GRD Journal for Engineering | Volume 1 | Issue 2 | January 2016
ISSN: 2455-5703

# Development Phases of Technologies in Face Recognition Systems

**Jesny Antony**
*PG Student*
*Department of Computer Science and Information Systems Engineering*
*Federal Institute of Science and Technology, Mookkannoor P O, Angamaly, Ernakulam, Kerala 683577, India*

**Prasad J.C.**
*Associate Professor*
*Department of Computer Science and Engineering*
*Federal Institute of Science and Technology, Mookkannoor P O, Angamaly, Ernakulam, Kerala 683577, India*

## Abstract

Face recognition is for recognizing human faces from single images out of a large database. The task is difficult because of image variation in terms of position, size, expression, and pose and it is important because this uses in different areas such as surveillance at airports, border crossings, security systems etc. There are two underlying motivations for us to write this survey paper: the first is to provide an up-to-date review of the existing literature, and the second is to offer some insights into the studies of machine recognition of faces. There are many methods exist to perform face detection, on this, Some use flesh tones, some use contours, and other are even more complex involving templates, neural networks, or filters. This all methods generally utilize 2D images for feature extraction and matching. Nowadays some of the algorithms make use of 3D images for the same. It will give more accurate and higher resilience towards covariates, such as expression, illumination and pose variations. But it is challenging because of the high cost of 3D sensors. RGB-D face recognition is introduced to reduce this problem of high cost of 3D sensors, which makes use of the features of both original image and depth image for the face recognition. This survey focuses on different 2D or 3D or combination of both face recognition techniques under different covariates, such as illuminations, expressions, noise, disguised conditions and different pose variations.
**Keywords- Face recognition, RGB-D, LBP, SURF features, Kinect, Entropy, Saliency**

## I. INTRODUCTION

Face Recognition, though an easy task for humans is a long standing research area due to the challenging covariates, such as disguise and aging, which make it very accurately verify the identity of a person. Some major approaches proposed for face recognition includes, Principal component analysis(PCA) [1], Fisher's Linear Discriminant analysis (FLDA) [2], Local Binary patterns (LBP) [3], Scale Invariant Feature Transform (SIFT) [4] and Sparse Representation Classifier (SRC) [5]. In all these cases, face recognition will be accurate only when complete face is visible in the image and when all the features can be extracted. There are lots of challenges in this area such as occlusion, rotation, overlapping, noise, pose variations, illuminations, expression differences, disguised images or faces that undergoes some surgical procedures such as plastic surgery and all are make the face recognition task becomes more difficult. To overcome these problems different sources of information have been used like skin color, geometry of the face, 3D information of the face etc. Among all these information depth information is proven to be very useful.

The human face plays an important role in our social interaction, conveying people's identity. Using the human face as a key to security, biometric face recognition technology has received significant attention in the past several years due to its potential for a wide variety of applications in both law enforcement and non-law enforcement. Biometric technologies that offer greater accuracy, like fingerprint, iris, but it require much greater cooperation from the user. For example, fingerprint requires each user make physical contact with the sensor surface and iris imaging requires that each user carefully position their eye relative to their position. So these biometric technologies are good in the case of authentication and security areas. While this is not that efficient in the case of investigation, human computer interactions and surveillance areas.

Three-dimensional face recognition technique uses 3D sensors to capture information about the shape of a face. This information is then used to identify distinctive features on the surface of a face, such as the contour of the eye sockets, nose, and chin. One advantage of 3D facial recognition is that it is not affected by changes in lighting like other techniques. Three-dimensional data points from a face vastly improve the precision of facial recognition. 3D research is enhanced by the development of sophisticated sensors that do a better job of capturing 3D face imagery. The sensors work by projecting structured light onto the face.

Recent RGB-D cameras, such as the Microsoft's Kinect camera, acquire both visual information (RGB) like regular camera systems, as well as depth information at high frame rates. The images in the 2D cameras are not enough for the better and accurate face recognition. Even though the 3D sensors give more accurate result that are not practical because of the high cost.

RGB-D cameras are the solution to this problem, which give more accurate result than the 2D sensors by the use of RGB image and depth image.

## II. Technological Development in Face Recognition

The subject of face recognition is as old as computer vision, both because of the practical importance of the topic and theoretical interest from cognitive scientists. Despite the fact that other methods of identification such as fingerprints, or iris scans can be more accurate, face recognition has always remains a major focus of research because of its non-invasive nature and because it is people's primary method of person identification.
The different such techniques are explained here:

### A. Face Recognition using Elastic Graph Matching:

Kohonen [6], who demonstrated that a simple neural net could perform face recognition for aligned and normalized face images. The type of network he employed computed a face description by approximating the eigenvectors of the face image's autocorrelation matrix, these eigenvectors are now known as Eigen faces. Kohonen's system was not a practical success, however, because of the need for precise alignment and normalization. In following years many researchers tried face recognition schemes based on edges, inter-feature distances, and other neural net approaches. While several were successful on small databases of aligned images, none successfully addressed the more realistic problem of large databases where the location and scale of the face is unknown.

Kirby and Sirovich (1989) [7] later introduced an algebraic manipulation which made it easy to directly calculate the Eigen faces, and showed that fewer than 100 were required to accurately code carefully aligned and normalized face images. Turk and Pentland [8] then demonstrated that the residual error when coding using the Eigen faces could be used both to detect faces in cluttered natural imagery, and to determine the precise location and scale of faces in an image. They then demonstrated that by coupling this method for detecting and localizing faces with the Eigen face recognition method, one could achieve reliable, real-time recognition of faces in a minimally constrained environment. This demonstration that simple, real-time pattern recognition techniques could be combined to create a useful system sparked an explosion of interest in the topic of face recognition.

### B. Principal Components Analysis (PCA):

The PCA technique18 converts each two dimensional image into a one dimensional vector. This vector is then undergoes several steps such as Detect Face in Image, Normalize Facial landmarks, Extract Facial Features and then Recognize Face Image. The technique selects the features of the face which vary the most from the rest of the image. In the process of decomposition, a large amount of data is discarded as not containing significant information since 90% of the total variance in the face is contained in 5-10% of the components. This means that the data needed to identify an individual is a fraction of the data presented in the image. Each face image is represented as a weighted sum (feature vector) of the principle components (or Eigen faces), which are stored in a one dimensional array. Each component represents only a certain feature of the face, which may or may not be present in the original image. A probe image is compared against a gallery image by measuring the distance between their respective feature vectors. For PCA to work well the probe image must be similar to the gallery image in terms of size (or scale), pose, and illumination. It is generally true that PCA is reasonably sensitive to scale variation.

Kyong I. Chang Kevin W. Bowyer Patrick J. Flynn [9] face recognition performs the comparison and combination of 2D and 3D face recognition algorithm. This algorithm makes use of a PCA-based approach tuned separately for 2D and for 3D. The first step of this algorithm is normalization which omits the background and leave only the extract the Eigen face from the image. Next step is the data collection of gallery and probe images. A gallery image is an image that is enrolled into the system to be identified. A probe image is a test image to be matched against the gallery images. After that the distance matrices are computed. 2D data represents a face by intensity variation whereas 3D data represents a face by shape variation. Next, the data fusion is performed. The images can simply be concatenated together to form one larger aggregate 2D-plus-3D face image. Metric level fusion combines the match distances that are found in the individual spaces. Having distance metrics from two or more different spaces, a rule for combination of the distances across the different biometrics for each person in the gallery can be applied. The ranks can then be determined based on the combined distances. However, we also find a recognition rate of 98.5% in a single probe study and 98.8% in multi-probe study, which is statistically significantly greater than either 2D or 3D alone. In general, the results show that the path to higher accuracy and robustness in biometrics involves use of multiple biometrics rather than the best possible sensor and algorithm for a single biometric.

### C. Biometrics Recognition:

Biometrics is used in the process of authentication of a person by verifying or identifying that a user requesting a network resource is who he, she, or it claims to be, and vice versa. It uses the property that a human trait associated with a person itself like structure of finger, face details etc. By comparing the existing data with the incoming data we can verify the identity of a particular person. There are many types of biometric system like fingerprint recognition, face detection and recognition, iris recognition etc., these traits are used for human identification in surveillance system, criminal identification. Advantages of using these traits for identification are that they cannot be forgotten or lost. These are unique features of a human being which is being used widely.

 

### *D. Multi objective Evolutionary Algorithm:*

Himanshu S. Bhatt, Samarth Bharadwaj, Richa Singh and Mayank Vatsa, [10] face recognition algorithm which makes use of a Multi objective Evolutionary Algorithm. In this research, the algorithm is proposed to match face images before and after plastic surgery. The face of Michael Jackson is a better example of this method. Plastic surgery procedures provide a proficient and enduring way to enhance the facial appearance as well as, plastic surgery procedures are beneficial for patients suffering from several kinds of disorders caused due to excessive structural growth of facial features or skin tissues. The algorithm first generates non-disjoint face granules at multiple levels of granularity. The granular information is assimilated using a multi objective genetic approach that simultaneously optimizes the selection of feature extractor for each face granule along with the weights of individual granules. In feature extraction, two feature extractors are used, EUCLBP and SIFT. Then, multi-objective evolutionary algorithm is applied. (Assign proper weight for each granule). This Unifies diverse information from all granules to address the nonlinear variations in pre- and post-surgery images. Aging and occlusions are a challenge to this algorithm.

### *E. Local Binary Patterns*

A local binary pattern (LBP) is a type of feature used for classification in computer vision. LBP is the particular case of the Texture Spectrum model proposed in 1990. It has since been found to be a powerful feature for texture classification. it has further been determined that when LBP is combined with the Histogram of oriented gradients (HOG) descriptor, it improves the detection performance considerably on some dataset.

T. I. Dhamecha, R. Singh, M. Vatsa, and A. Kumar [11] face recognition mainly concentrating on disguised faces. In this, it uses localized feature descriptors which can identify the disguised face patches and this makes the improves matching accuracy. Disguise can be due to many reasons such as, variations in hair style, variations due to beard and mustache, variations door to glasses, variations due to cap or hats, variations due to mask etc. The facial part or patches which are under the effect of disguise, are the least useful for face recognition, and may also provide misleading details. It is this misrepresentation that a person uses to hide his/her own identity and/or to impersonate someone else. This proposed framework is called Anavrta. There are mainly two stages for this framework, patch classification and patch based face recognition. Patch classification divides the image into patches and classify them with biometric and non-biometric classes. In patch based face recognition, biometric patches are matched using Local Binary Patterns (LBP) based face recognition algorithm. The recognition accuracy of familiar and same ethnicity subjects is found to be significantly better than that of unfamiliar and different ethnicity. The faces with similar disguise accessories account considerably high error rates.

### *F. RGB-D Based Face Recognition*

N. Engelhard, F. Endres, J. Hess, J. Sturm, and W. Burgard [12] face recognition present a RGB-D SLAM system. That is, this method generates colored 3D models of objects and indoor scenes using the hand-held Microsoft Kinect sensor. The algorithm first, extract SURF features from the incoming color images. Then the extracted features are match against features from the previous images. By evaluating the depth images at the locations of these feature points, we will get a set of point-wise 3D correspondences between any two frames. Based on these correspondences, we estimate the relative transformation between the frames using RANSAC algorithm. The third step is to improve this initial estimate using a variant of the ICP algorithm. As the pair-wise pose estimates between frames are not necessarily globally consistent, we optimize the resulting pose graph in the fourth step using a pose graph solver. The output of our algorithm is a globally consistent 3D model of the perceived environment, represented as a colored point cloud. This approach enables a robot to generate 3D models of the objects in the scene. But also applications outside of robotics are possible. For example, this system could be used by interior designers to generate models of flats and to digitally refurbish them and show them to potential customers.

B. Y. L. Li, A. S. Mian, W. Liu, and A. Krishna [13] face recognition system that makes use of low resolution 3D sensor for face recognition under challenging conditions such as varying poses, expressions, illuminations and disguise. Simultaneously dealing with different variations is a challenging task for face recognition. A preprocessing algorithm is proposed which exploits the facial symmetry at the 3D point cloud level to obtain a canonical frontal view, shape and texture, of the faces irrespective of their initial pose. This algorithm also fills holes and smooth's the noisy depth data produced by the low resolution sensor. The canonical depth map and texture of a query face are then sparse approximated from separate dictionaries learned from training data. The texture is transformed from the RGB to Discriminant Color Space before sparse coding and the reconstruction errors from the two sparse coding steps are added for individual identities in the dictionary. The query face is assigned the identity with the smallest reconstruction error. This algorithm requires only the nose tip position. An efficient Iterative Closest Point (ICP) method and facial symmetry are used to canonicalize non-frontal faces. A multi-modal sparse coding based approach is proposed to utilize Kinect color texture and depth information (RGB-D). Ultimately, we can recognize faces under different poses, expressions, illumination and disguise using a single algorithm, which is compact and scalable. This gives a recognition rate of 96.7% for th RGB-D data and 88.7% for the noisy depth data alone.

R. I. Hg, P. Jasek, C. Rofidal, K. Nasrollahi, T. B. Moeslund, and G. Tranchet [14] face recognition system based on the Microsoft's Kinect for Windows Face detection. This algorithm also makes use of both color image and depth image for the purpose face recognition. In this, a scene is set to produce different face positions. The experiment for this method uses thirteen points on a wall behind the Kinect were chosen and each person looked on that points sequentially to achieve roughly the same angles for each person. Then the face detection is performed by using the HK-classification method. Then based on the two simple rules such as Both eyes have to be above the nose and the triangle should be approximately equilateral the complete face is detected.

Then the depth image registration is performed. After that face validation using Principal Component Analysis is done to check whether the registered depth images actually contain a face image or not. Triangle detection is the method used for the PCA validation. This means the system detects the face 93.62 % of time for pose direct looking into the camera. However, this detection rate drops for the other poses which makes sense, because the employed method for face detection in this algorithm is based on finding facial triangles, which is obviously not visible in rotated images where the user does not face Kinect.

Gaurav Goswami, Mayank Vatsa, and Richa Singh [15] face recognition, uses the RGB image and depth image for the purpose of face recognition. It is basically based on the texture and attributes features of the image. Because of the cost of 3D sensors are very high, use of 3D sensors becomes a challenge in this area. These RGB-D images can be captured using low cost RGB-D sensors such as Kinect. The proposed algorithm initially undergoes preprocessing in which it detects the face in the given image using Viola Jones face detector algorithm. Then the algorithm computes the HOG descriptor based on entropy map of RGB-D faces along with the saliency feature obtained from 2D face. The combination of the five HOG descriptors provides the texture feature descriptor which is used as the input to trained Random Decision Forest (RDF) classifier to obtain the match score. This method is called RGB-D Image Descriptor Based on Saliency and Entropy (RISE) Along with this texture features, it also computes the attribute features of the image. Geometric features means the distance between various facial features such as eyes, noses and chin. This attributes are computes from the Depth map. There are mainly two stages are there in this attribute based on Depth map (ADM) method, First is Key point Labeling means to extract the geometric features a few facial key points are located such as nose tip, eye sockets and chin. Second is Geometric Attribute Computation means various distances between this landmark points are computes such as inter-eye distance, eye to nose bridge distance, nose bridge to nose tip distance, nose tip to chin distance, nose bridge to chin distance, chin to eye distance, eyebrow length, nose tip distance to both ends of both eyebrows, and overall length of the face. Then these two methods RISE and ADM methods are combined using Match Score Level Fusion or Rank Level Fusion to calculate the total match score value of the image.

## III. CONCLUSIONS

This paper attempts to provide a comprehensive survey of research on face detection. Although significant progress has been made in the last two decades, there is still work to be done, and we believe that a robust face detection system should be effective under full variation in lighting conditions, orientation, pose, and partial occlusion, facial expression, and presence of glasses, facial hair, and a variety of hair styles. Face detection is a challenging and interesting problem in and of itself. However, it can also be seen as a one of the few attempts at solving one of the grand challenges of computer vision, the recognition of object classes. The class of faces admits a great deal of shape, color, and variability due to differences in individuals, no rigidity, facial hair, glasses, and makeup. Images are formed under variable lighting and 3D pose and may have cluttered backgrounds. Hence, face detection research confronts the full range of challenges found in general purpose, object class recognition. However, the class of faces also has very apparent regularities that are exploited by many heuristic or model-based methods or are readily learned in data-driven methods. One expects some regularity when defining classes in general, but they may not be so apparent.

## REFERENCES

[1] Turk M, Pentland A, “Eigenfaces for recognition.” Journal of Cognitive Neuroscience 3(1), 1991, 71–86.
[2] Belhumeur P, Hespanha J, Kriegman D, “Eigenfaces vs. fisherfaces: recognition using class specific linear projection.” IEEE, 1997
[3] Ahonen T, Hadid A, Pietika¨ inen M “Face recognition with local binary patterns.” In: Proceedings of European Conference on Computer Vision. pp., 2004, 469–481.
[4] Bicego M, Lagorio A, Grosso E, Tistarelli M,”On the use of SIFT features for face authentication.” In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. pp. 2006, 35–42.
[5] Wright J, Yang A, Ganesh A, Sastry S, Ma Y “Robust face recognition via sparse representation.”, IEEE Transactions on Pattern Analysis and Machine Intelligence 31(2), 2009, 210–227.
[6] T. Kohonen, “Self-organization and Associative Memory”, Springer-Verlag, Berlin, 1989.
[7] M. Kirby and L. Sirovich, “Application of the karhunen-loeve procedure for the characterization of human faces,” IEEE Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pp. 103-108, 1990.
[8] M. Turk and A. Pentland, “Eigenfaces for recognition,” J. Cog. Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.
[9] K. Chang, K. Bowyer, and P. Flynn. Face recognition using 2D and 3D facial data. 2003 Multimodal User Authentication Workshop, pages 25–32, December 2003
[10] H. S. Bhatt, S. Bharadwaj, R. Singh, and M. Vatsa, “Recognizing surgically altered face images using multiobjective evolutionary algorithm,” IEEE Trans. Inf. Forensics Security, vol. 8, no. 1, pp. 89–100, Jan. 2013.
[11] T. I. Dhamecha, R. Singh, M. Vatsa, and A. Kumar, “Recognizing disguised faces: Human and machine evaluation,” PLoS ONE, vol. 9, no. 7, p. e99212, Jul. 2014.
[12] N. Engelhard, F. Endres, J. Hess, J. Sturm, and W. Burgard, “Real-time 3D visual slam with a hand-held RGB-D camera,” in Proc. RGB-D Workshop, Eur. Robot. Forum, vol. 180. 2011.
[13] B. Y. L. Li, A. S. Mian, W. Liu, and A. Krishna, “Using Kinect for face recognition under varying poses, expressions, illumination and disguise,” in Proc. IEEE Workshop Appl. Comput. Vis., Jan. 2013, pp. 186–192.
[14] R. I. Hg, P. Jasek, C. Rofidal, K. Nasrollahi, T. B. Moeslund, and G. Tranchet, “An RGB-D database using microsoft’s Kinect for windows for face detection,” in Proc. 8th Int. Conf. Signal Image Technol. Internet Based Syst., Nov. 2012, pp. 42–46.
[15] Gaurav Goswami, Mayank Vatsa, and Richa Singh, “RGB-D Face Recognition With Texture and Attribute Features” in Proc IEEE Transactions On Information Forensics And Security, Vol. 9, No. 10, October, 2014.