Research Article **Open Access**

Using Support Vector Machine and Local Binary Pattern for Facial Expression Recognition

Ayeni Olaniyi Abiodun [1], Alese Boniface Kayode[1], Dada Olabisi Matemilayo[2]

[1.] Department of Computer Science, Federal University Technology Akure, PMB 704, Akure, Nigeria.

[2.] Department of computer science, Kwara state polytechnic Ilorin, Kwara-State, Nigeria.

. *Corresponding author: Ayeni Olaniyi Abiodun Mail Id: oaayeni@futa.edu.ng

ABSTRACT

Facial expressions are natural ways by which people can express their feelings and emotions. In the field of affective computing and human computer interactions, a better result will be achieved when we have an intelligent interface(s) that could act and behaves in a way similar to that of human being. This research intends to bring about the development of a face recognition model and applying it to a real-data set of expressions. Five expressions will be classified which include: fear, happiness, disgust, sadness and surprise using the innovations of support vector machine (SVM) and local binary pattern (LBP). The students of Federal University of Technology, Akure (FUTA) will be used as a case study. LBP will be used for feature extraction while SVM will be used for classification and recognition of expressions.

   Keywords: facial expression, detection, recognition, support vector machine, principal local binary pattern & face feature.

INTRODUCTION

Facial expressions are the changes in the appearance of a face according to a person's internal emotional states, intentions, or social communications (Kamal et al, 2012). Facial expressions play an important role wherever Human-computer interfaces is required.

Facial expression is one of the most powerful and immediate means for human beings to communicate their emotions, intentions, and opinions to each other. Facial expressions also provide information about cognitive state, such as interest, boredom, confusion, and stress. Facial expressions are natural and can express emotions sooner than people verbalize their feelings. It conveys non-verbal cues, which play an important role in interpersonal relations. Facial expressions recognition technology helps in designing intelligent human computer interfaces. . Recognizing the expression of a man can help in many of the areas like in the field of medical science where a doctor can be alerted when a patient is in severe pain. It helps in taking prompt action at that time. Expression is a basic way to express mankind's feelings and is one kind of effective communication. Since last three decades of face recognition technology, there exists many commercially available systems to identify human faces, however face recognition is still an outstanding challenge against different kinds of variations like facial expressions, poses, non-uniform light illuminations, occlusions and aging effects.

The facial expressions have corresponding change before people expressing their emotions. The facial expression can not only express their thoughts and feelings accurately and subtly, but also describe the others' cognitive attitudes and inner world. Facial expression contains rich human behaviours and is a kind of information resources in human-computer interaction

with more effective, natural and direct way. If computers and robots have the ability to understand and express feelings as men's adapting to the environment, it will change the relationship between the computer and human fundamentally. If that, the computer can offer a better service to mankind. Detection of face features such as eyes and mouth have been major issues of facial image processing which may be required for various areas such as emotion recognition and face identification. Face feature detection can be used to determine the face features from images to be used later as input for other functions like face and emotion recognition.

Facial expression plays an important role in smooth communication among individuals. The extraction and recognition of facial expression has been the topic of various researches subject to enable smooth interaction between computer and their users. In a way, computers in the future will be able to offer advice in response to the mood of the user. Computer-based recognition of facial expressions goes a long way to help in this area and various methods have been proposed. All the method can be classified into two broad-based categories: probabilistic approach and feature based approach. The feature-based method utilizes the Facial Action Coding System (FACS) designed by Ekman and Friser . In FACS, the motions of the face are divided into 44 action units (AU), and their combinations may describe any facial expression. More than 7,000 combinations of AU have been observed.

Hehua Chi, Lianhua Chi, Meng Fang, Juebo Wu**: Facial Expression Recognition based on Cloud model.**

This work was based on the transformation from images into grids with M by N, where M and N denote the actual image

positioning of the grid. Each grid is a gray value (0-255) and the grids stand for the data from data points to data sets based on cloud model. Cloud model has three characteristics: Expectation (Ex) is the prototype value (centre or standard value) of concept, and is the most representative value of the qualitative concept. Entropy (En) is the measurement of concept uncertainty while Hyper-entropy (He) is the measurement of entropy uncertainty, that is, the entropy of entropy. The quantitative numerical characteristics {Ex, En, He} of facial expressions were mined by the backward generator of cloud model. In this work, the hidden knowledge in facial expression images were obtained with the numerical characteristics {Ex, En, He} of cloud model. Ex is the characteristics of the facial image in common, En is the personality deviation of general common knowledge, and He is the discrete level of knowledge. In analyses of facial image knowledge, by the numerical characteristics {Ex, En, He}, the facial expression can be realized

Salih Burak Gokturk, Jean-Yves Bouguet, Carlo Tomasi & Bernd Girod (2002): **Model-Based Face Tracking for View-Independent Facial Expression Recognition.**

This research work proposes a new method for robust, view-independent recognition of facial expressions that allows the user faces the camera to change his head pose. The system uses a novel 3D model-based tracker to extract simultaneously and robustly the pose and shape of the face at every frame of a monocular video sequence. There are two main contributions of this work. First, it demonstrate that the 3D information extracted through 3D tracking enables robust facial expression recognition in spite of large rotational and translational head movements (up to 90 degrees in head rotation). Secondly, it shows that Support Vector Machine is a suitable engine for robust classification.

Future work. In the future, more work on performing another set of experiments with more subjects and expressions is encouraged. One important objective is then to build a generic parameterized static face model and use it along with the principal movement directions for tracking and expression recognition of any generic person. Another direction is to investigate applications of the tracking approach to different recognition problems such as face recognition, and lip reading.

Kamal Kumar Ranga and Arpita Nagpal (2012): Facial Expression Recognition System: A practical Implementation.

In this research work, facial expression recognition technique has been performed on the Indian faces extracted from a video. Initially, a live video of Indian college students is given as input to Haar classifier which traces out the faces from it. Then 42 facial feature points are detected using Active Appearance Model (AAM) technique using which it extract the facial features that are to be mapped on the extracted faces. In the last step four primary facial expressions (happy, sad, surprise, angry) have been classified using the technique support vector machine (SVM).

Haar basis function uses 3 kinds of features:

1. Two- rectangle feature- Difference between sums of pixels within two rectangular regions. Type-1 & 2 in fig.

2. Three rectangle features- Computes the sum within two outside rectangle and subtracts it with the sum in the center rectangle. Type 3 & 4 in fig.

3. Four-rectangle features- difference between diagonal pairs of rectangle. Type 5 in fig.

Each feature results in a single value which is calculated by subtracting the sum of the white rectangle(s) from the sum of the black rectangle(s).
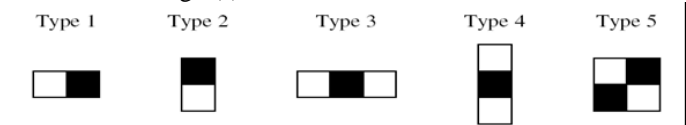
The limitation of this work is that the number of training images were too small. In future, Support vector machine classifier can be replaced by Neural Network and it may give different performance**.**

L H Koh, S Ranganath, M W Lee and Y V Venkatesth: An integrated Face Detection and Recognition System.

This work presents an integrated approach to unconstrained face recognition in arbitrary scenes. The front end of the system comprises of a scale and pose tolerant face detector. Scale normalization is achieved through novel combination of a skin color segmentation and log-polar mapping procedure. Principal component analysis was used with the multi-view approach to handle the pose variations. For a given color input image, the detector encloses a face in a complex scene within a circular boundary and indicates the position of the nose. Next, for recognition, a radial gridmapping centered on the nose yields a feature vector within the circular boundary. As the width of the color segmented region provides an estimated size for the face, the extracted feature vector is scale normalized by the estimated size. The feature vector is input to a trained neural network classifier for face identification. The system was evaluated using a database of 20 person's faces with varying scale and pose obtained on different complex backgrounds. Experimental face detection rates of more than 95% was achieved. The face recognition scheme considered uses the detected face image and extracts gradient features that are encoded by a novel radial grid method and subsequently input to neural network classifiers for classification. The number of faces used need to be increased.

Zahid Riaz, Christopher Mayer, Matthias Wimmer, Michael Beetz and Bernd Radig (2009) : A Model Based Approach for Expressions Invariant Face Recognition

This research work describes an idea of recognizing the human face in the presence of strong facial expressions using model based approach. The features extracted for the face image sequences can be efficiently used for face recognition. The approach follows in 1) modeling an active appearance model (AAM) parameters for the face image, 2) using optical flow based temporal features for facial expression variations estimation, 3) and finally applying classifier for face. The novelty lies not only in generation of appearance models which is obtained by fitting active shape model (ASM) to the face image using objective functions but also using a feature vector which is the combination of shape, texture and temporal

parameters that is robust against facial expression variations. Experiments have been performed on Cohn-Kanade facial expression database using 62 subjects of the database with image sequences consisting of more than 4000 images. This achieved successful face recognition rate up to 91.17% using binary decision tree (BDT), 98.6% using Bayesian Networks (BN) with 10-fold cross validation in the presence of six different facial expressions. However the benchmarked database consists of only frontal view of faces. This technique is capable of working in real time environment. This system is a constituent of HRI system. It can keep the person identity information even under the presence of facial expressions which could originate under human machine interaction scenarios. However, in real time environment the system can only work by further improving it for light illuminations and using 3D information.

MATERIALS AND METHOD

The proposed system

Based on the literature reviewed above and putting into consideration the weakness and the limitations of these works, a new system is hereby proposed.
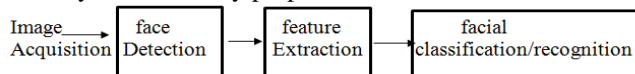
Figure 1: overview of the proposed system.

The system will involves pre-processing image data by normalizing and applying a simple mask, extracting certain (facial) features using principal component analysis (PCA) and Gabor filters and then using support vector machine (SVM) for classification and recognition of expressions. Eigenfaces for each class will be used to determine class-specific masks which will then be applied to the image data and used to train multiple, one against the rest, SVMs.

| Happy | Disgust | Fear | Sad | Surprise |

Figure 2: five different facial expressions of the same person

a. With local binary pattern (LBP), it is possible to describe the texture and shape of a digital image. This is done by dividing an image into several small regions from which the features are extracted (figure 3). These features consist of binary patterns that describe the surroundings of pixels in the regions. The obtained features from the regions are concatenated into a single feature histogram, which forms a representation of the image. Images can then be compared by measuring the similarity (distance) between their histograms
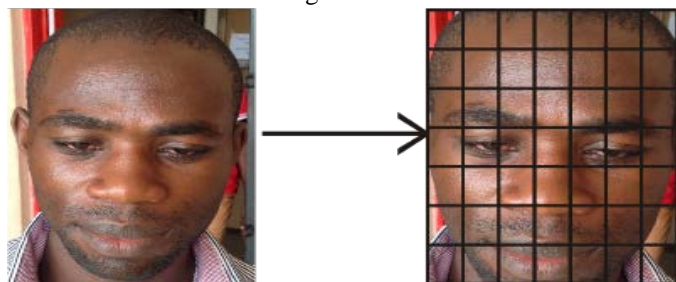

Figure 3: facial image divided into 49 regions (7x7)

Given a pixel position (x_c, y_c), LBP is defined as an ordered set of binary comparisons of pixel intensities between the central pixel and its surrounding pixels. The resulting decimal label value of the 8-bit word can be expressed as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(l_n - l_c)2^n \qquad (1)$$

Where l_c correspond to the grey value of the centre pixel (x_c, y_c), l_n to the grey values of the 8 surrounding pixels, and function $s(k)$ is defined as:

$$s(k) = \begin{cases} 1 & if \ k \geq 0 \\ 0 & if \ k < 0 \end{cases} \qquad (2)$$

If the coordinates of the center pixel are (x_c, y_c) then the coordinates of his P neighbors (x_p, y_p) on the edge of the circle with radius R can be calculated with the sines and cosines:

$X_p = x_c + Rcos(2\pi p/P)$ (3)
$Y_p = y_c + Rsin(2\pi p/P)$ (4)

In the last step to produce the LBP for pixel (x_c, y_c) a binomial weight 2^p is assigned to each sign $s(g_p - g_p)$. These binomial weights are summed:

$$LBP(x_c, y_c) = \sum_{p=0}^{p-1} s(g_p - g_c)2^p \qquad (5)$$

Support Vector Machine as a Classifier

SVM is based on the principle of structural risk minimization (minimizing classification error). A SVM is binary classifier that optimally separates the two classes of data (Burges, 1998). Two major phases are required in the development of SVM as classifier. The first phase involves the determination of the optimal hyperplane which will optimally separate the two classes and the other is transformation of non-linearly separable classification problem into linearly separable problem. Figure 4 below shows linearly separable binary classification problem with no possibility of miss-classification data.

Let m and n be a set of input feature vector and the class label respectively. The pair of input feature vectors and the class label can be represented as tuples $\{m_i, n_i\}$ where i =1,2,…,N and n =±1. In the case of linear separable problem, there exists a separating hyperplane which defines the boundary between class 1 (labelled as n = 1) and class 2 (labelled as n = -1). The separating hyperplane is,

$$w.x + b = 0 \qquad (6)$$

which implies

$$n_i(w.x + b) \geq 1, i = 1, 2, …, N \qquad (7)$$

We have several possible values of {**w**,b} that create separating hyperplane, but in SVM only hyperplane that maximizes the margin between two sets is used. Margin is the distance between the closest data to the hyperlane
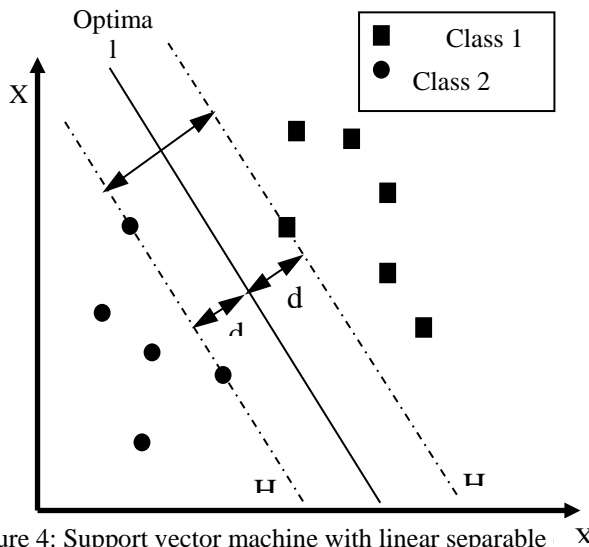

Figure 4: Support vector machine with linear separable

Considering the Figure 4 above the margins are defined as d+ and d-. The margin will be maximized in the case d+ = d−. Furthermore, training data in the margins will lie on the hyper-planes H+ and H-. The distance between hyperplane H+ and H- is,

$$d_+ + d_- = \frac{2}{\|w\|} \qquad (8)$$

There is no training data which fall between H+ and H- as H+ and H- are the hyperplane which is the closest training data to the optimal hyperplane. This means the hyperplane that separates optimally the training data is the hyperplane which minimizes $\|w\|^2$ so that the distance of the equation (8) is maximized. However, the minimization of $\|w\|^2$ is constrained

by equation (7). When the data is non-separable, slack variables, ξi, are introduced into the inequalities for relaxing them slightly so that some points are allow to lie within the margin or even being misclassified completely.

Support Vector Algorithm
a. Support Vector Machine (SVM), a machine learning technique with high recognition rate for facial recognition will be used. The goal of a support vector machine is to find the particular hyperplane for which the *margin of separation ρ* is maximized.

Consider training sample $(\mathbf{x}_i, d_i)^N$, where $\mathbf{x}_i$ is the input pattern for the ith example and d_i is the corresponding desired response(target output). To begin with, we assume that the pattern (class) represented by the subset $d_i = +1$ and the pattern represented by the subset $d_i = -1$ are *linearly separable*. The equation of surface in the form of a hyperplane that does the separation is

$$\mathbf{W^T} \mathbf{x} + b = 0 \qquad (.9)$$

where **x** is an input vector, **w** is an adjustable weight vector, and b is a bias. Then it could be written as

$$W^T x_i + b \geq 0 \; for \; d_i = +1$$
$$W^T x_i + b < 0 \; for \; d_i = -1$$

subject to $y_i(w.x_i + b) - 1 \geq 0$

The separating hyperplane which maximizes the distance between the closest datapoints and itself is then obtained. By using Lagrangian functions, equation above can be transform into the dual form as follow:

$$LP(w, b, \propto) = \frac{\|w\|^2}{2} - \sum_{i=1}^{m} \alpha_i(y_i(w.x_i + b) - 1) \qquad (10)$$

solving the SVM problem is equivalent to finding a solution to the Karuch Kuhn Tucker (KKT) conditions. This becomes a dual problem L_D as given by:

maximize L_D (w,b) =

$$L_D(w, b) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i}^{m}\sum_{j}^{m} \alpha_i \alpha_j y_i y_j x_i.x_j \qquad (11)$$

subject to:

$$\alpha_i \geq 0$$

and

$$\sum_i \alpha_i y_i = 0$$

where α_i is a Lagrange multiplier, there is one Langrange multiplier for each training sample

CONCLUSION

At the end of this research work, a facial expression detection and recognition system will have been developed using Viola & Jones algorithm for detection, Local binary pattern (LBP) for feature extraction and support vector machine (SVM) for classification. The system will involves pre-processing image captured by extracting certain facial features and reducing the dimension of the feature vectors using PCA.

REFERENCES

1. Abrantes G.A and F. Pereira, (1999) "MPEG-4 Facial Animation Technology: Survey Implementation, and Results," IEEE Trans. Circuits and Systems for Video Technology,vol. 9, no. 2, March.

2. Andrew J. Calder, A. Mike Burton , Paul Miller , Andrew W. Young , Shiger Akamatsu (2001) "A Principal Component Analysis of Facial Expression"

3. Bartlett. M, P. Viola, T.Sejnowski, L. Larsen, J. Hager, and P. Ekman (1996). Classifying facial action. In D. Touretzky, M. Mozer, and M. Hasselmo, editors, Advances in Neural Information Processing Systems 8. MIT Press, Cambridge, MA.

4. Bartlett M.S, G. Littlewort, I. Fasel, and R. Movellan,(2003) "Real Time Face Detection and Facial Expression Recognition: Development and Application to Human Computer Interaction," Proc. CVPR Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction,vol. 5.

5. Burges C. J. (1998). A Tutorial on Support vector Machines for Pattern Recognition. The book on Knowledge Discovery and Data Mining 2

6. Cohn J.F, Z. Ambadar and P. Ekman, (2005) "Observer-Based Measurement of Facial Expression with the Facial Action Coding System", in J.A. Coan and J.B. Allen (Eds.), The handbook of emotion elicitation and assessment, Oxford University Press Series in Affective Science, Oxford.

7. Cottrell.G and J. Metcalfe. (1991) Face, gender and emotion recognition using holons. In D. Touretzky, editor, Advances in Neural Information Processing Systems 3, pages 564–571. Morgan and Kaufman, San Mateo.

8. Fasel. B and J. Luttin ,(2003) "Automatic Facial Expression Analysis: a survey," Pattern Recognition, vol. 36, no. 1, pp. 259-275.

9. Samal. A and P.A. Iyengar, (1992) "Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey," Pattern Recognition, vol. 25, no. 1, pp. 65-77.

10. Ekman.P and W. Friesen (1975). Unmasking the Face: A guide to recognizing emotions from facial expressions. Consulting Psychologists Press, Palo Alto, CA.

11. Ekman.P and E.L. Rosenberg (1997), "What the face reveals: basic and applied studies of spontaneous expression using the facial action coding system (FACS)", Illustrated Edition, Oxford University Press.

12. Ekman. P,(1994) "Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique," Psychological Bulletin,vol.115, no. 2, pp. 268-287.

13. Essa I and A. Pentland.(1997) Coding, analysis, interpretation, and recognition of facial expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):757–763.

14. Irene Kotsia and Ioannis Pitas (2007) "Facial Expression Recognition in Image Sequences using Geometric Deformation Features and Support Vector Machines".

15. Lanitis. A, C. Taylor, and T. Cootes(1997). Automatic interpretation and coding of face images using flexible models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):743–756.

16. Lee. T(1996) Image representation using 2d Gabor wavelets. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(10):959–971.

17. Lyons M.J, S. Akamatsu, M. Kamachi and J. Gyoba,(1998) "Coding Facial Expressions with Gabor Wavelets," Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp. 200-205, Japan.

18. Mase.K.(1991) Recognition of facial expression from optical flow. IEICE Transactions E, 74(10):3473–3483.

19. Ojala,T.,Pietikäinen,M.,1999.Unsupervised texture segmentation using feature distributions.Pattern Recognition 32,477–486.

20. Ojala,T.,Pietikäinen,M.,Harwood,D.,1996.Acomparative study oftexture measures with classificationbased onfeatured distribution.Pattern Recognition 29,51–59.

21. Ojala,T.,Pietikäinen,M.,Mäenpää,T.,2002.Multiresolution gray-scale and rotation invariant texture classificationwith local binary patterns.IEEE Transaction onPattern Analysis and Machine Intelligence 24,971–987.

22. Ojansivu,V.,Heikkilä,J.,2008.Blur insensitive texture classificationusing local phase quantization.In:International conference onImage and Signal Processing.

23. Osuna, E., Freund, R., and Girosi, F. 1997a. Training support vector machines: An application to face detection. In Proceedingsof the IEEE Conference on Computer Vision and Pattern Recognition.

24. Osuna, E., Freund, R., and Girosi, F. 1997b. Training support vector machines: an application to face detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition

25. Pantic M and I. Patras,(2006) "Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments Form Face Profile Image Sequences," IEEE Trans. Systems, Man, and Cybernetics Part B,vol. 36, no. 2, pp. 433-449.

26. Rowley.H, S. Baluja, and T. Kanade(1995). Human face detection in visual scenes. Technical Report CMU-CS-95-158R, School of Computer Science, Carnegie Mellon University.

27. Sayette M.A, J.F. Cohn, J.M. Wertz, M.A. Perrottand D.J. Parrott(2001), "A psychometric evaluation of the facial action coding system for assessing spontaneous expression," Journal of Nonverbal Behavior ,vol. 25, no. 3, pp. 167-185.

28. S. Dubuisson, F. Davoine, and M. Masson (2002), "A solution for facial expression representation and recognition," Signal Process.: Image Commun., vol. 17, no. 9, pp. 657–673.

29. X.-W. Chen and T. Huang,(2003) "Facial expression recognition: A clustering-based approach," Pattern Recognit. Lett., vol. 24, no. 9–10, pp. 1295–1302.

30. Wiskott.L, J.-M. Fellous, N. Kr¨uger, and C. von der Malsburg(1997). Face recognition by elastic bunch graph matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7):775–779.

31. Yacoob.Y and L. Davis(1994). Recognizing facial expressions by spatio-temporal analysis. In Proceedings of the International Conference on Pattern Recognition, volume 1, pages 747–749, Jerusalem, Israel.