Face Recognition In Presence Of Occlusion Using Machine Learning Classifier

Vandana P, Manjunath C N

Abstract– **Sensor technology is getting advanced day-by-day Which aids for the use of biometric technology in almost every area and wide variety of types of biometric systems are being used to achieve various functionalities. Human face is one of the highly preferred means to authenticate an individual. Face recognition is sensitive to aspects like pose variation, illumination and Occlusion.In this paper, we propose a fully automatic face recognition system which is robust to occlusions. We basically consider two problems: 1)occlusion handling for surface registration and 2) recognition with missing data. Both in the pre-processing and processing stage registration scheme make use of Iterative Closest Point (ICP) technique and in the final classification stage Principal Component Analysis (PCA) is used to deal with incomplete data.Experimental results confirm that registration based on the ICP together with the PCA classification offer an occlusion robust face recognition system.**

***Index Terms*— Face recognition, occlusions, Iterative closest point, Principal component analysis.**

I. INTRODUCTION

In biometric systems, human beings are identified by distinctive features, such as physiological and behavioral characteristics. As a biometric modality, the human face is widely preferred because of several advantages: Due to its contactless acquisition, it is well accepted among users. Furthermore, its applicability to non-cooperative scenarios makes it suitable for a range of applications such as surveillance systems. However, in non-cooperative and uncontrolled scenarios, recognizing individuals from their faces is a challenging task. The factors that degrade the performance of a face recognizer, include presence of illumination differences, in-depth pose variations, facial expression variations, and the presence of occlusions.

II. RELATED WORK

Generally, when the type of occlusion is known, a specific strategy can be used to eliminate the occlusion and asuitable face recognition strategy can be used. For example,park et al. [1] proposed a method which could eliminate theocclusion caused by glasses in the frontal face. Kim et al. [2] have proposed a part-based local representation methodcalled Locally Salient Independent Component Analysis (LS-ICA) that works for specific occlusions. The problem of detecting occlusions in 2D faces has been investigated by Lin and Tang [8]. They derived a Bayesian formulation unifying the occlusion detection and recovery stages. Colombo et al. [3] proposed a fully automated system for face detection and recognition which are occluded by unknown objects where the occluded regions are first detected by considering their effects on the projections of the faces in a suitable face space. A different approach has been investigated by Tarres and Rama [9]. Instead of searching for local non-occluded features, they try to eliminate some features which may hinder recognition accuracy in the presence of occlusions or changes in expression. De Smet et al. [10] proposed an algorithm which iteratively estimates the parameters of a 3D morphable face model to approximate the appearance of a face in a 2D image. Simultaneously, a visibility map is computed which segments the image into visible and occluded regions. Zhang et al. [12] proposed a method based on local Gabor Binary patterns where the face is divided into rectangular region and, on the basis of their gray level histograms, the probability of occlusion is estimated.

III. PROPOSED SYSTEM

The input to the proposed face recognition system is a video stream and the output is identification or verification of the subject. We can define this face recognition system as a 3step process as in Fig.1.The system used in our approach is outlined in Fig.2.

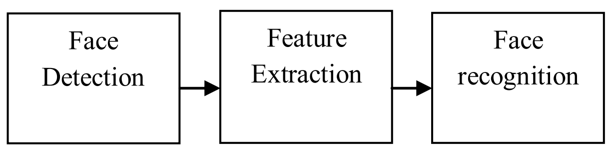

Fig.1. Face Recognition system

ISSN:2278 – 7798

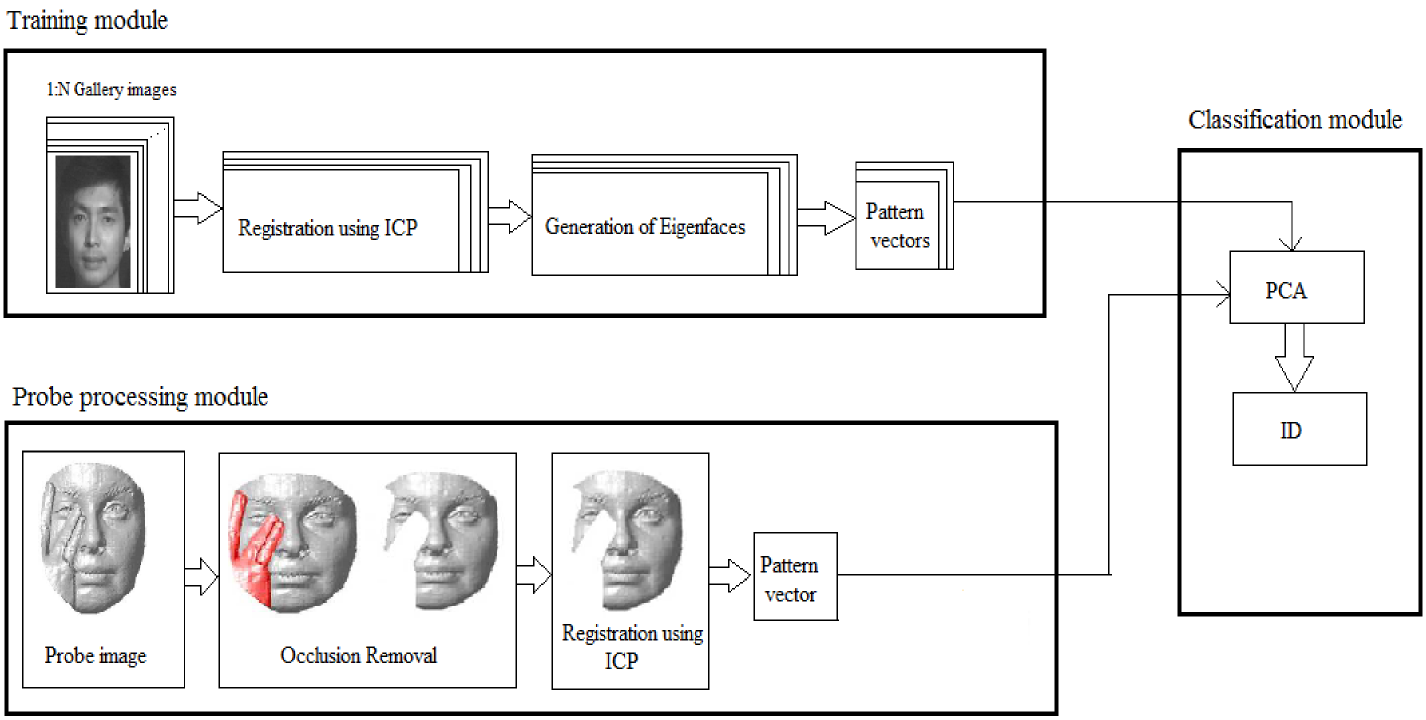

Fig.2. Illustrative diagram of the proposed face recognition approach

A. Face detection

In this system we are using Viola Jones Detection technique. The technique relies on the use of simple Haar-like features in Fig.3 that are evaluated quickly through the use of a new image representation. Based on the concept of an “Integral Image” it generates a large set of features and uses the boosting algorithm AdaBoost to reduce the over-complete set and the introduction of a degenerative tree of the boosted classifiers provides for robust and fast interferences. The detector is applied in a scanning fashion and used on gray-scale images, the scanned window that is applied can also be scaled, as well as the features evaluated.

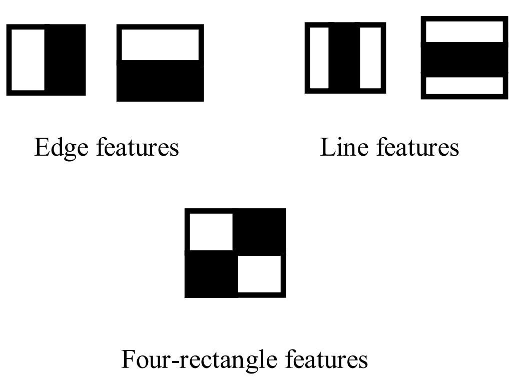

Fig.3 Haar features

B. Face registration

Once generated, candidate face images are normalized in pose and orientation in order to perform a final classification into faces or non-faces. The normalization is computed in two steps. First, a rough normalization is performed using the position of the candidate facial features. A reference position is defined aligning the eyes positions with the x axis and the plane passing through the eyes and the tip of the nose rotated by 45 degrees around the same axis. Finally, the tip of the nose is translated to the origin. In all those cases where one feature is missing (for example in case of occlusions) a degree of freedom is left undetermined. The rough registration generates a good starting position for the following ICP [4] based fine registration. The idea here is to refine the position registering the candidate face with a mean face template. We used a customized version of the ICP algorithm aimed to handle the presence of extraneous objects. The algorithm has been inspired by the variants presented in [5]. The ICP algorithm requires a matching criteria in order to find correspondences between the points of the surfaces to be registered. In our implementation we used a projective matcher . Based on the assumption that the rough registration computes a good registration (i.e. with a low error) between the mean face and the candidate face surface, the projective matcher tries to find correspondences using orthographic projections of each vertex. For each point located at coordinates (i, j) in the data image space, the correspondent point is searched in the model image space in locations (i ± r, j ± r); where r ≥ 1 is an integer defining a square region around the current location. The correspondence criterion is the point at minimum Euclidean distance.

C. Face recognition

In the appearance-based methods, Principal Component Analysis (PCA) is the fundamental technique for image reconstruction. Image reconstruction using PCA computes the principal component scores from a part of the input image and reconstructs the whole input image according to the principal component scores. Using only effective pixels which are not occlusion, occluded regions can reconstructed by using PCA. The reconstruction accuracy depends on the detection accuracy of occluded regions, since it is necessary to accurately detect occluded regions and compute principal component scores only from the effective pixels.

a) Calculating the 'Eigenfaces':

Let each face image, I (x, y) , be a two dimensional w × h= P array of intensity values. Each image, written as a P × vector, represents a point in P-dimensional space. Therefore, the collection of images in the training set constitutes a collection of points in a huge space. Again, due to the similarity of faces in their overall configuration, these points will not be randomly distributed in this immense space, but are likely to be close and occupy only a small portion of the space. Thus, the collection of points can be described by a relatively low dimensional subspace.

Let the training set of images be Γ1,Γ2,….ΓM. The average face of this set is then defined by

$$\bar{\Gamma} = \frac{1}{M}\sum_{i=1}^{M} \Gamma_i$$

Now, let each face differ from the mean face by $\Phi_i = \Gamma_i - \bar{\Gamma}$. For each location in an image, we have one sample for each of the M images. We can study the intensity relations between any two sample points analyzing their covariance. The covariance between points i and j, denoted C_{ij} , can be approximated by

$$C_{ij} = \frac{1}{M}\sum_{K=1}^{M} \Phi_k(i)\,\Phi_k(j)$$

Therefore, the covariance matrix C can be obtained as follows

$$C = AA^T = \frac{1}{M}\sum_{i=1}^{M} \Phi_i \Phi_i^T$$

Where $A = [\Phi_1 \Phi_2 \ldots\ldots \Phi_M]$ By using principle component analysis analysis on the set of large vectors, we have obtained a set of *M* orthonormal vectors $\vec{U}_n$and their associated eigenvalues λ_n , which best describe the spread of the data in the *P*-dimensional space. These orthonormal vectors, i.e. the principle components, turn out to be the eigenvectors of the covariance matrix *C*. The matrix C, however, is P × P and determining the P eigenvectors and eigenvalues is an intractable task for typical image sizes . Nevertheless, when M is very small compared to P, like the case in our experiments, a smaller M × M problem can be solved instead. Consider the eigenvectors V_i of AA^T such that

$$AA^T V_i = \mu_i V_i$$

Multiplying each side from the right by A yields

$$AA^T AV_i = \mu_i AV_i$$

From this we see that AV_i are th eigenvectors of C .
If we let$V = [v_1 v_2 \ldots\ldots\ldots v_m]$be thematrix formed from the eigenvectors of $AA^T U = [u_1 u_2 \ldots\ldots\ldots u_p]$be the matrix formedfrom the eigenvectors ofA^T, then U=AV.

Although *M* eigenvectors ("eigenfaces")are necessary to encode each image of the trainingset without loss of information, *M' <M* aresufficient enough for recognition. Therefore, fromthe *M* eigenvectors of *V*, we pick the *M'*eigenvectors that account for the most variation,i.e. the *M'* eigenvectors having the highesteigenvalues.

b) Identifying Faces:

Once the basis vectors for the "face space" have been constructed, all that remains is to project all the images in the training set onto the "face space". Any image G can be expressed in terms of the *M'* "eigenfaces", using *M'* weights calculated as follows

$$W_k = U_k^T(\Gamma - \bar{\Gamma})$$

These *M'* weights form a vector $U^T = [W_1 W_2 \ldots\ldots W_M]$, quantifying the contribution of each of the "eigenfaces" in representing the input face image, treating the "eigenfaces" as a basis set for the face images. The weight vector is then used to determine which of the predetermined number of faces matches best the query image. The easiest method to identify an image in the training set that provides the best description of the input image is to choose the face image that minimizes the Euclidean distance between weight vectors, that is

$$\varepsilon_k = |(\Omega - \Omega_k)|$$

whereΩ_k is the vector describing the kth image in the training set. The algorithm proposed by Turk and Pentland makes the distinction between a "face class" and a face image. A "face class" consists of the collection of face images belonging to anindividual, and in the case where a "face class" contains more than one image, $\Omega_k$is calculated as the average of the weights obtained when projecting each image of a class k. For reasons that we will explain soon, we chose to have only one image per "face class". The query image belongs to an individual k in the training set only when the minimum ε_k is below a threshold θ_ε . On the other hand, if the minimum distance is above this fixed threshold, then the queried face is classifiedas unknown to the system. The validity of the threshold relies on the assumption that faces from the same person map close to each other in the "face space". In other words, it relies on the assumption that the "face space" consists of a series of small clusters distant from each other, with each cluster representing faces from one individual, and where each cluster approximately has the same dimensions. However, several times in our experiments, two sample faces from two different people mapped closer onto the "face space" than two faces from the same person, thus breaking the cluster assumption. Therefore, we could not establish the face identification threshold as proposed originally. Images from the same individual seemed scattered across the "face space", hence we chose to only put one image per face class. We fixed the value of the threshold "instinctively" so as to get a reasonable ratio of correct/incorrect positive identifications. Any image, given it has the right dimensions, can be projected onto the "face space". More specifically, any input image can be more or less approximated as a linear combination of the "eigenfaces". Since the *M'* largest eigenvectors were chosen to span the *M'*-dimensional "face space", they capture the most variation for the face images. This implies that projecting any non-face image onto this *M'* subspace is likely to yield weights for which no face would have mapped to.

IV. EXPERIMENTAL RESULT

A. Real time video is given as input

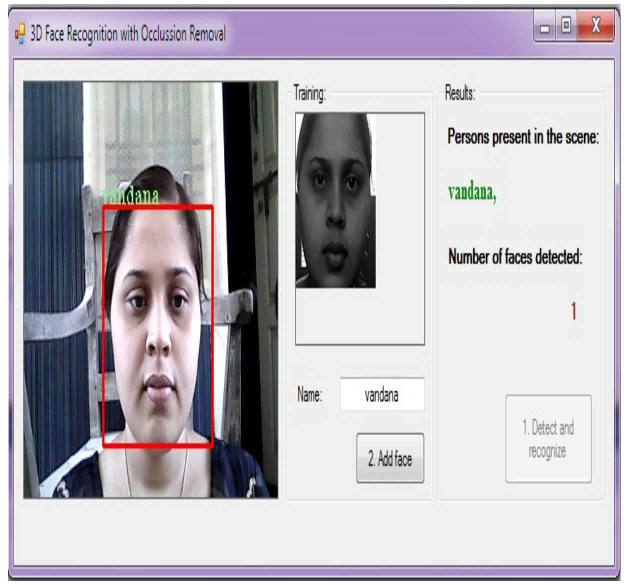

Fig.4.1 snapshot of input Video

We are providing video as a input to our system in the real time and Fig.4.1 is the snapshot of the given video at particular time instant .We can also notice in the Fig.4.1 that the detected face in the video as been marked by Red square box.

B. Recognized face even in the presence of Occlusion

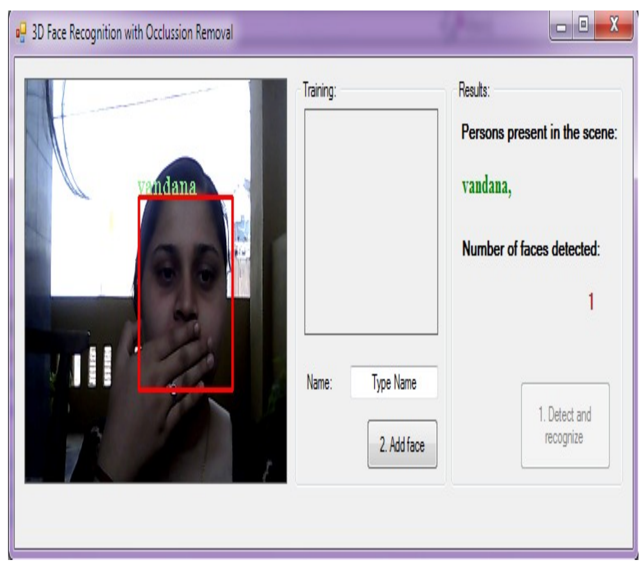

Fig.4.2 Occlusion of mouth area by hand

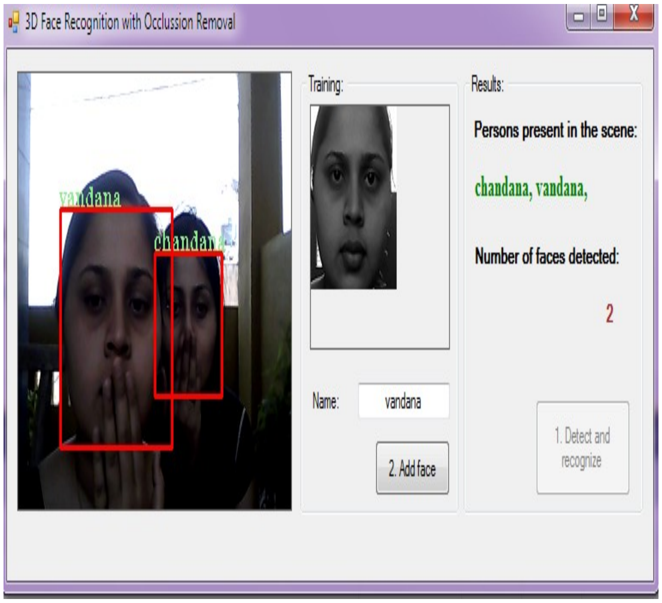

Fig.4.3 Multi face Recognition under occlusion

Fig.4.2 and Fig.4.3 shows the successful recognition of face even when some parts of the face are occluded or hidden by exterior object(by hand in this case).We can notice the correct name of the person being displayed even when the face has been occluded.

V. CONCLUSION

In this study, occlusion problem, which has been researched relatively less than illumination and pose problems in face recognition, is discussed. Locally occluded areas in faces are detected using thresholding the difference map obtained by computing the absolute difference between the average face and the input face and hence classified as occluded and non-occluded face images.

Following the occlusion detection stage, the facial parts detected as occluded are removed to obtain occlusion-free surfaces. Classification is handled on these occlusion-free faces. In this work, we are using PCA algorithm for classification. Experimental results demonstrated that the proposed algorithm could reliably recognize partially occluded faces with higher recognition rate than the existing methods.

REFERENCES

[1] Park, J.S., Oh, Y.H., Ahn, S.C., Lee, S.W.: Glasses removal from facial image using recursive error compensation.IEEE Trans. Pattern Anal. Mach. Intell. 27(5), 805–811 (2005)

[2] Kim, J., Choi, J., Yi, J., Turk, M.: Effective representation using ica for face recognition robust to local distortion and partial occlusion.IEEE Trans. Pattern Anal. Mach. Intell. 27(12), 1977–1981(2005)

[3] Colombo, A., Cusano, C., Schettini, R.: A 3d face recognition system using curvature-based detection and holistic multimodal classification. In: Proc. 4th Int’l Symp. On Image and Signal Processing and Analysis, pp. 179–184(2005)

[4] Paul J. Besl and Neil D. McKay, “A Method for Registration of 3-D Shapes”, IEEE Transactions on PAMI, vol 14, pp. 239–256, 1992.

[5] Rusinkiewicz, S., Levoy, M., ”Efficient Variants of the ICP Algorithm”,Third International Conference on 3D Digital Imaging andModeling (3DIM), 2001

[6] NeseAlyuz, BerkGokberk, and LaleAkarun, “3-D Face Recognition Under Occlusion Using Masked Projection” (2013).A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, “2D and 3D face recognition: A survey,” Pattern Recognit. Lett., vol. 28, no. 14, pp. 1885–1906, 2007.

[7] Lin, D., Tang, X.: Quality-driven face occlusion detection and recovery. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1–7 (2007)

[8] Tarrés, F., Rama, A.: A novel method for face recognition under partial occlusion or facial expression variations. In: Proc. 47th Int’l Symp. ELMAR, pp. 163–166 (2005)

[10] De Smet, M., Franses, R., Van Gool, L.: A generalized em approach for 3D model based face recognition under occlusions. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 1423–1430 (2006)

[11] Zhang, W., Shan, S., Chen, X., Gao, W: Local Gabor binary patterns based on KullbackÓleibler divergence for partiallyoccludedface recognition. IEEE Signal Process. Lett. **14**(11), 875–878 (2007)

Vandana P
The author is currently pursuing M.tech in Digital Electronic and Communication Engineering at Akshaya Institute of Technology, Tumkur , Karnataka , affiliated to Visvesvaraya Technological University.

Manjunath C N
The author is Lecturer in Dept. of Electronic and Communication Engineering at Akshaya Institute of Technology, Tumkur , Karnataka , affiliated to Visvesvaraya Technological University.

