

Texture based Emotion Recognition from Facial Expressions using Support Vector Machine

A. Punitha

Dept. of Computer Science and Engg.,
Annamalai University

M. Kalaiselvi Geetha

Dept. of Computer Science and Engg.,
Annamalai University

ABSTRACT

The mission of automatically recognizing different facial expressions in human-computer environment is significant and challenging. This paper presents a method to identify the facial expressions by processing images taken from Facial Expression Database. The approach for emotion recognition is based on the texture features extracted from the gray-level co-occurrence matrix (GLCM). The results show that the features are highly efficient to discriminate the expressions and require less computation time. The extracted GLCM features are trained with Support Vector Machine using different kernels to recognize the basic emotions Happy, Disgust, Surprise and Neutral.

Keywords:

Gray Level Co-Occurrence Matrix (GLCM), Texture Feature, Support Vector Machine

1. INTRODUCTION

Emotions are normally displayed by visual, vocal, and other physiological means. One of the important way humans display emotions is through facial expressions. Facial expression is one of the most powerful ways that people bring together conversation and communicate emotions and other mental, social, and physiological cues. The real start of automatic facial expression recognition systems was initiated in the early 1990s. Since the early 1990s, Facial Expression Recognition System have enthralled many researchers from numerous disciplines, such as the fields of robotics, psychology, and computer vision. Moreover, it has gained prospective applications in areas such as human-computer interaction systems, image retrieval, face modeling, patient monitoring systems for pain and depression detection, face animation, drowsy driver detection in vehicle surveillance etc.,

1.1 Related Work

Over the last two decades, intensive research has been conducted to develop more robust person-independent facial expression recognition systems that work in real-time and in difficult scenarios with lighting changes, facial makeup, different ages and races, low resolution images, and facial occlusions. Generally existing approaches for facial

expression recognition systems can be divided into three categories, based on how features are extracted from an image for classification. The categories are geometric-based [1, 2] appearance-based [3, 4, 5] and hybrid-based. The geometric feature measures the variations in shape, location, distance of facial components like mouth, eyes, eyebrows, nose, etc., The appearance features present the appearance (skin texture) variations of the face, such as wrinkles, furrows, etc. It can be extracted on either the whole face image or specific regions in a facial image. Hybrid features uses both geometric and appearance based approaches to extract features. Detailed review of existing methods on facial expression is seen in [6, 7, 8].

Many methods have been proposed for face tracking which include active contours [9], robust appearance filter [10], adaptive active appearance model [11] and active appearance model [12]. The emotion informations captured from facial images by different multiresolution algorithms such as Wavelet, curvelet and contourlet transform is presented in [13]. [14] investigates the efficiency of performance of three texture based feature extraction methods Grey Level Co-occurrence Matrix (GLCM), Local Binary Pattern (LBP) and Elliptical Local Binary Template (ELBT) for face recognition. Local Binary Pattern (LBP) feature for texture recognition and application in recognizing facial expression is presented in [15]. A wide range of classification algorithms (e.g. Hidden Markov Models (HMMs) [16], Bayesian network classifier [17], Support Vector Machine (SVM) [18][19], Principal Component analysis (PCA) [19], Linear Discriminant Analysis (LDA)) have been applied to the facial expression recognition problem.

1.2 Outline of the work

This paper deals with emotion recognition from face expressions. The method has been evaluated with Facial Expression database [24] where the subjects expresses emotions such as Normal, Smile, Disgust and Surprise. All the images in the database are of size 63 X 63 pixels. Gray Level Co-Occurrence Matrix has been computed for all the images and second-order statistical texture features that are extracted from the GLCM matrix is given as input to the Support Vector Machine for classification. The rest of the paper is organised as follows :

Section 2 describes the Feature Extracion. Expression

recognition using SVM is presented in Section 3. Section 4 shows the Experimental results of our approach and Section 5 concludes this work.

2. FACIAL EXPRESSION RECOGNITION SYSTEM

Most systems of recognition for facial expressions can be divided into the following phases:

- 1) Face localization: To detect the face present in an image.
- 2) Region of Interest Detection: To divide the face into regions of interest for feature extraction in each region.
- 3) Feature Extraction: To detect and extract the features present in each region of interest.
- 4) Expression Recognition: To identify the expressions using Classifiers.

2.1 Feature Extraction

Feature is a descriptive portion extracted from an image or a video stream. Visual data exhibit numerous types of features that could be used to recognize or represent the information it reveals. These features exhibit both static and dynamic properties. Classification or recognizing an appropriate video relies on competent use of these features that provide discriminative information useful for high-level classification. The following subsection present the description of the feature used in this study.

2.1.1 Gray Level Co-Occurrence Matrix. Texture analysis aims in finding a distinctive way of representing the essential characteristics of textures and represent them in some simpler but unique form, so that they can be used for robust, accurate classification. The texture features of face contain significant expression information and can represent some subtle facial expression changes. Gray Level Co-Occurrence Matrix (GLCM) has proved to be a popular second order statistical method of extracting textural feature from images. This feature represents the distribution of intensities and information about the relative positions of neighboring pixels of an image. Gray level co-occurrence matrix (GLCM) was firstly introduced by Haralick [20]. [21] use the improved approach based on generalized image co-occurrence matrix for image retrieval. The theory and techniques behind the Gray Level Cooccurrence Matrix (GLCM) method is presented in [22]. GLCM is obtained by calculating how often a pixel with gray-level (grayscale intensity) value i occurs adjacent to a pixel with the value j . Each element (i,j) in GLCM specifies the number of times that the pixel with value i occurred adjacent to a pixel with value j . The adjacency can be defined to take place in each of four directions 0° , 45° , 90° and 135° degrees in a two-dimensional square pixel image (directions horizontal, vertical, left and right diagonals) Fig 1.

Given an Image I , of size $N \times N$, the co-occurrence matrix P Eqn. 1 can be defined as

$$p(i, j) = \sum_{x=1}^N \sum_{y=1}^N \begin{cases} 1, & \text{if } I(x, y) = i \text{ and } I(x + \Delta_x, y + \Delta_y) = j; \\ 0, & \text{Otherwise} \end{cases} \quad (1)$$

where the offset (Δ_x, Δ_y) , specifies the distance between the pixel of interest and its neighbour.

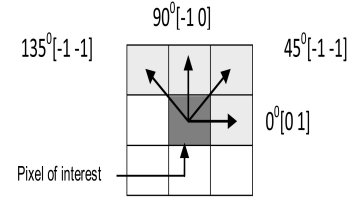


Fig. 1. The four directions of adjacency that are defined for calculating the texture features

A number of texture features may be extracted from the GLCM. This work uses texture descriptors i.e., Contrast, Correlation, Cluster shade, Cluster prominence, Homogeneity and Energy [Eqn. 2 - Eqn. 7]. These measures provide high discrimination accuracy required for emotion recognition. The following are the texture features used in this work :

$$Contrast = \sum_{i,j} |i - j|^2 p(i, j) \quad (2)$$

$$Correlation = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \quad (3)$$

$$Energy = \sum_{i,j} p(i, j) \quad (4)$$

$$Homogeneity = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (5)$$

$$Cluster shade = \sum_{i,j} (i - \mu_x + j - \mu_y)^3 p(i, j) \quad (6)$$

$$Cluster Prominence = \sum_{i,j} (i - \mu_x + j - \mu_y)^4 p(i, j) \quad (7)$$

3. EXPRESSION RECOGNITION USING SVM

Feature classification is performed in the last stage of an automatic facial expression analysis system. Support Vector Machines are a learning method for pattern recognition problem introduced by V. Vapnik et al [23]. Support Vector Machines classify data through determination of a set of support vectors, through minimization of the Structural Risk. The support vectors are members of the set of training inputs that outline a hyperplane in feature space. This

l-dimensional hyperplane, where l is the number of features of the input vectors, defines the boundary between the different classes Fig 2 The classification task is simply to determinate which side of the hyperplane the testing vectors reside in. Minimizing the structural risk reduces the average error of the inputs and their target vectors. The support vector algorithm approximately performs Structural Risk Minimization. Given a set of training examples $(x_1, y_1), (x_2, y_2) \dots (x_l, y_l)$, if there is a hyperplane that separates the positive and negative examples, then the points x which lie on the hyperplane satisfy $w \cdot x_i + b = 0$, where w is normal to the hyperplane and b is the distance from the origin. The margin of a separating hyperplane is defined as the shortest distance to the closest positive or negative example. The support vector algorithm looks for the separating hyperplane with the largest margin. Researchers have expanded from two-class classification to multiclass classification. Different types of SVM are used depending upon the distribution of input patterns. A linear maximal margin classifier is used for linearly separable data, a linear soft margin classifier is used for overlapping classes, and a nonlinear classifier is used for classes that are overlapped as well as separated by nonlinear hyperplanes.

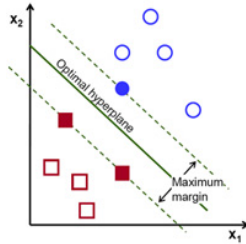


Fig. 2. Optimal Separating Hyperplane

An example for SVM kernel function maps 2-dimensional input space (x_1, x_2) to higher 3-dimensional feature space $(x_1^2, x_2^2, 2x_1x_2)$ as shown in Fig3. SVM was originally developed for two class classification problems. The N class classification problem can be solved using N SVMs. Each SVM separates a single class from all the remaining classes (one-vs-rest approach).

3.1 Inner Product Kernels

SVM generally applies to linear boundaries. In the case where a linear boundary is inappropriate SVM can map the input vector into a high dimensional feature space. By choosing a non-linear mapping, the SVM constructs an optimal separating hyperplane in this higher dimensional space. The function K is defined as the kernel function for generating the inner products to construct machines with different types of non-linear decision surfaces in the input space. The Table1 shows the SVM kernel functions used in this work.

The dimension of the feature space vector (x) for the polynomial kernel of degree p and for the input pattern

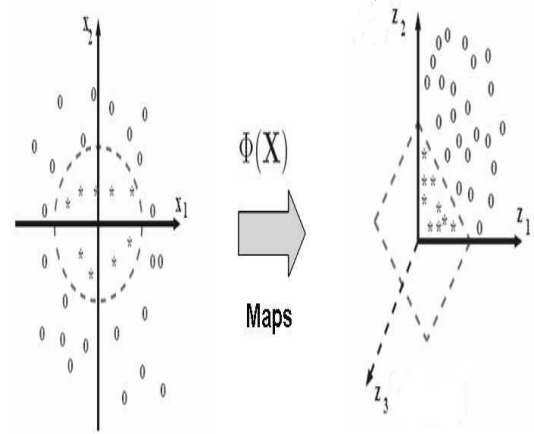


Fig. 3. Mapping of 2 dimensional data to 3 dimensional data

Table 1. Types of SVM inner product kernels

Types of Kernel	Inner Product Kernel	Details
Polynomial	$(x^T x_i + 1)^p$	where x is input patterns, x_i is support vectors, p is degree of polynomial
Gaussian	$\exp[-\frac{\ x - x_i\ ^2}{2\sigma^2}]$	σ^2 is variance, $1 \leq i \leq N_s$, N_s is number of support vectors

dimension of d is given by

$$\frac{(p+d)!}{p!d!}$$

and dimension for gaussian kernel is shown to be infinite. There are several approaches to adopting SVMs to classification problems with three or more classes:

- Multiclass ranking SVMs, in which one SVM decision function attempts to classify all classes.
- One-against-all classification, in which there is one binary SVM for each class to separate members of that class from members of other classes.
- Pairwise classification, in which there is one binary SVM for each pair of classes to separate members of one class from members of the other. In this work, LIBSVM [a library of Support Vector Machine] [25] a simple, easy-to-use, and efficient software for SVM classification and regression is used.

4. EXPERIMENTAL RESULTS

In this section, we present the experimental results obtained on facial expression dataset. The experiments are conducted in C++ with OpenCV 2.2 in Ubuntu 12.04 operating system on a computer with Intel Xeon X3430 Processor 2.40 GHz with 4 GB RAM. The extracted features are fed to LIBSVM for training. Polynomial and RBF kernels are used for experimental purpose.

4.1 Dataset

In the Facial Expression database, there are 13 subjects and each subject has 75 images showing different expressions. These face images are collected in the same lighting condition using CCD camera. Around 700 images are used for training and 250 images are used for testing. Fig 4 shows the sample images taken from the facial expression database.



Fig. 4. Sample images taken from the database[24]

4.2 Evaluation metrics

Precision (P) and Recall (R) are the commonly used evaluation metrics and these measures are used to evaluate the performance of the proposed system. The measures are defined as follows:

$$Precision = \frac{\text{no. of true positives}}{\text{no. of true positives} + \text{false positives}}$$

$$Recall = \frac{\text{no. of true positives}}{\text{no. of true positives} + \text{false negatives}}$$

The work used F-score as the combined measure of Precision (P) and recall (R) for calculating accuracy which is defined as follows:

$$F_{\alpha} = \frac{2PR}{P + R}$$

where α is weighting factor and $\alpha = 0.5$ is used. Recognition accuracy has been measured for the four expressions Table.2 shows the recognition accuracy using RBF kernel and among the expressions considered, Happy is detected with 94% accuracy in our work and the confusion matrix for the proposed feature with SVM Guassian kernel is presented in Table.3 The number of misclassifications under each classes are shown in the confusion matrix.

Table 2. Recognition accuracy for expressions

Expression	Precision %	Recall %	F- Score %
Happy	96	92	94
Suprise	93	91	92
Disgust	90	84	86
Neutral	90	84	86

Table 3. Confusion matrix for the tested samples %

Expression	Happy	Surprise	Disgust	Neutral
Happy	94	0	0	6
Suprise	3	92	1	4
Disgust	2	2	86	10
Neutral	10	1	3	86

4.3 SVM Kernel Functions

4.3.1 Polynomial Kernel. The proposed work is evaluated with various degree of the polynomial kernel. It is observed that the classification performance improves as the degree of the polynomial kernel is increased and the kernel with degree 3 performs better than the other kernel degree. Further, it has been noticed that the performance almost saturates and does not change much as the degree is further increased. This kernel gives a recognition accuracy of 84%.

4.3.2 RBF kernel . The system is evaluated with RBF kernel and it has been noticed that performance of SVM with RBF kernel gives better recognition accuracy when compared to Polynomial kernel. This seems to indicate that polynomial kernel best fit the data for recognition. The approach using this kernel gives an accuracy of 90%. Table 4 shows the recognition rate with various SVM kernel functions.

Table 4. Accuracy rates for facial expression recognition

SVM Kernel	Polynomial (degree 3)	RBF
Recognition Rate	84%	90%

Further the performance of the system in terms of accuracy with respect to the number of images is depicted in Fig 5. It shows that, if the number of images is less than 100, then SVM is not able to capture the sequence information needed for expression recognition and the recognition accuracy is also very less. If the number of frame is 100 or greater, then there is a significant improvement in the recognition performance of SVM.

5. CONCLUSION

This work concentrated on classifying facial expressions into four emotions: Happy, Disgust, Neutral and Surprise using texture features extracted from Gray Level Co-occurrence Matrix. The results are proving that GLCM features based SVM is giving higher classification rate of 90 %. The system can be extended to extract higher-order statistical texture feature from images and taking into account some of the strange facial expressions.

6. REFERENCES

- [1] Gokturk, S.B., Bouguet, J.Y., Tomasi, C. and Girod, B., "Model-based face tracking for view independent facial expression recognition," *Proc. IEEE International Conference in Face and Gesture Recognition*, 2002, pp. 272-278.
- [2] Chang, Y., Hu, C., Feris, R. and Turk, M., "Manifold based analysis of facial expression", *Journal of Image and Vision Computing*, 2006, Vol. 24, No. 6, pp. 605-614.

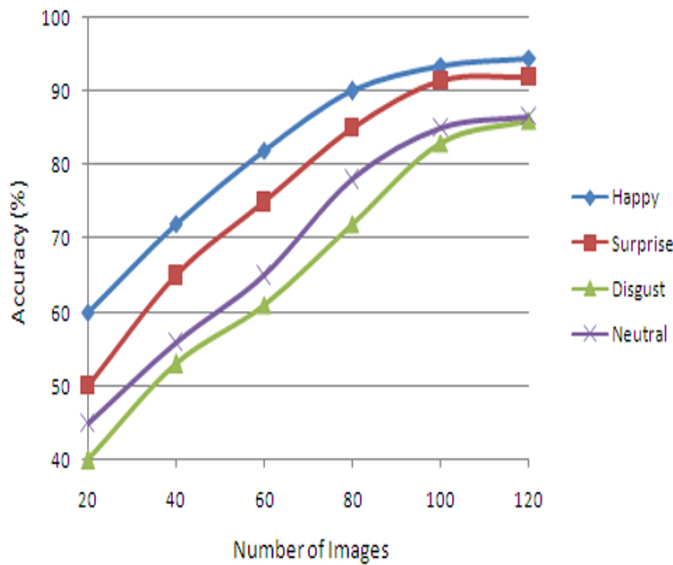


Fig. 5. Expression Recognition Accuracy using LIBSVM

- [3] Guo, G. and Dyer, C.R., "Learning From Examples in the Small Sample Case - Face Expression Recognition. ", *IEEE Trans. Systems, Man, and Cybernetics - Part B* , Vol. 35, No. 3, pp. 477-488.
- [4] Bartlett, M.S., Littlewort, G., Frank, M.G., Lainscsek, C., Fasel, I. and Movellan, J., "Fully automatic facial action recognition in spontaneous behavior ", *Proc. IEEE Conf. Automatic Face and Gesture Recognition* , 2006, pp. 223-230.
- [5] Anderson, K. and McOwan, P.W., "A Real-Time Automated System for Recognition of Human Facial Expressions ", *IEEE Trans. Systems, Man, and Cybernetics - Part B*, 2006, Vol.36, No. 1, pp. 96-105.
- [6] Lienhart, R., Fasel, B., and Luetttin, J., "Automatic facial expression analysis: a survey", *Pattern Recognition* , 2003, pp. 259-275.
- [7] Pantic, M., and Rothkrantz, L., "Automatic analysis of facial expressions: the state of art", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, Vol. 22 (12) , pp. 1424-1445.
- [8] S.-S. Liu, Y.-T. Tian, and D. Li, "New Research Advances of Facial Expression Recognition", *International Conference on Machine on Machine Learning and Cybernetics*, Baoding, 2009, Vol.2, pp. 1150-1151.
- [9] Freedman, D., "Active Contours for Tracking Distributions", *IEEE Transactions on Image Processing*, 2003, Vol. 13, No. 4, pp. 518-526.
- [10] Nguyen, H.T., and Smeulders, A.W.M., "Fast Occluded Object Tracking by a Robust Appearance Filter", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, Vol. 26, No. 8, pp 1099-1104.
- [11] Batur, A.U., and Hayes, M.H., " Adaptive Active Appearance Models", *IEEE Transactions on Image Processing*, 2005, Vol. 14, No. 11, pp. 1707-1721.
- [12] Corcoran, P., Ionita, M.C., and Bacivarov, I., " Next Generation Face Tracking Technology Using AAM TechniquesA", *Proceedings of International Symposium on Signals, Systems and Circuits*, 2007, Vol. 1, pp. 1-4.
- [13] Gyanendra. K. Verma, Tiwary V.S, and Mahendra. K. Rai, " Facial Emotion REcognition using Different Muti Resoluiton Transforms", *Advances in Computing and Communication*, 2011, Vol. 192, pp. 469-477.
- [14] Reena Rose, R., and Suruliandi, A., " Improving Performance of Texture Based Face Recognition Systems by Segmenting Face Region", *ACEEE Int. Journal. on Network Security* , 2011, Vol. 02, No. 03.
- [15] Guoying Zhao and Matti Pietikainen, "Dynamic Texture Recognition using Local Binary Patterns with an Application to Facial Expressions", *IEEE transactions on Pattern Analysis and Machine Intelligence*, 2007.
- [16] Mohanta, P., P, Saha. S.K, ChandaPardàs, M., Bonafonte, A. and Landabaso, J, B, "Emotion recognition based on MPEG4 facial animation parameters, " *Proceedings of IEEE ICASSP*, 2002.
- [17] Cohen, I., Sebe, N., Cozman, f., Cirelo, M., and Huang, T., "Learning Bayesian network classifiers for facial expression recognition using both labelled and unlabeled data", *Proceedings of the 2003 IEEE CVPR*, 2003.
- [18] Susskind, J.M., Littlewort, G., Bartlett, M.S., Movellan, J., and Anderson, A.K., " Human and computer recognition of facial expressions of emotion, Neuro psychologia", 2007, Vol. 45, pp. 152-162.
- [19] Geetha, A., Ramalingam, V. and Palanivel, S., "Facial expression recognition, A real time approach", *Expert Systems with Applications*, 2009, Vol. 36, pp. 303-308.
- [20] Haralick, R. M., Shanmugam, K., ItsŠ hak Dinstein, " Texture features for image classification", *IEEE Transactions on Systems, Man and Cybernetics*, 1973, 3(6), pp. 610-621.
- [21] Hong, Q. Q., and Wang, B. Z., " An improved approach for image retrieval based on generalized co-occurrence matrix", *Journal of Computational Information Systems*, 2008, 4: 1, pp. 97-104.
- [22] Fritz Albrechtsen, "Statistical Texture Measures Computed from Gray Level Cooccurrence Matrices ", *Image Processing Laboratory Department of Informatics University of Oslo*, 2008, November 5.
- [23] Vapnik, V. N., "The nature of statistical learning theory". New York, Springer, 1995.
- [24] <http://chenlab.ece.cornell.edu/projects/FaceAuthentication>
- [25] Chih-Chung Chang and Chih-Jen Lin "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, 2011, 2, pp. 1-27.