© 2018 IJNRD | Volume 3, Issue 3 March 2018 | ISSN: 2456-4184

FACIAL EMOTION RECOGNITION USING LBP AND SVM TECHNIQUE

[1]Shubham Meshram, [2]Shashipal Panwar, [3]Mustak Soloanki, [4]Akshay Waghmare, [5]Prof. Muzzafar Khan
1,2,3,4B.E. Student, Department of ETC , ACET, Nagpur, Maharashtra, India
5Asst. Professor, Department of ETC , ACET, Nagpur, Maharashtra,India

Abstract: Facial expressions convey non-verbal cues, which play an important role in interpersonal relations. Automatic recognition of facial expressions can be an important component of natural human-machine interfaces; it may also be used in behavioural science and in clinical practice. Automatic facial expression recognition system has many applications including, but not limited to, human behavior understanding, detection of mental disorders, and synthetic human expressions. Two popular methods utilized mostly in the literature for the automatic FER systems are based on geometry and appearance. Even though there is lots of research using static images, the research is still going on for the development of new methods which would be quiet easy in computation and would have less memory usage as compared to previous methods. This paper presents a quick survey of facial expression recognition. A comparative study is also carried out using various feature extraction techniques on JAFFE dataset. Support Vector Machine (SVM) is used as a classifier for classifying the expressions of supplied face into five basic expressions like surprise, neutral, sad,happy and angry.

IndexTerms—Facial,Emotiong,LBP,SVM,Jaffe Database.

I. INTRODUCTION

In the early few years several papers have been published on face detection in the community which discusses different technique like neural network, edge detectors and many more. There is a good survey by Chellapa, Wilson and Sirohey (1995) which tells about the trends of paper in face detection [11]. Previously, many researchers and engineers have designed different purpose specific and application specific detectors. The main goal of this kind of classifiers was to achieve a very high detection rate along with low computational cost. Few examples of different detectors are corner detectors, AdaBoost Haar Cascade detector by Jones and Viola [10] are very useful. This kind of detectors mostly use simple and fast classifiers that reject the most common negative samples and then they use progressively more complex classifiers to deal with the more difficult and odd negative samples. Another approach to detect the face is through skin color classification algorithm [4]. Kjeldsen and Kender used the concept of different color space model to separate the skin patches from the image [18]. But if those algorithms are used solely, then it becomes difficult to detect more than one face from the image. Hsu, Abdel-Mottaleb, and Jain came up with the tactics of calculating the eye, mouth and nose map values [4]. Eyes, nose and mouth are significant features of face which distinguishes face from other 12 parts of the body. Studies have shown that they exhibit unique properties in the YCbCr color model, so it becomes easy to detect the face region using this feature extraction. Poggio, Heisele and Ho presented a component based approach of locating facial components, extracting them and combining into a single feature vector which is used for classification of faces by Support Vector Machine. The authors used the gray scale image to define the feature vector for classification. Authors' using the approach of Support Vector Machine iterates through the whole image and compares it with face template to classify the region of interest. This takes very high computation time and error rate. Also, some of them can detect faces only from gray scale images. But our algorithm overcomes these drawbacks of classification. In our approach, preprocessing stages are applied which extracts the region of interest and feature vector is applied on this ROI instead of whole image for classification of the face region.

In FER, face detection is a crucial task. There are different approaches for this, namely appearance based approach, template based approach, feature based approach, and the local-global graph approach [9].. Nevertheless, one limitation of this approach is that it is only accurate in frontal images that are well illuminated and have a simple background. In the template based approach, a standard face pattern is generated and correlated with the image to find the face. The limitation with this approach is the difficulty to generalize to different shapes, sizes and poses. In the feature based approach, not-changing facial features are detected. Then, candidate faces are grouped and verified. The limitation with this approach is it is difficult to find features in pictures that do not have simple backgrounds or are well-illuminated. Finally, the local-global graph approach generates a graph with nodes created according to colour similarity which are, to an extent, invariant to translation, rotation and scale. The limitations with this approach are the low-image size, and the poor-image quality. In the feature extraction phase, the objective is to obtain discriminatory and stable facial features according to one of two perspectives, geometric feature-based methods, and appearance-based methods. On the one hand, geometric feature-based methods involve defining geometric relationships of facial components in terms of their location and shape.

I. LITERATURE REVIEW

The use of fiducial points on the face with the geometric positions and multi-scale and multi-orientation Gabor wavelet coefficients have been investigated by [Zhang, 1998; Zhang, 1999]. The papers describe the integration within an architecture based on a twolayer perceptron. According to the results reported by the author, the Gabor wavelet coefficients show much better results for facial expression recognition, when compared to geometric positions. In the paper [Fellenz et al., 1999] the authors compare the performance and the generalizing capabilities of several low-dimensional representations for facial expression recognition in static pictures, for three separate facial expressions plus the neutral class. Three algorithms are presented: the template-based determined by computing the average face for each emotion class and then performing matching of one sample to the templates, the multi-layered perceptron trained with the back-propagation of error algorithm and a neural algorithm that uses six odd-symmetric and six even-symmetric Gabor features computed from the face image. According the authors, the template-based approach presented 75% correct classification while the generalization achieved only 50%, the multilayered perceptron has 40% to 80% correct recognition, depending on test the data set. The third approach did not provide an increase in the performance of the facial expression recognition. Several research works study the facial dynamics of recognition of facial expressions. The work of [Yacoob and Davis, 1994] uses optical flow to identify the direction of rigid and non-rigid motions shown by facial expressions. The results range from 80% for sadness to 94% for surprise emotion on a set of 46 image sequences recorded from 30 subjects, for six facial expressions. Some attempts to automatically detect the salient facial features implied computing descriptors such as scale-normalized Gaussian derivatives at each pixel of the facial image and performing some linear-combinations on their values. It was found that a single cluster of Gaussian derivative responses leads to a high robustness of detection given the pose, illumination and identity [Gourier et al., 2004]. A representation based on topological labels is proposed [Yin et al., 2004]. It assumes that the facial expression is dependent on the change of facial texture and that its variation is reflected by the modification of the facial topographical deformation. The classification is done by comparing facial features with those of the neutral face in terms of the topographic facial surface and the expressive regions. Some approaches firstly model the facial features and then use the parameters as data for further analysis such as expression recognition. The system proposed by [Moriyama et al., 2004] is based on a 2D generative eye model that implements encoding of the motion and fine structures of the eye and is used for tracking the eye motion in a sequence. As concerning the classification methods, various algorithms have been developed, adapted and used during time [Pantic and Rothkrantz, 2000]. Neural networks have been used for face detection and facial expression recognition [Stathopoulou and Tsihrintzis, 2004] [deJong and Rothkrantz, 2004]. The second reference directs to a system called Facial Expression Dictionary (FED) [deJong and Rothkrantz, 2004] that was a first attempt to create an online nonverbal dictionary. The work of [Padgett et al., 1996] present an algorithm based on an ensemble of simple feed-forward neural networks capable of identifying six different basic emotions. The initial data set used for training their system included 97 samples of 6 male and 6 female subject, considered to portray only unique emotions. The overall performance rate of the approach was reported to be 86% on novel face images. The authors use the algorithm for analyzing sequences of images showing the transition between two distinct facial expressions. The sequences of images were generated based on morph models. [Schweiger et al., 2004] proposed a neural architecture for temporal emotion recognition. The features used for classification were selected by using optical flow algorithm in specific bounding boxes on the face. Separate Fuzzy ARTMAP neural networks for each emotional class were trained using incremental learning. The authors conducted experiments for testing the performance of their algorithms using Cohn-Kanade database. Other classifiers included Bayesian Belief Networks (BBN) [Datcu and Rothkrantz, 2004], Expert Systems [Pantic and Rothkrantz, 2000] or Support Vector Machines (SVM) [Bartlett et al., 2004]. Other approaches have been oriented on the analysis of data gathered from distinct multi-modal channels. They combined multiple methods for processing and applied fusion techniques to get to the recognition stage [Fox and Reilly, 2004]. The work of [Bourel et al., 2001] presents an approach for recognition of facial expressions from video in conditions of occlusion by using a localized representation of facial expressions and on data fusion. In [Wang and Tang, 2003] the authors proposed a combination of a Bayesian probabilistic model and Gabor filter. [Cohen et al., 2003] introduced a Tree-AugmentedNaive Bayes (TAN) classifier for learning the feature dependencies. The system presented in [Bartlett et al., 2003] is able to automatically detect frontal faces in video and to code the faces according to the six basic emotions plus the neutral. For the detection of faces, the authors use a cascade of feature detectors based on boosting techniques. The face detector outputs image patches to the facial expression recognizer. A bank of SVM classifiers make use of Gabor-based features that are computed from the image patches. Each emotion class is handled by a distinct SVM classifier. The paper presents the results for the algorithms that involved linear and RBF kernels for being used by the SVM classifiers. Adaboost algorithm was used for select the relevant set of features from an initial set of 92160 features. [Jun et al., 2000] propose a system for the recognition of facial expressions based on Independent Component Analysis - ICA algorithm and Linear Discriminant Analysis - LDA. The algorithm implies the use of ICA for obtaining a set of independent basis images of the face image and the use of LDA for selecting features obtained by ICA. The authors provide the results for each experimental setup for the use of the two methods. They report the highest recognition ration for the case of using LDA and ICA together (95.6%). [Feng et al., 2000] proposes a sub-band approach in using Principal Component Analysis – PCA. In comparison with the traditional use of PCA namely for the whole facial image, the method described in the paper gives better recognition accuracy. Additionally, the method achieves a reduction of the computational load in the cases the image database is large, with more than 256 training images. The facial expression accuracy ranges from 45.9% in the case of using 4X4 wavelet transform to 84.5% for a size of 16X16. The accuracy of the recognition is improved by 6%, from 78.7% in case of using the traditional approach to 84.5% using sub-band 10. The analysis are carried for wavelet transform ranging from sub-band 1 to sub-band 16 and full size original image.

II. FACIAL EXPRESSION RECOGNITION

The Second phase is the facial expression recognition which is accomplished in three stages. First the facial image is preprocessed by resizing it and converting it to grayscale. Then ,feature extraction is done using SVM Support Vector Machine with Kernel algorithm. A.Facial Image Preparation In order to recognize the facial expression in the cropped image of the previous phase, the image has to be resized to 64 x 64 pixels. Next the RGB image is converted into grayscale by eliminating the hue and saturation information while retaining the luminance. B. System Configuration and Support vector Machine Learning Procedure. In machine learning, task of deducing a category from supervised training data is known as Supervised Learning. In supervised learning the training data consist of a set of training examples, where each example is a pair consisting of an input and an anticipated output value. A supervised learning algorithm analyzes the training dataand then predicts the correct output categorization for given data-set input. For e.g. Teacher teaches student to identify apple and oranges by giving some features of that. Next time when student sees apple or orange he can easily classify the object based on his learning from his teacher, this is called supervised learning. He can identify the object only if it is apple or orange, but if the given object was grapes the student cannot identify it.

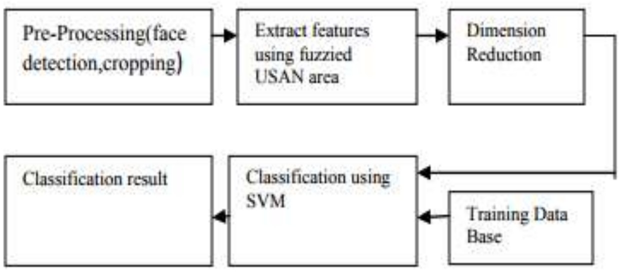

Figure 1: Block Diagram of Face Recognition using Support Vector machine

There are many folds advantages of using the supervised learning approach of Support Vector Machine (SVM). They are very effective when we have very high dimensional spaces. Also, when number of dimensions becomes greater than the existing number of samples, in such cases too SVM is found to be very effective. SVM uses a subset of training point also known as support vectors to classify different objects hence it is memory efficient. Support Vector Machines are versatile, for different decision function we can define different kernel as long as they provide correct result. Depending upon our requirement and application we can choose types of kernel which is most productive for our application.

The facial expression recognition system is proposed to recognize facial expressions to identify five universal human expressions: neutral, happy, anger, sad and surprise using the face images. Figure 2 illustrates the flow chart of the facial expression recognition system. This system is divided into four phases: face detection, mouth segmentation, feature extraction from the mouth region and finally classification

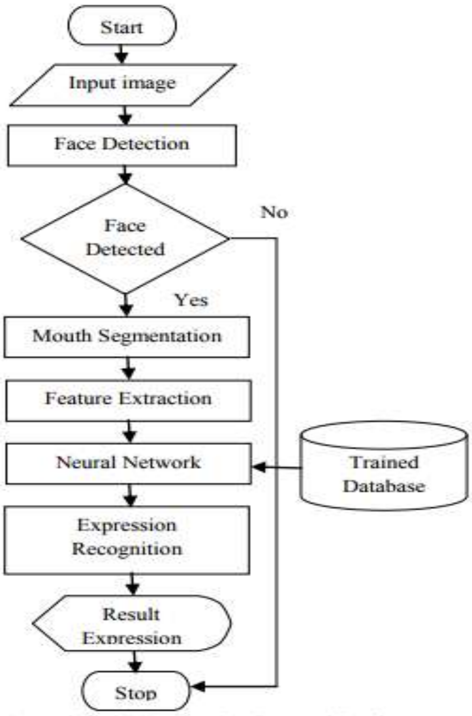

Figure 2: Flow chart of proposed System

III. SUPPORT VECTOR MACHINE

SVMs can be considered as a new paradigm to train polynomial function, neural networks, or radial basis function (RBF) classifiers. While most of the techniques used to train the above mentioned classifiers are based on the idea of minimizing the training error, which is usually called empirical risk, SVMs operate on another induction principle, called structural risk minimization, which minimizes an upper bound on the generalization error. An SVM classifier is a linear classifier where the separating hyper plane is chosen to minimize the expected classification error of the unseen test patterns. This optimal hyper plane is defined by a weighted combination of a small subset of the training vectors, called support vectors. Estimating the optimal hyper plane is equivalent to solving a linearly constrained quadratic programming problem.

3.1 Classification using SVM Suppose there exist a dataset D = {(xi , yi)}l i=1 of labeled examples and yi Є {-1,1}, the aim is to determine, among the infinite number of linear classifiers that separate the data, which one will have the smallest generalization error [61]. In this step a hyperplane that leaves the maximum margin between the two classes, where the margin is defined as the sum of the distances of the hyper plane from the closest point of the two classes, can be used. (Fig 3.3)

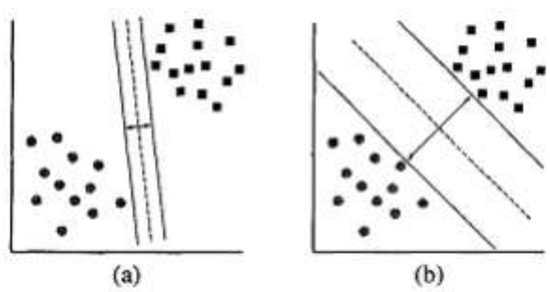

Figure 3: Separating hyperplane with a small margin (a) and a larger margin (b)

Even if the two classes are non-separable, it can be still searched for the hyper plane that maximizes the margin and that minimizes a quantity proportional to the number of misclassification errors. A positive constant C that is chosen before hand controls the trade off between margin and misclassification error. In this case it can be shown that the solution to this problem is a linear classifier

IV. RESULT AND CONCLUSION

The performance was through MATLAB. The proposed algorithm was trained and evaluated on the dataset of around 125 images containing 305 face images. This dataset was build from my collection of photographs and some random images from internet. The test images consisted of images with different lumination condition night time, daytime and combination of them. The image formats acceptable to the algorithm are jpeg, png, bmp, etc. The dataset consist of images of size ranging from 400x320 to 2000x1800. If the size of the image is more than 2000x1800 then it would create problem in processing the image. In termination face detection algorithm using the skin color detection, edge detection, facial feature extraction and using the concept of different color space. After these pre processing stages, the algorithm utilizes the highly powerful concept of Support Vector Machine (SVM) to classify the image into face and non-face region. We have significantly reduced the misclassification errors. The computation time for our algorithm is very less and the accuracy on the image data set of 125 images with 305 face image is around 90% with error rate of approximate 16%. We overcame the limitation of detecting one face from image using skin color algorithm; by combining the concept of different color space and face feature extraction process.

V. REFERENCES

[1] (1999). Garcia, C., & Tziritas, G,"Face detection using quantized skin color regions merging and wavelet packet analysis". Multimedia, IEEE Transactions on, 1(3), 264-277.

[2] (2002). Hsu, R. L., Abdel-Mottaleb, M., & Jain, A. K. "Face detection in color images. Pattern Analysis and Machine Intelligence", IEEE Transactions on, 24(5), 696-706.

[3] (1997)Osuna, E., Freund, R., & Girosi, F. "Training support vector machines: An application to face detection". Paper presented at the Cvpr, 130.

[4] (2002) Rein-Lien Hsu, Abdel-Mottaleb, M., & Jain, A. K." Face detection in color images.", Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(5), 696-706.

[5] (1998) Rowley, H. A., Baluja, S., & Kanade, T. " Neural network-based face detection",. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 20(1), 23-38

[6] (2004) Yang, G., & Huang, T. S. (1994)." Human face detection in a complex background",. Pattern Recognition, 27(1), 53-63.

[7] (2006). Shih, P., & Liu, C.," Face detection using discriminating feature analysis and support vector machine". 60

[8] (2001)KADIR, T., BRADY, M.. Saliency, Scale and Image Description. International Journal of Computer Vision, Volume 45, Number 2.

[9] (2001 P. Viola and M. Jones. Rapid Object Detection Using a Boosted Cascade of Simple Features. IEEE Conf. CVPR, Vol. 1, pp. 511-518.