ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 11, November 2013

Facial Emotion Recognition Using Optimization Technique

S.Ramya, G.Sophia Reena

Research Scholar, IT Department, PSGR Krishnammal College for Women, Coimbatore, Tamil Nadu, Coimbatore

Assistant Professor, IT Department, PSGR Krishnammal College for Women, Coimbatore,Tamil Nadu ,Coimbatore

Abstract: Facial Emotion Expressions plays an important role in interpersonal relations.This is because human convey lot of information visually than verbally.To automate recognition of emotion state,machine should taught to understand facial guestures.In this paper we classify emotion expression through Support Vector Machine(SVM) & Hidden Markov Model (HMM),then Hidden Markov Model is optimized using Glowworm Swarm Optimization(GSO)to get better accuracy.The system was tested on Japanese Female Facila Expression dataset of frontal view facial expression.

Index Terms: Feature Extraction, SVM Classification, HMM Classification, HMM with GSO

I. INTRODUCTION

Human Being express their emotions in everyday interactions.Emotions are frequently reflected on face,in hand & body gestures.Recent Psychology research has shown that the most expressive way humans display emotion is through facial expressions.Mehrabian[1] indicated that verbal part of a message contribute only for 7% to effect of message as a whole,vocal part 38%,while facial expression for 55%to effect of speaker's message.Ekman & Friesen [1] represent 6 basic facial expressions (emotions). The process of expression Recognition involves processing images by detectingthe face, extracting the facial features, and then usingan algorithm to identify the expressions made basedon the movements of the feature made

In this paper,there are five phase for Facial Feature Recognization they are Face Detection, Feature Extraction, Selection ,Classification Technique.In the First phase Face Detection extract the face from scenes.So system positively identifies a certain image region as face.In Second phase Feature extraction involves extracting feature from the image using Log Gabor filter. In Third stage, for feature selection we use Independend Component Analysis used for selecting unique feature from extacted feature.In fourth stage we use to recognize the face by using classification technique they are SuppotiveVector Machine & Hidden Markov Model.Hidden Markov Model is then optimized using Glowworm Swarm Optimization Algorithm.Experiment result that classification using Glowworm Swarm Optimization with Hidden Markov Model give better Accuracy rate for Facial Recognization

The rest of this paper is organized as follows Section 2 describes Related Work, Section 3 describes Methodology, Section 4 describes Conclusion

II. RELATED WORK

In the past decade, many of researches have been done on machine recognition of human facialexpressions. Convention approaches extracted features of facial components, such as eyes and mouth, in gray or color images of frontal view faces. Properties and relations (e.g., areas, distances, angles) between the features were used as descriptors of faces for recognizing the expressions fromtheir changing properties or their geometrical relationships by different particular facial expressions [3]

More work has been made due to advances in the appearance-based and learning-based approaches.Viola and Jones used the Adaboost algorithm to scan an image by passing multiple scales sub-windowfor rapid face detection [4]. Littlewort used a bank of forty Gabor wavelet filters at different scalesand orientations to perform a data extraction, then generated image sequences image-by-image to train two stages of support vector machines from Gabor filter jets [5].

Ekman and Friesen developed the Facial Action Coding System (FACs) tocode expressions as a combination of forty four Action Units, and defined six basic emotions: angry, disgust, fear, happy, sad, and surprise [6]. Cohn used displacement vectors of facial feature points torepresent the extracted expression information, applied separate discriminant functions together withvariance-covariance matrices to different facial regions displacements as predictors for classification [7]. Dailey used a six unit, a single layer neural network to classify into six basicemotion categories given Gabor jets extracted from static images [8].

III. METHODOLOGY

A. Database Preparation

Database is prepared using JAFFE Face Database.Initial conditions for selection & preparation of faces databases can be summarized for the proposed system are as follows:

- All images must be taken in similar & uniform illumination conditions
- No physical obstruction
- Camera at same distance from face
- We need to take Unique Feature from face
- Image must be in gray format or changed to gray format
- Images should be of equal size & small,for best result & fast processing.

B. Face Detection

Face detection is the first stage of an automated face recognition system, since a face has to be located before it is recognized. The face-detection component of our method is based on the well-known Viola-Jones detector. This detector is known for its computational efficiency and performance. Viola and Jones employed efficient machine learning.

C. Feature Extraction

Salient features are extracted from biometric images to uniquely represent the acquired image in the form of a numeric template which can be enrolled in a system databasefor comparison, matching or

Classification purposes. Other advantages of feature extraction are dimensionality reduction, normalisation, security and faster recognition of individuals.To achieve thebroad spectral information and to overcome the bandwidth limitation of the traditional Gabor filter, Field proposed Log-Gabor filter. The response of the Log-Gabor filteris Gaussian when viewed on a logarithmic frequency scale instead of a linear one. This allows more information to becaptured in the high-frequency areas with desirable high pass characteristics. In this contribution, a bank of 24 Log-Gaborfilters is employed to extract the facial features. The polarform of 2-D Log-Gabor filters in frequency domain is given by

$$H(f,\theta) = \exp\left\{\frac{-\left[\ln\left(\frac{f}{f_0}\right)\right]^2}{2\left[\ln\left(\frac{\sigma_f}{f_0}\right)\right]^2}\right\} \exp\left\{\frac{-(\theta-\theta_0)^2}{2\sigma_\theta^2}\right\} \quad (1)$$

where $H(f, \theta)$ is the frequency response function of the2-D Log-Gabor filters, f and θ denote the frequency and the phase/angle of the filter, respectively, f_0 is the filter'scenter frequency and θ_0 the filter's direction. The constant σ_f defines the radial bandwidth B in octaves and the constant $\sigma\theta$ defines the angular bandwidth $\Delta\Omega$ in radians

$$B = 2\sqrt{\frac{2}{\ln 2}} \times \left|\ln\left(\frac{\sigma_f}{f_0}\right)\right|, \quad \Delta\Omega = 2\sigma_\theta\sqrt{\frac{2}{\ln 2}}. \quad (2)$$

the ratio $\sigma f / f0$ is kept constantfor varying $f0$, B is set to one octave and the angularbandwidth is set to $\Delta\Omega = \pi/4$ radians. This left only σf to be determined for a varying value of $f0$. Six scales andfour orientations are implemented to extract features fromface images. This leads to 24 filter transfer Functions representing different scales and orientations. The image filteringis performed in the frequency domain making the processfaster compared with the spacial domain convolution. Afterthe 2-D fast Fourier transform (FFT) into the frequencydomain, the image arrays, **x**, are changed into the spectralvectors, **X** and multiplied by the Log-Gabor transfer functions $\{H1, H2. H24\}$, producing 24 spectral representations foreach image. The spectra are then transformed back to the spatial domain via the 2-D inverse FFT.

D.Feature Selection

In feature selection which is based on independent component analysis one can consider an independent component si as the i-th feature of the recognizedobject represented by the observed pattern vector x. The feature patterncan be formed from m independent components of the observed data pattern.

The use of ICA for feature extraction is mainly used for revealing the similar principle of pattern dimensionality can be found in the early processing of sensory data by the brain.

In order to form the ICA patterns the following procedure should be followed

1. Extraction of nf element feature patterns xf from the recognition objects.Composing th original data set Tf containing N cases {xTf ,i, ci}. The featurepatterns are represented by matrix Xf and corresponding categorical classes are represented by column c.

2. Heuristic reduction of feature patterns from the matrix Xf into nfr elementreduced feature patterns xfr (with resulting patterns Xfr). This step could be directly possible for example for features computed as singular values of image matrices.

3. Pattern forming through ICA of reduced feature patterns xfr from the data set Xfr.

(a) Whitening of the data set Xfr reduced feature patterns ofdimensionality nfr into nfrw element whitened patterns xrfw (projectedreduced feature patterns into nfrw principal directions).

(b) Reduction of the whitened patterns x*frw* into first *nfrwr* element reducedwhitened patterns x*frwr* through projection of reduced feature patternsinto first principal directions of data.

4. Computing the unmixing matrix W and computing reduced number *nicar*of independent components for each pattern x*frwr* obtained from whiteningusing ICA (projection patterns x*frwr* into independent component space).

5. Forming *nicar* element reduced ICA patterns x*icar* from corresponding independentcomponents of whitened patterns, with the resulting data set X*icar*.Forming a data set *Ticar* containing pattern matrix X*icar* and original classcolumn c

6. Providing rough sets based processing of the set *Ticar* containing ICA patterns x*icar*. Discretizing pattern elements and finding relative reducts fromset *Ticar*. Choosing one relevant relative reduct.

Selecting the elements ofpatterns x*icar* corresponding to chosen reduct and forming the final patternx*fin*. Composing the final data set *Tfinal,d* containing discrete final patternsx*fin,d* and class column. Composing the real valued data set *Tfin* from the set *Ticar* choosing elements of real-valued pattern using selected relative reduct.

E.Requirements Analysis: Image-based Classification Application

This section details a requirements analysis for the image-based classification application. An Overview of the application is illustrated below:

1. Allow user to select images to be used for training
2. Perform feature extraction on the images to gather numerical SVM input
3. Allow labeling of training examples with the corresponding emotion classes (e.g. joy, sorrow, anger, etc.), defined by the user
4. Make SVM parameters (kernel function type, SVM algorithm used, etc.) user-controllable
5. Train an SVM instance:
(a) Gather training data from processed input images
(b) Create an SVM model from the data
6. Classify unseen examples:
(a) Allow selection of an unseen example & perform feature extraction on it
(b) Classify unseen example according to model
7. Report results back to user
8. Code should be modular enough to allow possible later use of a variety of feature extraction Methods

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 11, November 2013

F.Facial Expression Classification: SVMs Classifier

Support Vector Machines (SVMs) are a machine learning system that is directlybasedon a branch of mathematics called statistical learning theory, which became properly formalized in thelast decade. While SVMs have successfully been applied to a number of classificationt asks and tend to often outperform classic neural network approaches, their application to facial expression analysis and emotion classification so far has been limited. The aim of this study is to apply SVMs to automated emotion classification.Machine learning algorithms are systems that receive input data during a training phase, then build a decision model according to the input and generate a function that can be used to predict a futuredata.

By given a set

$$D=\{(X_{i,}\ Yi)\}^{l}{}_{i}=1 \qquad (3)$$

Of labeled training examples, where

$$Y_i \in \{-1, 1\} \qquad (4)$$

Learning system typically try to find a decision function of the form

$$f(x)= sgn([w.x]+b) \qquad (5)$$

where *w* is a vector of weights and *b* is called bias, that yields a label ∈ {-1, 1} (for the basic case ofbinary classification) for an unseen example *x*, which have the smallest generation error.

The SVMs technique performs projection the original set of variables *x* in higher dimensional Feature space:

$$\mathrm{x} \in R^{d} \rightarrow z(x) \equiv (\emptyset_1(x), \ldots, \emptyset_n(x)) \in R^{2n} \qquad (5)$$

where a linear algebra and a geometry may be used to separate data that is only separable withnonlinear rules in an input space. By formulating the linear classification problem in the feature space, the solution will be in the form

$$f(x) = sgn\left(\sum_{i=1}^{l} \alpha_i y_i K(x, x_i\right) + b) \qquad (6)$$

associates with using the kernel functions, allowing an efficient computation of inner products directly in a feature space, by given a nonlinear mapping Φ that embeds input vectors into a feature space, kernels will have the form

$$K(x,z) = [\Phi\ (\mathrm{x}).\Phi\ (\mathrm{z})] \qquad (7)$$

Where the αi are Lagrange multipliers of a dual optimization problem. It is possible to show that onlysmall number coefficients αi is different from zero, and since every coefficient corresponds to aparticular data point, this means that the solution is determined by the data points associated to the nonzerocoefficients. These data points, which are called, Support Vector. These induce sparseness in the solution and give rise to efficient approaches to optimization.

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 11, November 2013

Once a decision function has been obtained, classification of an unseen example x amounts to checking which of the two subspaces defined by theseparating hyperplane the example lies in. Figure 1 illustrates using a simple linear classifier.

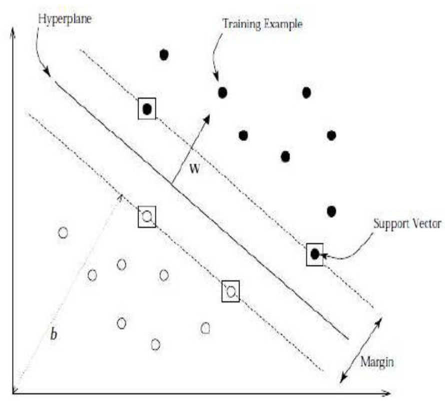

Figure: 1 A linear maximal margin classifier.Thehyperplane's normal vector w and the bias b define the decision function $f(x)= sgn([w.x]+b)$. The support vector lies on the margin

G. HMM Classifier

Each prototypic expression is modeled as an HMM. Let $\lambda = [A,B, \pi]$ denote an HMM to b trained and N be the number of hidden states in the model, we denote the statesas $S = \{S1, S2, ..., SN\}$and the state at t is qt.$A = \{aij\}$isthe state transition probability distribution, where $aij = P [qt+1 = Sj \mid qt = Si]$, $1 \leq i, j \leq N$. $B = \{bj (k)\}$is the observation probability distribution in state j, k is and observation . We use Gaussian distributions to estimate each $B = \{bj (k)\}$, where $bj (k) = P [k|qt = Sj] \sim N (\mu j,\Sigma j)$, $1 \leq j \leq N$. Let $\pi = \{\pi i\}$ be the initial state distribution, where $\pi i = P [q0 = Si]$, $1 \leq i \leq N$. Then, given an observationsequence, $O = O1O2...OT$, where Oi denote an observation at times i, the training procedure is:

Step1: Take the optimized feature representation OG of each observed facial Emotion Expression as an observation

Step2: Initialize the HMM model λ. Eachobserved model of a sequence corresponds to one state and is used to estimate the parameters in the observation matrix B. Set initial values of A and π based on observations.

Step 3: Use the forward-backward algorithm to derive an estimation of the model parameter $\lambda = [A, B, \pi]$ when $P (O|\lambda)$ is maximized. Finally, we derived6 HMMs; each represents one of the six prototypic expressions

Given a query model sequence, we follow the Step 1 of the training procedure to represent it as $Q = Q1Q2...QT$, where the optimized representation of each frame is one observation, denoted as$OG =(OG,1,OG,2,OG,3,OG,4,OG,5)$. Using the forwardbackwardmethod, we compute the probability of the observation sequence given a trained $HMMi$ as $P (Q|\lambda i)$

H.HMM with GSO

Step 1: Initialize GSO

GSO→i…….n, Best_ fitness =0; Where i=1

Step 2: initialize solution of each GSO with HMM for training dataset

Step 3: Evaluate solution of each GSO

Evaluate GSOi ----> Accuracy

Step 4: If max (GSOi ----> Accuracy) > Best_fitness

Best_fitness ← max (GSOi ----> Accuracy)

Step 5: Update Luciferin value of glowworm i

$$l_i(t) = (1-\rho)l_i(t-1) + \gamma J(x_i(t));$$

Step 6: Probability for position update

$$p_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)}; \quad (8)$$

Step 7:Update Movement update

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t) - x_i(t)}{\left\|x_j(t) - x_i(t)\right\|}\right); \quad (9)$$

Step 8: Goto Step2 till calculating the best fitness value of GSO

Step 9: Best fitness ←max (fitness)

Optimized value←best fitness

Step 10: Return Optimized fitness value, luciferin update, position update, movement update

Step 11: The calculated transition probability and observation probability of Best HMM of trained result will be used for prediction.

I. Experimental Results

The database we use in our experiments contains 181 images of female facial expressions. They were collected by Kamachi and Gyoba at Kyushu University, Japan. Ten expressors were asked to pose several different facial expressions. Each expressor, when ready, took pictures of herself, through remote control, while looking towards the camera through a semi-reflective plastic sheet. Original images have been rescaled and cropped such that the eyes are roughly at the same position with a distance of 60 pixels in the final images (resolution: 256 pixels *256pixels). The number of images corresponding to each of the 6 categories of expression (neutral, happiness, sadness, surprise, anger, and fear) is roughly the same. For Testing we take 121 images (1/3) for Training 60 images .A few of them are shown in Fig. 2

FIG: 2 JAFFE DATABASE

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 11, November 2013

Table 1: shows Confusion Matrix of Six Facial Emotions for SVM Classification

Emotion	Angry (22)	Fear (21)	Happy (22)	Neutral (18)	Sad (17)	Surprise (20)
Angry	19					3
Fear		18	1	1	1	
Happy	1		20	1	1	
Neutral	1	1		14	1	1
Sad	1	1	1	2	14	2
Surprise		1				14

Table: 2 shows Confusion Matrix of Six Facial Emotion for HMM Classification with GSO

Emotion	Angry (23)	Fear (21)	Happy (22)	Neutral (18)	Sad (17)	Surprise (20)
Angry	23					
Fear		19			2	
Happy		1	22	1		
Neutral				16		
Sad		1		1	15	1
Surprise						19

Table: 3 shows Confusion Matrix of Six Facial Emotions for HMM Classification

Emotion	Angry (23)	Fear (21)	Happy (22)	Neutral (18)	Sad (17)	Surprise (20)
Angry	21		1	1	1	
Fear		18	1			1
Happy	2	1	20	1	2	
Neutral				16		
Sad		1			14	1
Surprise		1				18

Table:4 shows Recognition Rate of SVM, HMM, HMM with GSO

Emotion	Percentage of Accuracy SVM	Percentage of Accuracy of HMM	Percentage of Accuracy of HMM with GSO
Anger	86%	91%	100%
Fear	86%	86%	90%
Happy	91%	91%	100%
Neutral	78%	89%	89%
Sad	82%	82%	88%
Surprise	70%	90%	95%

IV. CONCLUSION

Emotion classification is a challenging application for much future human-computer interaction. This paper describes the emotion classification through facial expressions using the SVM, HMM and HMM with GSO .The system extracts the facial features, and feed it into the SVM, HMM,HMM with GSO to categorize result expression. Two main contribution of this paper. First, the demonstration of the facial feature extraction scheme enables robust emotion classification .Second HMM with GSO is a suitable engine for reliable Classification. The classification Rate of all 6 basic emotion of SVM is82% , HMM is 88%and HMM with GSO is 94% approximately, high percentage of accuracy can be achieved at classifying two distinct emotion expression (Happy & Anger). This clearly shows that the proposed technique classifies facial emotion with better accuracy than Facial emotion with SVM and HMM.

Future Work

- Develop New Feature Selection Method to cope with large Database & improve overall System Performance
- There are strong Classifier that have achieved good Performance in Face Reorganization problem .Nowadays, there is a trend to combine different Classifier in order to get the best performance.
- Use Better Optimization Technique that can be combined with Classifier to give better Accuracy Rate

Reference

[1] Mehrabian .A, “Communication without Words”, Psychology Today, Vol 2(4), PP 53-56, 1968

[2] M. Pantic and L. Rothkrantz,”Automatic Analysis of Facial Expressions: The State of the Art”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22. no.12.December 2000

[3] K. Matsuno, C-W Lee, S. Kimura& S. Tsuji ,”Automatic Recognition of Human Facial Expressions”, In Proceeding of the Fifth International Conference on Computer Vision (ICCV’95),1995

[4] P. Viola and M. Jones, “Robust Real-Time Object Detection”, Technical Report 2001/01.Compaq Cambridge Research Lab,2001.

[5] G. Littlewort, I. Fasel, M. Stewart Bartlett and J. R. Movellan, “Fully Automatic Codingof Basic Expressions from Video”, Technical Report 2002.03, UCSD INC MPLab.2002

[6] P.Ekman and W. Frisen, “Facial Action Coding System (FACS)”, Manual. ConsultingPsychologists Press. Palo Alto. CA. USA, 1978

[7] J. F. Cohn, A. and T. Kanade,” Automated face analysis by feature-point tracking has high concurrent validity with manualfacescoding”, Psychophysiology. 36:35-43, 1999

[8] M. N. Dailey, G. W. Cottrell, and R. Adolphs,” A six-unit network is all you need to discover happiness”, In Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society. Mahwah. NJ. USA, 2000