International Journal of Psychophysiology 96 (2015) 29–37

Contents lists available at ScienceDirect

International Journal of Psychophysiology

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

Randomly dividing homologous samples leads to overinflated accuracies for emotion recognition

Shuang Liu [1], Di Zhang [1], Minpeng Xu, Hongzhi Qi *, Feng He, Xin Zhao, Peng Zhou, Lixin Zhang, Dong Ming *

Neural Engineering & Rehabilitation Lab, Department of Biomedical Engineering, College of Precision Instruments and Optoelectronics Engineering, Tianjin University, China

ARTICLE INFO

Article history:
Received 1 November 2014
Received in revised form 15 February 2015
Accepted 16 February 2015
Available online 21 February 2015

Keywords:
Emotion recognition
Electroencephalography (EEG)
Overinflated accuracies
Affective computing
Feature selection
Valence

ABSTRACT

There are numerous studies measuring the brain emotional status by analyzing EEGs under the emotional stimuli that have occurred. However, they often randomly divide the homologous samples into training and testing groups, known as randomly dividing homologous samples (RDHS), despite considering the impact of the non-emotional information among them, which would inflate the recognition accuracy. This work proposed a modified method, the integrating homologous samples (IHS), where the homologous samples were either used to build a classifier, or to be tested. The results showed that the classification accuracy was much lower for the IHS than for the RDHS. Furthermore, a positive correlation was found between the accuracy and the overlapping rate of the homologous samples. These findings implied that the overinflated accuracy did exist in those previous studies where the RDHS method was employed for emotion recognition. Moreover, this study performed a feature selection for the IHS condition based on the support vector machine-recursive feature elimination, after which the average accuracies were greatly improved to 85.71% and 77.18% in the picture-induced and video-induced tasks, respectively.

1. Introduction

Emotion is a psycho-physiological process triggered by the conscious and/or unconscious perception of an object or a situation, which is often associated with mood, temperament, personality disposition and motivation (Koelstra et al., 2012). Emotion recognition has received an increasing amount of attention since the accumulation of negative emotions would impede the work of the immune system, which would make people more vulnerable to viral infections, increase the time for patients recovering from surgery or disease (Picard et al., 2001), and even severely affect people's performance and efficiency. Recently, researchers focus on recognizing human emotions by using the EEGs (Wang et al., 2014). Compared to the audiovisual recordings such as facial expressions, voice and gestures, EEG provides the most direct, nonintrusive and portable measurements for the central nervous system which is where emotions originate (Laine et al., 2002). EEG can also directly reflect the emotional state with a high temporal resolution. Therefore the EEG measurement possesses crucial advantages for practical use, which makes it a primary option for developing the online emotion recognition system.

A well-established mechanism for emotion induction, is to trigger emotions by utilizing affective pictures (Brown et al., 2011; Hidalgo-Muñoz et al., 2013a; Horlings et al., 2008; Peng et al., 2005) or affective video clips (Murugappan et al., 2011; Nie et al., 2011; Soleymani et al., 2012). Affective sounds have also been used in the existing emotional studies (Bradley and Lang, 2000). The affective pictures and sounds are usually picked out from the international affective picture system (Brown et al., 2011; Hidalgo-Muñoz et al., 2013a, 2013b) or the self-improved IAPS's (Balconi and Mazza, 2009) and the international affective digital sounds system (IADS) (Bradley and Lang, 2000), while the video clips are mostly selected by researchers, as no standard video library exists. The classification process, however, varies greatly among differing studies depending on the diverse factors that are performance, including the number of emotional categories, presentation, type of stimuli and the number of available trials. In general, the video induced task outperformed the picture induced task in classification.

To date, different EEG features have been used to recognize emotions. Previous studies suggested that the spectral power in various frequency bands was a distinguishable indicator of emotion. The alpha power varies with valence states (Verma and Tiwary, 2014) and also with some discrete emotions such as happiness, sadness and fear (Balconi and Lucchiari, 2006). Specifically, the frontal asymmetry of the alpha power has been repeatedly reported to steadily correlate with the valence states (Gotlib, 1998; Lee et al., 2014). In addition, the interactive properties between a pair of EEG oscillations, such as phase synchronization and coherence (Balconi and Mazza, 2009; Martini

* Corresponding authors at: Neural Engineering & Rehabilitation Lab, Department of Biomedical Engineering, College of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China. Tel.: +86 02227408718.
E-mail addresses: qhz@tju.edu.cn (H. Qi), richardming@tju.edu (D. Ming).
[1] Equal contributors.

http://dx.doi.org/10.1016/j.ijpsycho.2015.02.023
0167-8760/

et al., 2012; Miskovic and Schmidt, 2010; Wyczesany et al., 2011), have also been used as a feature for the recognition of emotion. It is generally accepted that the brain is an incredibly complex system, and therefore the nonlinear features should be taken into account in its modeling and analysis. Current research highlights that many nonlinear methods can distinguish different emotions, such as fractal dimension (an estimate of the degrees of system freedom) (Liu et al., 2010; Sourina et al., 2009a, 2009b; Sourina et al., 2009a, 2009b), Hurst exponent (Wang et al., 2014) and entropy (Duan et al., 2013; Wang et al., 2014) (reflecting the unpredictability of dynamics due to the sensitive dependence on the initial conditions). In terms of the classification, various machine learning methods can be used, such as the support vector machine (SVM) (Brown et al., 2011; Duan et al., 2013; Hidalgo-Muñoz et al., 2013a, 2013b; Nie et al., 2011; Soleymani et al., 2012), k-nearest neighbor (Guyon and Elisseeff, 2003; Murugappan et al., 2011), multi-layer perceptron (Yoon and Chung, 2013), linear discriminant analysis (Murugappan et al., 2011) and Bayesian network (Hidalgo-Muñoz et al., 2013a, 2013b). However, it should be noted that there is no single best classification algorithm, and it greatly depends on the characteristics of the data set to be classified when choosing the appropriate algorithm to use.

The recorded EEG responses contain not only the information related to the emotional state, but also others related to the basic sensory information processing and information processing delivered by the stimulus material which may play a significant role in the classification. However, this was ignored in the previous studies. For example, besides the emotional differences, a faster changing video would induce a different response in the occipital cortex than would a static video (Soleymani et al., 2012). In the pattern recognition theory, parts of samples are used to build the emotion classifier and the rest are used to predict the emotional labels that are needed to get a classification rate, respectively named the training samples and testing samples. And in the traditional emotion classification strategy, parts of the mixed samples from all the stimulus material are used for building a classifier, while the rest are used for the testing samples. It takes a big chance that the testing samples are partly recognized by other responses, but not completely by the emotional features. So it is necessary to remove the non-emotional features of the brain response before the recognition of emotional states, which is however neglected in the traditional method. Therefore, in previous studies, it was risky that some non-emotional features played an important role in the recognition of emotional states. So the partition of the samples from the same video, i.e., the partition of the homologous samples, would affect the final result, as they share some of the non-emotional information. As depicted in Fig. 1, A and B, or C and D are the homologous samples, whereas A and C are the heterogeneous samples, i.e., samples from a different picture. When the homologous samples are distributed in both the training and testing sets, for example A is sent to the training set while B is sent to the testing set, the shared non-emotional information from A will make it easier to correctly recognize B for the classifier, thereby resulting in an overinflated accuracy. Additionally, time overlap between the extracted homologous samples may bring about a more serious overinflated effect on the emotion recognition (Peng et al., 2005). Furthermore, the matter of dividing homologous samples in the pattern recognition step may also be essential in other fields related to the extremely complex cognitive processes, such as studies focusing on mental workload, mental fatigue and vigilance.

This work proposed a modified method, i.e., the integrating homologous samples method (IHS) in emotion recognition, and compared the classification results between the IHS and the previous method, i.e., the randomly dividing homologous samples method (RDHS). Additionally, the problem of the overinflated effect has been proven to exist with the RDHS method, which was generally used in the previous studies. Thus, a feature selection based on the recursive feature elimination (RFE) was employed to optimize the accuracies obtained with the IHS method, for finding the most relevant EEG features with the brain emotion.

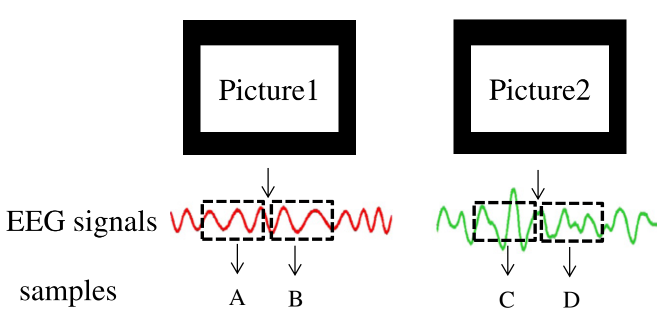

Fig. 1. Homologous samples and non-homologous samples. Samples A and B, also C and D, are extracted from the EEG signal that was evoked by the same affective picture, literally they are called homologous samples, alternatively we call them non-homologous (heterogeneous) samples, such as A and C.

2. Materials and methods

2.1. Emotion models

As all people express their emotions differently, it is not an easy task to judge or to model human emotions. Researchers often use two methods to model emotions. One approach is to label the emotions as discrete categories, like anger, happiness, love, etc. (Hoseingholizade et al., 2012). Some psychologists, however, pointed out that the emotions are not discrete but rather continuous phenomena. The other way is to have multiple dimensions or scales to categorize emotions. A common continuous model is the 2-D model, which consists of two affective dimensions, the valence and the arousal (Scherer, 2005). Emotions could be identified by their position in the two-dimensional space as shown in Fig. 2. For example, happiness has a positive valence, while disgust has a negative valence. On the other hand, sadness has a low arousal level, whereas disgust has a high arousal level. The emotions employed in this work are shown in Fig. 2. The valence-arousal space can be subdivided into four quadrants, namely, EQ1 (high arousal/positive), EQ2 (low arousal/positive), EQ3 (low arousal/negative) and EQ4 (high arousal/negative). It should be noted that there is an absence of emotional states in EQ2 in this study due to the difficulty of inducing emotions with a strong valence but a low arousal (Köchel et al., 2011), which is also observed in the well-validated ratings for the IAPS (Lang et al., 2005) and the international affective digital sounds system

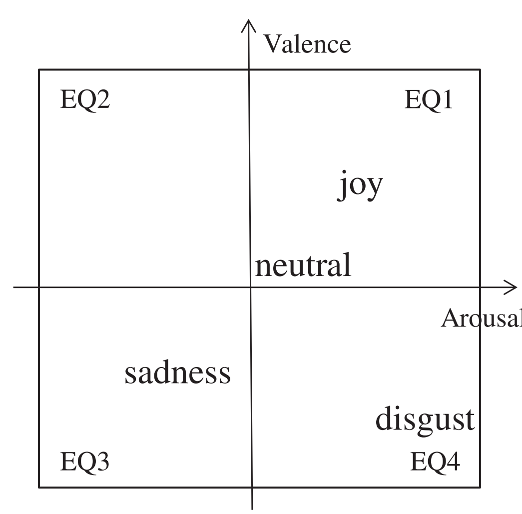

Fig. 2. Emotions based on 2D emotion model. Valence represents the degree of personal pleasure, from negative to positive; arousal expresses the degree of excitement felt by subjects, from calmness to excitement.