IJDIAS International Journal of Discoveries and Innovations in Applied Sciences

| e-ISSN: 2792-3983 | www.openaccessjournals.eu | Volume: 3 Issue: 6

A LBP and SVM Based Face Expression Classification System

Sandeep Kumar
P.G. Student, Department of CSE, Sat Kabir Institute of Technology and Management, Haryana, India

Rishabh, Kirti Bhatia
Assistant Professor, Department of CSE, Sat Kabir Institute of Technology and Management, Haryana, India

Abstract:

This work presents support vector machine (SVM)-based emotion detection and multi-class facial expression categorization. By traversing each bin in both a clockwise and an anticlockwise orientation, the Local Binary Pattern (LBP) Histogram can be used to generate facial feature vectors in double format. The LBP pictures in double format are used to determine the Histogram feature descriptors, which are then combined to produce the features of the full-face image. The suggested algorithm is evaluated using the conventional Japanese Female Facial Expression Database (JFFED) and the Taiwanese facial expression database, and the outcomes are confirmed using a locally created student face database in India. The suggested algorithm functions noticeably better than traditional LBP-based techniques.

Keywords: Facial Expression Perception, Support Vector Machine, Local Binary Pattern.

I. Introduction

In light of its significant possibilities in multimedia applications, such as streaming media, service to customers, driver surveillance, and other areas, facial expression recognition (FER) has gained a lot of popularity as an important area of study in human-computer interaction (HCI) [1]. If computers could recognise users as people who can gain from resolving FER challenges, HCI would become more approachable and intuitive. The goal of FER is to analyse and categorise a given facial image into one of the 6 frequently expressed emotion types: anger, contempt, fear, happiness, sadness, and surprise. Over the past few years, a number of FER algorithms, including recognising expressions from front and non-frontal facial photos, have been suggested in the literature [2]. According to research by Ekman and Friesen, facial expressions are inherent and global. Facial variations in reaction to an individual's inner emotional states, goals, or messages are referred to as facial expressions. A computer vision system can communicate with people by naturally reading facial expressions. The most obvious and potent indicators of an individual's emotional condition are their facial expressions.

Yet, only a small portion of the algorithms among the numerous Methods suggested actually address this difficult problem. A generic recognition approach that has been used in most prior investigations may be broken down into two main components for both frontal and non-frontal FER challenges: feature extraction and classifier development. In the earlier publications, a variety of image features were used for capturing facial features, including scale-invariant feature transform (SIFT), histograms of oriented gradients (HOG), local binary pattern (LBP), and local phase quantization. SIFT has shown outstanding results among the many face features because of its robustness to image scaling, motion, obstruction, and lighting variation [3]. The challenge of classifying emotions involves two classes. The person can be in either of two emotional states [4].

IJDIAS International Journal of Discoveries and Innovations in Applied Sciences
| e-ISSN: 2792-3983 | www.openaccessjournals.eu | Volume: 3 Issue: 6

First, a happy or surprised expression, which is a positive emotion. Negative emotions include expressions of disgust, unhappiness, fear, and anger. A multilevel categorization challenge is the detection and recognition of distinct facial expressions. For the recognition of many facial expression classes, multiple studies have been undertaken.

For the classification of face expressions, many databases were used. Both the Japanese Female Facial Expression (JFFE) [5] and the Taiwanese Facial Expression Image Database (TFEID) [6] are common databases that researchers frequently utilise to test and validate their findings. Three key pieces make up the Basic Facial Expression Categorization system. Face detection from the input image comes first, followed by the extraction of facial features from the trimmed face pattern and the categorization of facial expressions. The FER receives noise-filtered pixel image data as input. The clipped face pattern and the face in the image data are both detected by the face detection module. First, the detected face is normalised. The feature extraction module extracts facial features that define the pattern of the face using discriminating criteria that are most important to the expression of the face. The final stage is to identify the person's emotional state by categorising their expression into pre-established facial expression classifications. Six categories are used to categorise facial expressions: ecstatic, shocked, disgusted, unhappy, fearful, and indignant

II. RELATED WORK

The picture facial features vector is extracted using a variety of approaches in the current system, which exhibits minimal inter-person variation. A multilayer perceptron receives this feature vector as input to perform tasks like face recognition or identity verification. The suggested technique combines Gabor and Eigen faces to produce the feature vector. The outcomes of the evaluation demonstrate the suggested system's robustness against variations in lighting, clothing, facial expressions, scale, and position within the collected image, as well as desire, noise pollution, and filtering. The suggested scheme also offers some latitude for variations in the subject's age. The suggested scheme's evaluation findings with identification and verification setups are presented, and they are contrasted with those of other feature extraction techniques to highlight the most desirable aspects of an algorithm.

For the purpose of identifying six fundamental facial expressions, two image representation techniques dubbed non-negative matrix factorization (NMF) and local non-negative matrix factorization (LNMF) have been applied to two facial databases. Using principal component analysis (PCA), fared similarly for the comparison of facial expression recognition. For the first database, we discovered that LNMF performs better than both PCA and NMF, with NMF producing the worst recognition performance. For the second database, the outcomes are essentially identical, with a little boost to NMF's efficiency. It is suggested to use the Local Fisher Discriminate Analysis (LFDA) to recognise face expressions. Fig. 1 shows the basic expression identification system.

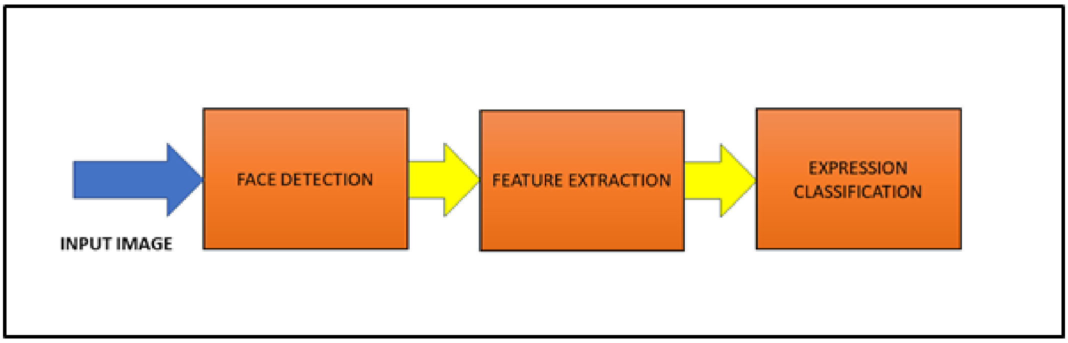

Fig.1: Basic expression identification system

IJDIAS International Journal of Discoveries and Innovations in Applied Sciences

| e-ISSN: 2792-3983 | www.openaccessjournals.eu | Volume: 3 Issue: 6

III. Face Detection

Face detectors are used to retrieve the face pattern. The Viola-Jones face detector and the Kanade-Lucas-Tomasi tracker are popular face detectors. The Adaboost approach is used by the Viola-Jones face detector. AdaBoost algorithms offer a straightforward and efficient method for learning a nonlinear categorization function stage by stage [7]. AdaBoost incrementally improves just a few of poor classifiers to create a stronger classifier with higher accuracy. At each iteration, the distribution is modified to raise the weights of the incorrectly categorised samples, and a weak classifier that minimises the weighted error rate is chosen.

IV. Facial Feature Extraction

Discriminating elements of the face are extracted using a facial feature extraction process. The primary goal of feature extraction is the discriminatory parameterization of a vast volume of pixel data. The input space's dimensionality is significantly decreased during feature extraction. The attributes that were retrieved are then used for categorization. Global feature descriptors and local feature descriptors are the two different forms of feature descriptors. While local descriptors are based on the physical characteristics of the face pattern [8], global descriptors are based on the geometry of the pattern [9]. Global feature descriptors use the shape and placement of facial features including the mouth, chin, and brows to characterise the geometrical characteristics of the facial pattern [9]. To create a feature vector that reflects the face geometry, the facial parts, or facial feature points, are retrieved. For the full facial pattern, geometrical characteristics emerge. The feature vectors that were thus collected were then utilised to categorise facial emotion.

Local characteristics descriptors highlight changes in the face's look by textually describing the skin's wrinkling and deformation [8]. Applying texture extraction techniques to different areas of the face allows for the creation of appearance-based features. The collection of characteristics is aggregated to describe facial expression, which is then further classified. Micro patterns in skin texture can be captured by appearance-based characteristics. The most common method for representing textual information about facial pattern is called LBP base [1]. The most widely used and effective method in computer vision applications, including face recognition and recognition of facial expressions, has been facial image analysis utilising the LBP descriptor. The calculation of recognition efficiency is closely related to the features extraction method chosen.

V. Expression Classification and Emotion Detection

Machine learning is used to classify facial features. By employing a known collection of data, computers can be programmed to do classification tasks more efficiently. Regarding the input data, there are two main categories of learning: supervised learning and unsupervised learning. The objective of supervised learning is to develop a mapping from an input to an output whose correct values are supplied by a supervisor. There is no formal supervisor and merely input data in unsupervised learning. There are numerous machine vision methods for learning and categorization, with K-nearest Neighbour (K-NN) , Support Vector Machine (SVM) , and Artificial Neural Networks (ANN) being a few examples. Vapnik [9] introduced the supervised binary classification approach known as the SVM. The fundamental concept behind SVM is to utilise a linear model to implement boundaries by performing a nonlinear input vector to high-dimensional feature space mapping. SVM is divided into two sections: training and testing. Six common expressions, including Happy, Surprise, Disgust, Unhappy, Fear, and Angry, are taken into account for emotion detection.

VI. PROPOSED APPROACH

This method proposes a face descriptor, via the technique of (LBP), for the recognition of facial expressions. Hence, LBP is used to derive the description of emotion-related characteristics through the use of the directional information and ternary structure in order to identify the fine edge in the face area while the face having the smooth zones. The grid is then categorised while sampling expression-related data at various scales to create the face descriptor. The goal of dimension reduction through the extraction of distinctive characteristics is to increase the overall scatter of the data while reducing variation within classes. It is clear that the feature values for the six classes have a strong tendency to combine, which may lead to a high percentage of misclassification. The real number of elements may be greater than three; nevertheless, the first three features were chosen to construct for the purpose of visualisation. As a result, this work makes use of a strong characteristic. This is simple to understand, has strong predictability, and costs less to compute than other approaches now in use.

Regarding the classification component, numerous techniques have been used to classify expressions accurately. Some authors used (ANNs) to identify various facial expressions, and they were successful in achieving a high recognition rate. Yet, ANN is a "black box" and only partially capable of categorising potential basic linkages. Additionally, ANNs could take a while to train and might fall victim to poor local minima. The (SVMs) were also used by the authors to create their FER system. But with SVMs, there is no direct estimation of the observation probability; instead, the observation probability is calculated indirectly. Each frame is anticipated to be statistically independent from the others since SVMs simply ignore temporal relationships between video frames. In order to classify crops and weeds for real-time selective herbicide systems, we evaluated and confirmed the accuracy of wavelet transform combined with support vector machines (SVMs). The proposed approach differs from prior systems in that it includes a pre-processing step that helps to reduce lighting effects and assure high accuracy in real-world circumstances. In order to separate the classes of weeds with broad leaves from those with narrow leaves, we examined a huge number of wavelets and decomposed them up to four layers. This was used to condense the feature space by just extracting the most important features. The features offered by SVMs for classification, lastly.

The term "pre-processing" refers to the "preparation" of the sample or picture before it is fed into an algorithm to perform a specific task, such as feature extraction, monitoring targets, or recognition. A data mining approach called data pre-processing entails putting raw data into a comprehensible format. Real-world data is frequently inaccurate and lacking in specific behaviours or trends. It is also often unreliable and imprecise. Pre-processing data is a tried-and-true way to fix these problems.

The following procedure can be used to build the LBP feature vector in its simplest version.

Cellularize the window being examined. Compare each pixel in a cell to its eight neighbours. Move either clockwise or anticlockwise through these pixels on a circular course. If the value of the middle pixel exceeds that of any neighbouring pixels, mark 0. If not, mark 1. It produces a binary number with a 1 byte output that is frequently translated to a decimal value. For every layout pixel that is lower or larger than the midway, calculate the histogram (256-dimensional feature vector) to represent the regularity of each occurrence number in the cells. Make the histogram normal. Integrate and normalise the histogram of each cell. This displays the feature vector for the full window. The feature vector that has been gathered in this way can now be created using an SVM or similar ML technique to categorise the images. These classifiers can be used for recognising faces or textural analysis. The uniform pattern is a helpful addition to the main LBP operator that may be

IJDIAS International Journal of Discoveries and Innovations in Applied Sciences
| e-ISSN: 2792-3983 | www.openaccessjournals.eu | Volume: 3 Issue: 6

used to reduce the dimension of the feature vector and apply straightforward substitution consistent descriptors. In texture pictures, some binary patterns might be seen more frequently than others. To generate the LBP descriptor, transform the image to grayscale, choose a locality of dimension r close to the centre pixel, produce an LBP value for it, and then save the result in a 2D output array with the identical dimensions as the source image.

Up until this point, the algorithm was trained using the training collection, which resulted in one histogram for each image. In order to create a histogram that accurately depicts the image, perform the next stages for the new image given a source image. In order to create the image with the closest histogram, two histograms are compared in order to locate the image that is identical to the input image. Applying the Euclidean distance, chi-square, absolute value, etc., two histograms can be compared.

The process determines which image produces the closest histogram. The technique also yields the estimated distance, commonly referred to as the confidence level. The threshold and the confidence value serve to define the successfully detected image. The algorithm has successfully detected the image if the confidence is less than the stated threshold. We altered LBP to produce histograms of the input image. We looked at an oval-shaped neighbourhood pixel trajectory as opposed to a circular one centred on the central pixel.

VII. EXPERIMENTAL RESULTS

Analysis relies on the JAFFE Collection and the (HOG) methodology. In the pre-processing stage, the face region is isolated and the rest of the image is ignored using the face detection approach. This makes ignoring the useless information simpler. Thus, the feature information extraction stages' time to implement is reduced. likewise the dimension alignment method helps with any necessary image size adjustments. The histogram equalisation method, on the other hand, uses a distribution of the image's density value to specify how bright the image should be.

The JAFFE library has 213 images for 7 expressions that were collected from 10 Female. In our study, all other individuals are always included in the training set, but only one individual is present at a time in the testing set, therefore this operation has been performed (N-1) times, where N is the total number of participants in each collection. The research projects are also divided into groups based on the suggested methodology. Furthermore, six databases are used to implement each strategy; as a result, each technique's results are independently reported according to the datasets that were utilised. Additionally, "Cell Size" describes how many shape data points will be represented in a specific retrieved feature procedure's measurements. For instance, a cell size of [8X 8] denotes a high level of shape information encoding, but a cell size of [64 X64] denotes a lower level of information encoding.

This method extracts facial attributes from facial photos using the LBP method. An SVM classifier is used to categorise these properties. In order to show how cell size affects classification models, tests are also done on six different datasets utilising varying cell sizes in each collection. The precision of the LBP+SVM method was 77.46% with cell size=32.

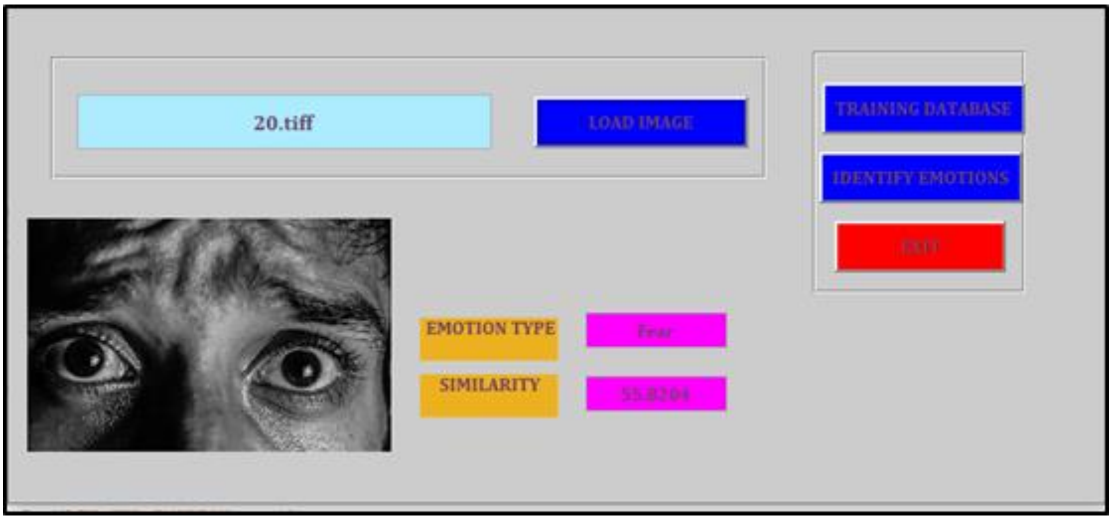

Fig. 2: Identification of Fear Emotion

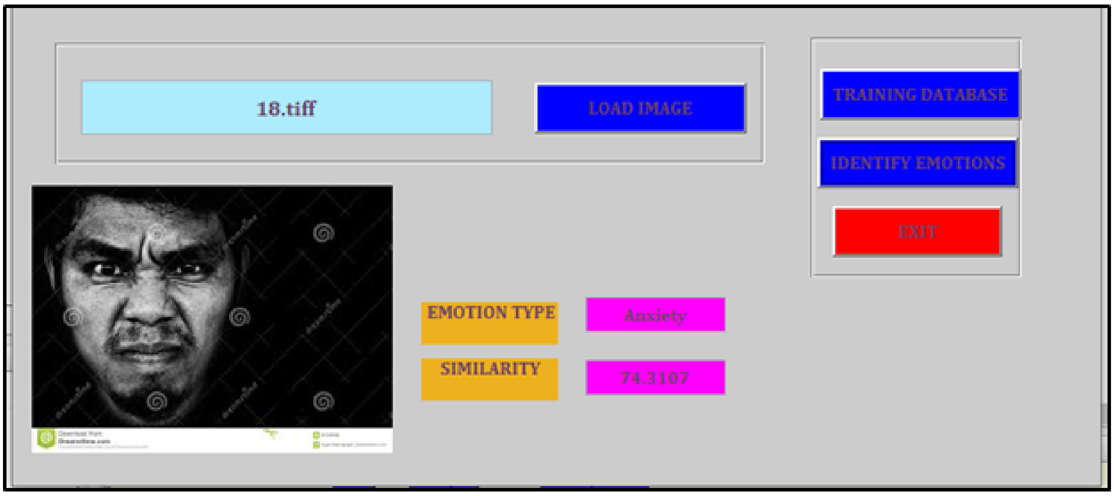

Fig. 3: Identification of Anxiety Emotion

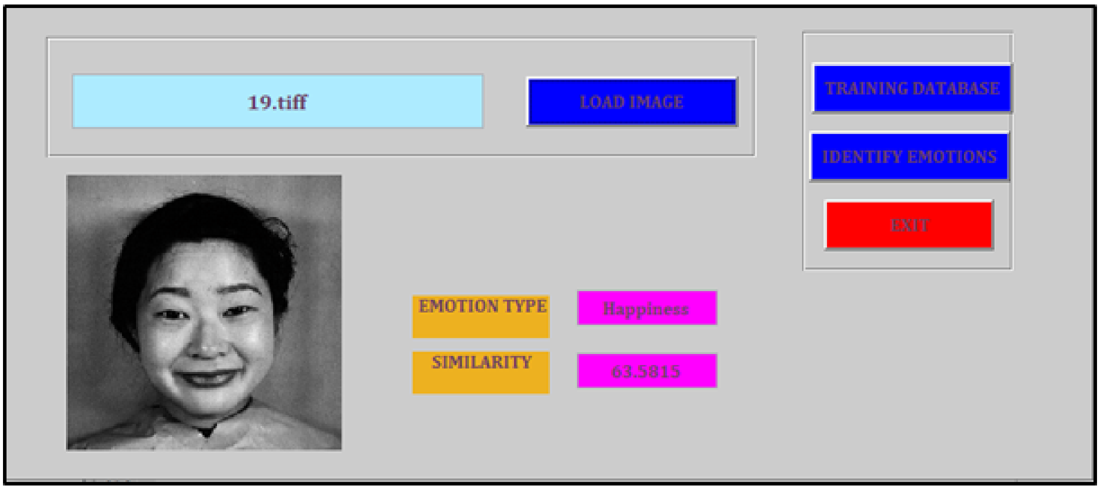

Fig. 4: Identification of Happy Emotion

IJDIAS International Journal of Discoveries and Innovations in Applied Sciences

| e-ISSN: 2792-3983 | www.openaccessjournals.eu | Volume: 3 Issue: 6

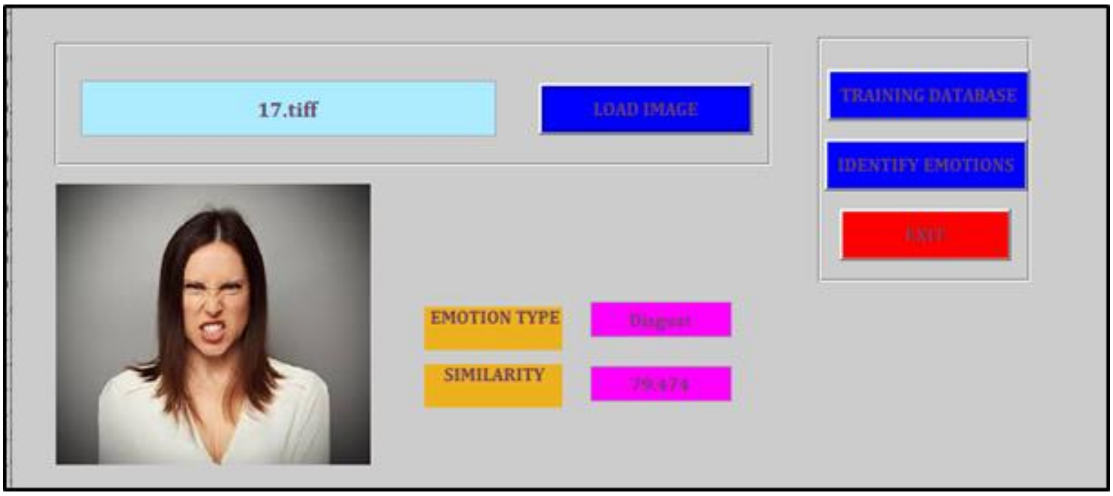

Fig. 5: Identification of Disgust Emotion

VIII. CONCLUSION

We have suggested a novel method that totally and accurately encodes the facial textual pattern. The creation of a straightforward technique to encode textual information of facial pattern is the primary accomplishment of this work. Comparing the experimental results to other traditional LBP-based algorithms, recognition rates exhibit enhanced accuracy. With the suggested strategy, the intrinsic benefit of LBP is preserved with an additional benefit. Both common databases and an obscure Indian picture database were used for the classification of the photographs. When using SVM Classification, we took the correlation of the pixel data into account when selecting the class. Happy and Surprise, Angry and Disgust, and Unhappy and Fear are all substantially connected in the multiclass method.

References

1. Ankit Jain, Kirti Bhatia, Rohini Sharma, Shalini Bhadola, An Overview on Facial Expression Perception Mechanisms, SSRG International Journal of Computer Science and Engineering (SSRG - IJCSE) - Volume 6 Issue 4 - April 2019, pp. 19-24.
2. Ankit Jain, Kirti Bhatia, Rohini Sharma, Shalini Bhadola, An emotion recognition framework through local binary patterns, Journal of Emerging Technologies and Innovative Research, Vol -6, Issue-5, May 2019.
3. Sandeep Kumar, Rishabh, Kirti Bhatia, Rohini Sharma, A Study on Face Expression Observation Systems, American Journal of Science and Learning for Development, Volume 2, No 6, Jun -2023, pp.153-158.
4. Porawat Vistutsak, "Emotion Classification through Lower Facial Expressions using Adaptive Support Machine", Journal of Man, Machine and Technology, Vol 2 No 1, June 2013.
5. Michael J. Lyons, Shigeru Akemastu, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998).
6. Li-Fen Chen and Yu-Shiuan Yen. (2007). Taiwanese Facial Expression Image Database Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan.

7. Bin Jiang and Kebina Jia, " Semi Supervised Facial Expression Recognition Algorithm on The Condition of Multi-Pose", Journal of Information Hiding and Multimedia Signal Processing", VOL. 4, No. 3, July 2013.

8. Xinbo Gao, Ya Su Xuelong Li, and Dacheng Tao, "A Review of Active Appearance Models", IEEE Transaction on System, Man and Cybernetics, Vol 40, No 2 March 2010.

9. Irene Kotsia and Ioannis Pitas, " Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines", IEEE Transaction on Image Processing, Vol 16, No 1, January 2007.