International Journal of Multimedia Data Engineering and Management
Volume 12 • Issue 2 • April-June 2021

Improving Emotion Analysis for Speech-Induced EEGs Through EEMD-HHT-Based Feature Extraction and Electrode Selection

Jing Chen, Harbin Institute of Technology, China
Haifeng Li, Harbin Institute of Technology, China
Lin Ma, Harbin Institute of Technology, China
Hongjian Bo, Shenzhen Academy of Aerospace Technology, China

ABSTRACT

Emotion detection using EEG signals has advantages in eliminating social masking to obtain a better understanding of underlying emotions. This paper presents the cognitive response to emotional speech and emotion recognition from EEG signals. A framework is proposed to recognize mental states from EEG signals induced by emotional speech: First, speech-evoked emotion cognitive experiment is designed, and EEG dataset is collected. Second, power-related features are extracted using EEMD-HHT, which is more accurate to reflect the instantaneous frequency of the signal than STFT and WT. An extensive analysis of relationships between frequency bands and emotional annotation of stimulus are presented using MIC and statistical analysis. The strongest correlations with EEG signals are found in lateral and medial orbitofrontal cortex (OFC). Finally, the performance of different feature set and classifier combinations are evaluated, and the experiments show that the framework proposed in this paper can effectively recognize emotion from EEG signals with accuracy of 75.7% for valence and 71.4% for arousal.

KEYWORDS

Electro-Encephalograph (EEG), Emotion Recognition, Ensemble Empirical Mode Decomposition (EEMD), Hilbert-Huang Transform (HHT), Maximal Information Coefficient (MIC), Time-Frequency Analysis

INTRODUCTION

Emotion plays an important role in human mental life. It is a conscious mental experience reflecting the personal significance of internal and external events. Human speech conveys not only linguistic messages, but also emotional information. Speech is one of the principal conveyers to express one's emotion during human social communication. The ability to identify vocal expressions of emotion or attitude in speech is one of the basic cognitive functions of human beings. Although emotional speech are as ubiquitous as facial expressions, far less is known about the brain mechanisms of emotion perception in the speech than emotional facial expressions.

Emotion detection using physiological responses have advantages on eliminating social masking to obtain a better understanding of underlying emotions. Electroencephalography (EEG) signals are the summation of the activities of billions of neurons in the cerebral cortex, which can directly reflect brain activities. EEG is noninvasive and provides high temporal resolution which can reflect effective emotional changes in brain. Furthermore, advances in wearability, price, portability and ease-to-use,

DOI: 10.4018/IJMDEM.2021040101

emotion recognition based on EEG has received extensive attention in numerous fields such as affective brain-computer interface (BCI), neuroscience, health care, emotional companionship and e-learning.

Multi-channel EEG provides population measures of neurons that allow us to uncover the complex cognitive processes of emotional information integration and processing. The fundamental goal of decoding mental states from EEG recordings is to identify emotions in EEG data, which correspond to the experimental task or stimulus by using machine learning classifiers. The research about emotion recognition from EEG is still progressing and few studies regard to emotion detection from speech-evoked EEG signals. This paper focuses on EEG-based mental state recognition induced by emotional speech.

Our EEG-based emotion recognition framework mainly includes two modules: feature extraction and classification. Various features have been extracted from different type of domains (time-domain, frequency-domain, time-frequency domain) for EEG analysis. This research will limit the scope of discussion to time-frequency analysis and recognition. Power-related features from different frequency bands of EEG are often used during classification. The most commonly used analytical technique is short-time Fourier transform(STFT), wavelet transform(WT).

In this work, Ensemble Empirical Mode Decomposition and Hilbert-Huang Transform (EEMD-HHT) are first applied to analyze the time-frequency distribution of EEG signals and corresponding features are extracted for classification. Compared with the traditional Pwelch frequency feature extraction method based on STFT, HHT can process non-stationary signals and obtain practical instantaneous frequencies, which is more accurate to reflect the actual frequency of the signal. EEMD is an adaptive decomposition method, which avoids the limitation of selecting the mother wavelet and setting the number of decomposition layers in WT. Then, maximal information coefficient (MIC) is applied to measure the relationship between frequency band power and the level of arousal or valence. Next, the authors reduce the feature dimensions by using statistical test to determine whether the energy of electrodes significantly vary under different conditions. Finally, three classifiers are used for classification.

The remaining of this paper is organized as follows: First, the authors briefly introduce relevant researches on emotion recognition based on EEG signals . Second, the authors describe the speech-evoked emotion cognitive experiment protocol and EEG dataset collection, followed by the feature extraction method based on EEMD-HHT. Third, the relationship between band power and the level of emotion rating will be presented. Fourth, the classification results are provided and discussed in Results and Discussion section. Finally, the authors conclude the findings and discuss future directions.

RELATED WORK

EEG signals can provide insights into relationship between emotional states and brain activities. Numerous studies have estimated the changes in human emotional state respond to multimedia stimulus(Alarcao & Fonseca, 2017). Common stimulus for inducing emotion include images (International Affective Picture System and Geneva Affective Picture Database), sound (International Affective Digitized Sound System) and music or film video. There have been several widely used EEG datasets for emotion analysis, including DEAP(Koelstra et al., 2011), MAHNOB-HCI(Soleymani, Lichtenauer, Pun, & Pantic, 2011), SEED(Zheng, Liu, Lu, Lu, & Cichocki, 2018; Zheng & Lu, 2015). DEAP utilized 40 one-minute long music videos as stimulus. The accuracies of 2-classes emotion recognition on DEAP dataset were 57.7% for valence and 62.5% for arousal(Koelstra et al., 2011). (Thammasan, Moriyama, Fukui, & Numao, 2017) used PSD feature and SVM to classify emotion and achieved performance of 73.3% for arousal and 72.5% for valence on DEAP dataset. (Bo, Ma, Liu, Xu, & Li, 2019) designed a three-stage experimental paradigm of long-time music stimuli and got an accuracy of 66.8% for valence and 59.5% for arousal. However, none of these datasets are derived from emotional speech stimulus. Lack of EEG dataset induced by emotional speech limits the exploration of cognitive mechanisms of language's role in emotion concept acquisition.

Related Content

Application of Error Control Coding for Multimedia Watermarking Technologies
Mehul S. Raval (2010). *Advanced Techniques in Multimedia Watermarking: Image, Video and Audio Applications (pp. 407-424).*
www.irma-international.org/chapter/application-error-control-coding-multimedia/43480

Evolution of Big Data in Medical Imaging Modalities to Extract Features Using Region Growing Segmentation, GLCM, and Discrete Wavelet Transform
Yogesh Kumar Gupta (2021). *Advancements in Security and Privacy Initiatives for Multimedia Images (pp. 41-78).*
www.irma-international.org/chapter/evolution-of-big-data-in-medical-imaging-modalities-to-extract-features-using-region-growing-segmentation-glcm-and-discrete-wavelet-transform/262067

Automation of Explainability Auditing for Image Recognition
Duleep Rathgamage Don, Jonathan Boardman, Sudhashree Sayenju, Ramazan Aygun, Yifan Zhang, Bill Franks, Sereres Johnston, George Lee, Dan Sullivanand Girish Modgil (2023). *International Journal of Multimedia Data Engineering and Management (pp. 1-17).*
www.irma-international.org/article/automation-of-explainability-auditing-for-image-recognition/332882

Generating Personalized Explanations for Recommender Systems Using a Knowledge Base
Yuhao Chen, Shi-Jun Luo, Hyoil Han, Jun Miyazakiand Alfrin Letus Saldanha (2021). *International Journal of Multimedia Data Engineering and Management (pp. 20-37).*
www.irma-international.org/article/generating-personalized-explanations-for-recommender-systems-using-a-knowledge-base/301455

A Randomized Framework for Estimating Image Saliency Through Sparse Signal Reconstruction

Kui Fuand Jia Li (2018). *International Journal of Multimedia Data Engineering and Management (pp. 1-20).*

www.irma-international.org/article/a-randomized-framework-for-estimating-image-saliency-through-sparse-signal-reconstruction/201913