EXPRESSION RECOGNITION USING SUPPORT VECTOR MACHINES

Jyothi L[1], Arun Anathanarayanan[2]
[1]UG Student, [2]Assistant Professor Department of ECE
MVJ College of Engineering, Near ITPB, Whitefield,
Bangalore, Karnataka, India

ABSTRACT

Facial expression detection has many applications in many fields, where it is used for learning the behavior of human at the particular situations, in medical field used for mental disorders detection and synthetic human expressions. This paper discusses about the position of one of application of the artificial intelligence techniques, support vector machines (SVM) for recognizing the expressions in the human face. This paper discusses the detection of the key expressions on human face like anger, sadness, neutrality, happiness, disgust, fear, surprise. The database is created based on the key expressions and characteristics are obtained from the database. The SVM is trained using these features and SVM trained model is generated. When the test image is given to the SVM classifier, it extracts the characteristics of that test image and compares those features with the trained features and takes the decision and provides the result

KEYWORDS: SVM (Support Vector Machine), face feature tracker, Feature extraction, Training and classification.

I. INTRODUCTION

Human beings use facial expressions as an essential and powerful tool to convey their emotions and to interact. Understanding human emotions is the primitive area of research, as ability of recognizing the emotion gives access to excess of opportunities and applications, ranging from friendlier human-computer interactions to be better targeted advertising campaigns, and conclude with an improved communication among humans, by improving the impassioned intelligence. We use facial expressions not only to express our emotions, but also to provide important communicative cues during social interaction. Many ways are present to know the human commotions, ranging from facial expressions, and posture of the body, the human voice’s speed and tone. We focus here only on facial expressions. One of the main reasons on which we focus on the facial expressions is because of the certain expressions have universal meaning and these emotions have been documented from many years.

Face detection and recognition are related to facial expression recognition in their methods used for characteristic extraction and classification. The face as a biometric plays a major role in modern day security procedures as it is a unique feature among people. Face detection is an independent field of research and is used for facial expression and face recognition systems. Those systems require face detection as the first step to supplement their operation in a real world environment. Various methods and techniques for solving complex task of facial expression recognition have been proposed. A generic facial expression recognition system can be decomposed in the following procession stages: firstly, facial features are located in the face to centralize on the details in the face. Next facial motion or dislocations of Facial features are extracted. The extracted features are often modeled prior to recognition with different representations. Finally, classification takes place to classify the expressions by action units.

In the paper written by MD. ZIA UDDIN, WERIA KHAKSAR AND JIM TORRESEN, they mentioned about extracting the features by developing the local directional rank histograms patterns(LDRHP) and local directional pattern, where local directional strength patterns are expanded with LDRHP features proceeded by Kernel principal component and generalize discriminatant analysis to inaugurate enormous robust characteristics. Then the characteristics are trained with a deep learning approach and convolution neural network. Most approaches to robotize expression recognition require to deal with the issues of the face localization, facial characteristic extraction and training as well as the segregation stages of the learning method used (Pantic & Rothkrantz 2000).

Most of the systems previously consider the presence of a full frontal face view in the image series being studied and some philosophy of the universal face location is been gained. The definite location of the face is given by; Viola & Jones use the Ad-a boost algorithm to completely pass an exploration sub-window over the image at numerous scales for getting the rapid face detection. From image sequences, Essa and Pentland executed spatial and temporal filtering together with thresholding to select motion blobs. To detect existence of a face, then these blobs are graded using Eigen faces method via principal component analysis to calculate the distance of the realized region from a face space of 128 sample images. Using spatio-temporal filtering the Person Spotter system tracks the bounding box of the head in the video and stereo disparity in pixels due to motion, thus selecting image regions of interest. Then it employs skin color and convex region detectors in these areas to check for face presence.

By evaluating the distance of an image form the feature space given a set of sample images through Fast Fourier Transform and a local energy computation, Essa & Pentland extend their face detection approach to extract the positions of remarkable facial features using Eigen features as well as PCA. In the last step of expression analysis, expressions are classified according to some pattern. Essa & Pent land calculate ideal motion energy templates for every expression category and take the Euclidean norm of the difference between the observes motion energy in a sequence of images and each motion energy in a series of images and each motion energy template as a relative metric.

In this paper, the method proposed is for automatically deducing emotions by detecting facial expressions. We base the method on the machine learning system of Support Vector Machine (SVMs). SVM is a supervised machine learning algorithm which can be used for segregation and reverting. It is mostly probably used in classification problems. SVM can solve non-linear problems with minimum amount of samples and high dimensions features. The features that are captured are used to train an SVM classifier to detect the unseen expressions.

II. SUPPORT VECTOR MACHINE

A Support Vector Machine (SVM) is a selective classifier formally defined by a detaching hyper plane. For the given labeled training data (supervised learning), the algorithm provides an optimal hyper plane that classifies new examples. In two dimensional spaces the hyper plane is a line that divides the plane into two parts where in each class lay in either side.

The hyper plane in linear SVM is learnt by transforming the problem using some linear algebra. A powerful perception is that the linear SVM can be recasted using the inner product of any two given information, rather than the observations themselves. The inner product between two vectors is the sum of the multiplication of each pair of input values. There will be many hyperplanes, the right hyperplane should be selected that will segregate the objects better. The right hyperplane can be selected by maximizing the distances between nearest data point (either class) and hyper-plane which is called as Margin. SVM selects the hyper-plane which classifies the classes accurately prior to maximizing margin.

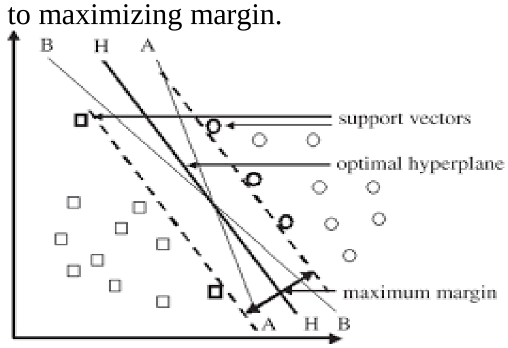

Fig 1: Selection of the optimal hyper plan

The real time facial feature tracker is used to pack with the complications of face localization and feature extraction on uncontrolled expressions. And also the training as well as the classification stages of the SVM is also some problems that are dealed by tracker. The area of facial feature is been extracted by the tracker.

The calculated displacements for each characteristic between a neural and a prototypal frame of expression. These are used together with the label of the expression as the input to the training stage of SVM classifier. Upon request of the user or continuously, for each frame in the video stream the trained SVM model is used to classify unseen feature displacements in real time. The Fig 2 dynamically illustrates the implementation.

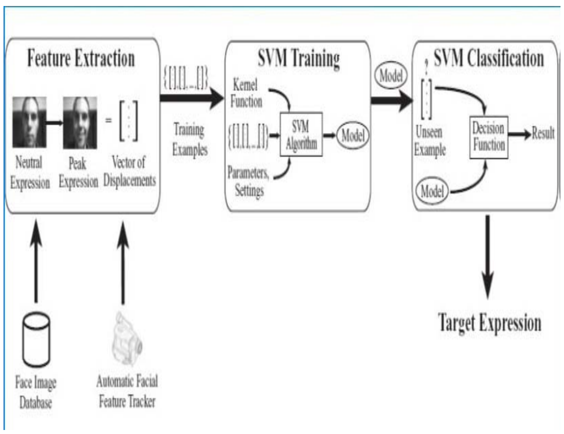

Fig2: Block diagram of the SVM automated expression recognition

The FIg 3 explains the stages involved in implementation.

The images of different expressions are captured and the database of all those images is generated. The preprocessing which will remove the noises in the images that is there in the database. The neutral faces without any expressions are captured of the faces that are in the database. The Eigen vectors of expression images are extracted as the feature vectors and classifier is trained with these features and is ready for testing or recognizing the expression of the unseen image.

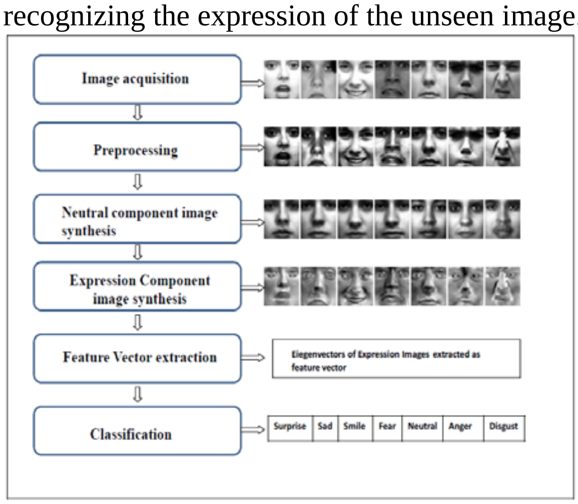

Fig3: Flow chart that describes the implementation steps

III. FEATURE EXTRACTION

The feature displacement approach is chosen due to its appropriateness for a real time video system, in which the motion is latent and that places a strict upper bound on the computational complexity of methods that are used so that they meet time constraints. The face template is used by the tracker to initially locate the area of certain facial features they can be of 22 or 58 features of the face model in the video stream and uses a filter to track the position over successive frames. For every expression, feature displacements vector is evaluated by taking the euclidean distance between feature location in a neutral and a peak frame representative of the expression, as shown in Fig 5 and 6. Feature locations are automatically secured when the amount of motion is at a minimal phase or the final phase of uncontrolled expression,when motion has settled corresponding to either the initial neural around its peak frame.

Fig 5: Facial Features localized of the peak frames in the sequence of frames in video

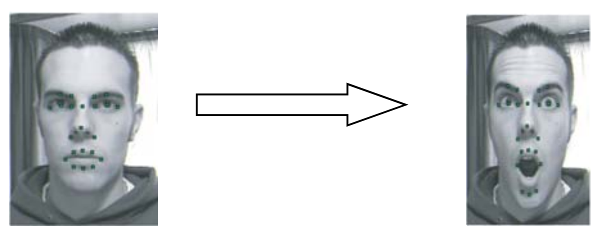

Fig 6: Feature motion pattern

IV. TRAINING AND CLASSIFICATION

The labelled vector of displacments of each expession aupplied is used as input to an SVM classifier, culminate in a model of the training data, which is subsequently used to actively classify unseen feature displacements. The output is the returned to the user.

We assume that the input data is to be linearly separable. Real world data, however is generally not linear separable. To address this issue, SVMs separate non linear data using a trick. The trick is to extend linear SVMs to nonlinear SVMs by grading the input data nonlinearly into ahigher dimensional space called feature space. In sufficient high dimensions, the data becomes linearly detachable. With this mapping in mind, the SVM can solve the optimisation problem in the feature space as it would do in the input space and find an optimal separating hyperplane. Once the optimal hyperplane is found , it is mapped back into the input space resulting in a non linear decision surface.

Classification methods emulated via kernel archive better performance due to the optimal separating hyperplane the SVM pursue to find. It provides a solution for the problem with a high probable global optimum. SVMs are maximal margin hyperplane classifiers that exhibit high classification accuracy for small training sets and good generalisation performance on very variable and difficult to separate data. This makes them particularly suitable to expression recognition in video.

V. EVALUATION

We evaluate our system by considering classification performance for the basic emotions. This approach makes no assumption about the specific emotions used for training or classification and works for arbitrary, user-defined emotion categories. To establish an upper bound on the recognition accuracy achievable for a combined feature displacement / SVM approach to expression recognition, on still images initially it is evaluated. Features were manually defined for each image and displacements were subsequently extracted from pairs of images consisting of a neural and a representative frame for each expression. We use the standard SVM classification algorithm together with a linear kernel.

The Table II gives the percentage of correctly classified examples per basic emotion and the overall recognition accuracy. Particular emotions (e.g.: joy) or indeed particular combinations (e.g.: fear vs. joy) are inherently harder to distinguish than others. However, the results also expose certain motion properties of expressions which seem to be universal across subjects (e.g.: raised mouth corners for ‘joy’, corresponding to a smile).

Emotions	Percentage correct
Anger	82.3%
Disgust	84.7%
Fear	71.8%
Joy	93.06%
Sorrow	85.47%
Surprise	99.5%
Average	86.1%

Table II: Recognition accuracy of SVM classification on displacements extracted from the still image

Kernel choice is the most important customizations that can be made when adjusting an SVM classifier to a particular application domain. Then the test samples are been SVM classified using kernel. To evaluate expression recognition in still images, while live sessions were carried out to evaluate video-based recognition of spontaneous expressions. Table I gives the improvements including selection of a kernel function customized to the training data boosting recognition accuracy up to 87.9%. Incorporating further possible enhancements such as adjusting data to account for head motion or performing automatic SVM model selection is likely to yield even better performance and further increase the suitability of SVM based expression recognition approaches in building affective and socially intelligent human-computer interfaces.

Emotion	**Anger**	**Disgust**	**Fear**	**Joy**	**Sorrow**	**Surprise**	*Overall*
Anger	-	72.4%	61.9%	97.2%	93.9%	97.2%	84.2%
Disgust	72.6 %	-	63.2%	94.8%	88.8%	100.0%	83.8%
Fear	61.9%	63.3%	-	90.8%	66.8%	100.0%	76.3%
Joy	97.2%	94.8%	90.8%	-	96.9%	97.3%	95.4%
Sorrow	93.9%	88.8 %	66.8%	96.9%	-	100.0%	89.5%
Surprise	97.2%	100.0%	100.0%	97.1%	100.0%	-	98.9%

TABLE1: Classification accuracy for each pairing of basic emotions. Rightmost column gives the overall mean accuracy of each row (the confusion matrix)

VI. CONCLUSION

In this paper, Support Vector Machine classifier is used for classifying the test images based on the features that are extracted from it and taking the decision based on the data that is trained. Adding on emotive information in computer – human interfaces will allow for much more natural and efficient interaction models to be established. It is our belief that methods for emotion recognition through facial expression that work on the real time video without preprocessing while remaining cheap in terms of equipment and unassuming for the user will play an increasing role in building such affective and intelligent multimodal interfaces.

VII. REFERENCES

[1] B. Boser, I. Guyon, and V. Vapnik. Training Algorithm for optimal margin classifiers. In D. Haussler, editor, Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, Pages 144–152, 1992.

[2] Cristianini, N. & Shawe-Taylor, J. (2000) an Introduction to Support Vector Machines. Cambridge: Cambridge University Press

[3] Ekman, P. (1982) Emotion in the Human Face. Cambridge: Cambridge University Press Analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00).

[4] Kanade, T., Cohn, J.F., & Tian, Y. (2000) Comprehensive Database for Facial Expression Support Vector Machines and other Kernel-basedLearning Methods. Cambridge University Press,Cambridge, UK, 2000.

[5] Pantic, M. & Rothkrantz, L.J.M. (2000) Automatic Analysis of Facial Expressions: The State of the Art. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1424–1445.

[6] 'Support Vector Machines for Machine Learning',https://machinelearningmastery.com/support-vector-machines-for-machine-learning/

[7] Rana El Kaliouby, Philipp Michel,' Real Time Facial Expression Recognition in Video using Support Vector Machines',

[8] Anvita Bajpai, Kunal Chadha,' Real-time Facial Emotion Detection using Support Vector Machines', *(IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 1, _o. 2, 2010]*

[9] Hehua Chi a, Lianhua Chi b *, Meng Fang a, Juebo Wu c, 'FACIAL EXPRESSION RECOGNITION BASED ON CLOUD MODEL', The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part II

[10] *Philipp Michel and Rana El Kaliouby*, 'Facial Expression Recognition Using Support Vector Machines'

[11] Uzair Asad, Nirbhay Kashyap, Shailendra Narayan Singh, 'Recent Advancements in Facial Expression Recognition Systems: A Survey'