Covariates of Face Recognition

Submitted by: Himanshu Sharad Bhatt
Advisors: Dr. Richa Singh and Dr. Mayank Vatsa

***Abstract*—Face recognition has found several applications ranging from cross border security, surveillance, access control, multimedia to forensics. Face recognition under variations due to pose, illumination, and expression has been extensively studied in literature and several approaches have been proposed to address these covariates. Many applications of face recognition require matching face images with variations in age and disguise such as matching a recent photo with your passport image or image on driver's license. In literature, techniques have also been proposed to recognize face images with variations in age and disguise. These challenges can be grouped as *existing covariates* of face recognition. However, with ever increasing applications of face recognition there has emerged a need to understand new fascinating challenges in face recognition, *emerging covariates* of face recognition. Covariates such as forensic sketches, surgically altered faces, low resolution faces, and look-alikes or twins are some of the challenges that have emerged as new covariates of face recognition. These covariates have important law enforcement applications; therefore, it has now become imperative for current face recognition systems to be robust to these challenges. This report focuses on three different aspects. First, it presents a review of different techniques proposed to address the *existing covariates*, limitations of current techniques and future scope of advancements. Second, it presents how the *emerging covariates* have evolved, what are the challenges, proposed techniques, and future research directions for each of these covariates. Finally, the report presents an evolutionary granular approach to address one of the *emerging covariate*, plastic surgery.**

I. Introduction

Biometrics, a part of Identification Science, measures physical and behavioral characteristics of human body, such as fingerprint, retina, iris, voice, face, hand geometry, and gait. Automatic verification of individuals based on these physical (e.g., fingerprint, iris) or behavioral (e.g., gait, key-strokes) characteristics is referred to as biometric verification. Id cards or passwords that are generally used for verifying individuals can be stolen, forgotten, forged, or lost, whereas, biometric authentication is more robust to theft and forgery. It is based on what an individual inherently posses rather than what the individual carries (ID cards) or knows (passwords). Biometrics has found several applications in both civilian and government undertakings. Law enforcement agencies such as homeland security and border security require automated systems to determine the identity of an individual for apprehending criminals and finding missing individuals. Enterprize solutions require biometrics based access control, surveillance, visitor management authentication system, and customer verification at the point of sale. With advancement in technology and its deep penetration among large masses, biometric based authentication also includes e-commerce applications where extra security is required in order to avoid frauds in business transactions and identity thefts.

In general, a biometric system compares the given biometric trait of an individual with stored templates in the database and computes a match scores for the comparison. Depending on the context, a biometric system can operate either in a verification (1:1 matching) or an identification (1:N matching) mode. Verification mode involves confirming or denying the identity claimed by an individual. On the other hand, identification involves recognizing an individual from a list of N individuals in the database. As shown in Fig.1, a typical biometric system comprises of five main components, namely acquisition, preprocessing, feature extraction, template database, and matching.

A. Biometric Modalities

As shown in Fig. 2, different biometric modalities have been used in several applications based on its suitability and user convenience. Fingerprint, face and iris are the most widely used biometric modalities. Other biometric modalities that have gained sufficient popularity includes gait, hand geometry, palm-prints, and voice. In general, there is no single biometric modality that is best for all applications. Moreover, in some applications, more than one biometrics is used to attain higher security and to address failure to enroll. Such systems are called multimodal biometric systems. Several factors have to be considered while designing any biometric system such as location, security risks, identification or verification task, expected number of users and their characteristics. This section provides the review of some of the most widely used and accepted biometric modalities.

1) Fingerprints: Uniqueness of a fingerprint is determined by the patterns made by the ridges and valleys. A ridge is a series of dark lines that represents the high, peaking portion of the skin. On the other hand, valleys are the low, shallow portion of the skin that appear between the ridges. Fingerprint identification algorithms are generally categorized into minutia based algorithms and ridge flow based algorithms. Minutia-based algorithms first locate minutia points and then map their relative placement on the finger. Ridge flow based algorithms focus on the location and direction of the ridge endings and bifurcations along the ridge path.

2) Iris: Iris recognition [1] is the process of recognizing a person by analyzing the random pattern of an iris. The first stage of iris recognition is segmentation, *i.e.* to isolate the actual iris region in a digital eye image. Once the iris is successfully segmented, normalization is performed to transform the iris region so that it has fixed dimensions in order to

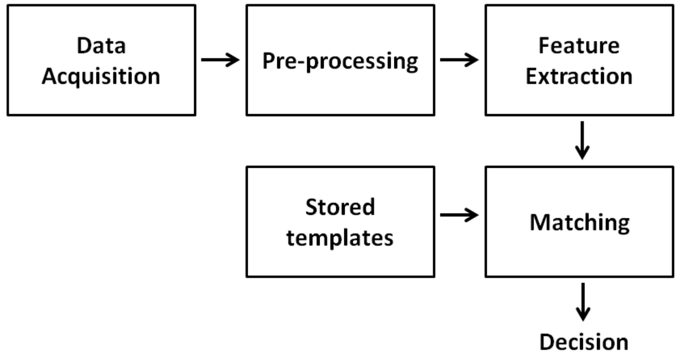


Fig. 1. Block diagram illustrating different modules in a biometric system.

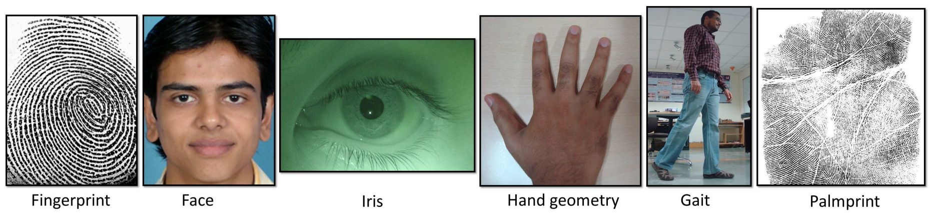


Fig. 2. Biometric modalities that are widely used in several applications.

allow efficient comparison among different iris images. The next stage is the feature encoding and matching stage. The most discriminating information present in an iris pattern is extracted using band pass decomposition of the iris image to get a compact and discriminating representation of the iris patterns. Finally, hamming distance gives a measure of how many bits are same between any two bit patterns.

3) Face: Face as a biometric modality has an inherent advantage of being non-invasive and can be easily captured from a distance without much co-operation from the user. As shown in Fig 3, a face recognition algorithm [2] starts by detecting the facial regions in an image, face detection. Once the faces are detected, feature extraction is performed to obtain a template that captures the discriminating information from the face image. After face detection and feature extraction, matching algorithms are used to match the template obtained from probe with the stored templates in database.

4) Other Biometric Modalities: The need for security in every day life is continuously increasing and the need for different biometric modalities and authentication approaches is also increasing. Commonly used (i.e. face, fingerprints and iris) biometric modalities are not applicable in every possible application. Therefore, several other biometric modalities have also been proposed such as gait, hand geometry, palm prints, hand veins, hand knuckle, periocular, ear, key-strokes, signature, and voice.

B. Contributions of the Report

This report provides a review of several techniques proposed for face recognition to address different challenges caused due to variations in pose, illumination, expression, aging, and disguise. The report also summarizes existing surveys on face recognition and their contributions in providing different outlook to the research in face recognition. This report attempts to provide a new ideology in face recognition by categorizing several face recognition algorithms based on covariates of face recognition. Further, covariates of face recognition are classified as existing and emerging covariates based on how well a covariate has been documented and studied in literature. The report also states the limitations of existing techniques, and possible future directions in each covariate of face recognition. Finally, the report presents an evolutionary granular approach for matching face images alerted due to plastic surgery, one of the emerging covariate of face recognition.

II. Face Recognition

Face, as a biometric, has the benefit of being non-intrusive and passive as compared to other biometric modalities (such as fingerprints and iris). Many applications require face images to be captured outdoors where the lighting conditions are unpredictable, the subjects may not be cooperative, the poses may vary, or the angles and distances from the camera may not be controlled. A robust face recognition system should be able to identify a face captured in an uncontrolled environment. Performance of current face recognition systems significantly deteriorates for such imperfect and challenging cases. Such generality in the application of face recognition systems has brought several covariates such as pose, illumination, expression, aging, and disguise. The effect of these challenges is compounded when two or more covariates are present in a single image. Several approaches have also been proposed to address these challenges. It is observed that these covariates often result in variations between the images of same individual (intra-class variations) to be larger than variations in the images from different individuals (inter-class variations).

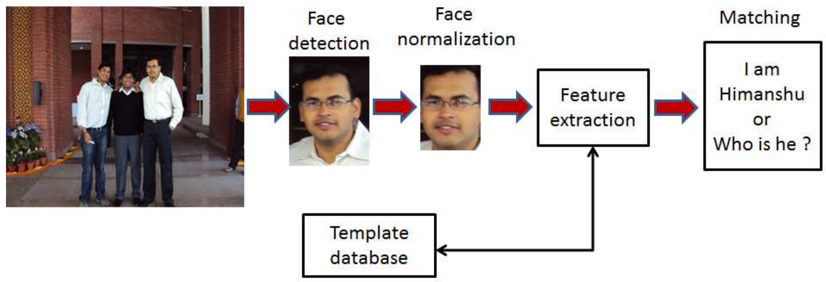


Fig. 3. Illustrating different stages in a face recognition system.

Face recognition has received a lot of attention from both academicians as well as from the industry because of its ever increasing applications in surveillance, access control, law enforcement, cross border security, multimedia applications, forensics, and many more. With advancement in technology and reduction in cost of cameras (sensors), new applications of face recognition has now become prevalent. Verification based on face images captured through built-in cameras is used to allow access to personal devices such as laptops and mobile phones. With development in face recognition technology, it is now used for cross border security. Hong Kong-SAR border has the world's first drive-thru face recognition system. Smart-Gate at Australia, US visit, and Japan visit programs also collect face for all visitor. Face recognition is also used in kiosk applications to allow access to ATM machines, server rooms, and e-commerce applications (online banking). Face recognition has found applications in large social welfare programs where a new user is matched against all existing users to check for duplicates. Currently, two states in United States (Massachusetts and Connecticut) use face identification for large scale duplication check. In India, UIDAI is also collecting face biometric (along with other biometric modalities) to issue a unique identification number to all the citizens. Face, being a non-invasive biometric, is widely used for surveillance. In surveillance applications, face images are captured without active co-operation from the user and are matched to a watch-list database of individuals. Surveillance cameras now have a profound presence at public places like airports, railway stations, shopping malls, and banks.

A. First Step in Face Recognition: Detection

Face detection is the first stage in an automated face recognition system. Given a single image, the goal of face detection is to identify all image regions which contain a face regardless of its three-dimensional position, orientation, and lighting conditions. Face detection from a single image is a challenging task because of the variability in scale, location, orientation, pose, facial expression, occlusion, and lighting conditions. Yang *et al.* [3] categorized the techniques for face detection into four classes, namely, knowledge-based, feature invariant, template matching, and appearance-based methods. Knowledge-based methods comprise a set of rules that encode human knowledge of what constitutes a face and generally consists of relationships between facial features. Feature invariant methods aim to find structural features that exist even when the pose, viewpoint, and lighting conditions vary, and then use these to locate faces. In template matching methods, standard patterns of a face are stored and the correlations between an input image and the stored patterns are used for detection. In appearance-based methods, models (or templates) learnt from a set of training images that capture the representative variability of facial appearance are used for face detection. In literature, there has been two widely used face detectors: Rowley [4] face detector and Adaboost proposed by Voila and Jones [5]. Face detector proposed by Rowley is a neural network based architecture for detecting faces. Rowley face detector is fast and efficient than previous feature and appearance based approaches but, it is computationally expensive. The face detector proposed by Viola and Jones [5] uses Haar-like features and a cascade of boosted decision tree classifiers as a statistical model. Each Haar feature is essentially a scalable template that can be applied to the search window on the image. Detection is accomplished by sliding a search window through the image and checking the response of the classifier to decide whether a certain location looks like a face or not. Viola proposes the building of a cascade of boosted weak classifiers with increasing complexity. This technique greatly increases the detection speed because maximum time is spent on detecting faces while, majority of the non-face regions are rejected in the first few levels of the boosted tree. Zhang and Zhang [6] presented a survey on recent advances in face detection where several techniques are categorized based on the feature extraction and learning algorithms utilized for robust face detection.

B. Feature Extraction and Matching

Face recognition, being a long standing problem, has attracted researchers from different domains such as psychology, pattern recognition, neural networks, computer vision, and computer graphics. It is due to this fact that the literature on face recognition is vast and diverse. In an attempt to categorize the techniques proposed for face recognition, Zhao *et al.* [7] presented a survey about different algorithms, their detailed descriptions, and existing challenges. According to [7], the techniques proposed for face recognition in still face images can be categorized into holistic, feature based, and hybrid approaches. Holistic approaches use the global appearance of the face image and extract features from full face, whereas in features based approaches, local features such as the eyes, nose, and mouth are extracted and their characteristics such as local geometry or appearance are utilized. Based on human perception, hybrid approaches use both local features and the whole face region for recognition. They have identified pose and illumination variations as the two major issues in face recognition. Kong *et al.* [8] divided techniques for face recognition into visible and infrared domain. They presented a review of 2D face recognition techniques in visible spectrum and show that these algorithms can achieve significant performance in controlled settings with cooperative users. However, the performance of these algorithms degrade when face images are captured in uncontrolled environment with large variations in pose, illumination, and expression. Their survey also presents a comprehensive review of algorithms proposed for robust face recognition in infrared imagery. Several approaches such as detecting disguise variations using thermal imagery and multi-spectral fusion for illumination normalization are presented. Face recognition techniques in infrared imagery have shown to improve the overall performance in uncontrolled environments, however, one limitation of infrared sensing methods is their high dependency on the environmental illumination. Belhumeur [9] presents some on going challenges in face recognition such as pose, illumination, and expression and describes several techniques proposed to address these challenges. In [9], techniques proposed for matching face images across these variations are categorized as feature based, appearance based, and 3D face recognition techniques. Feature based methods using geometric relations (e.g. distances and angles) between facial features such as eyes, mouth, nose, and chin are used for efficient face recognition because of their economical representation. However, due to insensitivity to small variations in illumination and viewpoint, feature-based methods are quite sensitive to the feature extraction process. The subspace based methods differ from feature-based techniques in their low-dimensional representation. However, recognition using subspace based methods under a particular lighting condition, pose, and expression can be performed reliably provided the face has been previously seen under similar circumstances. In 3D face recognition, images acquired during enrollment are used to estimate models of the 3D shape of each face. These 3D models can then be used to synthetically render images of each face under arbitrary pose and lighting conditions, effectively increasing the gallery set for each subject. Out of several approaches proposed for face recognition, 3D face recognition have gone a long way towards addressing challenges due to pose, lighting, and expression. Abate *et al.* [10] presents a comprehensive review of techniques for 2D and 3D face recognition.

Klare and Jain [11] proposed the taxonomy of facial features by grouping the salient information available in 2D face images into feature categories: level 1, level 2, and level 3. Level 1 facial features captures the holistic nature of the face such as skin color, gender, and the general appearance of the face (such as PCA and LDA) that can be extracted from low resolution face images. These appearance based features can be efficiently used to differentiate a probe from individuals having very different appearances, but cannot be used to identify a probe from a more similar looking individual in the gallery. These features can be easily computed even from a low resolution face image and used for indexing or reducing the search space. Level 2 features are locally derived and describe structures in the face that are relevant for face recognition. Level 2 features are (such as Gabor wavelets, LBP, and SIFT) are the most discriminative face features and are predominantly used for face recognition. However, there are few applications of face recognition where level 2 features alone are not sufficient for efficient matching. These applications include matching look-alike faces [12], biologically identical twins [13], [14], [15] and faces across different age. Level 3 features contain unstructured micro level features on the face, which includes certain irregularities in the facial skin such as scars and facial marks. It is quite recent that level 3 information has been used in such challenging applications of face recognition where level 2 features alone cannot perform efficient face recognition. Fig. 4 illustrates the three levels of facial features. Another important application of face recognition is in forensics, Jain *et al.* [16] discussed some recent developments in automated face recognition that impact the forensic face recognition community. Ongoing research in facial aging, facial marks, forensic sketch recognition, face recognition in video, near-infrared face recognition, and use of soft biometrics for enhancing forensic face recognition is discussed along with current limitations and future research directions.

C. Covariates of Face Recognition

Existing literature discusses current face recognition techniques from different points of view. One view categorizes face recognition algorithms into holistic, feature-based, and hybrid techniques, another categorizes them as visible and infrared domain techniques; one divides them as 2D and 3D face recognition techniques whereas, the other proposes a grouping based on the salient information into hierarchical feature category. There are few papers that review face recognition across pose variations [17], illumination variations [9], [18], and forensic applications [16]. However, to the best of our knowledge, there is no literature that categorizes different techniques based on the covariates of face recognition. In this report, advances in face recognition are categorized based on the covariates of face recognition. *A covariate is defined as an effect that independently increases the intra-class variability or decreases the inter-class variability or both*. Fig. 5 illustrates the intra-class and inter-class variability between two subjects.

Variations due to pose and illumination hide/alter some of the features that make faces of different subjects appear more similar to each other than their own actual faces, thus decreasing the inter-class separation. On the other hand, variations due to expression, aging, plastic surgery, and disguise increase the difference between the face images of the same subject, affecting the intra-class similarity. Several covariates that affect the intra-class and inter-class variability are identified and studied in face recognition. As shown in Fig. 6, the covariates in face recognition can be categorized as: 1) existing covariates, and 2) emerging covariates.

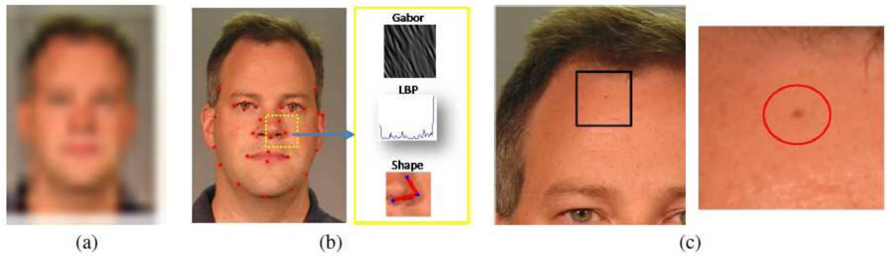


Fig. 4. Examples of three levels of facial features. (a) level 1 features, (b) level 2 features, and (c) level 3 features. Image from [11].

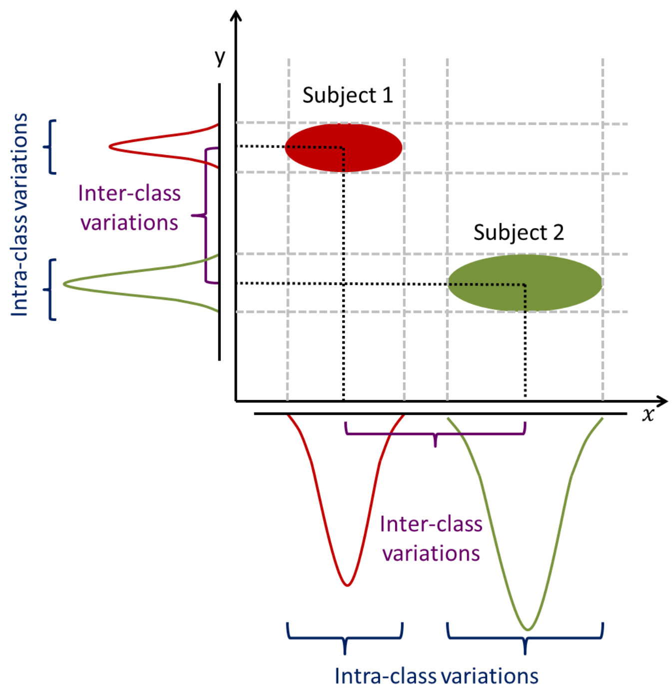


Fig. 5. Illustrating the inter-class and intra-class variations.

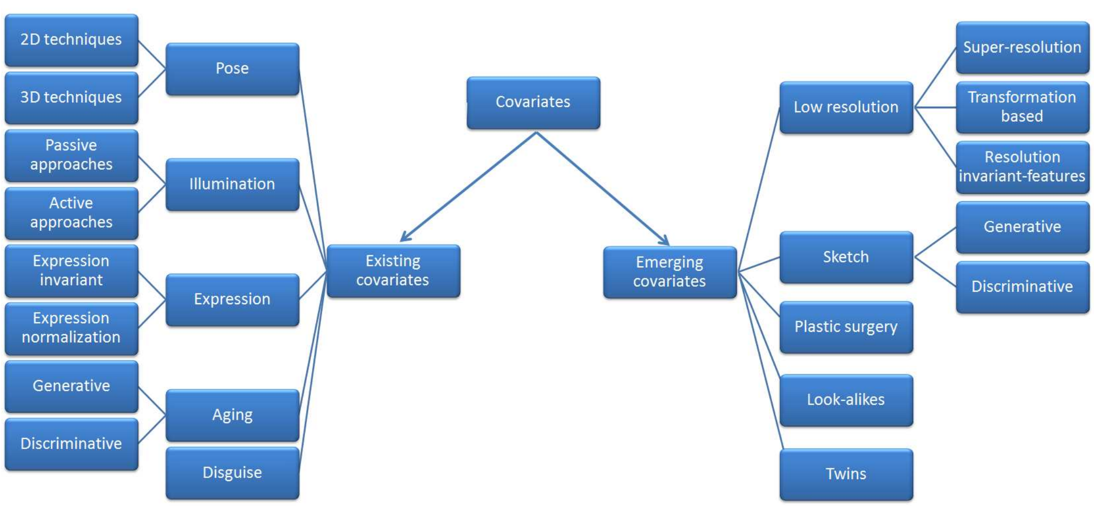


Fig. 6. Illustrating categorization of techniques proposed for different covariates in face recognition.

III. Existing Covariates

As shown in Fig. 7, there are several covariates that have been well established and extensively studied in literature. Covariates such as pose, illumination, and expression significantly affect the facial appearance and extensive research has been dedicated to develop algorithms that can efficiently address these covariates. Apart from these challenges, face recognition with aging and disguise variations have also been studied and several techniques have been proposed to address them. Most of the existing algorithms are designed to address an isolated covariate and their performance in a real world unconstrained environment is still far from satisfactory. To develop an algorithm that can address the problem of face recognition in an unconstrained environment (when two or more covariates are present in the same face image) is still an open research problem. In this section, existing covariates and different approaches proposed to address them are reviewed.

A. Pose

Face recognition across pose has great potential in many applications dealing with uncooperative subjects in an unconstrained environment. Frontal face images have a lot more information than profile or semi-profile face images. In face images with pose, some of the features are not be visible and can reduce the performance of face recognition systems. Face recognition across pose variations has gained huge interest in the research community and a few promising approaches have been proposed for addressing this problem. Fig. 8 shows examples of face images across different pose. As shown in Table I, face recognition techniques for matching face images across pose variations can be broadly classified[1] into two categories: 1) 2D techniques, and 2) 3D techniques.

1) 2D techniques: In 2D imaging, several approaches have been proposed to address pose variation. According to [17], these approaches can be classified into three groups: 1) Pose-tolerant feature extraction, 2) View-based matching, and 3) 2D pose transformation.

Pose tolerant feature extraction builds a classifier or finds a linear or non-linear mapping in image space that is insensitive to pose variations. Several approaches based on extracting features from local facial regions have been utilized for robust face recognition in any arbitrary pose. Local approaches such as Elastic Bunch Graph Matching [20] and Local Binary Patterns [21], [22] have shown good performance across pose variation. Unlike holistic approaches (such as PCA and LDA), local approaches are independent of pixel-wise correspondence between the gallery and probe images. These pixel-wise correspondences are adversely affected by pose variations. In view-based matching approaches, multiple gallery images per person are collected from different viewing angles to cover exhaustive range poses. Generally, it is observed that the tolerance of a face recognition system across pose variations increases with more gallery images per person because it increases the probability that the probe pose lies close to one

[1]The classification structure is adapted from [17] and advances thereafter are combined within the same classification.

TABLE I
Broad classification of approaches proposed for pose invariant face recognition. Adapted from [17].

2D techniques	Pose-tolerant features	EBGM [20]
		LBP [21]
		Eigenfaces [23]
	View-based matching	Mosaicing [24]
	2D pose transformation	Parallel deformation [25]
		AAM [26], [27]
3D techniques	Generic shape methods	Cylindrical face [28], [29]
		Texture synthesis [30]
	Feature based reconstruction	Deformable models [31]
		Jiang's method [32]
	Image based reconstruction	Morphable models [33]
		Stereo matching [34]

of the gallery poses. Singh *et al.* [24] proposed a mosaicing scheme to form a panoramic view from multiple gallery images to cover all possible poses under all horizontal in-depth rotations. Availability of multiple poses of an individual's face is assumed during enrollment. A pair of face images, representing the frontal and profile views of an individual, are mosaiced after aligning them using image registration.

As shown in Fig. 9, registered images are mosaiced using the multi-resolution splines algorithm based on Gaussian and Laplacian pyramids [35]. Multi-resolution splines also perform blending as an integral part of mosaicing thereby offering some inherent advantages. Finally, texture based algorithm is used to match face images in any arbitrary pose to the mosaiced image in the gallery. However, in many real world applications it is computationally expensive and difficult to collect multiple gallery images, e.g. passport photo database or police mugshot database that has one frontal and two profile face images. Synthesizing virtual views to substitute multiple real views using pose transformation techniques is another feasible alternative that is used in literature. Beymer and Poggio [25] proposed parallel deformation to generate virtual views from a single view using feature-based 2-D warping. The virtual views thus generated covers all possible poses. Gonzalez-Jimenez and Alba-Castro [36] proposed an active shape model (ASM) with manual facial component locating to synthesize virtual views in different poses using the point distribution model. In ASM, principal component analysis is applied on the location of facial components (such as facial contours, eyes, eyebrows, and lips) and presented as connected point distribution from a variety of manually labeled images. As an extension to ASM, active appearance models (AAM) [26] simultaneously model the variations in shape represented by point distributions and texture represented by pixel intensities. Vetter [37] further extended the concept of AAM by replacing the point distribution with a pixel-wise correspondence between two images in different poses using optical flow. Kahraman *et al.* [27] synthetically generates images to cover large pose variations and recorded the displacements of all landmarks of the AAM using a reference face. Synthetic images in different pose are then generated from single frontal images by moving the landmarks along the recorded displacements.

Similar to transformation in image space, there are some approaches to transform the feature space for robust face

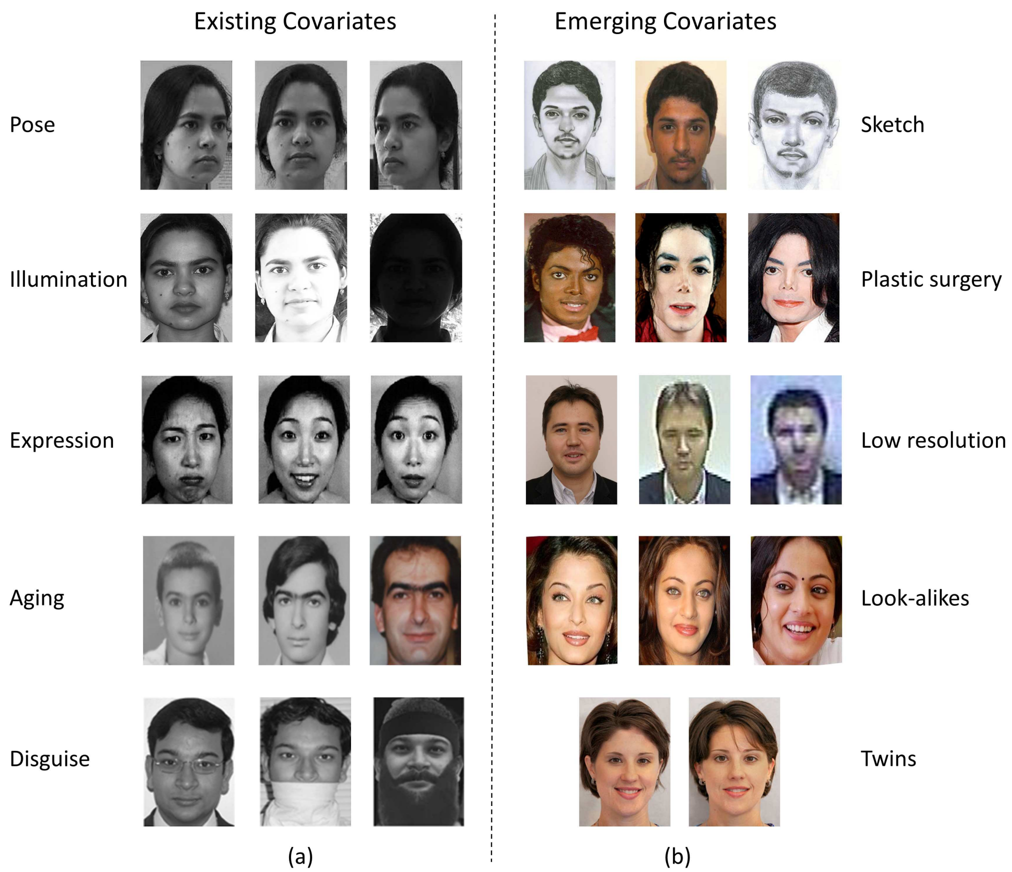


Fig. 7. Covariates in face recognition: (a) existing covariates, and (b) emerging covariates.

Fig. 8. Images with pose variations from the CMU-PIE [19] database.

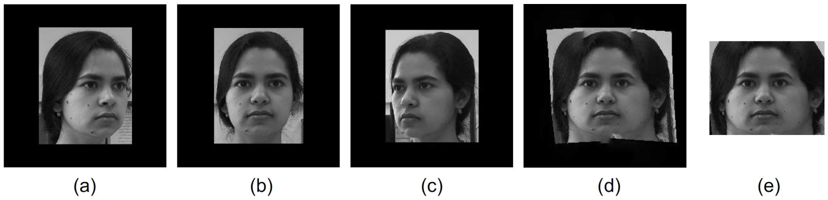


Fig. 9. (a), (b), (c) Frontal and profile input images. (d) Mosaiced face generated using (a), (b) and (c). (e) Final mosaiced and cropped face image. Image from [24].

recognition across pose variations. Pose transformations in feature space tend to implicitly improve linear separability of face images under pose variations by non-linear mapping prior to recognition. Kim and Kittler [38] proposed a hybrid approach fusing four different systems to tolerate pose variations. The first two systems belong to linear pose-tolerant feature extraction, the third system belongs to non-linear pose-tolerant feature extraction, and the fourth system is 2D transformation approach using a $3D$ generic face model.

2) $3D$ techniques: Large pose variations bring image discontinuities that cannot be addressed within a 2D space. $3D$ approaches approximate the image variations caused by pose variations in the 3D space rather than limiting them within the image plane. 3D approaches for pose invariant face recognition can be studied as :1) generic shape-based methods, 2) feature based 3D reconstructions, and 3) image-based 3D reconstruction.

Generic shape models are the simplest and quite efficient methods based on the cylindrical face shape. Arbitrary face images are mapped onto a generic cylindrical face shape and frontal views are reconstructed [28], [29]. Liu and Chen [28] proposed a probabilistic geometry assisted (PGA) face recognition algorithm to address pose variations. PGA approximates human heads as an ellipsoid whose radius, locations, and orientations are estimated based on a universal mosaic model. Instead of using simple geometries as generic 3D face models, Zhang *et al.* [30] proposed automatic texture synthesis to reconstruct rotated face views from a single frontal view for recognition using generic shape models.

Feature-based 3D face reconstructions [39], [40] estimate personalized face shapes from the 2D locations of facial features (facial components such as eyes, nose, etc. and image features such as edges or corners). Prabhu *et al.* [41] constructed a 3D model for each subject in the database from a single 2D image. For matching, probe pose is estimated based on automatic facial landmark annotations and then the 3D model is used to synthesize corresponding 2D pose view of gallery. Lee and Ranganath [31] presented a composite 3D deformable face model for pose estimation and synthesis based on template deformation which maintained connectedness and smoothness. Using five images of the same person with different poses, a complete 3D face model of the person can be generated. Matching is performed by comparing the synthesized image with the probe pixel-wisely in Euclidean distance. For feature based 3D reconstructions, Jiang *et al.* [32] used facial features to efficiently reconstruct personalized 3D face models from a single frontal face image. Their method is based on automatic detection of facial features on the frontal views using Bayesian shape localization. A set of 100 3D face scans is used as prior knowledge of human faces.

Image-based reconstructions depend on pixel-wise appearance of face images and are generally capable of generating more detailed face structures. Blanz and Vetter [33] utilize the relation between the image pixel intensities and its corresponding shape/texture properties for image based 3D reconstructions. Passalis *et al.* [42] proposed a face recognition method using facial symmetry. An annotated face model is constructed from estimated pose and detected landmarks. The model is fitted to input $3D$ scans which results in a pose invariant geometry image. Stereo vision techniques [34] are also used to reconstruct 3D face models from two face images in different poses. Feature-based reconstructions are accurate near facial features and could be inaccurate in other non-feature regions because they are usually interpolated from adjacent facial features. Unlike image-based reconstructions, feature-based reconstructions also suffer from the inaccuracy of feature detections. However, image-based 3D reconstruction are sensitive and vulnerable to image variations, such as shadows and spatial misalignment because it considers pixel-wise reflection mechanisms in estimating shape and texture information.

A number of promising techniques have been proposed to tolerate and compensate image variations brought by pose changes and Table II reports the results on different databases across pose variations. However, achieving pose invariance in face recognition remains an unsolved challenge. In 2D techniques, prior knowledge of human faces plays an important role in handling pose variations in face recognition [26], [27], and [25]. But, inclusion of this prior knowledge requires extensive training and the performance is dependent on training data. View based techniques [24], [25] require multiple gallery images in different pose, however, it is pragmatic in many applications. 3D approaches looks more promising to achieve better performance in face recognition across pose. However, 3D modeling has its own limitations in terms of computation as it needs to estimate a large number of unknown parameters [33] and [32].

B. Illumination

Images with proper and uniform illumination are ideal for face recognition applications. Since illumination variations may cause some features to be hidden or altered, images with illumination variations reduce the performance of recognition algorithms. The situation becomes more challenging when illumination variations are combined with other covariates such as pose and expression. According to [18] , algorithms proposed for matching faces with illumination variations can be broadly classified[2] as passive or active approaches as shown in Table III.

1) Passive approaches: The problem of illumination variation is addressed in visible spectrum and can be broadly classified as 1) modeling based, 2) illumination invariant feature extraction, 3) photometric normalization, and 4) morphable models.

First approach is to model face images under varying illuminations based on statistical models. In statistical modeling techniques, such as PCA and LDA, training face images captured under different illumination conditions are used to obtain a subspace. The subspace thus obtained covers all possible variations due to illumination and can be used to efficiently match an image with any arbitrary illumination. Hallinan [43] used five eigenfaces of an individual to represent

[2]The classification structure is adapted from [18] and advances thereafter are combined within the same classification.

TABLE II
A COMPARISON OF POSE INVARIANT FACE RECOGNITION METHODS.

Algorithm	Database	Pose variation	Gallery/ Probe	Accuracy
LBP [21]	CMU-PIE	13 poses within $\pm 66°$ in yaw and $\pm 15°$ in tilt	2(0°, 66°)/ 11 remaining	74.2%
Mosaicing [24]	CMU-PIE	5 poses:0°, $\pm 30°$, $\pm 60°$ in yaw	3(0°, $\pm 30°$)/ 5 (0°, $\pm 30°$, $\pm 60°$)	96.8%
PDM [36]	CMU-PIE	5 poses:0°, $\pm 15°$, $\pm 30°$ in yaw	1(0°)/ 4 remaining	97.4%
Cylindrical 3D [28]	Bern university	5 poses:0°, $\pm 20°$ in yaw & tilt	1(0°)/ 4 remaining	80.0%

TABLE III
BROAD CLASSIFICATION OF APPROACHES PROPOSED FOR ILLUMINATION INVARIANT FACE RECOGNITION.

Passive approaches	Illumination variational modeling	Linear subspace [43]
		Illumination cone [44]
	Illumination invariant features	Features derived from image derivatives [45], [46]
		Quotient image [47], [48]
		Transformation domain features [49], [50]
	Photometric normalization	Histogram equalization [51]
		Homomorphic filtering [51]
		Local illumination normalization [52]
	3D morphable model	PCA analysis of the shape and texture [33]
Active approaches	3D information	Fusion of 3D and 2D information [53]
	Infrared	Thermal imaging [54], [8], [55]
		Active Near-IR [56]
		Active differential imaging [57]

face images under a wide range of possible illumination. Similarly, Belhumeur [23] used three images per subject under different illumination to construct basis for a linear subspace. Belhumeur and Kriegman [44] also proposed that images of a convex lambertian object from the same viewpoint but illuminated by an arbitrary number of distant point sources form a convex illumination cone. The dimensions of the cone is same as the number of distinct surface normals. Using the illumination cone, their approach can match images in any arbitrary pose and illumination. In another approach known as spherical harmonic representation, a set of images from a convex lambertian object obtained under a wide variety of lighting conditions can be approximated by a low-dimensional linear subspace. Zhang and Samaras [58] used spherical harmonics representation for face recognition under arbitrary unknown lighting. In one of their methods, basis images for a face is estimated based on maximum a posterior estimation. In another method, they combined 3D morphable model with harmonic representation to perform face recognition with both illumination and pose variations. These modeling techniques largely depends on the training data and perform well under seen illumination conditions.

Second approach in passive methodology is to extract illumination invariant features for recognition such as edge map and image intensity derivatives. These features are quite stable (or less affected) across different illumination conditions. Gao and Leung [45] proposed line edge map representation for face images and used Hausdorff distance to measure the similarity between two faces. Surface geometry and reflectance are the intrinsic properties of face and the probability distribution of image gradient is a function of these attributes. Therefore, the direction of image gradient [59] is insensitive to illumination changes and can be used for face recognition across varying illumination conditions. Human faces follow a general shape, however, each face differs in surface albedo (reflecting power of a surface). The Quotient image, Q_y, [47] of a human face is independent of illumination changes and depends only on the relative surface texture information which makes it useful for recognition. Chen and Chen [48] proposed an illumination invariant face recognition by constructing an illumination subspace using quotient images. The relative phase information is robust against large illumination variations. Therefore, methods based on frequency domain representation for illumination invariant face recognition have also gained attention in the research community. Savvides *et al.* [49] performed PCA in the phase domain and showed that their approach is robust to wide range of illumination variations present in the CMU-PIE database [19]. Heo *et al.* [50] also achieved significantly better results as compared to the eigenphase approach proposed in [49] by applying Support Vector Machines on phase features for illumination invariant face recognition. Extracting illumination invariant representation is efficient in terms of computational time and can also be used to represent unseen illumination variations.

Third approach involves photometric normalization [51] such as histogram equalization, gamma correction and ratio images to normalize and compensate illumination variations in face images. In homomorphic filtering [51], a reflectance model is used to separate the reflectance and luminance, $I(x,y) = R(x,y) \times L(x,y)$, where $I(x,y)$ is the intensity of the image, $R(x,y)$ is the reflectance function, and $L(x,y)$ is the luminance function. Xie and Lam [52] proposed an illumination normalization method called Local Normalization. They split the face region into a set of triangular regions and the intensity values within each facial region are normalized to zero mean and unit variance. Recently, Wang *et al.* [60] proposed an illumination normalization approach based on Weber's law. They represent face images in an illumination insensitive representation, Weber faces. The Weber face is calculated by computing differential excitation of every pixel in the face image as shown in Eq. 1.

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

where x_c is the intensity value at central pixel, P is the number of neighbors and α is the parameter for adjusting the intensity difference between neighboring pixels. Fig. 10 represents a face under different illuminations and the corresponding Weber-faces which gives good illumination normalization results. Recognition using Weber's face yields an average accuracy of 94.7% on the CMU-PIE [19] database.

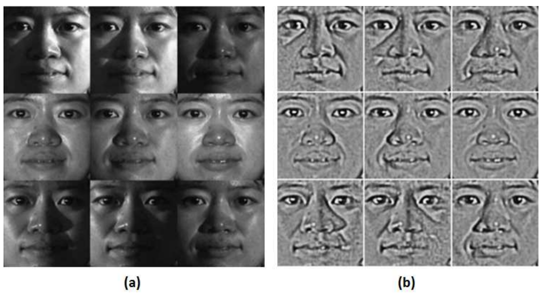


Fig. 10. Showing (a) illumination variations in face images from the CMU-PIE database [19], and (b) corresponding Weber's face normalizing the illumination effect. Image from [61].

Finally in 3D morphable models [33], shape and texture of face images are described separately based on the PCA analysis obtained from a database of 3D scans. To fit a face image under arbitrary pose and illumination to the model, shape coefficients, texture coefficients along with other rendering parameters are optimized to minimize the difference between the input image and the rendered image. The recognition is performed based on the model coefficients for shape and texture. These morphable model are the most accurate among all the passive approaches because they accurately model any arbitrary illumination using the shape and texture information. However, a large training data in different illumination conditions is required during training. These models are also computationally expensive because it requires optimizing shape and texture coefficients corresponding to each gallery image.

2) Active approaches: Illumination has a drastic effect on appearance of a face image in visible spectrum as it causes some features to be hidden. Capturing face images with additional optical sensors to capture images that are insensitive or independent of illumination changes has also been used in face recognition. For example, 3D face, thermal face, and near infrared face. 3D structure of the face can be used for recognition across illumination variations. The texture information in 3D images can be fused with 2D images [62] to achieve better performance across illumination changes in uncontrolled environment. However, these approaches are susceptible to the effect of illumination on 2D images. Face recognition using thermal images [54] is proposed to addresses the challenges pertaining to visible spectrum images. Thermal images are related to thermal radiations from an object. These radiations depend on the temperature and emissivity of the object. For scenarios with large illumination changes and facial expressions, thermal face recognition outperforms face recognition algorithms in visible spectrum [54], [63]. Several schemes for fusing thermal images with visible spectrum images for illumination normalization [8], [55], [64], [65] and [66] have shown to achieve better performance than recognition based on either modality. Li *et al.* [56] proposed a face recognition system based on active near infrared (NIR) lighting to capture good quality face images independent of environmental illumination. To obtain an illumination invariant face representation, local binary patterns (LBP) are used. NIR images are used for robust face recognition in indoor environment. For outdoor environment, the infrared energy is very strong and near infrared imaging may not help, therefore, active differential imaging [67] is used for the outdoor application. In active differential imaging, an image is obtained by the pixel-to-pixel differentiation between two successive frames, one with the active illuminant on and the other with the illuminant off. Face recognition based on active NIR differential imaging system achieved lower error rates even in the scenario with drastic illumination changes as reported in [57].

Several techniques proposed for illumination invariant face recognition have fairly good performance as reported in Table IV, however, each technique has its own drawbacks. The illumination modeling methods [43] require multiple training samples with different illumination conditions. Their performance degrades for unseen illumination variations. The performance of many photometric normalization [52], [51] methods is sensitive to the choice of parameters and may drastically vary across databases because of different capture conditions. However, normalization methods are efficient and does not require extensive training. In general, the performance of active approaches is better than the performance of passive approaches, but, active sensing methods dependent on the environmental illumination. NIR based face recognition systems [56] also encounters the problem of specular reflections of active NIR lights on eye/ eyeglasses which is a critical issue in active NIR image based face detection. As suggested in [18], an interesting future research direction can be combining active approaches and passive approaches for illumination invariant face recognition.

C. Expression

Variations in expression can cause deformation in local facial structure and also change the facial appearance and local geometry of the face as shown in Fig. 11. Facial expressions effect the performance of a face recognition system and to address these challenges, existing methods rely on extracting stable face features such as extracted line segment and geometric invariants [68]. It is observed that geodesic distances [69] on the facial surface are significantly less sensitive to facial expressions compared to Euclidean ones. Bronstein *et al.* [69] proposed an expression invariant signature based on the geodesic distance on the facial surface. Nagesh and Li [70] observed that changes due to variation in expressions are sparse with respect to the whole image. Using distributed compressive sensing theory, they represented the training images of a given subject by only two feature images: one that captures the global features of the face, and the other that captures the different expressions in all training samples.

TABLE IV
A COMPARISON OF ILLUMINATION INVARIANT FACE RECOGNITION METHODS. FAR REPRESENTS FALSE ACCEPT RATE.

Algorithm	Database (# gallery/ # probe)	Accuracy
LBP+AdaBoost [56] NIR Images	CASIA NIR (1000 /3,237) (training/ testing)	91.8% at 0.1% FAR
Fusion of visible and and thermal images [66]	Notredame (159,1815)	95.8% at 0.01% FAR
	Equinox (95/18715)	94.9% at 0.01% FAR
Weber's face [60]	CMU-PIE (68/1357)	94.7% (rank-1 accuracy)

Wiskott and Malsburg [71] proposed a dynamic link matching which is robust under face rotation and deformation. Martinez [72] proposed an expression invariant face recognition system based on a probabilistic approach. Their approach independently assigns more weights to those local areas which are less sensitive to expressional changes. The effect of expression on different local facial regions for each specific expression is learned during training. Amberg *et al.* [73] proposed an expression invariant method for face recognition by fitting an expression separated 3D Morphable Model to shape data. Compared to a single face image, a video sequence consists of a large number of successive images and temporal information that boosts the recognition task. Recently, video based face recognition systems have shown to address a wider range of facial expressions [74].

Variations caused due to expressions change local facial regions anywhere and in any size or shape. Tan *et al.* [75] proposed a non-metric partial similarity measure, inspired by human perception, to capture the prominent partial similarity. Tsai and Jan [76] used a subspace based approach to develop a face recognition system that is insensitive to local facial deformations. However, like any other subspace based approach, their method also requires multiple training images in each class and with different expression variations. Morphing probe images to neutral the shape of training has also used for expression invariant face recognition. Active appearance model (AAM) [26] separates the shape and texture information that can be used for warping. Using this information, Ramachandran *et al.* [78] processed the image to convert a smiling face to a neutral face. Facial expressions change the facial geometry, therefore, affecting the performance of face recognition systems. Li *et al.* [79] applied a face mask for face geometry normalization and further calculated the eigenspace for geometry and texture separately.

Fig. 11. The left most image is the neutral face. The others are the face images with angry, disgust, fear, happy, sad, and surprise expressions in columns from left to right with increasing levels in rows from top to bottom. Images from BU-3FEDB [77].

Another approach is to use optical flow, Yacoob *et al.* [80] recognized facial expressions from image sequence using the facial dynamics. Their algorithm utilizes optical flow to identify the direction of rigid and nonrigid motions caused due to human facial expressions. However, it is difficult to learn the local motions within feature space to determine the expression changes of each face because different individuals express in different ways. Martinez [81] proposed a weighting method that independently weighs the local areas which are less sensitive to expressional changes. Optical flow between the gallery and probe image is used and pixels with small optical flow have high weights while pixels with high optical flow have lower weights. Recently, Hseih *et al.* [61] proposed an expression normalization technique using optical flow based on pixel deformations and intensity variations. Fig. 12 shows examples of expression normalization using optical flow. Their expression normalization technique is computationally efficient as the optical flow for each input face is computed only once with reference to a given standard neutral face. In a recent approach [82], they used optical flow to transform the neutral images in the database to the exact expression of probe using warping.

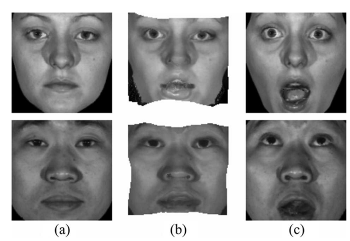


Fig. 12. Illustrates expression normalization using constrained optical flow: (a) original neutral faces, (b) expression normalized faces, and (c) the face images with surprise expression. Images from [61]

Table V reports the performance of algorithms proposed for expression invariant face recognition. The best algorithm for expression invariant face recognition gives about 98% rank1 identification accuracy, however, there are few limitations to the existing solutions. The warping approach [79] cannot warp all images to a neutral image because of the lack of texture in certain regions, like closed eyes. Also the linear warping is not consistent to the nonlinear characteristics of facial expression movements. Therefore, a non-linear warping approach that can efficiently map the expression movements can result in better invariance to expression. Recently, optical flow based approaches [61] have shown promising results for expression invariance. The performance of these algorithms is highly dependent on number of training images, but in a practical scenario, training database contains only neutral face image

TABLE V
A COMPARISON OF EXPRESSION INVARIANT FACE RECOGNITION METHODS. EER REPRESENTS EQUAL ERROR RATE.

Algorithm	Database (# gallery/ # probe)	Accuracy
Optical Flow-based Normalization [61]	BU-3FEDB (100/2400)	12.3% (EER)
2D image warping [79]	Yale (25/ 25)	96% (rank-1 accuracy)
Morphable model [73]	GavabDB (1281/ 1708)	98.1% at 0.1 FAR
Deformation modeling [83]	BU-3FEDB (100/800)	13.3% (EER)

per subject.

D. Aging

Facial aging in human faces is a complex process and it brings major structural and texture (skin wrinkles and other artifacts) variations in human face. Fig. 13 shows the effect of age variations in an individual. These variations cause major difference in both the geometry of facial features and the texture of skin which are difficult to model for any face recognition system. According to [84], [85] face recognition approaches proposed to address age variations can be categorized as generative or discriminative as shown in Table VI.

1) Generative approaches: It simulates aging and then apply subsequent recognition algorithms for matching. Burt and Perrett [87] considered face images from seven age groups and identified a set of fiducial features from every face to characterize the facial shape. They proposed using composite face for each age group obtained by averaging the shape and skin color from faces belonging to different age groups. Craniofacial growth being the primary source of facial aging, Ramanathan and Chellappa [88] proposed a facial growth model that adapts to the craniofacial transformations and fused age based anthropometric measurements in predicting appearance across ages. Next, Ramanathan and Chellappa [89] proposed a two-step approach for modeling shape and texture variations separately. The shape based model learns the elastic nature of face while the texture based model characterizes facial wrinkles in different regions. Lanitis *et al.* [90], [91] devised a combined shape and intensity model to represent face images in terms of principal components from the eigen space that correspond to facial shape and intensity. They built functions to describe the relationship between facial age and the AAM parameters. These functions can estimate the age from a child image and predict face growth. Geng *et al.* [92] observed that similar faces age in similar ways and based on this observation, they proposed to learn a subspace of aging patterns. Similar to AAM, face representation is composed of face texture and shape. To compensate for age variations in face recognition, Park *et al.* [84] also proposed an age modeling technique that generates 3D face models for a given 2D face database.

2) Discriminative approaches: It concentrates on deriving age-invariant representation from faces. Age estimation based approaches use anthropometric distances extracted from different facial regions [93]. Kwon and Lobo [93] proposed an age classification approach that identifies the age group based on a face image using the anthropometry of the face and the density of wrinkles. It is well established in literature that the lower and upper halves of faces grow at different rates [94]. Based on this observation and ratio of facial measurements, face images can be divided among different age groups. Ramanathan and Chellappa [97] created intra-personal subspace based on the age separation between a pair of images. They used a Bayesian framework to classify face images into different age categories and further extended it to perform face verification across age progression. Facial feature drifts observed in face images across different ages of the same individual follow a coherent drift pattern, the same is not true in face images of two different individuals across different ages. Based on this observation, Biswas *et al.* [95] used coherency in facial feature drifts as a measure to perform face verification across ages. Guo *et al.* [96] adopted a manifold learning approach and used support vector regression to estimate the age from low-dimensional representation of faces. Ling *et al.* [99] represented face images in a hierarchical manner using the gradient orientation pyramid. Their approach for effective representation combined with support vector machines demonstrates good results for face verification. Recently, Li *et al.* [85] proposed using densely sampled local feature descriptors to capture the age invariant and discriminative information such as edge direction in face images. Further, multi-feature discriminant analysis (MFDA) is used to process the two local feature space and obtain an age invariant representation for the face image as shown in Fig. 14. 47.5% rank-1 identification performance is reported on the FG-Net aging database using the MFDA-based approach for recognizing faces across age variations.

Face recognition across different age groups is a very challenging problem and the techniques proposed to address these challenges are efficient only for age variations in adults up to $\pm$ 10 years. Facial growth [88] depends on various factors such as gender, ethnicity, and age group. Moreover, facial features grow at different rates during different ages. During adolescence, there is a drastic change in the facial features and texture. The facial growth during $20-45$ years is subtle, however after the age of 45 years, the effects of aging starts appearing in terms of face wrinkles. Table VII reports the performance of different techniques proposed for age invariant face recognition. Human face being a 3D object, further development in 3D facial aging models such as [91], [92], [84] can provide possible solutions for efficient face recognition across age variations.

E. Disguise

Disguise is one of the covariates that has received relatively less attention from the research community and only a few

Fig. 13. Illustrating the effect of structural and textual changes with aging in face images of a person from childhood to old age. Images are taken from the FG-NET database [86].

TABLE VI
BROAD CLASSIFICATION OF APPROACHES PROPOSED FOR AGE INVARIANT FACE RECOGNITION.

Generative approaches	Linear subspace [90], [91]
	Active appearance models [26]
	Composite face [87]
	Craniofacial growth models [88]
	Model-based approaches [89], [90], [91], [92], [84]
Discriminative approaches	Anthropometry [93], [94]
	Feature coherency [95]
	Manifold-learning [96]
	Discriminant analysis [85]

TABLE VII
A COMPARISON OF AGE INVARIANT FACE RECOGNITION METHODS. ADAPTED FROM [85].

Authors	Database (# subjects, # images)	Rank-1 accuracy
Lanitis *et al.* [91]	Private database (12,85)	68.5%
Ramanathan *et al.* [97]	Private database (109,109)	15.0%
Wang *et al.* [98]	Private database (NA,2000)	63.0%
Geng *et al.* [92]	FG-NET (10,10)	38.1%
Park *et al.* [84]	FG-NET (82,82)	37.4%
	MORPH Album 1 (612,612)	66.4%
	MORPH Album 2 (10000,20000)	79.8%
Li *et al.* [85]	FG-NET (82,82)	47.5%
	MORPH Album 2 (10000,20000)	83.9%

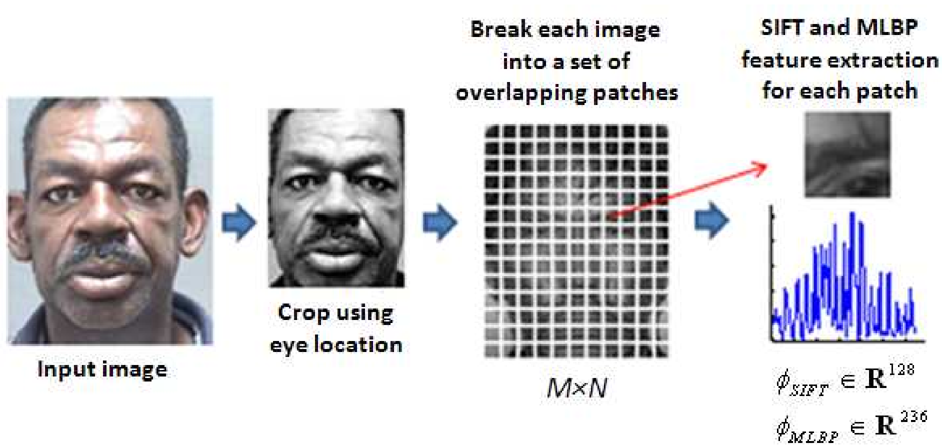


Fig. 14. Illustrating the process of extracting age invariant representation of face images using MFDA [85].

solutions have been proposed to address face recognition with disguise. Disguise can be intentional or unintentional. An individual can knowingly use some accessories (beard moustaches, hair etc.) and makeup to hide the identity in order to escape from law enforcement or to intrude in some unauthorized premises. As shown in Fig. 15, disguise may drastically increase the intra-class variations and degrade the performance of a face recognition system. A few approaches [100], [101] have been proposed to address the challenge of recognizing faces with disguise. Ramanathan *et al.* [100] studied the facial similarity for several variations including disguise by forming two eigen spaces from two halves of the face, one using the left half and other using the right half. Retrieval of images from the database is more efficient when an optimally illuminated half face is used rather than the full face. Silva and Rosa [101] proposed using Eigen-eyes for face recognition across different covariates including disguise. This approach is more robust than eigenface based approach as the variations in different facial regions did not effect the performance of Eigen-eyes. Pamudurthy *et al.* [102] used dynamic features obtained from the skin correlation and matched them using nearest neighbor classifier. Singh *et al.* [103] proposed a face recognition algorithm based on the phase features extracted from a face image using 2D log polar Gabor wavelet. They used a dynamic neural network architecture to extract discriminating 2D Log Gabor transform coefficients from the face image. Gabor phase features were further divided into frames before matching using the hamming distance. Experiments performed on real face disguise database showing 18 variations of an individual and synthetic face disguise database showing 40 variations of the same face show the efficacy of their algorithm for matching face images with disguise even with limited samples per person in the gallery.

Fig. 15. Illustrates the change in appearance of an individual due to disguise using accessories such as beard, moustaches, caps, and glasses.

Deceiving a face recognition system using facial disguise is one of the prevalent tactics and several techniques to match face images with disguise have been proposed. Table VIII reports the performance of face recognition algorithms for matching face images with disguise. However, matching face images with disguise have its own limitations. First being the absence of an extensive facial disguise database where these algorithms can be trained and tested. Another limitation of these algorithms [101] is that the performance degrades when important regions such as eye and mouth are covered using accessories, and some of these algorithms require multiple gallery images. Developing disguise database with different variations and designing face recognition algorithms that can perform well with a few gallery images are the major challenges in face recognition with disguise.

IV. Emerging Covariates

Covariates such as matching forensic sketches with digital face images, matching faces altered due to plastic surgery, matching low resolution faces and matching look-alikes or twins are some of the challenges that have emerged as new covariates in face recognition. After extensive research in face recognition, research community identified the necessity to define some new covariates that need to be addressed in face recognition. These covariates have very important law enforcement applications; however, research in this direction is limited because these covariates have been realized in recent times. With ever increasing applications of face recognition, the need to address these covariates has grown and it has now become imperative for current face recognition systems to be robust to these challenges. Recently, some databases have also been prepared by researchers to study these covariates in greater detail as shown in Table IX. The research in these covariates has just instigated and some preliminary research directions have been set to address these covariates, however, there is no technique that can address these covariates up to a satisfactory level.

A. Low Resolution Images

An important covariate of face recognition that has gained much attention in recent years is matching low resolution (surveillance quality) face images with high resolution gallery images. Surveillance systems installed at public places, airport gates, security checkpoints, and government buildings are primarily designed to cover a large area from a single location. For the sake of wider coverage, low storage and operational costs, the resolution of the images is reduced.

Most of the approaches for low resolution face recognition use super resolution to enhance the low quality probe image before recognition [125], [126]. Huang and He [127] proposed a super-resolution method that uses nonlinear mappings to infer coherent features that favor higher recognition of a single low resolution (LR) face image. They proposed to build a coherent subspace between the PCA features of high resolution (HR) and low resolution (LR) images mapped using radial basis functions. Baker and Kanade [128] proposed an algorithm to a priori learn the spatial distribution of image gradients for frontal facial images. Their approach attempts to recognize local features in low resolution images and then enhances their resolution. Chakrabarty *et al.* [129] proposed a learning based method to super-resolve face images with kernel principal component analysis-based prior model. Chang *et al.* [130], [131] used local facial patches in the low and high resolution images to form geometrically similar manifolds. They used training images to estimate the high-resolution embedding and construct a smooth super-resolved image. Yang *et al.* [132] proposed a super resolution approach by representing local patches as a sparse linear combination of elements from high resolution images. In addition to these local models, Liu *et al.*

TABLE VIII
A COMPARISON OF FACE RECOGNITION METHODS FOR DISGUISE.

Algorithm	Database (# subjects/ # images)	Accuracy
Half eigen face [100]	AR (126/ 4000)	66.0% (rank-10 accuracy)
Phase features [103]	AR (126/ 4000)	81.2% (rank-1 accuracy)
	Disguise (16/ 196)	74.3% (rank-1 accuracy)
	Synthetic (100/ 4000)	83.3% (rank-1 accuracy)

TABLE IX
LIST OF WIDELY USED FACE DATABASES FOR DIFFERENT COVARIATES.

Covariate	Database	Description
Pose	CMU-PIE [19]	13 poses within $\pm 66°$ in yaw and $\pm 15°$ in tilt
	At&T (ORL) [104]	10 random poses within $\pm$ 20° in yaw and tilt
	XM2VTS [105]	5 poses:0° $\pm 30°$ in yaw and tilt
	Multi-PIE [106]	15 poses within $\pm 66°$ in yaw and $\pm 15°$ in tilt
	FERET [107]	18 poses, $0° to \pm 90°$
	Yale-B [23]	9 different poses $0°, 12°$ and $24°$
	FRGC [108]	50,000 recordings divided into training and validation
	CAS-PEAL [109]	21 different poses
Illumination	CMU-PIE [19]	Different illumination from 13 light sources
	AR [110]	Left, right, and all side lights on
	Yale-B [23]	64 lighting conditions and 1 ambient illumination
	CASIA NIR [56]	3,940 images of 197 subjects
	Notredame collection B [111]	33,287 images of 487 subjects
	Notredame collection C [111]	2,492 LWIR face images from 241 subjects
	PolyU NIR [112]	34,000 images of 335 subjects
Expression	JAFFE [113]	7 different facial expressions
	Cohn-Kanade [114]	Neutral to a peak expression
	CMU-AMP [115]	75 images showing different expressions
Aging	FG-NET [86]	6-18 images per subject from 0-69 years of age
	Morph [116]	46 days to 29 years
Disguise	AR [110]	accessories like glasses and scarf
Sketch	CUHK face sketch [117]	606 viewed sketches
	CUHK face sketch FERET [118]	1194 viewed sketches
Plastic surgery	IIIT-D plastic surgery [119]	900 subjects with different plastic surgery cases
Look-alikes & Twins	IIIT-D look alike [12]	50 subjects with 5 genuine and 5 look-alikes
	3D twins expression challenge [120]	428 images of 107 twin pairs
Low resolution face	SCface [121]	4160 surveillance images of 130 subjects
	MBGC [122]	399 walking sequence and 202 standard sequence (720×480)
	ND-QO-Flip crowd video [111]	14 crowd videos of 90 subjects
Unconstrained	Labeled faces in the wild [123]	13233 images of 5749 subjects
	MBGC [122]	Unconstrained face from still & video
	PubFig [124]	58,797 images of 200 subjects

[133] integrated a holistic parametric model and a local non-parametric model using two-step statistical modeling for face hallucination. Global linear model learns a relationship between the high resolution and low resolution images, whereas, the high-frequency content residue is modeled using a patch-based non-parametric Markov network. Super-resolution approaches are more susceptible to environmental variations and introduce distortions that affect recognition. Moreover, the primary objective of super-resolution is to obtain a good visual reconstruction from multiple low resolution faces, and these algorithms are generally not intended for recognition.

There are some approaches that simultaneously optimize super resolution as well as face recognition. Jia and Gong [134] combined super-resolution and face recognition by computing a maximum likelihood identity parameter vector in high-resolution tensor space for recognition. Hennings-Yeomans *et al.* [135] proposed an approach where facial features are included in a super-resolution method as the prior information for simultaneous reconstruction of super resolved images as well as recognition. They also proposed to use models from an image formation process, super-resolution priors, and face feature extraction methods. The matching algorithm extracts new features based on how well an intermediate super-resolution reconstruction of the probe image fits into the models used in the process. Recently, Biswas *et al.* [136] proposed a multidimensional scaling transformation approach to approximate the distance between low quality probe and high quality gallery image and makes it comparable to the distance had the probe been captured in the same conditions as the gallery. Let $f : R^d \to R^m$ denote the mapping from the input feature space R^d to the embedded Euclidean space R^m. Here m is the dimension of the transformed space and d denotes the input dimension. The mapping $f = (f_1, f_2, \cdots, f_m)^T$ is considered to be a linear combination of p basis functions of the form

$$f_i(x; W) = \sum_{j=1}^{p} \omega_{ji} \phi_j(x) \tag{2}$$

where $\phi_j(x); j = 1, 2, \cdots, p$ is a linear or non-linear function of the input feature vectors. $|W|_{ij} = \omega_{ij}$ is a $p \times m$ matrix of weights and the mapping function can also be defined as:

$$f(x; W) = W^T \phi(x) \tag{3}$$

The goal now is to simultaneously transform the feature vectors from high resolution gallery images and low resolution

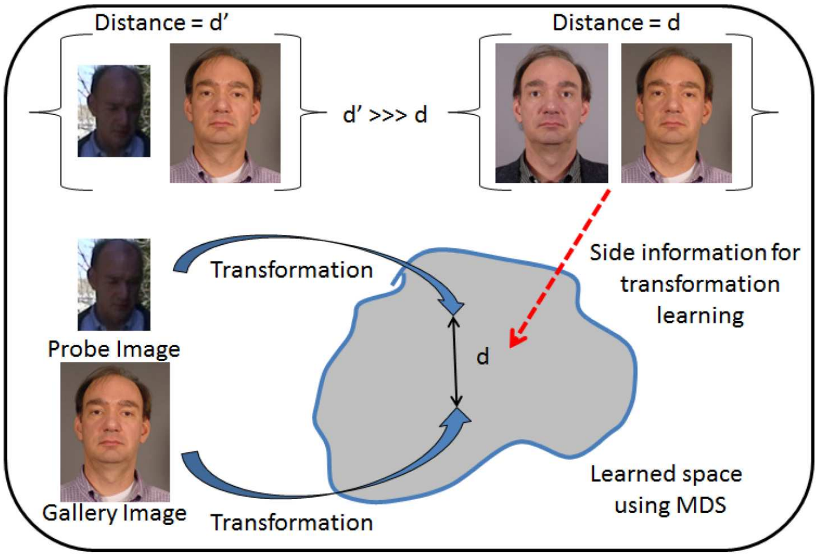


Fig. 16. Illustrates the multidimensional scaling transformation. Image from [136].

probe images such that the Euclidean distance between the transformed feature vectors approximates the distance had the probe images were captured in similar conditions. Fig. 16 illustrates the process of multidimensional scaling transform for matching low resolution probe images. In their experimental evaluation with surveillance quality face images, the gallery resolution is fixed at 60×55 and four different probe resolutions are considered, 32×28, 21×19, 16×14, and 13×11. Down sampling the gallery images to match them with low resolution query images has also been explored by researchers. However, down sampling face images results in loosing discriminating information useful for face recognition such as texture and other high frequency information. To this end, Lei *et al.* [137] proposed using magnitude and phase information to build a local frequency descriptor in frequency domain. Li *et al.* [125] used coupled mappings to project face images with different resolutions into a unified feature space to minimize the difference between the low-resolution image and its high-resolution counterpart.

Table X reports the performance of different approaches proposed for matching low resolution face images. Although, research in matching low resolution face images in not as matured as research in other covariates such as pose, illumination, and expression, some approaches have shown fairly good performance. In general, it has been observed that approaches based on local facial regions [137], [125] have achieved significantly better results as compared to the global approaches. Extracting resolution invariant features from images that can be matched across different resolutions can be another intersecting research direction.

B. Sketch Recognition

To apprehend individuals eluding from law enforcement, agencies match forensic sketches with the database comprising digital face images of known individuals. Generally, methods used by law enforcement agencies require manual matching of sketches with a large digital face (mugshot) database which is not pragmatic in real world applications. However, an automatic sketch to digital face image matching system can assist these agencies and make the recognition process efficient and relatively fast. Existing 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 do not generate the digital image from sketches or the sketch from digital images, but, perform feature extraction and matching using the given digital image and sketch.

1) Generative Approaches: Wang and Tang [138] 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 [139]. Liu *et al.* [140] proposed a non-linear discriminative classifier based approach for synthesizing sketches by preserving face geometry. Li *et al.* [141] matched sketches and photos using a method similar to the Eigen-transform after converting sketches to photos. Wang and Tang [117] further proposed using Markov Random Fields to automatically synthesize sketches from digital face images and vice-versa.

2) Discriminative Approaches: Uhl and Lobo [142] proposed photometric standardization of sketches to compare it with digital photos. They further geometrically normalized sketches and photos to match them using Eigen analysis. Yuen and Man [143] used local and global feature measurements between sketches and mug-shot images. Zhang *et al.* [144] compared the performance of humans and PCA-based algorithm for matching sketch-photo pairs with variations in gender, age, ethnicity, and inter-artist difference. They also discussed about the quality of sketches in terms of artist's skills, experience, exposure time, and distinctiveness of features [145]. Similarly, Nizami *et al.* [146] analyzed the

TABLE X
A COMPARISON OF DIFFERENT APPROACHES PROPOSED FOR MATCHING LOW RESOLUTION FACE IMAGES.

Authors	Approach	Database (# gallery/ # probes)	Gallery/ probe size	Rank-1 accuracy
Huang and He [127]	Subspace of coherent features	FERET(1196/1195)	$72 \times 72/\ 12 \times 12$	84.4%
Jia and Gong [134]	Maximum-likelihood identity parameter	FERET(1475/1475)	$56 \times 36/\ 14 \times 9$	86.2%
Hennings-Yeomans *et al.* [135]	Simultaneous super-resolution and matching	Multi-PIE (224/2912)	$24 \times 24/\ 6 \times 6$	62.8%
Biswas *et al.* [136]	Multidimensional scaling	Multi-PIE(236/236)	$36 \times 30/\ 12 \times 10$	76.5%
Lei *et al.* [137]	Low frequency descriptor	FERET(1196/1195)	$88 \times 80/\ 33 \times 30$	87.6%
Li *et al.* [125]	Coupled mapping projection	FERET(1196/1195)	$72 \times 72/\ 12 \times 12$	90.1%

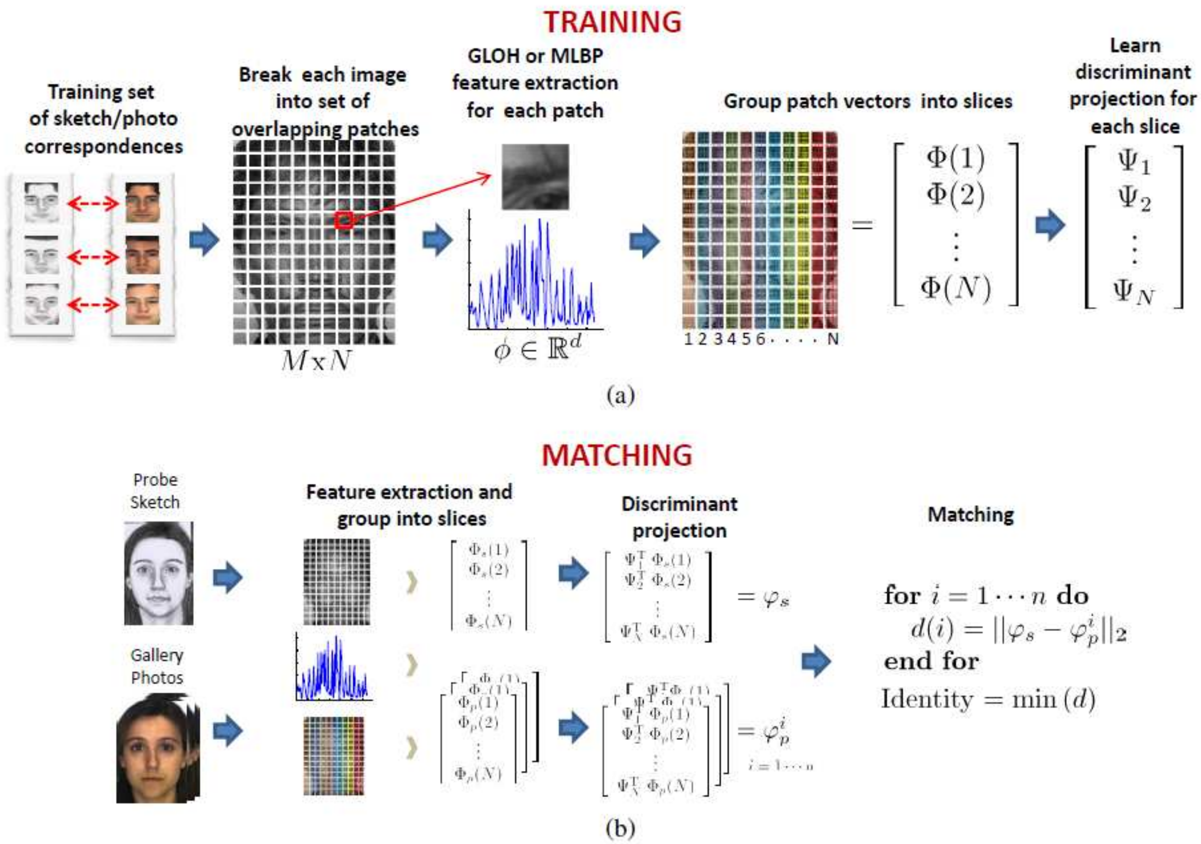


Fig. 17. Illustrating the steps involved in (a) training and (b) matching using the LFDA framework. Image from [149].

effect of matching sketches drawn by different artists. Klare and Jain [147] 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.* [148] 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.* [149] extended their approach using local feature discriminant analysis (LFDA) to match forensic sketches. As shown in Fig. 17, sketch and face images are first partitioned into N slices. Scale-invariant feature transform (SIFT) and multi-scale local binary pattern (MLBP) descriptors are computed for each slice, which remain stable between sketches and photos. Next, local-feature-based discriminant analysis (LFDA) is used to extract the most salient features for each slice. Finally, the system measures the similarity between feature vectors to match sketches with photos. The LFDA systems accuracy is further improved using subjects demographic information such as race, gender, age, and height. In their recent approach, Klare and Jain [150] proposed a framework for heterogeneous face recognition where both probes and gallery images are represented in terms of non-linear kernel similarities. Zhang *et al.* [151] analyzed the psychological behavior of humans for matching sketches drawn by different sketch artists. Zhang *et al.* [118] 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.

Table XI reports rank-1 identification accuracy of different approaches for matching sketches with digital face images. State-of-the-art in matching viewed sketches is about 99%, however, current techniques to match forensic sketches yield rank-1 identification accuracy of about 16%. Forensic sketches are drawn based on the recollection of an eye-witness from the crime scene and the expertise of sketch artist. Forensic

TABLE XI
A COMPARISON OF DIFFERENT APPROACHES PROPOSED FOR MATCHING SKETCHES WITH DIGITAL FACE IMAGES.

Authors	Approach	Database (# gallery/ #probes)	Image representation	Rank-1 accuracy
Wang and Tang [138]	Eigen-Transformation	CUHK (300/ 300)	Pixels	90.0%
Liu *et al.* [140]	LLE Transformation	CUHK (300/ 300)	Pixels	87.7%
Wang and Tang [117]	MRF Transformation	CUHK (300/ 300)	Pixels	96.3%
Klare *et al.* [147]	Direct matching	CUHK (300/ 300)	SIFT	97.8%
Bhatt *et al.* [148]	Genetic Algorithm	CUHK (233/ 233)	EUCLBP	94.1%
Klare *et al.* [149]	LFDA	Forensic Sketch (10,100/ 49)	SIFT, MLBP	16.3%

sketches [149] include several inadequacies because of the incomplete or approximate description provided by the eye-witness. As shown in Fig. 18, this may result in exaggeration of facial features leading to a sketch that does not resemble the actual person. Therefore, existing state-of-the-art face recognition algorithms cannot be used directly and require additional mechanism to address non-linear variations present in sketches and digital face images. Availability of large forensic sketch database and continued efforts from the research community is required to efficiently match a forensic sketch to a digital face database.

C. Plastic Surgery

The popularity of plastic surgery is driven by factors such as the availability of advanced technology, affordable cost and the speed with which these procedures are performed. These surgical procedures prove beneficial for patients suffering from structural or functional impairment of facial features, but these procedures can also be misused by individuals who are trying to conceal their identity with the intent to commit fraud or evade law enforcement. These surgical procedures may allow anti-social elements to freely move around without any fear of being identified by any face recognition system. Plastic surgery results being long-lasting or even permanent, provide an easy and robust way to evade law and security mechanism. As shown in Fig. 19, these procedures modify both the shape and texture of facial features to varying degrees, therefore, it is challenging for existing face recognition algorithms to match pre and post surgery images. Plastic surgery has been recently established as a new and important covariate of face recognition alongside pose, expression, illumination, aging and disguise [119]. Facial plastic surgery is a complex subtle process and, unlike aging, is not continuous. In our opinion, such non-uniform face transformations are not generic and difficult to be modeled. Singh *et. al.* [119] analyzed the effects of several types of local and global plastic surgery procedures and their effect on different face recognition algorithms. Their results showed that plastic surgery procedures reduce the performance of face recognition systems significantly.

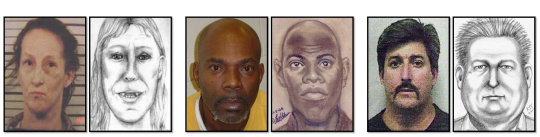


Fig. 18. Sample images showing exaggeration of facial features in forensic sketches.

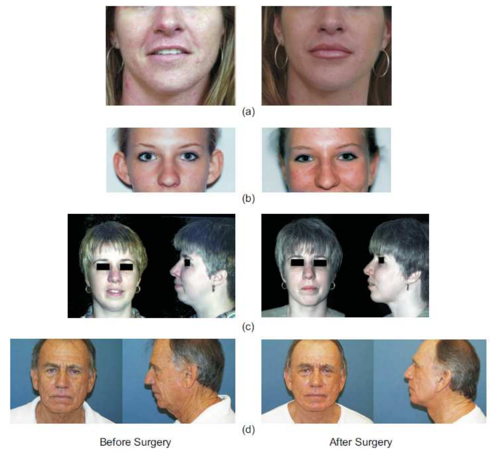


Fig. 19. Illustrating the example of (a) lip augmentation, (b) otoplasty or ear surgery, (c) liposubmental chin implant and liposuction of chin/neck, and (d) face resurfacing. Image from [119].

Bhatt *et al.* [152] proposed an evolutionary granular computing based algorithm for recognizing faces altered due to plastic surgery. The algorithm starts with generating non-disjoint face granules with each face granule having different information at varying size and resolution. Further, two feature extractors were used for extracting discriminating information from face granules. Finally, different responses are unified in an evolutionary manner using genetic algorithm for improved performance. Marsico *et al.* [153] also proposed an approach that integrates information derived from local region to match pre and post surgery face images. Plastic surgery also raises some social and ethical issues [154], being related to the medical history of an individual which is secure under law, invasion of privacy is an important constraint in this research. Apart from effecting the face recognition algorithms, plastic surgery procedures may also lead to identity theft. Identity theft can be intentional when a person consciously attempts to resemble someone by undergoing facial plastic surgery procedures or unintentional where he/she may resemble someone else after the surgery. Plastic surgery procedures modifying the facial geometry and texture along with associated privacy issues makes it an arduous research problem that seeks more attention from the research community.

These procedures can significantly change the facial regions both locally and globally, altering the appearance, facial features and texture, thereby posing a serious challenge to face recognition systems. Existing face recognition algorithms generally rely on local and global facial features and any variation can affect the recognition performance. More research is required to design optimal face recognition algorithms that can account for the challenges due to plastic surgery.

D. Look-alikes and Twins

Face recognition aims at finding a discriminating representation of face images that reduces the intra-class variations wile maximizing the inter-class variations. However, a face recognition algorithm can be deceived by low inter-class variations in look-alike faces. The problem of automatically recognizing look-alike faces has recently gained attention form the research community. Recognizing identical twins can be considered as a subclass of this problem as twins are biological look-alikes. Fig. 20 shows some examples of look-alike faces and identical twins.

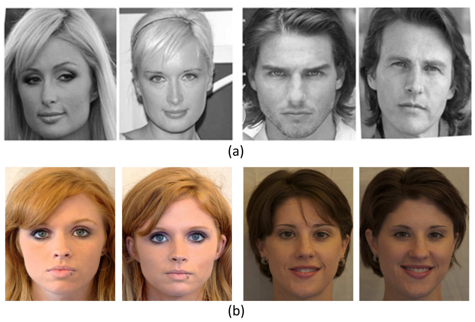


Fig. 20. Sample images showing (a) Look-alike faces (not twins) and (b) identical twins. Images from [12], [14], and [13].

Phillips *et al.* [14] performed a study on analyzing the performance of face recognition algorithms on twins under different lighting conditions, expressions, gender and age. Biswas *et al.* [155] investigated human capability to distinguish between identical twins based on facial traits. Lamba *et al.* [12] compared the performance of humans and automated face recognition algorithms for recognizing look-alike faces. They also proposed an approach to enhance the verification performance by extracting features from overlapping region around eyes, nose and mouth. Bowyer [15] performed experiments with identical twin data and provided research directions in analyzing twin data using face and iris biometrics. Klare *et al.* [13] also proposed a component based analysis of facial features in identical twins using discriminant learning methods. They also showed that using soft biometric features such as scars, skin marks can further enhance the performance of face recognition systems.

The research in matching look-alike faces is still in its preliminary stages, however, feature extraction and matching based on local facial regions have shown useful research direction in addressing the challenging problem of match look-alikes and twins.

Discussion: The emerging covariates discussed in this report bring out new research directions in face recognition. These covariates have gained a lot of attention in recent times from the research community because of their critical applications. A few techniques have also been proposed for recognizing faces under these covariates; however, the general applicability of these techniques still need to be validated. Since most of these techniques are trained and tested on a single database, they do not account for large inter-class variations present in different databases and thus may not generalize well. One of the limitations in developing robust solutions for face recognition is the lack of large databases for these emerging covariates. Publically available large databases will allow better understanding and characterization of these variations thus leading to better quality solutions. Several face recognition techniques often rely on large number of training images so as to efficiently address the variations caused due to the covariates. However, such solutions are not feasible in real world where few images per individual are available during training (or in gallery). Moreover, the emerging covariates such as sketch, plastic surgery, and low resolution face have important law enforcement applications; therefore the solutions for these covariates should be reliable and must have a quick response time. These observations suggest that further research is needed to develop robust face recognition algorithms to address the new emerging covariates.

V. Evolutionary Granular approach for matching surgically altered face images

Plastic surgery procedures are performed either for medical or cosmetic reasons. These surgical procedures alter the appearance, texture and shape of different facial regions. Plastic surgery is now established as a new and challenging covariate in face recognition alongside aging and disguise. Generally, facial appearances can be modified by using disguise accessories such as beard, moustache, and make up or by undergoing plastic surgery. As shown in Fig. 21, variations caused due to plastic surgery have some intersection with the variations caused due to aging and disguise. However, the overall impact of plastic surgery on face is rather diverse from the variations caused due to disguise and aging. 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 change that is typically 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. Variations caused due to disguise are temporary and reversible; however, variations caused due to plastic surgery are long-lasting and may not be reversible. Plastic surgery procedures may also lead to identity theft. Identity theft can be intentional when a person consciously attempts to resemble someone by undergoing facial plastic surgery procedures or unintentional where he/she may resemble someone else after the surgery.

The popularity of plastic surgery is increasing because of reduction in cost and speed with which these procedures can be performed. Even the wide spread acceptability of these surgical procedures in the society, to rejuvenate facial appearance, encourages individuals to undergo plastic surgery for cosmetic reasons. The following statistics provided by the American Society for Aesthetic Plastic Surgery [156]

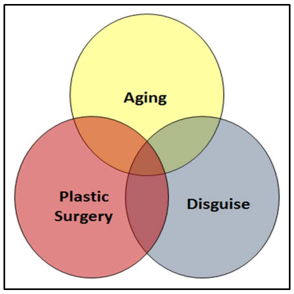


Fig. 21. Plastic surgery as an intersection of aging and disguise variations.

substantiate the facts about the popularity of plastic surgery procedures.

- From 2009 to 2010, there was almost a 9% increase in the total number of cosmetic surgical procedures, with over 1.6 million surgical procedures in 2010.
- In the age group of $35-50$ years, more than 4 million people (the maximum) underwent surgery which contributes to 44% of the total surgical procedures. People in the age group of $19-34$ years had 20% of the total procedures; age group $51-64$ years had 28%; people with age above 65 years had 7%, and under 18 years of age had 1.3%.
- Women had almost 8.6 million cosmetic procedures while men had over 750,000 cosmetic procedures costing nearly $\$10.7$ billion in 2010.

It is quite recent that research community has realized the necessity for face recognition systems to be robust and reliable in matching faces that are altered due to plastic surgery procedures. Singh *et al.* [119] analyzed the effects of 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 be addressed using current face recognition algorithms. Marsico *et al.* [153] proposed an approach that integrates information derived from local region to match pre- and post-surgery face images. The increasing popularity of plastic surgery and wide acceptance in the society has also nurtured several social and ethical issues. Bhatt *et al.* [154] discussed various social, ethical and engineering challenges unveiled by plastic surgery. This research proposes an approach for recognizing faces altered due to plastic surgery procedures and analyzes the effect of different surgical procedures on the performance of face recognition algorithms. Generally, face recognition algorithms either use facial information in a holistic way or extract features and process them in parts. On the other hand, cognitive neuroscientists have observed that humans solve problem using perception and knowledge represented at different levels of information granularity [157]. Humans 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 them to recognize a face even with altered appearances. Inspired from these observations, the algorithm starts with generating non-disjoint face granules with each face granule having different information at varying sizes and resolution. Further, two feature extractors, namely Extended Uniform Circular Local Binary Pattern (EUCLBP) [148] and Scale Invariant Feature Transform (SIFT) [158], are used for extracting discriminating information from face granules. Finally, different responses are unified in an evolutionary manner using genetic algorithm for improved performance. The major contributions of this research can be summarized as follows:

1) An evolutionary granular computing based algorithm for recognizing faces altered due to different plastic surgery procedures.
2) Analysis of different granular levels and their combination. These granular levels include discriminating information from multi-resolution gaussian and laplacian pyramids, different inner and outer facial regions, and local facial fragments.
3) Performance comparison of the proposed approach with the commercial-of-the-shelf (COTS) face recognition system and also presents results on matching surgically altered face images against large scale gallery.
4) Presents the effect of plastic surgery on periocular region as a biometric.
5) Presents a comprehensive analysis on the effect of different types of plastic surgery on the proposed face recognition algorithm.

A. Granular Computing Approach for Face Recognition

Sinha *et al.* established 19 results based on face recognition capabilities of the human mind [157]. They suggested that humans can efficiently recognize familiar face images even with low resolution and noise. Moreover, high and low frequency facial information are processed both holistically and locally. Further, Campbell *et al.* [159] reported that inner and outer facial regions represent distinct information which is helpful for face recognition. Moreover, researchers from cognition suggest that local facial fragments can provide robustness against partial occlusion and change in viewpoints [157], [160], [161]. It is our hypothesis that if these capabilities can be encoded in an automatic face recognition algorithm, then the recognition performance of the algorithm can be comparable to the performance of human mind for matching faces.

To incorporate the above mentioned research findings, we propose a granular approach [162], [163] for facial feature extraction and matching. In the granular computing approach a unified framework is used to extract non-disjoint features at different granularity levels. These features are then synergistically combined to obtain a more comprehensive information set. With granulated information, more flexibility is achieved in analyzing underlying information such as nose, ears, forehead, cheeks, or combination of two or more features. The proposed granulation process is described as follows:

1) Face Image Granulation: Let F be the detected frontal face image of size $n \times m$. Face granules are generated pertaining to three different levels of granularity. The first level of granularity provides global information at multiple levels of resolution. 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.* [159], at the second level of granularity, different inner and outer facial information are extracted. Since, local facial features play an important role in face recognition by human mind. Therefore, at the third level of granularity, we extract facial features from the local facial fragments.

First Level of Granularity: In the first level, face granules are generated by applying the Gaussian and Laplacian operators [35]. The Gaussian operator generates a sequence of low pass filtered images by iteratively convolving each of the constituent images 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 and reduction factor 4. The $reduce$ operation, Re, can be written as,

$$Re[F(p,q)] = \sum_{x=1}^{5}\sum_{y=1}^{5}\overline{w}(x,y)F(2p+x, 2q+y) \tag{4}$$

A Gaussian pyramid I_B is defined as,

$$I_0 = F \tag{5}$$

$$I_B = Re[I_{B-1}], \quad 0 \ < \ B \ < \ A \tag{6}$$

Further, the Laplacian operator generates band-pass images and the process can be summarized as follows:

$$L_B = I_B - Ex[I_{B+1}], \quad 0 \ \leq \ B \ < \ A \tag{7}$$

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

$$Ex[I_{B,D}(p,q)] = 4\sum_{x=-2}^{2}\sum_{y=-2}^{2}\overline{w}(x,y)I_{B,D-1}\left(\frac{p-x}{2}, \frac{q-y}{2}\right) \tag{8}$$

Note that $I_{B,D}$ in Equation 8 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. 22 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. This level of granular information thus provides resilience to variations in facial features such as eyes, mouth, and nose.

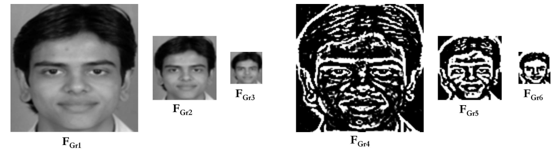


Fig. 22. Face granules in the first level of granularity. $F_{Gr1}, F_{Gr2}, and F_{Gr3}$ are generated by the Gaussian operator, and F_{Gr4}, F_{Gr5}, and F_{Gr6} are generated by the Laplacian operator.

Second Level of Granularity: To accommodate Campbell *et al'.s* [159] findings, in the second level of granularity, we generate horizontal and vertical granules by dividing the face image F into different regions as shown in Figs. 23 and 24. 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} have the same 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. 23 and 24 show horizontal and vertical granules when the size of normalized face image is 196×224 and $\epsilon = 15$[3]. This level of granularity provides resilience to variations in different inner and outer facial regions.

Fig. 23. Horizontal face granules from the second level of granularity ($F_{Gr7} - F_{Gr15}$).

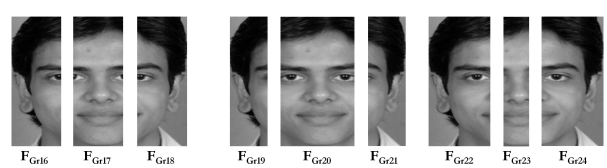


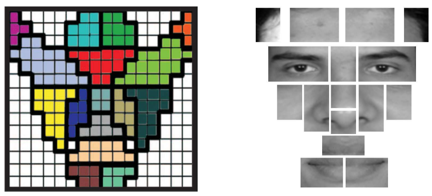
Fig. 24. Vertical face granules from the second level of granularity ($F_{Gr16} - F_{Gr24}$).

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 face granules in the third level of granularity. Given the eye coordinates, 16 local facial fragments are extracted using the golden ratio face template [164] shown in Fig. 25(a). Each of these fragments is a granule representing local information that

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

provides unique features for handling variations due to plastic surgery. Fig. 25(b) shows an example of local facial fragments used as face granules in the third level of granularity.

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. 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.



(a) (b)

Fig. 25. (a) Golden ratio face template [164], and (b) face granules from third level of granularity ($F_{Gr25} - F_{Gr40}$).

B. *Evolutionary Approach for Selection of Feature Extractor and Weight Optimization*

Psychological studies in face recognition [157] have shown that some facial regions are more discriminating than others and hence, contribute more towards the recognition accuracy. Moreover, humans [165] 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. To capture these observations, the proposed approach incorporates selection of optimal feature extractor to encode diverse information and weights for matching each face granules using weighted χ^2 distance.

1) Feature Extraction: For extracting facial features Extended Uniform Circular Local Binary Patterns (EUCLBP) and Scale Invariant Feature Transform (SIFT) are used. Both these feature extractors are fast, discriminating, rotation invariant, and robust to changes in gray level intensities due to illumination. They also efficiently use information assimilated from global as well as local facial regions. 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.

Extended Uniform Circular Local Binary Patterns: Local Binary Patterns (LBP) based descriptor [21], [166] is a widely used texture operator because of its robustness to gray level changes and high computational efficiency. In Circular Local Binary Patterns (CLBP), texture descriptor is computed based on the neighboring pixels well separated on a circle around a central pixel [21], [22]. The circle can have different diameters and varying number of neighbors to account for texture at different scales. CLBP is extended to Uniform Circular Local Binary Patterns [21] to achieve robustness to rotation variations and dimensionality reduction. A binary pattern is called uniform binary pattern if it has at most two bitwise transitions from 0 to 1 or vice-versa. A descriptor is computed where every uniform pattern has a separate bin and all non-uniform patterns are assigned to a single bin. The concatenation of all the histograms pertaining to each grid constitutes the image signature. Uniform CLBP is described using Eqs. 9 and 10.

$$C_{N,R}^{riu2}(p,q) = \begin{cases} \sum_{i=0}^{N-1} f(n_i - n_c)2^i & if \;\; U(C_{N,R}) \leq 2 \\ N+1 & otherwise \end{cases} \tag{9}$$

where,

$$U(C_{N,R}) = \begin{array}{l} \sum_{i=1}^{N-1} |f(n_i - n_c) - f(n_{i-1} - n_c)| \\ +|f(n_{N-1} - n_c) - f(n_0 - n_c)| \end{array} \tag{10}$$

where n_c corresponds to the gray-level intensity of center pixel of the circle and n_i corresponds to the gray-level intensities of N evenly spaced pixels on a circle of radius R. $riu2$ represents the use of rotation invariant uniform patterns.

Encoding difference of signs 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 values. Huang *et.al.* proposed a method to encode the exact difference of gray-level intensities and reported a marked improvement in the performance of texture descriptors [123]. This forms the motivation to further extend Uniform CLBP to encode exact gray-level difference along with the original encoding. The proposed descriptor is called Extended Uniform Circular Local Binary Pattern (EUCLBP). It provides information assimilated from the exact gray-level difference and adds a complimentary layer of discrimination on top of the original descriptor. Fig. 26 explains feature extraction using the proposed EUCLBP. Layer 1 is Uniform CLBP that encodes difference of signs while the other three layers encode the exact gray-level differences. We experimentally observed that Layer 1 and Layer 2 of EUCLBP are the most discriminating. Therefore, the final descriptor is the concatenation of Layer 1 and Layer 2 histograms.

Scale Invariant Feature Transform: SIFT is a scale and rotation invariant descriptor [158] that generates a compact representation of an image based on the magnitude, orientation, and spatial vicinity of the image gradients. For computing the SIFT descriptor, an image is tessellated into 7×6 local patches. The intensity image is used to compute the gradient image, which is weighted by a Gaussian kernel. The spatial coordinates in the gradient image are then quantized into $m \times n$ values. Gradient orientation ranges from $[0, \pi)$ for each gradient image pixel and is further quantized into one of the k orientation bins. At each of the $m \times n$ spatial coordinates, the sum of the Gaussian weighted gradient magnitude values is computed for each of the k orientations. This yields a feature descriptor of $m \times n \times k$ dimension, where each component contains the sum of weighted gradient magnitudes at the given location and orientation. The SIFT descriptor is normalized to unit length. Any component larger than 0.2 is truncated to 0.2

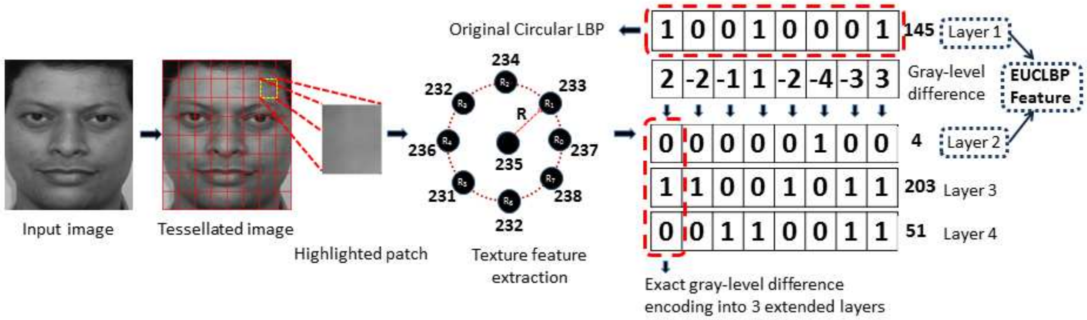


Fig. 26. Feature extraction using EUCLBP.

and the descriptor is re-normalized to unit length. In this paper, the parameters for computing SIFT descriptor are set to $m = 4$, $n = 4$, and $k = 8$, which results in a 128-dimensional SIFT descriptor. This process is repeated for all the tessellated regions and the final descriptor for a face images is obtained by concatenating all the descriptors.

2) Genetic Optimization: The uniqueness and discriminability of EUCLBP and SIFT features depends 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, the experiments are conducted to determine the performance of EUCLBP and SIFT for each of the granules. Feature extractor for each granule is selected depending on the reliability of features for that particular granule. Further, it is observed that every face granule has different contribution towards the recognition accuracy as shown in Table XIV. It is our assertion that giving higher preference to face granules that have more contribution towards the recognition performance and selecting the correct feature extractor for each granule should improve the overall accuracy.

Based on the above observations, the next task is simultaneously optimizing the selection of feature extractor and weights associated with every face granule for matching. The problem of finding the optimal feature extractor and the weights for each granule embroils searching very large spaces and finding several suboptimal solutions. Genetic algorithms are well proven in searching very large spaces to quickly reach to the near optimal solution [167]. Therefore, we propose an evolutionary genetic approach to select feature extractor and corresponding weights for each face granule. Fig. 27 represents the genetic search process to find optimum feature extractor and weights for each face granule. 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: 1) for selecting feature extractor (referred to as chromosome $type1$) and 2) for assigning weights to each face granule (referred to as chromosome $type2$). Each 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. Chromosome $type2$ has real valued numbers associated with corresponding weights of the 40 face granules.

Initial Population: Initially, two generations of 100 chromosomes corresponding to two different optimization functions are populated.

1) For selecting feature extractor ($type1$ chromosome), half the initial generation i.e. 50 chromosomes; is set with all bits as 1 representing EUCLBP as the feature extractor for all 40 face granules. Remaining 50 chromosomes in the initial generation have all bits as 0 representing SIFT as the feature extractor for all 40 face granules.
2) For assigning weights to each face granule ($type2$ chromosome), a chromosome with weights proportional to the identification accuracy of individual face granule (as proposed by Ahonen [22]) is used as the seed chromosome. The remaining 99 chromosomes are generated by randomly changing one or more units in the initial chromosome. The weights are normalized such that the sum of all 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 is computed on a training set and 10 best performing chromosomes are selected for crossover and mutation to populate the next generation.

Crossover: A set of uniform crossover operations is performed on 10 best performing chromosomes to populate a new generation of 100 chromosomes. A set of uniform crossover operations is performed to populate the next generation. 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 previous generation. For chromosome $type1$, mutation is performed by randomly inverting the bits of the chromosome.

The search process is repeated till convergence, i.e. till the identification accuracy for new generation is better than previous generations. At this point, the optimum feature extractor and weights for each face granule pertaining to the best performing chromosomes (i.e. chromosomes giving best recognition accuracy on the training data) are obtained. Genetic optimization also enables to discard redundant and non-discriminating face granules whose contribution towards recognition accuracy is very low (i.e. the weight for that face granule is close to 0). This leads to 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 loss of diversity in a population that leads to decrease in the quality of solution. In this research, adaptive mutation rate [168] and random offspring generation [169] 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 population diversity decreases and vice-versa. Population diversity is measured as the overall standard deviation of the weights assigned to each unit in a chromosome. Further, random offspring generation is used to produce random offsprings if there is a high degree of similarity among 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 the offsprings are generated randomly.

C. Combining Face Granules with Evolutionary Learning for Recognition

The granular approach for matching faces altered due to facial 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) To match corresponding features extracted from the gallery and probe images, descriptors for each face granule are first normalized. The weighted χ^2 distance measure is used to the compute the dissimilarity score. Here, the weights for each face granule are learnt using the genetic approach.

$$\chi^2(a,b) = \sum_{i,j} \omega_j \frac{(a_{i,j} - b_{i,j})^2}{a_{i,j} + b_{i,j}} \tag{11}$$

where a and b are the normalized descriptors (EUCLBP or SIFT descriptors), i and j correspond to the i^{th} bin of the j^{th} face granule, and ω_j is the weight for the j^{th} face granule.
4) In identification mode ($1:N$), this procedure is repeated for each gallery-probe pair and top matches are obtained based on the dissimilarity scores.

D. Experimental Results

To evaluate the performance of the proposed algorithm two different databases are used. First is the plastic surgery face database [119] that comprises images from 900 subjects who have undergone plastic surgery. Second database comprises of the plastic surgery face database in addition to 1800 non-plastic surgery images from another 900 subjects. Section V-D1 provides details about the databases used in this research, section V-D2 elaborates the experimental protocol whereas section V-D3 presents the experimental analysis.

1) Database: In this research, the plastic surgery face database [119] is used which comprises of 1800 pre- and post-surgery images for 900 subjects with frontal pose, proper illumination and neutral expression. It is a real world database that consist of different types of facial plastic surgery cases such as rhinoplasty (nose surgery), blepharoplasty (eyelid surgery), brow lift, skin peeling, and rhytidectomy (face lift). In real world applications, 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 plastic surgery face database is appended with 1800 non-surgery images pertaining to another 900 subjects from different face databases. This database is termed as the *combined heterogeneous face database* and comprises 3600 images pertaining to 1800 subjects. The non-surgery images pertaining to 900 subjects are from the same database used by Singh *et. al.* [119] which consists of two frontal, proper illumination and neutral expression images from different face databases.

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 make them zero mean and unit variance. Then histogram equalization is applied to maximize the image contrast by applying a gray level transform which tries to flatten the resulting histogram. Further, wiener filtering is applied to restore the blurred edges. Finally, face images are geometrically normalized and size of each detected face image is 192×224 pixels.

2) Experimental Protocol: To evaluate the efficacy of the proposed approach, different experiments are performed with 10 times non-overlapping random cross validation. In each experiment, 40% of the database is used for training and the remaining 60% is used for testing [4]. 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 are used. 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 probe.

[4] Samples in these cross validation trials were same as Singh *et. al'.s* experiments.

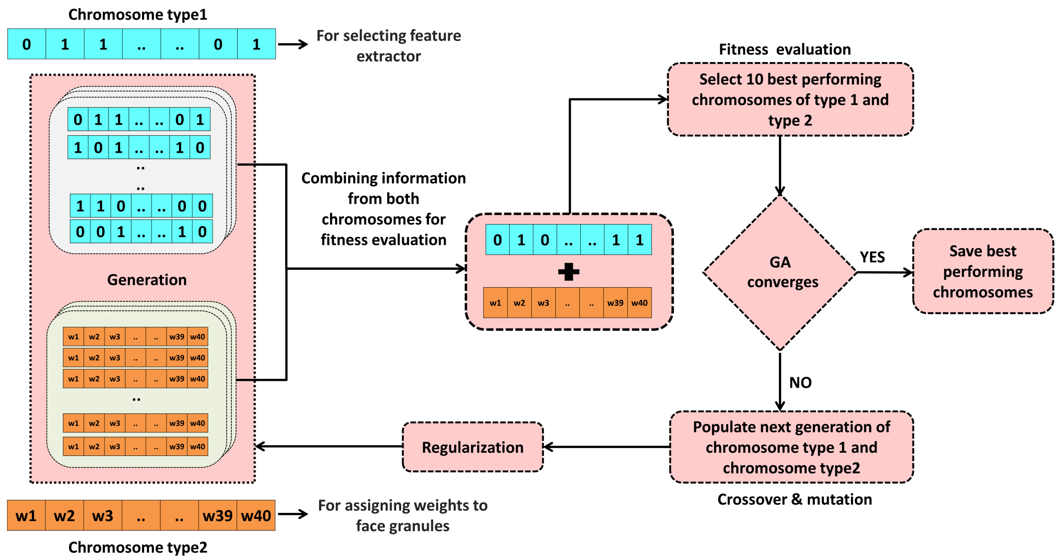


Fig. 27. Genetic optimization process for selecting feature extractor and weight for each face granule.

- *Experiment 2:* Out of 1800 subjects from the combined heterogeneous face database, 720 subjects are used for training and remaining the 1080 subjects 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 completely unaware of any plastic surgery cases.
- *Experiment 3:* To evaluate the effectiveness of the proposed approach for matching individuals with large sized 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. This increases the gallery size to 8124.
  - 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 8124 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 8124 subjects.

3) Analysis: The performance of the proposed approach is compared with the sum-rule fusion [170] of SIFT and EUCLBP on different granules and the commercial-of-the-shelf (COTS) face recognition system for matching face images altered due to plastic surgery procedures. Key results and observations are summarized below.

- The CMC in Fig. 28 shows rank-1 identification accuracy of all the algorithms for Experiments 1 and 2. The proposed approach outperforms other algorithms by at least 5.27% on the plastic surgery face database and 5.02% on the combined heterogeneous face database. The proposed approach 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 data-sets consist of pre- and post-surgery images along with non-surgery images. It closely resembles the condition that real world face recognitions systems encounter. 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.
- CMC curves in Fig. 29 show the performance of the proposed algorithm and the commercial system on large sized gallery (experiment 3). The proposed algorithm yields rank-1 identification accuracy of 83.88% which is about 4.6% better than the performance of the commercial system for matching probes from the plastic surgery face database. The proposed approach gives 86.33% rank-1 identification accuracy for matching probes from the combined heterogeneous face database.
- Table XIV shows individual rank-1 identification accuracies 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 performed better on granules that comprise fiducial features such as eyes, nose, and mouth, however its performance on flat facial 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 from flat regions.
- The performance of EUCLBP when applied on full face image is compared with the performance when it is

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

Database	Training/Testing Images	Algorithm	Rank-1 Identification Accuracy	Standard Deviation
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 XIII
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

TABLE XIV
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%

applied on face granules. The results show that, applying EUCLBP on face granules improves the rank-1 accuracy by 7-8% as compared to a full face image. This improvement in recognition accuracy can be attributed to the discriminating information obtained from face granules varying in size and resolution.

- To show the efficacy of evolutionary approach for selecting feature extractor and weight optimization using genetic algorithm, the performance is compared with sum-rule fusion [170] 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 was selected for 22 face granules whereas EUCLBP was selected for 18 face granules.
- Different types of plastic surgery procedures have varying effect on one or more facial regions. In experiment-2, training-testing data-sets have both pre- and post-surgery images and non-surgery images. The proposed algorithm inherently provides the benefit of addressing the non-linear variations induced by different plastic surgery procedures.
- Experimental results strengthen our assertion that local

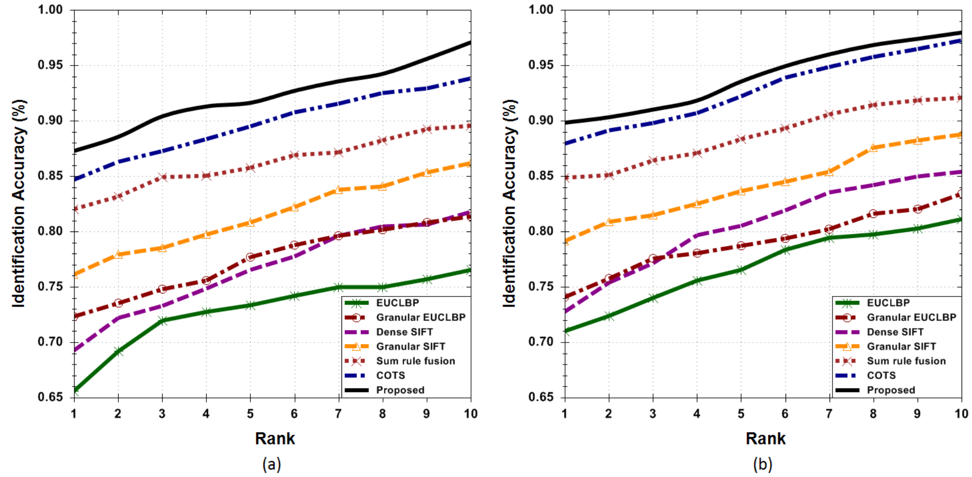


Fig. 28. CMC curves for the proposed and existing algorithms on the (a) plastic surgery face database, and (b) combined heterogeneous face database.

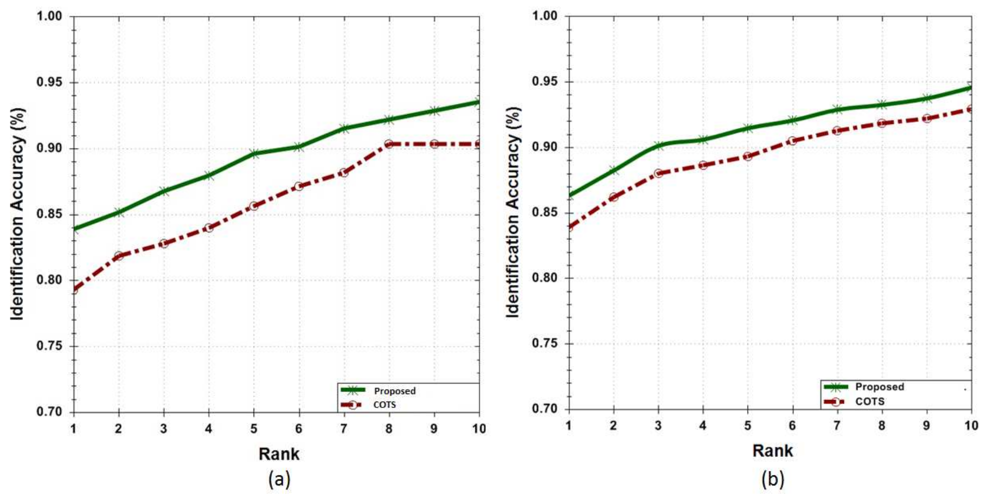


Fig. 29. CMC curves for the proposed and commercial algorithms for large scale evaluation on probe images from the (a) plastic surgery face database, and (b) combined heterogeneous face database.

information based approaches such as face granules can handle the variations introduced by plastic surgery procedures. The ability to encode local features at different resolution and size allows the proposed algorithm to be resilient to such non-linear variations.

4) Analysis of Different Types of Plastic Surgery Procedures: Global plastic surgery [119] can completely transform the face and is recommended for patients where functional damage is to be cured such as patients with fatal burns or trauma. In these kinds of surgeries, facial appearance, skin texture and feature shapes may vary drastically, thus, making it arduous for any face recognition system to recognize faces before and after surgery. Rhytidectomy (facelift) is used to treat patients with severe burns on face and neck. It can also be employed to confront aging and get a younger look by treating the face skin. In our experiments, rhytidectomy has the maximum influence on performance of the proposed approach as it modifies the appearance and texture of the whole face. Skin peeling procedures such as laser resurfacing and chemical peel are used to treat wrinkles, stretch marks, acne and other skin damages caused due to aging and sun burn. Using skin resurfacing for gaining suave skin alters the texture information that affects the performance of the proposed approach. As shown in Fig. 30(a), these two global plastic surgery procedures have severe impact on the performance of proposed approach. Rank-1 identification accuracy of 71.76% and 85.09% is obtained for subjects who have undergone rhytidectomy and skin peeling respectively. Local plastic surgery [119] is aimed at reshaping and restructuring facial features to improve the aesthetics. These surgical procedures result in varying amount of changes in the geometric distance between facial features but the overall texture and appearance may look 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 local plastic surgery procedures listed in [119], dermabrasion has the most prominent effect on the performance of the

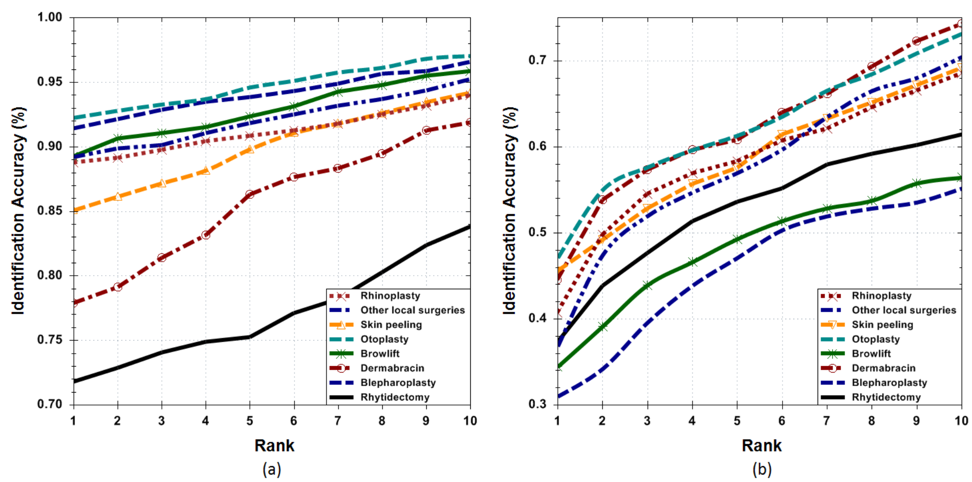


Fig. 30. CMC curves on different types of local and global plastic surgery procedures for (a) the proposed algorithms, and (b) sum-rule fusion of SIFT and EUCLBP on left and right periocular region.

proposed approach as it drastically changes the face texture. The proposed approach yields rank-1 identification accuracy of 77.89% for dermabrasion cases. Otoplasty involves bringing the ears closer to the face thus reducing the size of ears and orienting structural ear elements. Ears are not considered in most of the face recognition algorithms; therefore, the proposed approach is not affected by this type of plastic surgery. Other local plastic surgery procedures also affect the performance of the proposed approach to varying degrees. The performance of the proposed approach across images with different kind of local plastic surgery procedures is shown in Fig. 30.

5) Analysis of Different Granules: To understand the effect of different granules for recognizing face images altered due to plastic surgery, a detailed experimental study on individual level of granulation is performed. The correlation analysis of all three granules are reported in Table XIII. CMC curves in Fig. 31(a) and (b) show the identification accuracy for individual levels of granulation 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 pyramid at level-1 and level-2 provide high level prominent features of a face. Different physiological studies illustrate that humans use different inner and outer facial features to identify individuals [165]. The inner facial features include nose, eyes, eye brows, and mouth while the outer facial region comprises face outline, structure of jaw/chin, and forehead. Therefore, granular level-2 extracts information from different inner and outer facial regions representing discriminating information which is useful for face recognition. Local facial fragments such as nose, eyes, and mouth provide robustness to partial occlusion and change in viewpoints. Human mind can efficiently distinguish and classify individuals with their local facial fragments. Therefore, granular level-3 assimilates discriminating information from these fragments. The performance of the proposed evolutionary granular approach is optimized for a particular granular level by assigning null weights to the face granules corresponding to other granularities during genetic optimization. Further to analyze the improvement using complimentary information provided by different face granules, performance is evaluated for different combinations of granular levels. The proposed approach is optimized for different combinations of granular levels by assigning null weights to granules in the remaining granular levels. CMC curves in Figs. 31(c) and (d) show the results for different combinations of granular levels on the two databases.

Periocular Region: Recent studies have shown that periocular region can be used as a biometric [171] and is the least invasive among all eye based biometrics. Eyelid is the thin skin that covers and protects our eyes and is a major feature in periocular based recognition algorithm. In the proposed granulation scheme, granules 29 and 31 provide right and left periocular regions respectively. Following the experimental protocol of Experiment 1 in Section V-D2, periocular region is used to identify individuals who have undergone plastic surgery. CMC curves in Fig. 32 show the performance of periocular based biometrics for matching surgically altered faces from the plastic surgery face database. The performance is calculated based on the sum-rule fusion [170] of SIFT and EUCLBP on left and right periocular regions.

Blepharoplasty (Eyelid surgery) is identified as one of the top five surgical procedures performed in year 2010 [156] and is used to reshape both upper as well as lower eyelids to treat excessive growth of skin tissues obstructing the vision. Further, to analyze the effect of blepharoplasty on periocular region, experiments are performed with periocular region for different types of local and global plastic surgeries. Table XV and CMC curves in Fig. 30(b) show that rank-1 identification accuracy using periocular region for matching faces altered due to specific types of plastic surgery. Blepharoplasty alters periocular region, therefore, it adversely affects the performance of periocular based biometrics. Moreover, it is observed that the performance of periocular biometrics is

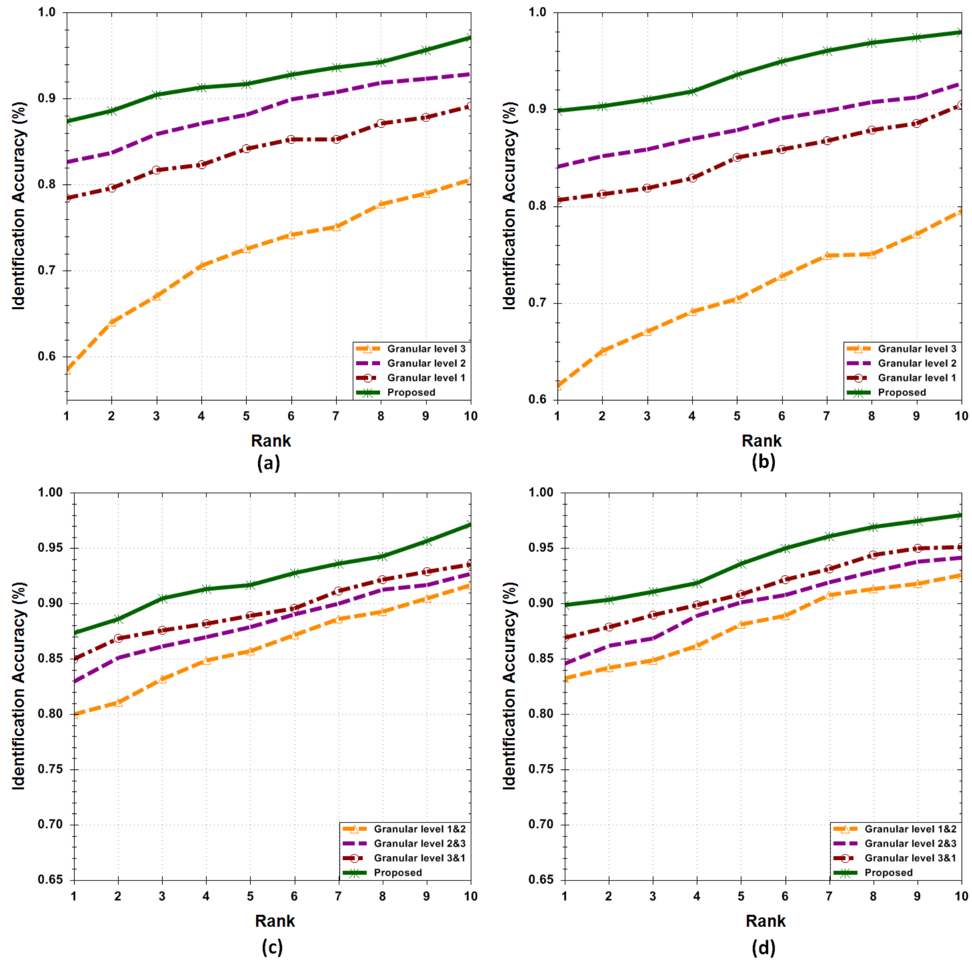


Fig. 31. CMC curves showing the performance of individual granular levels on (a) the plastic surgery face database, (b) the combined heterogeneous face database, performance of combination of granular levels on (c) the plastic surgery face database, and (d) the combined heterogeneous face database.

TABLE XV
RANK-1 IDENTIFICATION PERFORMANCE OF THE PROPOSED APPROACH AND THE PERIOCULAR BASED METHOD ON DIFFERENT TYPES OF PLASTIC SURGERY PROCEDURES.

Type	Surgery	Accuracy of Proposed Approach	Accuracy of Periocular Region
Local	Browlift	89.22%	34.42%
	Dermabrasion	77.89%	44.56%
	Otoplasty	92.25%	47.25%
	Blepharoplasty	91.42%	30.96%
	Rhinoplasty	88.85%	40.71%
	Other	89.17%	35.81%
Global	Rhytidectomy	71.76%	37.27%
	Skin peeling	85.09%	45.83%
	Overall	87.32%	40.11%

affected when a local region close to periocular region (such as nose and forehead) is altered due to plastic surgery. This is mainly because modifying a local feature also bring some changes in the adjacent facial regions. The results suggest that plastic surgery is also an important challenge for periocular biometrics; though, periocular based algorithms have shown robustness to aging and occlusion.

VI. CONCLUSION

Generally, face recognition systems and algorithms are designed to recognize faces of cooperative individuals in controlled environment. However, it becomes a challenging problem when faces are captured in uncontrolled non-ideal conditions. This report presents a review of existing techniques categorized based on different covariates of face recognition. It provides discussions on the major challenges posed by these covariates, techniques proposed to address these challenges, current limitations, and future research directions. It divides these covariates into existing and emerging covariates of face recognition based on how well a covariate is studied and documented in literature. The report also presents an evolutionary granular approach for matching surgically altered face images,

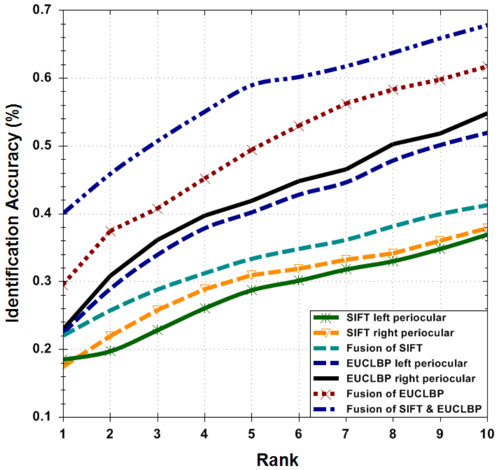


Fig. 32. CMC curves showing the performance of periocular region on the plastic surgery face database.

an *emerging covariate*. It presents an algorithm for selecting better feature extractor and optimal weights for each face granule. The proposed algorithm automatically evolves itself to address the non-linearity induced by different types of plastic surgery procedures and outperforms commercial algorithm on different databases.

REFERENCES

[1] J. Daugman, "How iris recognition works", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004.
[2] A. Jain and S. Li, *Handbook of face recognition*, Springer-Verlag New York, 2005.
[3] M. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: A survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34–58, 2002.
[4] H. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 1, pp. 23–38, 1998.
[5] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2001, vol. 1, pp. 511–518.
[6] C. Zhang and Z. Zhang, "A survey of recent advances in face detection", Tech. Rep. MSR-TR-2010-66, 2010.
[7] W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face recognition: A literature survey", *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.
[8] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi, "Recent advances in visual and infrared face recognition: A review", *Computer Vision and Image Understanding*, vol. 97, pp. 103–135, 2005.
[9] P. Belhumeur, "Ongoing challenges in face recognition", 2004.
[10] A. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey", *Pattern Recognition Letters*, vol. 28, pp. 1885–1906, 2007.
[11] B. Klare and A. Jain, "On a taxonomy of facial features", in *Proceedings of International Conference on Biometrics: Theory Applications and Systems*, 2010, pp. 1–8.
[12] H. Lamba, A. Sarkar, M. Vatsa, and R. Singh, "Face recognition for look-alikes: A preliminary study", in *Proceedings of International Conference on Biometrics: Theory, Applications and Systems*, 2011, pp. 1–6.
[13] B. Klare, A. Paulino, and A. Jain, "Analysis of facial features in identical twins", in *Proceedings of International Conference on Biometrics: Theory, Applications and Systems*, 2011, pp. 1–6.
[14] P. Phillips, P. Flynn, K. Bowyer, R. Bruegge, P. Grother, G. Quinn, and M. Pruitt, "Distinguishing identical twins by face recognition", in *Proceedings of International Conference on Automatic Face Gesture Recognition and Workshops*, 2011, pp. 185–192.
[15] K. Bowyer, "What surprises do identical twins have for identity science?", *IEEE Computer*, vol. 44, pp. 100–102, 2011.
[16] A. Jain, B. Klare, and U. Park, "Face recognition: Some challenges in forensics", in *Proceedings of International Conference on Automatic Face Gesture Recognition and Workshops*, 2011, pp. 726–733.
[17] X. Zhang and Y. Gao, "Face recognition across pose: A review", *Pattern Recognition*, vol. 42, pp. 2876–2896, 2009.
[18] X. Zou, J. Kittler, and K. Messer, "Illumination invariant face recognition: A survey", in *Proceedings of International Conference on Biometrics: Theory, Applications, and Systems*, 2007, pp. 1–8.
[19] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1615–1618, 2003.
[20] L. Wiskott, J. Fellous, N. Krger, and C. Malsburg, "Face recognition by elastic bunch graph matching", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 775–779, 1997.
[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] P. Bellhumer, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 7, pp. 711–720, 1997.
[24] R. Singh, M. Vatsa, A. Ross, and A. Noore, "A mosaicing scheme for pose-invariant face recognition", *IEEE Transactions on Systems, Man, and Cybernetics-B*, vol. 37, no. 5, pp. 1212–1225, 2007.
[25] D. Beymer and T. Poggio, "Face recognition from one example view", in *Proceedings of International Conference on Computer Vision*, 1995, pp. 500–507.
[26] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
[27] F. Kahraman, B. Kurt, and M. Gokmen, "Robust face alignment for illumination and pose invariant face recognition", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–7.
[28] X. Liu and T. Chen, "Pose-robust face recognition using geometry assisted probabilistic modeling", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 502–509.
[29] F. Yang and A. Krzyzak, "Face recognition under significant pose variation", in *Proceedings of Canadian Conference on Electrical and Computer Engineering*, 2007, pp. 1313–1316.
[30] X. Zhang, Y. Gao, and M. Leung, "Automatic texture synthesis for face recognition from single views", in *Proceedings of International Conference on Pattern Recognition*, 2006, vol. 3, pp. 1151–1154.
[31] M. Lee and S. Ranganath, "Pose-invariant face recognition using a 3D deformable model", *Pattern Recognition*, vol. 36, pp. 1835–1846, 2003.
[32] D. Jiang, Y. Hu, S. Yan, L. Zhang, H. Zhang, and W. Gao, "Efficient 3D reconstruction for face recognition", *Pattern Recognition*, vol. 38, pp. 787–798, 2005.
[33] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1063–1074, 2003.
[34] C. Castillo and D. Jacobs, "Using stereo matching for 2-D face recognition across pose", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.
[35] P. Burt and E. Adelson, "A multiresolution spline with application to image mosaics", *ACM Transaction on Graphics*, vol. 2, no. 4, pp. 217–236, 1983.
[36] D. Gonzalez-Jimenez and J. Alba-Castro, "Toward pose-invariant 2-D face recognition through point distribution models and facial symmetry", *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 3, pp. 413–429, 2007.
[37] T. Vetter, "Synthesis of novel views from a single face image", *International Journal on Computer Vision*, vol. 28, pp. 103–116, 1998.
[38] T. Kim and J. Kittler, "Design and fusion of pose-invariant face-identification experts", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 9, pp. 1096–1106, 2006.

[39] Y. Hu, D. Jiang, S. Yan, L. Zhang, and H. Zhang, "Automatic 3D reconstruction for face recognition", in *Proceedings of International Conference on Automatic Face and Gesture Recognition*, 2004, pp. 843–848.

[40] A. Lanitis, C. Taylor, and T. Cootes, "Automatic interpretation and coding of face images using flexible models", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 743–756, 1997.

[41] U. Prabhu, J. Heo, and M. Savvides, "Unconstrained pose-invariant face recognition using 3D Generic Elastic Models", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1952–1961, 2011.

[42] G. Passalis, P. Perakis, T. Theoharis, and I. Kakadiaris, "Using facial symmetry to handle pose variations in real-world 3D face recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1938–1951, 2011.

[43] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 1994, pp. 995–999.

[44] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[45] Y. Gao and M. Leung, "Face recognition using line edge map", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 764–779, 2002.

[46] W. Zhao and R. Chellappa, "Illumination-insensitive face recognition using symmetric shape-from-shading", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2000, vol. 1, pp. 286–293.

[47] T. Riklin-Raviv and A. Shashua, "The quotient image: Class based recognition and synthesis under varying illumination conditions", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 1999, vol. 2, pp. 637–643.

[48] C. Chen and C. Chen, "Lighting normalization with generic intrinsic illumination subspace for face recognition", in *Proceedings of International Conference on Computer Vision*, 2005, vol. 2, pp. 1089–1096.

[49] M. Savvides, V. Bhagavatula, and P. Khosla, "Eigenphases vs eigenfaces", in *Proceedings of International Conference on Pattern Recognition*, 2004, vol. 3, pp. 810–813.

[50] J. Heo, M. Savvides, and V. Bhagavatula, "Illumination tolerant face recognition using phase-only support vector machines in the frequency domain", in *Proceedings of International Conference on Advances in Pattern Recognition*, 2005, vol. 3687, pp. 66–73.

[51] R. Gonzalez and R. Woods, *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., 2nd edition, 2001.

[52] X. Xie and K. Lam, "An efficient method for face recognition under varying illumination", in *Proceedings of International Symposium on Circuits and Systems*, 2005, pp. 3841–3844.

[53] K. Chang, K. Bowyer, and P. Flynn, "An evaluation of multimodal 2D+3D face biometrics", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 619–624, 2005.

[54] D. Socolinsky and A. Selinger, "A comparative analysis of face recognition performance with visible and thermal infrared imagery", in *Proceedings of International Conference on Pattern Recognition*, 2002, vol. 4, pp. 217–222.

[55] G. Bebis, A. Gyaourova, S. Singh, and I. Pavlidis, "Face recognition by fusing thermal infrared and visible imagery", *Image and Vision Computing*, vol. 24, pp. 727–742, 2006.

[56] S. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 627–639, 2007.

[57] X. Zou, J. Kittler, and K. Messer, "Ambient illumination variation removal by active Near-IR imaging", in *Proceedings of International Conference on Biometric*, 2006.

[58] Lei Zhang and D. Samaras, "Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 3, pp. 351–363, 2006.

[59] S. Wei and S. Lai, "Robust face recognition under lighting variations", in *Proceedings of International Conference on Pattern Recognition*, 2004, vol. 1, pp. 354–357.

[60] B. Wang, W. Li, W. Yang, and Q. Liao, "Illumination normalization based on Weber's law with application to face recognition", *IEEE Signal Processing Letters*, vol. 18, no. 8, pp. 462–465, 2011.

[61] C. Hsieh, S. Lai, and Y. Chen, "Expression-invariant face recognition with constrained optical flow warping", *IEEE Transactions on Multimedia*, vol. 11, no. 4, pp. 600–610, 2009.

[62] J. Kittler, A. Hilton, M. Hamouz, and J. Illingworth, "3D assisted face recognition: A survey of 3D imaging, modelling and recognition approachest", in *Proceedings of International Conference on Computer Vision and Pattern Recognition Workshops*, 2005, pp. 114–117.

[63] D. Socolinsky, A. Selinger, and J. Neuheisel, "Face recognition with visible and thermal infrared imagery", *Computer Vision and Image Understanding*, vol. 91, pp. 72–114, 2003.

[64] D. Socolinsky and A. Selinger, "Thermal face recognition in an operational scenario", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2004, vol. 2, pp. 1012–1019.

[65] X. Chen, P. Flynn, and K. Bowyer, "Visible-light and infrared face recognition", in *Proceedings of Workshop on Multimodal User Authentication*, 2003, pp. 48–55.

[66] R. Singh, M. Vatsa, and A. Noore, "Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition", *Pattern Recognition*, vol. 41, pp. 880–893, 2008.

[67] Y. Ni, E. Krichen, W. Hizem, S. Garcia-Salicetti, and B. Dorizzi, "Active differential CMOS imaging device for human face recognition", *IEEE Signal Processing Letters*, vol. 13, no. 4, pp. 220–223, 2006.

[68] E. Bronstein, M. Bronstein, and R. Kimmel, "Expression-invariant 3D face recognition", in *Proceedings of Conference on Audio and Video-based Biometric Person Authentication*, 2003, pp. 62–69.

[69] A. Bronstein, M. Bronstein, and R. Kimmel, "Expression-invariant representations of faces", *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 188–197, 2007.

[70] P. Nagesh and B. Li, "A compressive sensing approach for expression-invariant face recognition", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1518–1525.

[71] L. Wiskott and C. Malsburg, "Face recognition by dynamic link matching", in *Proceedings of Lateral Interactions in the Cortex: Structure and Function*, 1996, pp. 6–9.

[72] A. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 748–763, 2002.

[73] B. Amberg, R. Knothe, and T. Vetter, "Expression invariant 3D face recognition with a morphable model", in *Proceeding of International Conference on Automatic Face Gesture Recognition and Workshops*, 2008, pp. 1–6.

[74] D. Gorodnichy, "Video-based framework for face recognition in video", in *Proceedings of Canadian Conference on Computer and Robot Vision*, 2005, pp. 330–338.

[75] X. Tan, S. Chen, Z. Zhou, and J. Liu, "Face recognition under occlusions and variant expressions with partial similarity", *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 2, pp. 217–230, 2009.

[76] P. Tsai and T. Jan, "Expression-invariant face recognition system using subspace model analysis", in *Proceedings of International Conference on Systems, Man and Cybernetics*, 2005, vol. 2, pp. 1712–1717.

[77] L. Yin, X. Wei, Y. Sun, J. Wang, and M. Rosato, "A 3D facial expression database for facial behavior research", in *Proceedings of International Conference on Automatic Face and Gesture Recognition and Workshops*, 2006, pp. 211–216.

[78] M. Ramachandran, S. Zhou, D. Jhalani, and R. Chellappa, "A method for converting a smiling face to a neutral face with applications to face recognition", in *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, 2005, vol. 2, pp. 977–980.

[79] X. Li, G. Mori, and H. Zhang, "Expression-invariant face recognition with expression classification", in *Proceedings of Canadian Conference on Computer and Robot Vision*, 2006, pp. 77–80.

[80] Y. Yacoob and L. Davis, "Recognizing human facial expressions from long image sequences using optical flow", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 6, pp. 636–642, 1996.

[81] M. Martinez, "Recognizing expression variant faces from a single sample image per class", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 353–358.

[82] C. Hsieh, S. Lai, and Y. Chen, "An optical flow-based approach to robust face recognition under expression variations", *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 233–240, 2010.

[83] D. Smeets, T. Fabry, J. Hermans, D. Vandermeulen, and P. Suetens, "Isometric deformation modeling using singular value decomposition for 3D expression-invariant face recognition", in *Proceedings of*

International Conference on Biometrics: Theory, Applications, and Systems, 2009, pp. 1–6.

[84] U. Park, Y. Tong, and A. Jain, "Age-invariant face recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 947–954, 2010.

[85] Z. Li, U. Park, and A. Jain, "A discriminative model for age invariant face recognition", *IEEE Transactions on Information Forensics and Security*, 2011.

[86] "FG-Net aging database - http://www.fgnet.rsunit.com/".

[87] M. Burt and D. Perrett, "Perception of age in adult caucasian male faces: Computer graphic manipulation of shape and colour information", *Journal of Royal Society*, vol. 259, pp. 137–143, 1995.

[88] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2006, vol. 1, pp. 387 – 394.

[89] N. Ramanathan and R. Chellappa, "Modeling shape and textural variations in aging faces", in *Proceedings of International Conference on Automatic Face Gesture Recognition and Workshops*, 2008, pp. 1–8.

[90] A. Lanitis, C. Taylor, and T. Cootes, "Toward automatic simulation of aging effects on face images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 442–455, 2002.

[91] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation", *IEEE Transactions on Systems, Man, and Cybernetics-B*, vol. 34, no. 1, pp. 621–628, 2004.

[92] X. Geng, Z. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2234–2240, 2007.

[93] Y. Kwon and N. Lobo, "Age classification from facial images", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 1994, pp. 762–767.

[94] R. Shaw J. Todd, L. Mark and J. Pittenger, "The perception of human growth", *Scientific American*, vol. 242, no. 2, pp. 132–144, 1980.

[95] S. Biswas, G. Aggarwal, N. Ramanathan, and R. Chellappa, "A non-generative approach for face recognition across aging", in *Proceedings of International Conference on Biometrics: Theory, Applications and Systems*, 2008, pp. 1–6.

[96] G. Guo, Y. Fu, C. Dyer, and T. Huang, "Image-based human age estimation by manifold learning and locally adjusted robust regression", *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1178–1188, 2008.

[97] N. Ramanathan and R. Chellappa, "Face verification across age progression", *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3349–3361, 2006.

[98] J. Wang, Y. Shang, G. Su, and X. Lin, "Age simulation for face recognition", in *Proceedings of International Conference on Pattern Recognition*, 2006, vol. 3, pp. 913–916.

[99] H. Ling, S. Soatto, N. Ramanathan, and D. Jacobs, "Face verification across age progression using discriminative methods", *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 1, pp. 82–91, 2010.

[100] N. Ramanathan, R. Chellappa, and A. Roy Chowdhury, "Facial similarity across age, disguise, illumination and pose", in *Proceedings of International Conference on Image Processing*, 2004, vol. 3, pp. 1999–2002.

[101] P. Silva and A. Rosa, "Face recognition based on eigeneyes", *Pattern Recognition and Image Analysis*, vol. 13, no. 2, pp. 335–338, 2003.

[102] S. Pamudurthy, E. Guan, K. Mueller, and M. Rafailovich, "Dynamic approach for face recognition using digital image skin correlation.", in *Proceedings of International Conference on Audio and Video based Biometric Person Authentication*, 2005, pp. 1010–1018.

[103] R. Singh, M. Vatsa, and A. Noore, "Face recognition with disguise and single gallery images", *Image and Vision Computing*, vol. 27, no. 3, pp. 245–257, 2009.

[104] "ORL - http://www.cl.cam.ac.uk/research/dtg/facedatabase.html".

[105] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre, "Xm2vtsbd: The extended m2vts database", in *Proceedings of International Conference on Audio and Video-base Biometric Personal Verification*, 1999, vol. 3.

[106] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-pie", *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010.

[107] P. J. Phillips, H. Moon, S. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face recognition algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[108] "FRGC - http://www.nist.gov/itl/iad/ig/frgc.cfm".

[109] W. Gao, B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, and D. Zhao, "The cas-peal large-scale chinese face database and baseline evaluations", *IEEE Transactions on Systems, Man and Cybernetics-A*, vol. 38, no. 1, pp. 149–161, 2008.

[110] A. Martinez and R. Benevento, "The AR face database", *CVC Technical Report #24*, 1998.

[111] "CVRL Data Sets - http://www.nd.edu/c̃vrl/UNDBiometricsDatabase.html".

[112] B. Zhang, L. Zhang, D. Zhang, and L. Shen, "Directional binary code with application to polyu near-infrared face database", *Pattern Recognition Letters*, vol. 31, pp. 2337–2344, October 2010.

[113] M. Lyons, J. Budynek, and S. Akamatsu, "Automatic classification of single facial images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 12, pp. 1357–1362, 1999.

[114] P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression", in *Proceedings of International Conference on Computer Vision and Pattern Recognition Workshops*, 2010, pp. 94–101.

[115] "CMU AMP - http://amp.ece.cmu.edu/faceauthentication/download.htm".

[116] K. Ricanek and T. Tesafaye, "Morph: a longitudinal image database of normal adult age-progression", in *Proceedings of International Conference on Automatic Face and Gesture Recognition*, 2006, pp. 341–345.

[117] 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.

[118] 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, pp. 513–520.

[119] R. Singh, M. Vatsa, H. Bhatt, S. Bharadwaj, A. Noore, and 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.

[120] "3D Twins Expression Challenge - http://www.nd.edu/c̃vrl/datasets.html".

[121] M. Grgic, K. Delac, and S. Grgic, "Scface - surveillance cameras face database", *Multimedia Tools and Applications Journal*, vol. 51, no. 3, pp. 863–879, 2011.

[122] "Mbgc v2 - http://www.nist.gov/itl/iad/ig/mbgc.cfm".

[123] G. Huang, M. Ramesh, T. Berg, and E. Leonard, "Labeled faces in the wild : A database for studying face recognition in unconstrained environment", Tech. Rep. 07-49, University of Massachusetts, 2007.

[124] N. Kumar, A. Berg, P. Belhumeur, and S. Nayar, "Attribute and Simile Classifiers for Face Verification", in *Proceedings of International Conference on Computer Vision*, 2009.

[125] B. Li, H. Chang, S. Shan, and X. Chen, "Low-resolution face recognition via coupled locality preserving mappings", *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 20–23, 2010.

[126] P. Hennings-Yeomans, V. Bhagavatula, and S. Baker, "Robust low-resolution face identification and verification using high-resolution features", in *Proceedings of International Conference on Image Processing*, 2009, pp. 33–36.

[127] H. Huang and H. He, "Super-resolution method for face recognition using nonlinear mappings on coherent features", *IEEE Transactions on Neural Networks*, vol. 22, no. 1, pp. 121–130, 2011.

[128] S. Baker and T. Kanade, "Limits on super-resolution and how to break them", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1167–1183, 2002.

[129] A. Chakrabarti, A. Rajagopalan, and R. Chellappa, "Super-resolution of face images using kernel PCA-based prior", *IEEE Transactions on Multimedia*, vol. 9, no. 4, pp. 888–892, 2007.

[130] H. Chang, D. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2004, pp. 275–282.

[131] B. Li, H. Chang, S. Shan, and X. Chen, "Locality preserving constraints for super-resolution with neighbor embedding", in *Proceedings of International Conference on Image Processing*, 2009, pp. 1189–1192.

[132] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation", *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.

[133] C. Liu, H. Shum, and W. Freeman, "Face hallucination: Theory and practice", *International Journal of Computer Vision*, vol. 75, no. 1, pp. 115–134, 2007.

[134] K. Jia and S. Gong, "Multi-modal tensor face for simultaneous super-resolution and recognition", in *Proceedings of International Conference on Computer Vision*, 2005, pp. 1683–1690.

[135] P. Hennings-Yeomans, S. Baker, and V. Bhagavatula, "Simultaneous super-resolution and feature extraction for recognition of low-resolution faces", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.

[136] S. Biswas, G. Aggarwal, and P. Flynn, "Pose-robust recognition of low-resolution face images", in *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2011, pp. 601 –608.

[137] Z. Lei, T. Ahonen, M. Pietikainen, and S. Li, "Local frequency descriptor for low-resolution face recognition", in *Proceedings of International Conference on Automatic Face Gesture Recognition and Workshops*, 2011, pp. 161–166.

[138] X. Tang and X. Wang, "Face photo recognition using sketch", in *Proceedings of International Conference on Image Processing*, 2002, vol. 1, pp. 257–260.

[139] X. Tang and X. Wang, "Face sketch synthesis and recognition", in *Proceedings of International Conference on Computer Vision*, 2003, vol. 1, pp. 687–694.

[140] 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.

[141] 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.

[142] 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.

[143] 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.

[144] 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.

[145] 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.

[146] 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.

[147] 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.

[148] 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, pp. 1–7.

[149] 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.

[150] B. Klare and A. Jain, "Heterogeneous face recognition using kernel prototype similarities", Tech. Rep., Michigan State University, 2011.

[151] 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 and Workshops*, 2011.

[152] H. Bhatt, S. Bharadwaj, R. Singh, M. Vatsa, and A. Noore, "Evolutionary granular approach for recognizing faces altered due to plastic surgery", in *Proceedings of International Conference on Automatic Face Gesture Recognition and Workshops*, 2011, pp. 720–725.

[153] 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.

[154] H. Bhatt, S. Bharadwaj, R. Singh, M. Vatsa, and A. Noore, "Face recognition and plastic surgery: Social, ethical and engineering challenges", in *Proceedings of International Conference on Ethics and Policy of Biometrics and International Data Sharing*, 2010, vol. 6005, pp. 70–75.

[155] S. Biswas, K. Bowyer, and P. Flynn, "A study of face recognition of identical twins by humans", in *Proceedings of International Workshop on Information Forensics and Security*, 2011.

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

[157] 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.

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

[159] R. Campbell, M. Coleman, W. Michael, B. Jane, J. Philip, 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.

[160] W. 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.

[161] A. Schwaninger, J. Lobmaier, and S. 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.

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

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

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

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

[166] T. Ojala and M. Pietikainenand D. Harwood, "A comparative study of texture measures with classification based on feature distributions", *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, January 1996.

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

[168] 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.

[169] 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.

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

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