ABSTRACT
Face recognition refers to an individual's understanding and interpretation of the human face especially in relation to the associated information processing in the brain. Autism Spectrum Disorder (ASD) is a comprehensive neural developmental disorder that produces many deficits including social, communicative and perceptual. Individuals with autism exhibit difficulties in various aspects of facial perception, including facial identity recognition and recognition of emotional expressions. Autism Spectrum Disorders (ASD) are characterized by atypical patterns of behaviours and impairments in social communication. Traditional intervention approaches often require intensive support and well-trained therapists to address core deficits. People with ASD have tremendous difficulty accessing such care due to lack of available trained therapists as well as intervention costs. Thus a Human Facial Emotions based image processing system is to be developed which processes autistic people’s expressions and enables them to access PC applications based on their expressions.


I. INTRODUCTION
Research on Facial Emotion Recognition (FER) is a very challenging field that targets methods to make Human Computer Interaction (HCI) effective. Facial expressions not only to express our emotions but also to provide important cues during social interactions such as level of interest, our desire to take a speaking turn and to provide continuous feedback on the understanding of the information conveyed. Identification and classification of emotions by computers has been a research area since Charles Darwin’s age. Facial emotion recognition is a field where lot of work has been done and a lot more can be done.

Face recognition refers to an individual's understanding and interpretation of the human face especially in relation to the associated information processing in the brain. Autism Spectrum Disorder (ASD) is a comprehensive neural developmental disorder that produces many deficits including social, communicative and perceptual. Individuals with autism exhibit difficulties in various aspects of facial perception, including facial identity recognition and recognition of emotional expressions.

Image processing is a rapidly growing area of computer area of computer science. Its growth has been fuelled by the technological advances in digital imaging, computer processors and mass storage devices.

There are five stages in any digital image processing application. They are broadly classified as:

- Image Acquisition
- Image Pre-processing
- Image Segmentation
- Features Extraction
- Classification and Prediction

A. Image Acquisition
The image is captured by a sensor (e.g. Camera), and digitized if the output of the camera or sensor is not already in digital form, using analogue-to-digital convertor.

B. Image Pre-processing
Pre-processing methods use a small neighbourhood of a pixel in an input image to get a new brightness value in the output image. Such pre-processing operations are also called filtration. Image pre-processing tool uses many useful pre-processing operations which suppress information that is no relevant to the specific image processing and enhance some image features important for further processing.

Some of the pre-processing includes image enhancement, cropping, de-noising, etc.

C. Image Segmentation
Segmentation is the process of partitioning an image into non-intersecting regions such that each region is homogeneous and the union of no two adjacent regions is homogeneous. The goal of segmentation is typically to locate certain objects of interest which may be depicted in the image. Segmentation could therefore be seen as a computer vision problem.

There are four popular segmentation approaches: threshold methods, edge-based methods, region-based methods and the connectivity-preserving relaxation methods.

D. Features Extraction
Feature extraction is a special form of the dimensionality reduction. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. Feature extraction methods can be supervised or unsupervised, depending on whether or not class labels are used.

Among the unsupervised methods, Principal Component Analysis (PCA), Independent Component Analysis (ICA), Multi-dimensional scaling (MDS) are the most popular ones. Supervised FE methods (and also FS methods) either use information about the current classification performance called wrappers, or use some other, indirect measure, called filters.

E. Classification and Prediction
This final step involves classification of segmented image under various labels based on the features generated. This classification is done using the various data mining techniques. Classification consists of assigning a class label to a set of unclassified cases. There are two different classes of classification. They are supervised classification where the set of possible classes is known in advance and unsupervised classification where the set of possible classes is not known and after classification we can try to assign a name to the class. Unsupervised classification is also known as clustering.

The paper is organized as follows. Literature survey is presented in Section II. The proposed work is...
presented in Section III. The implementation and results are analysed in Section IV and future work is presented in Section V. We conclude in Section VI.

II. LITERATURE SURVEY
This section we will discuss some methods which are presently used for the human facial emotion recognition along with their advantages and disadvantages. The methods are explained below with their features and drawbacks.

Global face recognition methods are based on statistical approaches wherein features are extracted from the entire face image. In this, every element in the feature vector refers to some global characteristics of face images. Subspace based methods, spatial-frequency techniques, and moment based methods are examples of most frequently used global methods. Among subspace-based methods, Principal Component Analysis (PCA), Fisher Linear Discriminant (FLD), Two-Dimensional PCA (2DPCA), and Two-Dimensional Two-Directional PCA (2D2PCA) are the most widely used and successful face recognition approaches as discussed by Chandan Singh, Ekta Walia and Neerja Mittal (2011). The spatial-frequency techniques such as Fourier transform as discussed by Singh C and Walia E (2010) and Discrete Cosine Transform (DCT) as discussed by Soyle H and Demirel H (2010) are useful in extracting the facial features at some preferred frequency. In these methods, firstly the images are transformed to the frequency domain, and thereafter, the coefficients of low frequency band are taken as the invariant image features. Furthermore, the moment invariants are the most widely used image descriptors in many pattern recognition applications such as character recognition, palm print verification, etc. Some of these moments as discussed by Neerja and Walia. E (2008) such as Hu moment invariants and radial moments such as the Zernike moments (ZMs), pseudo Zernike moments (PZMs), and orthogonal Fourier–Mellin moments possess the property of being invariant to image rotation and can be made invariant to translation and scale after applying the geometric transformations.

Although the global face recognition techniques are most common and well-liked in face recognition, recently lots of work is being done on local feature extraction methods as these are considered more robust against variations in facial expressions, noise, and occlusion. These structure-based approaches deal with local information related to some interior parts of face images, i.e., features of nose patch, distance between the eye-centers, mouth width or height, etc. These methods can be classified in two categories: firstly, the sparse descriptor which initially divides face images into patches and then illustrates its invariant features and, secondly, a dense descriptor for scale, illumination, noise and rotation variations. The experimental results devoted to the texture analysis and face detection prove the robustness of this descriptor for scale, illumination, noise and rotation variations.

III. PROPOSED METHOD
For this study, video containing static images of different human beings with different facial expressions is considered.

The flow of the proposed system is represented in the figure 1 below:

Fig 1: Flowchart of the Proposed System

Methodology

1) Image acquisition
Every image processing application always begins with image acquisition. The images of different human beings with different facial expressions are considered as input. The images are saved in JPEG format.
format. The facial images are captured using a webcam.

2) Image pre-processing
Image pre-processing creates an enhanced image that is more useful or pleasing to a human observer. The image pre-processing steps used in the system are: a) Filtering of the image and b) Skin Tone Detection.

a) Filtering of the image
Filtering in image processing is a process that cleans up appearances and allows for selective highlighting of specific information. The Prewitt operator is used in image processing, particularly within edge detection algorithms. Technically, it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the Prewitt operator is either the corresponding gradient vector or the norm of this vector. The Prewitt operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively insensitive in terms of computations. The purpose of smoothing is to reduce noise and improve the visual quality of the image.

b) Skin Tone Detection
A skin detector typically transforms a given pixel into an appropriate colour space and then uses a skin classifier to label the pixel whether it is a skin or a non-skin pixel. A skin classifier defines a decision boundary of the skin colour class in the colour space. Important challenges in skin detection are to represent the colour in a way that is invariant or at least insensitive to changes in illumination. Another challenge comes from the fact that many objects in the real world might have skin-tone colour. This causes the skin detector to have much false detection in the background. The simplest way to decide whether a pixel is skin colour or not, is to explicitly define a boundary. RGB matrix of the given colour image converted into different colour space to yield distinguishable region of skin or near skin-tone.

3) Image Post-Processing
Once the image has been enhanced and segmented, the interesting part can be extracted and features can be analysed. The feature statistics include mean, variance, range, quantile maximum, quantile minimum, and quantile range. The quantile features were used instead of the maximum, minimum, and range because they tend to be less noisy. The pitch features were extracted only over the voiced regions of the signal. The video motion-capture derived features were occasionally missing values due to camera error or obstructions. To combat this missing data problem, the features were extracted only over the recorded data.

4) Feature Extraction
In feature extraction, Weber’s Law Descriptor (WLD) based on Weber’s Law is used. It represents an image as a histogram of differential excitations and gradient orientations, and has several interesting properties like robustness to noise and illumination changes, elegant detection of edges and powerful image representation.

5) Classification
The features classification is done by the Fuzzy C-Means (FCM) classifier. Fuzzy clustering plays an important role in solving problems in the areas of pattern recognition and fuzzy model identification. It uses reciprocal distance to compute fuzzy weights. The left eye, right eye and lips of the human being in the image, which are the key features needed to deduce a facial expression, are extracted.

6) Edge Detection
The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient magnitude at each point in an input gray scale image. Mathematically, the operator uses two 3×3 kernels which are convolved with the original image to calculate approximations of the derivatives - one for horizontal changes, and one for vertical. The result of the Sobel operator is a 2-dimensional map of the gradient at each point.

7) Database Training
Naive Bayesian Classifiers have been used for Database Training along with C4.5 algorithm. This method is important for several reasons. It is very easy to construct, not needing any complicated iterative parameter estimation schemes. This means it may be readily applied to huge data sets. It is easy to interpret, so users unskilled in classifier technology can understand why it is making the classification it makes. It is robust and is known to do quite well.

8) Emotion Detection – PC Access
Based on final feature set analysis, different kinds of facial emotions are identified using classification techniques and related PC (Personal Computer) applications will be processed. In our application, four different kinds of emotions are analyzed i.e. Smile, Surprise, Sad and Neutral which will help Autism People to survive on their own.

IV. IMPLEMENTATION AND RESULTS
The screenshots of the application after each operation is as follows:

1) Prewitt Filtering
2) Skin Tone Detection
3) **Face Detection**

![Fig 4: Facial region detection](image)

4) **Features Extraction**

![Fig 5: Extraction of key features](image)

Figure 4 and 5 show the output of the Face Detection and Features Extraction operations respectively.

5) **Edge Detection**

![Fig 6: Sobel Operator-Edge Detection](image)

Figure 6 shows the output of Edge Detection process.

6) **Emotion Detection**

The database contains 50 actors (25 male, 25 female) of age 20 to 50 in 4 classes of emotions and each image of size 8 KB in JPEG format. Based on the emotions detected, relevant PC applications can be accessed. Figure 7 shows how an emotion is detected and Figure 8 shows how the related application is accessed.

![Fig 7: Emotion Detection-Smile](image)

![Fig 8: Related PC Applications access-MS Paint](image)

### V. FUTURE WORK

In this paper, we have studied the four basic emotions. As a future enhancement we would study and implement further emotions. Further, continually changing emotions in motion pictures can be sorted as a future enhancement. Further, a mobile application can be developed.

### VI. CONCLUSION

Image processing technique plays an important role in the detection of the human facial emotions. Virtual Reality (VR)-based facial expression system is to be developed that is able to collect eye tracking and peripheral psycho-physiological data while the subjects are involved in emotion recognition tasks. It enables individuals with Autism Spectrum Disorder (ASD) to access PC by processing nonverbal communication in a virtual reality environment in spite of their potent impairments.

### REFERENCES