SURVEY ON FACE EXPRESSION RECOGNITION FROM HUMAN EMOTIONS BY USING SVM CLASSIFIER

Suma K[1], Lakshminarayana. M[2]
[1]M.Tech, 4th Sem, DEC, Dept. of ECE, MVJCE, Bangalore
[2]Asst. Professor, Dept of ECE, MVJCE, Bangalore

ABSTRACT
A facial expression is one or more motions or positions of the muscles beneath the skin of the face. Facial Feature extraction is done by a new feature descriptor method to identify the proper facial expression from the facial images. Then computes different directional motion information and encodes the directional flow information with enhanced local patterns. As it captures the spatial temporal changes of facial movements through by using pyramidal lucas kanade method and it shows its robustness in recognizing facial information along with that SURF (Speeded-Up Robust Features) is used for feature extraction. Finally, the extracted features are used to train the expression model through the Support Vector Machine (SVM). The performance of the proposed method has been measured by accuracy. Experimental results may demonstrate that the proposed system is more robust in extracting facial information and provides a higher classification rate than other existing promising methods.

Keywords: Adaboost Face Detection Technique, Pyramidal Lucas Kanade, SURF and SVM Classifier

I. INTRODUCTION

Face expression is used to convey information to humans without using words and it is natural way to convey any information. In past years people use to express their opinion in textual form but nowadays it changes, people are expressing their views in video format through whatsapp, face book etc.Face detection is an active research in an image sequence for computer vision because it has many applications like monitoring classifying images, smart rooms, biomedical, image analysis, human computer interface in robots, ete. Facial expressions are recognized by three expressions such as anger, happy, and fear. Support vector machine classifier is used to recognize different facial expressions. Support vector machine is a machine learning technique used for both classification and training, but most widely used in classification.

There are few methods for face associated problems were distinct face feature extraction methods are introduced based upon feature types. Facial feature extraction method divided into two categories: geometric feature based method and appearance based method. Feature vectors are formed based on geometric relationships. In geometric approach feature vectors are position,distance,angle between eye,ears,nose etc.Facial action coding system is a geometric feature based method this facial action coding system represents different expressions by using group of action units. Each and every action unit indicates behaviour of facial muscle. Geometric based method depends on error free detection of facial part. Appearance based method .Appearance based method takeout facial aspect by putting image filter or filter bank on entire face image.

Face attributes are used in many applications they are for face verification, identification and retrieval. Prediction of face attributes from an image is challenging because of poses, lightings and occlusions. Attribute recognition methods are generally categorized into two groups: global method and local method. In global method it extract the features from entire object but accurate location and landmark does not required. And this method proposes frame work to integrate CNNs i.e LNet and ANet.LNet is used to recognize entire face and ANet extract high-level face. Local method it first detect part of the object and it extract the features.

Due to it's the wide range applications, the facial expression recognition system has become more vital role and its efforts are taken for increase the efficiency. Facial expressions are recognized by three expressions such as anger, happiness and fear. To recognize these expressions we are using SVM classifier.

Emotions are part and parcel of human life and among other things, highly influence decision making. Computers have been used for decision making for quite some time now but have traditionally relied on factual information. Recently, interest has been growing among researchers to find ways of detecting subjective information used in images and videos. These six basic emotions are namely, happiness, sadness, surprise, fear, disgust and anger.The human face is central to our identity. Human beings interact with one another not only through words of mouth but also through other methods such as gestures and facial emotions. If computers can also be intelligent enough to perceive human emotions, then human computer interaction will change for the better, leading to new ways of interacting with computers and perhaps new set of computer applications could emerge.

One of the major development in machine learning is Ensemble method, which find high accurate classifier by combining many moderately accurate classifier by component classifiers. Support vector machine was developed based on the theory of structural risk minimization. By using a kernel trick to map the training samples from an input space to a high dimensional feature space, SVM finds an optimal separating hyper plane in the feature space and uses a regularization parameter, C, to control its model complexity training error.

Support vector machines (SVMs) are formulated to solve a classical two class pattern recognition problem. We adapt SVM to face recognition by modifying the interpretation of the output of a SVM classifier and devising a representation of facial images that is concordant with a two class problem. Traditional SVM returns a binary value.

II PROPOSED METHOD

In this proposed method mainly we are focusing on facial expression detection from human emotions. Adaboost face detection technique is using to detect the human face and SURF and Pyramidal Lucas Kanade are generated from the features to produce the expression feature vectors. Finally, the objective sequences of the feature vectors are trained through the Support Vector Machine (SVM) to produce the expression model. It represents only the changing information from an image, the significant changes can be easily captured which occur because of the facial expression. Results of proposed model gives better recognition rate for depth based facial expression recognition system. Fig. 1 represents proposed system architecture.

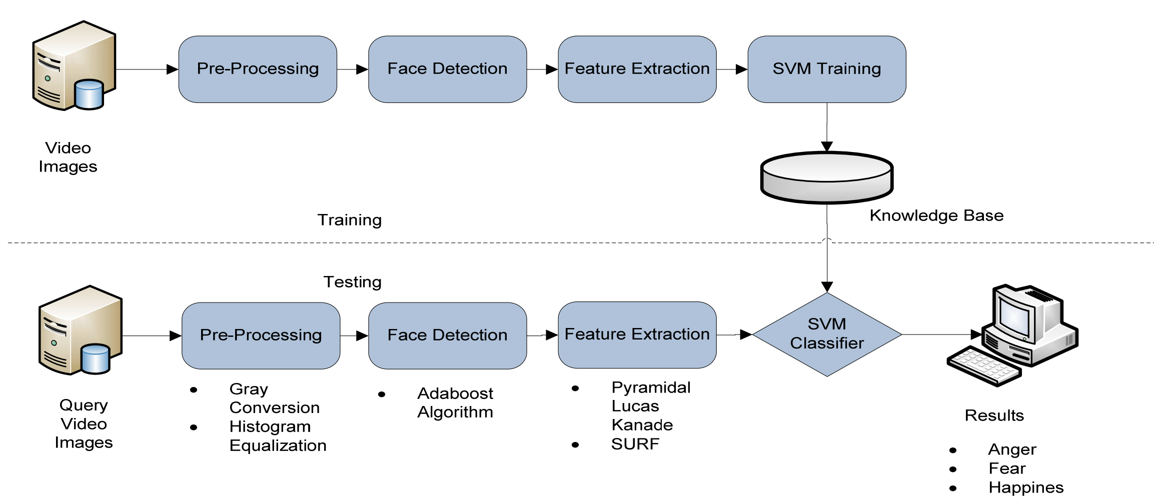

Fig. 1: Block Diagram of Proposed System

2.1Pre-processing

In this pre-processing step, we are applying two techniques such as gray conversion and histogram equalisation. To decrease the information of input images, image must convert into gray scale. Usually, Input images are in the form of RGB, {(R1, G1, B1),….(Rn, Gn, Bn)} where R, G, B depicts the value of red, green and blue respectively and n represents the total number of pixel in given image. 2. Thus the new gray scale images are a pixel from G1…Gn. The formula of gray scale image is as follows,

$$G = 0.21R + 0.71G + 0.07\,B \qquad (1)$$

Then apply the histogram equalization technique, this technique is works statistically of average scattered in histogram.

2.2Face Detection

The face detection method is used for our implementation. This method looks for exact Haar features of face image. These features are found and allow the human face to move to next stage of detection. The human face is the rectangular part of original image which can known as sub window. These sub-windows are fixed size that is 24x24 pixels. This sub-window frequently scaled to obtain the variety of various face size. The face detection technique is scaned complete image with sun window and defines each individual part of face image.

A Adaboost Algorithm

The integral image can be defined as summation of pixels of input original image. The location (x, y) value of integral image is equal to the addition of image pixels and left of the location (x,y).

The integral image consists sum of the pixel of x and y at location to the left of x and y, general

$$\mathrm{ii}(x,y) = \sum_{x' \le x,\, y' \le y} \mathrm{i}(x',y') \qquad (2)$$

Where $\mathrm{i}(x,y)$ represents the pixel value of original image

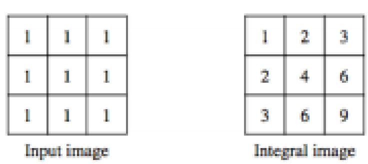

Fig 2.2: Example of Integral Image

$$X\,(x,y) \in ABCD\ \mathrm{i}(x,y) = \mathrm{ii}(D) + \mathrm{ii}(A) - \mathrm{ii}(B) - \mathrm{ii}(C) \quad (3)$$

$\mathrm{ii}(x',y')$ is corresponding their image integral value. By using the integral value to calculate the sum of rectangle area which can be more efficient is depicts in above figure. The sum pixel in rectangle of ABCD is computed by only four values from an integral image;

$$\sum_{(x,y)\in ABCD} \mathrm{i}(x,y) = \mathrm{ii}(D) + \mathrm{ii}(A) - \mathrm{ii}(B) - \mathrm{ii}(C) \qquad (4)$$

These features can be composed of two or three rectangles. Human face images are scanned also found Haar features of present stage. The weights were generated constantly by learning technique.

Each Haar feature value can be computed by taking the region of each rectangle and multiply their respective weights also adding the results. The area of rectangle can easily calculated by integral image. Then coordinate of any corner of rectangle is used to get all pixel values for summation and using integral image to left of that location. Thus L1 subtracted twice it should add back to correct the rectangle area. This area of rectangle is depicts the integral rectangle and calculated the locations of integral image i.e L4-L3-L2+L1.

2.3Feature Extraction

A. PyramidalLucasKanade

high level of the lowest detail image (image pyramid) also works down to finer detail (low levels). Tracking image pyramids are permits large motions is to be caught the local windows.

Suppose the local patch of the pixel is moved coherently so that we can easily find for the motion of center pixel by surrounding pixels to set up the system equation. Hence, the recommended method is to solve for optical flow at top layer then is use to resulting motion estimation as the starting point for next down layer.

Lucas Kanade technique attains an optimization velocity by least square estimation principle.

$$\nabla I^T U = -I_t \qquad (5)$$

$$\nabla I = [I_x, I_y]^T \; and \; U = [u \; v]^T \quad (6)$$

Therefore the least square $U = [u \; v]^T$ estimation of the function error can be written as

$$\xi(u, v) = \sum\sum [I(x + \delta x, y + \delta y, t) - I(x, y, t)]^2 \quad (7)$$

Consider small displacement and constant by the neighbouring object point. The Taylor expansion of $I(x + \delta x, y + \delta y, t + \delta t)$ can be written as,

$$I(x + \delta x, y + \delta y, t + \delta t) = I(x, y, t) + u\frac{\partial I}{\delta x} + v\frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} \quad (8)$$

Substitute this eq.2 in eq.1 can be give as,

$$\xi(u, v) = \sum\sum uI_x + vI_y + {I_t}^2 \quad (9)$$

The optimized rates of the flow velocities are achieved by minimize the function error with respect to u and v. This may leads to the system of linear formulas; representation of matrix is generating the standard forms of the LK optical flow expression.

$$U = E^{-1}d$$

Where, $E = \begin{bmatrix} \sum\sum I_x^2 & \sum\sum I_x I_y \\ \sum\sum I_x I_y & \sum\sum I_y^2 \end{bmatrix}$

and $d = -\begin{bmatrix} \sum\sum I_x I_t \\ \sum\sum I_y I_t \end{bmatrix}$ intrinsically global nature of the horn and schunck algorithm. The drawbacks of using this small local window is in lucas kanade is that large motions are move the outside points of local window. Therefore it becomes impossible for the algorithm to find. This difficulty leads to develop the pyramidal lucas kanade technique which can tracks from algorithm. It can be applied in sparse context thus it relies only when local information is derived from some window surrounding each points of interest. This method is easily applied to the subset of points in input image and it becomes more important sparse This is in contras

B. SURF

$$DoH(x, y, \sigma) \approx \frac{D_{xx}(x, y, \sigma). D_{yy}(x, y, \sigma) - [\beta. D_{xy}(x, y, \sigma)]^2}{\sigma^2} \quad (13)$$

$$\beta = \frac{\|G_{xy}(x, y, \sigma)\|_F . \|D_{yy}(x, y, \sigma)\|_F}{\|G_{yy}(x, y, \sigma)\|_F . \|D_{xy}(x, y, \sigma)\|_F} \approx 0.9 \quad (14)$$

The term of compensation β is inserted to compensate for error which introduced by approximating the true Gaussian derivative masks in computation of DoH. This multiplicative factor is defined as the ratio of Frobenius norms $\left(\|A(x, y, \sigma)\|_F \equiv \sqrt{\sum_{x,y}|A(x, y, \sigma)|^2}\right)$ between the box filter estimates with their true counterparts. Even though β dependent on scale size σ, it turns in practice β is estimated by static constant of 0.9.

2.3SVM Classifier

Classification is the collection of various techniques for separating the data. Classify the data is to sort out the features that can separate each group of the information into clusters. Before the classification this is common to practice to normalize the information. By assuming all features has same importance so that will make the distance in different ways to carry out the same weight.

SVM is also known as maximum margin classifier. Consider one simple example of two linear separable clusters in 2D space. This happens to be the solution which can maximize the margin. This margin could be the distance between black and gray lines is shows in Fig. 5. Consider an equation for hyperplane to understand how it can be done is shows below,

$$w \cdot x + b = 0 \quad (15)$$

Where w represents the non-zero vector normal to hyperplane, x presents the point in same space as hyperplane and b represents the scalar. This indicates the equation remains same without dimensionality. Then consider two extra hyperplanes which can normalized is given as

$$w \cdot x + b = \pm 1 \quad (16)$$

These hyperplanes can be dashed gray lines. The distance from given xi point is shows below.

$$d(w, b; xi) = |w \cdot xi + b| \, ||w|| \quad (17)$$

Support vectors can be described as points on normalized hyperplane is shows in Eq. 15 and Eq. 16 is set to as 1.points which cannot be a support vectors doesn't have influence on classification. Eq. 17 is simplified to

$$d(w, b; |w \cdot x + b| = 1) = 1 \, ||w|| \quad (18)$$

Remember that the distance is margin thus should maximize this term in maximize the distance.

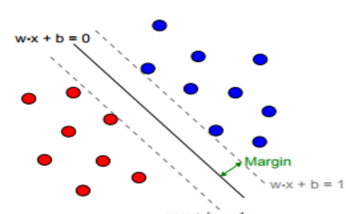

Fig. 2.3: The hyper plane separating two classes, and its equation.

Maximize the term in Eq. 18 which is not friendly optimization problem. However, the maximizing ||w||−1 can

III.COMPARISION TABLE

Author	Problem Definition	Method	Performance Parameter	Limitations
Valstar et .al(1) 2013	Facial expression recognition	LBP,SVM	Average, Accuracy	Cannot fully capture expression sequence
Li et.al(2) 2014	Facial action unit Recognition using Data Free Prior Model	ADABOOST,BN/DBN prior model	Average, Confusion Matrix	Poor Representation of Facial Expression
Chongliang et.al(3) 2015	Facial Expression Recognition	HMM,SVM	Accuracy, Confusion Matrix	More error rates
Wang et.al(4) 2013	Simultaneous Facial Feature Tracking & Facial Expression Recognition	DBN	Accuracy	Poor Representation of Facial Expression
Majumder A et.al(5) 2016	Automatic facial expression Recognition system	LBP	Accuracy, Recognition rates	Poor Representation of Facial Expression
Zhang et.al(6) 2017	Facial Expression Recognition	SVM	Accuracy, Average	More error rates
Kabir et.al(7) 2017	Facial Expression Recognition	POMF	Accuracy, Confusion matrix, Recognition rates, Average	POMF uses only linearity

III CONCLUSION

In this paper, facial expression recognition system is proposed. This proposed system is recognized with the help of SVM Classifier. This recognition is increasing application areas and requires more accurate and reliable system. The results are shows that increase the accuracy rate of predicting the expressions. Our proposed system is gives better accuracy than the existing system that is 90%.

VI REFERENCES

[1]Valstar MF, Mehu M, Jiang B, Pantic M and Scherer K, "Meta-Analysis Of The First Facial Expression Recognition Challenge", IEEE, Vol. 42, No.4 , pp. 966-79, 2013.

[2]Li Y, Chen J, Zhao Y and Ji Q "Data-Free Prior Model for Facial Action Unit Recognition", IEEE Transactions on affective computing, Vol. 4, No. 2, pp. 127-41, 2013.

[3] Li Y, Wang S, Zhao Y and Ji Q, "Simultaneous Facial Feature Tracking And Facial Expression Recognition", IEEE Transactions on Image Processing, Vol. 22, No. 7, pp. 2559-73, 2013.

[4]Almaev TR and Valstar MF, "Local Gabor Binary Patterns From Three Orthogonal Planes for Automatic Facial Expression Recognition", IEEE on Humaine Association Conference, pp. 356-361, 2013.

[5]Agrawal S and Khatri P "Facial Expression Detection Techniques: Based On Viola And Jones Algorithm And Principal Component Analysis" IEEE on Advanced Computing & Communication Technologies, pp. 108-112, 2015.

[6]Wu C, Wang S. and Ji "Multi-Instance Hidden Markov Model For Facial Expression Recognition" IEEE on Automatic Face and Gesture Recognition, Vol.1, pp. 1-6, 2015.

[7] Facial Expression Recognition From Depth Video With Patterns of Oriented Motion Flow MD. HASANUL KABIR1, (Member, IEEE), MD SIRAJUS SALEKIN1,MD. ZIA UDDIN2, (Member, IEEE), AND M. ABDULLAH-AL-WADUD3, (Member, IEEE)