Expert Systems with Applications 42 (2015) 1261–1277

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Facial emotion recognition using empirical mode decomposition

Hasimah Ali *, Muthusamy Hariharan, Sazali Yaacob, Abdul Hamid Adom

School of Mechatronic Engineering, University Malaysia Perlis, 02600 Arau, Perlis, Malaysia

ARTICLE INFO

Article history:
Available online 16 September 2014

Keywords:
Facial emotion
Empirical mode decomposition
PCA + LDA
LFDA
KLFDA
Extreme learning machines

ABSTRACT

This paper proposes a new method of using empirical mode decomposition (EMD) technique for facial emotion recognition. The EMD algorithm can decompose any nonlinear and non-stationary signal into a number of intrinsic mode functions (IMFs). In this method, the facial signal obtained from successive projection of Radon transform of 2-D image is decomposed using EMD into oscillating components called IMFs. The first IMF (IMF1) was extracted and considered as features to recognize the facial emotions. Three dimensionality reduction algorithms: Principal Component Analysis (PCA) + Linear Discriminant Analysis (LDA), PCA + Local Fisher Discriminant Analysis (LFDA), and Kernel LFDA (KLFDA) were independently applied on EMD-based features for dimensionality reduction. These dimensionality reduced features were fed to the k-Nearest Neighbor (k-NN), Support Vector machine (SVM) and Extreme Learning Machines with Radial Basis Function (ELM-RBF) classifiers for classification of seven universal facial expressions. The proposed method was evaluated using two benchmark databases JAFFE and CK. The experimental results on both facial expression databases demonstrate the effectiveness of the proposed algorithm.

1. Introduction

Facial emotion or synonymously facial expression is one of the most powerful and natural means for human beings to deliver their messages or emotional states. A study based on communications of feelings and attitudes conducted by psychologist (Mehrabian, 1968) founded that 55% of messages pertaining to feeling and attitude have contributed via facial expression. This implies that the facial expression forms a major channel in human social interaction. Generally, facial emotion is generated by contraction of one or more of facial muscles that temporarily deformed the facial components such as eyelids, eyebrows, nose, mouth, chin and lips due to emotion instability. Cross-culture research has shown that six basic emotions i.e. anger, disgust, fear, happiness, sadness and surprise are universal across human ethnicities and cultures (Ekman, 1992). In Ekman and Friesen (1978) developed Facial Actions Coding Systems (FACS) used to measure the facial expression by using action units (AUs) based on facial anatomy. FACS also provides tools for behavioral science research, facial behavior and becomes a strong basis for human–computer interaction (HCI). Over the last decade, facial expression recognitions have received a great attention among computer vision, pattern recognition and its related field. A detailed survey on facial expression recognition can be referred to Pantic and Rothkrantz (2000), Fasel and Luettin (2003).

Nowadays, with the advanced of technology the facial expression based emotion recognition has shown fast growing in the development of robust HCI. This HCI has focused on the invention of social welfare robot in order to help the physically disable person who is either confined on bed or wheelchair. As the robot becomes part of our living space, it is important for a robot to understand human's mood and emotion so that natural interaction between man and machine can be achieved. For example, the invention of nurse robot could assists a patient in hospital via estimating his/her emotion and create a better response for the particular emotions displayed. Other several application domains including intelligent tutoring system (Whitehill, Bartlett, & Movellan, 2008), *Emochat* (Anderson & McOwan, 2004), facial animation, virtual reality of facial emotion recognition have been built based on the natural form of human communication.

There are various techniques of facial emotion recognition are constantly proposed in the past. Some of them have utilized pixel based (Rahulamathavan, Phan, Chambers, & Parish, 2013; Wang & Ruan, 2010), Gabor filter (Deng, Jin, Zhen, & Huang, 2005; Donato, Bartlett, Hager, Ekman, & Sejnowski, 1999; Owusu, Zhan, & Mao, 2014), wavelet transform (Kazmi, Qurat-ul-Ain, & Jaffar, 2012; Shih, Chuang, & Wang, 2008), facial contour (Gu, Venkatesh, &

* Corresponding author.
E-mail addresses: hasimahali@unimap.edu.my (H. Ali), hari@unimap.edu.my (M. Hariharan), s.yaacob@unimap.edu.my (S. Yaacob), abdhamid@unimap.edu.my (A.H. Adom).

http://dx.doi.org/10.1016/j.eswa.2014.08.049
0957-4174/

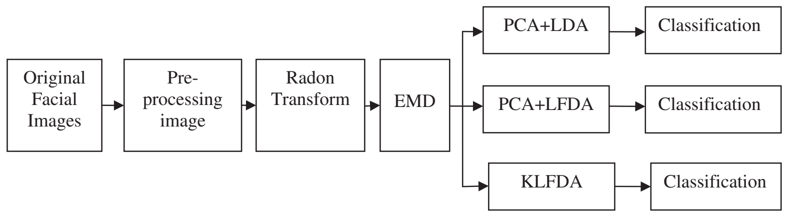

Fig. 1. Block diagram of the proposed method.

Fig. 2. Face model (left) and example of cropped facial emotions of JAFFE subject (right).

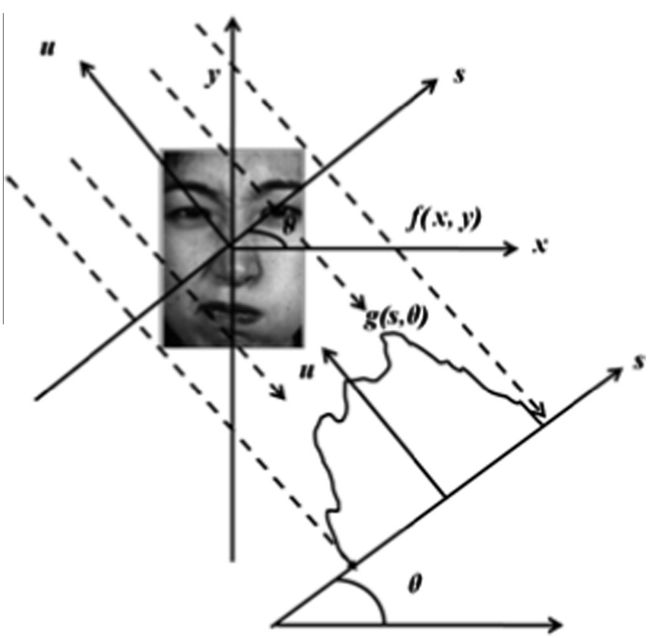

Fig. 3. Projection of 2D facial images into 1D signal using Radon transform.

Xiang, 2010), edge and skin detection (Ilbeygi & Hosseini, 2012), discrete cosine transform (Gupta, Agrwal, Meena, & Nain, 2011; Kharat & Dudul, 2009) and local binary pattern (Feng, Hadid, & Pietikäinen, 2005; Liu, Yi, & Wang, 2009; Luo, Wu, & Zhang, 2013; Moore & Bowden, 2009; Shan, Gong, & McOwan, 2009; Zhang, Zhao, & Lei, 2012; Zhao & Zhang, 2011) have gained lots of successful experiences. Eventhough facial emotion recognitions have achieved a certain level of success, however the performance is far from human perception. The ability of facial emotion recognition to operate in fully automated with high accuracy are remains challenging due to complexity, variability, subtle changes of non-linear emotional facial features.

Therefore, this paper proposes an application of nonlinear and non-stationary data analysis technique called, empirical mode decomposition (EMD) for classifying the facial emotion. Recent studies show that EMD algorithm was used to detect epileptic seizure using EEG signals (Kaleem, Guergachi, & Krishnan, 2013; Orosco, Lacia, Correa, Torres, & Graffigna, 2009; Pachori, 2008) and diagnose Alzheimer disease (Gallix, Gorriz, Ramirez, & Lang, 2012). Also, EMD algorithm is extended into bi-dimensional EMD (BEMD) and has been used for texture analysis (Nunes, Bouaoune, Delechelle, Niang, & Bunel, 2003), image compression (Linderhed, 2005), skeletonization pruning (Krinidis & Krinidis, 2013), image fusion (Hariharan, Koschan, Abidi, Gribok, & Abidi, 2006), face recognition (Bhagavatula & Savvides, 2007; Zhang & Tang, 2009) and facial pose-estimation (Qing, Jiang, & Yang, 2010). EMD is a multi-resolution decomposition technique suitable for nonlinear data has been used to decompose any complicated signal into frequency components called IMFs. The major advantage of using EMD is that the basis functions can be directly derived from the signal itself based on the local characteristic time scale of the data which provides fully data-driven approach and often brings not only high decomposition efficiency but also sharp frequency and time localizations (Nunes et al., 2003).

Due to the advantages offer by EMD, therefore this paper introduces the application of nonlinear technique of using EMD for facial emotion recognition using static images. In our proposed method, the facial signal obtained from successive projection of Radon transform of 2-D image is decomposed using EMD and subsequently three dimensionality reduction methods were applied independently to extract the features. Three dimensionality reduction methods were used: PCA + LDA, PCA + LFDA and KLFDA were independently applied on EMD-based features. The obtained feature vectors then were subjected to statistical analysis using ANOVA test. These features with reduced dimensionality that statistically significant were fed to the k-NN, SVM and ELM-RBF classifiers for classification of seven types of facial emotions.

The paper is organized as follows: Section 2 presents database used, Section 3 deals with methods and classifiers used, Section 4 provides the experimental results and discussion, Finally Section 5 concludes the paper.

2. Database

Two benchmark databases were used in this work: JAFFE and CK database. First, Japanese Female Facial Expression (JAFFE) data-