Internal Guide:
Prof. P. M. Jadav

EMOTION CLASSIFICATION FROM FACIAL EXPRESSION

Presented and Prepared By:
Janki K. Patel (14MTPOS014)
4th Sem M.Tech

Outline

- Introduction
- Proposed System
- Experimental results
- Conclusion
- Future work

Introduction

- In the area of Computer Vision & Artificial Intelligence facial expression classification is becoming one of the foremost challenges that can be used for Human–Computer Interaction (HCI).

- Ekman's [1] studies of human facial expression and showed that there are six basic emotions that are independent of the cultures and races of people, and these emotions have the same appearances in every human being.

➢ These emotions are: Anger, Disgust, Fear, Happiness, Sadness, and Surprise

Introduction

- Mehrabian [2] reported that facial expressions have an important effect on the person we are talking to;
- for facial expression about 55%,
- for vocal part 38%,
- for verbal part 7%.

- **Facial expression recognition:**

Given an image, the goal of facial expression detection is to locating faces, extracting facial features from the detected face, classifying this information into some facial expression-interpretative categories.

Introduction

- **<u>Applications</u>**

- ➢ Human-Computer-Interaction
- ➢ Driver Monitoring
- ➢ lie detection
- ➢ Patient Monitoring
- ➢ surveillance and security
- ➢ computer games

Introduction

- **Challenges**

➢ faces vary from one individual to another quite considerably due to different ethnicity & age,

➢ presence of several factors like facial hair, glasses etc.

➢ various size and orientation of the face in input images,

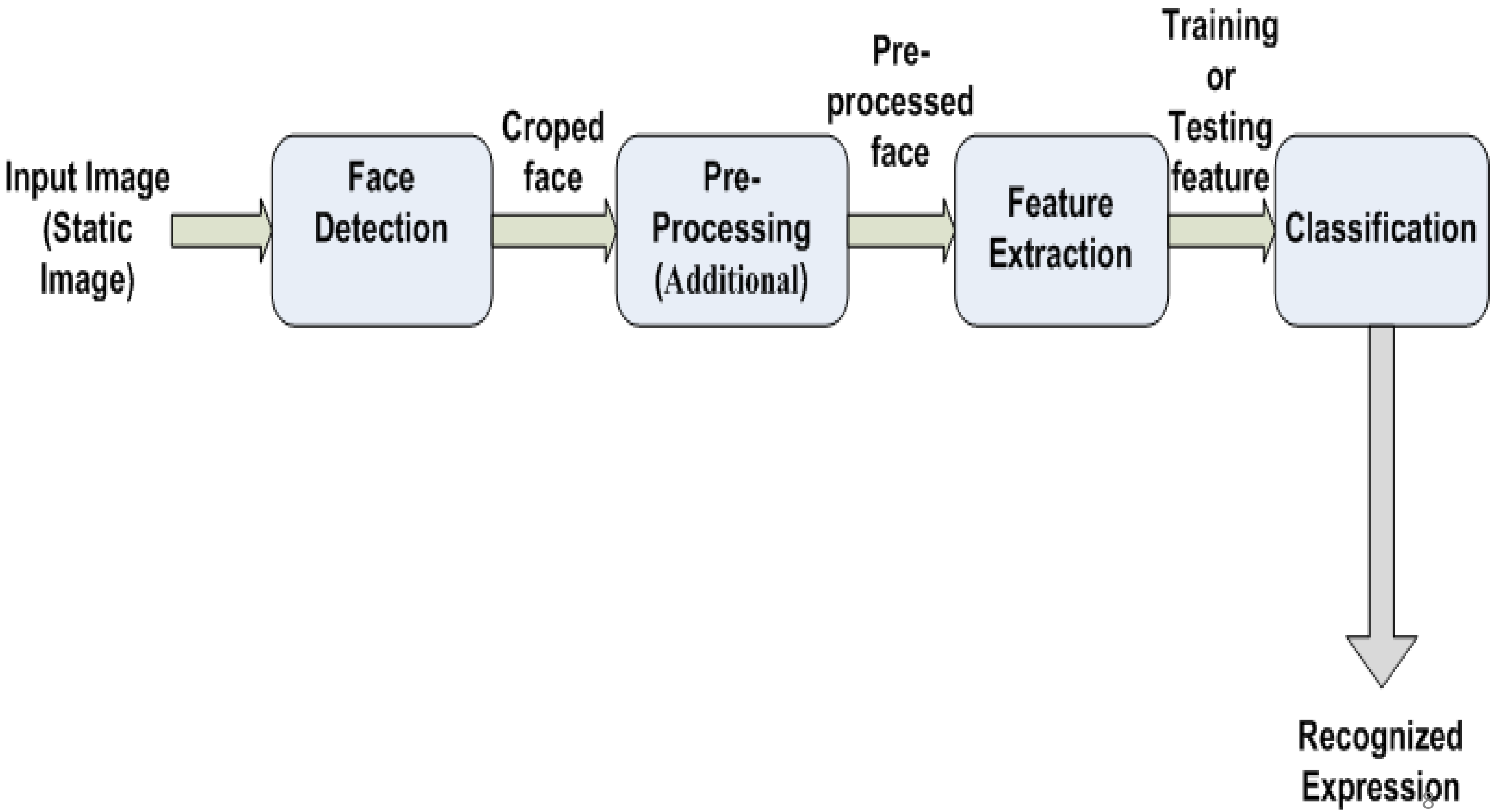
➢ non-uniform illumination facial point ,

➢ Pose variation.

Problem Definition & Objective

- Detection of facial expression from image irrespective of presence of noise, rotation and scaling.
- The system should be invariant to distraction as glasses, changes in hair style, facial hair, moustache, beard etc .
- It should detect expression in the least time.

Proposed System



Flow of Work

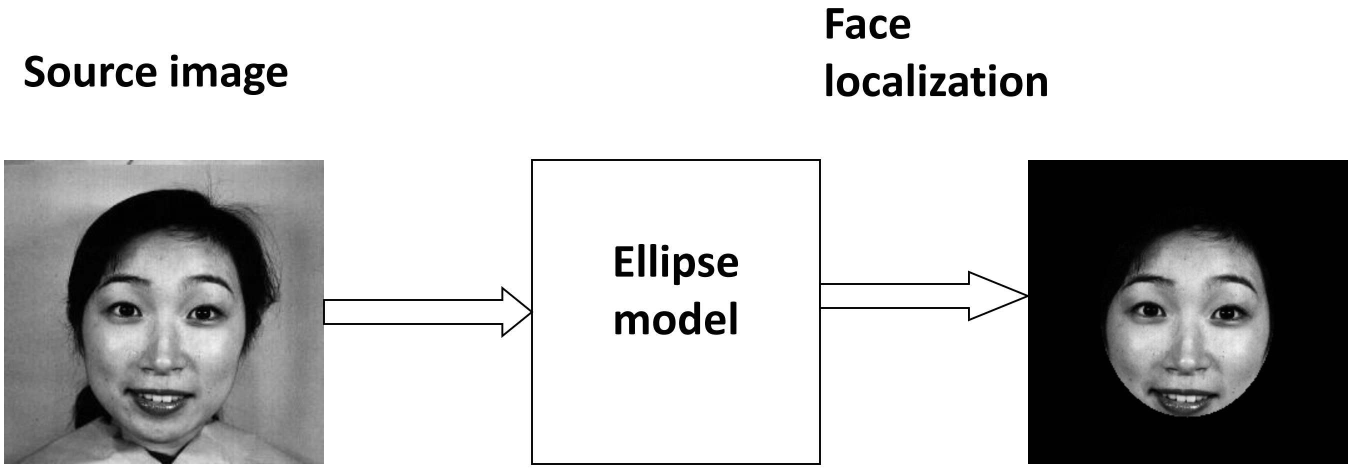
- Creation of image database
- Face localization and pre-processing
- Feature Extraction(PCA,ZM,PZM)
- Feature Vector Generation
- Classification of Images using NN
- Performance evaluation(Using Precision, Recall and Accuracy and Recognition rate)

Dataset detail

Sample of JAFFE dataset [7]

Sample of CMU AMP dataset [8]

Face detection & Pre-processing



- In order to detect the face region, an ellipse model with three parameters is used: ***X*0, *Y*0 are the centers** of the ellipse.

- The ellipse mask defined as:

$$\frac{(X-X0)^2}{a^2} + \frac{(Y-Y0)^2}{b^2} <= 1$$

where **X,Y are the coordinates** of any point on the ellipse, **a, b are the radius on the x and y axes** respectively.

Feature Extraction

- From literature survey we finalize below two methods,

  1) Zernike moments(ZM)

  2) Pseudo Zernike moments(PZM)

- ZM is a set of orthogonal polynomials on the unit disk.
- To compute Zernike moments of order **n** with repetition **l** of a function f (x, y) are defined as:

$$A_{nl} = \frac{n+1}{\pi} \sum_{x} \sum_{y} f(x,y)\, Z^{*}_{nl}\, (x,y) \qquad \textbf{(1)}$$

Feature Extraction

- where **circular Zernike polynomials in a unit circle** are defined as:

$$Z_{nl}(x,y) = Z(r,\theta) = Z(r\cos\theta, r\sin\theta) = R_{nl(r)}e^{il\theta}$$

- The real valued radial polynomials are given by,

$$R_{nl(r)} = \sum_{s=0}^{\frac{n-|l|}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|l|}{2}-s\right)!\left(\frac{n-|l|}{2}-s\right)!} r^{n-2s} \qquad \textbf{(2)}$$

where l=−∞, . . . ,−2,−1, 0, 1, 2, 3, . . . ,∞; the integer $n \geq 0, n \geq |l|$ and $n - |l|$ always even.

Feature Extraction

- Discrete approximation of the continuous Zernike integral based on Eq. (1) and Eq. (2) for image function I(i, j) with spatial dimension M × N written as follows:

$$A_{nl} = \frac{n+1}{\pi} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} I(i,j) R^{*}{}_{nl(r_{ij})} e^{-il\,\theta_{ij}}$$

- where the discrete polar coordinates:

$$r_{ij=}\sqrt{x^2{}_j + y^2{}_i} \qquad \theta_{ij} = \arctan\left(\frac{y_i}{x_j}\right)$$

- To mask the **pixels lying inside or on unit circle**, compute the unit disk with (x, y) as the center of the unit disc and **r is the polar value** and **θ is the polar coordinate**

Feature Extraction

- **Limitation of ZM:**

- number of feature elements produced by ZM is lesser due to condition of n-|l|=even,
- so ZM contains less information,
- Due to this it is not that much sufficient for recognition of any same expression(twins).

Feature Extraction

- **Pseudo-Zernike Moment (PZM):**

➢ Number of features in PZM is more than the ZM.

➢ PZM is more appropriate than ZM for recognition[10].

➢ Because in eq.(2) n-|l| = even condition is eliminated for **PZM**.

$$R_{nl(r)}=\sum_{k=0}^{\frac{n-|l|}{2}}(-1)^k\frac{(n-k)!}{k!\left(\frac{n+|l|}{2}-k\right)!\left(\frac{n-|l|}{2}-k\right)!}r^{n-2k}$$

Feature Calculation

Order	ZM feature elements	No. of FE ZM	PZM feature elements	No. of FE PZM
n=1,2,...6	n=1, l=1 n=2, l=0,2 n=3, l=1,3 n=4, l=0,2,4 n=5, l=1,3,5 n=6, l=0,2,4,6	**15**	n=1, l=0,1 n=2, l=0,1,2 n=3, l=0,1,2,3 n=4,l=0,1,2,3,4 n=5,l=0,1,2,3,4,5 n=6,l=0,1,2,3,4,5,6	**26**
n=6,7,8	n=6, l=0,2,4,6 n=7, l=1,3,5,7 n=8, l=0,2,4,6,8	**13**	n=6, l=0,1,2,3,4,5,6 n=7, l=0,1,2,3,4,5,6,7 n=8, l=0,1,2,3,4,5,6,7,8	**24**
n=9,10	n=9, l=1,3,5,7,9 n=10, l=,0,2,4,6,8,10	**11**	n=9, l=0,1,2,3,4,5,6,7,8,9 n=10,l =0,1,2,3,4,5,6,7,8,9,10	**21**

Feature Extraction

- PZM **allows the feature extractor to have a lower-dimensional vector** using below eq.

$$F_{vj} = \{PZM_{kl}\}, \qquad k = j, j+1, \ldots, N,$$

- number of **feature elements decrease** the **error rate is not changed ,**

- also the **higher orders of the PZMI contain more and useful information** for expression recognition process *[11].*

Classification

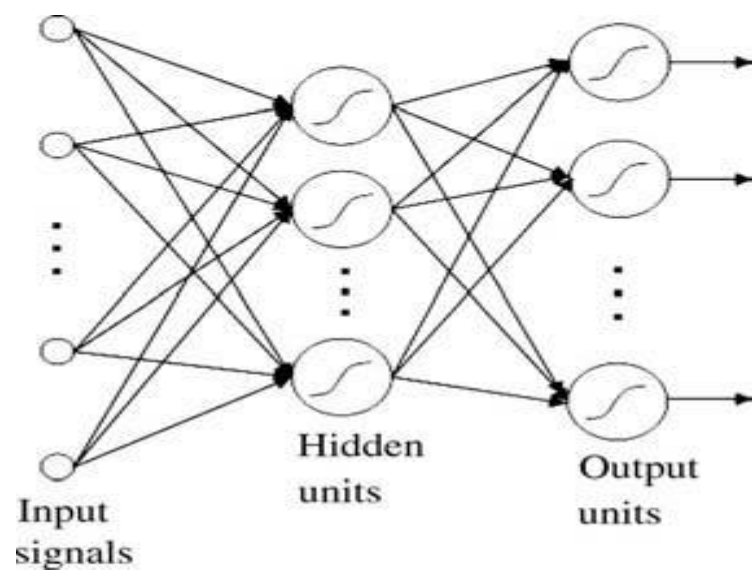
- Due to the generality, simplicity, and good learning ability of the neural networks we have used NN for Classification purpose.
- In NN we have used Patternnet: A two-layer feed-forward network

- ***<u>Property of Patternnet</u>* :**

✓ simplest model,

✓ pattern recognition,

✓ fast training speed



Patternnet: A two-layer feed-forward network(12)

Classification

- **Training Algorithm:**

  PatternNet required lowest number of iterations (in between 7 to 20) for the training process especially when it was trained using **TRAINLM algorithm**.

- **Patternnet** is a two-layered network,

1. hidden layer
2. output layer

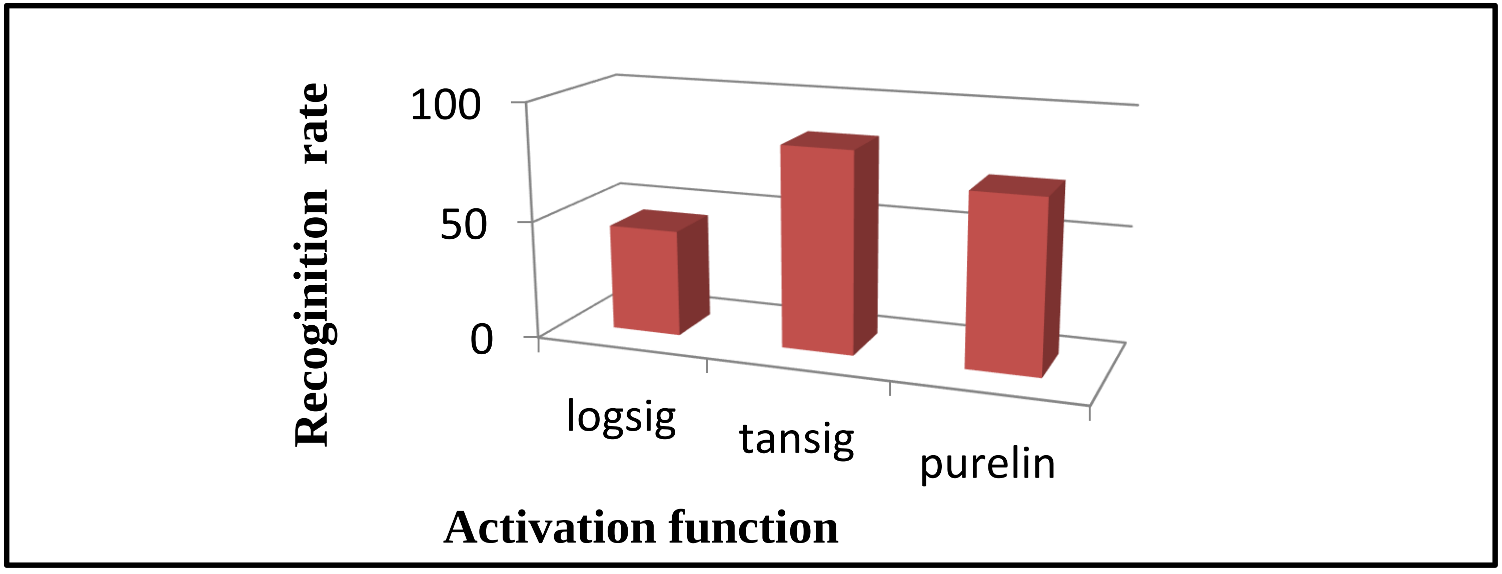
- ➢ At hidden layer, the input vector is transformed by **tansig** as **activation function.**
- ➢ output layer is a **competitive layer** as one of the expression is required to be identified at any point in time.

Experimental results

- **Initial Experiment**

➢ **Selection of Activation Function**

Activation Function	RR(%)
logsig	45%
tansig	**85%**
purelin	72.5%



Experimental results

- **Selection of Hidden Neurons**

Number of Hidden Neuron	RR(%) for Training	RR(%) for Testing
10	96.9	52.5
15	99.4	60
20	100	62.5
25	99.2	65
30	99.6	65
50	100	72
60	100	70.5
70	100	76
78	100	85
80	**100**	**90**

Experiment 1

- **Experiment Setup**

Dataset	**JAFFE [7]**	**CMU AMP [8]**
Training /Testing Ratio	60:40	80:20
Size of images	80×80	64×64
No. of Class	7	5
No. of Images	200	
Feature	PCA using eigen vector	
No. of Features	50	
Classifier	Neural network	

Experiment 1

- **JAFFE dataset**

Output/ Desired	Angry [10]	Disgust [11]	Fear [12]	Happy [12]	Neutral [10]	Sad [18]	Surprise [7]	Recall
Angry	**10**	0	0.25	0.25	0	1	0	**0.869**
Disgust	0	**11**	1	0	0	0	0	**0.916**
Fear	0	0	**9.25**	0	0	1.5	0	**0.860**
Happy	0	0	0.5	**11**	0	1	0	**0.888**
Neutral	0	0	0	0.25	**10**	0	0.5	**0.930**
Sad	0	0	0	0.5	0	**14.5**	0	**0.966**
Surprise	0	0	1	0	0	0	**6.5**	**0.866**
Precision	**1**	**1**	**0.771**	**0.916**	**1**	**0.856**	**0.928**	
Accuracy	**97.96**	**98.60**	**94.44**	**96.65**	**98.97**	**94.75**	**97.90**	

Experiment 1

- **CMU AMP dataset**

Output/ Desired	Angry [9]	Fear [6]	Happy [9]	Sad [6]	Surprise [10]	Recall
Angry	**8**	0	0	1	0	**0.833**
Fear	1	**5**	0	0	0	**0.833**
Happy	0	0	**9**	0	1	**0.9**
Sad	0	1	0	**5**	0	**0.833**
Surprise	0	0	0	0	**9**	**1**
Precision	**0.889**	**0.833**	**1**	**0.833**	**0.9**	
Accuracy	**94.73**	**94.73**	**97.29**	**94.73**	**97.29**	

Experiment 1

- **Observation Chart**

Experiment 1

➢Observation

PCA/NN	JAFEE dataset	CMU AMP dataset
emotion recognition varied	in the range of 85% to 91%	in the range of 87% to 90%
average classification rate	90.31%	87% to 90%
best training performance	0.00028593(MSE) at epoch 7	0.0008481(MSE) at epoch 13
highest recognition rate	disgust, angry, neutral	happy, surprise
lowest recognition rate	fear	fear and sad

➢Issue:

- Non-frontal view of the face (head movement).
- Computationally expensive and complex with the increase in data size.

Experiment 2

- **Experiment Setup**

Dataset	**JAFFE [7]**	**CMU AMP [8]**
Training /Testing Ratio	80:20	80:20
Size of images	80×80	64×64
No. of Class	7	5
No. of Images	200	
Feature	Zernike Moments	
No. of Features	21	
Classifier	Neural network	

Experiment 2

- **JAFFE dataset**

Output/ Desired	Angry [3]	Disgust [6]	Fear [10]	Happy [6]	Neutral [4]	Sad [9]	Surprise [2]	Recall
Angry	**2.25**	1.5	0	0.25	0	0.25	0	**0.542**
Disgust	0	**2.75**	0	0	0	0.25	0	**0.916**
Fear	0.25	0.25	**7.5**	0	0	2.25	0.75	**0.681**
Happy	0.5	0	0	**5.5**	0	0.25	0	**0.888**
Neutral	0	0.25	0.5	0	**4**	1	0.25	**0.666**
Sad	0	1.25	1.75	0	0	**5**	0	**0.625**
Surprise	0	0	0.25	0.25	0	0	**1**	**0.666**
Precision	**0.75**	**0.458**	**0.75**	**0.916**	**1**	**0.555**	**0.5**	
Accuracy	**91.05**	**88.88**	**82.35**	**95.7**	**93.33**	**80.00**	**94.91**	

Experiment 2

- **CMU AMP dataset**

Output/ Desired	Angry [8]	Fear [11]	Happy [10]	Sad [7]	Surprise [4]	Recall
Angry	**7**	1.33	0	1.66	0	**0.700**
Fear	1	**8.33**	0	0	0	**0.892**
Happy	0	0.33	**10**	0	0	**0.968**
Sad	0	1	0	**5.33**	0.33	**0.842**
Surprise	0	0	0	0	**3.66**	**1**
Precision	**0.875**	**0.757**	**1**	**0.761**	**0.916**	
Accuracy	**89.58**	**90.36**	**99.04**	**91.98**	**99.0**	

Experiment 2

- **Observation Chart**

Experiment 2

➢Observation

ZM/NN(9 & 10 order)	JAFEE dataset	CMU AMP dataset
emotion recognition varied	in the range of 65% to 77%.	in the range of 75% to 85%.
average classification rate	70%	85%
best training performance	0.00020983(MSE) at epoch 11	0.00069509(MSE) at epoch 12
highest recognition rate	neutral	happy
lowest recognition rate	disgust	fear

➢Issue:

- ✓ its condition p-|q| = even, are eliminated.
- ✓So extract less feature which contain less information.
- ✓it is not suitable for recognition of identical twins type problems.

Experiment 3

- **Experiment Setup**

Dataset	**JAFFE [7]**	**CMU AMP [8]**
Training /Testing Ratio	80:20	80:20
Size of images	80×80	64×64
No. of Class	7	5
No. of Images	200	
Feature	Pseudo Zernike Moments	
No. of Features	40	
Classifier	Neural network	

Experiment 3

- **JAFFE dataset**

Output/ Desired	Angry [8]	Disgust [4]	Fear [5]	Happy [9]	Neutral [6]	Sad [6]	Surprise [2]	Recall
Angry	**7**	1	0	0	0	0	0	**0.875**
Disgust	0.5	**2**	0.5	0	0	0.25	0	**0.615**
Fear	0	1	**4.25**	0	0	0.5	0	**0.739**
Happy	0	0	0	**8.75**	0	0	0	**1**
Neutral	0.25	0	0	0.25	**6**	1	0.75	**0.727**
Sad	0.25	0	0.25	0	0	**4.25**	0	**0.894**
Surprise	0	0	0	0	0	0	**1.25**	**1**
Precision	**0.875**	**0.500**	**0.850**	**0.972**	**1**	**0.708**	**0.625**	
Accuracy	**95.10**	**91.15**	**93.70**	**99.25**	**93.7**	**93.7**	**97.81**	

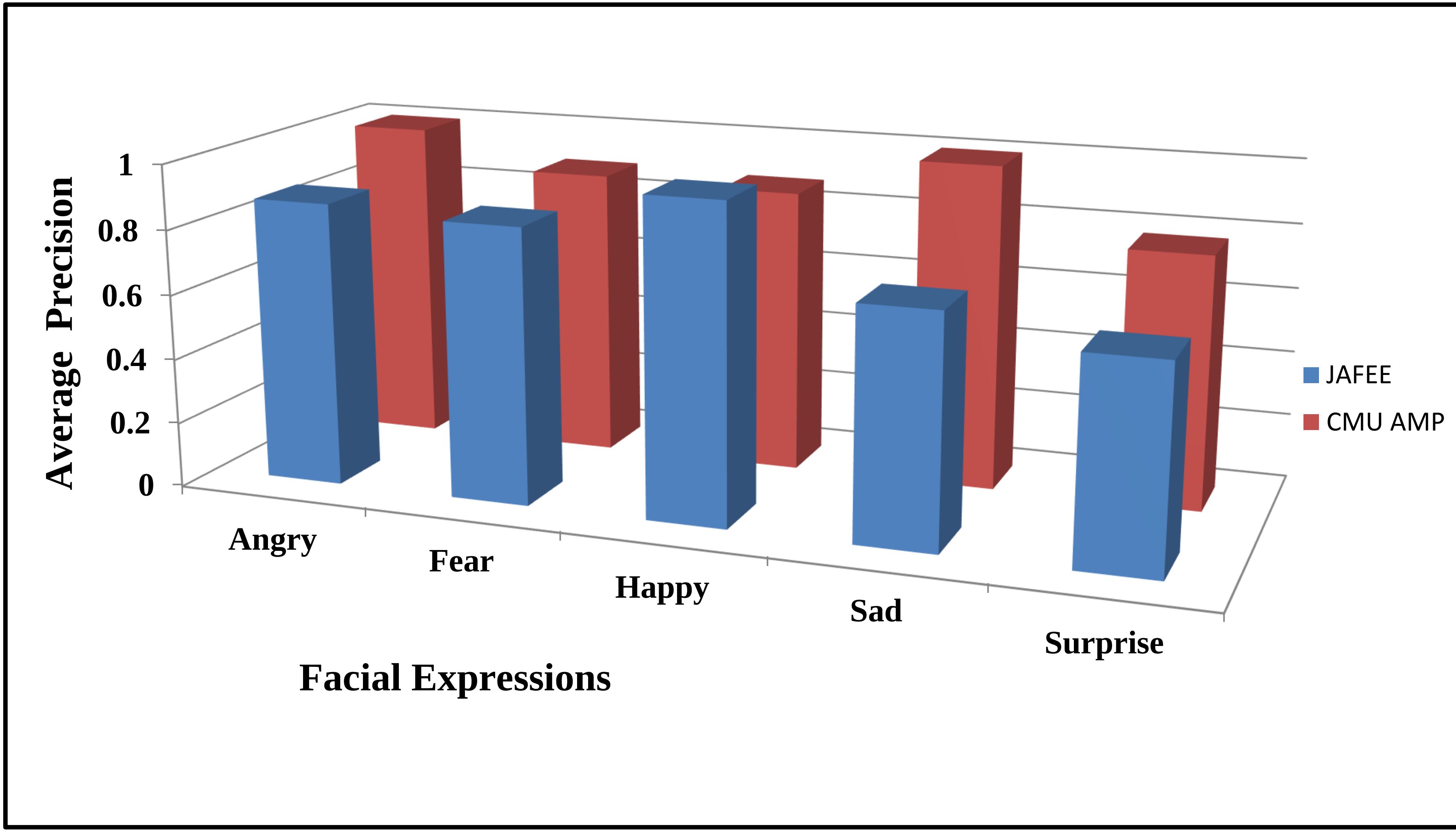
Experiment 3

- **CMU AMP dataset**

Output/ Desired	Angry [4]	Fear [9]	Happy [8]	Sad [7]	Surprise [12]	Recall
Angry	4	1	0	0	0.33	**0.750**
Fear	0	8	0	0	2	**0.857**
Happy	0	0	7	0	0.33	**0.954**
Sad	0	0	0	7	0	**1**
Surprise	0	0	1	0	9.33	**0.903**
Precision	**1**	**0.889**	**0.875**	**1**	**0.778**	
Accuracy	**96.37**	**92.17**	**96.37**	**100**	**90.61**	

Experiment 3

- **Observation Chart**



Experiment 3

➢Observation

PZM/NN(9 & 10 order)	JAFEE dataset	CMU AMP dataset
emotion recognition varied	in the range of 70% to 85%.	in the range of 80% to 90%.
average classification rate	<u>83.75%</u>	<u>88.32%</u>
best training performance	0.00028593(MSE) at epoch 7	0.00041909(MSE) at epoch 14
highest recognition rate	neutral	happy , surprise
lowest recognition rate	disgust	fear and sad

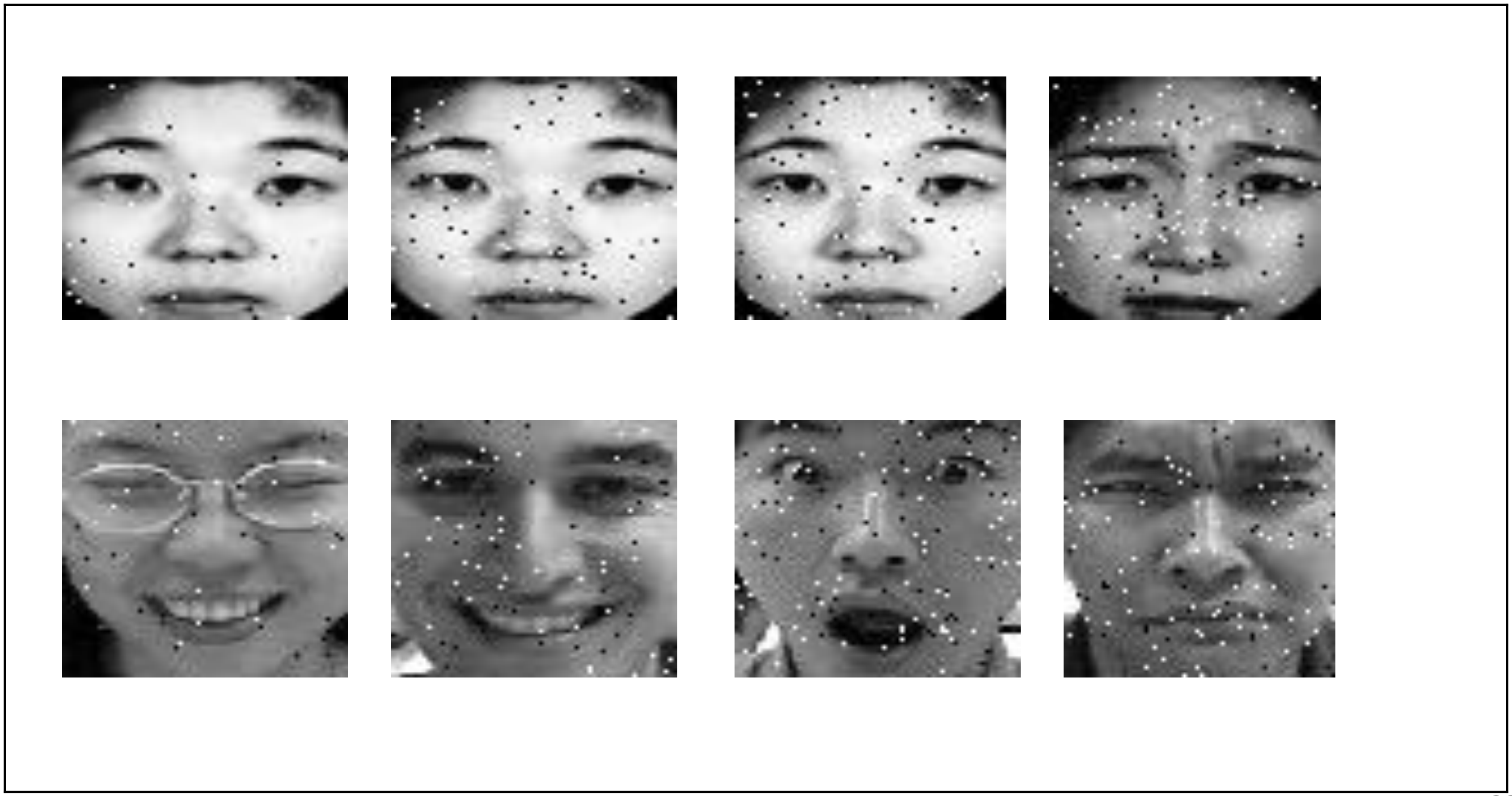
Experiment 4

- **Experiment Setup**

OBJECTIVE	See the affect of noise (salt and pepper)	
Dataset	**JAFFE [7]**	**CMU AMP [8]**
Training /Testing Ratio	80:20	80:20
Size of images	80×80	64×64
No. of Class	7	5
No. of Images	200	
Feature	Pseudo Zernike Moments	
No. of Features	40	
Classifier	Neural network	

Experiment 4

- **Facial expression with different impulsive noises**



Experiment 4

- **JAFFE dataset**

Noise density	Angry [4]	Disgust [5]	Fear [6]	Happy [6]	Neutral [4]	Sad [6]	Surprise [4]	Avg. accuracy	Error rate
0.01	91.07	94.44	92.72	86.44	86.44	89.47	92.72	**90.47**	**0.072**
0.02	90.19	86.79	85.18	83.63	85.18	82.14	97.87	**87.28**	**0.090**
0.03	95.45	97.67	76.36	75.0	79.24	79.24	91.30	**84.89**	**0.097**

- **CMU AMP dataset**

Noise density	Angry [9]	Fear [6]	Happy [9]	Sad [6]	Surprise [10]	Avg. accuracy	Error rate
0.01	91.04	82.43	91.04	89.7	91.04	**89.05**	**0.095**
0.02	89.2	85.2	86.5	84.0	89.2	**86.82**	**0.104**
0.03	83.3	73.5	83.3	83.3	80.6	**80.8**	**0.124**

Experiment 4

- **Observation**

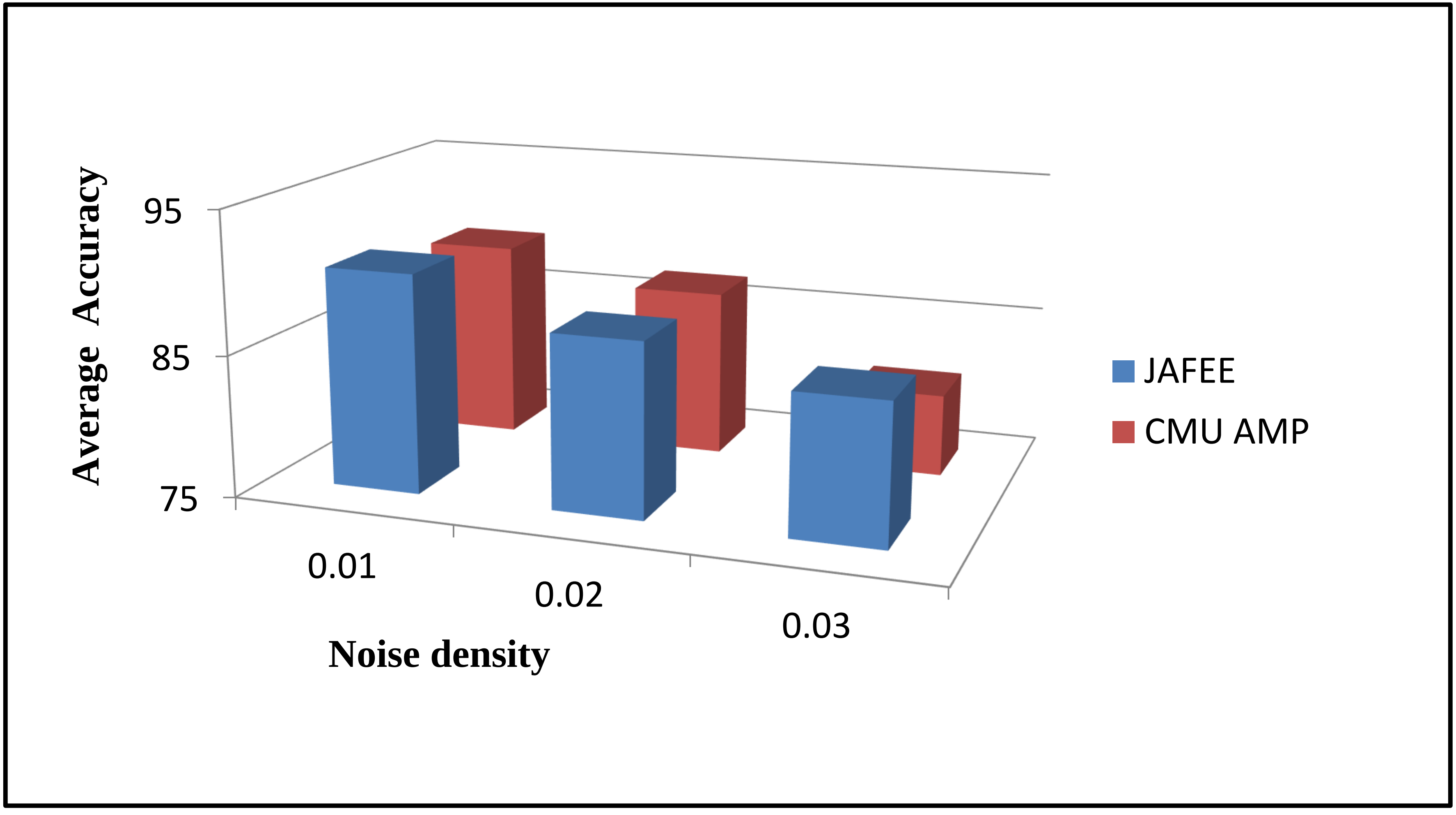
- ✓ In both dataset we observed the error rate,

$e = \frac{\text{number of misclassified}}{\text{Total number of samples}}$ for different noisy face images,

- ✓ and shows that the noise has less affect to the accuracy based on Pseudo Zernike moments.

Experiment 4

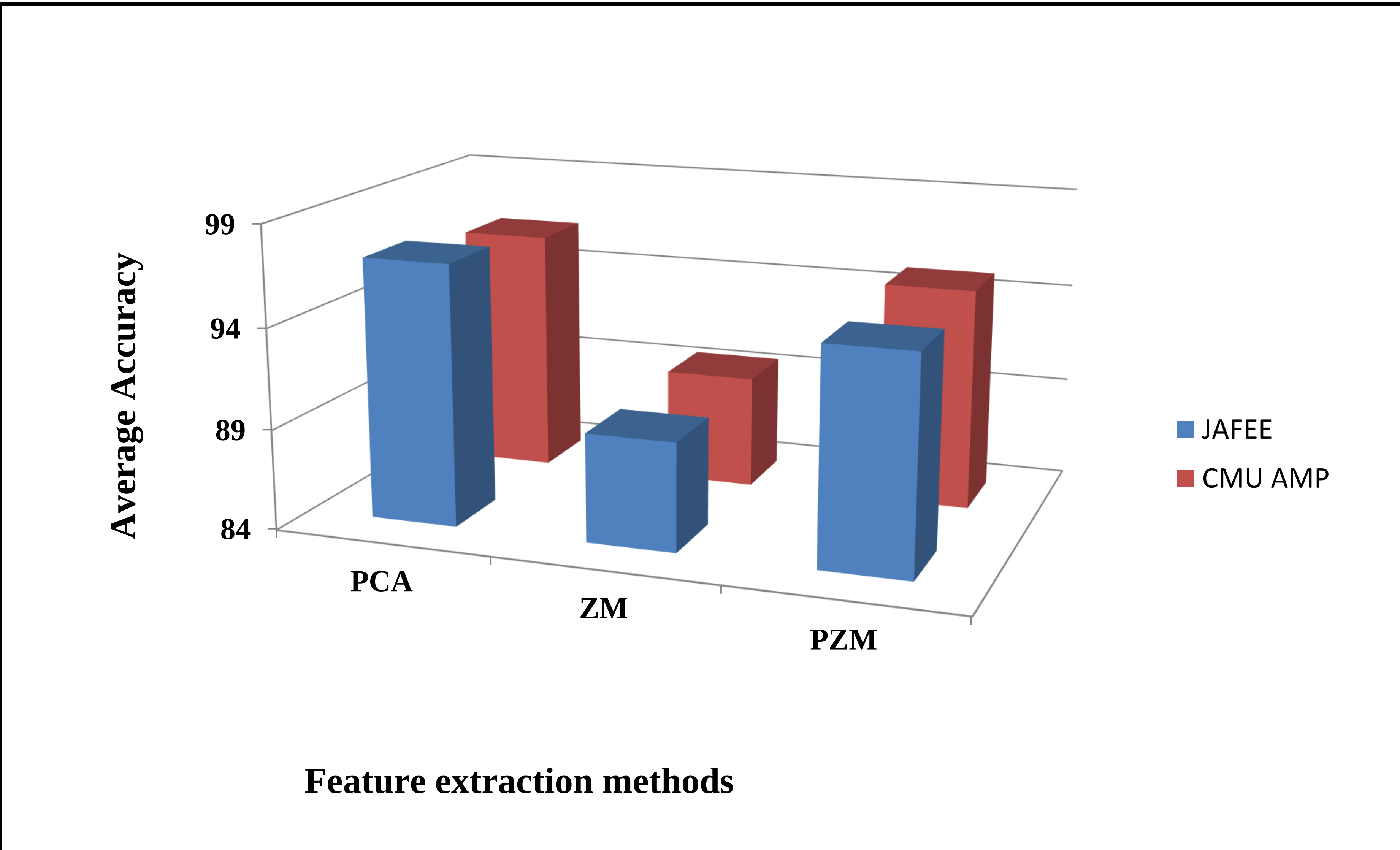
- **Observation Chart**



Summary of Experiment Result

Method	Dataset		RR (Avg.)	Precision (Avg.)	Recall (Avg.)	Accuracy (Avg.)
PCA	JAFFE		**90.31**	**91.74**	**89.84**	**97**
	CMU AMP		90	89.1	87.98	96.2
Zernike Moments (9 & 10 order)	JAFFE		**70**	**70.44**	**71.10**	**89.4**
	CMU AMP		85.82	86.22	88.05	89.62
Pseudo Zernike Moments (9 & 10 order)	JAFFE		**83.75**	**79.00**	**89.47**	**94.7**
	CMU AMP		88.32	90.834	89.32	95.06
Pseudo Zernike Moments with Noise	JAFFE	0.01	**72.86**	**74.52**	**77.09**	**90.47**
		0.02	**65.71**	**68.095**	**66.66**	**87.28**
		0.03	**60.0**	**61.31**	**60.93**	**84.89**
	CMU AMP	0.01	76.25	74.45	88.07	89.05
		0.02	72.5	71.76	73.56	86.82
		0.03	62.5	62.02	63.08	80.8

Observation Chart



Conclusion

- ✓ The performance of Pseudo Zernike Moment(PZM), Zernike Moment (ZM), Principal Component Analysis (PCA) and Neural Network (NN) in the facial expression classification system have been presented in this work.
- ✓ The result obtained using PCA is good, but to fulfill the rotation, noise, scale invariant properties we have implemented ZM and PZM.
- ✓ The combinations of PZM with Neural network contain more information about face image and improve the classification rate as compared to ZM.
- ✓ The average classification rate of **83.75%** *with JAFEE dataset* and **88.32%** *with CMU AMP dataset* are achieved using our system, which represents the overall performance of this facial expression classification system.
- ✓ According to the experimental results, the proposed PZM as feature extraction system is able to extract informative feature vector from input image and *also is robust to noise scaling and rotation*
- ✓ The presences of eyeglasses also have been handled by the system.

Future work

- ✓ In future work, accuracy for noisy image can be improved which we have presented in this work.
- ✓ Future work includes that the same technique can be used for detection of emotions in the presence of other noise such as speckle noise.
- ✓ Other neural network, which can be learning through optimization technique, can be used to improve overall significance of the system.

References

[1] Ekman, P, Friesen, “Constants across Cultures in the Face and Emotion”, J. Pers. Psycho. WV, 1971, vol. 17, no. 2, pp. 124-129.

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

[3] Ajit P. Gosavi, S. R. Khot,” Facial Expression Recognition Using Principal Component Analysis,” *International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3,* pp.258-262, 2013

[4] Bharati A. Dixit,” Statistical Moments Based Facial Expression Analysis,” *IEEE International Advance Computing Conference (IACC),* pp.552-557, 2015

[5] Maryam, Roja,” Local Weighted Pseudo Zernike Moments and Fuzzy classification for facial expression Recognition,” *IEEE 13th Iranian Conference on Fuzzy Systems (IFSC),* 2013

[6] Tran Binh Long, Le Hoang Thai, and Tran Hanh, “Facial Expression Classification Method Based on Pseudo Zernike Moment and Radial Basis Function Network,” International Journal of Machine Learning and Computing, pp.402-405, 2012

References

[7] http://www.kasrl.org/jaffe_download.html

[8] http://chenlab.ece.cornell.edu/_download/FaceAuthentication/faceExpressionDatabase.zip

[9] http://cvc.yale.edu/projects/yalefaces/yalefaces.html

[10] C.H. Teh and R.T. Chin, " On image analysis by the methods of moments," IEEE *Trans. Pattern Anal. Machine Intell, vol. 10*, pp.496-512, July 1988

[11] Javad Haddadnia, Majid Ahmadi, Karim Faez," An Efficient Feature ExtractionMethod with Pseudo-Zernike Moment in RBF Neural Network-Based Human Face Recognition System," *EURASIP Journal on Applied Signal Processing,* pp. 948–960, 2003

[12] Mohamed Cheriet "Character RecognitionSystems" *A John Wiley and Sons, Inc., Publication 1* edition 2007.

THANK YOU