## ABSTRACT

Undoubtedly homeland security is a noteworthy concern in today's increasingly connected world and there is a bevy of IT-based security solutions and services emerging and evolving to guarantee the safety and security of people and properties. Security and surveillance cameras are the prominent security solutions. People movements, gestures, and activities are being minutely monitored, captured as images and videos, and subjected to a variety of investigations in order to extract anything uncommon. As videos capture the movement, there is no requirement for the user to coordinate with the camera; video arrangements need not contain face images alone despite the fact that human appearances are vital articles in video sequences. Henceforth recognising a face from the sequence of a video turns into an essential undertaking in PC vision applications.

As the user is not totally coordinating with the camera, it is not necessary that the face captured in the video sequence is the frontal face. At times, it need not be a human face. Hence, it is necessary to detect the human face region in a frame. If a human face is detected in a video sequence, there are chances that the detected face can be masked. It is also possible that the faces can be captured in various poses. There is also a difficulty in identifying a face subjected to various illuminations. The above natural possibility of not capturing the frontal face of a human in a frame makes the identification task difficult. This problem prompts researchers working under face recognition technology to design an improved framework that increases the recognition rate in the above situations. From the above-mentioned possibilities that deteriorate the recognition rate of a face, in this thesis work, two major problems namely face with

varying pose and partially occluded faces are considered for recognition from a captured video.

The major challenge with this biometric is a way to handle pose variation between the input face image and gallery set images. When faces are captured in a video, there is a possibility that captured face can be oriented at any angle in the X-Y or Y-Z plane. By taking into account this preliminary idea, in this work in order to calculate the angle at which a face is oriented when captured, a geometric approach is considered. In this approach, the eye points of the face and nose tip are taken into consideration. Assuming that these points form a triangle; with respect to a frontal face, further study is carried in order to determine the angle of orientation of a face. In this case, three scenarios are considered to capture the angle of orientation of the face namely, considering the X-Y plane with rotation of the face along Z-axis in clockwise and anti-clockwise direction, considering the Y-Z plane with rotation of the face along X-axis in clockwise and anti-clockwise direction and considering X-Z plane with rotation of the face along Y-axis in clockwise and anti-clockwise direction. Depending on the angle of orientation of the face with respect to the geometrical points considered, the input face image is matched with the gallery sets. To extract the feature points, Curvelet transform is used. The Curvelet Transform is coined because they are good at extracting the edge features. The extracted features are classified using minimum distance classifier. With this, the recognition rate of a valid face image and the time required for recognition is calculated. MatLab is used to analyse the recognition rate and time for recognition. The method described is compared against the state of art methods for a comparative analysis. To analyse the above metric, two existing datasets namely YouTube dataset and Honda/UCSD datasets are used. These

datasets include subjects with varying poses. The result shows an improvement in recognition rate when compared with the existing pose invariant face recognition techniques and a decrease in time for recognition.

One of the unsolved problems with video-based face recognition system is recognising faces with occlusion. Faces captured in a video can be partially occluded either by an accessory or by some other object. This leads to the loss of feature values of the face and, in turn, deteriorates the recognition rate. Recovering lost region of a face is a tedious task. Inpainting is a technique used in recovering the lost region of an image by diffusing data from the known region. Criminisi algorithm that is based on exemplar-based inpainting gives the best result compared to existing inpainting techniques. This method chooses a patch from the contour of the missing region to be replaced by the patch on a known region depending on the confidence term and the data term of the patch. Confidence term defines the number of known pixels around a patch and data term is the gradient value to determine the most important pixel that is to be considered for filling. In this work, a modified exemplar method is proposed, where sobel operator is used to determine the gradient information. A video sequence is considered as the input. Occlusion is detected from a face image. Once occlusion is detected, the image is classified into two regions, known and unknown respectively. From the contour pixels of the unknown region, a most prioritised pixel is chosen based on the number of the known pixel within a patch considering the pixel at the centre. In this case, a 9×9 patch is considered. A pixel with the most number of known pixels and if the pixel is on any of the edges, then that pixel is given maximum priority. Once the most prioritised pixel enclosed in a patch of size 9×9 is chosen, the most similar patch need to

be selected from the known region. To check the most similar patch, the structural similarity measure is used. Once a similar patch is detected from the known region, the patch on the contour with the prioritised pixel is replaced with the most similar patch. Continuing with the same procedure, fills in the occluded region. The reconstructed face is checked against the gallery images. Features are extracted using curvelet transform. Minimum distance classifier is used to check for the matching face in the gallery. The experiment is tested on YouTube dataset. The result shows a better recognition rate when hair and glasses are considered while comparing with the existing methods. The time for recognition is reduced for this method compared to the state of art methods.