AUTOMATED RECOGNITION OF FACIAL EXPRESSIONS

[1] Pavan S. Ahire, PG Student, Dept. of Computer Engineering, METs Institute of Engineering, Adgoan,Nashik,Maharashtra.
[2] Prof. R. P. Dahake, Dept. of Computer Engineering, METs Institute of Engineering, Adgoan, Nashik, Maharashtra.

ABSTRACT

Automatic facial expression recognition is a research oriented subject on the basis of interesting area in human and computer interaction. The classification accuracy achieve by automatic system which uses static images as input have barrier by image quality, lighting conditions and the direction of the face. These problems can be overcome by fully automated facial expression detection and classification framework. A different classifier consists of particular sets of Random Forests that are equipped with support vector machine labelers. The system performs under real time conditions, The dynamic acted still-image database is used for training purposes and expression-oriented video databases is used for testing.

Keyword: - *Facial actions, random forest, support vector machine, video glossary.*

---

1 INTRODUCTION

Human face is elastic object that have various attributes related to the facial objects. Some behavior can extract human intention that exhibits some emotions. This can be pointed out through review of the state of facial expression research [1]: Use of various viewpoints is limited. Over the past two years however, more work has been dedicated to this topic as low cost, high quality consumer cameras came bundled with gaming devices.A totally automated complete facial expression detection and classification scheme is design and implemented in this paper. A typical facial expression recognition system starts with the detection and accessing of a face from an image/video. A professional detector held by PittPatt [5] and a set of random forests for classification purpose are paired with SVM labelers, it uses a face detector and a classifier including with no human interference. Various facial images and videos are used to recognize and detect expressions using expression labelers. The facial expressions classified by individuals in a random environment record by CCTV. The major assertion is that externally presented facial actions are an important window on subjective emotions.Find and crop every face in each images or video frames, to identify the center of the face.Formerly a face has been disclosed and assigns it with brightness, contrast, size and pose.Each possible expression yaw is analyzed using an individual RF[2].The RF collection occupied here is a novel design that advantageous of individual RF that distributes a multiple label classification problem into a set of binary sub-problems. The second level is an SVM labeler where it takes the output and excerpts a most likely expression.

2 LITERATURE SURVEY

The nature of various poses of human faces are random in nature.Sometimes the visibility of face according to their angles are differ.To overcome this problem Mostafa K. Abd El Meguidet.al.[1] has been developed fully automated recognition system with the help of random forest classifier. Random forests are a design that proposed by Leo Breiman et al. in the 2000s for producing a predictor combines with a number of determining trees that builds up in randomly choose subspaces of data. Growing interest and informative use, there has been little bit analysis of the graphical properties of random forests. An comprehensive reasoning of a random forests model implied by Breiman[2],shows that the procedure is constant and lies to sparsity, that its rate of convergence depends only on the number of strong build features and not on how many noise oriented variables are present. The purpose of this work is only not to identify some part of facial emotions which is implies with human facial action but to demonstrate its degree variation. It is based on that facial expression recognition that can be fulfilled by extracting a variation from unexpression face with concluding face area. Using a elastic net model, a variation of facial expression is represented as motion vectors of the de-formed Net from a facial edge image. Then, applying K-L expansion, the change of facial expression represented as the motion vectors of nodes is mapped into low dimensional Eigen space, and estimation is achieved by projecting input images on to the Emotion Space. A constructive three kinds of expression models are: happiness, anger, surprise, and experimental results are evaluated [3]. A computer vision system, including both facial feature extraction and recognition that automatically discriminates among subtly different facial expressions. Expression classification is based on Facial Action Coding System (FACS) action units (AUs), and discrimination is performed using Hidden Markov Models (HMMs). The first method is facial feature point tracking using a coarse-to-fine pyramid method. This method is sensitive to subtle feature motion and is capable of handling large displacements with sub-pixel accuracy. The second method is dense flow tracking together with principal component analysis (PCA).The third method is high gradient component analysis in the spatio temporal domain, which exploits the transient variation associated with the facial expression[4]. Upon extraction of the facial information, non-rigid facial expression is separated from the rigid head motion component, and the face images are automatically aligned and normalized using an affine transformation. This system also provides expression intensity estimation, which has significant effect on the actual meaning of the expression. Traditionally, human facial expressions have been studied using either 2D static images or 2D video sequences. The 2D-based analysis is incapable of handling large pose variations. Although 3D modeling techniques have been extensively used for 3D face recognition and 3D face animation, barely any research on 3D facial expression recognition using 3D range data has been reported. A newly developed 3D facial expression database, which includes both prototypical 3D facial expression shapes and 2D facial textures of 2,500 models from 100 subjects. The new database can be a valuable resource for algorithm assessment, comparison and evaluation[5]. Automated analysis of human affective behavior has attracted and increasing attention in psychology, computer science, linguistics etc.Deliberate behavior differs in visual appearance, audio profile, and timing from spontaneously occurring behavior. To address this problem, efforts to develop algorithms that can process naturally occurring human affective behavior have recently emerged [6]. Moreover an increasing number of efforts are reported toward multi modal fusion for human affect analysis, including audiovisual fusion, linguistic and paralinguistic fusion, and multi cue visual fusion based on facial expressions, head movements, and body gestures. Finally outline some of the scientific and engineering challenges to advancing human affect sensing technology [7]. Within the past decade, significant effort has occurred in developing methods of facial expression analysis. By describing the problem space for facial expression analysis, which includes level of description, transitions among expression, eliciting conditions, reliability and validity of training and test data, individual differences in subjects, head orientation and scene complexity and relation to non-verbal behavior. Presented CMU-Pittsburgh AU-Coded Face Expression Image Database, which currently includes 2105 digitized image sequences from 182 adult subjects of varying ethnicity, performing multiple tokens of most primary FACS action units. This database is the most comprehensive testbed to date for comparative studies of facial expression analysis [8]. CK database has become one of the most widely used test-beds for algorithm development and valuation. CK database has been used for both AU and emotion detection. The target expression for each sequence is fully FACS coded and emotion labels have been revised and validated[9].Facial Action Coding System Affect Interpretation Dictionary (FACSAID) is a project that links facial expressions with their psychological interpretations. FACSAID information is stored in a relational database, which models facial behaviors and the meanings of these expressions. The system is described as a dictionary because the user can look up meanings for a particular facial behavior, or look up the facial behaviors that imply a particular meaning or emotion, such as anger or happy. The database dictionary approach of FACSAID is different from that of a particular rule-based or expert system[10].Contraction of the facial muscles produce changes in both the direction and magnitude of the motion on the skin surface in the appearance of permanent and transient facial features[12]. An image may be defined as a two-dimensional function, where x and y are spatial (plane) coordinates, and the amplitude of at any pair of coordinates (x,y)is called the intensity or gray level of the image at that point[13] Some of the initial problems in improving the visual quality of these early digital pictures were related to the selection of printing procedures and the distribution of intensity levels. We had investigated the correspondence between perceived meanings of smiles and their morphological and dynamic characteristics. Morphological characteristics included co-activation of *Orbicularis oculi* (AU 6)[14],smile controls, mouth opening, amplitude, and asymmetry of amplitude Dynamic characteristics included duration, onset and offset velocity, asymmetry of velocity, and head movements. Quality data recorded in varied realistic environments is vital for effective human face related research. Currently available datasets for human facial expression analysis have been generated in highly controlled lab environments. We present a new static facial expression database Static Facial Expressions. The JAFFE database [15] is one of the earliest static facial expressions dataset. It contains 219 images of 10 Japanese females. The subjects better variant posed for six expressions(angry,disgust,fear)and the neutral expression. Quality data recorded in varied realistic environments is vital for effective human face related research. Facial expression analysis has been a very active field of research and many robust methods have been reported in the literature in the past[16] However these methods have been experimented on different databases and using different protocols within the same databases. The lack of a standard protocol makes it difficult to compare systems and acts as a hinderance in the progress of the field.

3 PROPOSED SYSTEM

Automatic facial expression identification System is applying an interest in various applications. Firstly perform Detection and acquisition of a face from an image or video after that track and crop each face in each video frame or images. Once a face has been detected. Enhance it with prior to brightness, contrast, size. Use feature analysis and dimensional reduction methodology to catch the face where expressions are identified using a classifier. A collection of binary random forests is represented. SVM labeler that takes the output of the RF collections and extracts a single, preferable expression from all available possibilities. As shown in Fig. 1, I/P training dataset is used in proposed system in which their purpose is for training the dataset. The following system architecture shows overall processing of input as well as output of the system.

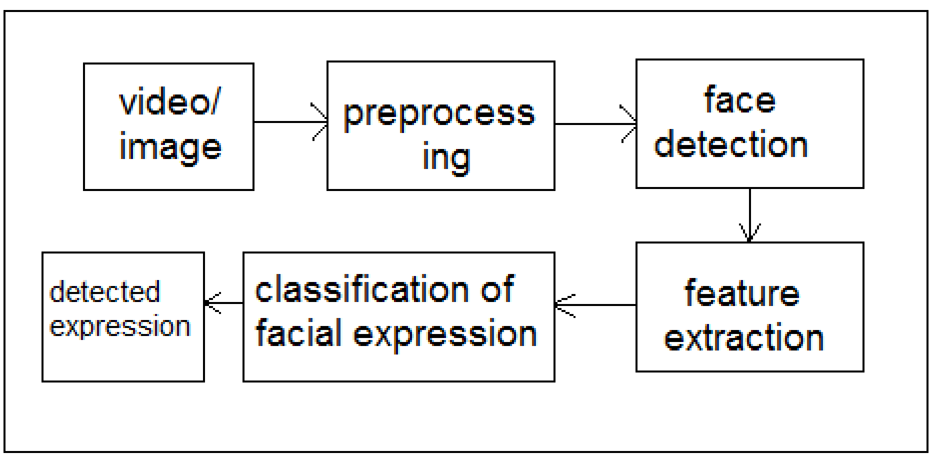

Fig -1 :Proposed Work for Facial Expression Recognition

3.1 Image Smoothing

Image Smoothing is a method by which we can blur the image and smoothing reduces noise giving us a more accurate intensity surface.

3.2 Gaussian filter

Gaussian filtering is applied to blur images and eliminate noise. It also illustrates a probable distribution for noise. It is a preprocessing method that need to use the two dimensional Gaussian function. It is a low pass filter specially used for images.

3.3 Edge detection

The primary objective is to minimize the amount or size of data in an image, as well as preserving the temporal attributes that to be used for image processing. For detecting edges A Canny Edge detection algorithm is used. There are different steps in canny edge in which they are processed.

3.4 Face Detection

Detecting and tracking faces in images is a task for realizing monitoring systems human computer intervention. Feature extraction starts with initial sets of measured data and produce derived values to be informative and subsequent learning is related to dimensionality reduction, for this Haar like feature method is used.

3.5 Feature Extraction

It is a normal function in analysis of image that compare similarity between two images. The features in the image utilizing its optical contents such as color signature, shape and color texture.

3.6 RF Classifier

Random decision forests are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

3.6 SVM Labeler

This is the final stage of given architecture in which it is output of implementation.For labeling the expressions support vector machine labeler is used.They primarily used to find out the Support Vectors for a given set of points.The algorithm flows by handling a candidate Support Vector set by using a greedy approach to pick up points for inclusion in that candidate set.

4. HAAR LIKE FEATURE

A face tracking system is basis on a quick directed face detection scheme using Haar-like features. It is instance of combined Haar-like features, which good catches the properties of human facial actions, possible to build a powerful classifier. The joint Haar-like feature can be measured very fast independently of image resolution and has robustness against addition of noise and change in illumination. A face detector is learned stage wise selection of weak classifiers basis on the joint Haar-like features.Figure 2 shows an overview of our method. The final strong classifier H(x) is a straight combination of weak classifiers $h^1(x)$ to $h_T(x)$.Each weak classifier contain multiple features. For example, $h_1(x)$ find F features and calculate joint statistics of this features. The identical similarities of faces, which cannot be measured using a single feature, are extracted from z_1,1 (eye regions are darker than neighboring regions), z_1,f (nostrils are dark) and z_1,F (the region between the eyes is brighter than the eyes).These combined features are selected in each round of the boosting process, such that the error on the training set is minimized. Feature concurrence makes it possible to classify difficult examples that are misclassified by weak classifiers using a single feature.To calculate the joint probability, quantize the feature value z to two levels. Every feature value is represented by a binary variable s, which is 1 or 0, specifying face or non face respectively. The variable s for an example x is calculated by

$$s(x) = \begin{cases} 1 & if p.z(x) > p.\theta \\ 0 & otherwise \end{cases}$$

Figure 2 shows an example of the joint Haar-like feature, which is based on the cooccurrence of three Haar-like features.

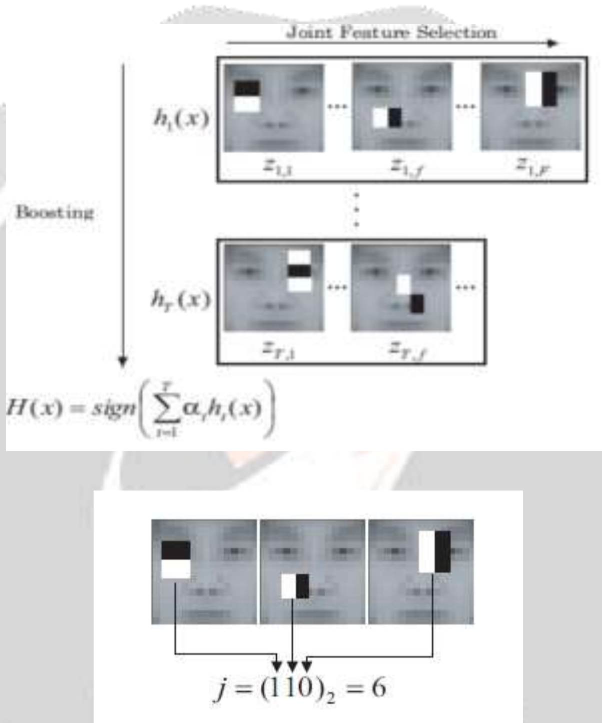

Fig. 3. Example of the joint Haar-like feature

the variables are 1, 1 and 0, the value of the joint Haar-like feature is calculated by

$$j = (110)^2 = 6$$

The feature value j as a binary number specifies an index for 2F different combinations, where F is the number of combined features. The feature represents the feature co-occurrence between different positions, resolutions and orientations. For each class, statistical dependencies between the features are obtained by observing j for each example.We Use such dependencies for classification. The sub window is Classified to be face or non face by evaluating from which class the feature value is likely to be observed. The combined Features are selected to capture distinctive structural Similarities of faces.

5. IMPLEMENTATION DETAILS

A. Dataset

A static image dataset is used which is mutually available for testing. Today many datasets are available like CohnKanade[15]CK+ image set that contains expressions like frames of faces include various actors(male/female) with various total size of 1.70GB and Jaffee[9]contain japnese female dataset size of 408 MB,also own self made videos(synthetic) is also taken that contain expressions.images also taken for input that are dynamic in nature.some are available on google source while others are self created or clicked one most important thing image dataset contain age wise collection of images in which child,adult and old like category is taken.

B. Video Result

Face tracking Detection Taking the image as a input in the following figure 4.It is the first step to ttrace the face from the given image by detecting their eyes and mouth areas with haar like features that points the face centralized the current regions. It shows analysing and detection through face regions in the result as the images in the face detection figure as cropped faces.

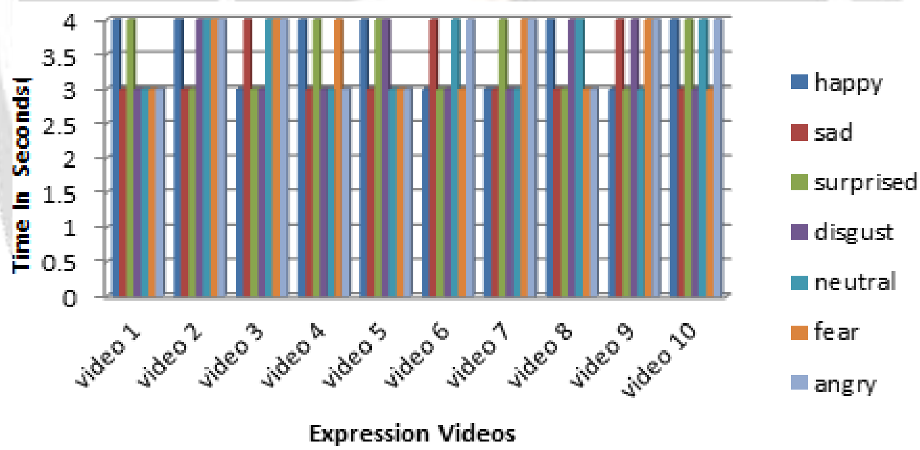

Figure 4 indicate the graph showing the expression detection on the basis of multiple expressions.

Expression with total image	Child	Adult	Old
Happy(15 image)	12	11	8
Angry(15 image)	10	9	11
Sad(15 image)	9	14	13
Neutral(15 image)	10	12	8
Surprised(15 image)	11	12	10

Expression wise accuracy for both same image in existing and proposed system

The given table shows the detection of child expression,adult and old person and comparing with each other by total image.

6. CONCLUSIONS

This paper carry out classifying expressions in various situations. The classifier obtained is trained on the Cohn Kanade & Jaffee Video database and also tested on other self made databases. The classification accuracies is close to expected result only some frame related obligations or resolution image have problem for detection of some expression like sad and neutral because they are correlated to each other.Sometimes other expressions are also deliberately collide with each other when we randomly trained and tested them.still some limitations are present while detecting and classifying the expressions.Future aspect will meet this requirementS for next research level for this topic.

5. ACKNOWLEDGEMENT

I am very much thankful to all the authors of the papers which we have referred and glad to express my sentiments of gratitude to all of them who rendered their valuable contribution and help to make this work successful.

6. REFERENCES

[1]. Mostafa K. Abd El Meguid and Martin D. Levine, Fully Automated Recognition of Spontaneous Facial Expressions in Videos Using Random Forest Classifiers", IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. 5, NO. 2, APRIL-JUNE 2014 pp. 141154.

[2]. L. Breiman, Random forests, Mach. Learn., vol. 45, pp. 532, 2001.

[3]. B. Seddik, H. Maamatou, S. Gazzah, T. Chateau, and N. E. Ben Amara, Unsupervised facial expressions recognition and avatar reconstruction from kinect, in Proc. 10th Int. Multi-Conf. Syst., Signals Devices, 2013, pp. 16.

[4]. Y.-L. Tian, T. Kanada, and J. F. Cohn, Facial expression recognition, in Handbook of Face Recognition, 2nd ed., New York, NY, USA: Springer, 2011. pp. 487519.

[5] Pittsburgh, Pattern Recognition. (2012, Nov. 4) PittPatt SDK v5.2.2 Documentation [Online]. Available: http://www.pittpatt. Com/documentation.

[6] F. de la Torre and J. F. Cohn, Facial expression analysis, in Guide to Visual Analysis of Humans: Looking at People. New York, NY, USA: Springer, 2011.

[7] P. Ekman, W. V. Friesen, and P. Ellsworth, Emotion in the Human Face: Guide-Lines for Research and an Integration of Findings. *New York, NY, USA: Pergamon, 1972.*

[8] P. Ekman, Universals and cultural differences in facial expressions of emotion, in Proc. Nebraska Symp. Motivation, Lincoln, NE, USA, 1971, pp. 207283.]

[9] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I.Matthews, “The extended Cohn-Kanade dataset (CKþ): A complete dataset for action unit and emotion-specified expression,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog. Workshops,2010, pp. 94–101.

[10] J. Hager, “FACSAID – A tool for interpreting facial expressions,”2007.

[11] T. Ying-Li, T. Kanade, and J. F. Cohn, “Recognizing upper face action units for facial expression analysis,” in Proc. IEEE Conf.Comput. Vis. Pattern Recog., 2000, vol. 1, pp. 294–301.

[12] P. Ekman and W. V. Friesen, “Measuring facial movement,” J.Nonverbal Behavior, vol. 1, pp. 56–75, 1976.

[13] R. C. Gonz_alez and R. E. Woods, Digital Image Processing. Englewood Cliffs, NJ, USA: Prentice-Hall, 2008.

[14] Z. Ambadar, J. Cohn, and L. Reed, “All smiles are not created equal: Morphology and timing of smiles perceived as amused,polite, and embarrassed/nervous,” J. Nonverbal Behavior, vol. 33,pp. 17–34, 2009.

[15] A. Dhall, R. Goecke, S. Lucey, and T. Gedeon, “Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark,” in Proc. IEEE Int. Conf. Comput. Vis. Workshop, Barcelona, Spain, 2011, pp. 2106–2112.

[16] A. Dhall, R. Goecke, S. Lucey, and T. Gedeon, “Acted facial expressions In in the wild database,” Australian Nat. Univ., Canberra, Australia, Tech. Rep. TR-CS-11-02, Sep. 2011.