University of Kent

# Kent Academic Repository

**Langroudi, George, Jordanous, Anna and Li, Ling (2018) *Music Emotion Capture: sonifying emotions in EEG data.* In: Emotion Modelling and Detection in Social Media and Online Interaction symposium at the AISB 2018 Convention, 4-6 April 2018, Liverpool, UK.**

## Downloaded from
https://kar.kent.ac.uk/66564/ The University of Kent's Academic Repository KAR

## The version of record is available from
http://collab.di.uniba.it/aisbemotions/

## This document version
Author's Accepted Manuscript

## DOI for this version

## Licence for this version
UNSPECIFIED

## Additional information

## Versions of research works

### Versions of Record
If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

### Author Accepted Manuscripts
If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal* , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

## Enquiries
If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).

# Music Emotion Capture: sonifying emotions in EEG data

**George Langroudi** and **Anna Jordanous** and **Ling Li**[1]

**Abstract.** People's emotions are not always obviously detectable, due to difficulties expressing emotions, or geographic distance (e.g. if people are communicating online). There are also many occasions where it would be useful for a computer to be able to detect users' emotions and respond to them appropriately. A person's brain activity gives vital clues as to emotions they are experiencing at any one time. The aim of this project is to detect, model and sonify people's emotions. To achieve this, there are two tasks: (1) to detect emotions based on current brain activity as measured by an EEG device; (2) to play appropriate music in real-time, representing the current emotional state of the user. Here we report a pilot study implementing the Music Emotion Capture system. In future work we plan to improve how this project performs emotion detection through EEG, and to generate new music based on emotion-based characteristics of music. Potential applications arise in collaborative/assistive software and brain-computer interfaces for non-verbal communication.

## 1 INTRODUCTION

It is not always straightforward to detect people's emotions; people may be unable to express their emotions in usual ways due to physical or other limitations, or emotions may be difficult to detect if people are geographically separated e.g. interacting online. There are also many occasions where it would be useful for a computer to be able to detect users' emotions, particularly where the computer is working collaboratively with a user. If a computer can accurately detect a user's emotional state, it may be able to respond appropriately, demonstrating some 'empathy' with its user - which is likely to lead to a more positive user experience overall.

The aim of this project is to detect, model and musically sonify people's emotions in real time. To achieve this, there are two steps: (1) to detect emotions based on current brain activity as measured by an EEG device; (2) to generate appropriately matching music in real-time, representing the current emotional state of the user.

The pilot study reported in this work-in-progress paper uses a version of Russell's circumplex model of emotions [14] updated by Scherer [16], alongside a music database tagged with relevant metadata, to implement a music player in C# that interacts with emotions experienced by someone wearing an Emotiv EEG headset. The Emotiv headset device simplifies EEG signal capture, making EEG data more accessible for experiments such as this. Section 2 describes the Emotiv headset, as well as previous research on modelling emotions and expressing emotions musically. Section 3 describes implementation of our proof-of-concept Music Emotion Capture system.

Evaluation of software which responds to people's emotions is a tricky issue that needs careful handling, for ethical reasons. Evaluation (Section 4) of this work in progress has currently been limited to system evaluation, as we would like to improve some aspects of the system implementation before we evaluate the functionality of Music Emotion Capture with users. Section 4 describes how, in future work, we plan to improve how this project performs emotion detection through EEG, and how new music can be generated based on emotion-related characteristics of music.

We see a large number of potential applications of the ideas in this work. Section 4.3 describes how this work could be applied in music therapy or interactive software, mainly centred around applications with brain-computer interfaces for enhancing non-verbal communication and interaction, e.g. for improving computer-human communication/empathy in collaboration, or for customising HCI to user experience, or enabling people in a vegetative state ('locked-in') with a new mechanism for expressing the emotions they are feeling.

## 2 BACKGROUND

### 2.1 Computational modelling of emotions

How can human emotions be modelled computationally? For computers to be able to use data about human emotions, a computational model needs to be able to represent emotions accurately in a machine-readable format amenable to reasoning and analysis.

There are two leading theories for representing emotion: Ekman's six basic emotions [2] and Russell's circumplex model [14].[2]

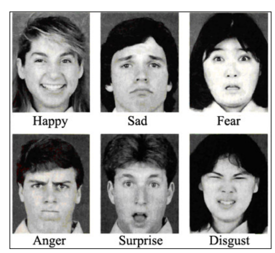


**Figure 1.** Ekman's basic emotions (from MIT Brain & Cognitive Sciences)

**Ekman's model of six basic emotions** Ekman has provided a simple set of six[3] dimensions for the study of emotions in psychology [2]: *happy, sad, fear, anger, surprise, disgust* (see Figure 1)[4].

Ekman's model is highly applicable for work focusing on detecting emotions through analysis of facial features. One issue is that, although now more widely interpreted, it was originally derived specifically around the study of emotions as revealed in facial expressions,

[1] School of Computing, University of Kent, Medway campus, UK, email: George.Langroudi@outlook.com ; [A.K.Jordanous / C.Li]@kent.ac.uk

[2] This is not to say that other models do not exist, e.g. [11].

[3] Though this model is widely treated as having six dimensions, we acknowledge that Ekman later argued for (*contempt*) to be added [10].

[4] Image source `https://ocw.mit.edu/courses/brain-and-cognitive-sciences/9-00sc-introduction-to-psychology-fall-2011/emotion-motivation/discussion-emotion/diss_img.jpg`.

as Figure 1 illustrates. It has been questioned[5] whether facial expressions (as analysed by Ekman) are fully representative of emotions, and whether Ekman's particular model is too focused on one culture's expression of emotions, a Western expression, without being applicable to other cultures [15]. Hence for this work's focus, on detection of emotions at the brain activity level, rather than through facial expression, alternative options should be considered too.

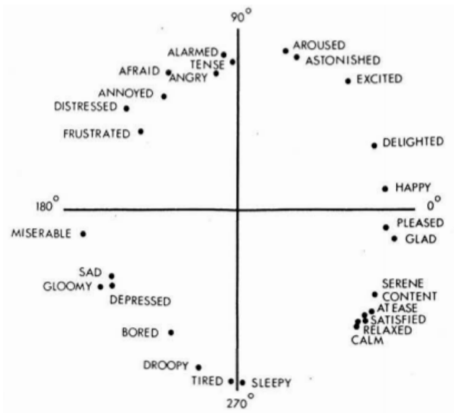


**Figure 2.** Russell's Circumplex model of affect [14]

**Russell's circumplex model** Predating Ekman's model, Russell's circumplex model [14] spatially represents emotions and their relations to each other. Emotions are plotted on a two-dimensional graph: see Figure 2. The two dimensions are *valence* and *arousal*:

- **valence**: the extent to which an emotion is considered positive or negative; e.g. a positive valence emotion on Russell's model is *HAPPY*, and a negative valence emotion is *MISERABLE*.
- **arousal** The extent to which an emotion represents a psychologically activated state; e.g. a high arousal emotion on Russell's model is *ALARMED*, and a low arousal emotion is *TIRED*.

Unlike Ekman's model, Russell's model acknowledges interrelations and dependencies between emotions; e.g., if one is feeling happy, there is an expectation that feelings of sadness are likely to be correspondingly low or non-existent. On the Russell model, similarity-based relationships between emotions can be analysed using distance. Similar emotions are plotted close to each other and dissimilar emotions far away from each other. For example, in Figure 2 the emotions *MISERABLE/SAD/GLOOMY*, all similar in sentiment, are positioned close to each other, but far away from a cluster of more positive emotions such as *HAPPY/PLEASED/GLAD*, which are opposed in meaning to the first cluster of emotions.

**Scherer's update to the Russell model** Russell's model gives emotions neatly arranged in a circular pattern, such that the only emotions expressed are those that fall in (x, y) co-ordinates in the circumplex graph that roughly satisfy the circle equation of $(xh)^2 + (yk)^2 = r^2$. No other emotions are represented in this Russell model. But what if we want to represent the emotions being experienced by someone at a state of, say, 0 arousal and with only a small positive level of valence?

[5] We note this comes from Russell, the author of another emotions model.

Russell's model was updated by Scherer [16] to represent a greater variety of emotions than those at the edge of the circle: see Figure 3.

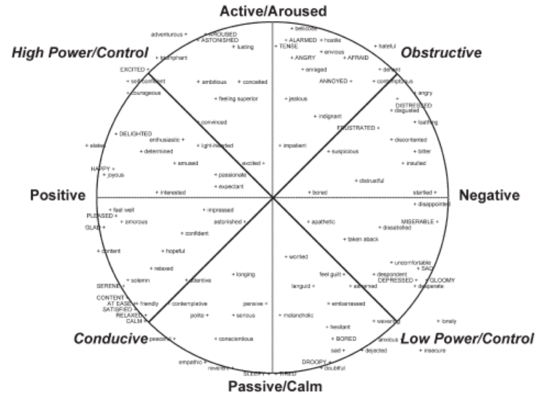


**Figure 3.** Scherer's updates to the Russell circumplex model [16]

Scherer's updated model was applied for measuring the emotional content of blog posts by [13]. During this application, the graphical model was converted into quantitative data and reported in [13]. This becomes very useful for computational implementation; we now have a quantified version of a widely-cited model of emotions that can be experienced, which has been successfully applied for modelling emotional content. The model is based on two simple dimensions of valence and arousal, and is representative of a wide range of emotions, across the full ranges of these dimensions. Hence we chose the Russell model (as updated by [16, 13]) for this work.

## 2.2 Detecting emotions through EEG

The Emotiv EEG headset detects and measures EEG signals from the scalp. It allows data acquisition of 14 channels at 128Hz per channel, which is sufficient to capture the brain activity below 64Hz. This frequency range covers the frequency bands alpha, beta, theta and delta; i.e. the majority of the EEG used in clinical practice. Unlike clinical EEG measuring equipment, the headset can be prepared and used fairly quickly for experimental purposes. This accessibility comes with a trade-off in accuracy: as it has fewer channels than some high spec EEG equipment, with less rich data (especially when the topological distribution patterns are of interest). This is, though, a more accessible way of obtaining EEG data.

Emotiv's SDK gives access to four measurements: *Engagement, Excitement, Meditation* and *Frustration*. Each measurement is given as a float between 0 and 1. The measurements offered by Emotiv are useful indicators to use for our pilot study, but there is some discussion concerning these detectors' validity and reliability [5]. We discuss alternative ways to detect emotion-related data in Section 4.

## 2.3 Musical expression of emotions

Work on expressing emotions through music has been a topic of study for decades, e.g. the seminal work by Meyer [12]. Juslin & Laukka [7] reviewed several studies covering the acoustic expression of emotions through music and through speech. They found evidence in the reviewed studies that music can be used to express emotions, in a manner similar to how we use vocal cues. For example, as shown in Figure 4, the expression of the emotion 'Anger' is often expressed via music with fast tempos,[6] high variability in volume (sound level) and in the range of musical notes (pitches) used, as well as other acoustic patterns. In contrast, musical expressions of the emotion ‘Sadness’ are typically within music with slow tempos and low variability in volume/sound level and in the range of notes/pitches, as well as having other musical features (see Figure 4).

For our work, generation of music through this set of musical patterns would be highly interesting to study: a valid standalone research project in its own right (e.g. [3]). For faster progress in this pilot study, though, we also investigated if there were precomposed, existing sources of musical data that we could use, with annotations representing emotions associated with each piece of music.[7] The *Emotion in Music Database* (http://cvml.unige.ch/databases/emoMusic/) [18] consists of 744 musical extracts (45 second snippets of music from the Free Music Archive (http://freemusicarchive.org/). As described by [18], each extract is provided with metadata annotations of crowd-sourced ratings of both valence and arousal, gathered by people submitting ratings as they listened to the music and these ratings being averaged out over the full extract and over different raters, and normalised to a range of [-1, 1]. The approach of [18] is not without its flaws: notably, taking average ratings without use of accompanying measures of standard deviation/variance hides subtleties in variation of ratings over time. For our work, though: this is a dataset that has been deployed in previous research, is openly available and fits with our decision to use a valence- and arousal-based model of emotion.

*Summary of Cross-Modal Patterns of Acoustic Cues for Discrete Emotions*

| Emotion | Acoustic cues (vocal expression/music performance) |
|---|---|
| Anger | Fast speech rate/tempo, high voice intensity/sound level, much voice intensity/sound level variability, much high-frequency energy, high F0/pitch level, much F0/pitch variability, rising F0/pitch contour, fast voice onsets/tone attacks, and microstructural irregularity |
| Fear | Fast speech rate/tempo, low voice intensity/sound level (except in panic fear), much voice intensity/sound level variability, little high-frequency energy, high F0/pitch level, little F0/pitch variability, rising F0/pitch contour, and a lot of microstructural irregularity |
| Happiness | Fast speech rate/tempo, medium–high voice intensity/sound level, medium high-frequency energy, high F0/pitch level, much F0/pitch variability, rising F0/pitch contour, fast voice onsets/tone attacks, and very little microstructural regularity |
| Sadness | Slow speech rate/tempo, low voice intensity/sound level, little voice intensity/sound level variability, little high-frequency energy, low F0/pitch level, little F0/pitch variability, falling F0/pitch contour, slow voice onsets/tone attacks, and microstructural irregularity |
| Tenderness | Slow speech rate/tempo, low voice intensity/sound level, little voice intensity/sound level variability, little high-frequency energy, low F0/pitch level, little F0/pitch variability, falling F0/pitch contours, slow voice onsets/tone attacks, and microstructural regularity |

*Note.* F0 = fundamental frequency.

**Figure 4.** How music and voice express different emotions (from [7])

## 3 IMPLEMENTATION

The Music Emotion Capture software was written in C# using Emotiv’s Professional SDK. The interface was made using the OpenTK library for C# to make a basic OpenGL display.

### 3.1 Detecting valence and arousal

The Emotiv SDK translates the signals from the Emotiv EEG headset to give measurements of the extent to which each of the four basic emotions (Excitement, Engagement, Meditation, Frustration) is being experienced by the user at a given time. Using Russell’s circumplex model, (only) Valence and Arousal measurements are needed. We used Engagement as a proxy measure of arousal, and used Excitement as a proxy measure for Valence. We acknowledge the potential flaws from these choices of proxy; while not perfect translations, these compromises allow us to progress in our aims.

[6] The tempo of a piece of music is the speed at which a regular metrical beat or pulse occurs; e.g. 180 beats per minute (bpm) would be a fast tempo in Western Classical music, whereas 45 bpm would be a slow tempo.

[7] However we discuss in Section 4 our intentions to develop this pilot study with real-time generation of music based on the findings of [7].

### 3.2 Mapping emotions to music

Emotiv will give us a constantly updated float between 0 and 1 for each of the four basic emotions. We mapped valence and arousal to the X and Y axis respectively of a graph (see Figure 5), plotting a point on the graph, and then used the emotion mapping from [13].

When the Music Emotion Capture software runs, the software continuously takes EEG measurements. The EEG measurements are continually translated into valence-arousal co-ordinates, to generate data about possible emotional states. In the terminal window in Figure 5 you can see this data: the emotions in an ordered list (the closest emotion at the top) as well as the ‘distance’ between that emotion and the user’s current emotion (see Section 2.1).

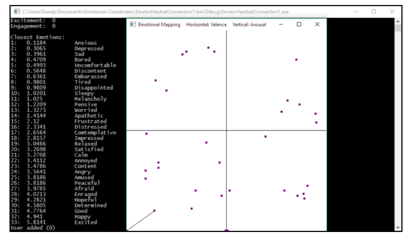

**Figure 5.** Output of the Music Emotion Capture pilot program, with the terminal showing emotions currently being detected by the system at a given point in time, ordered by relevance, and the graph plotting each of the Emotional Coordinates (x axis: valence, y axis: arousal)

Using the Emotion in Music Database described in Section 2.3 [18] we mapped each song to a location then checked for the song that matches the users emotion closest every 10 seconds

## 4 EVALUATION AND FUTURE WORK

Evaluation to date has been through white-box tests, e.g. testing with different settings/checking that the code functions as expected when different emotions are detected. As of yet, we have not yet run experiments with subjects; this will happen when the current pilot study is developed further. The evaluation methodology is crucial for this project, especially how emotions relate with music. At that point, we intend to evaluate the project by testing people with the headset, asking people whether they feel the music they hear matches their emotional state. We will also ask participants to describe their emotional state at various points in time during the study, so that we have data to evaluate whether self-reported emotion match what the headset was picking up at that time. Emotion databases are also available e.g. AAAC, RECOLA,[8] which we see as a foundation for data-driven evaluation, to augment user-reported emotion-music associations.

This pilot study represents work in progress; the above user evaluations form a key part of future work. Certain design choice compromises were made during this pilot study (as described above). We would like to carry out some of the following future work improving the system before we then conduct a formal user evaluation.

### 4.1 Improved Valence/Arousal detection

In Section 2.2 we noted that although Emotiv provide easy access to four emotional measurements via their headset, (1) questions have been raised regarding the validity of these measurements and (2) our

[8] AAAC: http://emotion-research.net/wiki/Databases; RECOLA: http://diuf.unifr.ch/diva/recola/index.html.

choice of two of these as valence and arousal proxies is potentially flawed. Research has been done on improving Emotiv-based emotion recognition by fusing biometric and eye tracking technologies [9]. Also, other EEG based arousal/valence work indicated 82% accuracy for automatic classification of positive, negative and neutral valence based on film clip viewing, using frequency feature and its dynamics in EEG [1]. We envisage new feature extraction and pattern recognition methods, which have updated the state-of-the-art in many areas. Therefore we would like to investigate work calculating valence and arousal from raw EEG signals (e.g. [4]). Moreover, we will investigate different approaches for quantifying EEG patterns in emotions with time-frequency features, synchrony-based features [8] and perform pattern recognition considering recent work achieved with deep neural network architectures and generative models in identifying emotional states [19] and generating emotion-specific music.

### 4.2 Generating music using the user's emotions

As described in Section 2.3, Juslin & Laukka [7] review how acoustic cues in music (and in speech) can be used to express different emotions. They conclude from their review that: "emotion-specific patterns of acoustic cues used to communicate each emotion ... are generally consistent with K. R. Scherer's (1986) theoretical predictions." [the predictions by Scherer that formed the basis for [16]]. Given that the pilot study reported above is based on Scherer's development of the Russell circumplex model of emotions (Section 3), the guidelines from [7] are a promising direction for future work.

### 4.3 Potential applications

Many practical applications of this work emerge from the potential of **brain-computer interfaces (BCI) for enhancing non-verbal communication of emotion**. Several interesting paths of investigation emerge for Human-Computer Interaction (HCI) issues around how collaborative software can be **better customised to the user experience**. If the user is currently experiencing, say, high levels of frustration, the software can interact differently to if the user is experiencing high levels of happiness.[9] For example, a BCI could be used to enhance the quality and accuracy of computer-human communication during collaboration, e.g. with a Creativity Support Tool [17], or with software designed to be a creative partner in a *co-creative* scenario (a scenario which users sometimes find more troubling to navigate [6]).

Other applications in the domain of **assistive technologies** include enabling people in a vegetative state ('locked-in') to communicate emotions they are feeling, either directly to software or to other people via the sonification of their emotions. **Music therapy** applications can also be envisaged: playing music to people which is aimed at bringing their emotional state into the region of positive emotions.

There is also potential application to the task of **playlist creation**. Being able to detect a person's current mood may allow musical software to better recommend songs. For example, if the software detected that the user was relaxed it may want to recommend a similarly relaxing playlist. With the possibility of more portable, functional EEGs[10] one can imagine music players on a user's phone that adjust its suggestions to adapt to the user's current state.

[9] This idea stems from conversations with composer Oded Ben Tal.

[10] Developments look promising: see https://kokoon.io/ ; https://spectrum.ieee.org/the-human-os/biomedical/devices/wireless-earbuds-will-record-your-eeg-send-brainwave-data-to-your-phone .

## 5 CONCLUSION

This work-in-progress paper reports a pilot study which enables us to detect, model and musically sonify the emotions of people who interact with our EEG-based Music Emotion Capture software. An Emotiv headset lets us gather EEG data for translation into music representing the user's current emotions, building upon the Russell circumplex model of emotions. There are many exciting avenues for potential applications, following further work and full evaluation.

## ACKNOWLEDGEMENTS

We would like to thank members of the Data Science group at the School of Computing, University of Kent, in particular Palani Ramaswamy, for their useful comments.

## REFERENCES

[1] L. Brown, B. Grundlehner, and J. Penders, 'Towards wireless emotional valence detection from EEG', in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2188–2191, Boston, MA, (2011).

[2] P. Ekman, 'An argument for basic emotions', *Cognition & Emotion*, **6**(3), 169–200, (1992).

[3] S. Elisha and T. Weyde, 'Exploration of emotion-based cross-modal mappings for generating music for videos', in *2nd Conference on Computer Simulation of Musical Creativity*, Milton Keynes, UK, (2017).

[4] G. Garcia-Molina, T. Tsoneva, and A. Nijholt, 'Emotional brain–computer interfaces', *International Journal of Autonomous and Adaptive Communications Systems*, **6**(1), 9–25, (2013).

[5] T Harrison. The Emotiv mind: investigating the accuracy of the Emotiv EPOC in identifying emotions and its use in an intelligent tutoring system. Honours report, University of Canterbury, New Zealand, 2013.

[6] A. Jordanous, 'Co-creativity and Perceptions of Computational Agents in Co-creativity', in *Eighth International Conference on Computational Creativity (ICCC'17)*, pp. 159–166, Atlanta, GA, (2017). ACC.

[7] P. N. Juslin and P. Laukka, 'Communication of emotions in vocal expression and music performance: different channels, same code?', *Psychological Bulletin*, **129**(5), (2003).

[8] L. Li, A. Witon, S. Marcora, H. Bowman, and D. P. Mandic, 'EEG-based brain connectivity analysis of states of unawareness', in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pp. 1002–1005. IEEE, (2014).

[9] J-M. López-Gil, J. Virgili-Gomá, R. Gil, T. Guilera, I. Batalla, J. Soler-González, and R. García, 'Method for improving eeg based emotion recognition by combining it with synchronized biometric and eye tracking technologies in a non-invasive and low cost way', *Frontiers in Computational Neuroscience*, **10**, 85, (2016).

[10] David Matsumoto and Paul Ekman, 'The relationship among expressions, labels, and descriptions of contempt.', *Journal of personality and social psychology*, **87**(4), 529, (2004).

[11] Albert Mehrabian, 'Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament', *Current Psychology*, **14**(4), 261–292, (1996).

[12] L. B. Meyer, *Emotion and meaning in music*, University of Chicago Press, Chicago, 1956.

[13] G. Paltoglou and M. Thelwall, 'Seeing stars of valence and arousal in blog posts', *IEEE Trans. Affective Computing*, **4**(1), 116–123, (2013).

[14] J. A. Russell, 'A circumplex model of affect', *Journal of Personality and Social Psychology*, **39**, 1161–78, (1980).

[15] J. A. Russell. Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies., 1994.

[16] K. R. Scherer, 'What are emotions? and how can they be measured?', *Social Science Information*, **44**(4), 695–729, (2005).

[17] B Shneiderman, 'Creativity support tools: Accelerating discovery and innovation', *Commun. ACM*, **50**(12), 20–32, (December 2007).

[18] M. Soleymani, M. N. Caro, E. M. Schmidt, C-Y. Sha, and Y-H. Yang, '1000 songs for emotional analysis of music', in *Procs 2nd ACM international workshop on Crowdsourcing for multimedia*, (2013).

[19] A. Supratak, S. Schneider, H. Dong, L. Li, and Y. Guo, 'Towards desynchronization detection in biosignals', in *NIPS Time Series Workshop*, Long Beach, CA, (2017).