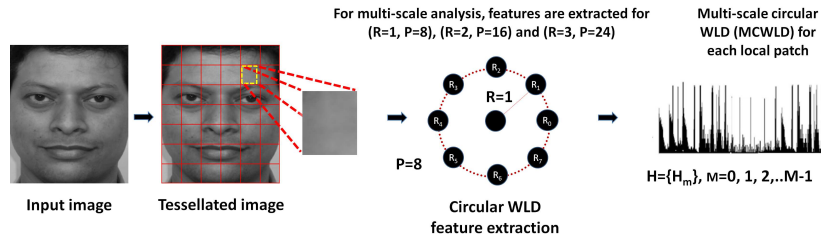


Fig. 4. Steps involved in the proposed algorithm for matching sketches with digital face images.

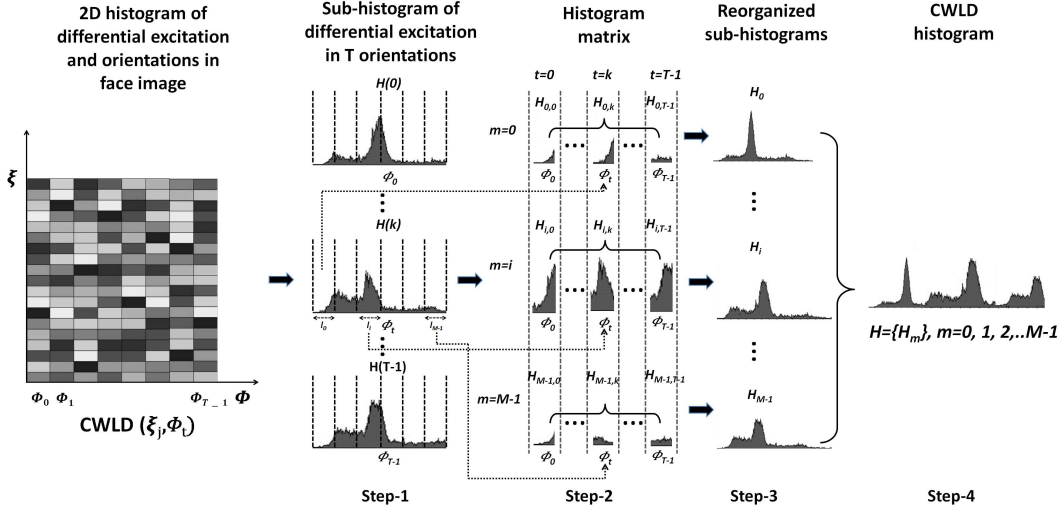


Fig. 5. Illustrating the steps involved in computing the circular WLD histogram (adapted from [18]).

recognition accuracy. Moreover, MCWLD histogram corresponding to each local facial region comprises of  $M$  sub-histogram segments (as shown in Step-3 of Fig. 5) representing different frequency information. Generally, the regions with high variance are more discriminating as compared to flat regions, therefore,  $M$  sub-histogram segments may also have varying contribution towards the recognition accuracy. It is our assertion that while matching MCWLD histograms, different weights need to be assigned to local regions and histogram segments for better performance. Here, the weights associated with 42 local facial regions and 6 sub-histogram segments at three different scales have to be optimized. Optimizing such large number of weights for best performance is a very challenging problem and requires a learning based technique.

Memetic algorithm (MA) [26] can be effectively used to optimize such large search spaces. It is a form of hybrid global-local heuristic search methodology. The global search is similar to traditional evolutionary approaches such as population-based method in a Genetic Algorithm (GA), while the local search involves refining the solutions within the population. From an optimization perspective, MAs have been found to be more efficient (i.e. require fewer evaluations to find optima) and effective (i.e. identify higher quality solutions) than traditional evolutionary approaches such as GA [27]. In this research, memetic algorithm is thus used for weight optimization.

1) *Weighted  $\chi^2$  Matching using Memetic Optimization:* For matching two MCWLD histograms, weighted  $\chi^2$  distance

measure is used.

$$\chi^2(x, y) = \sum_{i,j} \omega_j \left[ \frac{(x_{i,j} - y_{i,j})^2}{(x_{i,j} + y_{i,j})} \right] \quad (8)$$

where  $x$  and  $y$  are the two MCWLD histograms to be matched,  $i$  and  $j$  correspond to the  $i^{th}$  bin of the  $j^{th}$  histogram segment ( $j = 1, \dots, 756$ ), and  $\omega_j$  is the weight for the  $j^{th}$  histogram segment. As shown in Fig. 6, a memetic search is applied to find the optimal values of  $w_j$ . The steps involved in the memetic optimization process are described below:

*Memetic Encoding:* A chromosome is a string whose length is equal to the number of weights to be optimized i.e.  $42 \times 6 \times 3 = 756$ . Each unit or meme in a chromosome is a real valued number representing the corresponding weight.

*Initial Population:* MA is initialized with 100 chromosomes. For quick convergence, weights proportional to the rank-1 identification accuracy of each individual region are used as the initial chromosome [22]. The remaining 99 chromosomes are generated by randomly changing one or more units in the initial chromosome. Further, the weights are normalized such that the sum of all the weights in a chromosome is one.

*Fitness Function:* Each chromosome in a generation is a possible solution and the recognition is performed using the weights encoded by the chromosomes. The identification accuracy, used as fitness function, is computed on the training set and the 10 best performing chromosomes are selected as

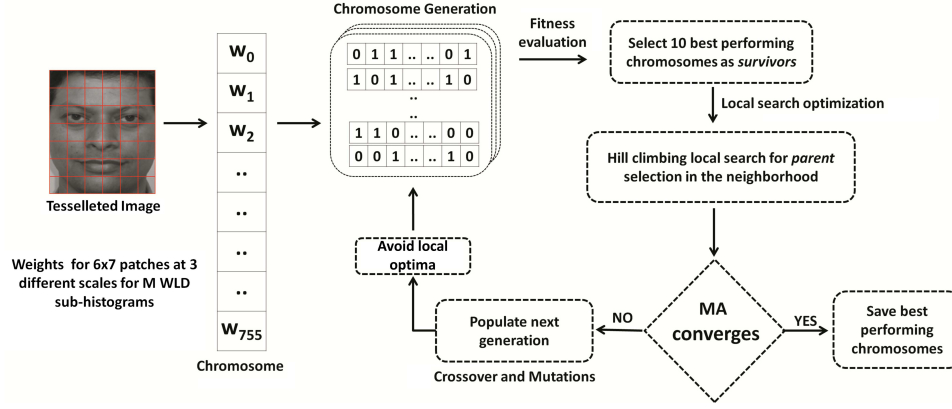


Fig. 6. Illustrating the steps involved in memetic optimization for assigning optimal weights to each tessellated face region.

*survivors*. These survivors are used for crossover and mutation to populate the next generation.

*Hill Climbing Local Search*: MA requires a local search on *survivors* to further refine the solution [27]. Two *survivors* are recombined to produce two candidate *parents*. Note that in a pair of two, this process is repeated for all 10 *survivors* to find better chromosomes. If the candidate *parents* have better performance than participating *survivors*, they replace the *survivors* to become *parents* and populate the next generation. This local search is performed at each generation to find better *parents* from the competing *survivors* which leads to quick convergence and better quality of solution.

*Crossover and Mutation*: A set of uniform crossover operations is performed on *parents* (obtained after local search) to populate a new generation of chromosomes. After crossover, mutation is performed by changing one or more weights by a factor of its standard deviation in previous generations. After mutation and crossover, 100 chromosomes are populated in the new generation.

The MA search process is repeated till convergence and terminates when the identification performance of chromosomes in the new generation does not improve compared to the performance of chromosomes in previous five generations. At this point, weights pertaining to the best performing chromosome (i.e. chromosome giving best recognition accuracy on training data) are obtained and used for testing. Thus, for a given data set, the MA search process finds optimal weights. It also enables to discard redundant and non-discriminating regions whose contribution towards recognition accuracy is very low (i.e. the weight for that region is zero or close to zero). This leads to dimensionality reduction and better computational efficiency because MCWLD histograms for poor performing facial regions are not computed during testing.

2) *Avoiding Local Optima*: Evolutionary algorithms such as MA often fail to maintain diversity among individual solutions (chromosomes) and cause the population to converge prematurely. This leads to decrease in the quality of solution. Different techniques have been proposed to maintain certain degree of diversity in a population, without affecting the convergence. In this research, *adaptive mutation rate* [28] and *random offspring generation* [29] are used to prevent

premature convergence at local optima.

- *Adaptive Mutation rate*: Mutation rate can be increased to maintain the diversity in population. However, higher value of mutation rate may introduce noise and affect the convergence process. Instead of using a fixed high or low mutation rate, an adaptive mutation rate, depending on population's diversity, is used. Population diversity is measured as the standard deviation of fitness values in a population as shown in Eq. 9:

$$stddev(P) = \sqrt{\frac{\sum_{i=1}^N (f_i - f_{mean})^2}{(N - 1)}} \quad (9)$$

where  $N$  is the population size and  $f_i$  is the fitness of the  $i^{th}$  chromosome in the population. The process starts with an initial value of mutation rate (probability 0.02), and whenever population diversity falls below a predefined threshold, mutation rate is increased.

- *Random Offspring Generation*: One of the reasons for evolutionary algorithms converging to local optima is high degree of similarity among participating chromosomes (*parents*) during crossover operation. Combination of such chromosomes is ineffective because it leads to offsprings that are exactly similar to the *parents*. If such a situation occurs where participating chromosomes (*parents*) are very similar, then crossover is not performed and offsprings are generated randomly.

The memetic algorithm for weight optimization is summarized in Algorithm 1.

### C. Proposed Algorithm for Matching Sketches with Digital Face Images

The process of matching sketches with digital face images is as follows:

- 1) For a given sketch-digital image pair, the pre-processing technique is used to enhance the quality of face images.
- 2) Both sketches and digital face images are tessellated into non-overlapping local facial regions.
- 3) For each facial region, MCWLD histograms are computed at three different scales. The facial representation is obtained by concatenating MCWLD histograms for every facial region.



**Algorithm 1** Memetic algorithm for weight optimization.

Step 1: Memetic Encoding: A chromosome of length  $42 \times 3 \times 6 = 756$  is encoded where each unit in the chromosome is a real valued number representing the corresponding weight.

Step 2: Initial Population: A population of 100 chromosomes is generated starting with a seed chromosome.

Step 3: Fitness Function: Fitness is evaluated by performing recognition using the weights encoded by each chromosome. 10 best performing chromosomes from a population are selected as *survivors* to perform crossover and mutation.

Step 4: Hill Climbing Local Search: The *survivors* obtained in Step 3 are used to find better chromosomes in their local neighborhood and *parents* are selected.

Step 5: Crossover and Mutation: New population is generated from *parents* obtained after local search in Step 4. A set of uniform crossover operations is performed followed by mutation. To avoid local optima, adaptive mutation and random offspring generation techniques are used.

Step 6: Repeat Steps 3-5 till convergence criteria is satisfied.

- 4) To match two MCWLD histograms, weighted  $\chi^2$  distance measure is used where the weights are optimized using Memetic algorithm.
- 5) In identification mode, this procedure is applied for each gallery-probe pair and top matches are obtained.

## IV. SKETCH DATABASES

To evaluate the performance of the proposed algorithm, three types of sketch databases are used: 1) Viewed Sketch, 2) Semi-forensic Sketch, and 3) Forensic Sketch database.

- 1) *Viewed Sketch Database*: It comprises a total of 549 sketch-digital image pairs from two sketch databases: the CUHK database [6] and the IIIT-Delhi Sketch database [13]. The CUHK database comprises 606 sketch-digital image pairs from CUHK students [6], the AR [30], and the XM2VTS databases. Since the XM2VTS database is not available freely, the remaining 311 sketch-digital image pairs are used in this research. Further, the authors have prepared a database of 238 sketch-digital image pairs. The sketches are drawn by a professional sketch artist for digital images collected from different sources. This database is termed as the IIIT-Delhi Viewed Sketch database.
- 2) *Semi-forensic Sketch Database*: As described earlier, sketches drawn based on the memory of sketch artist rather than the description of an eye-witness are termed as semi-forensic sketches. To prepare the IIIT-Delhi Semi-forensic Sketch database, the sketch artist is allowed to view the digital image once (for about 5 – 10 minutes) and is asked to draw the sketch based on his memory. The time elapsed between the artist viewing an image and starting to draw a sketch is about 15 minutes. Sketch artist is not allowed to view the digital image while preparing the sketch. These sketches are thus drawn based on the recollection of the sketch artist,

thus eliminating the effect of attrition based on how well the eyewitness remembers an individual's face and how well he/she is able to describe it to the sketch artist. 140 digital images from the IIIT-Delhi Viewed Sketch database are used to prepare the Semi-forensic Sketch database. Therefore, all images that are used to draw a semi-forensic sketch also have a corresponding viewed sketch. Fig. 7 presents samples of viewed and semi-forensic sketches corresponding to digital face images.

- 3) *Forensic Sketch Database*: Forensic sketches are drawn by a sketch artist from the description of an eyewitness based on his/her recollection of the crime scene. These sketches are based on (1) how well the eyewitness can recollect and describe the face and (2) the expertise of the sketch artist. In this research, a database of 190 forensic sketches with corresponding digital face images is used. This database contains 92 forensic sketch-digital image pairs obtained from Lois Gibson [31], 37 pairs obtained from Karen Taylor (published in [32]), and 61 pairs from different source on the internet. Fig. 8 shows sample images from the forensic sketch database.

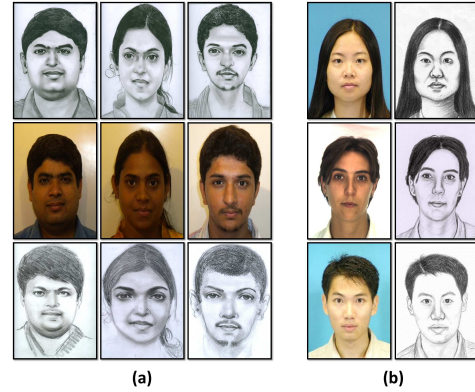


Fig. 7. (a) Sample images from the IIIT-Delhi Sketch database. The first row represents the viewed sketches, the second row represents the corresponding digital face images and the third row represents the corresponding semi-forensic sketches. (b) Sample images from the CUHK database.

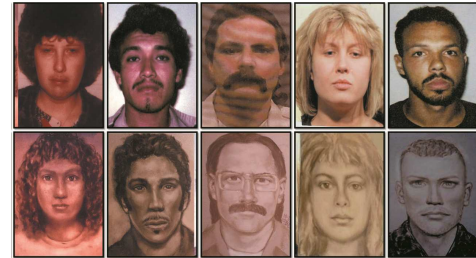


Fig. 8. Sample images from the Forensic Sketch database. Images are obtained from various sources such as [31], [32].

## V. VIEWED SKETCH MATCHING RESULTS

To establish a baseline, the performance of the proposed and existing algorithms are first computed on the viewed sketch database. Since the application of sketch recognition is dominant with identification scenario, the performance of the

TABLE I  
EXPERIMENTAL PROTOCOL FOR MATCHING VIEWED SKETCHES.

Experiment	Number of Sketch-Digital Image Pairs	Training Database	Testing Database
Experiment 1	311 from CUHK	125	186
Experiment 2	238 from IIIT-Delhi	95	143
Experiment 3	549 from Combined	220	329

TABLE II  
RANK-1 IDENTIFICATION ACCURACY OF SKETCH TO DIGITAL FACE IMAGE MATCHING ALGORITHMS FOR MATCHING VIEWED SKETCHES.  
IDENTIFICATION ACCURACIES ARE COMPUTED WITH FIVE TIMES RANDOM CROSS VALIDATION AND STANDARD DEVIATIONS ARE ALSO REPORTED.

Database (Training/Testing)	Algorithm	Rank-1 Identification Accuracy (%)	Standard Deviation (%)
CUHK (125/186)	COTS-1	91.25	0.83
	COTS-2	92.05	0.72
	WLD [18]	93.42	0.85
	MWLD [18]	94.14	0.82
	MCWLD	95.08	0.76
	SIFT [12]	94.36	1.03
	EUCLBP+GA [13]	95.12	0.93
	LFDA [14]	97.10	1.16
	<b>Proposed</b>	<b>97.28</b>	<b>0.68</b>
IIIT-Delhi Viewed Sketch (95/143)	COTS-1	71.46	0.87
	COTS-2	73.26	0.75
	WLD [18]	74.34	0.81
	MWLD [18]	75.68	0.83
	MCWLD	78.48	0.89
	SIFT [12]	76.28	1.33
	EUCLBP+GA [13]	79.36	0.87
	LFDA [14]	81.43	1.11
	<b>Proposed</b>	<b>84.24</b>	<b>0.94</b>
Combined (220/329)	COTS-1	80.14	0.78
	COTS-2	79.24	0.86
	WLD [18]	84.37	0.88
	MWLD [18]	85.32	0.86
	MCWLD	88.25	0.84
	SIFT [12]	85.86	1.01
	EUCLBP+GA [13]	88.75	0.87
	LFDA [14]	91.16	0.93
	<b>Proposed</b>	<b>93.16</b>	<b>0.96</b>

proposed algorithm is evaluated in identification mode. Three sets of experiments are performed using the viewed sketch databases. In all three experiments, digital images are used as gallery and sketches are used as probe. Further, 40% of the database is used for training and the remaining 60% pairs are used for performance evaluation. The protocol for all three experiments is described in Table I.

For each experiment, training is performed to compute the parameters of feature extractor and weights using the Memetic Optimization. This non-overlapping train-test partitioning is repeated five times with random sub-sampling and Cumulative Match Characteristic (CMC) curves are computed for performance comparison.

#### A. Experimental Analysis

The performance of the proposed approach is compared with existing algorithms designed for matching sketches with digital face images and two leading commercial face recognition systems<sup>3</sup>. Existing algorithms include SIFT [12], EU-

CLBP+GA [13], and LFDA [14]. Further, the performance gain due to multi-scale analysis and circular sampling is analyzed by comparing the performance of WLD, Multi-scale WLD (MWLD) algorithms with square sampling, and Multi-scale circular WLD (MCWLD). The same weighting scheme proposed by Chen *et al.* [18] is used in WLD, MWLD, and MCWLD algorithms. Further, to quantify the improvement due to memetic optimization of weights as compared to the weighting method proposed in [18], the performance of the proposed algorithm is compared with MCWLD. The pre-processing technique enhances the quality only when there are irregularities and noise in the input image and it does not alter good quality face images (i.e. sketch-digital image pairs from the viewed sketch database). Therefore in the experiments with *Viewed Sketch* database, no pre-processing is applied on sketch-digital image pairs. Key results and observations for matching viewed sketches are summarized below:

- The CMC curves in Fig. 9 show the rank-1 identification accuracy of sketch to digital face image matching algorithms. Table II summarizes the rank-1 identification accuracy and standard deviation for five times random subsampling (cross validations) on all three sets of experiments. On the CUHK database, the proposed approach yields rank-1 accuracy of 97.28% which is slightly better than LFDA [14] and is at least 2% better than MWLD [18], MCWLD, SIFT [12], and EUCLBP+GA [13]. The proposed approach also outperforms the two commercial systems by at least 5%.
- On comparing WLD with MWLD, it is observed that MWLD provides an improvement of about 1% on different viewed sketch databases due to multi-scale analysis. Further, compared to MWLD algorithm [18], the proposed MCWLD algorithm improves the rank-1 identification accuracy by about 1% on the CUHK, 2.8% on the IIIT-Delhi, and 2.9% on the combined databases. It suggests that circular sampling method yields more discriminative representation of the face image as compared to square sampling. Note that both MWLD and MCWLD are applied at three different scales with parameters as  $(R = 1, P = 8)$ ,  $(R = 2, P = 16)$  and  $(R = 3, p = 24)$ . The parameters for WLD are  $M = 6$ ,  $T = 8$ , and  $S = 3$ .
- Compared to the weighting scheme (proposed by Chen *et al.* [18]) used in MCWLD algorithm, the proposed memetic optimization improves the rank-1 identification accuracy by 2.2% on the CUHK, 5.7% on the IIIT-Delhi, and 4.9% on the combined databases. This improvement in rank-1 identification accuracy validates our assertion that assigning memetically optimized weights to local facial regions boosts the identification performance. This also corroborates with several psychological findings that different facial regions have varying contribution towards the recognition performance [25].
- The CUHK sketch database and the IIIT-D viewed sketch database have variations introduced by different drawing styles of artists. As discussed by Zhang *et al.* [10], drawing styles of different artists play an important role in how closely a sketch resembles the actual digital photo

<sup>3</sup>The license agreements of these commercial face recognition systems do not allow us to name the product in any comparison. Therefore, the two products are referred to as COTS-1 and COTS-2.



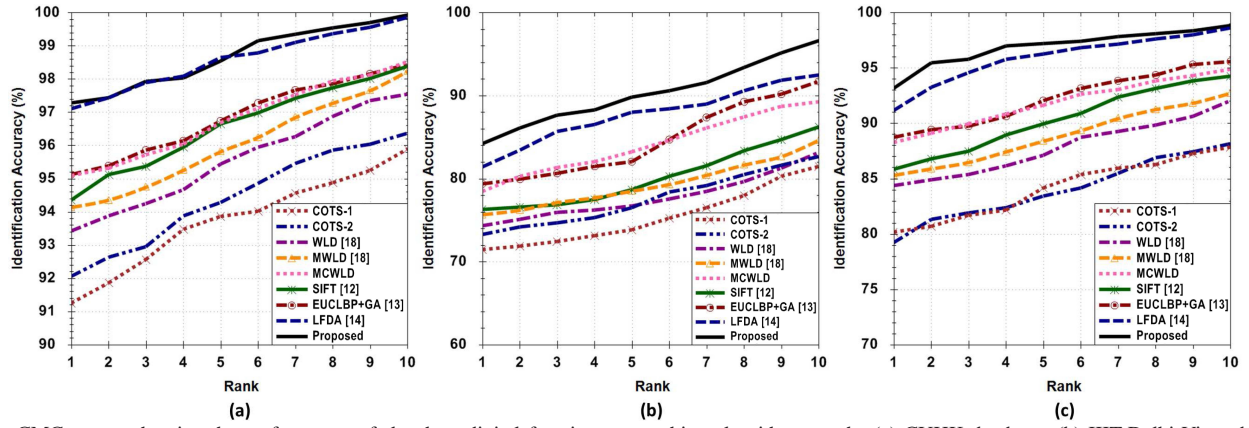


Fig. 9. CMC curves showing the performance of sketch to digital face image matching algorithms on the (a) CUHK database, (b) IIIT-Delhi Viewed Sketch database, and (c) Combined database.

thus influencing the performance of different algorithms.

- As shown in Fig. 9(c), the rank-1 identification accuracy of the proposed algorithm on the combined database is at least 2% better than existing approaches and outperforms the two commercial systems by 13%. The proposed approach represents the face image by combining MCWLD histograms obtained from every local facial region. The multi-scale analysis along with memetic optimization for assigning weights corresponding to each local facial region helps in capturing the salient micro patterns from both sketches and digital face images. Further, memetic optimization helps in dimensionality reduction; i.e. at the end of memetic optimization, on an average, 32 out of 126 ( $42 \times 3$ ) local facial patches at different scales are assigned null weights. Therefore, MCWLD histogram for these patches are not computed during testing.
- Experiments are performed by reducing the dimensionality of features using PCA; however, the results are not encouraging as it does not capture the observation that information vested in local regions have varying contribution in recognition accuracy and assigning optimal weights to these regions will enhance the performance. To incorporate this observation, MA is used that leads to dimensionality reduction and better computational efficiency because MCWLD histograms for poor performing facial regions are not computed during testing.

## VI. MATCHING FORENSIC SKETCHES WITH DIGITAL FACE IMAGES

Previous research [14] in matching forensic sketches suggests that existing sketch recognition algorithms trained on viewed sketches are not sufficient for matching forensic sketches with digital face images. Moreover, poor quality of forensic sketches further degrade the performance of sketch to digital image matching algorithms. This research attempts to analyze and evaluate the performance of semi-forensic sketches and use it for improving the training of algorithms for forensic sketch matching.

### A. Matching Semi-Forensic Sketches

Viewed sketches and forensic sketches are very different from each other. As shown in Fig. 2, the level of difficulty

increases from viewed to forensic sketch matching. In an attempt to bridge the gap between viewed and forensic sketches, semi-forensic sketches are introduced in this research. It is our assertion that training on semi-forensic sketches can improve modeling the variations for matching forensic sketches as compared to training on viewed sketches. Therefore, to better understand the progression from viewed to semi-forensic sketches, experiments are performed where training is done on viewed sketches and performance is evaluated on semi-forensic sketches.

To evaluate the performance on semi-forensic sketches, the algorithms are trained on the Viewed Sketch database. 95 sketch-digital image pairs from the IIIT-Delhi Viewed Sketch database are used for training and testing is performed with the remaining 454 digital face images as gallery and 140 semi-forensic sketches as probe. Fig. 10(a) shows the rank-1 identification accuracy of sketch to digital face image matching algorithms on semi-forensic sketches. The proposed approach that uses MCWLD and memetically optimized weighted  $\chi^2$  distance yields rank-1 identification accuracy of 63.24% and outperforms existing algorithms such as SIFT [12], EUCLBP+GA [13], and LFDA [14] by 2 – 5%. The proposed approach also outperforms the two commercial face recognition systems by at least 9%.

### B. Matching Forensic Sketches

Since forensic sketches are based on the recollection of an eyewitness, they are often inaccurate, incomplete, do not closely resemble the actual digital face image, and may be of poor quality. These concerns make the problem of matching forensic sketches with digital face images more challenging than matching viewed sketches. This section presents the evaluation of algorithms on the Forensic Sketch database.

1) *Experimental Protocol*: To evaluate the proposed approach for matching forensic sketches, four sets of experiments are performed. The performance of the proposed algorithm is also compared with existing algorithms and two commercial face recognition systems. The protocol for all the experiments are listed below:

- 1) *Training on IIIT-Delhi Viewed Sketch database*: Training is performed on 140 sketch-digital image pairs from

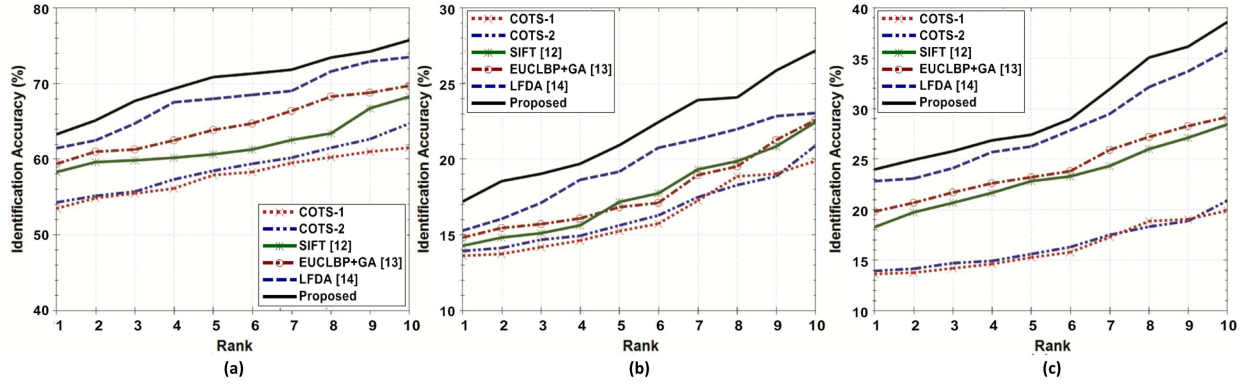


Fig. 10. CMC curves showing the identification performance when algorithms are (a) trained on viewed sketches and matching is performed on semi-forensic sketches, (b) trained on viewed sketches and matching is performed on forensic sketches, and (c) trained on semi-forensic sketches and matching is performed on forensic sketches.

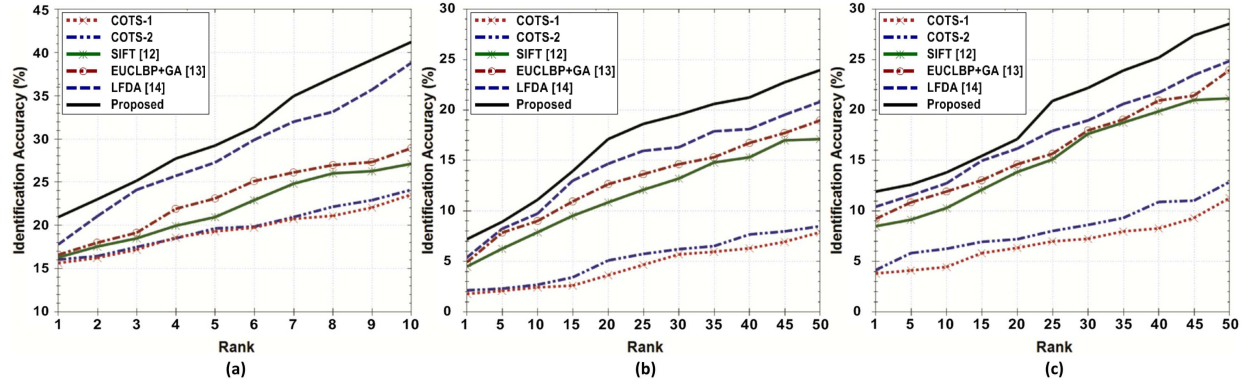


Fig. 11. CMC curves showing the identification performance when algorithms are (a) trained on viewed sketch-digital image pairs and testing is performed using pre-processed (enhanced) forensic sketch-digital image pairs, (b) trained on viewed sketch-digital image pairs and tested with large scale digital gallery and forensic sketch probes, and (c) trained on semi-forensic sketch-digital image pairs and tested with large scale digital (enhanced) gallery and pre-processed forensic sketch probes.

the IIIT-Delhi Viewed Sketch database. For testing, 190 forensic sketches are used as probe. The gallery comprises of 599 digital face images (remaining 409 digital face images from the IIIT-Delhi Viewed Sketch database and 190 digital face images from the Forensic Sketch database).

- 2) *Training on IIIT-Delhi Semi-forensic Sketch database:* Training is performed on 140 sketch-digital image pairs from the IIIT-Delhi Semi-forensic Sketch database. For testing, 190 forensic sketches are used as probe and 599 digital face images as gallery.
- 3) *Enhancing Quality of Forensic Sketches:* In this experiment, the quality of Forensic Sketch database is enhanced using the pre-processing technique described in Section II. Training is performed on 140 sketch-digital image pairs from the IIIT-Delhi Viewed Sketch database. 190 forensic sketches are used as probe and 599 digital face images are used as gallery.
- 4) *Large Scale Forensic Matching:* To replicate the real world scenario of matching forensic sketches to police mugshot database with large gallery size, 6324 digital face (frontal) images obtained from government agencies are appended to the gallery of 739 digital face images used in previous experiments. To evaluate the effect of training on semi-forensic sketches and

quality enhancement using the proposed pre-processing algorithm, two experiments are performed in large scale evaluation.

- Training is performed on 140 sketch-digital image pairs from the IIIT-Delhi Viewed Sketch database and no pre-processing is applied on the forensic sketches.
- Training is performed on 140 sketch-digital image pairs from the IIIT-Delhi Semi-forensic Sketch database and the forensic sketches are enhanced using the pre-processing technique.

2) *Experimental Analysis:* Figs. 10-11 and Tables III-IV illustrate the results of these experiments and the analysis is provided below.

- Table III and Fig. 10(b) show identification performance of the proposed and existing algorithms for matching forensic sketches when the algorithms are trained on the IIIT-Delhi Viewed Sketch database (Experiment 1). The proposed algorithm yields 17.19% rank-1 identification accuracy which is about 2% better than existing algorithms. The proposed approach also outperforms the two commercial face recognition systems by at least 3%.
- In Experiment 2, the training is performed on semi-forensic sketches for the same 140 subjects that are used

TABLE III

RANK-1 IDENTIFICATION ACCURACY OF SKETCH TO DIGITAL FACE IMAGE MATCHING ALGORITHMS FOR MATCHING FORENSIC SKETCHES.

Experiment	Gallery/ Probe Images	Algorithm	Rank-1 Identification Accuracy (%)
Experiment 1 Fig. 10(b)	599/190	COTS-1	13.62
		COTS-2	13.92
		SIFT [12]	14.26
		EUCLBP+GA [13]	14.81
		LFDA [14]	15.26
		<b>Proposed</b>	<b>17.19</b>
Experiment 2 Fig. 10(c)	599/190	COTS-1	13.62
		COTS-2	13.92
		SIFT [12]	18.26
		EUCLBP+GA [13]	19.81
		LFDA [14]	22.78
		<b>Proposed</b>	<b>23.94</b>
Experiment 3 Fig. 11(a)	599/190	COTS-1	15.62
		COTS-2	16.01
		SIFT [12]	16.26
		EUCLBP+GA [13]	16.54
		LFDA [14]	17.78
		<b>Proposed</b>	<b>20.94</b>

TABLE IV

RANK-50 IDENTIFICATION ACCURACY FOR LARGE SCALE FORENSIC SKETCH MATCHING AS SHOWN IN FIGS. 11(B)& (C).

Experiment 4	Gallery /Probe Database	Algorithm	Rank-50 Identification Accuracy(%)
Training on Viewed Sketch database without pre-processing applied on forensic sketches	6923/190	COTS-1	7.88
		COTS-2	8.46
		SIFT [12]	17.11
		EUCLBP+GA [13]	18.93
		LFDA [14]	20.81
		<b>Proposed</b>	<b>23.94</b>
Training on Semi-forensic database with proposed pre- processing on forensic sketches	6923/190	COTS-1	11.28
		COTS-2	12.86
		SIFT [12]	21.24
		EUCLBP+GA [13]	23.75
		LFDA [14]	24.62
		<b>Proposed</b>	<b>28.52</b>

for training in Experiment 1. The results in Fig. 10(c) show that there is an improvement of about 7% in rank-1 identification accuracy of the proposed algorithm and at least 4% for existing algorithms when the algorithms are trained using the semi-forensic sketches. This improvement in accuracy validates our assertion that training sketch recognition algorithms on viewed sketches is not sufficient for matching forensic sketches. The proposed algorithm performs better than LFDA based algorithm [14] because the proposed approach can be efficiently trained even with less number of sketch-digital image pairs whereas LFDA requires large number of training samples to compute the discriminant projection matrices.

- The forensic sketch database contains sketches and digital face images of poor quality. The pre-processing technique enhances the quality of forensic sketches by reducing noise and irregularities from the images. The CMC curves in Fig. 11(a) show the results for Experiment 3 where enhancing the quality of forensic sketches leads to an improvement of 2 – 3% in rank-1 identification accuracy for all algorithms (compared to CMCs in Fig. 10(b)).
- Experiment 4 demonstrates the scenario where a forensic sketch is matched against a large mugshot database. The

CMC curves in Fig. 11(b) show the results for large scale forensic sketch matching when algorithms are trained using viewed sketches without any pre-processing. In this case, rank-50 identification accuracy of the proposed algorithm is 23.9% which is at least 3% better than existing algorithms.

- Comparing the CMC curves in Fig. 11(c) show that the pre-processing technique along with training on semi-forensic sketches improves the identification accuracy of the proposed approach significantly (at least 4.72% improvement in rank-1 accuracy). Enhancing the quality of forensic sketch-digital image pairs improves the rank-1 identification accuracy of the two commercial face recognition systems also by at least 2%.
- The CMC curves in Figs. 11(b) and (c) suggest that existing algorithms for matching sketches to digital face images are still not able to achieve acceptable identification accuracy for large scale applications. However, the proposed algorithm still performs better than existing algorithms and commercial face recognition systems. As shown in Table IV, the proposed algorithm achieves rank-50 accuracy of 28.52% which is at least 4% better than existing algorithms and 15% better than the two commercial face recognition systems.
- It is to be noted that the performance of automated algorithms on semi-forensic sketches is better than the performance on forensic sketches. This improvement is attributed to the fact that semi-forensic sketches act like a bridge between viewed and forensic sketches. Therefore, training sketch recognition algorithms on semi-forensic sketches consistently improves the performance of all existing algorithms.
- At 95% confidence, non-parametric rank-ordered test (using the ranks obtained from the algorithms) and parametric t-test (using the match scores) suggest that the two top performing algorithms (i.e. the proposed and LFDA) are significantly (statistically) different.
- Finally, on a 2 GHz Intel Duo Core processor with 4 GB RAM under C# programming environment, the proposed algorithm requires 0.096 seconds to compute the MCWLD descriptor of a given probe sketch.

The proposed approach emphasizes on the discriminating information vested in local regions. To capture our assertion that every local region has varying contribution, memetic algorithm assigns optimal weights to each local facial region. Assigning discriminating weights to different facial regions also supports the conclusion made by Klare *et al.* [14] that different internal, external, and individual facial regions such as eyes, nose, mouth, and chin have significant contribution in sketch recognition. Next, Fig. 12(a) shows examples of sketch-digital image pairs that are correctly identified by the proposed approach as well as the LFDA [14] based approach (correctly identified in rank-50). Sketches that show high recognizability have some unique features such as beard, mustache, and soft marks on the face. Fig. 12(b) shows sample cases where LFDA based approach performed poorly while the proposed approach correctly identified the sketch. This is mainly because the



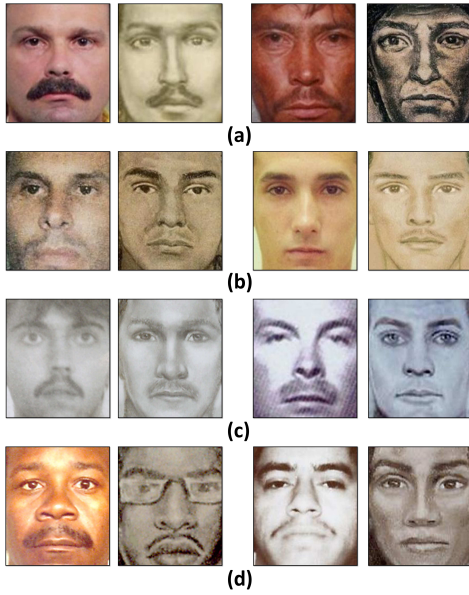


Fig. 12. Illustrating sample cases when (a) the proposed approach and LFDA [14] correctly recognize, (b) LFDA fails while the proposed algorithm correctly recognizes, (c) the proposed algorithm fails while LFDA correctly recognizes, and (d) both the algorithms fail to recognize.

proposed approach focuses on the structural details along with discriminating and prominent features of the face. Fig. 12(c) shows some examples of sketch-digital image pairs where the proposed approach performed poorly, whereas, LFDA based approach correctly matched sketches with digital face images. Finally, Fig. 12(d) shows some examples where both the proposed approach and LFDA based approach failed to match sketches with the correct digital face images. These sketches either do not resemble the actual digital face image or converge to an average face that resembles more than one digital face image in gallery because of the common features.

## VII. HUMAN ANALYSIS FOR MATCHING SKETCHES WITH DIGITAL FACE IMAGES

Several studies have analyzed human capabilities to recognize faces with variations due to illumination and expression [33]. Recently, Zhang *et al.* [16] performed an extensive study to analyze the human performance in matching sketches obtained from multiple artists. This section presents a study to understand the cognitive process of matching sketches with digital face images by humans on viewed, semi-forensic and forensic sketch databases. This examination of human responses also considers local region used by each subject while matching sketches with digital face images.

### A. Experimental Method

Since the validity of a psychological experiment is closely related to fatigue and interest level of the subject [16], human analysis is performed on a subset of 140 viewed, 140 semi-forensic and 190 forensic sketches.

1) *Participants*: A total of 82 subjects, largely undergraduate university students, volunteered to participate in the sketch to digital face image matching study. Some of the volunteers

may be familiar with few subjects in the IIIT-Delhi Viewed and Semi-forensic Sketch database but not with any of the sketches in the Forensic Sketch database.

2) *Questions*: In every question, a probe sketch must be matched to one of the 12 digital face images in the gallery. Since this is a web based application, we came up with 12 digital face images as gallery so as to properly layout the query sketch and digital face images on a computer screen. The gallery necessarily includes the correct matching digital face image and the remaining images in the gallery are the top retrieved digital face images for the probe sketch obtained using the proposed MCWLD algorithm. In the interest of fairness, un-cropped images that may include hair, ears, and neck are used for human evaluation. Automatic algorithms on the other hand, do not require this additional information.

3) *Procedure*: Each volunteer interacts with a web interface, where he/she is first authenticated. It is done to ensure that the user gets different questions in every session. Subsequently, the volunteer is presented with the questions, one at a time. Each question is selected randomly from a unique unanswered question bank comprising a mixture of viewed, semi-forensic and forensic sketches. Further, the user selects one of the gallery image as a suitable match for the query sketch. Along with this selection, the user marks the local region in the digital face image that he/she finds to be the most beneficial in recognizing the query sketch. This response is indicated by user's click on the most discriminating local facial region of the selected gallery image. A volunteer answers between 2 and 12 questions in a single session and can participate in up to four sessions.

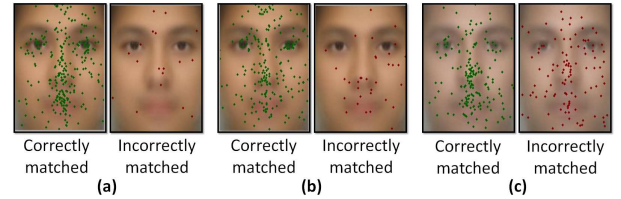


Fig. 13. Facial regions for correctly and incorrectly matched (a) viewed sketches, (b) semi-forensic sketches, and (c) forensic sketches. Dots represent the area that users found to be the most discriminating in matching the sketch with digital face images.

### B. Results and Analysis

A total of 1169 human responses are obtained for 470 probe sketches. Of these responses, 71.94% are found to be correct matches. Table V shows the total number of responses and individual accuracy of these responses across the three types of sketches. Further, Fig. 13 shows human response (clicks) that the participant deemed as important in matching the sketches with digital face images. These clicks are plotted over a mean face image to enable better visualization. The key observations from this study are listed below:

- The click-points, shown in Fig 13 indicate that the dominant local regions of a face image such as mouth, nose, and eyes (accurately depicted by the sketch artist), are used for matching.
- Fig. 13(a) shows the click-plot obtained when the user is presented viewed sketches. The high accuracy can be

TABLE V

DISTRIBUTION OF 1169 HUMAN RESPONSES OBTAINED FROM THE STUDY.

Type	Total Human Responses	% Correct
Viewed	403	80.4
Semi-forensic	334	79.6
Forensic	432	58.1

TABLE VI

DISTRIBUTION OF USER CLICKS BETWEEN PROMINENT FACIAL REGIONS.

	Viewed (%)	Semi-forensic (%)	Forensic (%)
Eyes	6.13	13.97	13.17
Nose	18.10	14.90	18.10
Mouth	10.58	10.56	14.76

attributed to the correct depiction of features by the artist. The user clicks are concentrated close to nose and mouth region.

- Fig. 13(b) shows the click-plot obtained when a user is presented semi-forensic sketches. The points seem to deviate towards the exaggerated features such as corners of eyes, nose and eyebrows.
- *Forensic Sketch* database contains poor quality sketches and *two-fold* exaggeration at witness description and artist depiction. The large differences in appearance, age, and high possibility of accessories result in user preference for nose and mouth regions, as shown in Fig. 13(c).
- As the difficulty of recognition task escalates from viewed to forensic sketches, there is a notable increase in the use of prominent facial features (eyes, nose, and mouth), as indicated in Table VI. This marked increase in user preference for local facial features when presented with unfamiliar sketches is a strong indication of their importance in the recognition task.
- This study supports our initial hypothesis that local regions provide discriminating information for matching sketches with digital face images. Finally, with 1169 sample size at 95% confidence level, confidence interval lies in 2 – 3% for the three types of sketches.

The accuracy claimed by humans for different types of sketches cannot be compared with the accuracy of automatic algorithms because of different experimental protocols. This analysis is to validate our assertion that discriminating patterns in local facial regions have major contribution in matching sketches with digital face images.

### VIII. CONCLUSION

Sketch to digital face matching is an important research challenge and is very pertinent to law enforcement agencies. This research presents a discriminative approach for matching sketch-digital image pairs using modified Weber's local descriptor and memetically optimized weighted  $\chi^2$  distance. The algorithm starts with the pre-processing technique to enhance sketches and digital images by removing irregularities and noise. Next, MCWLD encodes salient micro patterns from local regions to form facial signatures of both sketches and digital face images. Finally, the proposed (evolutionary) memetic optimization based weighted  $\chi^2$  distance is used to match two MCWLD histograms. Comprehensive analysis, including

comparison with existing algorithms and two commercial face recognition systems, is performed using the viewed, semi-forensic, and forensic sketch databases. Semi-forensic sketches are introduced to bridge the gap between viewed and forensic sketches. It is observed that sketch recognition algorithms trained on semi-forensic sketches can better model the variations for matching forensic sketches as compared to algorithms trained on viewed sketches. Analysis of results also suggest that local regions play an important role in matching sketch-digital image pairs and is effectively encoded in MCWLD and memetically optimized weighted  $\chi^2$  distance. The results also show that the proposed algorithm is significantly better than existing approaches and commercial systems. In future, we plan to extend the approach by combining generative and discriminative models at feature level. It may be difficult (computationally expensive) to completely transform sketches to digital face images and vice-versa using generative approaches; however, they can be easily transformed into an intermediate representation where discriminative approaches can be further used to classify them.

### IX. ACKNOWLEDGEMENT

The authors are thankful to Lois Gibson and Karen Taylor for providing the forensic sketch-digital image pairs. Special acknowledgement to Dr. A. Lanitis for providing the FG-Net face database, Dr. A. Martinez for the AR face database, Dr. X. Wang for the CUHK database, and Dr. G. Huang for the LFW database. The authors would also like to acknowledge the participants who contributed in the human analysis. Dr. B. Klare's help in implementation of his algorithm is also appreciated. The authors acknowledge reviewers and associate editor for constructive and useful feedback.

### REFERENCES

- [1] X. Tang and X. Wang, "Face photo recognition using sketch", in *Proceedings of International Conference on Image Processing*, 2002, vol. 1, pp. 257–260.
- [2] X. Tang and X. Wang, "Face sketch synthesis and recognition", in *Proceedings of International Conference on Computer Vision*, 2003, vol. 1, pp. 687–694.
- [3] Q. Liu, X. Tang, H. Jin, H. Lu, and Songde Ma, "A nonlinear approach for face sketch synthesis and recognition", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 1005–1010.
- [4] L. Yung-hui, M. Savvides, and V. Bhagavatula, "Illumination tolerant face recognition using a novel face from sketch synthesis approach and advanced correlation filters", in *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, 2006, vol. 2.
- [5] X. Gao, J. Zhong, J. Li, and C. Tian, "Face sketch synthesis algorithm based on E-HMM and selective ensemble", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 4, pp. 487–496, 2008.
- [6] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, pp. 1955–1967, 2009.
- [7] W. Zhang, X. Wang, and X. Tang, "Lighting and pose robust face sketch synthesis", in *Proceedings of European Conference on Computer Vision*, 2010, pp. 420–433.
- [8] R. Uhl and N. Lobo, "A framework for recognizing a facial image from a police sketch", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 1996, pp. 586–593.
- [9] P. Yuen and C. Man, "Human face image searching system using sketches", *IEEE Transactions on Systems, Man and Cybernetics - A*, vol. 37, no. 4, pp. 493–504, 2007.

- [10] Y. Zhang, C. McCullough, J. Sullins, and C. Ross, "Hand-drawn face sketch recognition by humans and a PCA-based algorithm for forensic applications", *IEEE Transactions on Systems, Man and Cybernetics - A*, vol. 40, no. 3, pp. 475–485, 2010.
- [11] H. Nizami, Adkins-Hill, P. Jeremy, Y. Zhang, J. Sullins, C. McCullough, S. Canavan, and L. Yin, "A biometric database with rotating head videos and hand-drawn face sketches", in *Proceedings of International Conference on Biometrics: Theory, applications and systems*, 2009, pp. 38–43.
- [12] B. Klare and A. Jain, "Sketch-to-photo matching: A feature-based approach", in *Proceedings of Society of Photo-Optical Instrumentation Engineers Conference Series*, 2010, vol. 7667.
- [13] H. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa, "On matching sketches with digital face images", in *Proceedings of International Conference on Biometrics: Theory Applications and Systems*, 2010.
- [14] B. Klare, L. Zhifeng, and A. Jain, "Matching forensic sketches to mug shot photos", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 639–646, 2011.
- [15] B. Klare and A. Jain, "Heterogeneous face recognition using kernel prototype similarities", Tech. Rep., Michigan State University, 2011.
- [16] Y. Zhang, S. Ellyson, A. Zone, P. Gangam, J. Sullins, C. McCullough, S. Canavan, and L. Yin, "Recognizing face sketches by a large number of human subjects: A perception-based study for facial distinctiveness", in *Proceedings of International Conference on Automatic Face and Gesture Recognition*, 2011.
- [17] W. Zhang, X. Wang, and X. Tang, "Coupled information-theoretic encoding for face photo-sketch recognition", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2011.
- [18] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao, "WLD - A robust local image descriptor", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1705–1720, 2010.
- [19] I. Daubechies, *Ten lectures on wavelets*, Society for Industrial and Applied Mathematics, 1 edition, 1992.
- [20] Z. Rahman, D. Jobson, and G. Woodell, "Multi-scale retinex for color image enhancement", in *Proceedings of International Conference on Image Processing*, 1996, vol. 3, pp. 1003–1006.
- [21] S. Chang, Bin Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression", *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.
- [22] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.
- [23] D. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [24] C. Geng and X. Jiang, "Face recognition using SIFT features", in *Proceedings of International Conference on Image Processing*, 2009, pp. 3313–3316.
- [25] P. Sinha, B. J. Balas, Y. Ostrovsky, and R. Russell, "Face recognition by humans: 19 results all computer vision researchers should know about", *Proceedings of IEEE*, vol. 94, no. 11, pp. 1948–1962, 2006.
- [26] N. Krasnogor and J. Smith, "A tutorial for competent memetic algorithms: model, taxonomy, and design issues", *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 5, pp. 474 – 488, 2005.
- [27] H. Wang, D. Wang, and S. Yang, "A memetic algorithm with adaptive hill climbing strategy for dynamic optimization problems", *Soft Computing*, vol. 13, pp. 763–780, 2009.
- [28] F. Vafaee and P. Nelson, "A genetic algorithm that incorporates an adaptive mutation based on an evolutionary model", in *Proceedings of International Conference on Machine Learning and Applications*, 2009, pp. 101–107.
- [29] M. Rocha and J. Neves, "Preventing premature convergence to local optima in genetic algorithms via random offspring generation", in *Proceedings of International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems: Multiple Approaches to Intelligent Systems*, 1999, pp. 127–136.
- [30] A. Martinez and R. Benevento, "The AR face database", *CVC Technical Report #24*, 1998.
- [31] L. Gibson, *Forensic Art Essentials*, Elsevier, 2008.
- [32] K. Taylor, *Forensic Art and Illustration*, CRC Press, 2001.
- [33] A. O'Toole, P. Phillips, F. Jiang, J. Ayyad, N. Penard, and H. Abdi, "Face recognition algorithms surpass humans matching faces over changes in illumination", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 9, pp. 1642–1646, 2007.



**Himanshu S. Bhatt** received the Bachelor in Technology degree in information technology in 2009 from the Jaypee Institute of Information Technology, Noida, India.

He is currently pursuing the Ph.D. degree from the Indraprastha Institute of Information Technology (IIIT) Delhi, India. His research interests include image processing, machine learning and their applications in biometrics. He is a recipient of IBM PhD fellowship 2011-13, best poster awards in IEEE BTAS 2010 and IJCB 2011.



**Samarth Bharadwaj** received the Bachelor in Technology degree in information technology in 2009 from the Jaypee Institute of Information Technology, Noida, India.

He is currently pursuing the Ph.D. degree from the Indraprastha Institute of Information Technology (IIIT) Delhi, India. His main area of interest are image processing, pattern classification and their application in the field of facial and ocular biometrics. He is also recipient of best poster awards in BTAS 2010 and IJCB 2011.



**Richa Singh** (S'04, M'09) received the M.S. and Ph.D. degrees in computer science in 2005 and 2008, respectively from the West Virginia University, Morgantown, USA.

She is currently an Assistant Professor at the Indraprastha Institute of Information Technology (IIIT) Delhi, India. Her research has been funded by the UIDAI and DIT, India. She is a recipient of FAST award by DST, India. Her areas of interest are biometrics, pattern recognition, and machine learning. She is also an editorial board member of Information

Fusion, Elsevier.

Dr. Singh is a member of the CDEFFS, IEEE, Computer Society and the Association for Computing Machinery. She is also a member of the Golden Key International, Phi Kappa Phi, Tau Beta Pi, Upsilon Pi Epsilon, and Eta Kappa Nu honor societies. She is the recipient of ten best paper and best poster awards in international conferences.



**Mayank Vatsa** (S'04, M'09) received the M.S. and Ph.D. degrees in computer science in 2005 and 2008, respectively from the West Virginia University, Morgantown, USA.

He is currently an Assistant Professor at the Indraprastha Institute of Information Technology (IIIT) Delhi, India. His research has been funded by the UIDAI and DIT. He is the recipient of FAST award by DST, India. His areas of interest are biometrics, image processing, computer vision, and information fusion.

Dr. Vatsa is a member of the IEEE, Computer Society and Association for Computing Machinery. He is also a member of the Golden Key International, Phi Kappa Phi, Tau Beta Pi, Sigma Xi, Upsilon Pi Epsilon, and Eta Kappa Nu honor societies. He is the recipient of ten best paper and best poster awards in international conferences. He is an area editor of IEEE Biometric Compendium.