ISSN (Online) 2395-2717
International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE) Vol 2, Issue 5, May 2016

Identification of Human Facial Expression Signal Classification Using Spatial Temporal Algorithm

[1] Mahanthesh U [2] Dr. H S Mohana
[1] Assistant Professor, [2] Professor & Dean
[1]GSSS Institute of Engineering & Technology for Women, Mysore [2] MCE Hassan

Abstract: -- **In the vision of computer Human Action Recognition (HAR) plays a significant role in the research area. The aim of the project is to recognize the Human Facial Expression like happiness, anger, fear, disgust, surprise and sadness from the input video sequence having homogeneous background. Here the pre loaded video sequence of the human facial expression is given as the input to the MATLAB software to obtain the desired output. This project can be implemented using High Configuration MATLAB simulation software.**

Index Items **- Human Facial Expression, Viola - Jones Algorithm, Spatial Temporal Algorithm, Matching algorithm Training, Testing**

I. INTRODUCTION

As the technology is growing rapidly, everything is getting automated. There is a huge need to recognize Human Facial Expression as it has wide range of applications such as information security, authentication, biometric identification, video surveillance, data privacy, Human Computer Interface (HCI), Human Behavior Interpretation (HBI), Health Care etc.,

In Human Machine Interaction (HMI) a natural way of communication is introduced by Facial Expression Recognition (FER). Facial Expression Recognition is categorized under Human Action Recognition (HAR). Emotions are the positive or negative state of a person's mind which is related with a pattern of any physiological activities. Mental state of the person is described by Emotions. Basics emotions are categorized into six classes: HAPPINESS, ANGER, FEAR, DISGUST, SURPRISE and SADNESS. These are the basic six emotions that blend to form complex emotions.

II. PROPOSED WORK

Any input video sequence that is been considered for the detection of the expression is first split into frames. These frames are converted into images. The converted images are pre processed in order to remove unwanted noise. The feature that are needed for the processing is been extracted by the Viola-Jones Algorithm. The desired expressions are identified and the output is obtained.

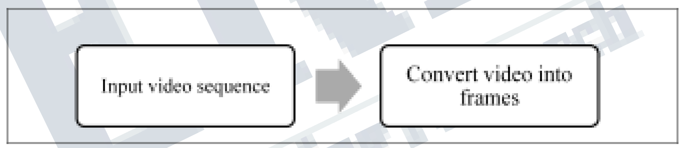

Fig 1: A System for Image Extraction

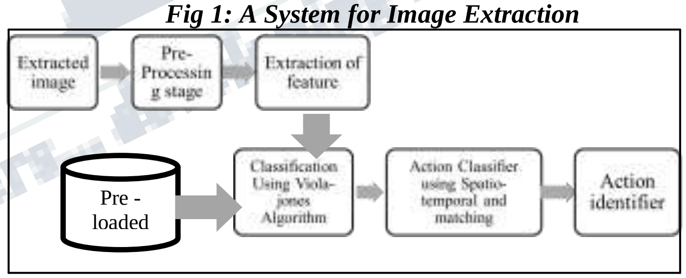

Fig 2: General Human Activity recognition System

The proposed block diagram of human activity recognition (HAR) system is as shown in Fig: 1 and 2. In order to have a robust recognition, the system should be invariant to scale, homogeneous background. There are two sets of data 1) Training and 2) Testing.
The training datasets are used to train the complete system with relevant data of the actions present in the current video sequences and the testing dataset is used for performance evaluation of the system.

The details of the each block of the proposed system are given below.

ISSN (Online) 2395-2717
International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE) Vol 2, Issue 5, May 2016

Image Extraction Model:

The video sequence is used as the input to the developed system. The video should be analyzed frame by frame to obtain the appearance of the object details such as shape and motion characteristics of the video. Hence, the video sequence is converted into frames and in turn into images.

Pre-processing:

In general, due to the presence of unwanted noise the input video, which is affected while capturing, lighting variations, changes in background clutters like, trees and so on, due to which the required details cannot to be extracted satisfactorily. Hence, before feature extraction process there is a need of preprocessing the input image to remove unwanted noise in order to preserve the required details.

Feature Extraction:

In order to perform the recognition task in HAR system there is a necessary for the details which characterize the motions of the object. There are many types of features which can be extracted to characterize the details of motions of human action. These features which represent the characteristics of the motions are necessary to train and test the algorithm which is used to recognize the different types of actions of human body.

III. VIOLA JONES ALGORITHM:

To identify whether it is face or not we are using viola jones algorithm. It consist of four step

1. Harr feature extraction.
2. Integral image
3. Adaboost method
4. Cascade classifier

Harr feature extraction:

All human faces have some similar property, to match these similarity we are using Harr feature. For ex: Eye region is darkening than the upper cheeks, Nose bridge region is brighter than the eyes. Each feature is correlate to a particular position in the sub- window.

We having three types of Rectangular feature i.e. Two, Three, Four rectangular feature. Viola jones uses two rectangular features.

Value=∑(pixel in black area) -∑(pixel in white area).

Integral image:

Image representation is known as integral image, it finds the value of rectangular feature in constant time. Integral image at location (X, Y), is the sum of the pixels above and to the left of (X, Y).

Adaboost method:

It is used to select the best feature and to train the classifiers. It built a "strong" classifier as a linear combination of weighted simple "weak" classifier.

$$H(x)= \text{sign}[\sum_{j=1}^{M} \alpha j\ hj(x)]$$

Each weak classifier is a threshold function based on the feature f j.

$$hj(x)=\begin{cases} -sj & if\ fj < \theta j \\ sj & otherwise \end{cases}$$

The threshold value θj and the polarity sj ∈ ±1 are determined in the training as well as the coefficient αj.

Cascade classifier:

The first classifier in the cascade known as attention operator. It uses two features to attain a false negative rate of about 0% and a false positive rate of 40%. In cascading, all phase consists of a strong classifier. So all the features are grouped into some phases where each phase has certain number of feature. The task of each phase is to find whether a given sub-window is face or not. A given sub-window is straightway discarded as not a face if it loss in any of the phases.

The false positive rate for an entire cascade is given by

$$F=\prod_{i=1}^{k} fi.$$

Detection rate is given by

$$D=\prod_{i=1}^{k} di$$

***Spatial Temporal Algorithm*:**

The variation of space with respect to time is given by Spatial Temporal Algorithm. The extension of spatial domain is Spatial Temporal, which includes geometry over time and location of object moving over invariant geometry.

Matching Algorithm

An **active appearance model (AAM)** is a computer vision algorithm in order to match a statistical model of an object shape and appearance to a new image. They are built during a training phase. A set of images, together with coordinates of landmarks that appear in all of the images, is provided to the training supervisor.

IV. RESULT

ISSN (Online) 2395-2717
International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE) Vol 2, Issue 5, May 2016

The following are the results obtained for the given input per loaded video sequence. The expression consider for processing was joy and the result obtained is the same. The input image is converted from RGB to gray scale and the features are extracted from the Gray Scale image to recognize the input expression of the video sequence. Following results were obtained.

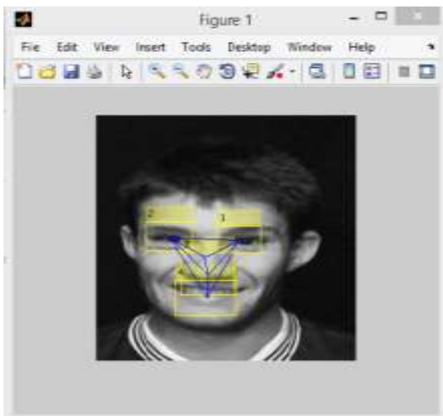

Fig 3(a): Extraction of features

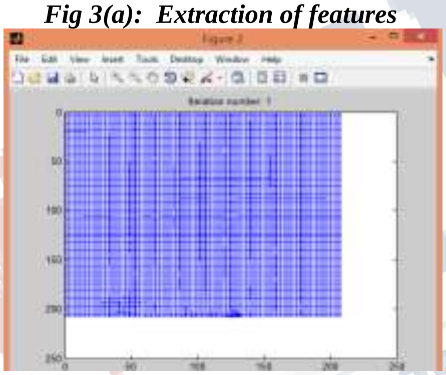

Fig 3(b): Vector for Extracted features

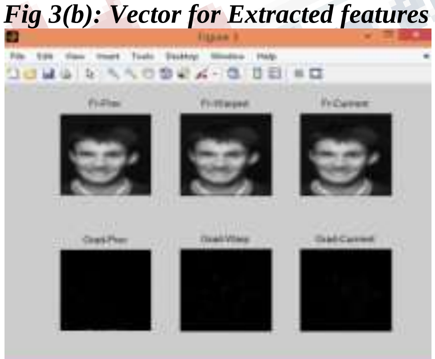

Fig 3(c): Preview, Warped & Current frames of the feature extracted frames

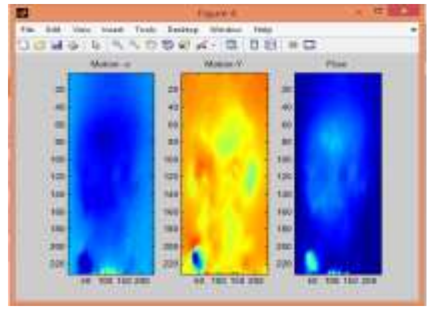

Fig 3(d): Motion with respect to X & Y axis

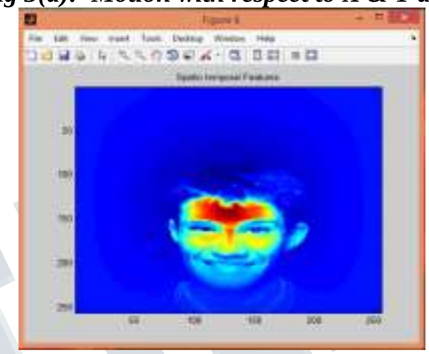

Fig 3(e): Spatio temporal features of the best feature extracted frame

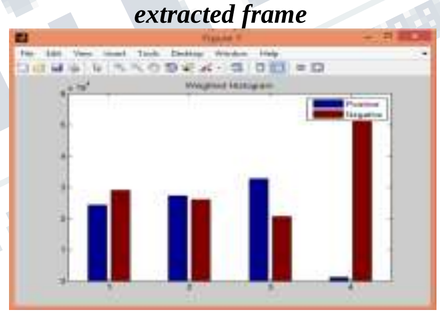

Fig 3(f): Weighted Histogram

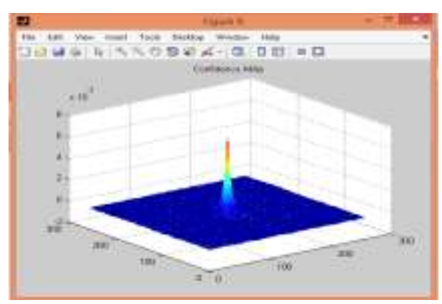

Fig 3(g): Confidence Map

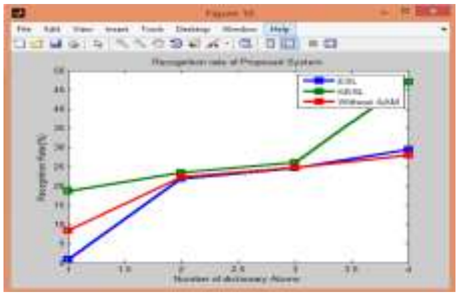

Fig 3(h): Recognition rate of proposed system

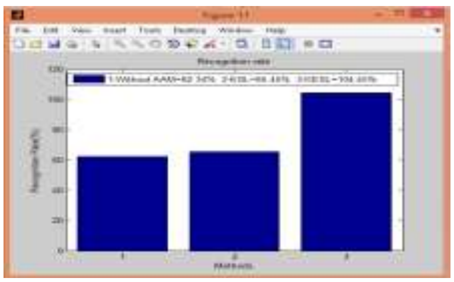

Fig 3(i): Recognition rate

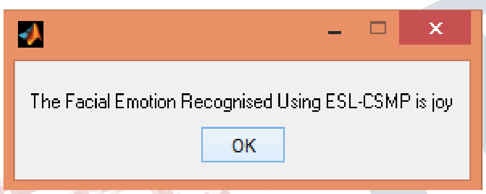

Fig 3(j): Expected Result for the preloaded video (joy expression)

IV. CONCLUSION

Human Facial Expression place a vital role in inters personal communication containing extremely abundant information of human behavior. This paper fulfills the requirement. Future work can de done by using a dynamic video sequence.

REFERENCE

[1] Pooja Yadav and Gaurav Aggarwal , "Speech Emotion Classification using Machine Learning" International Journal of Computer Applications (0975 – 8887) Volume 118 – No. 13, May 2015.

[2] R. V. Darekar, A. P. Dhande, "Emotion Detection with Multimodal Fusion Using Speech - A Review" International Journal of Computer Science and Communication Engineering Volume 3 issue 1(February 2014 issue)

[3] Michael B. Holte, Human Pose Estimation and Activity Recognition from Multi-View Videos: Comparative Explorations of Recent Developments, IEEE Journal of Selected Topics In Signal Processing, Vol. 6, No. 5, September 2012

[4] Lu Xia and J.K.Aggarwal,"Spatio-Temporal Depth Cuboid Similarity Feature for Activity Recognition Using Depth Camera", Computer & Vision Research Center / Department of ECE, The University of Texas at Austin, aggarwaljk@mail.utexas.edu.

[5] V. Kulkarni and Savitha S. Raut, "Emotion Recognition By Using Speech And Facial Expressions". Proceedings of 9th IRF International Conference, Pune, India, 18th May. 2014, ISBN: 978-93-84209-20-9.

AUTHOR'S BIOGRAPHY

Mr. Mahanthesha U Pursuing Ph.D in the field of "Image Processing". He Obtained M-Tech degree in Computer Networking Engineering at National Institute of Engineering, Mysuru under VTU Belgaum in 2014. He obtained his B.E degree in Electronics and Communication Engineering at HMS Institute of Technology, Tumkur under VTU Blegaum in 2009. He has published many National & International journal and conference papers. His areas of interest are AEC, C++, Computer Networks, Signals & Systems, DSP, Control Systems. Presently, he is working as Assistant Professor in the Dept. Of Electronics & Instrumentation Engineering at GSSS Institute of Engineering & Technology For Women, Mysuru.

Dr. H. S. Mohana, has a teaching experience of more than 28 years. At present, he is working as professor and dean in the department of instrumentation technology of malnad college of engineering, hassan. He has published several papers in international conferences and journals. His areas of interest include computer vision based instrumentation, digital image processing & pattern recognition etc.