Human Emotion Recognition System Using Optimally Designed SVM With Different Facial Feature Extraction Techniques

G.U.KHARAT
Professor & Dean (Research & Development)
Anuradha Engineering College, Chikhli
(Maharashtra),
INDIA.
gukharat@rediffmail.com

DR. S.V.DUDUL
Professor and Head, Applied Electronics Deptt.
S.G.B. Amaravati University Amaravati (Maharashtra),
INDIA.
dudulsv@rediffmail.com

Abstract: - This research aims at developing "Humanoid 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. In this paper all six universally recognized basic emotions namely angry, disgust, fear, happy, sad and surprise along with neutral one are recognized. Various feature extraction techniques such as Discrete Cosine Transform (DCT), Fast Fourier Transform (FFT), Singular Value Decomposition (SVD) are used to extract the useful features for emotion recognition from facial expressions. Support Vector Machine (SVM) is used for emotion recognition using the extracted facial features and the performance of various feature extraction technique is compared. Authors achieved 100% recognition accuracy on training dataset and 94.29% on cross validation dataset.

Keywords: - Discrete Cosine Transform (DCT), Fast Fourier Transform (FFT), Singular Value Decomposition (SVD) Support Vector Machine (SVM), Machine Intelligence.

1. Introduction

It is highly expected that computers and robots will be used more for betterment of our daily life [1]. Information – Computerized Society expect a harmonious interaction or heart to heart communication between computers and / or robots and human beings [2,3]. 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 humanoid robot is to recognize human emotions. Mehrabian [4] indicates 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 [5]. 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 [6]. The problem of recognizing facial expressions had attracted the attention of computer- vision community [7-14]. Bassili [15] 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.

FACS is developed by Ekman and Frison [21] using action potentials for emotion recognition. Essa and Petland [16] and Essa [17] 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 [8, 10, 12, 14 and 18 to 31]. Anthropometric facial points are used for feature extraction to recognize emotions [12, 14 and 32].

This paper provides the simplest approach of using DCT, FFT and SVD for extraction of facial features and their performance comparison with optimally designed SVM.

2. Facial Expression Database

Facial expression database in six universally recognized basic emotions and neutral one is collected from Japanese female database. 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

YM.SA2.56 YM.HA2.53 YM.SU3.60

YM.FE4.70 YM.DI3.66 YM.AN2.62

Fig. 1: Images of Japanese females in various emotions

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 1. Total 210 images are selected for our experiment.

3. Computer Simulation Experiment.

3.1 Feature Extraction Using DCT:

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

3.2 Feature Extraction Using FFT:

A program is developed to obtain FFT and statistical parameters namely energy, entropy, variance, standard deviation, contrast, homogeneity and correlation of an image. An optimal feature vector is obtained containing the features extracted by FFT and statistical parameters of each image. Thus dataset for all 210 images is prepared to feed to Neural Network for emotion recognition.

3.3 Feature Extraction Using SVD:

Similarly a program is developed to obtain SVD and statistical parameters namely energy, entropy, variance, standard deviation, contrast, homogeneity and correlation of an image. An optimal feature vector is obtained containing the features extracted by SVD and statistical parameters of each image. Thus dataset for all 210 images is prepared to feed to Neural Network for emotion recognition.

4 Support Vector Machine

Machine learning algorithms receive input data during a training phase, build a model of the input and output a hypothesis function that can be used to predict future data. Given a set of labeled training examples

$$S = ((x_1, y_1), \ldots, (x_i, y_i)),\ y_i \quad \{-1, 1\}$$

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

$$h(x) = \mathrm{sgn}((w \cdot x) + b)$$

That yields a label $\{-1, 1\}$ (for the basic case of binary classification) for a previously unseen example x.

Support Vector Machines are based on results from statistical learning theory. These results establish that the generalization performance of a learned function on future unseen data depends on the complexity of the class of functions it is chosen from rather than the complexity of the function itself. By bounding this class complexity, theoretical guarantees about the generalization performance can be made. SVMs perform an implicit embedding of data into a high dimensional feature space, where linear algebra and geometry may be used to separate data that is only separable with nonlinear rules in input space. To do so, the learning algorithm is formulated to make use of kernel functions, allowing efficient computation of inner products directly in feature space, without need for explicit embedding. Given a nonlinear mapping $\mathbf{\Phi}$ that embeds input vectors into feature space, kernels have the form

$$K(x, z) = (\mathbf{\Phi}(x) \cdot \mathbf{\Phi}(z))$$

SVM algorithms separate the training data in feature space by a hyperplane defined by the type of kernel function used. They find the hyperplane of maximal margin, defined as the sum of the distances of the hyperplane from the nearest data point of each of the two classes. The size of the margin bounds the complexity of the hyperplane function and hence determines its generalization performance on unseen data. The SVM methodology learns nonlinear functions of the form:

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

Where the α_i are Lagrange multipliers of a dual optimization problem. It is possible to show that only some of the α_i are non-zero in the optimal solution, namely those arising from training points nearest to the hyperplane, called support vectors. These induce sparseness in the solution and give rise to efficient approaches to optimization. Once a decision function is obtained, classification of an unseen example x amounts to checking on what side of the hyperplane the example lies.

The SVM approach is highly modular, allowing domain specific selection of the kernel function used. They deal with noisy data and over fitting (where the learned function perfectly explains the training set but generalizes poorly) by allowing for some misclassifications on the training set. This handles data that is linearly inseparable even in higher space. Multi-class classification is accomplished by a cascade of binary classifiers together with a voting scheme. Their high classification accuracy for small training sets and their generalization performance on data that is highly variable and difficult to separate make SVMs particularly suitable to a real time approach to expression recognition in video. They perform well on data that is noisy due to pose variation, lighting, etc. and where often minute differences distinguish expressions corresponding to entirely different emotions.

5 SVM for Human Emotion Recognition

The scheme for emotion recognition system from facial expressions using different feature extraction techniques is shown in figure 2. Authors have used DCT, FFT and SVD for feature extractions and SVM for emotion recognition.

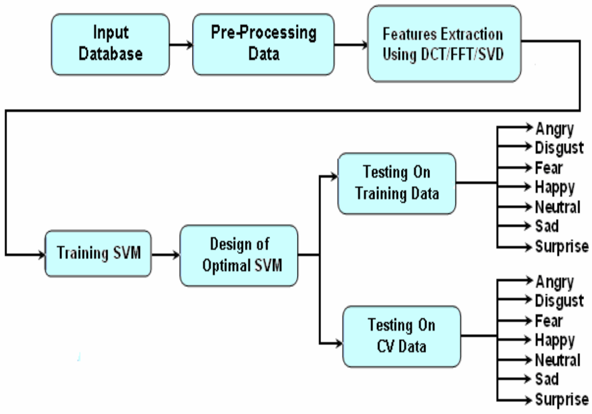

Fig.2: Scheme for Emotion Recognition system.

i) Human Emotion Recognition Using DCT:

The randomized data is fed to the SVM network. The network is trained three times by varying the number of exemplars for training and CV data. The average minimum MSE for train and CV data and percentage average classification accuracy is calculated and is shown in figure 3 & 4. The optimal results are obtained when 10% data is used for cross validation.

With 10% CV and 90% train data the step size for training the SVM is varied from 0.01 to 1.0 and each time network is trained and tested on training and CV dataset. The graph of average minimum MSE and % average classification accuracy is plotted against step size in figure 5 & 6 respectively. It is observed that optimal results are obtained for 0.5 step size. Optimally designed SVM is

Learning control	=	Supervised
Weight update	=	Batch
Step size	=	0.5
Number of epochs	=	1000
Number of runs	=	03
Termination of training	=	100 epochs without improve..

Time elapsed per epochs per exemplar = 0.564mSec.
Number of free parameters (P) of GFFNN = 14938
Number of exemplars in training dataset (N) = 189
(N/P) ratio = 0.0127

Finally designed SVM is tested on training and CV dataset and results are shown in table 1 to 4.

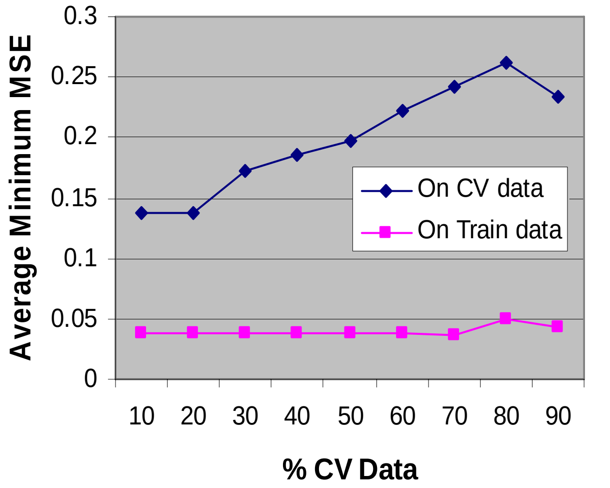

Fig. 3: Graph depicting variation of average minimum MSE on Training and CV dataset with % CV data.

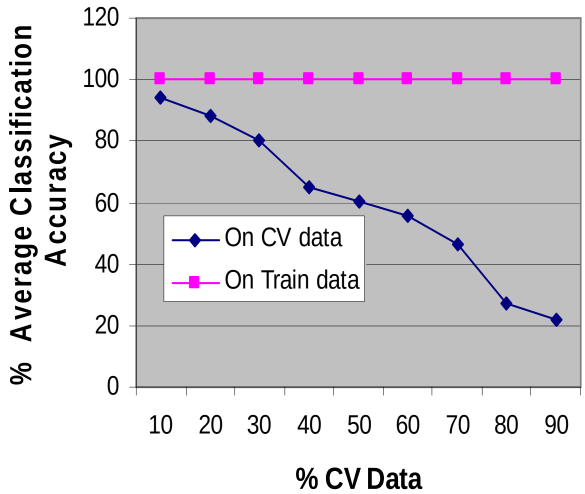

Fig. 4: Graph indicating variation of % average classification accuracy with % of CV data.

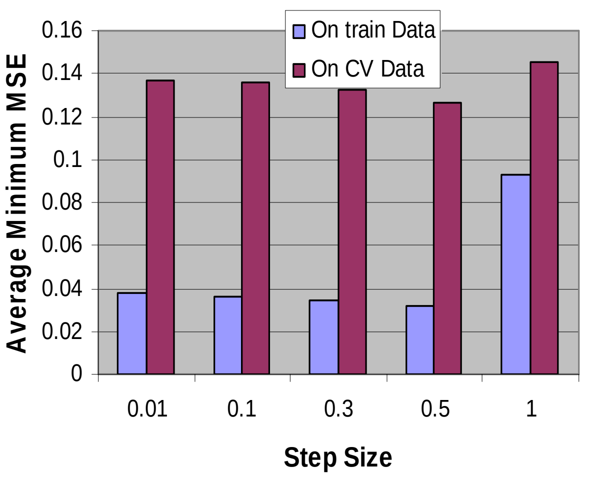

Fig. 5: Graph showing variation of average minimum MSE with Step size.

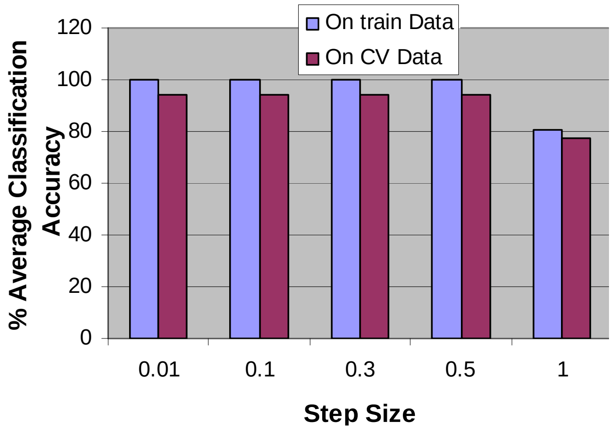

Fig. 6: Graph demonstrating variation of % average classification accuracy with Step size.

Table 1
Confusion Matrix for training data set using SVM

Output/ Desired	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	**28**	0	0	0	0	0	0
Disgust	0	**26**	0	0	0	0	0
Fear	0	0	**28**	0	0	0	0
Happy	0	0	0	**26**	0	0	0
Neutral	0	0	0	0	**29**	0	0
Sad	0	0	0	0	0	**27**	0
Surprise	0	0	0	0	0	0	**25**

Table 2
Performance parameters for training data sheet using SVM

Performance	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
MSE	0.00 457	0.00 494	0.00 514	0.00 542	0.00 512	0.00 598	0.00 536
NMS E	0.03 624	0.04 165	0.04 075	0.04 574	0.03 946	0.048 84	0.04 672
MAE	0.06 057	0.06 062	0.06 483	0.06 277	0.06 464	0.06 725	0.06 194
Min Abs. Error	0.00 138	0.00 025	0.00 157	0.00 359	0.00 011	0.00 013	0.00 219
Max Abs. Error	0.13 823	0.15 351	0.19 992	0.19 147	0.16 295	0.192 63	0.19 946
r	0.99 499	0.99 191	0.99 467	0.99 172	0.99 399	0.99 156	0.99 478

%Correct	100	100	100	100	100	100	100

The overall accurate recognition of emotion is = 100%

Table 3
Performance parameters for cross validation data sheet using SVM

Output/ Desired	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	**2**	0	0	0	0	0	0
Disgust	0	**4**	0	0	0	0	1
Fear	0	0	**2**	0	0	0	0
Happy	0	0	0	**4**	0	0	0
Neutral	0	0	0	0	**1**	0	1
Sad	0	0	0	0	0	**3**	0
Surprise	0	0	0	0	0	0	**3**

Table 4
Confusion Matrix for cross validation data set using SVM

Performance	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
MSE	0.02984	0.04580	0.04660	0.06540	0.05986	0.08515	0.09777
NMSE	0.34605	0.29705	0.54075	0.42412	1.31985	0.69538	0.53897
MAE	0.13638	0.16464	0.18095	0.19658	0.21184	0.21285	0.24894
Min Abs. Error	0.00158	0.00085	0.01385	0.00067	0.00054	0.00379	0.06507
Max Abs. Error	0.40028	0.48209	0.42234	0.57048	0.42576	0.83747	0.81586
r	0.87722	0.87419	0.82281	0.860489	0.57805	0.60562	0.77955
%Correct	100	100	100	100	100	100	60

Overall accurate recognition of emotion is = 94.29%

ii) Human Emotion Recognition Using FFT:

The randomized data is fed to the SVM network. The network is trained three times by varying the number of exemplars for training and CV data. The average minimum MSE for train and CV data and percentage average classification accuracy is calculated and is shown in figure 7 & 8. The optimal results are obtained when 10% data is used for cross validation.

With 10% CV and 90% train data the step size for training the SVM is varied from 0.01 to 1.0 and each time SVM is trained and tested on training and CV dataset. The graph of average minimum MSE and % average classification accuracy is plotted against step size in figure 9 & 10 respectively. It is observed that optimal results are obtained for 0.1 step size. Optimally designed SVM is

Learning control	=	Supervised
Weight update	=	Batch
Step size	=	0.1
Number of epochs	=	1000
Number of runs	=	03
Termination of training	=	100 epochs without improve.

Time elapsed per epochs per exemplar = 0.567ms
No. free parameters (P) of GFFNN = 14938
No. of exemplars in training dataset (N) = 189
(N/P) ratio = 0.0127

Finally designed SVM is tested on training and CV dataset and results are shown in table 5 to 8.

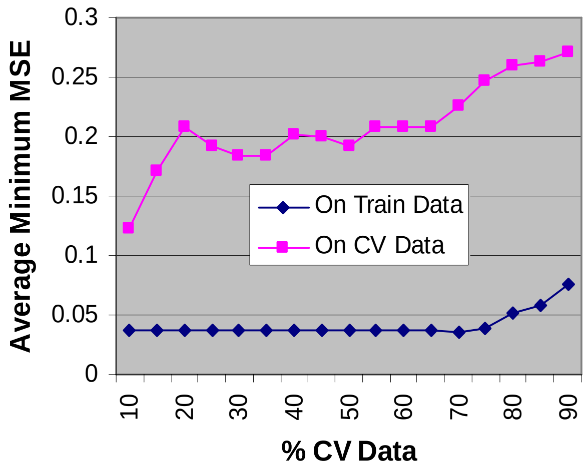

Fig. 7: Graph depicting variation of average minimum MSE on Training and CV dataset with % CV data.

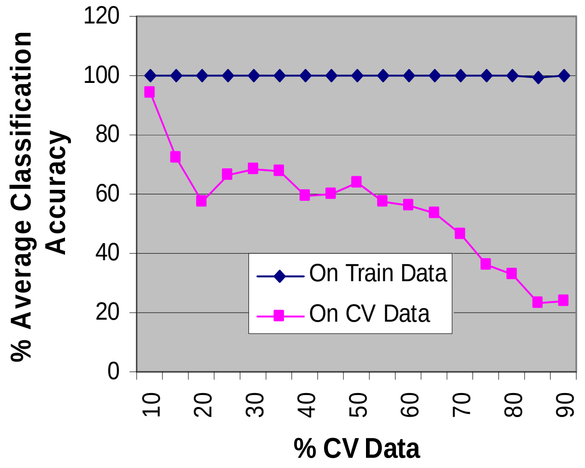

Fig. 8: Graph indicting variation of % average classification accuracy with % of CV data.

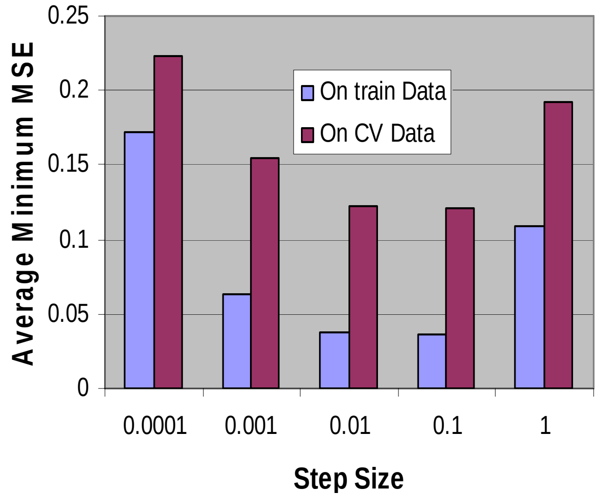

Fig. 9: Graph showing variation of average minimum MSE with Step size.

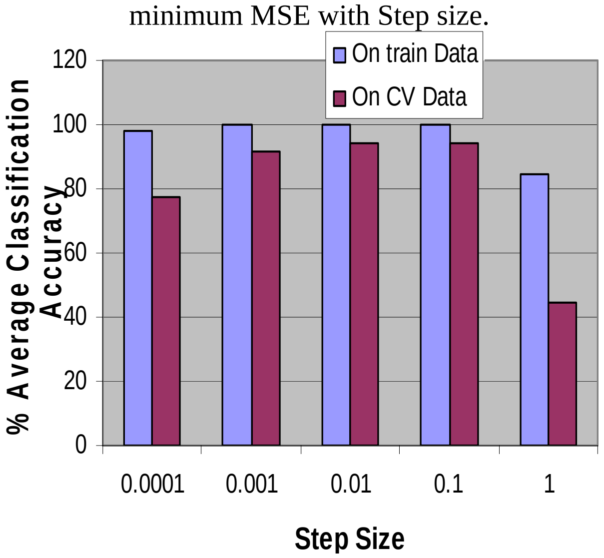

Fig. 10: Graph demonstrating variation of % average classification accuracy with Step size.

Table 5
Confusion Matrix for training data set using SVM

Output/ Desired	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	**27**	0	0	0	0	0	0
Disgust	0	**28**	0	0	0	0	0
Fear	0	0	**27**	0	0	0	0
Happy	0	0	0	**27**	0	0	0
Neutral	0	0	0	0	**27**	0	0
Sad	0	0	0	0	0	**25**	0
Surprise	0	0	0	0	0	0	**28**

Table 6
Performance parameters for training data sheet using SVM

Performance	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
MSE	0.00457	0.00494	0.00528	0.00692	0.00589	0.00743	0.00531
NMSE	0.03730	0.03911	0.04313	0.05655	0.04812	0.06472	0.04205
MAE	0.06125	0.06281	0.06471	0.07149	0.06776	0.07274	0.06632
Min Abs. Error	0.00116	0.00029	0.00319	0.00017	0.00093	0.00321	0.00547
Max Abs. Error	0.17522	0.16943	0.22347	0.26217	0.22085	0.25217	0.17040
r	0.99235	0.99234	0.99498	0.97740	0.99044	0.98668	0.994154
%correct	100	100	100	100	100	100	100

Overall accurate recognition of emotion is = 100%

Table 7
Performance parameters for cross validation data sheet using SVM

Output/ Desired	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	**3**	0	0	0	0	0	0

Disgust	0	**2**	0	0	0	0	0
Fear	0	0	**3**	0	0	1	0
Happy	0	0	0	**3**	0	0	0
Neutral	0	0	0	0	**3**	0	0
Sad	0	0	0	0	0	**3**	0
Surprise	0	0	0	0	0	1	**2**

iii) Human Emotion Recognition Using SVD

The randomized data is fed to the SVM network. The network is trained three times by varying the number of exemplars for training and CV data. The average minimum MSE for train and CV data and percentage average classification accuracy is calculated and is shown in figure 11 & 12. The optimal results are obtained when 10% data is used for cross validation.

Table 8
Confusion Matrix for cross validation data set using SVM

Performance	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
MSE	0.03984	0.04508	0.06365	0.03358	0.04937	0.07831	0.05033
NMSE	0.32537	0.52315	0.51980	0.27426	0.40317	0.43168	0.58404
MAE	0.12865	0.16048	0.19360	0.13538	0.15939	0.20551	0.18215
Min Abs. Error	0.00350	0.00058	0.01756	0.01029	0.00072	0.02360	0.01885
Max Abs. Error	0.53093	0.60236	0.62965	0.41561	0.58371	0.79588	0.50672
r	0.89761	0.74739	0.75228	0.88332	0.79074	0.81558	0.78396
%Correct	100	100	100	100	100	60	100

Overall accurate recognition of emotion is = 94.29%

With 10% CV and 90% train data the step size for training the SVM is varied from 0.01 to 1.0 and each time SVM is trained and tested on training and CV dataset. The graph of average minimum MSE and % average classification accuracy is plotted against step size in figure 13 & 14 respectively. It is observed that optimal results are obtained for 0.5 step size. Optimally designed SVM is

Learning control	=	Supervised
Weight update	=	Batch
Step size	=	0.5
Number of epochs	=	1000
Number of runs	=	03
Termination of training	=	100 epochs without improvement.

Time elapsed per epochs per exemplar = 0.495mSec.
Number of free parameters (P) of GFFNN = 14938
Number of exemplars in training dataset (N) = 189
(N/P) ratio = 0.0127

Finally designed SVM is tested on training and CV dataset and results are shown in table 9 to 12.

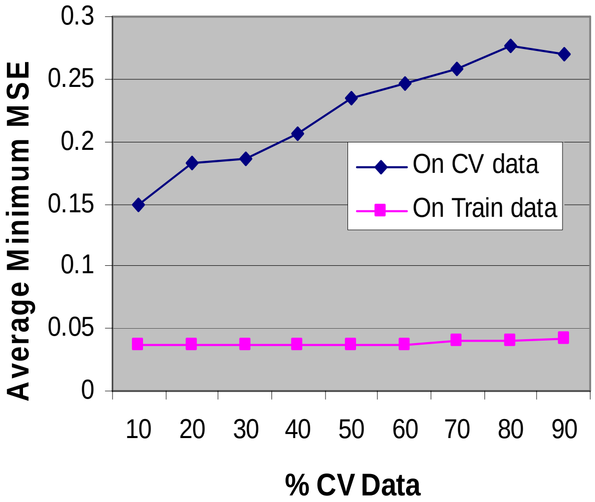

Fig. 11: Graph indicating variation of average minimum MSE on Training and CV dataset with % CV data.

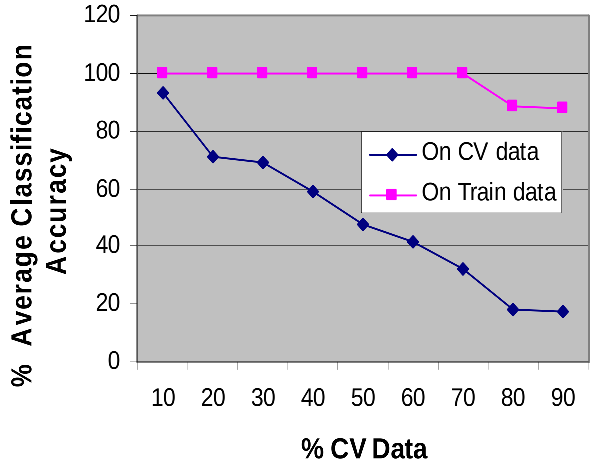

Fig. 12: Graph depicting variation of % average classification accuracy with % of CV data.

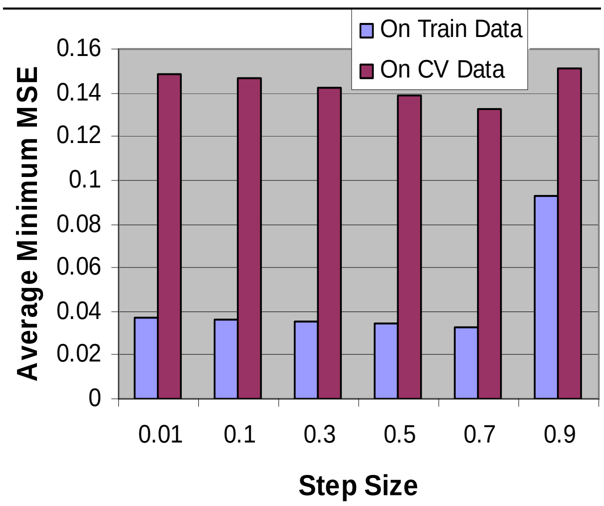

Fig. 13: Graph showing variation of average minimum MSE with Step size.

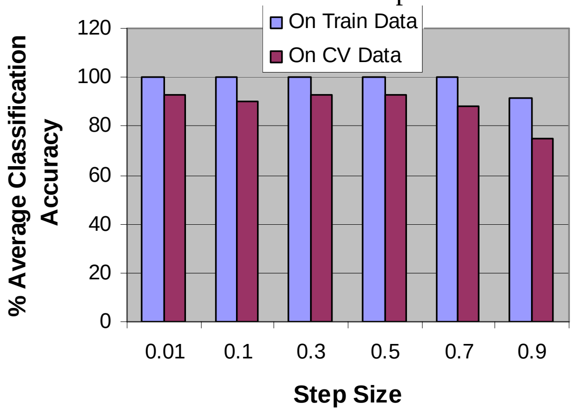

Fig.14: Graph demonstrating variation of % average classification accuracy with Step size.

Table 9
Confusion Matrix for training data set using SVM

Output/ Desired	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	**24**	0	0	0	0	0	0
Disgust	0	**26**	0	0	0	0	0
Fear	0	0	**29**	0	0	0	0
Happy	0	0	0	**27**	0	0	0
Neutral	0	0	0	0	**28**	0	0
Sad	0	0	0	0	0	**28**	0
Surprise	0	0	0	0	0	0	**27**

Table 10
Performance parameters for training data sheet using SVM

Performance	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
MSE	0.00	0.00	0.00	0.00	0.00	0.00	0.00

	712	721	717	756	636	717	733
NMSE	0.037 30	0.03 911	0.043 13	0.056 55	0.048 12	0.064 72	0.042 05
MAE	0.061 25	0.06 281	0.064 71	0.071 49	0.067 77	0.072 74	0.066 32
Min Abs. Error	0.001 16	0.00 029	0.003 19	0.000 17	0.000 93	0.003 21	0.005 48
Max Abs. Error	0.175 22	0.16 943	0.223 47	0.262 17	0.220 86	0.252 18	0.170 40
r	0.992 34	0.99 234	0.994 99	0.977 40	0.990 44	0.986 68	0.994 15
%Correct	100	100	100	100	100	100	100

Overall accurate recognition of emotion is = 100%

Table 11
Performance parameters for cross validation data sheet using SVM

Output/ Desired	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	**6**	0	0	0	0	0	0
Disgust	0	**2**	0	0	0	0	0
Fear	0	0	**1**	0	0	0	0
Happy	0	0	0	**3**	0	0	0
Neutral	0	0	0	0	**2**	0	0
Sad	0	1	0	0	0	**2**	0
Surprise	0	1	0	0	0	0	**3**

Table 12
Confusion Matrix for cross validation data set using SVM

Performance	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
MSE	0.06 545	0.11 597	0.04 042	0.07 887	0.04 667	0.05 351	0.07 241
NMSE	0.32 537	0.52 154	0.51 980	0.27 426	0.40 317	0.43 168	0.58 405
MAE	0.12 865	0.16 048	0.19 360	0.13 538	0.15 939	0.20 551	0.18 215
Min Abs. Error	0.00 350	0.00 058	0.01 756	0.01 029	0.00 073	0.02 361	0.01 885
Max Abs. Error	0.53 093	0.60 236	0.62 965	0.41 561	0.58 372	0.79 588	0.50 672

r	0.89761	0.74739	0.75228	0.88333	0.79075	0.81558	0.78396
% Correct	100	50	100	100	100	100	100

Overall accurate recognition of emotion is = 92.86%

6. Result

In this paper, the authors evaluated the performance of the three Feature Extraction Techniques namely DCT, FFT & SVD and compare their performance for the design of SVM to recognize human emotions.

Table 1 to 8 show emotion recognition results on training and testing data set for optimally designed SVM when DCT & FFT is used. The accuracy of recognition is 100% on train dataset and 94.29% on cross validation dataset for all the emotions namely angry, disgust, fear, happy, neutral, sad and surprise.

Table 9 to 12 show emotion recognition results on training and cross validation data set for optimally designed SVM when SVD is used. The accuracy of recognition is 100% on train dataset and 92.86% on cross validation dataset for all the emotions namely angry, disgust, fear, happy, neutral, sad and surprise.

6. Conclusion

The performance comparison for the various extraction methods is given below.

Neural Network Model	MSE		% Classification Accuracy		Time elapsed per epoch per exemplar (P4, 2.80 GB, 896MB RAM Computer)
	Train	CV	Train	CV	
DCT	0.0377	0.1370	100	94.29	0.564mS
FFT	0.0371	0.1221	100	94.29	0.567mS
SVD	0.0367	0.1484	100	92.86	0.495mS

It is observed that average minimum MSE on training dataset for SVM network is quite low for any of the methods for human emotion recognition and for all the methods average classification accuracy is 100%. Average minimum MSE on CV dataset is lower for FFT and higher for SVD. When average classification accuracy is calculated on CV data for SVM network, it is 94.29% when either DCT or FFT is used and 92.86% when SVD is used for emotion recognition.

When time elapsed per epoch per exemplar is calculated, it is lowest for SVD and highest for FFT, indicating that when SVD is used for feature extraction, the time required for training the SVM is lowest and when FFT is used for feature extraction time required for training the SVM is highest. The authors recommend DCT or FFT for the recognition of human emotions using Support Vector Machine (SVM).

References:

[1] Lotfi A. Zadeh, Towards Human Level Machine Intelligence, pp.15-16 procedding of 7th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering & Data Base (AIKED'08).

[2] L. Flores- pulido, O.Starostonko, D.flores-Quechol, Content Based Image Retrival Using Wavelets, pp 40-45, procedding of 2nd WSEAS International Conference on Computer Engineering & Applications (CEA'08).

[3] S. Chitwong, S.Witthayapradit, S.Intajag, Multispectral Image Classification Using Back Propogation Neural Network in Pca domain, pp 489-511, procedding of WSEAS International Multi Conference, Izmir, Turkey, Sep. 13-16,2004.

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

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

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

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

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

[9] 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.

[10] K. Mase, Recognition of Facial Expression from Optical Flow, *IEICE Trans., vol. E 74*, pp. 3,474-3483, 1991.

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

[12] 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.

[13] 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.
[14] 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.
[15] 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.
[16] 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.
[17] Analysis, Interpretation and Synthesis of Facial Expressions by Irfan Aziz Essa, Ph.D. thesis, MIT, February 1995.
[18] 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.
[19] 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.
[20] Mark Rosenblum, Yaser Yacoob, and Larry S. Davis, Human Expression Recognition from Motion Using; a Radial Basis Function Network Architecture. *IEEE Transactions On Neural Networks. Vol. 7, No. 5, September 1996*
[21] P. Ekman and W.Friesen, *The Facial Action Coding System.* San Francisco CA: Consulting Psychologists Press, 1978.
[22] 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.
[23] K. Mase and A. Pentland, Recognition of facial expression from optical flow, *IEICE Trans. E,* vol. 74, pp. 408–410, 1991.
[24] 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.
[25] B. Fasel and J. Luttin, Recognition of Asymmetric Facial Action Unit Activities and Intensities, IDIAP Research Rep., 1999.
[26] 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.
[27] M.Matsugu, K.Mori, Y.Mitari, and Y.Kaneda, Subject independent facial expression recognition with robust face detection using a convolutional neural network, *Neural Network.*, vol. 16, pp. 555–559, 2003.
[28] M.Gargesha and P.Kuchi, Facial expression recognition using artificial neural networks, Artificial *Neural Computing System*, pp. 1–6, 2002.
[29] 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.
[30] 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.
[31] 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.
[32] 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.