# FACIAL EMOTION RECOGNITION BASED ON TWO-DIMENSIONAL EMPIRICAL MODE DECOMPOSITION AND PCA PLUS LDA

Hasimah Ali, Muthusamy Hariharan, Sazali Yaacob and Abdul Hamid Adom
School of Mechatronic Engineering, University Malaysia Perlis, Arau, Perlis, Malaysia
E-Mail: hasimahali@unimap.edu.my

## ABSTRACT

This paper proposes a new approach of using nonlinear technique, two-dimensional empirical mode decomposition (2DEMD) and PCA plus LDA for facial emotion recognition. The EMD is a non-parametric data-driven analysis tools which decomposes any nonlinear and non-stationary signals into a number of intrinsic mode functions (IMFs). In this work we used the 2DEMD which is the extension of one dimensional EMD to extract the features at multiple scales or spatial frequencies from facial images. These features called IMFs that obtained by a sifting process. To reduce dimensional features, PCA plus LDA was applied on IMF features. The obtained features were classified using k-nearest neighbor classifier. To evaluate the effectiveness of the proposed method, Cohn-Kanade database was employed. A series of experiment shows that the proposed method achieves recognition rate of 98.28% thus demonstrates a promising result for classifying the facial emotions.

**Keywords:** facial emotion recognition, two-dimensional empirical mode decomposition, PCA plus LDA, k-nn.

## INTRODUCTION

Facial emotion or synonymously facial expression is generated by contractions of facial muscles which temporarily deformed the facial components such as eyes, brows, eyelids, cheeks and lasting for few seconds (Fasel and Luettin, 2003). Facial emotion plays an important role in human communication as they interact face-to-face to convey their messages or emotional states. Thus, people can realize and create better response toward the emotion displayed. Recent advances in image analysis and pattern recognition opened up the possibility of automatic detection and classification of facial emotion in achieving human-like interaction between man and machines. Such a system could bring a tool for research in behavioral sciences and human-computer interaction (Pantic and Rothkrantz, 2000).

Nowadays, with the advanced of technologies, applications of facial emotion recognition can be found in variety of contexts such as robotics, computer animated, computer vision and computer graphic. For instance, sociable humanoid robot called *Kismet* (Breazeal, 2003) perceives natural social cues via visual and auditory channel and delivers social signal using facial expression. Besides, the *EmotiChat* (Anderson and McOwen, 2003) also has been developed by employing emotion-based application in chatroom in which the user only needs to express the emotion while they are chatting e.g. happy,and emoticons (icon for emotion) appear automatically in the text. Recently, intelligent wheelchair-based expression (Luo, Wu and Zhang, 2013) has been desiged for disabled person in helping the disabled person to control the wheelchair using the facial expression. In behavioral science and medical research for instance, facial expression is used as pain monitoring (Hammal and Kunz, 2012) for individual with severe cognitive impairment like autism which cannot communicate pain verbally.

Therefore, due to important applications of such system in daily life, the past decades have witnessned the increasing number of approaches for automatic facial emotion recognition. The common approaches that widely applied in analyzing the facial expression recognition are Gabor filter (Deng, Jin, Zhen and Huang, 2005; Donato, Bartlett, Hager, Ekman and Sejnowski; 1999; Gu, Xiang, Venkatesh, Huang and Lin, 2012; Zavaschi, Britto, Oliveira and Koerich, 2013; Lyons, Budynek and Akamatsu, 1999), local binary pattern (Shan, Gong and McOwan, 2009; Feng, Hadid and Pietikäinen, 2004; Liu, Yi and Wang, 2009; Moore and Bowden, 2009; Zhao and Zhang, 2011), wavelet (Shih, Chuang and Wang, 2008; Kazmi, Ain and Jaffar, 2012; Ali, Hariharan, Yaacob and Adom, 2014). Eventhough facial emotion recognition has reached a certain level of success; however this system is far from human visual perception. As a result, recognition of facial emotion is still challenging task due to complexity and subtlety of facial features. The problems even more hampered by existing intra-class variations and also inter-class similarities of nonlinear emotional features.

To address this problem, therefore this paper proposes an application of nonlinear technique so called two dimensional empirical mode decomposition (2DEMD) using PCA plus LDA for facial emotion classification. In this work, the facial images were subjected to 2DEMD technique to produce a set of intrinsic mode functions (IMFs) via sifting process. The obtained features were further analyzed using dimensionality reduction technique before fed to k-nn classifier.

## METHODOLOGY

Figure-1 shows the generic framework of the proposed method. The descriptions of each block are discussed as follows.

VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



**Pre-processing of facial images**

To pre-process the image, face detection is used as a technique to extract the face region from a given image of having different backgrounds.

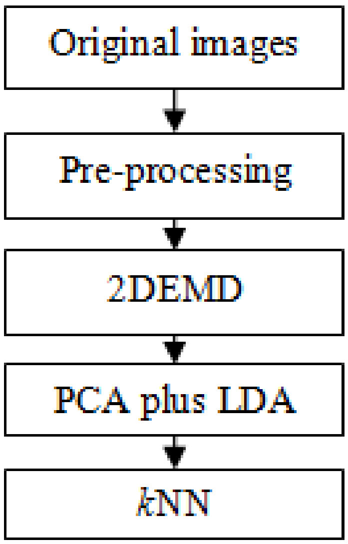


**Figure-1.** The framework of the proposed method.

Generally, most of facial expression images are obtained under controlled conditions (no lighting effects) and the acquired images are in frontal view. There are two common approaches in extracting the face region which is face as a whole unit (holistic approach) or face as a set of facial features (eyes, brows, nose and mouth). The location of the facial features in correspondence to each other sometimes determines the overall location of the face. In this work, the original facial image of size 640 x 490 pixels firstly is cropped and scaled into 132 x 132 pixels by removing the background influences.

**EMD**

The EMD is a new method used for analyzing nonlinear and non-stationary process in which any complicated data can be decomposed into a finite and often small number of IMF that admit well-behaved Hilbert transform (Huang, Shen, Long, Wu, Shih, Zheng, Yen, Tung and Liu, 1998). This method is considered as highly efficient due adaptiveness in decomposing the signal based on local characteristic time scale, thus making applicable to nonlinear and non-stationary process. It extracts the modes directly from the signal with no prior of assumption the nature of data. The usefulness of EMD technique can be seen in various research areas such as EEG, ocean, radar, earthquake, revealing the underlying oscillatory modes of real world signal. Due to advantages of EMD that is a fully data-driven method, no predetermined filter or wavelet function, thus (Nunes, Bouaoune, Delechelle, Niang and Bunel, 2003) have applied EMD in texture analysis.

**2DEMD applied to facial expression images**

The 2DEMD is the extension of one dimensional EMD widely used in image analysis such as texture analysis (Nunes, Bouaoune, Delechelle, Niang and Bunel, 2003, 2003), image compression (Linderhed, 2005), skeletonization pruning (Krinidis and Krinidis, 2013), image fusion (Hariharan, Koschan, Abidi, Gribok and Abidi, 2006), face recognition (Zhang and Tang, 2009; Bhagavatula and Savvides, 2007) and facial pose-estimation (Qing, Jiang and Yang, 2010). Thus, in this work facial images are converted into a one dimensional vector, x (t) on each row (or column) one by one. Based on (Huang, Shen, Long, Wu, Shih, Zheng, Yen, Tung and Liu, 1998) the algorithm of EMD can be summarized as follows:

1. Locate all local maxima and local minima of $x(t)$.
2. Generate upper envelope, $e_u(t)$ and lower envelope, $e_l(t)$ by cubic spline interpolation.
3. Calculate mean envelope, $m(t) = \dfrac{e_u(t) + e_l(t)}{2}$
4. Extract the details, $d(t) = x(t) - m(t)$ . In finding IMFs, two conditions should be satisfied
   a. The number extrema and number of zero crossings must be the same or differe by one.
   b. The mean value of upper envelope and lower envelope is zero.
5. Iterate on the residual $d(t)$.

The above procedure is called sifting process, which amounts to first iteration steps 1 to 4 upon detail signal $d(t)$, until this latter can be considered as zero-mean according to some stopping criterion. Once this is achieved the detail is referred to as an IMF, the corresponding residual is computed and step 5 applies.

Application of EMD to these vectors yields a set of vector IMF's which are then reshaped into matrix IMF's. Figure-2 illustrates the example of seven facial expressions and their corresponding modes and residues.

VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

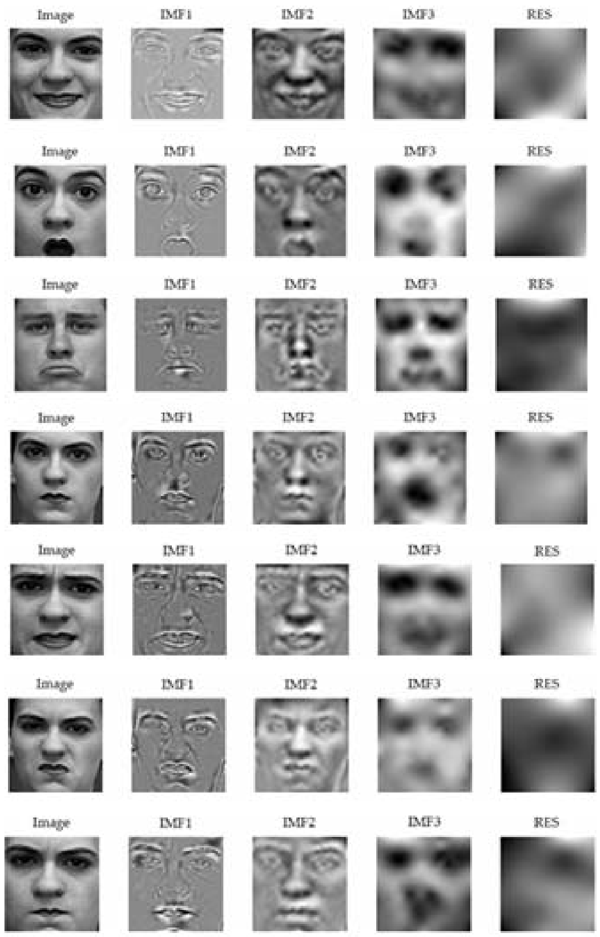


**Figure-2.** An example of seven facial expressions and their corresponding modes and residue.

We have considered IMF1 as features because the first IMF1 contains the largest magnitude extrema in which they contribute to the highest local information that describe the characteristic of facial emotion. In other words, IMF1 has highest frequency component. Then, this selected IMF1 are further analyzed using PCA plus LDA technique.

**PCA plus LDA**

Recently, PCA plus LDA has been widely applied as data dimensionality reduction techniques in pattern recognition (Belhumeur, Hespanha, and Kriegman, 1997; Deng *et al*, 2005; Martinez and Kak, 2001; Yang and Yang, 2003; Ali, Hariharan, Yaacob and Adom, 2014). Due to small sample size problem, PCA maps the original *d*-dimensional feature $x_k$ to the *f*-dimensional feature $y_k$ as an intermediate space, and then LDA projects the PCA output to a new *m*-dimensional feature vectors $z_i$ where $z_i = W_{lda}^T W_{pca}^T x_i$, for i=1, 2,…n. To apply PCA, let X is the data matrix of IMF1 features of entire dataset. Each row of X represents IMF1 facial image. Suppose $X=\{x_1, x_2, x_3 \ldots x_n\}$ be a feature vector of coefficient matrices of IMF1 with observations $x_i \in R^d$, Therefore, PCA can be calculated in following steps:

**Step-1:** Compute the mean of the data matrix $\mu$,

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

**Step-2:** Compute the covariance matrix $S$,

$$S = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)(x_i - \mu)^T \tag{2}$$

**Step-3:** Compute the eigenvalues $\lambda_i$ and eigenvectors $v_i$ of $S$,

$$Sv_i = \lambda_i v_i \tag{3}$$

, $i$=1,2, …, $n$

**Step-4:** Sort the eigenvectors corresponding to the $k$ largest eigenvalues. The projected sample in PCA space is given by:

$$y = W^T x \tag{4}$$

where $W$ is 404x404 dimension. In this work, 95% of variances (350 principal components) are kept to the energy level. The output of PCA then is projected to LDA subspace. The goal of LDA is to maximize between-*class* scatter matrix while minimize *within-class*r scatter matrix. In face recognition, LDA is more capable of distinguishing image variations due to different identities especially variations in illumination and also expression (Belhumeur, Hespanha, and Kriegman, 1997). To apply LDA, let $\{y_k | k=1, 2, \ldots, n\}$ is a PCA output with n samples in *f*-dimensional spaces. Let $l_i$ be the class label of $y_k$, where $l_i \in \{1,2,\ldots c\}$ and $c$ is number of class. Denote the $i$-th class samples by $N_i$. The scatter matrices $S_B$ and $S_W$ are given as,

$$S_B = \sum_{i=1}^{c} N_i (\mu_i - \mu)(\mu_i - \mu)^T \tag{5}$$

$$S_W = \sum_{i=1}^{c}\sum_{k=1}^{n} (y_k - \mu_i)(y_k - \mu_i)^T \tag{6}$$

Where $\mu$ is the global mean, $\mu_i$ is the mean of the $i$-th class:

$$\mu = \frac{1}{n}\sum_{k=1}^{n} y_k \tag{7}$$

VOL. 10, NO. 2, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences



$$\mu_i = \frac{1}{N_i}\sum_{k=1}^{N_i} y_k \tag{8}$$

If Sw is non-singular, the optimal projection $W_{opt}$ is the chosen matrix with orthonormal column, which maximize the ratio of determinant of between-class scatter of the projected samples to the determinant of within-class scatter matrix of the projected sample e.g.,

$$W_{lda} = \arg\max \frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|} = [w_1, w_2, w_3 ... w_m] \tag{9}$$

The projected sample in new PCA plus LDA space is given by:

$$z_i = W_{lda}^T W_{pca}^T x_i \tag{10}$$

Note that the dimension of LDA is bound to c-1, where c is the number of classes, thus in this work there are (7-1) classes is the final dimension. The first six columns of $z_i$ in Eqn. (10) are considered as six features and used to classify the seven facial emotions.

## RESULTS AND DISCUSSIONS

To evaluate the effectiveness of the proposed method, standard Cohn-Kanade Facial Expression Database (Kanade, Cohn and Tian, 2000) was used in this experiment. It consists of 500 images sequences having difference groups and ethnicities. The grayscale of the facial image start with neutral and slightly increase to the apex emotion. The subject was instructed to perform a series of six basic emotions (anger, disgust, fear, happiness, sadness and surprise). We have used 404 digitized facial images that consist of seven facial expressions (38 angry, 59 disgust, 52, fear, 75 happy, 54 sad, 46 surprise and 80 neutral). We note that, for some particular expressions of the particular subjects are missing. We also exluded facial images with high illumination effect in this experiment.

In extracting the features, firstly the facial images were decomposed using 2DEMD to produce IMF1, IMF2, IMF3 and residue. The 1st IMF (IMF1) was selected and considered as features since IMF1 contains significant features for classifying the seven facial emotions while IMF2 and IMF3 are discarded. The results of IMF1, IMF2, IMF3 and residue have been shown in Figure-2. As observed in Figure-2 that, all three modes (IMF1, IMF2, IMF3) and residue exhibit the pattern structures from finest to coarsest of the original expression. Due to nature of the EMD algorithm, as the order of the IMF increases, the relative mean of the data approaches to zero (Rilling, Flandrin and Goncalves, 2003). The first IMF (IMF1) is effectively has the largest magnitude extrema in these vectors which contribute to the highest local information that describes the characteristic of distinct facial emotion. So, we have chosen IMF1 as features for further analysis since IMF1 has the highest frequency component.

The IMF1 of size 132 x 132 pixels was concatenated into one dimensional array (1 x 17424) to form a feature vector. The obtained features were then subjected to PCA plus LDA for dimensionality data reduction to map high dimensional into a lower dimensional space. The reduced features finally were classify using *k*NN classifier. To evaluate the proposed method, we employed 10-fold cross-validation method. That is, the entire dataset was randomly divided into ten groups in which each group contains approximately the same proportion of class samples as the original dataset. Nine groups of the data were used as training, and the remaining one group (testing set) was used as testing. This procedure was repeated 10 times (fold) and average recognition of 10 folds is calculated. Besides, we also conduct using 5-fold cross-validation for comparison. Table-1 shows the recognition results of IMF1+PCA+LDA using k-nn classifier.

**Table-1.** Recognition rates IMF1+PCA+LDA features using k-nn classifier.

| k-nn | 5-fold (%) | 10-fold (%) |
|---|---|---|
| ***k*** = 1 | 97.31 | 96.07 |
| ***k*** = 2 | 97.89 | 98.28 |
| ***k*** = 3 | 97.71 | 97.10 |

As we can see from Table-1 the proposed IMF1+ PCA + LDA using simple k-nn classifier achieves the best accuracy of 98.28% using 10-fold cross validation in classifying the seven facial emotions. This result shows that the proposed 2DEMD + PCA + LDA is effective and highly efficient approach towards achieving automated facial emotion recognition.

## CONCLUSIONS

This paper has presented a new approach in classifying the seven facial emotions based on 2DEMD + PCA + LDA features using k-nn classifier. The 2DEMD has decomposed the facial images into IMF1, IMF2, IMF3 and residue before subjected to dimensionality reduction and classification. Based on the result obtained, the application of 2DEMD + PCA + LDA for facial emotion recognition is highly efficient since its demonstrated the recognition result of 98.28%. However, further studies has to be done to investigate the proposed method with different classifiers and tested using person-independent database as well as apply to generalization across-database before adopting into real-time application so that robust human-computer interface system can be achieved.

## REFERENCES

Ali H., Hariharan M., Yaacob S. and Adom A. H. 2014. Hybrid feature extraction for facial emotion recognition.

Int. J. Intelligent Systems Technologies and Applications. 13(3): 202-221.

Anderson K. and McOwan P.W. 2004. Robust real-time face tracker for use in cluttered environments. Computer Vision and Image Understanding. 95(2): 184-200.

Belhumeur P.N., J.P. Hespanha J.P. and D.J. Kriegman D.J. 1997. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 19(7): 711-720.

Bhagavatula R. and Savvides M. 2007. Analyzing facial images using empirical mode decomposition for illumination artifact removal and improved face recognition. Proc. of IEEE ICASSP2007. pp. 505-508.

Breazeal C. 2003. Emotion and sociable humanoid robots. Int. J. Human-Computer Studies. 59: 119-155.

Deng H.B., Jin L.W., Zhen L.X and Huang J.C. 2005. A new facial expression recognition method based on local gabor filter bank and PCA plus LDA. International Journal of Information Technology. 11(11): 86-96

Donato G., Bartlett M. S., Hager J. C., Ekman P. and Sejnowski T. J. 1999. Classifying facial actions. IEEE Trans. Pattern Anal. Mach. Intell. 21(10): 974–989

Fasel B. and Luettin J. 2003. Automatic facial expression analysis: A survey. Pattern Recognition. 1(30): 259-275.

Feng X.Y., Hadid A. and Pietik¨ainen M. 2005. Facial expression recognition with local binary patterns and linear programming. Pattern Recognition and Image Analysis. 15(2): 546-548.

Gu W. F., Venkatesh Y. V. and Xiang C. 2010. A novel application of self-organizing network for facial expression recognition from radial encoded contours. Soft Computing. 14, pp. 113-122.

Hammal Z. and Kunz M. 2012. Pain monitoring: A dynamic and context-sensitive system. Pattern Recognition. 45: 1265-1280.

Hariharan H., Koschan A., Abidi B., Gribok A. and Abidi M. 2006. Fusion of visible and infrared images using empirical mode decomposition to improve face recognition. Proc. of IEEE international Conference on Image Processing ICIP, Atlanta, GA. pp. 2049-2052.

Huang N. E., Shen Z., Long S., Wu M., Shih H., Zheng Q., Yen N., Tung C. and Liu H. 1998. The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis. Proc. R. Soc. London. 454: 903-995.

Kanade T., Cohn J. and Tian Y. 2000. Comprehensive database for facial expression analysis. In: IEEE International Conference on Automatic Face and Gesture Recognition (FG).

Kazmi S.B., Qurat-ul-Ain and Jaffar M. A. 2012. Wavelet-based facial expression recognition using a bank of support vector machines. Soft Computing. 16(3): 369-379.

Krinidis S. and Krinidis M. 2013. Empirical mode decompsition on skeletonization pruning. Image and Vision Computing. 31: 533-541.

Linderhed A. 2005. Compression by image empirical mode decomposition. IEEE International Conference on Image Processing. pp. 553-556.

Liu W. F., Yi S. J. and Wang Y. J. 2009. Automatic facial expression recognition based on local binary patterns of local areas. WASE International Conference on Information Engineering. pp. 197-200.

Luo Y., Wu C-M. and Zhang Y. 2013. Facial expression based on fusion feature of PCA and LBP with SVM. Optik. 124: 2767-2770.

Lyons M, Budynek J and Akamatsu S. 1999. Automatic classification of single facial images. IEEE Trans. Pattern Anal. Mach. Intell. 21(12): 1357-1362.

Martinez A. M. and Kak A. C. 2001. PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence. 23(2): 228-233.

Moore S. and Bowden R. 2009. The effects of pose on facial expression recognition. In: A. Cavallaro, S. Prince and D. Alexander. Proceedings of the British Machine Conference. pp. 79.1-79.11.

Nunes J. C., Bouaoune Y., Delechelle E., Niang O. and Bunel Ph. 2003. Image analysis by bidimensional empirical mode decomposition. Image and Vision Computing. 21: 1019-1026.

Pantic M. and Rothkrantz L. J. M. 2000. Automatic analysis of facial expressions: The state of the art. IEEE Trans. Pattern Analysis and Machine Intelligence. 22(12): 1424-1445.

Qing C., Jiang J. and Yang Z. 2010. Empirical mode decomposition-based facial pose estimation inside video sequences. Optical Engineering. 49(3): 037401(1-10).

Rilling G., Flandrin P. and Goncalves P. 2003. On empirical mode decomposition and its algorithms. Proc. of the IEEE-Eurasip Workshop on Nonlinear Signaland Image processing. 3: 8-11.

Shan C., Gong S. and McOwan P.W. 2009. Facial expression recognition based on local binary patterns: a comprehensive study. Image and Vision Computing. 27: 803-816.

Shih F.Y., Chuang C. F. and Wang P. S. P. 2008. Performance comparisons of facial expression recognition in JAFFE database. International Journal of Pattern Recognition. 22(3): 445-459.

Yang J. and Yang J-Yu. 2003. Why can LDA be performed in PCA transformed space?. Pattern Recognition. 36: 563-566.

Zavaschi T. H.H., Britto Jr. A. S., Oliveira L. E. S. and Koerich A. L. 2013. Fusion of feature sets and classifiers for facial expression recognition. Expert Systems with Applications. 40: 646-655.

Zhang D. and Tang Y. Y. 2009. Extraction of illumination-invariant features in face recognition by empirical mode decomposition. In: M. Tistarelli and M.S. Nixon (Eds), ICB, LNCS 5558. pp. 102-111.

Zhao X. and Zhang S. 2011. Facial expression recognition based on local binary patterns and kernel discriminant isomap. Sensors. pp. 9573-9588.