Recognizing Surgically Altered Face Images

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

***Abstract*—Widespread acceptability and use of biometrics for person authentication has instigated several techniques for evading identification. One such technique is altering facial appearance using surgical procedures that has raised a challenge for face recognition algorithms. Increasing popularity of plastic surgery and its effect on face recognition has attracted attention from the research community. However, the non-linear variations introduced by plastic surgery remain difficult to be modeled by existing face recognition systems. In this research, a multi-objective evolutionary granular algorithm is proposed to match face images before and after plastic surgery. The algorithm first generates non-disjoint face granules at multiple levels of granularity. The granular information is assimilated using an evolutionary genetic algorithm that simultaneously optimizes the selection of feature extractor for each face granule along with the weights of individual granules. The proposed algorithm presents significant improvements in matching surgically altered face images as compared to existing algorithms and a commercial face recognition system.**

I. Introduction

Plastic surgery procedures provide a proficient and enduring way to enhance the facial appearance by correcting feature anomalies and treating facial skin to get a younger look. Apart from cosmetic reasons, plastic surgery procedures are beneficial for patients suffering from several kinds of disorders caused due to excessive structural growth of facial features or skin tissues. Plastic surgery procedures amend the facial features and skin texture thereby providing a makeover in the appearance of face. Fig. 1 shows an example of the effect of plastic surgery on facial appearances. With reduction in cost and time required for these procedures, the popularity of plastic surgery is increasing. Even the widespread acceptability in the society encourages individuals to undergo plastic surgery for cosmetic reasons. According to the statistics provided by the American Society for Aesthetic Plastic Surgery for year 2010 [1], there is about 9% increase in the total number of cosmetic surgery procedures, with over $500,000$ surgical procedures performed on face.

Fig. 1. Illustrating the variations in facial appearance, texture, and structural geometry caused due to plastic surgery (images taken from INTERNET).

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.

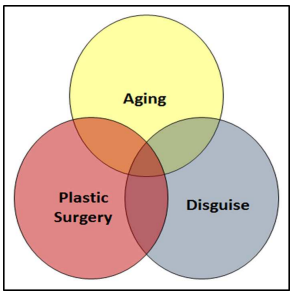

Fig. 2. Relation among plastic surgery, aging, and disguise variations with respect to face recognition.

Transmuting facial geometry and texture increases the intra-class variability between the pre- and post-surgery images of the same individual. Therefore, matching post-surgery images with pre-surgery images becomes an arduous task for automatic face recognition algorithms. Further, as shown in Fig. 2, it is our assertion that variations caused due to plastic surgery have some intersection with the variations caused due to aging and disguise. Facial aging is a biological process that leads to gradual changes in the geometry and texture of a face. Unlike aging, plastic surgery is a spontaneous process that is generally performed contrary to the effect of facial aging. Since the variations caused due to plastic surgery procedures are spontaneous, it is difficult for face recognition algorithms to model such non-uniform face transformations. On the other hand, disguise is the process of concealing one's identity by using makeup and other accessories. Both plastic surgery and disguise can be misused by individuals trying to conceal their identity and evade recognition. Variations caused due to disguise are temporary and reversible; however, variations caused due to plastic surgery are long-lasting and may not be reversible. Owing to these reasons, plastic surgery is now established as a new and challenging covariate of face recognition alongside aging and disguise. Singh *et al.* [2] analyzed several types of local and global plastic surgery procedures and their effect on different face recognition algorithms. They concluded that the non-linear variations induced by surgical procedures are difficult to address with current face recognition algorithms. De Marsico *et al.* [3] proposed an approach that integrates information derived from local regions to match pre- and post-surgery face images. Recently, Aggarwal *et al.* [4] proposed sparse representation on local facial fragments to match surgically altered face images. Though recent results suggest that the algorithms are improving towards addressing the challenge, there is a significant scope for further improvement.

This research presents an evolutionary granular computing

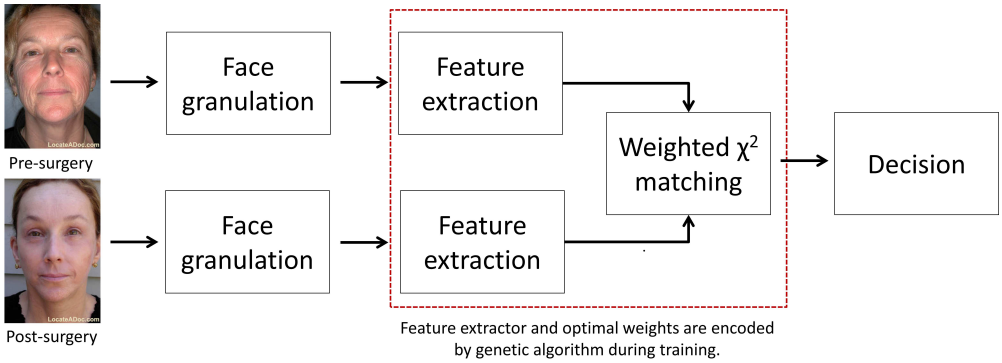

Fig. 3. Block diagram illustrating different stages in the proposed algorithm.

based algorithm for recognizing faces altered due to plastic surgery procedures. The proposed algorithm starts with generating non-disjoint face granules where each face granule represents different information at varying size and resolution. Further, two feature extractors, namely Extended Uniform Circular Local Binary Pattern (EUCLBP) [5] and Scale Invariant Feature Transform (SIFT) [6], are used for extracting discriminating information from face granules. Finally, different responses are unified in an evolutionary manner using a multi-objective genetic algorithm for improved performance. The performance of the proposed algorithm is compared with a commercial-off-the-shelf (COTS) face recognition system for matching surgically altered face images against large scale gallery.

II. Evolutionary Granular Computing Approach for Face Recognition

Singh *et al.* [2] have suggested that with large variations in the appearance, texture, and shape of different facial regions, it is difficult for face recognition algorithms to match a post-surgery face image with pre-surgery face images. Further, face recognition algorithms either use facial information in a holistic way or extract features and process them in parts. On the other hand, it is observed that humans solve problems using perception and knowledge represented at different levels of information granularity [7]. They recognize faces using a combination of holistic approach together with discrete levels of information or features. They can identify specific facial features and associate a contextual relationship among features to recognize a face even with altered appearance. Sinha *et al.* [7] established 19 results based on the face recognition capabilities of the human mind. They suggested that humans can efficiently recognize faces even with low resolution and noise. Moreover, high and low frequency facial information is processed both holistically and locally. Campbell *et al.* [8] reported that inner and outer facial regions represent distinct information that is helpful for face recognition. Researchers from cognitive science suggest that local facial fragments can provide robustness against partial occlusion and change in viewpoints [7], [9], [10]. Heisele *et al.* [11] proposed a component based face recognition approach where facial components are used to provide robustness to pose changes. Their approach suggests that assimilating information from different facial regions provides discriminating information useful for face recognition.

To incorporate the above mentioned research findings, this research proposes a granular approach [12], [13] for facial feature extraction and matching. In the granular computing approach, non-disjoint features are extracted at different granular levels. These features are then synergistically combined to obtain more comprehensive information. With granulated information, more flexibility is achieved in analyzing underlying information such as nose, ears, forehead, cheeks, or combination of two or more features. Fig. 3 shows the block diagram of the proposed algorithm.

A. Face Image Granulation

Let F be the detected frontal face image of size $n \times m$. Face granules are generated pertaining to three levels of granularity. The first level of granularity provides global information at multiple resolutions. This is analogous to a human mind processing holistic information for face recognition at varying resolutions. Next, to incorporate the findings of Campbell *et al.* [8], inner and outer facial information are extracted at the second level of granularity. Finally, local facial features play an important role in face recognition by human mind. Therefore, at the third level of granularity, features are extracted from the local facial fragments.

1) First Level of Granularity: In the first level, face granules are generated by applying the Gaussian and Laplacian operators [14]. The Gaussian operator generates a sequence of low pass filtered images by iteratively convolving each of the constituent image with a 2D Gaussian kernel. The resolution and sample density of the image is reduced between successive iterations and therefore the Gaussian kernel operates on a reduced version of the original image in every iteration. The resultant images $I_0, I_1, \ldots, I_A$ may be viewed as a 'pyramid' with I_0 having the highest resolution and I_A having the lowest resolution. Let $\overline{w}(x, y)$ represent the Gaussian kernel of dimension 5×5 with reduction factor 4. The $reduce$ operation, Re, can be written as,

$$Re[F(p, q)] = \sum_{x=1}^{5} \sum_{y=1}^{5} \overline{w}(x, y) F(2p + x, 2q + y) \qquad (1)$$

A Gaussian pyramid I_B is defined as,

$$I_B = Re[I_{B-1}], \quad 0 < B < A \tag{2}$$

where,

$$I_0 = F \tag{3}$$

Further, the Laplacian operator generates band-pass images using Eq. 4.

$$L_B = I_B - Ex[I_{B+1}], \quad 0 \leq B < A \tag{4}$$

Here, the $Ex[\cdot]$ operator interpolates a low-resolution image to the next higher resolution and is represented as,

$$Ex[I_{B,D}(p,q)] = 4 \sum_{x=-2}^{2} \sum_{y=-2}^{2} \overline{w}(x,y) I_{B,D-1}\left(\frac{p-x}{2}, \frac{q-y}{2}\right) \tag{5}$$

Note that $I_{B,D}$ in Equation 5 denotes 'expanding' I_B D number of times. Let the granules generated by Gaussian and Laplacian operators be represented by F_{Gri}, where i represents the granule number. For a face image of size 196×224, Fig. 4 represents the face granules generated in the first level by applying Gaussian and Laplacian operators. F_{Gr1} to F_{Gr3} are the granules generated by Gaussian operator and F_{Gr4} to F_{Gr6} are the granules generated by Laplacian operator. The size of the smallest granule in the first level is 49×56. In these six granules, facial features are segregated at different resolutions to provide edge information, noise, smoothness, and blurriness present in a face image. As shown in Fig. 4, the effect of facial wrinkles is lessened from granule F_{Gr1} to F_{Gr3}. The first level of granularity compensates for the variations in facial texture. Therefore, it provides resilience to plastic surgery procedures that alters the face texture (such as facelift, skin resurfacing, and dermabrasion).

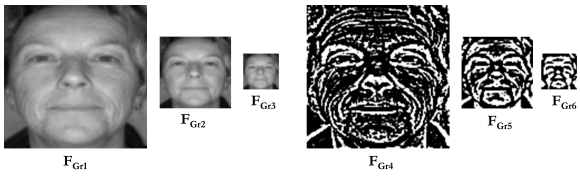

Fig. 4. Face granules in the first level of granularity. F_{Gr1}, F_{Gr2}, and F_{Gr3} are generated by the Gaussian operator, and F_{Gr4}, F_{Gr5}, and F_{Gr6} are generated by the Laplacian operator.

2) Second Level of Granularity: To accommodate the observations of Campbell *et al* [8], horizontal and vertical granules are generated by dividing the face image F into different regions as shown in Figs. 5 and 6. Here, F_{Gr7} to F_{Gr15} denote the horizontal granules and F_{Gr16} to F_{Gr24} denote the vertical granules. Among the nine horizontal granules, the first three granules i.e. F_{Gr7}, F_{Gr8}, and F_{Gr9} are of size $n \times m/3$. The next three granules, i.e., F_{Gr10}, F_{Gr11}, and F_{Gr12} are generated such that the size of F_{Gr10} and F_{Gr12} is $n \times (m-\epsilon)$ and the size of F_{Gr11} is $n \times (m+2\epsilon)$. Further, F_{Gr13}, F_{Gr14}, and F_{Gr15} are generated such that the size of F_{Gr13} and F_{Gr15} is $n \times (m+\epsilon)$ and the size of F_{Gr14} is $n \times (m-2\epsilon)$. Similarly, nine vertical granules, F_{Gr16} to F_{Gr24}, are generated. Figs. 5 and 6 show horizontal and vertical granules when the size of face image is 196×224 and $\epsilon = 15$[1]. The second level of granularity provides resilience to variations in inner and outer facial regions. It utilizes the relation between horizontal and vertical granules to address the variations in chin, forehead, ears, and cheeks caused due to plastic surgery procedures.

Fig. 5. Horizontal face granules from the second level of granularity ($F_{Gr7} - F_{Gr15}$).

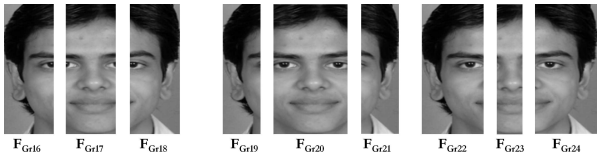

Fig. 6. Vertical face granules from the second level of granularity ($F_{Gr16} - F_{Gr24}$).

3) Third Level of Granularity: As mentioned previously, human mind can distinguish and classify individuals with their local facial fragments such as nose, eyes, and mouth. To incorporate this property, local facial fragments are extracted and utilized as granules in the third level of granularity. Given the eye coordinates, 16 local facial fragments are extracted using the golden ratio face template [15] shown in Fig. 7(a). Each of these fragments is a granule representing local information that provides unique features for handling variations due to plastic surgery. Fig. 7(b) shows an example of local facial fragments used as face granules in the third level of granularity. Generally, the impact of local plastic surgery procedures is confined to facial feature being operated on and its neighboring regions. The third level of granularity provides resilience to variations in local facial regions caused due to such plastic surgery procedures.

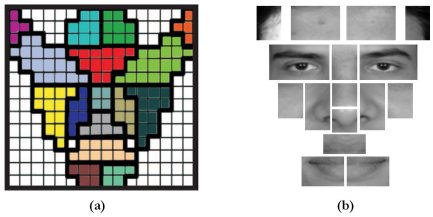

Fig. 7. (a) Golden ratio face template [15], and (b) face granules in the third level of granularity ($F_{Gr25} - F_{Gr40}$).

The proposed granulation technique is used to generate 40 non-disjoint face granules from a face image of size 196×224. The technique used for generating granules is based on fixed structure and no local feature based approach has been utilized.

[1]In the experiments, it is observed that ϵ = 15 yields the best recognition results when face image is of size 196×224.

For images captured from cooperative users, granulation can be performed according to the features. However, with non-cooperative users, identifying features can be challenging and hence feature-based partitioning may not yield accurate results compared to fixed structure partitioning.

B. Facial Feature Extraction

The proposed granulation scheme results in granules with different information content. Some granules contain fiducial features such as eyes, nose, and mouth while some granules predominantly contain skin region such as forehead, cheeks, and outer facial region. Therefore, different kinds of feature extractors are needed to encode diverse information from the granules. In this research, EUCLBP and SIFT are used for feature extraction. Both these feature extractors are fast, discriminating, rotation invariant, and robust to changes in gray level intensities due to illumination. However, the information encoded by these two feature extractors is rather diverse as one encodes the difference in intensity values while the other assimilates information from the image gradients. They efficiently use information assimilated from local regions and form a global image signature by concatenating the descriptors obtained from every local facial region. It is experimentally observed that among the 40 face granules, for some granules EUCLBP finds more discriminative features than SIFT and vice-versa (later shown in the experimental results).

1) Extended Uniform Circular Local Binary Patterns (EUCLBP): Circular Local Binary Patterns (CLBP) encode difference of sign between neighboring pixels that are well separated on a circle around the central pixel [16]. However, encoding difference of sign between the neighboring pixels is not sufficient for describing facial texture. Other important features could also be derived from the information that lies in the difference of the gray-level intensity values. Huang *et al.* [17] proposed a method to encode the exact difference of gray-level intensities and reported significant improvement in the performance of texture descriptor. Based on this observation, CLBP is extended to encode exact gray-level difference along with the original encoding. The modified descriptor is called Extended Uniform Circular Local Binary Pattern (EUCLBP) [5]. For computing EUCLBP descriptor, the image is first tessellated into non-overlapping uniform local patches of size 32×32. For each local patch, the EUCLBP descriptor is computed based on the 8 neighboring pixels uniformly sampled on a circle (radius=2) centered at the current pixel. The concatenation of descriptors from each local patch constitutes the image signature. Two EUCLBP descriptors are matched using the weighted χ^2 distance.

2) Scale Invariant Feature Transform (SIFT): SIFT [6] is a scale and rotation invariant descriptor that generates a compact representation of an image based on the magnitude, orientation, and spatial vicinity of image gradients. SIFT, as proposed by Lowe [6], is a sparse descriptor that is computed around detected interest points. However, SIFT can also be used in a dense manner where the descriptor is computed around pre-defined interest points. In this research, SIFT descriptor is computed in a dense manner. It is computed over a set of uniformly distributed non-overlapping local regions of size 32×32. SIFT descriptors computed at the sampled regions are then concatenated to form the image signature. Weighted χ^2 distance is used to compare two SIFT descriptors.

C. Evolutionary Approach for Selection of Feature Extractor and Weight Optimization

Psychological studies in face recognition [7] have shown that some facial regions are more discriminating than others and hence, contribute more towards the recognition accuracy. Moreover, humans [18] also emphasize on different internal and external facial regions for recognition. Every face granule has useful but diverse information, which if combined together can provide discriminating information for face recognition. The proposed algorithm therefore incorporates selecting feature extractor to encode diverse information and assigning optimal weights for matching each face granule. The uniqueness and discriminability of EUCLBP and SIFT features depend on the information present in the granules. Each feature extractor can better represent some granules as compared to the other feature extractor. Based on this hypothesis, feature extractor for each granule is selected depending on the reliability of features for that particular granule. It is our assertion that giving higher weight to face granules that have more contribution towards the recognition performance and selecting better feature extractor for each granule should improve the overall accuracy. (Both these assertions are later validated by the experimental results).

The next task is simultaneously optimizing the selection of feature extractor and weights associated with every face granule for matching. The problem of finding better feature extractor and optimal weights for each granule involves searching very large space and finding several suboptimal solutions. Genetic algorithms are well proven in searching very large spaces to quickly converge to the near optimal solution [19]. Therefore, a multi-objective evolutionary genetic approach is proposed to select feature extractor and corresponding weights for each face granule. Fig. 8 represents the genetic search process and the steps involved are elaborated as follows:

Genetic Encoding: A chromosome is a string whose length is equal to the number of face granules i.e. 40 in our case. For simultaneous optimization of two functions, two types of chromosomes are encoded: (i) for selecting feature extractor (referred to as chromosome $type1$) and (ii) for assigning weights to each face granule (referred to as chromosome $type2$). Each gene (unit) in chromosome $type1$ is a binary bit 0 or 1 where 0 represents the SIFT feature extractor and 1 represents the EUCLBP feature extractor. Genes in chromosome $type2$ have real valued numbers associated with corresponding weights of the 40 face granules.

Initial Population: Two generations with 100 chromosomes are populated. One generation has all $type1$ chromosomes while the other generation has all $type2$ chromosomes.

1) For selecting feature extractors ($type1$ chromosome), half the initial generation (i.e. 50 chromosomes) is set with all the genes (units) as 1, which represents EUCLBP as the feature extractor for all 40 face granules.

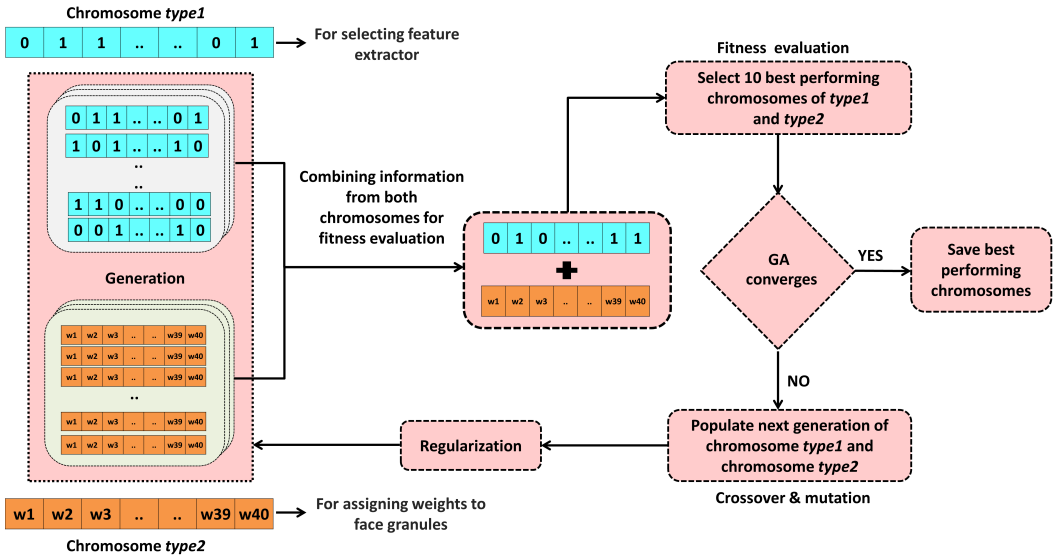

Fig. 8. Genetic optimization process for selecting feature extractor and weight for each face granule.

The remaining 50 chromosomes in the initial generation have all genes as 0 representing SIFT as the feature extractor for all 40 face granules.

2) For assigning weights to face granules ($type2$ chromosome), a chromosome with weights proportional to the identification accuracy of individual face granules (as proposed by Ahonen [16]) is used as the seed chromosome. The remaining 99 chromosomes are generated by randomly changing one or more genes in the initial chromosome. The weights are normalized such that the sum of all the weights in a chromosome is 1.

Fitness Function: Both chromosome $type1$ and chromosome $type2$ are combined and evaluated simultaneously. Recognition is performed using the feature extractor selected by chromosome $type1$ and weight encoded by chromosome $type2$ for each face granule. Identification accuracy, used as fitness function, is computed on the training set and 10 best performing chromosomes are selected as *parents* to populate the next generation.

Crossover: A set of uniform crossover operations is performed on *parents* to populate a new generation of 100 chromosomes. Crossover operation is same for both chromosome $type1$ and chromosome $type2$.

Mutation: After crossover, mutation for chromosome $type2$ is performed by changing one or more weights by a factor of its standard deviation in the previous generation. For chromosome $type1$, mutation is performed by randomly inverting the genes in the chromosome.

The search process is repeated till convergence and it is terminated 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, the feature extractor and optimal weights for each face granule (i.e. chromosomes giving best recognition accuracy on the training data) are obtained. Genetic optimization also enables discarding redundant and non-discriminating face granules that do not contribute much towards the recognition accuracy (i.e. the weight for that face granule is close to 0). This optimization process leads to both dimensionality reduction and better computational efficiency.

Evolutionary algorithms such as GA often fail to maintain diversity among individual solutions (chromosomes) and cause the population to converge prematurely. This problem is attributed to the loss of diversity in a population that leads to decrease in the quality of solution. In this research, adaptive mutation rate [20] and random offspring generation [21] are used to prevent premature convergence to local optima by ensuring sufficient diversity in a population. Depending on population's diversity, mutation is performed with an adaptive rate that increases if the diversity decreases and vice-versa. Population diversity is measured as the standard deviation of fitness values in a population. Further, random offspring generation is used to produce random offsprings if there is a high degree of similarity among the participating chromosomes (parents) during the crossover operation. Combination of such chromosomes is ineffective because it leads to offsprings that are exactly similar to parents. Therefore, under such conditions, crossover is not performed and offsprings are generated randomly.

D. Combining Face Granules with Evolutionary Learning for Recognition

The granular approach for matching faces altered due to plastic surgery is summarized below:

1) For a given gallery-probe face image pair, 40 face granules are extracted from each image.
2) EUCLBP or SIFT features are computed for each face granule according to the evolutionary model (learnt using the training data).
3) The descriptors extracted from the gallery and probe images are normalized and the weighted χ^2 distance

measure is used for matching.

$$\chi^2(a,b) = \sum_{i,j} \omega_j \frac{(a_{i,j} - b_{i,j})^2}{a_{i,j} + b_{i,j}} \quad (6)$$

where a and b are the normalized descriptors (EUCLBP and SIFT respectively), i and j correspond to the i^{th} bin of the j^{th} face granule, and ω_j is the weight of the j^{th} face granule. Here, the weights for each face granule are learnt using the genetic algorithm.

4) In identification mode ($1:N$), this procedure is repeated for each gallery-probe pair and top matches are obtained based on the match scores.

III. Experimental Results

Several experiments are performed to evaluate the performance of the proposed algorithm. The performance of the algorithm is also compared with SIFT and EUCLBP applied on full face image, SIFT and EUCLBP applied on the 40 face granules, sum-rule fusion [22] of SIFT and EUCLBP on face granules, and a commercial-off-the-shelf (COTS) face recognition system. Section III-A provides details about the databases used in this research, Section III-B elaborates the experimental protocol, and Sections III-C to III-E present comprehensive experimental analysis.

A. Database

In this research, experiments are performed on two databases: (a) plastic surgery face database [2] and (b) combined heterogeneous face database. The plastic surgery face database comprises 1800 pre- and post-surgery images corresponding to 900 subjects with frontal pose, proper illumination, and neutral expression. It is the first real world database that consists of different types of facial plastic surgery cases such as rhinoplasty (nose surgery), blepharoplasty (eyelid surgery), brow lift, skin peeling, and rhytidectomy (face lift). Since, it is difficult to isolate individuals who have undergone plastic surgery and use special mechanism to recognize them. Therefore, face recognition algorithms should be robust to variations induced by plastic surgery even in general operating environments. Considering such generality of face recognition, the second database is prepared by appending the plastic surgery face database with 1800 non-surgery images pertaining to other 900 subjects. This database is termed as the *combined heterogeneous face database* and comprises 3600 images pertaining to 1800 subjects. The non-surgery images are from the same databases used by Singh *et al.* [2] and consists of two frontal images per subject with proper illumination and neutral expression.

Images in the plastic surgery face database are collected from different sources on the internet and have noise and irregularities. The images in the database are first preprocessed to zero mean and unit variance followed by applying histogram equalization to maximize the image contrast by flattening the resulting histogram. Further, Wiener filtering is applied to restore the blurred edges. Finally, the face images are geometrically normalized to 100 pixel inter-eye distance and the size of each detected face image is 196×224 pixels.

B. Experimental Protocol

To evaluate the efficacy of the proposed algorithm, experiments are performed with 10 times repeated random subsampling validations. In each experiment, 40% of the database is used for training and the remaining 60% is used for testing. The training data is used to learn the model for (EUCLBP/SIFT) feature selection, weights for each face granule, and the testing data is used for performance evaluation. Experimental protocol for all the experiments are described here:

- *Experiment 1:* 1800 pre- and post-surgery images pertaining to 900 subjects from the plastic surgery face database [2] are used in this experiment. Images of 360 subjects are used for training and the performance is evaluated on pre- and post-surgery images of the remaining 540 subjects. Pre-surgery images are used as the gallery and post-surgery images are used as the probe.
- *Experiment 2:* Out of 1800 subjects from the combined heterogeneous face database, 720 subjects are used for training and the remaining 1080 subjects are used for testing. The training subjects are randomly selected and there is no regulation on the number of training subjects that have undergone plastic surgery. This experiment resembles real world scenario of training-testing where the system is unaware of any plastic surgery cases.
- *Experiment 3:* To evaluate the effectiveness of the proposed algorithm for matching individuals against large size gallery, two different experiments are performed. In both the experiments, 6324 frontal face images obtained from government agencies are appended to the gallery of 1800 face images used in other experiments.
  - Training is performed with images of 360 subjects from the plastic surgery face database. The performance is evaluated on post-surgery images from the remaining 540 subjects as probes against the large scale gallery of 7764 subjects.
  - Training is performed with images of 720 subjects from the combined heterogeneous face database. The performance is evaluated on images from the remaining 1080 subjects as probes against the large scale gallery of 7404 subjects.

C. Analysis

The proposed algorithm utilizes the observation that human mind recognizes face images by analyzing the relation among non-disjoint spatial features extracted at multiple granular levels. Further, simultaneously optimizing the feature selection and weights pertaining to each face granule allows for addressing the spontaneous and non-linear variations introduced by plastic surgery. Key results and observations from the experiments are summarized below.

- The CMC curves in Fig. 9 and Table I show rank-1 identification accuracy of the algorithms for Experiments 1 and 2. The proposed algorithm outperforms other algorithms by at least 5.27% on the plastic surgery face database and 5.02% on the combined heterogeneous face

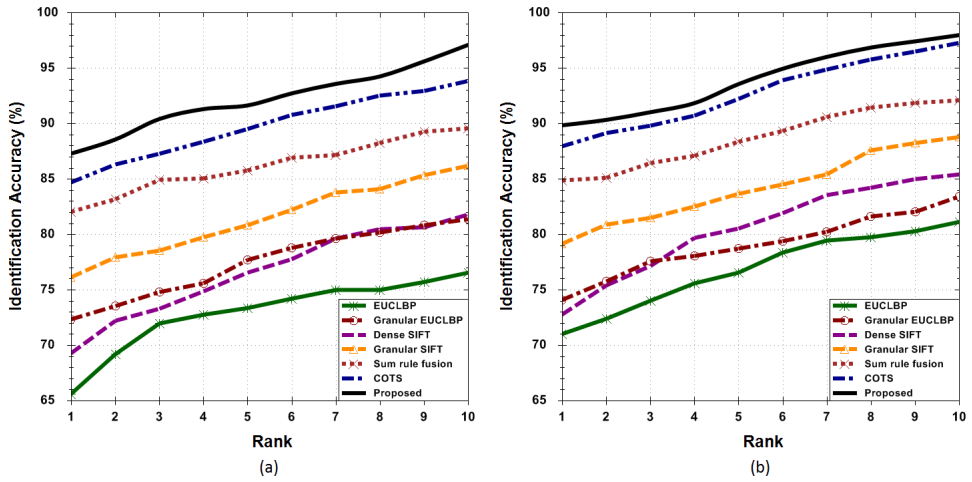

Fig. 9. CMC curves for the proposed and existing algorithms on the (a) plastic surgery face database, and (b) combined heterogeneous face database.

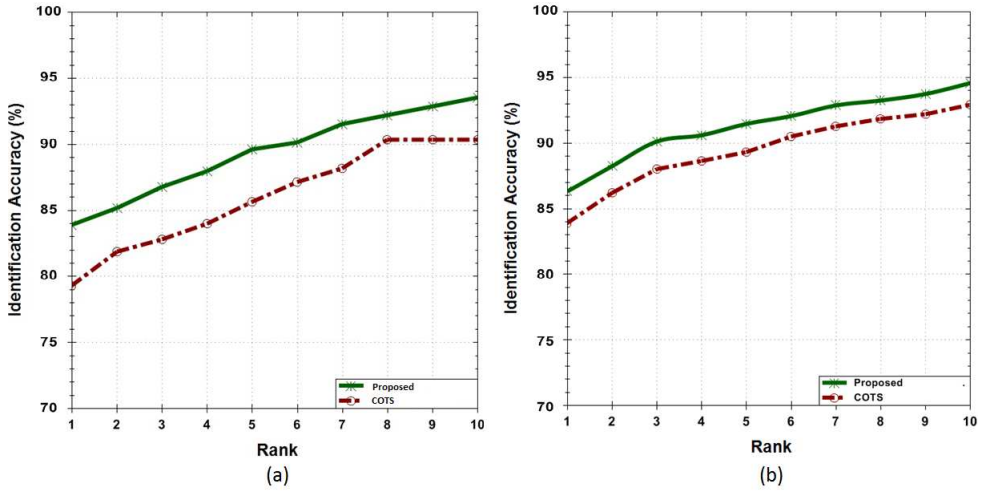

Fig. 10. CMC curves for the proposed and commercial algorithms for large scale evaluation on probe images from the (a) plastic surgery face database, and (b) combined heterogeneous face database.

TABLE I
RANK-1 IDENTIFICATION ACCURACY OF THE PROPOSED EVOLUTIONARY GRANULAR APPROACH AND OTHER FACE RECOGNITION ALGORITHMS. IDENTIFICATION ACCURACIES AND STANDARD DEVIATIONS (SD) ARE COMPUTED WITH 10 TIMES CROSS VALIDATION.

Database	#Train/#Test subjects	Algorithm	Rank-1 accuracy	SD
Plastic surgery face database	360/540	EUCLBP	65.56%	0.73
		SIFT	69.26%	1.08
		Granular EUCLBP	72.35%	0.64
		Granular SIFT	76.11%	1.33
		Sum Rule Fusion	82.05%	0.78
		COTS	84.66%	0.78
		Proposed	**87.32%**	**0.68**
Combined heterogeneous face database	720/1080	EUCLBP	70.98%	0.78
		SIFT	72.75%	1.28
		Granular EUCLBP	74.08%	0.68
		Granular SIFT	79.12%	2.03
		Sum Rule Fusion	84.85%	1.06
		COTS	87.94%	1.06
		Proposed	**89.87%**	**0.82**

TABLE II
RANK-1 IDENTIFICATION ACCURACY OF FACE GRANULES USING SIFT AND EUCLBP.

Granule	SIFT	EUCLBP	Granule	SIFT	EUCLBP
F_{Gr1}	69.26%	65.56%	F_{Gr21}	14.12%	22.08%
F_{Gr2}	51.42%	42.26%	F_{Gr22}	19.25%	23.96%
F_{Gr3}	46.18%	21.32%	F_{Gr23}	23.64%	19.25%
F_{Gr4}	22.86%	36.20%	F_{Gr24}	20.88%	23.94%
F_{Gr5}	20.15%	25.75%	F_{Gr25}	09.72%	5.50%
F_{Gr6}	16.26%	19.50%	F_{Gr26}	19.36%	8.85%
F_{Gr7}	10.46%	19.38%	F_{Gr27}	18.12%	12.50%
F_{Gr8}	39.06%	28.64%	F_{Gr28}	09.22%	7.25%
F_{Gr9}	17.85%	23.42%	F_{Gr29}	17.36%	22.50%
F_{Gr10}	13.14%	19.64%	F_{Gr30}	08.54%	6.48%
F_{Gr11}	41.43%	32.38%	F_{Gr31}	18.52%	22.86%
F_{Gr12}	28.20%	24.44%	F_{Gr32}	14.24%	6.48%
F_{Gr13}	16.88%	22.02%	F_{Gr33}	13.16%	11.24%
F_{Gr14}	33.06%	23.84%	F_{Gr34}	11.35%	05.65%
F_{Gr15}	30.56%	24.68%	F_{Gr35}	10.75%	7.94%
F_{Gr16}	15.76%	21.84%	F_{Gr36}	15.10%	13.54%
F_{Gr17}	33.12%	25.50%	F_{Gr37}	12.64%	6.28%
F_{Gr18}	15.64%	21.28%	F_{Gr38}	12.20%	10.38%
F_{Gr19}	11.82%	20.10%	F_{Gr39}	22.86%	12.82%
F_{Gr20}	51.60%	44.40%	F_{Gr40}	24.92%	11.18%

TABLE III
RANK-1 IDENTIFICATION PERFORMANCE ON DIFFERENT TYPES OF LOCAL AND GLOBAL PLASTIC SURGERY PROCEDURES.

Type	Surgery	#Cases	PCA	FDA	LFA	CLBP	SURF	GNN	Periocular	Proposed
Local	Browlift	60	28.5%	31.8%	39.6%	49.1%	51.1%	57.2%	34.42%	89.22%
	Dermabrasion	32	20.2%	23.4%	25.5%	42.1%	42.6%	43.8%	44.56%	77.89%
	Otoplasty	74	56.4%	58.1%	60.7%	68.8%	66.4%	70.5%	47.25%	92.25%
	Blepharoplasty	105	28.3%	35.0%	40.2%	52.1%	53.9%	61.4%	30.96%	91.42%
	Rhinoplasty	192	23.1%	24.1%	35.4%	44.8%	51.5%	54.3%	40.71%	88.85%
	Other	56	26.4%	33.1%	41.4%	52.4%	62.6%	58.9%	35.81%	89.17%
Global	Rhytidectomy	308	18.6%	20.0%	21.6%	40.9%	40.3%	42.1%	37.27%	71.76%
	Skin peeling	73	25.2%	31.5%	40.3%	53.7%	51.1%	53.9%	45.83%	85.09%
	Overall	**900**	**27.2%**	**31.4%**	**37.8%**	**47.8%**	**50.9%**	**53.7%**	**40.11%**	**87.32%**

database. The proposed algorithm also outperforms the commercial system by 2.66% and 1.93% on the plastic surgery face database and the combined heterogeneous face database respectively.

- In Experiment 2, the training-testing partitions have plastic surgery as well as non-surgery images. It closely resembles the condition which a real world face recognition system encounters in general operating environment. Unacquainted with specific plastic surgery cases, face recognition system has to be robust in matching surgically altered face images in addition to matching regular face images. Different types of plastic surgery procedures have varying effect on one or more facial regions. The proposed algorithm inherently provides the benefit of addressing the non-linear variations induced by different types of plastic surgery procedures.
- CMC curves in Fig. 10 show the performance of the proposed algorithm and the commercial system for matching probes against a large gallery (Experiment 3). The proposed algorithm yields rank-1 identification accuracy of 83.88% and 86.33% for matching probes from the plastic surgery and the combined heterogeneous face databases respectively. The proposed algorithm outperforms the commercial system by 4.6% and 2.21% on the plastic surgery and the combined heterogeneous face database respectively.
- Table II shows individual rank-1 identification accuracy of all 40 face granules using EUCLBP and SIFT on the plastic surgery face database. Face granules $4, 7, 19, 21, 29$, and 31 yield significantly better recognition performance with EUCLBP as compared to SIFT. On the other hand, face granules $2, 3, 8, 11, 14, 26, 39$, and 40 provide better recognition performance with SIFT as compared to EUCLBP. SIFT generally performs better on granules containing fiducial features such as eyes, nose, and mouth, however its performance on predominant skin regions such as forehead, cheeks, and outer facial region is not optimal. Since, EUCLBP is based on exact difference of gray level intensities, it can better encode discriminating micro patterns even from predominant skin regions.
- Table III shows a comprehensive breakup of rank-1 identification accuracy according to the type of surgeries performed. The proposed algorithm provides a significant improvement of at least 21.7% from the algorithms on different types of plastic surgery procedures[2] used in [2].
- Recently, Aggarwal *et al.* [4] proposed a sparse representation based approach to match surgically altered face images in a part-wise manner. The proposed evolutionary granular algorithm also outperforms the sparse representation [4] based approach by 9.4% on the plastic surgery face database under the same experimental protocol.
- The performance of EUCLBP and SIFT when applied on full face images is compared with the performance obtained when it is applied on face granules. The results show that applying EUCLBP and SIFT on face granules improves the rank-1 accuracy by at least 3% as compared to a full face image. The ability to encode local features at different resolutions and sizes (face granules) allows the proposed algorithm to be resilient to the non-linear variations induced by plastic surgery procedures.
- To show the efficacy of the evolutionary approach for selecting feature extractor and weight optimization using genetic algorithm, performance of the proposed algorithm is compared with sum-rule fusion [22] of SIFT and EUCLBP on face granules. The proposed algorithm outperforms sum-rule fusion by at least 5% on both the databases.
- Evolutionary approach for selecting feature extractor using genetic algorithm provides the advantage of choosing better performing feature extractor for each face granule. It is observed in our experiments that on average, SIFT is selected for 22 face granules whereas EUCLBP is selected for 18 face granules.
- From non-parametric rank-ordered test (Mann-Whitney test on the ranks obtained from the algorithms), it can be concluded that there is a statistically significant difference between the proposed algorithm and the commercial system. Further, at 95% confidence level, parametric t-test (using the match scores) also suggests that the proposed algorithm and the commercial system are statistically significantly different.

D. Analysis of Different Types of Plastic Surgery Procedures

According to Singh *et al.* [2], plastic surgery procedures can be categorized into global and local plastic surgery. Global plastic surgery completely transforms the face and is recommended in cases where functional damage is to be cured

[2]Since the experimental protocol in [2] and the current research is same, the results are directly compared.

TABLE IV
PEARSON CORRELATION COEFFICIENT BETWEEN DIFFERENT GRANULAR LEVELS ON THE PLASTIC SURGERY FACE DATABASE.

Database	Granules	Genuine correlation	Impostor correlation
Plastic surgery face database	Level 1 - Level 2	0.67	0.59
	Level 1 - Level 3	0.43	0.21
	Level 2 - Level 3	0.63	0.55
Combined heterogeneous face database	Level 1 - Level 2	0.81	0.78
	Level 1 - Level 3	0.38	0.20
	Level 2 - Level 3	0.42	0.26

such as patients with fatal burns or trauma. In these kind of surgeries, facial appearance, skin texture, and feature shapes vary drastically thus making it arduous for any face recognition system to recognize pre- and post-surgery faces. Rhytidectomy (full facelift) is used to treat patients with severe burns on face and neck. It can also be used to reverse the effect of aging and get a younger look by treating the face skin, thus modifying the appearance and texture of the whole face. Analogous to rhytidectomy, skin peeling procedures such as laser resurfacing and chemical peel alter the texture information thus affecting the performance of face recognition algorithms. These procedures are used to treat wrinkles, stretch marks, acne, and other skin damages caused due to aging and sun burns. These two global plastic surgery procedures severely impact the performance of the proposed algorithm which yields rank-1 identification accuracy of 71.76% and 85.09% for cases with rhytidectomy and skin peeling respectively, as shown in Fig. 11.

On the other hand, local plastic surgery [2] is meant for reshaping and restructuring facial features to improve the aesthetics. These surgical procedures result in varying amount of change in the geometric distance between facial features but the overall texture and appearance of the face remains similar to the original face. Dermabrasion is used to give a smooth finish to face skin by correcting the skin damaged by sun burns or scars (developed as a post-surgery effect), dark irregular patches (melasma) that grow over the face skin, and mole removal. Among all the local plastic surgery procedures listed in [2], dermabrasion has the most prominent effect on the performance of the proposed algorithm as it drastically changes the face texture. As shown in Fig. 11, the proposed approach yields rank-1 identification accuracy of 77.89% for dermabrasion cases.

Other local plastic surgery procedures also affect the performance of the proposed algorithm to varying degrees. Table III and CMC curves in Fig. 11 show the performance of the proposed algorithm across different types of global and local plastic surgery procedures. Table III also reports the performance of existing algorithms on several local and global plastic surgery procedures. Since plastic surgery procedures increase the difference between pre- and post-surgery images of the same individual (intra-class variations), it drastically reduces the performance of existing face recognition algorithms.

E. Analysis of Granules

To understand the contribution of different granules for recognizing face images altered due to plastic surgery, a detailed experimental study of individual granular levels is performed. The correlation analysis of all three granular levels is reported in Table IV. The complimentary information vested in different granular levels is utilized by the proposed algorithm for efficiently matching surgically altered face images. CMC curves in Figs. 12(a) and (b) show the identification accuracy of individual granular levels for the plastic surgery face database and the combined heterogeneous face database respectively. Granular level-1 has different levels of Gaussian and Laplacian pyramids that assimilate discriminating information across multiple resolutions. Pyramids at level-0 contain minute features whereas the pyramids at level-1 and level-2 provide high level prominent features of a face. Several psychological studies have shown that humans use different inner and outer facial features to identify individuals [18]. The inner facial features include nose, eyes, eyebrows, and mouth while the outer facial region comprises face outline, structure of jaw/chin, and forehead. Granular level-2 therefore extracts information from different inner and outer facial regions representing discriminative information that is useful for face recognition. Local facial fragments such as nose, eyes, and mouth provide robustness to variations in several local facial regions caused due to plastic surgery procedures. Human mind can efficiently distinguish and classify individuals with their local facial fragments. Therefore, granular level-3 assimilates discriminating information from these fragments. The proposed granular approach thus unifies diverse information from different granular levels that are useful for recognizing faces altered due to plastic surgery. Further, to analyze the complimentary information provided by different granular levels, performance is evaluated for different combinations of granular levels. The performance of the proposed evolutionary granular approach is optimized for a particular granular level or their combination by assigning null weights to the face granules corresponding to other granular levels during genetic optimization. CMC curves in Figs. 12(c) and (d) show the results for different combinations of granular levels on the plastic surgery face database and the combined heterogeneous face database respectively.

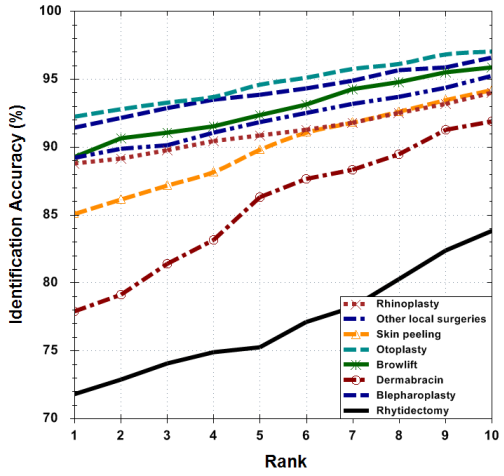

Fig. 11. CMC curves on different types of local and global plastic surgery procedures for the proposed algorithm.

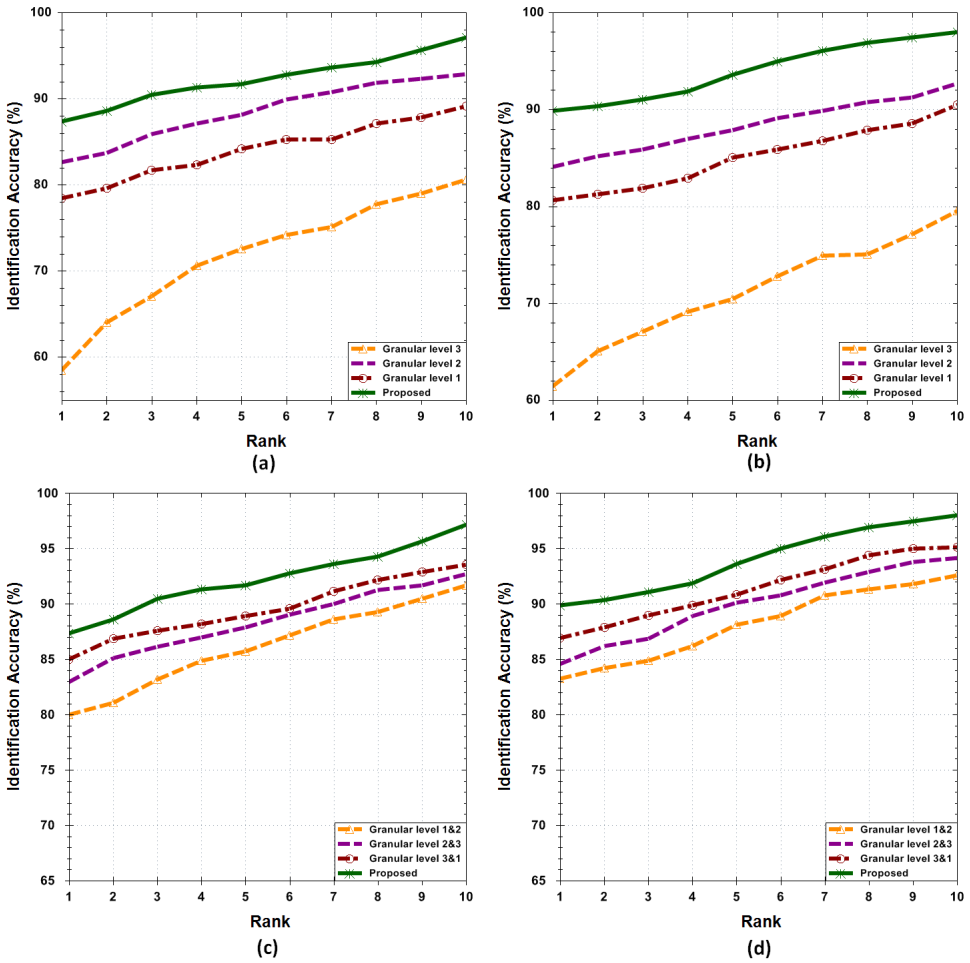

Fig. 12. CMC curves showing the performance of individual granular levels on the (a) plastic surgery face database, (b) combined heterogeneous face database, performance of combination of granular levels on the (c) plastic surgery face database, and (d) combined heterogeneous face database.

Fig. 13. F_{Gr29} represents the right periocular region and F_{Gr31} represents the left periocular region.

In the proposed granulation scheme, F_{Gr29} and F_{Gr31} represent the right and left periocular regions as shown in Fig. 13. The identification accuracies in Table II show that among all local facial granules in the third level of granularity, F_{Gr29} and F_{Gr31} provide good performance with both SIFT and EUCLBP features. Periocular region is used as a biometric when the face is occluded [23] or the iris cannot be captured [24]. Recently, Juefei-Xu *et al.* [25] proposed using periocular region for age invariant face recognition and reported substantial improvements in both verification and identification performance. Driven by the robustness of periocular biometric against occlusion and aging, this research also evaluates the performance of periocular biometrics for recognizing surgically altered face images. The experiments are performed using the experimental protocol of Experiment 1 in Section III-B. CMC curves in Fig. 14(a) show the performance of periocular region for matching surgically altered faces from the plastic surgery face database. The performance is computed individually for the left and right periocular region using SIFT and EUCLBP. Sum-rule fusion [22] of SIFT on the left and right periocular region (fusion of SIFT) and sum-rule fusion of EUCLBP on the left and right periocular region (fusion of EUCLBP) is also reported. Finally, the overall performance of periocular region is computed based on the sum-rule fusion of SIFT and EUCLBP on left and right periocular regions (fusion of SIFT and EUCLBP). The performance is also compared with an existing periocular based recognition algorithm, referred to as Bharadwaj *et al.* [24].

Eyelid is the thin skin that covers and protects our eyes and is a major feature in periocular recognition algorithms. According to the statistics provided by American Society for Aesthetic Plastic Surgery [1], blepharoplasty (eyelid surgery) is identified as one of the top five surgical procedures performed in 2010. It is used to reshape upper and lower eyelids to treat excessive growth of skin tissues obstructing vision. Some other global plastic surgery procedures such as rhytidectomy or skin peeling may also affect the periocular region. Therefore, it is important to analyze the effect of

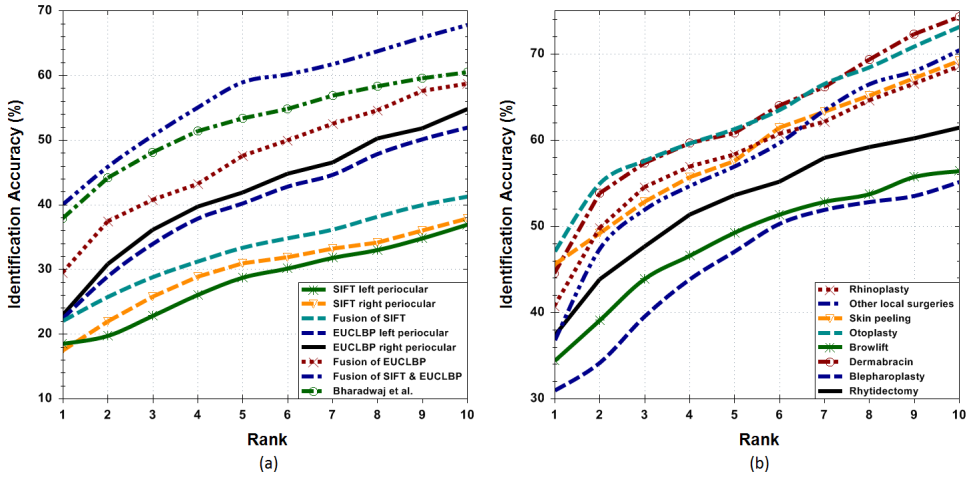

Fig. 14. CMC curves showing the performance of periocular regions on (a) the plastic surgery face database, and (b) different local and global plastic surgery procedures.

surgical procedures, including blepharoplasty, on periocular region as a biometric. Experiments are performed to analyze the effect of different global and local plastic surgery procedures (especially blepharoplasty) on periocular region. CMC curves in Fig. 14(b) and Table III report rank-1 identification accuracy of periocular region for matching faces altered due to specific types of plastic surgery. Blepharoplasty alters the periocular region thereby affecting the performance of periocular biometrics. Moreover, it is also observed that the performance of periocular biometrics is reduced when a local region neighboring the periocular region (such as nose and forehead) is transformed due to plastic surgery. This is mainly because modifying local features also transmits some vicissitudes in the adjacent facial regions. The results suggest that although, periocular biometrics has shown robustness to aging and occlusion, plastic surgery is an important challenge for periocular recognition algorithms.

IV. Conclusion

Plastic surgery has emerged as a new covariate of face recognition and its allure has made it indispensable for face recognition algorithms to be robust in matching surgically altered face images. This research presents an evolutionary granular algorithm that operates on several granules extracted from a face image. The first granular level processes the image with Gaussian and Laplacian operators to assimilate information from multi-resolution image pyramids. The second granular level tessellates the image into horizontal and vertical face granules of varying size and information content. The third granular level extracts discriminating information from local facial regions. Further, a multi-objective evolutionary genetic algorithm is proposed for selecting feature extractor and assigning optimal weights to each face granule for matching. The evolutionary selection of feature extractor allows switching between two feature extractors (SIFT and EUCLBP) that helps in encoding discriminating information complying with the face granules. The proposed algorithm utilizes the observation that human mind recognizes face image by analyzing the relation among non-disjoint spatial features extracted at different granularity. Experiments under different protocols, including large scale matching, show that the proposed algorithm outperforms existing algorithms including a commercial system for matching surgically altered face images. Further, experiments on different types of local and global plastic surgery also show that the proposed algorithm consistently outperforms other existing algorithms. Detailed analysis on the contribution of three granular levels and individual face granules corroborates the hypothesis that the proposed algorithm unifies diverse information from all granules to address the non-linear variations in pre-and post-surgery images.

V. Acknowledgement

A shorter version of this manuscript is published in the proceedings of IEEE Conference on Automatic Face and Gesture Recognition, 2011. H.S. Bhatt is partly supported by the IBM PhD fellowship. R. Singh is partially supported through a grant from DST under FAST track. This research is also supported in part by DIT, India.

References

[1] "American society for aesthetic plastic surgery 2010 statistics", *http://www.surgery.org/media/statistics*, 2010.

[2] R. Singh, M. Vatsa, H.S. Bhatt, S. Bharadwaj, A. Noore, and S.S. Nooreyezdan, "Plastic surgery: A new dimension to face recognition", *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 441–448, 2010.

[3] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Robust face recognition after plastic surgery using local region analysis", in *Proceedings of International Conference on Image Analysis and Recognition*, 2011, vol. 6754, pp. 191–200.

[4] G. Aggarwal, S. Biswas, P.J. Flynn, and K.W. Bowyer, "A sparse representation approach to face matching across plastic surgery", in *Proceedings of Workshop on the Applications of Computer Vision*, 2012, pp. 1–7.

[5] H.S. 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, pp. 1–7.

[6] D.G. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

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

[8] R. Campbell, M. Coleman, J. Walker, P.J. Benson, S. Wallace, J. Michelotti, and S. Baron-Cohen, "When does the inner-face advantage in familiar face recognition arise and why?", *Visual Cognition*, vol. 6, no. 2, pp. 197–216, 1999.

[9] W.G. Hayward, G. Rhodes, and A. Schwaninger, "An own-race advantage for components as well as configurations in face recognition", *Cognition*, vol. 106, no. 2, pp. 1017–1027, 2008.

[10] A. Schwaninger, J.S. Lobmaier, and S.M. Collishaw, "Role of featural and configural information in familiar and unfamiliar face recognition", in *Proceedings of International Workshop on Biologically Motivated Computer Vision*, 2002, pp. 245–258.

[11] B. Heisele, P. Ho, J. Wu, and T. Poggio, "Face recognition: component-based versus global approaches", *Computer Vision and Image Understanding*, vol. 91, pp. 6–21, 2003.

[12] A. Bargiela and W. Pedrycz, *Granular computing: An introduction*, Kluwer Academic Publishers, Boston, 2002.

[13] T.Y. Lin, Y.Y. Yao, and L.A. Zadeh, *Data mining, rough sets and granular computing*, Physica-Verlag, 2002.

[14] P.J. Burt and E.H. Adelson, "A multiresolution spline with application to image mosaics", *ACM Transaction on Graphics*, vol. 2, no. 4, pp. 217–236, 1983.

[15] K. Anderson and P.W. McOwan, "Robust real-time face tracker for cluttered environments", *Computer Vision and Image Understanding*, vol. 95, pp. 184–200, 2004.

[16] 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.

[17] G.B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild : A database for studying face recognition in unconstrained environment", in *Proceedings of Faces in Real-Life Images workshop at European Conference on Computer Vision*, 2008.

[18] A.W. Young, D.C. Hay, K.H. McWeeny, B.M Flude, and A.W. Ellis, "Matching familiar and unfamiliar faces on internal and external features", *Perception*, vol. 14, no. 6, pp. 737–746, 1985.

[19] E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Longman Publishing Co., Inc., 1989.

[20] F. Vafaee and P.C. 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.

[21] 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.

[22] A. Ross and A. Jain, "Information fusion in biometrics", *Pattern Recognition Letters*, vol. 24, pp. 2115–2125, 2003.

[23] U. Park, R.R. Jillela, A. Ross, and A.K. Jain, "Periocular biometrics in the visible spectrum", *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pp. 96–106, 2011.

[24] S. Bharadwaj, H.S. Bhatt, M. Vatsa, and R. Singh, "Periocular biometrics: When iris recognition fails", in *Proceedings of International Conference on Biometrics: Theory Applications and Systems*, 2010, pp. 1–6.

[25] F. Juefei-Xu, K. Luu, M. Savvides, T. Bui, and C.Y. Suen, "Investigating age invariant face recognition based on periocular biometrics", in *Proceedings of International Joint Conference on Biometrics*, 2011, pp. 1–7.