

Pattern Averaging Technique for Facial Expression Recognition Using Support Vector Machines

N. P. Gopalan¹, Sivaiah Bellamkonda²

^{1,2}Department of Computer Applications, National Institute of Technology, Tiruchirappalli, India Email: npgopalan@nitt.edu, sivaiah.bk@gmail.com

Received: 24 May 2018; Accepted: 03 July 2018; Published: 08 September 2018

Abstract—Facial expression is one of the nonverbal communication methods of identifying an emotional state of a human being. Due to its crucial importance in Human-Robot interaction, facial expression recognition (FER) is in the limelight of recent research activities. Most of the studies consider the whole expression images in their analysis, and it has several has several drawbacks concerning illumination, orientation, texture, zoom level, time and space complexity. In this paper, a novel feature extraction technique called the pattern averaging is studied on whole image data using reduction in the dimension of the image by averaging the neighboring pixels. The study is found to give better results on standard datasets using support vector machine classifier.

Index Terms—Facial expression recognition, pattern averaging, support vector machine, human-computer interaction

I. INTRODUCTION

Robots are used to replace humans in performing repetitive and dangerous tasks which humans consider laborious and dangerous at times. Several applications exist in the human-machine interaction that involves facial expression recognition [1,2] as one can understand the emotion of a person just by watching the expression on the face. This kind of ability needs to be given to the robots to make the human-robot interaction to be much humanistic than mechanical. There are a good number of feature extraction, and classification studies are available in literature [3-12] and SVM appears to be a popular classifier for FER systems although Neural Networks [13-16], Hidden Markov Models [17,18] and KNNs [19,20] have also extensively used in similar such studies.

SVM is a supervised learning technique with associated learning algorithms and used in this paper for classifying FER. It uses linear algebra and geometry to separate input data into a high dimensional feature space through a selected nonlinear mapping function called kernel, and a learning algorithm formed for its usage. A specific radial basis function used as the kernel in this paper. Local directional positional patterns proposed by Zia Uddin et al. [21] for FER using principal component analysis (PCA), generalized discriminant analysis, and deep belief network (DBN) in depth sense-based video camera images. For each pixel in the picture, eight distinct directional edge reaction values have calculated, and an average recognition accuracy of 91.67% and 96.67% are achieved using hidden Markov model and DBN classifiers respectively in 3D facial expression image videos.

Siyue Xie et al. [22] proposed a feature redundancy reduced convolution neural network (FRR-CNN) based FER and obtained an average recognition rate of 83.96% on combined datasets of Cohn Kanade and JAFFE.

Tsai et al. [23] used angular radial transform, the discrete cosine transform and the Gabor filter simultaneously in the design of the feature extraction for FER on Cohn Kanade and JAFFE datasets with an average recognition rate of 97.10% using SVM classifier.

Features extracted from active facial patches [3] Local Binary Pattern (LBP) and PCA are used for feature extraction and dimensionality reduction respectively for FER with Softmax regression classifier for classifying the six basic expressions from Cohn Kanade facial expression dataset. It achieved a recognition rate of 96.3%. The recognition rates in a similar study using SVM classifier was observed to result in 87% and 77% on JAFFE and MUFE datasets respectively [5, 6].

In this paper, a new technique proposed for FER using SVM for classification, and on benchmark datasets, it is found to perform well in lesser training durations.

II. METHODOLOGY

This paper introduces a novel technique for FER called pattern averaging. The input images used are still images taken from standard facial expression datasets such as Cohn-Kanade, JAFFE, and MMI and preprocessed before applying pattern averaging. It includes detection of facial region and its cropping forming a separate image. Pattern Averaging is applied to the preprocessed input image as described in the next section. Fig. 1 shows the proposed method in detail.

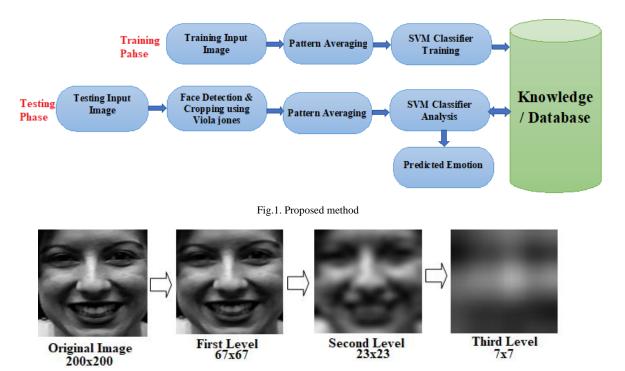


Fig.2. Images obtained after applying pattern averaging

Pattern averaging applies on each training input image for different block sizes and levels. A feature vector constructs by arranging each sample from training input data to a column vector.

A. Pattern Averaging

For an image of size 200x200, a total of 40000 data elements considered as a feature vector, which consumes time and space when considered a larger dataset of images requiring dimensionality reduction. Pattern averaging is applied to the entire image as a whole to reduce the dimension and construct feature vector.

Each grayscale image is a matrix of N x M size where N is the width and M is the height of the image. Pattern averaging is the process of averaging the pixel values of fixed sized blocks of the matrix. The block can be any square matrix of size ranging from 2x2, 3x3, and so on. These blocks averaged on their pixel values, and a new matrix constructed with these average values. The newly constructed matrix of the image is the reduced dimension version of the original image.

The average pixel calculated as follows for an image *x* with a block size of *nxn* as shown in equation 1:

$$p_{AVG}(k) = \frac{1}{n^2} \sum_{i=0}^{n} \sum_{j=0}^{n} x(i, j)$$
(1)

Where P_{avg} (k) is the new pixel value after averaging, *i* and *j* are variables to represent positions of width and height with in the block, and *n* is the value from block size *nxn*.

The following sample matrix with pixel values of an image shows pattern averaging by considering 3x3 block size:

[10	5	2	1	4	1]	[[10	5	2)	(1	4	1)]		
3	6	1	0	1	4	3	6	1	0	1	4		
4	6	1	0	2	1	4	6	1)	0	2	1)	[4	2]
1	4	9	1	7	0		4	9)	(1	7	0)	= 5	3
6	1	8	4	1	0	6	1	8	4	1	0		
8	3	1	6	8	1	$\begin{bmatrix} 10\\3\\4\\(1\\6\\8 \end{bmatrix}$	3	1)	6	8	1]]		

Figure 2 shows an image at different levels of pattern averaging by considering 3x3 blocks in the matrix.

The dimension of the image can be reduced further by repeating the same process. If the original image size is 200x200 pixels, after the first level of pattern averaging the reduced image size becomes 67x67 pixels which are 1/3rd of its original size. After the second level of pattern averaging over, the image size reduced to 23x23 pixels, which is 1/9th of original image size and after completion of the third level, its size will be of 7x7 pixels. Hence, if the total dimension of the original image is 40000 elements with 200x200 pixels, it is reduced to 49 elements with 7x7 pixels. The pattern averaging preserves important features of an image even after the dimension of the image reduced drastically and the newly constructed pixel values are the averages of the pixel values of the original image.

III. RESULTS AND DISCUSSION

A. Experimental Setup

The proposed model was implemented using MATLAB, and the testing conducted on benchmark datasets Cohn Kanade (CK) [24], Japanese Female Facial Expression (JAFFE) [25] and MMI [26] Facial

expression datasets. CK database is a benchmark dataset for FER. Images are saved with pixel sizes of 640x480 of 8-bits for grayscale values and resized to 200x200 pixels. Figure 3 shows sample images from the JAFFE dataset.



Fig.3. Sample expression images from Cohn-Kanade dataset

The JAFFE database contains 210 images of 10 persons, with seven different expressions. Each image saved with a resolution of 256x256. The original images are used without alterations (Cf. Fig 4).

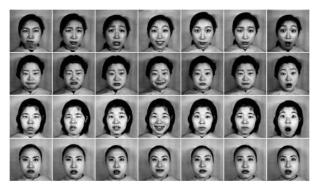


Fig.4. Sample expression images from JAFFE dataset

MMI Database contains images of 20 persons with 31 different expressions captured using FACS coding system for each of them. The images captured at a resolution of 1200x1900 pixels which resized to 200x200 pixels for experimentation purpose. (Cf. Fig. 5).

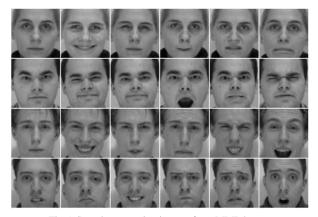


Fig.5. Sample expression images from MMI dataset

The accuracy of the system measured as a percentage

of correctly classified expressions. A confusion matrix with the list of emotions that are correctly classified or wrongly classified may also construct.

B. Pattern averaging for different block sizes

Pattern averaging applied on images for different block sizes. The resolution and the number of features vary with each block size. Table 1 shows various block sizes considered for the study along with the corresponding resolution and number of features obtained. The resolution and corresponding features of an image reduced for every level of pattern averaging. The original image of size 200x200 pixels with 40000 elements is reduced to 9801 at level 1 with block size 2x2 whereas, at level 3, it has only four features. As the number of features is very less, the recognition will be poor. It found that the recognition is fairly good when the number of features considered is above 400.

Table 1. No. of features obtained for corresponding block sizes

	Block S	ize		No. of	features	8
	2x2	3x3	4x4	2x2	3x3	4x4
Level 1	99x99	66x66	49x49	9801	4356	2401
Level 2	49x49	21x21	12x12	2401	441	144
Level 3	24x24	6x6	2x2	576	36	4

Figure 6 shows various resolutions of images obtained for different block sizes such as 2x2, 3x3, 4x4 at different levels of repetition such as level 1, level 2 and level 3.

Block2x2 Level1	Block3x3 Level1	Block4x4 Level1
)	(Y)	(Y)
Block2x2 Level2	Block3x3 Level2	Block4x4 Level2
*	(S)	130
Block2x2 Level3	Block3x3 Level3	Block4x4 Level3

Fig.6. Resolutions of different block size and levels

Figure 7 shows low-resolution images of various expressions, obtained after reduction of the dimension of 200x200 image to 21x21 resolution using a block of 3x3 at level 2. A total of 441 feature descriptors are generated using this method.



Fig.7. Low-resolution images at 21x21 using block 3x3 at level 2

C. Expression Specific Recognition Rate at Different Resolutions

I.J. Image, Graphics and Signal Processing, 2018, 9, 27-33

The confusion matrices obtained for a specific expression on different datasets shown in tables 2 - 6 for low-level resolution images of sizes 21x21, 24x24, 49x49,

66x66 and 99x99 pixels. The average accuracy of each method also listed in the table.

			JAF	FE			СК						ММІ					
	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU
AN	82	4	3	5	4	2	84	5	3	4	2	2	89	2	4	2	1	2
DI	5	83	5	1	4	2	5	84	5	1	3	2	3	88	4	1	1	3
FE	5	2	79	4	7	3	4	2	81	3	7	3	1	4	88	1	6	0
НА	3	2	6	85	1	3	3	2	4	88	1	2	2	1	2	92	2	1
SA	2	5	3	4	82	4	2	3	3	3	87	2	3	3	1	2	89	2
SU	3	3 4 4 1 2 86				86	2 4 4 1 0 89					89	2 2 1 2 1 92					
Avg.	82.83 %					85.50 %				89.66 %								

Table 2. Confusion matrices obtained using a total of 441 features

Table 3. Confusion matrices obtained using a total of 576 features

			JAF	FFE			СК						ММІ					
	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU
AN	82	4	3	5	4	2	85	5	2	4	2	2	89	2	4	2	1	2
DI	5						5 84 5 1 3 2					3	89	3	1	1	3	
FE	5	5 2 79 4 7 3					4 2 82 3 7 2					1	4	89	0	6	0	
НА	3	2	6	86	0	3	3	2	4	88	1	2	2	0	2	93	2	1
SA	2	5	3	3	83	4	2	3	3	3	87	2	3	3	1	2	89	2
SU	3	3 4 4 1 2 86				86	1 4 4 1 0 90				90	2	2	1	2	1	92	
Avg.	83.16 %					86.00 %					90.16 %							

Table 4. Confusion matrices obtained using a total of 2401 features

			JAF	FE			СК								М	MI	ММІ					
	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU				
AN	84	2	3	5	4	2	87	3	2	4	2	2	92	1	4	2	1	0				
DI	5						3 89 3 1 3 1					2	92	3	1	0	2					
FE	5	5 2 84 2 4 3					4 1 85 1 7 2				2	1	2	92	0	5	0					
НА	3	2	4	88	0	3	3	1	1	92	1	2	2	0	0	96	1	1				
SA	2	5	3	3	86	1	2	3	4	2	87	2	3	3	0	0	93	1				
SU	1	1 3 4 1 2 89				89	1 3 5 0 0 91				91	1 0 2 1 1 0 96										
Avg.	86.17 %					88.50 %						93.50 %										

Table 5. Confusion matrices obtained using a total of 4356 features

			JAF	FE					C	сĸ			MMI					
	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU
AN	86	3	2	3	4	2	91	1	3	1	3	1	92	1	4	2	1	0
DI	5	88	2	0	3	2	3	93	1	0	2	1	2	94	3	0	0	1
FE	5	2	85	1	4	3	2	1	89	2	5	1	1	2	92	0	5	0
НА	2	1	4	93	0	0	1	1	0	97	1	0	2	0	0	97	1	0
SA	1	4	3	3	88	1	2	3	4	0	89	2	3	3	0	0	93	1
SU	1	1 2 4 0 1 92				92	1	1	3	0	0	95	0	0	1	1	0	98
Avg.	g. 88.67 %					92.34 %					94.33 %							

			JAF	FE					C	K					М	MI		
	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU	AN	DI	FE	HA	SA	SU
AN	89	1	3	1	4	2	92	1	3	1	2	1	95	0	2	2	1	0
DI	3						3 94 1 0 1 1					0	97	3	0	0	0	
FE	5	2	85	1	4	3	1	1	90	2	5	1	1	1	94	0	4	0
HA	1	1	3	95	0	0	1	1	0	97	1	0	1	0	0	98	1	0
SA	1	4	3	3	88	1	2	2	4	0	91	1	3	2	0	0	94	1
SU	1	1 2 2 0 1 94				94	1 1 2 0 0 96					96	5 0 0 1 0 99					
Avg.	g. 90.17 %					93.34 %				96.16 %								

Table 6. Confusion matrices obtained using a total of 9801 features

It is clear from the confusion matrices that the emotions Surprise and Happy are found to give higher recognition accuracies as they have more distinguishable features compared with other emotions and the emotions Disgust and Angry are next in the similar study. The emotions Fear and Sad are confusing emotions with each other. It found that higher recognition rates recorded for a higher number of features. The summary of average accuracies for different resolutions of images listed in table 7. Figure 8 shows the comparative graph of table 7.

Table 7. Average-accuracies for different no. of features

Resolution	No. Of Features	Dataset							
Resolution	110. Of Features	JAFFE	CK	MMI					
21x21	441	82.83	85.50	89.66					
24x24	576	83.16	86.00	90.16					
49x49	2401	86.17	88.50	93.50					
66x66	4356	88.67	92.34	94.33					
99x99	9801	90.17	93.34	96.16					

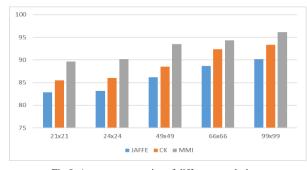


Fig.8. Average-accuracies of different resolutions

D. Performance comparison of expression recognition at different resolutions

The summary of the average performances listed in table 8. The results compared with the existing methods in the similar studies. Pattern averaging method at a resolution of 99x99 pixels can be seen to result in good recognition rates. The comparison of average accuracies obtained using pattern averaging method at a resolution of 99x99 pixels with the state of the art methods listed in table 8. The processing times of images with the corresponding block sizes for pattern averaging displayed in table 9.

Table 8. Comparison of recognition rates with similar studies

Method		Dataset	
Wieulou	JAFFE	СК	MMI
LBP + Softmax [3]	96.30	-	-
DWT + SVM [4]	96.67	-	-
PCA,LBP+SVM [6]	87.00	-	77.00
Gabor + ANN [9]	72.50	-	-
Proposed Pat-Avg +SVM	90.17	93.34	96.16

Table 9. Processin	ng time	for various	block	sizes an	d levels
--------------------	---------	-------------	-------	----------	----------

Level #	Block Size	Image Resolution	No. Of Elements	Processing Time in Sec.
2	3x3	21x21	441	0.033
3	2x2	24x24	576	0.026
2	2x2	49x49	2401	0.104
1	3x3	66x66	4356	0.574
1	2x2	99x99	9801	0.584

Figure 9 shows the comparative graphical analysis of the accuracies obtained by the proposed method with the similar studies done earlier. The proposed method can be seen to achieve competent recognition rates at lower resolutions also.

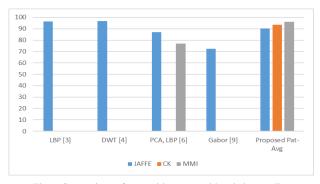


Fig.9. Comparison of recognition rates with existing studies

IV. CONCLUSION

In this paper, a novel pattern averaging technique is proposed for facial expression recognition and used for dimensionality reduction that drastically improves the processing speed of the system. SVM classifier has been found to improve the classification accuracy in comparison with other conventional classifiers found in the literature. The proposed model achieves an average accuracy of 90.17%, 93.34% and 96.16% on JAFFE, Cohn Kanade, and MMI datasets respectively at a low resolution of 99x99 pixels. The processing time is also considerably lower in the present analysis.

REFERENCES

- P. Ekman, and W. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movements", Consulting Psychologists Press, California, 1978.
- [2] Mehrabian.A, "Communication without Words", Psychology Today, 1968. Vol.2, No.4, pp 53-56.
- [3] Yanpeng Liu, Yuwen Cao, Yibin Li, Ming Liu, Rui Song, "Facial Expression Recognition with PCA and LBP Features Extracting from Active Facial Patches", *IEEE International Conference on Real-time Computing and Robotics June* 6-9, 2016, Angkor Wat, Cambodia, pp. 368-341, 2016.
- [4] Parth Patel, Khushali Raval, "Facial Expression Recognition Using DWT-PCA with SVM Classifier", International Journal for Scientific Research & Development, Vol. 3, Issue 03, pp. 1531-1537, 2015.
- [5] Yuan Luo, Cai-ming Wu, Yi Zhang, "Facial expression recognition based on fusion feature of PCA and LBP with SVM", Optik - *International Journal for Light and Electron Optics Volume* 124, Issue 17, pp. 2767-2770, 2013.
- [6] Muzammil Abdurrahman, Alaa Eleyan, "Facial expression recognition using Support Vector Machines", 23rd IEEE Conference on Signal Processing and Communications Applications Conference (SIU), 16-19 May 2015, Malatya, Turkey, 2015.
- [7] Anushree Basu, Aurobinda Routray, Suprosanna Shit, Alok Kanti Deb, "Human emotion recognition from facial thermal image based on fused statistical feature and multiclass SVM", 2015 Annual IEEE India Conference (INDICON), 17-20 Dec. 2015, New Delhi, India, 2015.
- [8] Mahesh Kumbhar, Manasi Patil, Ashish Jadhav, "Facial Expression Recognition using Gabor Wavelet", *International Journal of Computer Applications*, Volume 68, No.23, PP. 13-18, 2013.
- [9] Muzammil Abdulrahman, Tajuddeen R. Gwadabe, Fahad J. Abdu, Alaa Eleyan, "Gabor Wavelet Transform Based Facial Expression Recognition Using PCA and LBP", *IEEE 22nd Signal Processing and Communications Applications Conference* (SIU 2014), Trabzon, Turkey, 23-25 April 2014, pp. 2265 - 2268, 2014.
- [10] Shilpa Sharma, Kumud Sachdeva, "Face Recognition using PCA and SVM with Surf Technique", *International Journal of Computer Applications*, Volume 129, No.4, pp. 41-47, 2015.
- [11] Vinay A., Vinay S. Shekhar, K. N. Balasubramanya Murthy, S. Natarajan, "Face Recognition using Gabor Wavelet Features with PCA and KPCA - A Comparative Study, 3rd International Conference on Recent Trends in

Computing 2015 (ICRTC-2015), Procedia Computer Science 57, pp. 650–659, 2015.

- [12] Anurag De, Ashim Sahaa, Dr. M.C Pal, "A Human Facial Expression Recognition Model based on Eigen Face Approach", *International Conference on Advanced Computing Technologies and Applications* (ICACTA-2015), pp. 282-289, 2015.
- [13] Liangke Gui, Tadas Baltrusaitis, and Louis-Philippe Morency, "Curriculum Learning for Facial Expression Recognition", 12th *IEEE International Conference on Automatic Face & Gesture Recognition* (FG 2017), Washington, DC, USA, 2017
- [14] Zhiming Su, Jingying Chen, Haiqing Chen, "Dynamic facial expression recognition using autoregressive models", 7th International Congress on Image and Signal Processing (CISP), 14-16 Oct. 2014, Dalian, China, 2014.
- [15] Mao Xu, Wei Cheng, Qian Zhao, Li Ma, Fang Xu, "Facial Expression Recognition based on Transfer Learning from Deep Convolutional Networks", 11th *IEEE International Conference on Natural Computation (ICNC)*, 15-17 Aug. 2015, Zhangjiajie, China, 2015.
- [16] Pan Z., Polceanu M., Lisetti C., "On Constrained Local Model Feature Normalization for Facial Expression Recognition", in Traum D., Swartout W., Khooshabeh P., Kopp S., Scherer S., Leuski A. (eds) Intelligent Virtual Agents. IVA 2016, Lecture Notes in Computer Science, Vol. 10011. Springer, Cham, 2016.
- [17] Shaoping Zhu, "Pain Expression Recognition Based on pLSA Model", *The Scientific World Journal*, Volume 2014, 2014.
- [18] Myunghoon Suk and Balakrishnan Prabhakaran, "Realtime Mobile Facial Expression Recognition System – A Case Study", 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 23-28 June 2014, Columbus, OH, USA, 2014.
- [19] Zineb Elgarrai, Othmane El Meslouhi, Mustapha Kardouchi, Hakim Allali, "Robust facial expression recognition system based on hidden Markov models", *International Journal of Multimedia Information Retrieval*, Volume 5, Issue 4, pp. 229–236, 2016.
- [20] Guo Y., Zhao G., Pietikäinen M., "Dynamic Facial Expression Recognition Using Longitudinal Facial Expression Atlases", In Fitzgibbon A., Lazebnik S., Perona P., Sato Y., Schmid C. (eds) Computer Vision – ECCV 2012, Lecture Notes in Computer Science, vol. 7573. Springer, Berlin, Heidelberg, 2012.
- [21] M. Z. Uddin, M. M. Hassan, A. Almogren, A. Alamri, M. Alrubaian and G. Fortino, "Facial Expression Recognition Utilizing Local Direction-Based Robust Features and Deep Belief Network," in *IEEE Access*, vol. 5, pp. 4525-4536, 2017.
- [22] S. Xie and H. Hu, "Facial expression recognition with FRR-CNN," in *Electronics Letters*, vol. 53, no. 4, pp. 235-237, 2 16 2017.
- [23] Tsai, HH. & Chang, YC., "Facial expression recognition using a combination of multiple facial features and support vector machine", *Soft Comput* (2017).
- [24] Kanade, T., Cohn, J. F., Tian, Y., "Comprehensive database for facial expression analysis", *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition* (FG'00), Grenoble, France, pp. 46-53, 2010.
- [25] Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba, "Coding Facial Expressions with Gabor Wavelets", Proceedings of the third *IEEE International Conference on Automatic Face and Gesture Recognition*,

Nara Japan, IEEE Computer Society, pp. 200-205, 1998.

[26] M. Pantic, M.F. Valstar, R. Rademaker and L. Maat, "Web-based database for facial expression analysis", *Proceedings of IEEE International Conference on Multimedia and Expo* (ICME'05), Amsterdam, The Netherlands, July 2005.

Authors' Profiles



Dr. N. P. Gopalan is a professor in the department of computer applications, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India. He obtained his Ph.D. from Indian Institute of Science, Bangalore, India. His research interests are in Data Mining, Distributed Computing, Cellular Automata,

Theoretical Computer Science, Image Processing and Machine Intelligence.



B. Sivaiah is currently pursuing Ph.D. in the department of computer applications, National Institute of Technology, Tiruchirappalli. He obtained his B. Tech. in Computer Science & Engineering from Jawaharlal Nehru Technological University, Hyderabad, Andhra Pradesh, India in 2007 and M. Tech. in Computer

Science & Engineering from Jawaharlal Nehru Technological University, Kakinada, Andhra Pradesh, India in 2010. His research interests include image processing, neural networks, machine learning, and intelligent systems.

How to cite this paper: N. P. Gopalan, Sivaiah Bellamkonda, "Pattern Averaging Technique for Facial Expression Recognition Using Support Vector Machines", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.10, No.9, pp. 27-33, 2018.DOI: 10.5815/ijigsp.2018.09.04