Memetic Approach 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 the 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 algorithms and two commercial face recognition systems.**

***Index Terms*—Sketch Recognition, Face Recognition, Memetic Algorithms, Forensic Sketch.**

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. Existing state-of-the-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.

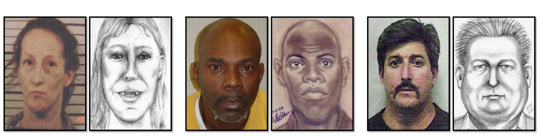

Fig. 1. Sample images showing exaggeration of facial features in forensic sketches.

H.S. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa are with the Indraprastha Institute of Information Technology (IIIT) Delhi, India, e-mail: (himanshub, samarthb, rsingh, mayank)@iiitd.ac.in.

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] matched sketches and photos using a method similar to the Eigen-transform after converting sketches to photos. Recently, Wang and Tang [5] proposed Markov Random Fields based algorithm to automatically synthesize sketches from digital face images and vice-versa.

2) Discriminative Approaches: Uhl and Lobo [6] proposed photometric standardization of sketches to compare it with digital photos. The sketches and photos were geometrically normalized and matched using Eigen analysis. Yuen and Man [7] used local and global feature measurements to match sketches and mug-shot images. Zhang *et al.* [8] 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 also discussed about the quality of sketches in terms of artist's skills, experience, exposure time, and distinctiveness of features [9]. Similarly, Nizami *et al.* [10] analyzed the effect of matching sketches drawn by different artists. Klare and Jain [11] 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.* [12] 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.* [13] extended their approach using Local Feature Discriminant Analysis (LFDA) to match forensic sketches. In their recent approach, Klare and Jain [14] 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.* [15] analyzed the psychological behavior of humans for matching sketches drawn by different sketch artists. Recently, Zhang *et al.* [16] proposed an information theoretic encoding band descriptor to capture discriminative information and random forest based matching to maximize the mutual information between the sketch and the 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 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) [17]. WLD is used for representing face 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 image. Further, memetically optimized χ^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 DWT fusion based pre-processing technique is presented to enhance 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 manually separate *good* and *bad* forensic sketches and generally focus on *good* forensic sketches only. Such classification is often based on the similarity between the sketch and the corresponding digital face image which is not available in real application. Therefore, this is not pragmatic at the time of matching a forensic sketch with a digital face database. 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 therefore, improves the performance by at least $2-3\%$.
2) Multi-scale Circular WLD and memetically optimized χ^2 based algorithm is proposed for matching sketches with digital face images. The proposed algorithm outperforms existing approaches on different sketch databases.
3) Human performance for matching sketches with digital face images is also analyzed. The information collected from the subjects corroborate with our initial observation that local regions in sketch provide discriminating information.
4) The paper also presents a part of the IIIT-Delhi database (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 χ^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, in the gallery, digital images may also be noisy and of sub-optimal quality because of the 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 described below:

- Let f be the color face image to be enhanced. Let f^r and f^y be the red and luma channels[1] respectively. These two channels are processed using the multi-scale retinex (MSR) algorithm [18], [19]. 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 [20] 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 may lead to burring of edges; therefore, a filter is applied to deblur those edges. Experiments show that Wiener filter can restore the genuine facial edges. Applying Wiener filter on $f^{rm^{'}}$ and $f^{ym^{'}}$ produces f^1 and f^2.

[1] 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.

- After computing 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 is applied on f^1 and f^2 to obtain the detail and approximation bands of these images. Let f^j_{LL}, f^j_{LH}, f^j_{HL}, and f^j_{HH} be the four bands and $j = 1, 2$. To preserve features of both the channels, coefficients from the approximation band of f^1 and f^2 are averaged.

$$f^e_{LL} = mean(f^1_{LL}, f^2_{LL}) \tag{1}$$

where f^e_{LL} is the approximation band of enhanced image. All three detailed subbands are divided into windows of size 3×3 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^e_{HL}. Similarly, enhanced subbands f^e_{LH} and f^e_{HH} are obtained. Finally, inverse DWT is applied on the four subbands to generate a high quality face image.

$$F = IDWT(f^e_{LL}, f^e_{LH}, f^e_{HL}, f^e_{HH}) \tag{2}$$

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 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 and the size of detected face region is 192×224.

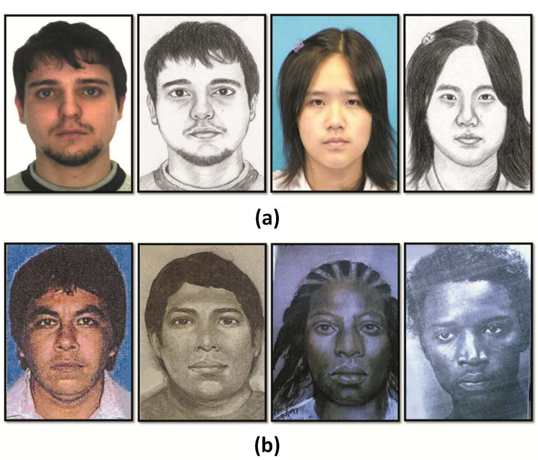

Fig. 2. Paper quality, sensor noise, and old photographs can affect the quality of sketch-digital image pairs and hence reduce the performance of matching algorithms. (a) Good quality sketch-digital image pairs and (b) poor quality sketch-digital image pairs.

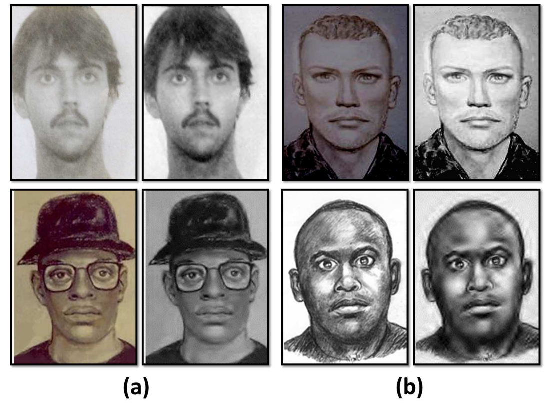

Fig. 3. Quality enhancement using the pre-processing technique. (a) represents digital face image before and after pre-processing and (b) represents forensic sketches before and after pre-processing.

III. Matching Sketches with Digital Face Images

Local descriptor based approaches have received attention in face recognition due to their robustness to scale, orientation, and speed. Local Binary Patterns (LBP) proposed by Ojala *et al.* [21] is one of the widely used descriptor for object as well as face recognition [22]. In face recognition literature, several variants of LBP have been proposed; Wolf *et al.* [23] proposed three patch and four patch LBP for recognizing faces in unconstrained settings. Zhao *et al.* [24] proposed LBP on three orthogonal planes and used it for dynamic texture recognition. Bhatt *et al.* [12] extended LBP to incorporate the 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 the texture features are computed for every pixel of the input face image. On the other hand, there are sparse descriptors such as Scale Invariant Feature Transform (SIFT) [25] that are based on interest point detection and computing the descriptor in the vicinity of detected interest points. SIFT is computed using gradient and orientation of neighboring points sampled around detected key point. As a sparse descriptor, SIFT has been used for face recognition by Bicego *et al.* [26] and Cong *et al.* [27]. Klare and Jain [11] 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.* [17] 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 extending it to multi-scale WLD. It is further optimized by computing the descriptor in a circular manner (in contrast to the originally proposed square neighborhood manner). Finally, matching of two multi-scale circular WLD (MCWLD) histograms is performed using memetically optimized weighted χ^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 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 radius R and number of neighbors P. Sketches and digital face images are represented using MCWLD as elaborated below:

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

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

where x_c is the intensity value at 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 current pixel. In contrast, if $\xi(x_c)$ is negative, it simulates the case that surroundings are darker than current pixel.

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

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

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 (ξ) and orientation (θ) 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 $j = 0, 1, ..., N-1$, $t = 0, 1, ..., T-1$, and N is the dimension of the image. Each column in the $2D$ histogram corresponds to a dominant orientation, θ_t, and each row corresponds to a differential excitation interval. Thus, intensity of each cell corresponds to the frequency of a certain differential excitation interval in a dominant orientation. Similar to Chen *et al.* [17], four step approach is followed to compute CWLD descriptor.

Step-1: The $2D$ histogram $CWLD(\xi_j, \theta_t)$ is further encoded into a $1D$ histograms. Differential excitations, ξ, are regrouped into T orientation sub-histograms, $H(t)$, where $t = 0, 1, ..., T-1$ 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 $m = 0, 1, ..., M-1$ 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 where S = 3 and is represented as:

$$H_{m,t} = h_{m,t,s} \tag{5}$$

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

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

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

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

Step-3: Sub-histogram segments $H_{m,t}$ across all dominant orientations are reorganized into M $1D$ histograms.

Step-4: Concatenating these M sub-histograms into a single histogram represents the final $6 \times 8 \times 3$ ($M \times T \times S$) circular WLD histogram. Segmenting the range of differential excitation into separate intervals accounts 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: Multi-scale analysis is performed by extracting CWLD descriptor with different values of P and R and concatenating the histograms obtained at different scales. 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 6×7 non-overlapping local facial regions and MCWLD histogram is computed for each region. MCWLD histograms for every region are then concatenated to form a global representation of the face image.

B. Memetic Optimization

According to psychological studies in face recognition [28], 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 recognition accuracy. Moreover, MCWLD histogram corresponding to each local facial region comprises of M ($M = 6$) 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 the flat regions, therefore, the M sub-histogram segments may also have varying contribution towards recognition accuracy.

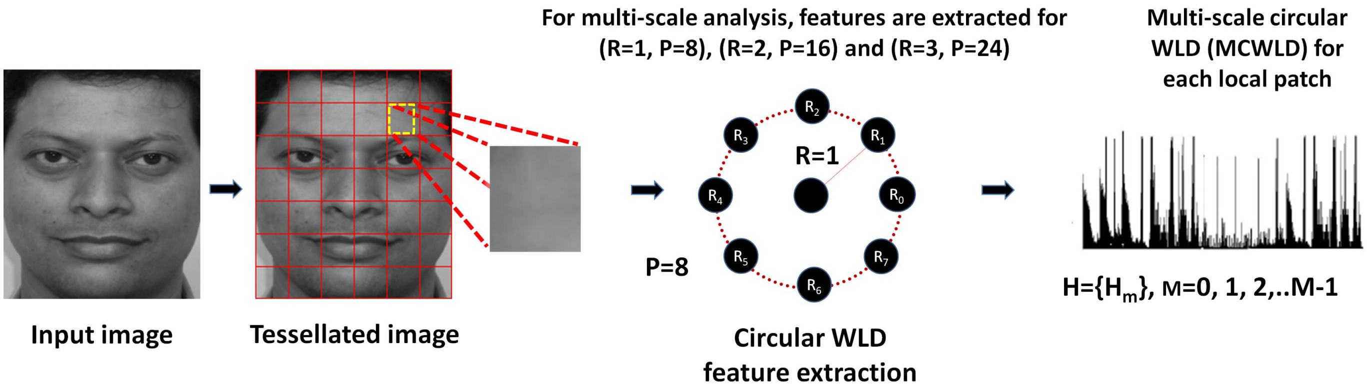

Fig. 4. Steps involved in the proposed algorithm for matching sketches with digital face images.

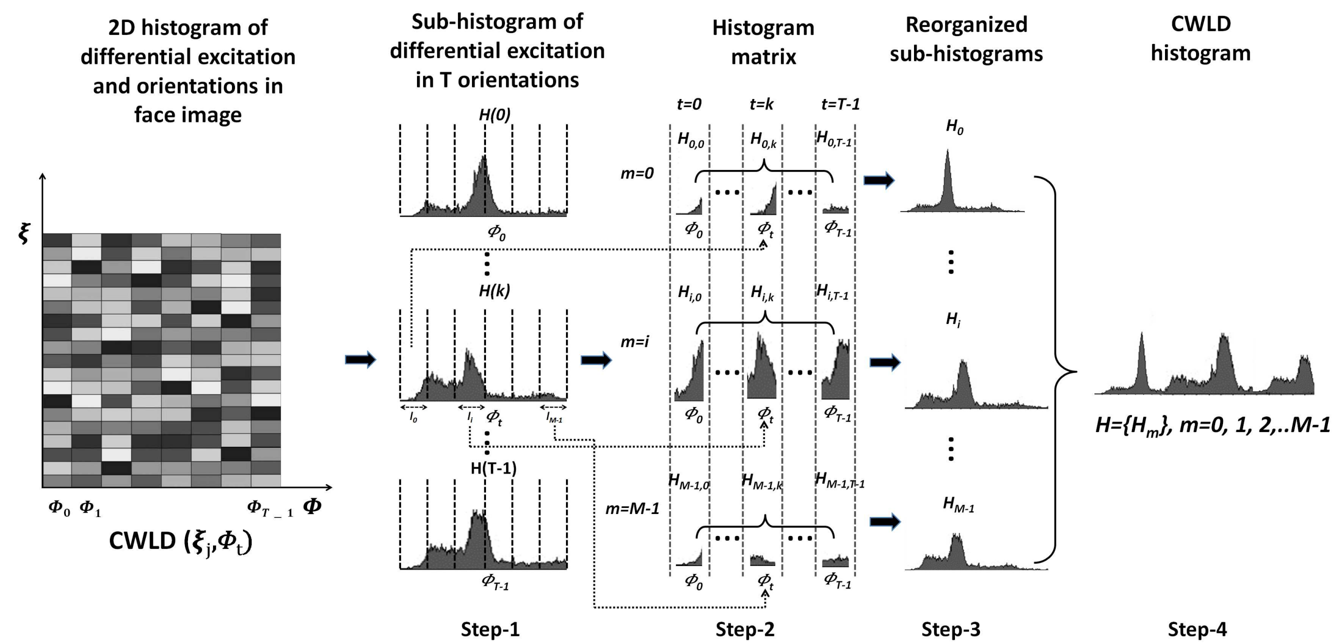

Fig. 5. Illustrating the steps involved in computing the circular WLD histogram (adapted from [17]).

It is our assertion that while matching MCLWD histograms, different weights need to be assigned to local regions, histogram segments, and scales for better performance. Here, the weights associated with 42 local facial regions and 6 sub-histograms segments at 3 different scales have to be optimized. Optimizing such large number of weights for best performance is a very challenging problem and requires machine learning based technique.

Memetic algorithm (MA) [29] can be effectively used to optimize 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 shown to be more efficient (i.e. requiring fewer evaluations to find optima) and effective (i.e. identifying higher quality solutions) than traditional evolutionary approaches such as GA [30]. In this research, memetic algorithm is used for optimizing the weights.

1) Weighted χ^2 Matching using Memetic Optimization: For matching two MCWLD histograms, weighted χ^2 distance measure is used.

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

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 ($j = 1, \cdots, 756$), and ω_j is the weight for the j^{th} histogram segment. As shown in Fig. 6, a memetic search is applied to find 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. $42 \times 6 \times 3$

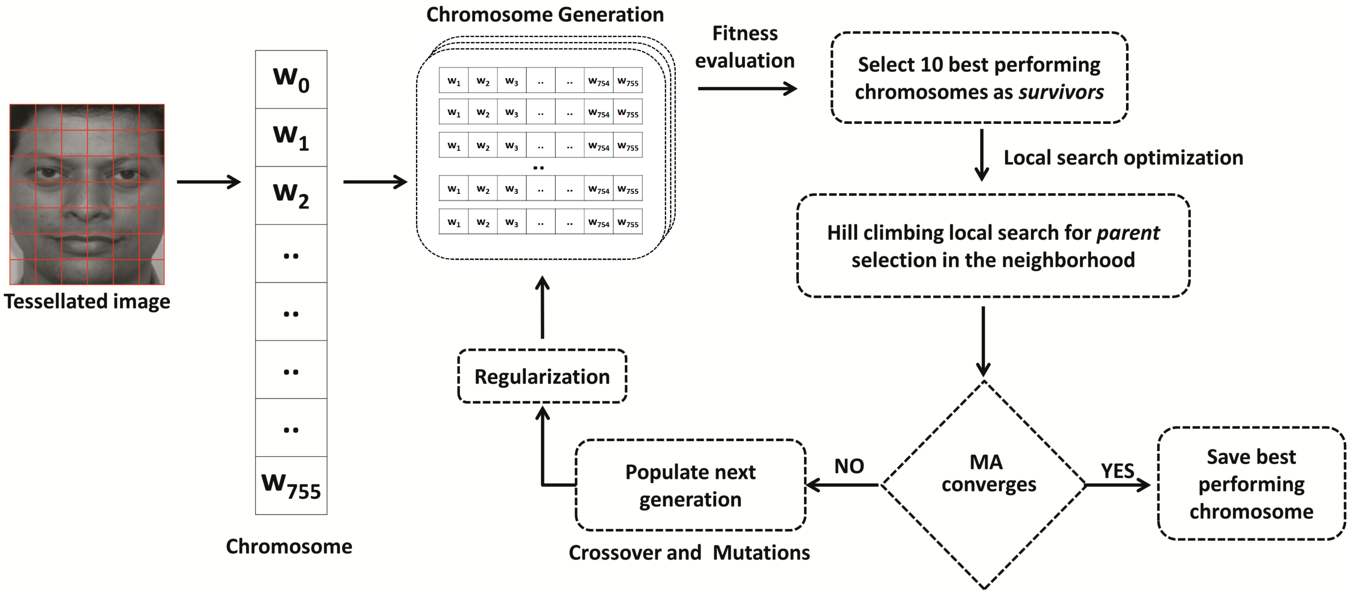

Fig. 6. Illustrating the steps involved in memetic optimization for assigning optimal weights to each tessellated face region.

= 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. In general, the first chromosome is generated randomly. However, for quick convergence in face recognition, weights proportional to the rank-1 identification accuracy of each 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 weights in a chromosome is 1.

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 *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 fine tune the solution [31]. 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*. Incorporating local search is essential for 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 the chromosomes in 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 and 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) *Regularization for 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* [32] and *random offspring generation* [33] are used to prevent premature convergence to local optima.

- *Adaptive Mutation rate:* To maintain diversity in the population, mutation rate can be increased. 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 fitness of the i^{th} chromosome in the population. The process starts with an initial value of mutation rate (probability of 0.02), and whenever population diversity falls below the 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 optimization for computing weights is summarized in Algorithm 1.

Algorithm 1 Memetic algorithm for optimizing weights.

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 chosen to populate next generation.
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 Step 3-5 till convergence criteria is satisfied.

C. Proposed Algorithm for Matching Sketches with Digital Face Images

The process for 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. A global representation is obtained by concatenating MCWLD histograms for every facial region.
4) To match two MCWLD histograms, weighted χ^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. DATABASE

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

1) *Viewed Sketch Database*: It comprises a total of 549 sketch-digital image pairs from two sketch databases: the CUHK database [5] and the IIIT-Delhi Sketch database [12]. In the CUHK database, that consists of the CUHK student database [5], the AR database [34], and the XM2VTS database, there are 606 sketch-digital image pairs. Since the XM2VTS database is not available freely, in our experiments the sketches corresponding to the XM2VTS database are removed and remaining 311 sketch-digital image pairs are used. 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 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 and is asked to draw the sketch based on his memory. 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 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[2].

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 [35], 37 forensic sketch-image pairs obtained from Karen Taylor (published in [36]), and 61 pairs from different source on the internet.

[2] The database will be made available to the research community.

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

Fig. 8. Sample images from the Forensic Sketch database.

V. Viewed Sketch Matching Results

To establish a baseline, the performance of the proposed and the 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 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 Characteristics (CMC) curves are computed for performance comparison.

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)	Commercial System-1	91.25	0.83
	Commercial System-2	92.05	0.72
	Original WLD [17]	93.40	0.85
	SIFT [11]	94.36	1.03
	EUCLBP [12]	95.12	0.93
	LFDA [13]	97.10	1.16
	Proposed	97.28	0.68
IIIT-Delhi Viewed Sketch (95/143)	Commercial System-1	71.46	0.87
	Commercial System-2	73.26	0.75
	Original WLD [17]	74.34	0.81
	SIFT [11]	76.28	1.33
	EUCLBP [12]	79.36	0.87
	LFDA [13]	81.43	1.11
	Proposed	84.24	0.94
Combined (220/329)	Commercial System-1	80.14	0.78
	Commercial System-2	79.24	0.86
	Original WLD [17]	84.37	0.88
	SIFT [11]	85.86	1.01
	EUCLBP [12]	88.75	0.87
	LFDA [13]	91.16	0.93
	Proposed	93.16	0.96

A. Experimental Analysis

The performance of the proposed approach is compared with the existing algorithms designed for matching sketches with digital face images and two leading commercial face recognition systems[3]. Existing algorithms include SIFT [11], EUCLBP [12] and LFDA [13]. To evaluate the effect of assigning memetically optimized weights to local facial regions, the performance of the proposed approach is also compared with original WLD algorithm [17]. Key results and observations for matching viewed sketches are summarized:

- 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 the 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 [13] and is at least 2% better than original WLD [17], SIFT [11], and EUCLBP [12]. The proposed approach also outperforms the two commercial systems by at least 5%.
- Compared to the original WLD algorithm [17], the proposed memetically optimized MCWLD algorithm improves the rank-1 identification accuracy by 3.88% on the CUHK database, 9.90% on the IIIT-Delhi database, and 8.79% on the combined database. This significant improvement in rank-1 identification accuracy validates our assertion that assigning optimal weights to local facial regions boosts the identification performance. This also corroborates with several physiological findings that

[3]The license agreements of these commercial face recognition systems does not allow us to name the product in any comparison. Therefore, the two products are referred to as Commercial System-1 and Commercial System-2.

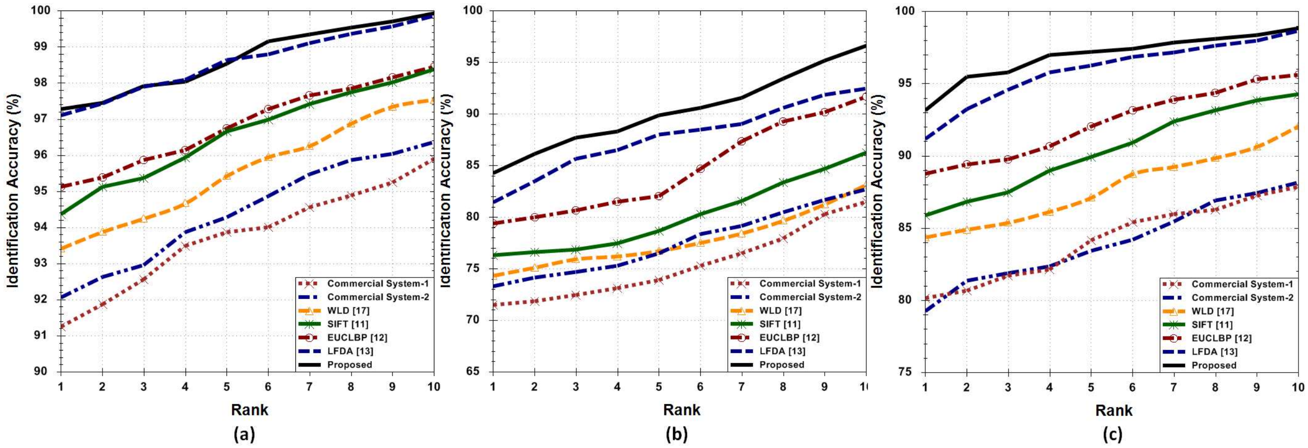

Fig. 9. CMC curves showing the performance of sketch to digital face image matching algorithms on the (a) CUHK database, (b) IIIT-Delhi Viewed Sketch database, and (c) Combined database.

different facial regions have varying contribution towards recognition performance [28].

- Unlike the CUHK sketch database, sketches and digital images in the IIIT-D Viewed Sketch database are not well registered and do not perfectly overlay. This leads to reduced performance of algorithms on the IIIT-Delhi Viewed Sketch database. Further, 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 the existing approaches and outperforms the two commercial systems by 13%.
- The proposed approach generates a holistic description of 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 the memetic optimization, on an average, 32 out of 126 (42×3) local facial patches at different scales are assigned null weights. Therefore, MCWLD histogram for these patches are not computed during testing.

VI. Matching Forensic Sketches with Digital Face Images

Previous research in matching forensic sketches suggests that the 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 degrades the performance of sketch to digital image matching algorithms. This research attempts to understand semi-forensic sketches (that are drawn based on the artist's memory). The performance of semi-forensic sketches is analyzed and then used for improving the training of the algorithms for forensic sketch matching.

A. Matching 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 remaining 454 digital face images as the gallery and 140 semi-forensic sketches as the probes. 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 χ^2 distance yields rank-1 identification accuracy of 63.24% and outperforms the existing algorithms such as SIFT [11], EUCLBP [12], and LFDA [13] 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 experimental evaluation of algorithms on the Forensic Sketch database.

1) Experimental Protocol: To evaluate the proposed approach for matching forensic sketches, four set 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 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

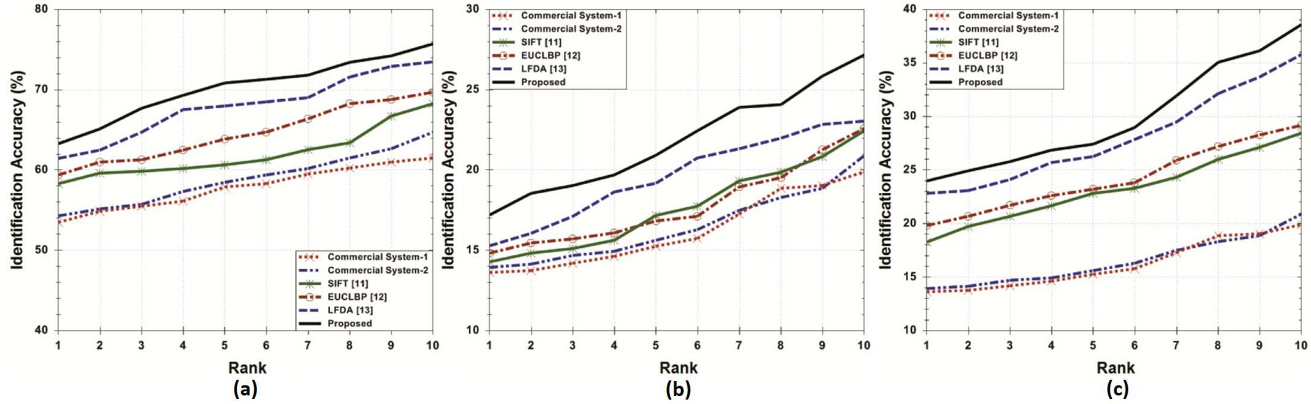

Fig. 10. CMC curves showing the identification performance when algorithms are (a) trained on viewed sketches and matching is performed on semi-forensic sketches, (b) trained on viewed sketches and matching is performed on forensic sketches, and (c) trained on semi-forensic sketches and matching is performed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

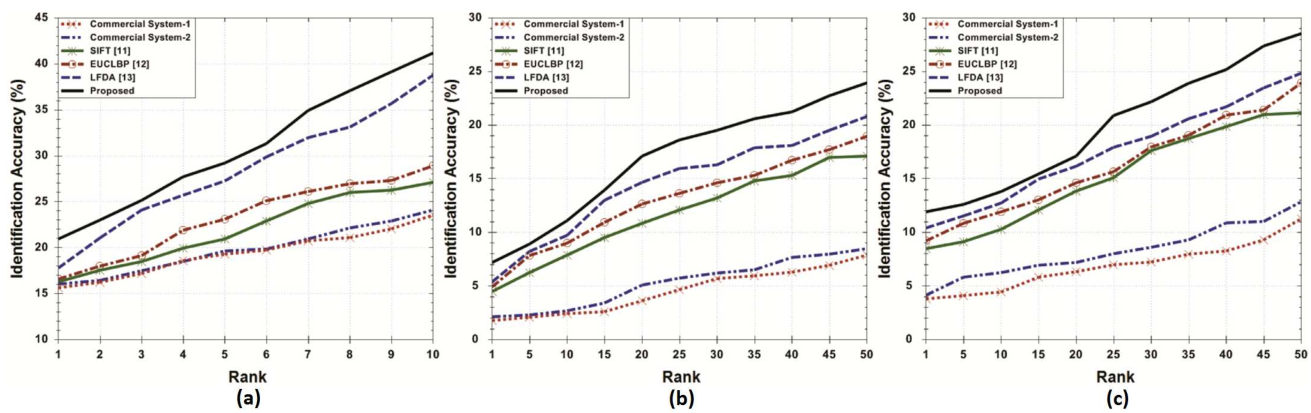

Fig. 11. CMC curves showing the identification performance when algorithms are (a) trained on viewed sketch-digital image pairs and testing is performed using pre-processed (enhanced) forensic sketch-digital image pairs, (b) trained on viewed sketch-digital image pairs and tested with large scale digital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

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, 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 other experiments. This increases the gallery size to 7063. To evaluate the effect of training on semi-forensic sketches and quality enhancement using the 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 Table III-IV illustrate the results of these experiments. The analysis of these results is provided below.

- Table III and Fig. 10(b) show identification performance of the proposed and the 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 the existing algorithms. The proposed approach also outperforms the

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	599/190	Commercial System-1	13.62%
		Commercial System-2	13.92 %
		SIFT [11]	14.26 %
		EUCLBP [12]	14.81 %
		LFDA [13]	15.26 %
		Proposed	17.19 %
Experiment 2	599/190	Commercial System-1	13.62%
		Commercial System-2	13.92 %
		SIFT [11]	18.26 %
		EUCLBP [12]	19.81 %
		LFDA [13]	22.78 %
		Proposed	23.94 %
Experiment 3	599/190	Commercial System-1	15.62%
		Commercial System-2	16.01 %
		SIFT [11]	16.26 %
		EUCLBP [12]	16.54 %
		LFDA [13]	17.78 %
		Proposed	20.94 %

TABLE IV
RANK-50 IDENTIFICATION ACCURACY FOR LARGE SCALE FORENSIC SKETCH MATCHING.

Experiment 4	Gallery /Probe Database	Algorithm	Rank-50 Identification Accuracy
Training on Viewed Sketch database without pre-processing applied on forensic sketches	7063/190	Commercial System-1	7.88%
		Commercial System-2	8.46 %
		SIFT [11]	17.11 %
		EUCLBP [12]	18.93 %
		LFDA [13]	20.81 %
		Proposed	23.94 %
Training on Semi-forensic database with proposed pre-processing on forensic sketches	7063/190	Commercial System-1	11.28%
		Commercial System-2	12.86 %
		SIFT [11]	21.24 %
		EUCLBP [12]	23.75 %
		LFDA [13]	24.62 %
		Proposed	28.52 %

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 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 the 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 achieving satisfactory results for matching forensic sketches. The proposed algorithm performs better than LFDA based algorithm [13] 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 the rank-1 identification accuracy for all algorithms, i.e. compared to Fig. **??**(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 the 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 also improves the rank-1 identification accuracy of the two commercial face recognition systems by at least 2%.
- The CMC curves in Figs. 11(b) and (c) suggest that the existing algorithms for matching sketches to digital face images are still not able to achieve acceptable identification accuracy for large scale application. 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 the 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 improved the performance for all the existing algorithms.
- At 95% confidence interval, 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, for a given probe sketch, the proposed algorithm requires 0.096 seconds to compute the MCWLD descriptor.

The proposed approach emphasizes on the discriminating information vested in the local regions. To capture our assertion that every local region has varying contribution, memetic algorithm assigns optimal weights to each local facial region. It also supports the conclusion made by Klare *et al.* [13] that different internal face, external face and individual face regions (eyes, nose, mouth, chin etc.) have significant contribution for sketch recognition. Next, Fig. 12(a) shows some examples of sketch-digital image pairs that are correctly identified by the proposed approach as well as the LFDA [13] based approach (correctly identified in rank-50). Sketches that show high

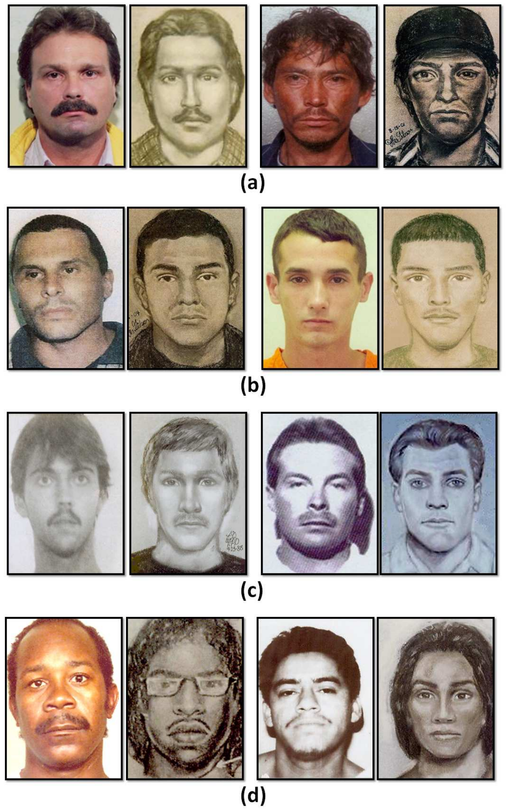

Fig. 12. Illustrating sample cases when (a) the proposed approach and LFDA [13] correctly recognize, (b) LFDA fails while the proposed algorithm correctly recognizes, (c) the proposed algorithm fails while LFDA correctly recognizes, and (d) both the algorithms fail to recognize.

recognizability have some peculiar features such as beard, mustache or soft marks on the face. Fig. 12(b) shows some examples where LFDA based approach performed poorly while the proposed approach correctly identified the sketch. This is mainly because the 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 [37]. Recently, Zhang *et al.* [15] performed an extensive study to analyze 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 [15], 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 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 include 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. The 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 the 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.

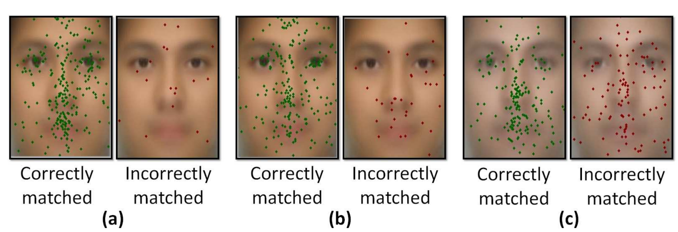

Fig. 13. Facial regions for correctly and incorrectly matched (a) viewed sketches, (b) semi-forensic sketches, and (c) forensic sketches. Dots represents the area that user found to be most discriminating in matching the sketch with digital face images.

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 %

B. Results and Analysis

A total of 1169 human responses are obtained for the 470 probe sketch images. 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 when the user is presented viewed sketches. The high accuracy can be attributed to the correct depiction of the features by the artist. The user clicks are concentrated close to nose and mouth region.
- Fig. 13(b) shows the click-plot when the 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 (i.e. witness description and artist depiction). The large differences in appearance, age and high possibility of accessaries result in user preference for nose and mouth regions, as shown in Fig. 13(c).
- As the difficulty of the recognition task escalates from viewed to forensic sketches, there is a notable increase in the use of the prominent facial features (eyes, nose and mouth) for recognition of sketches, 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 recognizing 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 Webers local descriptor and memetically optimized weighted χ^2 distance measure. 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 pertaining to both sketches and digital face images. Finally, the proposed (evolutionary) memetic optimization based weighted χ^2 distance measure 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. It is observed that local regions play an important role in matching sketch-digital image pairs and is effectively encoded in MCWLD and memetically optimized weighted χ^2 distance measure. 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.

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. B. Klare's help in implementation of his algorithm is also appreciated.

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. 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.
[6] 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.

[7] 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.
[8] Y. Zhang, C. McCullough, J. Sullins, and C. Ross, “Human and computer evaluations of face sketches with implications for forensic investigations”, in *Proceedings of the International Conference on Biometrics: Theory, Applications and Systems*, 2008, pp. 1–7.
[9] 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.
[10] 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.
[11] 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.
[12] 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.
[13] 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.
[14] B. Klare and A. Jain, “Heterogeneous face recognition using kernel prototype similarities”, Tech. Rep., Michigan State University, 2011.
[15] 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.
[16] 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.
[17] 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.
[18] 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.
[19] E. Land and J. McCann, “Lightness and retinex theory”, *Journal of Optical Society of America*, vol. 61, no. 1, pp. 1–11, 1971.
[20] 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.
[21] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
[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] L. Wolf, T. Hassner, and Y. Taigman, “Descriptor based methods in the wild”, in *Proceedings of Faces in Real-Life Images workshop at the European Conference on Computer Vision*, 2008.
[24] Z. Guoying and M. Pietikainen, “Dynamic texture recognition using local binary patterns with an application to facial expressions”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 915–928, 2007.
[25] D. Lowe, “Distinctive image features from scale-invariant keypoints”, *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
[26] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, “On the use of SIFT features for face authentication”, in *Proceedings of International Conference on Computer Vision and Pattern Recognition Workshop*, 2006, pp. 35–40.
[27] C. Geng and X. Jiang, “Face recognition using SIFT features”, in *Proceedings of International Conference on Image Processing*, 2009, pp. 3313–3316.
[28] 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.
[29] 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.
[30] J. Knowles and D. Corne, “Memetic algorithms for multiobjective optimization: Issues, methods and prospects”, in *Recent Advances in Memetic Algorithms*, vol. 166, pp. 313–352. 2005.
[31] 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.
[32] 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.
[33] 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.
[34] A. Martinez and R. Benevento, “The AR face database”, *CVC Technical Report #24*, 1998.
[35] L. Gibson, *Forensic Art Essentials*, Elsevier, 2008.
[36] K. Taylor, *Forensic Art and Illustration*, CRC Press, 2001.
[37] 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.