



Trends in  
**Applied Sciences  
Research**

ISSN 1819-3579



Academic  
Journals Inc.

[www.academicjournals.com](http://www.academicjournals.com)

## Facial Expression Recognition using Local Arc Pattern

Mohammad Shahidul Islam and Surapong Auwatanamongkol

Department of Computer Science, School of Applied Statistics, National Institute of Development Administration, Bangkok, Thailand

*Corresponding Author: Mohammad Shahidul Islam, Department of Computer Science, School of Applied Statistics, National Institute of Development Administration, Bangkok, Thailand*

### ABSTRACT

The success of a good facial expression recognition system depends on the facial feature descriptor. Features extracted from local region are widely used for facial expression recognition due to their simplicity but the long feature vector length produces by them makes the overall system slow for recognition. This study presents a unique local facial feature descriptor, the Local Arc Pattern (LAP) for facial expression recognition. Feature is obtained from a local 5×5 pixels region by comparing the gray color intensity values surrounding the referenced pixel to formulate two separate binary patterns for the referenced pixel. Each face is divided into equal sized blocks and histograms of LAP codes from those blocks are concatenated to build the feature vector for classification. The recognition performance of proposed method was evaluated on popular Japanese Female Facial Expression dataset using Support Vector Machine as the classifier. Extensive experimental results with prototype expressions show that proposed feature descriptor outperforms several popular existing appearance-based feature descriptors in terms of classification accuracy.

**Key words:** Facial expression recognition, image processing, local feature extraction, pattern recognition, JAFFE, computer vision

### INTRODUCTION

Facial expression is the natural and immediate means of human interaction. According to Mehrabian (1968) facial expression is more meaningful than only verbal communication. Due to this reason, it catches the researchers' eye to build accurate and automatic facial expression recognition system. Many real life applications like human-computer-interaction, video indexing, driver state identification, pain assessment, patient condition observation, lie detection etc., demand more research for fast and accurate expression recognition systems. Some applications demand high accuracy and some demand real-time recognition. A facial expression recognition system works in four phases. (1) Detecting the face, (2) Extracting features related to facial expression from the face (3) Building a model for facial expression classification based on the extracted features and (4) Recognizing test images using the model. The vital part of a good facial expression recognition system is the second phase or feature extraction.

Mainly two types of facial feature extraction approaches are found (Tian *et al.*, 2003): Geometric-based approach that uses position, distance, angle and other relations between the facial components and appearance-based approach that uses texture or color combinations from the full or part of the image. Both the methods are equally popular in this field of research. In the geometric-based methods, it is necessary to find the exact location of the facial components

(Shan *et al.*, 2005, 2009). Most of the previous works on geometric-based methods were based on Facial Action Coding System where facial expressions were coded using one more Action Units (Ekman and Friesen, 1978). AUs were based on one more facial muscle movements. Kotsia and Pitas (2007) manually placed some of the Candide grid nodes to the face landmarks to create facial wire gframe model for facial expressions and used a Support Vector Machine (SVM) for classification. Valstar *et al.* (2005) and Valstar and Pantic (2006) used some fiducial points on the face to create geometric features and claimed that, “geometric approaches are better in feature extraction than appearance-based approaches”. Zhang and Ji (2005) proposed IR illumination camera for facial feature detection and tracking. To recognize the facial expressions they used Dynamic Bayesian networks (DBNs). They marked facial expressions by detecting 26 facial features around the regions of eyes, nose and mouth.

Besides geometric-based methods using AUs, some local appearance-based feature representations were also proposed. The local features are much easier for extraction than those of AUs are. Ahonen *et al.* (2006) suggested a new technique of facial feature representation for static images based on Local Binary Pattern (LBP). In this method, the LBP value at the referenced center pixel of a M×M pixel region is computed by thresholding the neighboring pixels gray color intensity value with the value of the center pixel as follows:

$$LBP = \sum_{i=1}^N 2^{i-1} f(g(i) - c) \quad (1)$$

Where:

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

‘C’ and g(i) denotes the gray color intensity value of the center pixel and i-th neighboring pixel, respectively, N stands for the number of neighbors, i.e., 8 or 16. The Local Binary Pattern was first proposed by Ojala *et al.* (1996). Ojala *et al.* (1996) used LBP for texture analysis and got a very good result. Since then it was used in many researches in many areas. Another popular holistic and appearance-based feature extraction method was Gabor Filter, named after Dennis Gabor. Facial future representations using Gabor filter is time and memory intensive (Bartlett *et al.*, 2005). Lajevardi and Hussain (2012) solved some limitations of Gabor-filter using log-Gabor filter but the dimensionality of resulting feature vector was still high. Local Phase quantization-LPQ was proposed by Ojansivu and Heikkila (2008) but like Gabor Filter, it was also very time and memory expensive. Ahsan *et al.* (2013) proposed a new feature extraction method from a local 5×5 pixels region and achieved very high accuracy but the cost of their method was very high.

Keeping all these sensitive issues in mind, this study presents a new feature extraction technique LAP (Local Arc Pattern) which overcomes almost all those cost problems and weaknesses mentioned above. It considers local 5×5 pixel region to compute two separate patterns which together represents the local pattern at the center pixel. Proposed method LAP is an extension of the method GDP (Gradient Direction Pattern) used in Islam and Auwatanamongkol (2013). Unlike Local Transitional Pattern (LTP)+Gabor Filter proposed by Ahsan *et al.* (2013), it considers almost all the gray color intensity values of pixels in 5×5 pixel region. The local pattern at a pixel identifies the changes in the gray color intensities of its neighboring in all possible directions.

**METHODS**

A 5×5 pixels local region is used to calculate LAP pattern for the center pixel of the region, ‘C’, as shown in Fig. 1. The gray color intensity values of a1, a2, a3, a4, b1, b2, b3, b4, c1, c2, c3 and c4 are used to formulate the LAP binary pattern. LAP pattern consists of one 4-bit binary pattern and one 8-bit binary pattern, say Pattern-1 (P1) and Pattern-2 (P2). P1 is computed using the gray color intensity values of a1, a2, a3, a4, a5, a6, a7 and a8 and, P2 is computed using the gray color intensity values of b1, b2, b3, b4, b5, b6, b7, b8, c1, c2, c3, c4, c5, c6, c7 and c8, as shown in Fig. 2. P1 can have at most  $2^4 = 16$  bit combinations and P2 can have at most  $2^8 = 256$  bit combinations. For each combination, a bin is created to count the number of occurrences of the combination within a given block. Sixteen bins for P1 and 256 bins for P2 are concatenated to build the LAP histogram for a block. Therefore, the feature vector length for the proposed method is 16+256 (272) per block. A detailed, (a) 4-bit binary pattern, a(1-8) represents the corresponding gray color value of the pixels as shown in Fig. 1(c) and (b) 8-bit binary pattern, b(1-8) and c(1-8) represents the corresponding gray color value of the pixels as shown in Fig. 1(d) example of obtaining LAP patterns from a 5x5 pixels region in shown in Fig. 3. Once histograms of all blocks in an image have been computed, they are concatenated to form a final feature vector of the image, as shown in Fig. 4.

**Feature dimensionality reduction using variance:** For the LAP representation, feature vector dimension is 272 per block before feature selection. Not all the features in this feature vector are necessary for the classification if they can not quite differentiate faces of different facial expression classes. The features having higher variances values would have higher power to differentiate

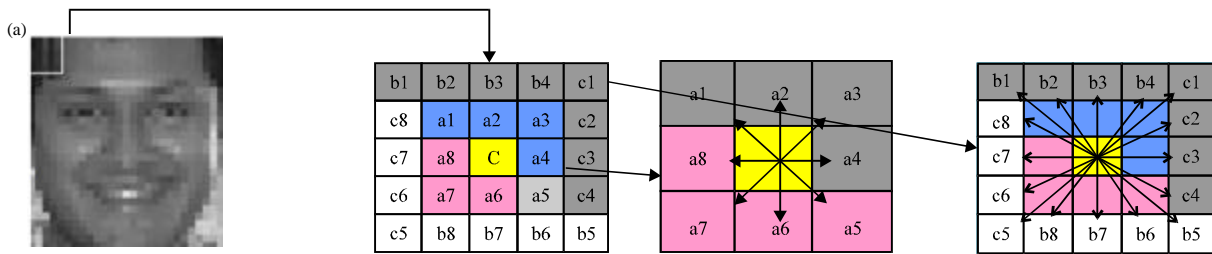


Fig. 1(a-d): Local pixels notation to formulate three different feature patterns for a single pixel e.g., “C”, (a) Facial image, (b) Local 5×5 pixels region, (c) Pixels used for pattern-1 and (d) Pixels used for pattern-2

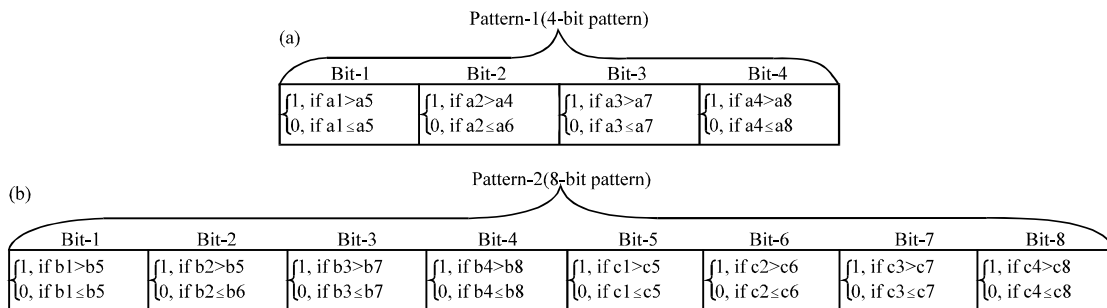


Fig. 2(a-b): Local Arc Pattern (LAP) formulation using the pixel’s gray color value of Fig. 1(b), (a) Pattern-1: 4-bit binary pattern and (b) Pattern-2: 8-bit binary pattern

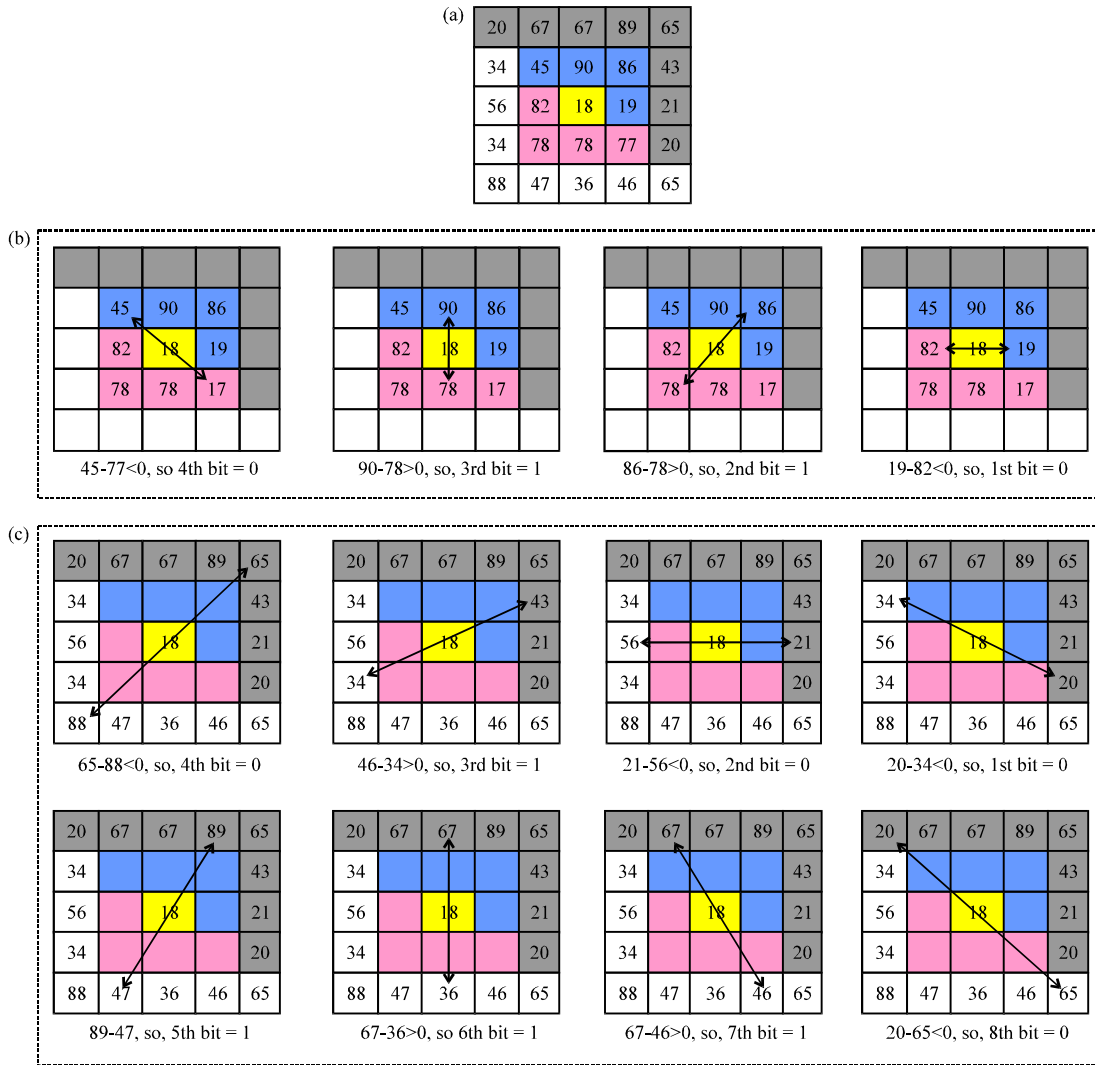


Fig. 3(a-c): Example of obtaining Local Arc Pattern (LAP) from a 5x5 pixels local region, here the referenced pixel is '18' and the Local Arc Pattern (LAP) code for that pixel is calculated as P1 = 0110 and P2 = 01110100 (a) 5x5 pixels local region, (b) 4-bit LAP code pattern 1 = 0110 and (c) 8-bit LAP code, pattern 2 = 01110100

faces of different facial expression classes than ones with lower variance values. So, variance values of the features can be used as indicators for feature selection. The variance value for each feature  $t$  of the feature vector can be calculated using Eq. 2:

$$VAR_t = \frac{1}{N} \sum_{j=1}^N (a_{jt} - \mu_N)^2 \quad (2)$$

where,  $a_{jt}$  denotes the value of the feature  $t$  of the  $j$ -th training sample,  $\mu_N$  represents the mean value of the feature  $t$ .  $N$  is the number of total training samples. The features are then sorted in descending order according to their variance values. The top  $M$  features with the highest variances values are then selected as the most contributing features to be used for the classification.

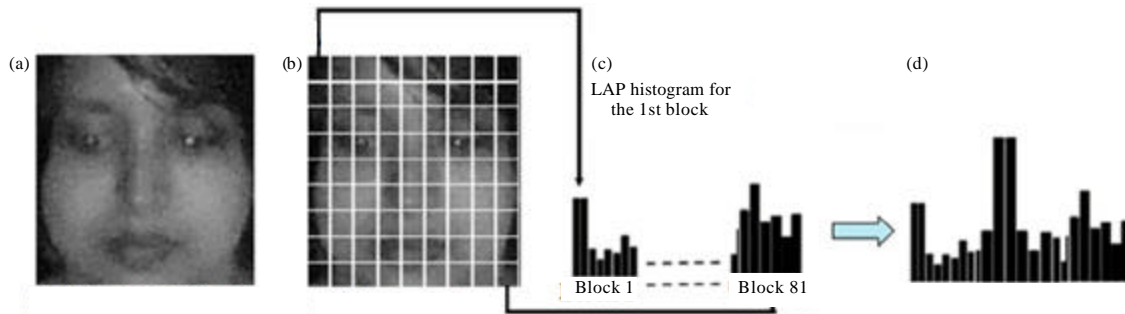


Fig. 4(a-d): Steps for facial feature extraction using proposed method. (a) Detected square sized facial region from a input gray color image (b) The face region is further divided into 81 square sized sub blocks (c) Local Arc Pattern (LAP) is applied to each pixel of each block and (d) Histogram of each block is concatenated to build the feature vector that uniquely represents that face

**Expression classifier:** For differentiating facial expression, the researchers used several classification techniques. Shan *et al.* (2005) did a comparative analysis on four machine learning techniques, namely Template Matching, Linear Discriminant Analysis (LDA), Linear Programming and Support Vector Machine. The author showed that SVM was the best in terms of classification accuracy. In this study, SVM is adopted as the classifier for facial expression.

## RESULTS AND DISCUSSION

In general, six or seven type of facial expressions are used to evaluate the facial expression recognition system (Shan *et al.*, 2005). The performance of the proposed local descriptor was evaluated on the well-known Japanese Female Facial Expression (JAFFE) Dataset (Lyons *et al.*, 1997). The dataset contains 213 images of 7 facial expressions (6 basic facial expressions +1 neutral) posed by 10 Japanese female models. The images in the dataset were taken at the Psychology Department in Kyushu University. An unpublished matlab code “fdlibmex” was used to detect the face from an image and re-dimension it to 99×99 pixels as a part of preprocessing. The image was then divided into 9×9 blocks. Each contains 11×11 pixels. No further alignment or no attempt was made to remove illumination changes. Linear, polynomial and Radial Basis Function (RBF) kernels were used in LIBSVM to classify the testing images. A ten-fold none overlapping cross validation was performed. The 90% of the images from each expression were used for training LIBSVM. The remaining 10% of the images were used for testing. For each fold, different 10% of the images were chosen for testing and it is user-dependent. Ten rounds of training and testing were performed and the average confusion matrix for proposed method was reported. The kernel parameters for the classifier were set to:  $s = 0$  for SVM type C-Svc,  $t = 0/1/2$  for linear, polynomial and RBF kernel, respectively,  $c = 100$  is the cost of the SVM,  $g = 1/(\text{length of feature vector dimension})$ ,  $b = 1$  is for probability estimation. This setting of LIBSVM was found to be suitable for JAFFE dataset with seven classes of data. The RBF kernel normally achieves slightly better recognition accuracy than linear or polynomial kernels (Chang and Lin, 2011). However, proposed

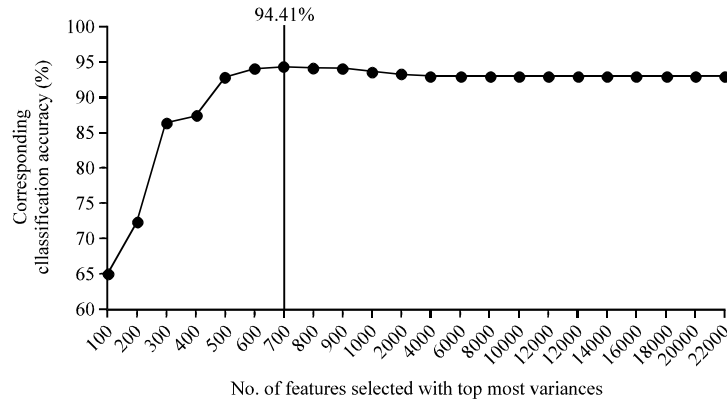


Fig. 5: No. of features selected with top most variances vs. corresponding classification accuracy (%), the highest accuracy of 94.41% is obtained at 700 features with top ranked variances

Table 1: Average confusion matrix for facial expression recognition system using proposed feature extraction technique Local Arc Pattern (LAP) (classification accuracy achieved 94.41%)

Prediction	Actual						
	Angry	Disgust	Fear	Happy	Neutral	Sad	Surprise
Angry	<b>100</b>	0.0	0.0	0.0	0.0	0.0	0.0
Disgust	0	<b>93.1</b>	6.9	0.0	0.0	0.0	0.0
Fear	0	3.1	<b>84.4</b>	0.0	3.1	6.3	3.1
Happy	0	0.0	0.0	<b>96.8</b>	0.0	3.2	0.0
Neutral	0	0.0	0.0	0.0	<b>100.0</b>	0.0	0.0
Sad	3.2	0.0	6.5	3.2	0.0	<b>87.1</b>	0.0
Surprise	0	0.0	0.0	3.3	0.0	0.0	<b>96.7</b>

system achieved better accuracy using polynomial kernel. The accuracy achieved using polynomial kernel is 94.41%. No fine-tuning of the “C” and “g” parameter of SVM had been performed. Further tuning may increase the RBF kernel performance substantially.

Figure 5 shows the plot between accuracy rate and the number of the top most contributing features selected for the classification. It is found that the number of selected features at 700 gives the highest accuracy of 94.41%. Therefore these 700 features are used to build the classification model and to validate the test images.

To get a better picture of the experimental results of individual facial expression types, the confusion matrix of 7-class expression recognition is given in Table 1. It is clear from the table that the expressions sad and fear are the most confusing in compare to the others. The results are compared with those of some previous works on JAFFE dataset, as shown in Table 2. All of the previous works are based on appearance methods. The feature extraction running time of all the methods cannot be directly comparable due to different experimental setup and execution environments.

Experiments were also performed on JAFFE dataset using other well known local feature methods, i.e. LBP, LBU<sub>v2</sub>. Table 3 compares the feature dimension feature extraction time and accuracy achieved from using the LAP and the other methods as the feature representations. It can be seen from the experimental results that the LAP outperforms the other methods on both accuracy and feature extraction time.

Table 2: Comparison of facial expression recognition accuracy of the proposed system and by some other recent works on JAFFE dataset

Author	Method	Classifier	Classification accuracy (%)
Proposed	Local Arc pattern	Multi class SVM (Poly)	94.41
Islam and Auwatanamongkol (2013)	Gradient direction pattern	Multi class SVM (Poly)	92.11
Subramanian <i>et al.</i> (2012)	Local binary pattern	SVM	88.09
Lyons <i>et al.</i> (1999)*	Gabor filter	LDA-based classification	92.00
Zhang <i>et al.</i> (1998)	Gabor filter	NN	90.10
Guo and Dyer (2003)	Gabor filter	Linear programming	91.00

NN: Neural network, LDA: Local discriminant analysis, SVM: Support vector machine

Table 3: Classification accuracy, feature dimension and feature extraction time (form a single image) comparisons between facial expression recognition systems that uses proposed local arc pattern and popular method local binary pattern as their feature extraction process

Parameter	Local arc pattern (LAP) (proposed)	Local binary pattern (LBP)	Uniform LBPU2
Classification accuracy (%)	94.410	91.140	90.120
Feature dimension	700.	20736.	4779.
Feature extraction time (sec)	0.035	0.068	0.065

## CONCLUSION

A novel feature representation for facial expression recognition is proposed in this study. Facial feature pattern at a pixel is extracted from pixels gray color intensity values of its neighboring pixels in 5x5 pixels region. To reduce the dimension of the feature vector, features are then selected based on their variance values. The experiments, conducted on JAFFE dataset, demonstrate the superiority of the proposed method over several other appearance-based methods. The proposed method successfully classifies nearly 95% of facial expressions accurately. Extensive experiments illustrate that the proposed method is more effective and takes less extraction time for facial expression recognition than the others. Future works include working with motion pictures where face registration is necessary.

## ACKNOWLEDGMENT

This study was supported by the 2012-2013 research fund of National Institute of Development Administration (NIDA), Bangkok.

## REFERENCES

- Ahonen, T., A. Hadid and M. Pietikainen, 2006. Face description with local binary patterns: Application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28: 2037-2041.
- Ahsan, T., T. Jabid and U.P. Chong, 2013. Facial expression recognition using local transitional pattern on gabor filtered facial images. *IETE Tech. Rev.*, 30: 47-52.
- Bartlett, M.S., G. Littlewort, M. Frank, C. Lainscek, I. Fasel and J. Movellan, 2005. Recognizing facial expression: Machine learning and application to spontaneous behavior. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Volume 2, June 20-25, 2005, San Diego, CA, USA., pp: 568-573.
- Chang, C.C. and C.J. Lin, 2011. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, Vol. 2.
- Ekman, P. and W. Friesen, 1978. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, CA., USA.



- Guo, G.D. and C.R. Dyer, 2003. Simultaneous feature selection and classifier training via linear programming: A case study for face expression recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 1, June 18-20, 2003, Madison, WI., USA., pp: 346-352.
- Islam, M.S. and S. Auwatanamongkol, 2013. Gradient direction pattern: A gray-scale invariant uniform local feature representation for facial expression recognition. *J. Applied Sci.*, 13: 837-845.
- Kotsia, I. and I. Pitas, 2007. Facial expression recognition in image sequences using geometric deformation features and support vector machines. *IEEE Trans. Image Process.*, 16: 172-187.
- Lajevardi, S.M. and Z.M. Hussain, 2012. Feature extraction for facial expression recognition based on hybrid face regions. *Adv. Electr. Comput. Eng.*, 9: 63-67.
- Lyons, M., M. Kamachi and J. Gyoba, 1997. Japanese Female Facial Expressions (JAFFE). Database of Digital Images. [http://www.kasrl.org/jaffe\\_info.html](http://www.kasrl.org/jaffe_info.html).
- Lyons, M.J., J. Budynek and S. Akamatsu, 1999. Automatic classification of single facial images. *Trans. Pattern Anal. Mach. Intell.*, 21: 1357-1362.
- Mehrabian, A., 1968. Communication without Words. *Psychol. Today*, 2: 53-56.
- Ojala, T., M. Pietikainen and D. Harwood, 1996. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.*, 29: 51-59.
- Ojansivu, V. and J. Heikkila, 2008. Blur insensitive texture classification using local phase quantization. Proceedings of the 3rd International Conference on Image and Signal Processing, July 1-3, 2008, Cherbourg-Octeville, France, pp: 236-243.
- Shan, C., S. Gong and P.W. McOwan, 2005. Robust facial expression recognition using local binary patterns. *Image Process.*, 2: 370-373.
- Shan, C., S. Gong and P.W. McOwan, 2009. Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vision Comput.*, 27: 803-816.
- Subramanian, K., S. Suresh and R.V. Babu, 2012. Meta-cognitive neuro-fuzzy inference system for human emotion Proceedings of the International Joint Conference on Neural Networks, June 10-15, 2012, Brisbane, QLD, pp: 1-7.
- Tian, Y.L., T. Kanade and J.F. Cohn, 2003. Facial Expression Analysis. Handbook of Face Recognition, Springer, USA..
- Valstar, M.F., I. Patras and M. Pantic, 2005. Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern, June 25, 2005, San Diego, CA, USA., pp: 76-84.
- Valstar, M. and M. Pantic, 2006. Fully automatic facial action unit detection and temporal. Proceedings of the Conference on Computer Vision and Pattern Recognition, June 17-22, 2006, Minneapolis, USA., pp: 149-156.
- Zhang, Y. and Q. Ji, 2005. Active and dynamic information fusion for facial expression understanding from image sequences. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27: 699-714.
- Zhang, Z., M. Lyons, M. Schuster and S. Akamatsu, 1998. Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron. Proceedings of the International Conference Automatic Face and Gesture Recognition, April 14-16, 2005, IEEE Computer Society Washington, DC., USA., pp: 454-459.