Brain and Cognition 101 (2015) 1–11

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Music-induced emotions can be predicted from a combination of brain activity and acoustic features

Ian Daly[a,*], Duncan Williams[b], James Hallowell[a], Faustina Hwang[a], Alexis Kirke[b], Asad Malik[a], James Weaver[a], Eduardo Miranda[b], Slawomir J. Nasuto[a]

[a] Brain Embodiment Lab, School of Systems Engineering, University of Reading, Reading, UK
[b] Interdisciplinary Centre for Music Research, University of Plymouth, Plymouth, UK

ARTICLE INFO

Article history:
Received 23 January 2015
Revised 3 August 2015
Accepted 4 August 2015
Available online 3 November 2015

Keywords:
Music
Affective state prediction
EEG
Acoustic features
Machine learning

ABSTRACT

It is widely acknowledged that music can communicate and induce a wide range of emotions in the listener. However, music is a highly-complex audio signal composed of a wide range of complex time- and frequency-varying components. Additionally, music-induced emotions are known to differ greatly between listeners. Therefore, it is not immediately clear what emotions will be induced in a given individual by a piece of music.

We attempt to predict the music-induced emotional response in a listener by measuring the activity in the listeners electroencephalogram (EEG). We combine these measures with acoustic descriptors of the music, an approach that allows us to consider music as a complex set of time-varying acoustic features, independently of any specific music theory. Regression models are found which allow us to predict the music-induced emotions of our participants with a correlation between the actual and predicted responses of up to $r = 0.234, p < 0.001$.

This regression fit suggests that over 20% of the variance of the participant's music induced emotions can be predicted by their neural activity and the properties of the music. Given the large amount of noise, non-stationarity, and non-linearity in both EEG and music, this is an encouraging result. Additionally, the combination of measures of brain activity and acoustic features describing the music played to our participants allows us to predict music-induced emotions with significantly higher accuracies than either feature type alone ($p < 0.01$).

1. Introduction

Music is widely acknowledged to be a powerful method for emotional communication, capable of eliciting a range of different emotional responses in the listener, such as joy, excitement, and fear (Scherer, 2004). Subsequently, music therapy may be used as a tool for treatment of emotional disorders such as depression (Maratos, Gold, Wang, & Crawford, 2008).

Music therapy is a health intervention in which the music therapist uses music as a tool to help their patient with their physical and/or mental health problems (Bradt, Magee, Dileo, Wheeler, & McGilloway, 2010; Erkkilä et al., 2011; McDermott, Crellin, Ridder, & Orrell, 2013). For example, in the treatment of depression music therapy has been reported to significantly improve mood when compared to standard care alone (Maratos et al., 2008) (for example antidepressant drugs alone vs. antidepressant drugs and music therapy (Chen, 1992)).

The music used in music therapy is selected by the therapist based upon a combination of the therapists evaluation of their patients current psychological state, the therapists expertise and experience, and the properties of the music that the therapist judges will be beneficial to the patient (Tamplin & Baker, 2006).

In order to select an appropriate piece of music for use in music therapy it is necessary to predict how the individual is likely to react to that piece of music. However, it is a considerable challenge to predict the potential reaction of an individual to a piece of music they have not previously heard before. There are large inter-personal differences in emotions induced by listening to a piece of music, which result from both the music itself and the participant's own previous and current mental states (Hunter, Schellenberg, & Schimmack, 2010).

These inter-person differences are a result of a wide range of influences and include the individuals prior experiences, their current mood, and a range of other factors both internal to the person

* Corresponding author.
E-mail address: i.daly@reading.ac.uk (I. Daly).

http://dx.doi.org/10.1016/j.bandc.2015.08.003
0278-2626/© 2015 Elsevier Inc. All rights reserved.

and external to them. Broadly speaking, a persons emotional response to a piece of music can be said to be a function of both the music itself and of the individual.

When considering the piece of music, a number of models have been proposed for the relationships between musical structure and syntax and both the perceived and/or induced emotional responses of a listener (for example, Thompson & Robitaille (1992), Schubert (1999), Gabrielsson & Lindström (2001), Gabrielsson & Juslin (2003)).

For example, in Russell (1980) and Livingstone and Thompson (2009) the circumplex model of affect and its relationship to musical descriptors is described. In this model, emotional responses have been plotted across two continuous axes, arousal (excitement) and valence (pleasantness), ranging from low to high. Musical descriptors drawn from music theory, such as tempo or modality, are plotted in this space.

However, while this model is intuitive and can be informative about perceptions of the role individual features of music theory may play in emotional responses, it is not complete. First, due to the very large inter-person differences in music-induced emotions, music features are not likely, by themselves, to be good predictors of emotional responses to music. This may be due to a variety of factors, including individual preferences for particular pieces of music, prior experience of music induced emotions, or a participants physiological state as they listen to music (Peretz, Aube, & Armony, 2013).

An alternative approach that has been adopted is to attempt to use physiological measurements of the participant as correlates of their emotional responses (Craig, 2005; Kim & André, 2008; Schmidt & Trainor, 2001). Patterns in these physiological measurements can be identified and used to attempt to identify a participants emotional response to a piece of music.

An example of this is the use of electrocardiogram (ECG) signals to identify emotional responses to music (Kim & André, 2008). As music causes listeners to become more excited, this can lead to increases in heart rate, which is reflected in the recorded ECG signal and, subsequently, classified (Kim & André, 2008). Other physiological measures which have been adopted to identify music-induced emotions include the galvanic skin response (GSR) (Craig, 2005), the electromyogram (EMG) recorded from the facial muscles (Lundqvist, Carlsson, Hilmersson, & Juslin, 2008), and respiration rate (Etzel, Johnsen, Dickerson, Tranel, & Adolphs, 2006).

Alternatively, a number of researchers have explored various indices of neural activity as a measure of music-induced emotion. This may be done by, for example, the use of the electroencephalogram (EEG) (Schmidt & Trainor, 2001).

Measures of activity in the EEG which have been reported to relate to music-induced emotion include asymmetry of activity within the alpha band over the prefrontal cortex (Schmidt & Trainor, 2001), measures of prefrontal asymmetry in the beta frequency band, and measures of connectivity between prefrontal and occipital cortical areas (Daly et al., 2014).

The influence of music-induced emotion on the EEG is derived from the neurobiological mechanisms mediating interactions of music with emotions (Peretz, 2009, chap. 5). Music is thought to engage a diverse network of neural structures, with no single pathway bearing responsibility for music-induced emotions. This is evidenced by the lack of reports of selective loss of all music-induced emotions due to brain injury, contrasting with the prevalence of evidence for selective loss of some music-induced emotions. For example, 'scary' and 'sad' music-induced emotions may be lost after damage to the amygdala (Gosselin, 2005; Gosselin, Peretz, Johnsen, & Adolphs, 2007) and impaired by Parkinson's disease (van Tricht, Smeding, Speelman, & Schmand, 2010). This is also evidenced by findings that preferred musical styles engage a listener's default mode network most strongly (Wilkins, Hodges, Laurienti, Steen, & Burdette, 2014).

As a consequence, music-induced emotions relate to a range of particular effects in the EEG. These include inter-hemispheric differences in EEG activity levels (Daly et al., 2014; Flores-Gutiérrez et al., 2007; Schmidt & Trainor, 2001) or changes in EEG over specific regions, such as the pre-frontal cortex (Lin et al., 2010). Taken together, it has been suggested that musical emotions engage a network of both cortical and sub-cortical regions, which produces a range of effects in the EEG (Peretz, 2009, chap. 5).

These effects are widely known to differ between individuals (Hunter et al., 2010). This can occur for a variety of reasons, including musical preferences (Bauer, Kreutz, & Herrmann, 2015), age (Daly et al., 2014), and emotional intelligence (Jausovec & Jausovec, 2005). Additionally, the EEG is known to be a noisy, non-stationary signal (Daly et al., 2012). Taken together this makes reliable identification of music-induced emotions from the EEG a very challenging problem.

Therefore, we suggest that a combination of both physiological measures of the listener and acoustic properties of the music may be used to effectively predict emotional responses to a piece of music. Specifically, we hypothesise that a combination of EEG measures and the acoustic properties of the music may be used to predict the emotional response they will report experiencing while listening to the music.

We play a series of musical clips to a group of participants, while recording their EEG. We then extract descriptive features of both the acoustic properties of the music and the participant's EEG. We attempt to use these features to train a regression model to predict the music-induced emotional responses of the participants.

2. Methods

2.1. Measurements

Thirty-one individuals between the ages of 18–66 (median 35, 18 female) participated in the study (previously detailed in (Daly et al., 2014)). All participants were healthy adults who did not report having any mental health, mood, or psychiatric problems. All participants had normal, or corrected to normal, hearing and vision. Twenty-nine of the participants were right handed (no significant differences were found in the results from the two left handed participants). The electroencephalogram (EEG) was recorded from each participant from 19 channels positioned according to the International 10/20 system for electrode placement.

The participants each listened to 40 pieces of music, which were uniformly drawn from a set of 110 excerpts from film scores. The stimuli were taken from a dataset of musical pieces chosen with the specific purpose of inducing emotional responses in the listener (Eerola & Vuoskoski, 2010).

Each musical clip was played for 12 s, as described in Daly et al. (2014), during which the participants were instructed to look at the screen and listen to the music without moving. They were then asked a series of 8 randomly-ordered Likert questions designed to identify the level of emotional response along 8 axes induced in them by the music.

These 8 axes allowed the participants to report their music-induced emotions in terms of pleasantness, energy, sadness, anger, tenderness, happiness, fear, and tension. However, as several of these categories are likely to be highly correlated, a principal component analysis (PCA) was used in order to identify a reduced set of categories. Three principal components (PCs) were identified, which explained >75% of the variance of the participant's