Emotion Recognition from Facial Expression Using Neural Networks

Govind U.Kharat †, and Sanjay V. Dudul ‡
† Anuradha Engineering College, Chikhli, India
‡ VIIT, Pune (Maharashtra), India
gukharat@rediffmail.com, dudulsv@rediffmail.com

Abstract - **This research aims at developing "Mind Implemented Robots" that can carry out intellectual conversation with human beings. The first step in this direction is to recognize human emotions by a computer using neural network. Various neural networks are employed and their performance is compared. We achieved 100% recognition rate on training data set and test data set.**

Keywords - **Multilayer Perceptron (MLP), Principal Component Analysis (PCA), Generalized Feed Forward Neural Network (GFFNN), Features Extraction**.

I. INTRODUCTION

It is highly expected that computers and robots will be used more for betterment of our daily life. Information-Computerized Society expects a harmonious interaction or heart to heart communication between computers and / or robots and human beings. For its realization it seems to be necessary that computers and robots will be implemented with artificial mind that enables them to communicate with human beings through exchanging not only logical information but also emotional one. The first step to realize mind implemented robot is to recognize human emotions. Mehrabian [1] indicated that the verbal part (i.e. spoken words) of a message contributes only for a 7% of the effect of the message, the vocal part (i.e. voice information) contributes for 38% while facial expressions of the speaker contributes for 55% of the effect of the spoken message. Hence in order to develop "Active Human Interface" that realizes heart to heart communication between intelligent machine and human beings we are implementing machine recognition of human emotions from facial expressions.

Affective computing addresses issues relating to emotion in computing and has been pioneered by the work of Picard at MIT [2]. Picard describes how "Affective interaction can have maximum impact when emotion recognition is available to both man and machine" and goes on to say if one party can not recognize or understand emotion then interaction is impaired [3]. The problem of recognizing facial expressions had attracted the attention of computer- vision community [4]-[11]. Bassili [12] suggested that motion in the image of the face would allow emotions to be identified even with minimal information about the spatial arrangement of features. Essa and Petland [13] and Essa [14] proposed FACS+ model extending Facial Action Coding System (FACS) model to allow combine spatial and temporal modeling of facial expressions. Optical flow computations for recognizing and analyzing facial expressions are used by [5], [7], [9], [11], [15]-[28].

Anthropometric facial points are used for feature extraction to recognize emotions [9], [11], [29]. Automatic recognition of facial expression [29]-[36] and their usage [37], [38] presents a number of different challenges. In general five main processes are distinguished in tackling the problem.

1. Robust and automated face detection system for the segmentation of face region.
2. Image analysis to extract meaningful information from video sequences to construct feature vectors for the recognition of facial expressions.
3. Data representation techniques that can keep useful information at the same time reduce the dimensionality of the feature vector to design better classifier.
4. Design and implementation of robust classifiers to learn the underlying models of facial expressions.

This paper provides a simplest approach of using an image and obtain its Discrete Cosine Transform (DCT) to extract features along with physical parameters namely energy, entropy, variance, standard deviation, contrast, homogeneity and correlation to obtain intelligent feature vector for the recognition of facial expression. MLP, PCA and GFFNN are used for recognition of emotions.

II. FACIAL EXPRESSION DATABASE

Facial expression database in six emotions and neutral one is collected for Japanese females. Ten expressers posed 3 to 4 examples of each of the six emotions along with neutral one for a total of 219 images of facial expressions. This data was prepared when expresser look into the semi reflective plastic sheet towards camera. Hairs were tied away to expose all expressive zones of the face. Tungsten lights were positioned to create even illumination on the face. The box enclosed the region between camera and plastic sheet to reduce back reflections. The images were printed in monochrome and digitized using flatbed scanner. Sample images are shown in figure 2. This authentic database is collected from data repositories. 210 images are selected for our experiment.

III. COMPUTER SIMULATION EXPERIMENT

A. *Formation of Feature Vector*

We have developed a program in MATLAB to obtain DCT and physical parameters namely energy, entropy, variance, standard deviation, contrast, homogeneity and correlation of 210 images. An intelligent feature vector is obtained containing the features extracted by DCT and physical parameters of each image. Thus dataset for all 210 images is prepared to feed to Neural Network for emotion recognition.

1-4244-1543-8/08/$25.00 ©2008 IEEE

B. *Neural Networks for Emotion Recognition*

The generalized procedure for emotion recognition from facial expressions using different neural networks is shown in figure 1. We have used Multilayer Perceptron (MLP), Principal Component Analysis (PCA), and Generalized Feed Forward Neural Network (GFFNN) one by one for emotion recognition using facial expressions. Number of input Processing Elements (PE) must be equal to that of input data of facial information. Since we have used 64 DCT & 07 physical parameters of an image, 71 input Processing Elements are used in input layer. Seven Processing Elements are used in output layer for six emotions and neutral one.

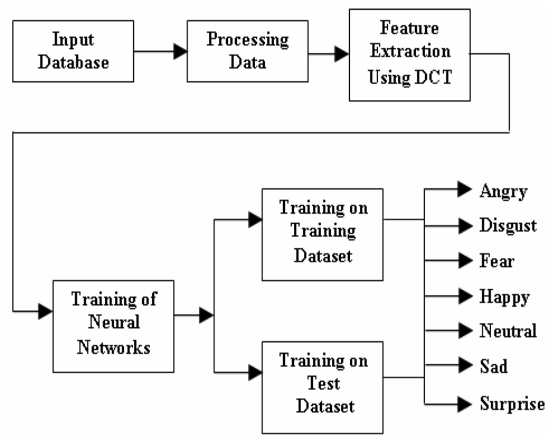


Fig. 1. Procedure for Emotion Recognition.

I. MLP for Emotion Recognition

The randomized dataset is fed to the MLP. Vigorous experimentation is done by varying the number of hidden layers, number of Neurons in a hidden layer, step size, learning rule, type of axon, number of epochs, number of runs and termination criteria, percentage of data for training and testing to obtain optimal selection of MLP as follows:

Tag data CV = 10% and train = 90%
Input P.Es. = 71
Output P.Es. = 07
Exemplars = 189
Hidden layer = 01

Hidden layer 1 –
Number of P.Es. = 10
Transfer function = tanhAxon
Learning rule = Momentum
Step size = 0.1
Momentum = 0.7

Output layer –
P. Es. = 07
Transfer function = tanhAxon
Learning rule = Momentum
Step size = 0.01
Momentum = 0.7

Supervised learning control –
Maximum epochs = 5000
Termination – MSE, CV set, increase
Weight update = Batch

Training of Network –
Number of epochs = 5000
Number of runs = 03

Termination is after 100 epochs without improvement.
Time elapsed per epoch per examplars = 0.1117 mSec.
Number of pre parameters (P) for MLP = 797
No. of examplars in training dataset (N) = 189
N/P ratio = 0.2371

Network is tested on training and test dataset. The results are shown in table1 to 4.

II. PCA for Emotion Recognition

Tag data CV = 15% and train = 85%
Input P.Es. = 71
Output P.Es. = 07
Exemplars = 178
Hidden layer = 01
Principal Components = 21
Learning rule = Sangers full

Hidden layer 1 –
P. Es. = 19
Transfer function = tanhAxon
Step size = 0.1
Momentum = 0.7

Output layer –
P. Es. = 07
Transfer function = tanhAxon
Learning rule = Momentum
Step size = 0.001
Momentum = 0.7

Unsupervised learning –
Maximum epochs = 100
Weight update = 0.0001

Learning rate starts at 0.1 and decay to 0.00001

Supervised learning control –
Maximum epochs = 3000
Termination- MSE, CV set
Minimum threshold = 0.000001

Training the network-
Number of epochs = 3100
Number runs = 03

Termination is after 100 epochs without improvement
Time elapsed per epoch per examplars = 0.0129 mSec.
Number of pre parameters (P) for MLP = 1508
No. of examplars in training dataset (N) = 178
N/P ratio = 0.118

Network is tested on training and test dataset, the result are as shown in table 5 to 8.

III. GFFNN for Emotion Recognition

Tag data CV = 10% and train = 90%
Input P.Es. = 71
Output P.Es. = 07
Exemplars =189
Hidden layer = 01

Hidden layer 1 –
Number of P.Es. = 29
Transfer function = tanhAxon
Learning rule = Momentum
Step size = 0.1
Momentum = 0.7

Output layer –
P. Es. = 07
Transfer function = tanhAxon
Learning rule = Momentum
Step size = 0.01
Momentum = 0.7

Supervised learning control –
Maximum epochs = 3000
Termination – MSE, CV set, increase
Weight update = Batch
Training of Network –
Number of epochs = 3000
Number of runs = 03
Termination is after 100 epochs without improvement.
Time elapsed per epoch per examplars = 0.0494 mSec.
Number of pre parameters (P) for MLP = 1508
No. of examplars in training dataset (N) = 189
N/P ratio = 0.1253

Network is tested on training and test dataset and results are shown in table 9 to 12.

IV. RESULT

In this paper, we evaluated the performance of the three Neural Networks MLP, PCA and GFFNN for recognition of emotions from facial expressions to develop a "mind implemented computers and robots" and examined the recognition results.

Table 1 to 4 shows emotion recognition results on training and testing data set using MLP. The accuracy of recognition is 100% on test dataset. On training dataset accuracy of recognition is 100% on angry, disgust, fear, happy, neutral and sad emotions and overall efficiency is 99.47%.

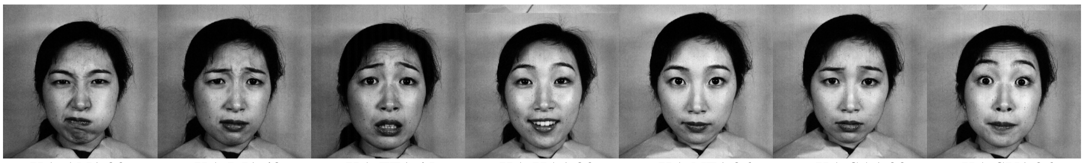
Table 5 to 8 indicate that for PCA the accuracy of recognition on training dataset is 95.46% but gives 100% recognition result on test dataset. Table 9 to 12 show that the result of emotion recognition on training and test dataset using GFFNN. The accuracy of recognition on training and testing dataset is 97.42 and 100% respectively. So we conclude that MLP is an excellent classifier.

V. CONCLUSION.

The performance comparison for the various Neural Networks for emotion recognition is shown below.

Neural Network Model	Average Minimum MSE		Average % Classification Accuracy		Time elapsed per epoch per exemplar	N/P Ratio
	Train	CV	Train	CV		
MLP	0.0028	0.0358	99.47	100	0.1117 mSec.	0.2371
PCA	0.0087	0.0803	95.47	100	0.0129 mSec.	0.118
GFFNN	0.0048	0.0831	97.42	100	0.0496 mSec.	0.1253

It is observed that time elapsed per epoch per exemplar is lowest for PCA, indicating that it is fastest network. N/P ratio is highest for MLP, is an indication of simplicity of MLP neural network in design and synthesis. The average minimum MSE is lowest and average percentage accuracy is highest for MLP. It means MLP is most suitable for human emotion recognition from facial expressions using DCT.



KA.AN1.39 KA.DI1.42 KA.FE1.45 KA.HA1.29 KA.NE1.26 KA.SA1.33 KA.SU1.36

Fig.2. Images of Japanese female KA in various emotions from database.

TABLE 1: CONFUSION MATRIX FOR TRAINING DATASET USING MLP

Output/Desired	Oa	Od	Of	Oh	On	Os	Osu
Oa	**27**	0	0	0	0	0	0
Od	0	**26**	0	0	0	0	1
Of	0	0	**27**	0	0	0	0
Oh	0	0	0	**29**	0	0	0
On	0	0	0	0	**25**	0	0
Os	0	0	0	0	0	**28**	0
Osu	0	0	0	0	0	0	**26**

TABLE 2: PERFORMANCE PARAMETERS FOR TRAINING DATASET USING MLP

Perform.	Oa	Od	Of	Oh	On	Os	Osu
MSE	0.002259705	0.002384747	0.002210249	0.002191703	0.002843431	0.002440459	0.008097807
% Correct	100	100	100	100	100	100	96.2962963

The overall accurate recognition of emotion is = 99.47%

TABLE 3: CONFUSION MATRIX FOR TEST DATASET USING MLP

Output / Desired	Oa	Od	Of	Oh	On	Os	Osu
Oa	**3**	0	0	0	0	0	0
Od	0	**4**	0	0	0	0	0
Of	0	0	**3**	0	0	0	0
Oh	0	0	0	**1**	0	0	0
On	0	0	0	0	**5**	0	0
Os	0	0	0	0	0	**2**	0
Osu	0	0	0	0	0	0	**3**

TABLE 4: PERFORMANCE PARAMETERS FOR TEST DATASET USING MLP

Perform.	Oa	Od	Of	Oh	On	Os	Osu
MSE	0.002895283	0.045778014	0.009529069	0.006264617	0.029641739	0.038633208	0.028475979
% Correct	100	100	100	100	100	100	100

The overall accurate recognition of emotion is = 100%

TABLE 5: CONFUSION MATRIX FOR TRAINING DATASET USING PCA

Output / Desired	Oa	Od	Of	Oh	On	Os	Osu
Oa	**24**	0	0	0	0	0	0
Od	0	**22**	1	0	0	0	0
Of	1	1	**25**	0	0	0	0
Oh	0	0	0	**29**	0	1	1
On	0	0	0	0	**27**	0	1
Os	0	0	0	0	0	**18**	1
Osu	0	0	1	0	0	0	**25**

TABLE 6: PERFORMANCE PARAMETERS FOR TRAINING DATASET USING PCA

Perform.	Oa	Od	Of	Oh	On	Os	Osu
MSE	0.011987136	0.023911139	0.031023655	0.013048545	0.018419962	0.022382219	0.023720358
% Correct	96	95.65217391	92.59259259	100	100	94.73684211	89.28571429

The overall accurate recognition of emotion is = 95.46%

TABLE 7: CONFUSION MATRIX FOR TEST DATASET USING PCA

Output / Desired	Oa	Od	Of	Oh	On	Os	Osu
Oa	**5**	0	0	0	0	0	0
Od	0	**7**	0	0	0	0	0
Of	0	0	**3**	0	0	0	0
Oh	0	0	0	**1**	0	0	0
On	0	0	0	0	**3**	0	0
Os	0	0	0	0	0	**11**	0
Osu	0	0	0	0	0	0	**2**

TABLE 8: PERFORMANCE PARAMETERS FOR TEST DATASET USING PCA

Perform.	Oa	Od	Of	Oh	On	Os	Osu
MSE	0.004538276	0.008625015	0.019240703	0.005533798	0.013757429	0.017308529	0.003722688
% Correct	100	100	100	100	100	100	100

The overall accurate recognition of emotion is = 100%

TABLE 9: CONFUSION MATRIX FOR TRAINING DATASET USING GFFNN

Output / Desired	Oa	Od	Of	Oh	On	Os	Osu
Oa	**25**	0	0	0	0	0	0
Od	0	**24**	0	0	0	0	1
Of	0	2	**27**	0	0	1	0
Oh	0	0	0	**26**	0	0	0
On	0	0	0	0	**27**	0	0
Os	0	0	0	0	0	**27**	0
Osu	0	0	0	0	0	1	**28**

TABLE 10: PERFORMANCE PARAMETERS FOR TRAINING DATASET USING GFFNN

Perform.	Oa	Od	Of	Oh	On	Os	Osu
MSE	0.003894195	0.016804945	0.008204842	0.002288221	0.008074253	0.016625264	0.013031139
% Correct	100	92.30769231	100	100	100	93.10344828	96.55172414

The overall accurate recognition of emotion is = 97.42%

TABLE 11: CONFUSION MATRIX FOR TEST DATASET USING GFFNN

Output / Desired	Oa	Od	Of	Oh	On	Os	Osu
Oa	**5**	0	0	0	0	0	0
Od	0	**4**	0	0	0	0	0
Of	0	0	**3**	0	0	0	0
Oh	0	0	0	**4**	0	0	0
On	0	0	0	0	**3**	0	0
Os	0	0	0	0	0	**1**	0
Osu	0	0	0	0	0	0	**1**

TABLE 12: PERFORMANCE PARAMETERS FOR TEST DATASET USING GFFNN

Perform.	Oa	Od	Of	Oh	On	Os	Osu
MSE	0.004986289	0.002692714	0.002150516	0.003050586	0.003722369	0.002523586	0.003095122
% Correct	100	100	100	100	100	100	100

The overall accurate recognition of emotion is = 100**%**

REFERENCES

[1] A.Mehrabian, "Communication without words," *Psychology today*, vol.2, no.4, pp.53-56, 1968.

[2] R.W. Picard, *Affective Computing*. Cambridge.MA : MIT Press, 1997.

[3] R.W.Picard, "Toward agents that recognize emotion", in *Proc. IMAGINA*, 1998, pp. 153–155.

[4] D. Beymer, A. Shashua, and T. Poggio, Example Based Image Analysis and Synthesis, M.I.T. A.I. Memo No. 1431, 1993.

[5] I.A. Essa and A. Pentland, "A Vision System for Observing and Extracting Facial Action Parameters", Proc. *IEEE CVPR*, pp.76-83, 1994.

[6] H. Li, P. Roivainen, and R. Forcheimer,"3D Motion Estimation in Model-Based Facial Image Coding", *IEEE Trans. Pattern Analysis and Machine intelligence,* vol. 15, pp. 545-555,1993.

[7] K. Mase, "Recognition of Facial Expression from Optical Flow", *IEICE Trans*., vol. E 74, pp. 3474-3483, 1991.

[8] K. Matsuno, C. Lee, and S.Tsuji, "Recognition of Human Facial Expressions without Feature Extraction", *Proc. ECCV,* pp. 513-520, 1994.

[9] M. Rosenblum, Y. Yacoob, and L.S. Davis, "Human Emotion Recognition from Motion Using a Radial Basis Function Network Architecture", IEEE Workshop Motion of Non-Rigid and Articulated Objects, Austin, Texas, pp. 43-49, Nov. 1994.

[10] D. Terzopoulos and K. Waters, "Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, pp. 569-579, 1993.

[11] Y.Yacob and L Devis, "Recognizing Human facial expression from long image sequences using optical flow", *IEEE transaction on Pattern Analysis and Machine Intelligence [PAMI]*, 18{6}: 636-642, 1996.

[12] J. N. Bassili, Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face, ***J.*** *Personality and Social Psych.,* vol. 37, pp. 2049-2059, 1979.

[13] Irfan A. Essa and Alex P. Pentland, **"**Coding, Analysis, Interpretation and Recognition of Facial Expressions", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, July 1997.

[14] Analysis, Interpretation and Synthesis of Facial Expressions by Irfan Aziz Essa, Ph.D. thesis, MIT February 1995.

[15] Gianluca Donato, Marian Stewart Bartlett, Joseph C. Hager, "Paul Ekman, and Terrence J. Sejnowski, Classifying Facial Actions", *IEEE Transaction Pattern Analysis and Machine Intelligence,* vol. 21, No. 10 October 1999.

[16] I. Essa and A. Pentland, "Coding, Analysis, Interpretation and Recognition of Facial Expressions", IEEE *Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 757-763, July 1997.

[17] Mark Rosenblum, Yaser Yacoob, and Larry S. Davis, Senior Member, IEEE, "Human Expression Recognition from Motion Using a Radial Basis Function Network Architecture", *IEEE Transactions on Neural Networks*, Vol. 7, No. 5, September 1996

[18] P. Ekman and W.Friesen. *The Facial Action Coding System.* SanFrancisco CA : Consulting Psychologist Press, 1978.

[19] Y. Yacoob and L. S. Davis, "Computing Spatio-temporal representations of human faces", in *IEEE trans. on Computer Vision and Pattern Recognition,* 1994, pp. 70-75.

[20] K. Mase and A. Pentland, "Recognition of facial expression from optical flow", *IEICE Trans*. E, vol. 74, pp. 408–410, 1991.

[21] C. Padgett and G. W Cottrell, "A simple neural network models categorical perception of facial expressions", in *Proc. 20th Annu. Cognitive Science Conf.*, 1998, pp. 806–807.

[22] B. Fasel and J. Luttin, Recognition of Asymmetric Facial Action Unit Activities and Intensities, IDIAP Research Rep., 1999.

[23] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models", *IEEE Trans. Pattern Analysis and Machine Intelligence.*, vol.19, no.7, pp.743-756, Jul.1997.

[24] M.Matsugu, K.Mori, Y.Mitari, and Y.Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network", J. *Neural Network*, vol. 16, pp. 555–559, 2003

[25] M.Gargesha and P.Kuchi, "Facial expression recognition using artificial neural networks," J. *Artificial Neural Computing System*, pp. 1–6, 2002.

[26] A.Johnston, P. W. McOwan, and C. P. Benton, "Robust velocity computation from a biologically motivated model of motion perception", in *Proc. Roy. Soc.* Lon., vol. 266, 1999, pp. 509–518.

[27] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques", Int. *Journal of Computer Vision*, vol. 12, pp. 43–77, 1994.

[28] Mohammed Yeasin, Baptiste Bullot, and Rajeev Sharma, "Recognition of Facial Expressions and Measurement of Levels of Interest from Video," IEEE *Transactions on Multimedia*, Vol. 8, No. 3, June 2006.

[29] M.Black and Y.Yacoob, "Recognizing facial expressions in image sequences using local parameterized models of image motion," presented at the Int. Conf. Computer Vision, 1995.

[30] C. Padgett and G. W. Cottrell, *Representing Face Image for Emotion Classification*. Cambridge, MA : PUBLISHER, 1997.

[31] J. Cassell, C. Pelachaud, N. Badler, M. Steedman, B. Achorn, T. Becket, B. Douville, S. Prevost, and M. Stone, "Animate conversation: rule-based generation of facial expression, gesture & spoken into nation for multiple conversation agents", in *Siggraph'94*, Orlando, FL,1994.

[32] G. Cottrell and M. Dailey, The face of fear, in APA Talk, 2000.

[33] T. Darrell, I. Essa, and A. Pentland, "Correlation and interpolation networks for real-time expression analysis/synthesis," in *Advances in Neural Information Processing Systems*, pp. 909–916, 1995.

[34] I. Essa and A. Pentland, Facial expression recognition using a dynamic model and motion energy, presented at the Int. Conf. on Computer Vision, 1995.

[35] N. Oliver, A. Pentland, and F. Berard, LAFTER: lips and face real time tracker with facial expression recognition, presented at the IEEE Conf. Computer Vision and Pattern Recognition, 1997.

[36] Y.Azoz, L. Devi, M. Yeasin, and R. Sharma, "Tracking human amusing constraint fusion and multiple cue-based localization", *Machine Vision Application.*, vol. 13, pp. 286–302, 2003.

[37] A. Kapoor, S. Mota, and R. Picard, "Toward a learning companion that recognizes affect", presented at the Amer. Assoc. Artificial Intelligence, 2001.

[38] R. W. Picard and T. Kabir, "Finding similar patterns in large image databases", in *Proc. IEEE ICASSP*, Mineaopolis, MN, 1993, pp. 161–6