International Journal of Science Engineering and Advance Technology, IJSEAT,Vol.3,Issue 4 | ISSN 2321-6905 March 2015

International Journal of Science Engineering and Advance Technology

# Emotional Based Smart Accessing and Controlling For Autism Spectrum Disordered People

**M.Indumathi[1], D.Durai kumar[2]**

[1] M.Tech- Information Technology, Ganadipathy Tulsi's Jain Engineering College, Vellore, Tamil Nadu, India.

[2]Associate Professor, Dept. Of Information Technology, Ganadipathy Tulsi's Jain Engineering College, Vellore, Tamil Nadu, India.

## ABSTRACT

Autism (or) Autism Spectrum Disorder (ASD) may be viewed as a neural developmental disability that can affect social interaction, language (or) behavioural skills of a person. Most autistic persons show symptoms of withdrawal from social interaction and a lack of emotional empathy towards others. This behaviour is usually attributed to their inability in understanding or expressing emotions. The underlying causes of autism spectrum disorder (ASD) are still not well understood but an alarming number of persons are diagnosed with this disorder. Among the fundamental social impairments in the ASD are challenges appropriately recognizing and responding to nonverbal cues and communication. Several existing assistive devices mostly serve as remedial tool that provide a learning environment for autistic persons to learn about the norms of social behaviour. However this tool lacks the capability to operate in conjunction with real world scenarios. We propose a new intervention paradigm that act as a portable system called facial expression recognition system that recognizes Virtual Reality (VR) based facial expressions in a synchronous manner and also to break the dependency of an autistic person by enhancing expression based accessing and controlling process in this modern environment.

## 1. Introduction

Autism may be viewed as a neural developmental disability that can affect social, language or behavioural skills of a person. The underlying causes of Autism Spectrum Disorders (ASD) are still not well understood but an alarming number of persons are diagnosed with this disorder. These disorders are characterized by difficulties in verbal and non-verbal communication, and restricted behaviour. It is otherwise called as pervasive development disorder (or) neuro development disorder.

The human face is an important human body part which plays an extraordinary role in the human to human or human to machine communications. It is important to design robust emotion detection system for real world applications like human decision making and effective human computer interaction. Our project mainly focuses on human computer interaction via facial expressions. Social interaction is the use of non-verbal or verbal behaviour in order to interact with people. Which involve eye contact, speech, gestures and facial expressions used to initiate and respond to interaction with other. It can be classified as

(i) **Verbal communication**: speech or spoken languages are used to share ideas, exchange information and regulate interactions.

(ii) **Non verbal communication:** express emotions without use of words. Messages can be communicated through body language or posture, gestures and touch, facial expression and eye contact.

Autistic people face tremendous difficulties in understanding social cues and conventions; they are unable to properly express non-verbal communication and body language. These inabilities hinder them from understanding verbal and non-verbal communications, as well as reading human facial expressions effectively.

### 1.1 Image processing system

The process of approaching image processing system is followed by the steps of

**Image Acquisition:** Image sensor captured by visual data.

**Discretization / Digitization, Quantization, compression**: Convert data in to discrete form and compress for efficient storage / transmission

**Image enhancement and restoration:** Improving image quality i.e low contrast, blur, noise

**Image segmentation:** Partition image in to objects (or) constituent parts

**Feature selection**: Descriptors from an image those are important for differentiating one class of objects from another

**Image representation:** Labels the object based on information provided by descriptors

**Image interpretation**: Labels meaning to an ensemble of recognized objects

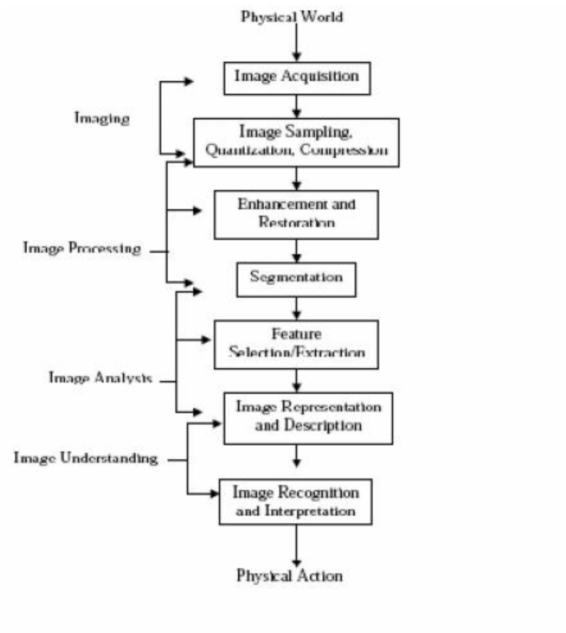


Fig 1.1 steps in an image processing

## 2 .SYSTEM ARCHITECTURE

### 2.1 Block diagram

Initially an image is captured by web camera; face detection is done using skin tone detection by finding skin-colored pixels and regions in an image then converting the image into the RGB color space and classifies to either skin or non-skin those results in much stretched skin color cluster image. The face clipped image takes as input for feature extraction that reduces dimensionality as well as extracts both transient and intransient features i.e. eyes, eyebrows and mouth to perform the desired task. In this module Weber's law descriptor is used to represent an image as a histogram of differential excitations and gradient orientation based on Weber's law. After applying a textual descriptor, facial features become darker. For expression analysis, intransient features are selected for the final emotion-specific feature set.

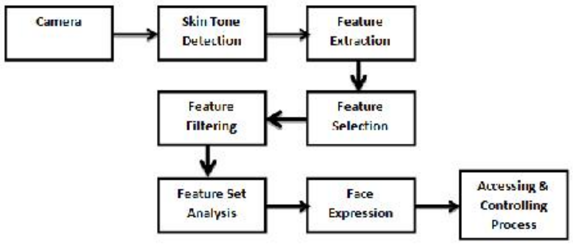


Fig 2.1 VR-based facial expression recognition system

The feature set was reduced using information gain on per emotion class basics and permit as ranking of the features was implemented using fuzzy C means clustering. Then filtered feature set is analyzed and classified to maintain trained dataset .Finally, VR-based facial expressions are recognized there by emotions are identified for accessing and controlling in the modern environment.

### 2.3 Architectural design

When an image is captured by a web camera; it can be further processed to identify different kind of facial emotions such as surprise, smile, sad, happy, ambiguous etc and also to break the dependency of an autistic persons by enhancing expression based accessing and controlling in modern environment.

In skin tone detection, it can isolate the presence of faces, arms, hands and gestures of the autistic person. Skin detector typically transforms a given pixel in to RGB color space and then uses skin classifier to label the pixel whether it is skin or non-skin pixel there by confirm face acquisition through common features on the face.RGB color space also eliminates the influence of varying illumination to the best extent. Then face clipped image is processed by texture descriptor known as Weber's local descriptor (WLD) exploits textual nature of human face as well as relationship between component features via eyes, nose and mouth to detect face patterns and reduce the dimension of the feature space.

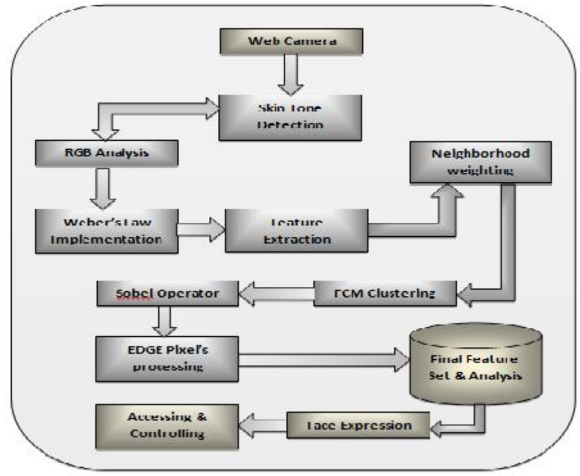


Fig 2.2 Architectural diagram-automatic emotion recognition system

This descriptor makes features darker for efficient feature extraction. Then feature extraction module partition the image in to constituent parts to extract facial features. The end result of extraction task is a set of features commonly called feature vector which constituents a representation of the image.

Then feature vector can be coarsely classified in to transient features (cheeks, nose, and ears) and intransient features (eyes, eyebrows, mouth) which are analyzed by neighbourhood weighting. Then feature selection is used to reduce the feature space which improves prediction accuracy and minimizes computation time. The goal of feature selection is to choose only a subset of intransient features by eliminating unnecessary features. The selected intransient features are used for finding out the facial expression vectors by calculating their difference from average image of the facial feature. Then image representation can be done based on fuzzy c means clustering (FCM) algorithm by assigning labels to an object based on information provided by descriptors (or) pertinent features. Then feature filtering is done via sobel operator to detect edges of intransient features that significantly reduce the amount of data and filters out useless information while preserving the important structural properties in an image. Sobel operator integrates smoothing and gradient calculation for detecting edges.

After edge detection, edge pixels are processed to maintain final feature set. The final feature set included the top 85 features for each emotion class. Based on final feature set analysis different kinds of facial emotions is identified using classification technique. Finally, a smart audio application is played corresponding to the emotion present on the face. For a happy mood, some hearty songs are played to maintain it and for a sad mood some rejuvenating ones are played.

## 3. CONSTRUCTIVE ALGORITHM FOR FACE RECOGNITION

### 3.1 skin colour tone detection

When an image has been captured by camera, it is further pre-processed using skin tone detection. A skin detector typically transforms a given pixel into an appropriate color space and then uses a skin classifier to label the pixel whether it is a skin or a non-skin pixel. Skin tone color detection is a technique that can isolate the presence of faces, arms, hands and gestures. After isolation, skin-colored pixels and regions in an image are identified then converting the image into RGB color space in which skin pixel can be clustered as compact as possible as well as eliminate the influence of varying illuminations to the best extent.

### 3.1.1 Normalizing colour

In RGB-encoded digital image, the color of a pixel is represented by the values ranging from 0 to 255, represent one color channel-red, green and blue. These values are affected by the condition of "illumination" then the RGB encodings can be normalized as

Intensity $I= (R+G+B)/3$

Normalized red $r=R/(R+G+B)$

Normalized green $g=G/(R+G+B)$

Normalized blue $b=B/(R+G+B)$

Fig 3.1 Normalizing colors on the image

### 3.1.2 Apply a skin color model

It translates the normalized R, G, B values to RGB values in [0,255] space if the color falls in the range or model otherwise it sets target pixel to black. From applying skin color model, it eliminates background, cloths and unwanted regions [cleanup image a little for further processing]

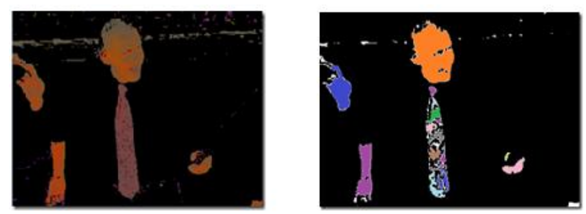

Fig 3.2 Mark pixels on the image using established skin color model

### 3.1.3 Remove regions that are less likely face

Binary erosion is applied to remove some small image then label the connected components of the image that labelling process identifies disconnected regions on the image. Bounding boxes of each region eliminates the regions that are most unlikely to be a face by comparing width-to-height ratio of the regions.

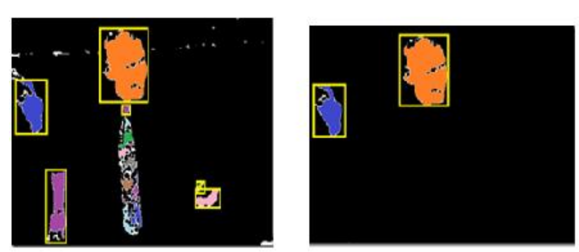

Fig 3.3 Remove regions those are unlikely to represent faces

### 3.1.4 Finding face

Hole-counting algorithm has been used. The hole indicates location of eye sockets and mouth of face region has to contain at least three holes. Human skin usually presents a very smooth pattern so likelihood of introducing artificial holes during image processing is not high.

Fig 3.4 Confirm face acquisition

### 3.2 Feature filtering

Weber's local descriptor (WLD) act as a texture descriptor plays an important role as it exploits textural nature of the human face via human visual system ( HVS) as well as relationship between component features(eyes, nose and mouth) to detect and recognize faces. this descriptor represent an image as a histogram of differential excitation and gradient orientations and has several interesting properties like robustness to noise and illumination changes ,powerful image representation etc. finally, facial features become darker and used for feature extraction analysis. The end result of extraction task is a set of features commonly called Feature vector which constitutes a representation of the image. This feature vector is further classified as transient (wrinkles, cheeks and nose) and intransient features (eyes, mouth, eyebrows) to reduce the dimension of the feature space as well as computational complexity.

Ernst Weber, an experimental psychologist in the 19th century, shows relationship can be expressed as:

$$\frac{\Delta I}{\Delta} = k$$

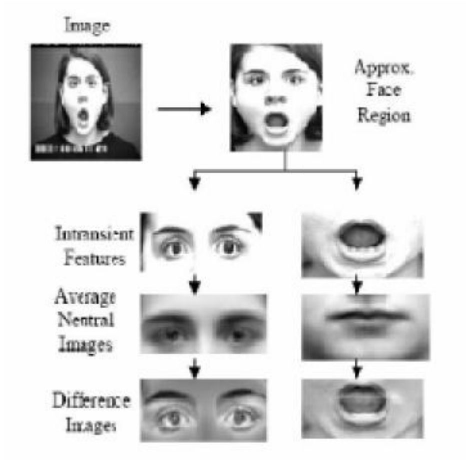


Fig 3.5 Facial feature extraction

Where ΔI represents the increment threshold .I represents the initial stimulus intensity and k signifies that the proportion on the left side of the equation remains constant despite variations in the I term. The fraction ΔI/I are known as the Weber fraction. Weber's Law, more simply stated, says that the size of a just noticeable difference (i.e., ΔI) is a constant proportion of the original stimulus value.

The computation of WLD descriptor involves three steps

1. Finding differential excitations

2. Finding gradient orientations

3. Building the histogram

### 3.2.1 Differential excitation:

For calculating differential excitation (x) of a pixel first intensity differences of with its neighbours xi, i = 1, 2… p are calculated as follows:

$$\Delta I_i = I_i - I_c$$

Then the ratio of total intensity difference of xc with its neighbours xi to the intensity of xc is determined as follows:

$$f_{ratio} = \sum_{i=0}^{p-1} \left(\frac{\Delta I_i}{I_c}\right)$$

Arctangent function is used as a filter to enhance the robustness of WLD against noise which results in:

$$\varepsilon(x_c) = \arctan\left[\sum_{i=0}^{p-1} \left(\frac{\Delta I_i}{I_c}\right)\right]$$

The differential excitation may be positive or negative. The positive value indicates that the current pixel is darker than its surroundings and negative value means that the current pixel is lighter than the surroundings.

### 3.2.2 Gradient orientation

Next main component of WLD is gradient orientation. The gradient orientation is calculated as follows:

$$\theta(x_c) = \arctan\left[\frac{I_{73}}{I_{51}}\right]$$

Upper pixel shows the intensity difference of two pixels on the left and right of the current pixel and lower pixel shows the intensity difference of two pixels directly below and above the current pixel,

### 3.2.3 Building histogram and histogram matrix

After calculating differential excitation and gradient orientation WLD descriptor is build. Corresponding to each gradient orientations, differential excitations are organized as histogram.Then each histogram Ht: t = 0, 1, 2… T-1 is evenly divided into M sub histograms Hm, t: m = 0, 1, 2… M-1, each with S bins. These histograms form a histogram matrix, where each column corresponds to a dominant direction Each row of this matrix is concatenated as a histogram Hm = { Hm,t: t = 0, 1, 2, …, T-1}. Subsequently, histograms Hm: m = 0, 1, 2… M-1 are concatenated into a histogram H = { Hm: m = 0, 1, 2, …, M-1}. This histogram is referred to as WLD descriptor.

### 3.3 Feature Selection

### 3.3.1 Fuzzy C Means Clustering (FCM) Algorithm

Feature selection helps to reduce the feature space which improves prediction accuracy and minimizes the computation time. It can be done based on clustering process. Clustering is the process of partitioning or grouping a given set of unlabelled patterns into a number of clusters such that similar patterns are assigned to one cluster. There are two main approaches to clustering-crisp (hard) and fuzzy (soft) clustering. In crisp clustering, boundary between clusters is fully defined but in some cases, the boundaries between clusters cannot be defined clearly. Some patterns may belong to more than one cluster where as fuzzy clustering provides better and more useful method to classify these patterns.

To remove prevalent, redundant and noisy features (i.e.) it selects the subset of features that can achieve the best performance in terms of accuracy and also performs dimensionality reduction using fuzzy c means clustering (FCM) algorithm. It is based on the concept of “fuzzy c-partition”. It allows one piece of data that belong to two or more clusters then data are bound to each cluster by means of membership function which represents fuzzy behaviour of algorithm.

The FCM algorithm and its derivatives have been successfully used in pattern recognition, classification, data mining and image segmentation. The FCM algorithm consists of several execution steps

FCM is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{j-1}^{C} u_{ij}^{m} \|x_i - c_j\| \quad 1 \le m < \infty$$

where m is any real number greater than 1,uij is the degree of membership of xi in the cluster j, xi is the ith of d-dimensional measured data, cj is the d-dimension centre of the cluster and ||*|| is any norm expressing the similarity between any measured data and the centre.

Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership uij and the cluster centres cj by:

$$u_{ij} = \frac{1}{\sum_{k-1}^{c} \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

,

This iteration will stop when $max_{ij}\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right|\right\} < \varepsilon$, where $\varepsilon$ is a termination criterion between 0 and 1, whereas k is the iteration steps. This procedure converges to a local minimum.

### 3.4 Feature Filtering

Facial features can be filtered out using edge detection technique. Edges detecting an image significantly reduce the amount of data and filters out useless information while preserving the important structural properties in an image.

Edge detection is a very important area in the field of Image processing. Edges define the boundaries between regions in an image, that helps with the segmentation and object recognition. Edges are significant local changes of intensity in an image.

#### 3.4.1 Sobel operator implementation

Sobel operator is used in image processing particularly within edge detection algorithm. Edge detecting an image significantly reduces the amount of data and filters out useless information while preserving important structural properties in an image.

The kernels can be applied separately to the input image, in order to produce separate measurements of the gradient component in each orientation. These can then be combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient.

The gradient magnitude is given by

$$|G| = \sqrt{Gx^2 + Gy^2}$$

The angle of orientation of the edge (relative to the pixel grid) giving rise to the spatial gradient is given by

$$\theta = \arctan(Gy/Gx)$$

Edges can be detected by Smoothing out image noise. Estimating the gradient of the image at every point to generate a “gradient” image thresholding the gradient image -Remove lighting effects and threshold value should be set between 50 to 100.

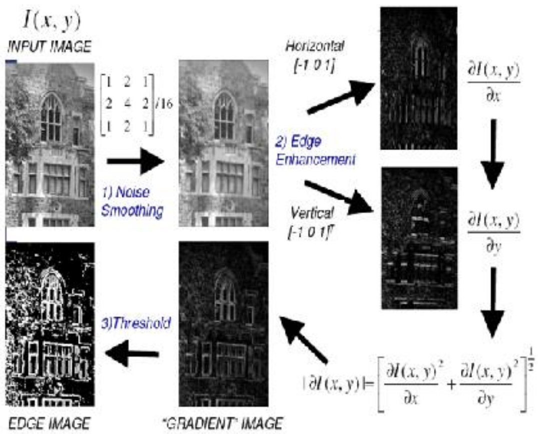


Fig 7.5 Sobel operator implementation

## 4. CONCLUSION

We have developed a VR-based controllable facial expression presentation system that was able to collect facial features data while the subjects were involved in emotion recognition tasks. A usability study involving typical controls was performed to evaluate the efficacy of the system as well as to study behavioural pattern differences in how individuals processed different valences of these expressions. Such capabilities are expected to be useful in understanding the underlying heterogeneous deficits individuals with ASD often display in processing and responding to nonverbal communication of others. Results indicated interesting differences in performance regarding identification of certain emotions as well as differences in how individuals with ASD often processed emotions. The system successfully presented the facial emotional expressions and collected the synchronized facial features.

## 6. REFERENCE

[1] L. Ma and K. Khorasani “Facial Expression Recognition Using constructive Feed forward Neural Networks” ieee transactions on systems, man, and cybernetics—part b: cybernetics, vol. 34, no. 3, june 2004

[2] Marian Stewart Bartlett, Gwen Littlewort, Ian Fasel , Javier R. Movellan “Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction” Proceedings of the 2003 Conference on Computer Vision and Pattern Recognition.

[3] Muzammil Abdulrahman, Tajuddeen R. Gwadabe, Fahad J. Abdu, Alaa Eleyan Electrical & Electronics Engineering Department, “Gabor Wavelet Transform Based Facial Expression Recognition Using PCA and LBP” 2014 IEEE 22nd Signal Processing and Communications Applications Conference (SIU 2014)

[4] Symeon Nikitidis, Anastasios Tefas, Nikos Nikolaidis “facial expression recognition using clustering discriminant non-negative matrix factorization” 2011 18th IEEE International Conference on Image Processing.

[5] Wang Ye, Wu Xiaojun Wang Shitong, Yang Jingyu “Facial Expression Recognition Using RBF Neural Network Based On Improved Artificial Fish Swarm Algorithm” Proceedings of the 27th Chinese Control Conference July 16-18, 2008, Kunming,Yunnan, China

[6] Wei-Lun Chao, 2 Jun-Zuo Liu, 3 Jian-Jiun Ding, 4 Po-Hung Wu “Facial Expression Recognition Using Expression-Specific Local Binary Patterns and Layer Denoising Mechanism” 978-1-4799-0434-1/13/$31.00 ©2013 IEEE