ARTICLE IN PRESS

Neural Networks ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Neural Networks

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

2017 Special Issue

Prediction of advertisement preference by fusing EEG response and sentiment analysis

Himaanshu Gauba [a,*], Pradeep Kumar [a], Partha Pratim Roy [a], Priyanka Singh [a], Debi Prosad Dogra [b], Balasubramanian Raman [a]

[a] Department of Computer Science and Engineering, Indian Institute of Technology, Roorkee, India

[b] School of Electrical Sciences, Indian Institute of Technology, Bhubaneswar, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
EEG signal analysis
Multimedia indexing and retrieval
Predictive modeling
Multimodal rating
Sentiment analysis

ABSTRACT

This paper presents a novel approach to predict rating of video-advertisements based on a multimodal framework combining physiological analysis of the user and global sentiment-rating available on the internet. We have fused Electroencephalogram (EEG) waves of user and corresponding global textual comments of the video to understand the user's preference more precisely. In our framework, the users were asked to watch the video-advertisement and simultaneously EEG signals were recorded. Valence scores were obtained using self-report for each video. A higher valence corresponds to intrinsic attractiveness of the user. Furthermore, the multimedia data that comprised of the comments posted by global viewers, were retrieved and processed using Natural Language Processing (NLP) technique for sentiment analysis. Textual contents from review comments were analyzed to obtain a score to understand sentiment nature of the video. A regression technique based on Random forest was used to predict the rating of an advertisement using EEG data. Finally, EEG based rating is combined with NLP-based sentiment score to improve the overall prediction. The study was carried out using 15 video clips of advertisements available online. Twenty five participants were involved in our study to analyze our proposed system. The results are encouraging and these suggest that the proposed multimodal approach can achieve lower RMSE in rating prediction as compared to the prediction using only EEG data.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

With the advancement in information technology, there has been a rapid growth in the use of digital multimedia contents for various applications. Due to the availability of huge amount of multimedia contents, individual's choice and interest play an important role in personalized selection (Hsu, Khabiri, & Caverlee, 2009; Murakami & Ito, 2011). Thus, it is important to develop an advanced rating models to evaluate and predict ratings or rankings of media materials, namely music, videos or television contents. The advertisement clips, available on the internet, must be rated based on people's attraction because majority of the marketing industries rely on these data. However, most of the existing multimedia rating prediction systems are based on viewers comments, interviews or likes/dislikes (Chikkerur, Sundaram, Reisslein, & Karam, 2011; Pang, Lee, & Vaithyanathan, 2002; Shamma, Yew, Kennedy, & Churchill, 2011). Hence, a robust system that can better predict the rating of an advertisement clip is of utmost importance.

The field of neurophysiology explores the signals activated in the brain as a response to multimedia contents (Yang et al., 2015). Electroencephalogram (EEG) is considered as a promising tool for measuring cognitive workload at lower cost with easy handling, wireless connectivity and lower maintenance cost (Stytsenko, Jablonskis, & Prahm, 2011). Such characteristics extend the use of technology from laboratory to general practices. Different prediction models are based on the underlying idea where known emotional states are designed according to the valence scores and arousal of feelings in the viewer's mind while watching the advertisements (Lang, 1990). Various neuroimaging techniques have already been applied to study the cortical and subcortical portions of the brain that get activated while watching pleasant or unpleasant video contents (Wang, Zhu, Wu, & Ji, 2014). However, these works have mostly focused on the analysis of user emotions while watching the videos instead of assessing the quality of videos and user preferences using physiological analysis of brain signals.

* Corresponding author.
E-mail address: gauba.himanshu@gmail.com (H. Gauba).

http://dx.doi.org/10.1016/j.neunet.2017.01.013
0893-6080/© 2017 Elsevier Ltd. All rights reserved.

Please cite this article in press as: Gauba, H., et al., Prediction of advertisement preference by fusing EEG response and sentiment analysis. Neural Networks (2017), http://dx.doi.org/10.1016/j.neunet.2017.01.013

This paper proposes a multimodal approach where we capture the brain signal (EEG data) as a physiological response to the content being watched by the user and it is used as the primary feature. To improve the user prediction, a classical feature is incorporated by analyzing comments posted by global participants. Usually, three major approaches are adopted to predict the ratings of advertising clips. One of them uses the physiological response from the user, where the EEG signals of the users are analyzed while they watch the advertising clips. EEG signals have been found to be quite useful for emotion analysis (Ansari-Asl, Chanel, & Pun, 2007; Chanel, Ansari-Asl, & Pun, 2007; Chanel, Kronegg, Grandjean, & Pun, 2006). Another approach is using multimedia content to analyze the comments published by the users and then predict the nature of the comments to conclude the rating of the observed video (Chikkerur et al., 2011; Siersdorfer, Chelaru, Nejdl, & San Pedro, 2010). The last one is a hybrid approach where features generated from the EEG signals and sentiment analysis of comments are fused to predict the final ratings. Sentiment analysis is the process of computationally identifying the writer's attitude with respect to a given piece of text (Pang & Lee, 2008). The attitude may be his/her judgment or evaluation, affective state (emotional state of the author when writing), or the intended emotional communication (the emotional effect the author wishes to have on the reader) towards a product, or topic.

Usually, the ratings provided by users play an important role in media industry, where these ratings are used to compare different videos and accordingly offer the best rated video content to the perspective viewers. Also, it helps physically challenged persons to understand their emotions better (Kumar, Gauba, Roy, & Dogra, 2016). If the viewer is illiterate or physically challenged (hearing impaired), then a rating of the video can be predicted by showing the video clip and analyzing the corresponding EEG response. Our objective is to analyze and predict the best possible rating of the advertisement videos. Measurements and comparisons drawn using categorical scales are often inappropriate as they may create inconvenience to the user in expressing the popularity of a video (Filippova & Hall, 2011). Meanwhile, it is possible to compare and rate videos on an ordinal scale to express the degree of popularity by the user. The main contributions to the paper are as follows:

1. Firstly, we present a multimodal framework for rating prediction of advertisement videos by fusing EEG signals and sentiment analysis of comments together.
2. Secondly, a rating evaluation has been performed separately for EEG signals and comments data. A multimodal scheme has been proposed to improve the overall rating performance.

The rest of this paper is organized as follows. Section 2 provides an overview of the related work. Next, we detail the methodology used for feature extraction in each modality. Different techniques used for the fusion task are discussed in Section 3. Results obtained using various regression techniques are presented in Section 4. In Section 5, we conclude the paper along with future possibilities.

2. Related work

The field of neurophysiology substantiates the significance of Human Visual System (HVS) and its sensitivity towards the visual fidelity in day-to-day environment. Variety of studies have already been carried out to understand the responses of a human brain towards different emotion reflexes and photo-realistic image sequences (Koelstra et al., 2012; Soleymani, Pantic, & Pun, 2012; Wang, Sourina, & Nguyen, 2010). Researchers have explored EEG analysis and NLP based sentiment analysis for designing efficient and robust prediction models of human preference and scoring. NLP is a technique which allows computers to analyze, understand and extract meaning from linguistic communications of the humans (Cambria & White, 2014). Using NLP, important tasks can be performed easily including automatic summarization, translation, named entity recognition, relationship extraction, sentiment analysis, speech recognition, or topic segmentation. Following sections describe some of the work performed in this field of research.

2.1. Physiological analysis using EEG signal

EEG signals have been widely used by researchers for building brain control interface (BCI) applications including biometric (Kaur, Singh, & Roy, 0000; Kostílek & Šťastný, 2012), gaming (Wang et al., 2010), emotion analysis (Koelstra et al., 2012), and medical applications (Dixit, Popescu, Bagić, Ghearing, & Hendrickson, 2013). Yazdani, Lee, Vesin, and Ebrahimi (2012) proposed an implicit emotional tagging framework for videos using EEG signals. The authors recorded EEG signals of 25 subjects while the subjects were watching video clips. Next, they classified six emotions using Bayesian Linear Discriminant Analysis (BLDA) classifier with an average accuracy of 80.19%. In Mustafa, Guthe, and Magnor (2012), the authors proposed a methodology to predict the presence of artifacts in videos using EEG event-related potentials. The authors used wavelet based classification for evaluating the performance of EEG signals where an accuracy of 85% was recorded in prediction of artifacts. This study revealed that, EEG signals are very useful in physiological assessment of the multimedia contents. Scholler et al. (2012) proposed an approach for measuring video quality using EEG signals. The authors obtained Event-Related Potentials (ERPs) by aligning the EEG signals and next they analyzed the positive voltage change in EEG data whenever there was a change in the video quality with a latency range of 400–600 ms.

A music prediction system using EEG signals was proposed in Hadjidimitriou and Hadjileontiadis (2012). The authors recorded EEG signals of 9 subjects while listening to music, where an accuracy of 86.53% was recorded on time–frequency based features when tested with k-Nearest Neighbor (k-NN) classifiers. Khushaba et al. (2013) explored the brain signals while performing a choice task designed to elicit product preferences. The authors recorded EEG signals of 18 participants with Tobii-Studio eye tracker to relate the EEG data to a specific choice option. They performed power spectral analysis with Fast Fourier Transform (FFT) features on EEG signals where a change in frontal, temporal and occipital regions was observed while users indicated their preferences. In Daimi and Saha (2014), authors proposed an emotion recognition approach using Dual-Tree Complex Wavelet Packet Transform (DT-CWPT) time–frequency features from EEG signals while watching music videos involving 32 participants. The authors used Singular Value Decomposition (SVD), QR factorization with column pivoting (QRcp) and F-Ratio based feature selection approaches before classifying the emotions using Support Vector Machine (SVM) classifier where an average classification rate of 65.3% was reported.

2.2. Sentiment analysis

Various researchers have worked on rating based sentiment analysis. In Thelwall, Buckley, Paltoglou, Cai, and Kappas (2010), Thelwall et al. developed an algorithm popularly known as SentiStrength to detect the strength of sentiments on MySpace comments. They used an initial set of 2600 human classified MySpace comments and evaluated the model using 10-fold cross validation on a test set of 1041 comments which recorded accuracies of 60.6% and 72.8% to predict positive emotion and negative emotion respectively, both based on strength between 1 and 5. The authors in Shieh and Zamoshchin (0000) proposed a study focusing on

Please cite this article in press as: Gauba, H., et al., Prediction of advertisement preference by fusing EEG response and sentiment analysis. Neural Networks (2017), http://dx.doi.org/10.1016/j.neunet.2017.01.013

Download English Version:

https://daneshyari.com/en/article/4946670

Download Persian Version:

https://daneshyari.com/article/4946670

Daneshyari.com