University of Twente

Brain Computer Interfacing
201600078

PORTFOLIO

Authors:
Qiurui Chen (1988476)

Course instructor:
dr. M. Poel

July 8, 2019

UNIVERSITY
OF TWENTE.

Contents

1 Research proposal **1**
1.1 Introduction . . . 1
1.2 Research questions . . . 1
1.3 Theoretical background and literature . . . 2
1.4 Research design . . . 3
1.5 Data collection and analysis . . . 4
1.6 Resources . . . 4
1.7 Timeline . . . 4

2 Experiments **4**
2.1 Data Processing . . . 5

3 Results **5**
3.1 Emotion classification . . . 5
3.2 Music generation . . . 6

4 Conclusion and Discussion **6**

5 Suggestions for future work **7**

This section describes the final project that is conducted in the last part of the BCI course. First, a research proposal was written. Then the actual research was carried out, results are documented, and conclusions are drawn later in this section.

1 Research proposal

1.1 Introduction

A lot if information about brain processes can be interpreted from EEG signals. Emotion recognition from EEG signals requires advanced interpretation methods, since emotions are elicited deep inside the brain. Emotions reflect mental states and psycho-physiological expressions, and therefore have a strong connection with physiological signals including brainwaves. Previous research has shown that various algorithms can reveal meaningful information about the emotional state of subjects [3]. Emotions can be evoked by various types of stimuli, including music. Music is known to elicit a wide variety of emotions very powerfully, according to Koelsch [4]. The goal of this research is to detect, interpret, and sonify people's emotions. The aim is to achieve this by composing music based on a person's emotional state extracted from EEG signals.

This research is considered relevant for multiple reasons. First, we will explore how emotional states can be detected from EEG signals. Previous research proposes various and inconclusive ways to determine emotions from brain signals (see 1.3). So for us the challenge is to find a robust, accurate and reliable way to detect emotions. Second, music generation according to a one's emotional state is potentially useful in music therapy, brain-computer interfaces for non-verbal communication and assistive technologies like creative music composition and automatic playlist creation.

1.2 Research questions

As mentioned, the aim of the research is to offline compose music based on the emotional state of a person, extracted from EEG signals. To achieve this, the following two sub-questions will be addressed:

1. How can emotions be best detected and classified from EEG signals?
2. How can music be composed from a classified emotional state in an offline BCI setting?

The focus of the project will lie on solving the first sub-question. Since the research and implementation of techniques of the second sub-question is less related to this BCI course, we will not develop a music generation model ourselves. Instead, we will reuse and implement an open source recurrent neural network to generate music in our offline BCI model.

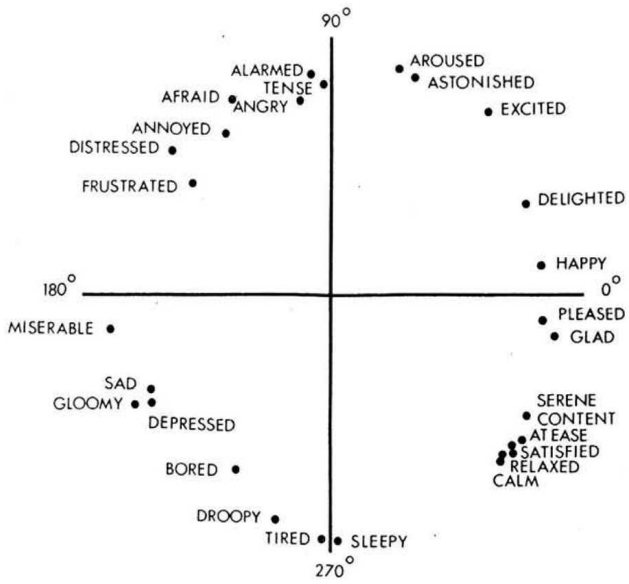


Figure 1: Russell's circumplex model of affect

1.3 Theoretical background and literature

Emotion classification with EEG data
To model emotions computationally, emotions need to be represented in a machine-readable format. Two theories to represent emotions dominate: Ekman's six basic emotions [2], and the circumplex model of Russell [10]. Russell's circumplex model of affect is displayed in figure 1. Here it is visible that emotions are plotted on two axes, valence and arousal. Valence indicates the extend to which an emotion is positive or negative. Arousal is the extend to which an emotion represents a psychologically activated state. We choose to adopt this model in our research, because it represents the relation between emotions, in contrast to Ekman's representation of the six basic emotions.

Emotions are elicited deep in the brain. Change of cognitive state causes alteration of amplitudes and frequencies of EEG signals, but since emotions originate so deep in the brain, no clear patterns or responses have been distinguished. Previous research present various activity changes in EEG signals due to emotional change. Multiple ways for detecting affective states from EEG data have therefore also been presented. Some research report higher activity in the left frontal lobe in comparison with the right hemisphere during positive feelings, mainly in alpha powerbands [11]. Other research found significant differences in theta and alpha bands in several brain areas [6]. Researchers seem to be unanimous about the fact that emotions could be identified by spatio-temporal connectivity patterns, although more research to a good detection method is necessary. The most reliable methods nowadays rely on machine learning methods, where patterns are automatically found for different emotions.

Music-induced emotion recognition based on EEG has been done before. Music is known to influence the emotional state of people, which is why music is used as elicitor of emotions to be measured by EEG. Examples are [13] and [1]. In the former study, Thammasan et al. used fractal dimension (FD) and power spectral density (PSD) to extract features from raw EEG signals, and applied emotion classification algorithms (SVM, MP and C4.5) to binary classify emotions. They found that FD value features outperform PSD, and that support vector machine (SVM) achieved better classification results than the other methods. This combination achieved an accuracy of 82.8%. The latter research was conducted by Daly et al., who achieved an emotion classification accuracy of 79.5%. Here, band power features were used in classification by LDA and SVM. Here, SVM resulted in the highest accuracy. Mehmood and Lee [8] analysed SVM and KNN as emotion classification methods. In their research, KNN (61%, k=3) resulted in higher classification accuracy compared to SVM (32 and 37%).

Feature extraction from EEG data must be done before classification. Various feature extraction methods exist and are used for emotion classification (adopted from the slides of lecture 4):

- Time-domain features
  - 6 Statistics of signal
  - Hjorth features: mobility and complexity
  - Non-Stationary Index (NSI)

- Fractal Dimension (FD)
  - Higher Order Crossing (HOC)

- Frequency-domain features
  - Band power (Fourier transform): Theta, Alpha, Beta and Gamma
  - Bin power: 4-40 Hz with length of bin = 2 Hz.

- Time-frequency domain features
  - Discrete Wavelet Transform

- Cross-channel features
  - Band power differences (left - right)
  - Band power ratio (left / right)
  - Magnitude squared coherence estimate (MSCE)

Music composition from emotion
No online BCIs for music generation exist. Yuksel et al. [14] developed a BCI application for music adaptation. They use fNIRS to classify the cognitive workload of subjects, which determines the if and how the music should be adapted. The adaptive musical interface BRAAHMS did not rely on machine learning methods, but simply provides harmonic additions depending on the BCI.

SentiMozart is a music generation application based on emotions, developed by Madhok, Goel and Garg [7]. They use a double stacked LSTM architecture to generate music, which takes one of seven emotions as impot. Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN). RNNs are useful machine learning architectures to generate sequences at output. A RNN should be used to generate music, since notes and chords in a song depend on previous notes and chords, and thus a song is a sequence of notes and chords.

Skuli presents in a blog post [12]. He has published his code on Github[1]. This model is build using the high level API Keras, to interact with machine learning methods of Tensorflow. A LSTM network is used. The model can be trained by feeding a set of MIDI files as input.

1.4 Research design

First, a classifier should be trained to classify emotions based on EEG signals. Music labelled to trigger certain emotions are taken from a dataset from the University of Jyväskylä [9]. This dataset consists of 360 music snippets of 12 different emotions or feelings.

To goal for this part of the experiment is to classify emotions anger, surprise, happy and sad, so only songs classified as eliciting one of these four emotions are used. For each class, x song fragments are taken, where each music fragment is approximately 20 seconds. All songs are concatenated in random order, with a small break of 5 seconds between the songs. The participant will be asked to label the emotion he/she feels. The participant also indicates when he/she starts feeling this emotion, which indicated the start of the event which we will need later for data processing. The participant is instructed to move as minimal as possible, because every movement will cause artifacts or noise in the EEG signals. Therefore, the participant only needs to touch a few buttons with one hand, so only movement of the fingers of one hand is required during the experiment.

The channels that will be recorded are Fp1, Fp2, F7, F3, Fz, F4, F8, T7, C3, C4, T8 and Pz, because these electrodes are located closely to the frontal lobe, which is believed to play a central role in emotion regulation [5]. [13] uses the same channels in their continuous EEG music-emotion recognition model.

The following steps will be carried out with two participants:

1. Inform the participant about the aim of the research and what steps will the role of the participant is.

2. Set up tMSI equipment for the participants, make sure he/she is relaxed and sit comfortably. Check the signal to make sure all electrodes are connected well.

3. Ask participants to fill questionnaire: Profile of Mood States (POMS), to capture the mood of the subjects before the experiment. The mood of the subjects before the experiment may influence the emotions felt during the experiments.

4. Start recording EEG data.

[1] https://github.com/Skuldur/Classical-Piano-Composer

5. Record 1 minute baseline: fixation cross on screen, and ask participant to relax.
6. Record EEG data of participants:
    (a) Start music playlist
    (b) For each song the subject indicates when he/she feels an emotion, and which emotion. The subject can choose from the four emotions listed.
    (c) After each song there is a break of 5 seconds, in which the participants hears no music and is asked to look at a fixation cross on the screen.
    (d) The EEG signals are saved into a gdf data file.

1.5 Data collection and analysis

For our project, we will need at least 2x2 subjects. First, we need training data for the emotion classification model. We plan to record EEG signals from two subjects for this. We will expose music (which is pre-classified to certain emotions) to the subjects, and the subject will be asked to note the emotion they feel during each music snippet. The pre-classified emotion of the audio snippet will be compared with the indicated emotions of the subjects. If there is a match, the EEG data will be used in training. It will be discarded otherwise.

1.6 Resources

For capturing EEG data, we will use tMSI equipment[2]. To model the BCI EEG measurement loop, OpenViBE[3] will be used. The emotion classification model will be built in Matlab[4]. Preprocessing will be done with similar commands EEGLab[5] uses. EEGLab itself will not be used, since this tool does not offer an efficient data processing method for a multiple datafiles. The recurrent neural network for music composition out of emotions will be implemented in Python[6]. Generated music in MIDI format can be played with Windows Media Player.

1.7 Timeline

In week 13 (week 7 of the course), literature research is done and a research proposal draft is made. In week 14 (week 8 of the course) we aim to answer the first research sub-question. In the beginning of the week, we will record training data from two subjects. We will design, implement and test a emotion classification from EEG data model in Matlab, using knowledge obtained from relevant literature. The model will be trained and tested using the data recorded earlier that week. In week 15 (week 9 of the course), we will address the second sub-question. We will implement, train and test an (open source) music generation model. Since this model needs classified emotions as input, we will need to adapt the input of the neural network, and train it again. If tests are successful, we will connect this to the emotion classification model we designed earlier. At the end of week 15 or the beginning of week 16, we will run experiments with our model with two more test subjects. Here, we will test if our model behaves as expected and ask the participant for his or her experience.

2 Experiments

We did two attempts to gather training data that would give us a sufficient accuracy rate on our emotion classification model.

The quantity and quality of the data captured in the first try appeared to be not sufficient to use for training. In this first attempt, we set up a EEG experiment with the tMSI equipment, prepared a playlist of 40 songs (10 of each emotion category) in random order, and prepared a data entry form for the participants to note their emotion. We did this first experiment with ourselves as subjects, so that we could explore the process. We used the four emotions 'sad', 'happy', 'surprised' and 'anger', because these emotions are most different from each other in terms of valence and arousal. When we analysed the data, it turned out that the overall quality and quantity was not high enough to use it for training yet. A few channels were so much affected by noise that the signals could not be used. Then, since not all songs were classified with the right emotion by the subjects (us),

[2] https://www.tmsi.com/
[3] http://openvibe.inria.fr/
[4] https://www.mathworks.com/products/matlab.html
[5] https://sccn.ucsd.edu/eeglab/index.php
[6] https://www.python.org/

and these trials thus had to be taken out of the dataset. In the end, we had 58 trials, of which 32 trials had 3 channels that were unusable.

For training the classifier, data should be split in training and test data. If we would use only 58 samples with a 80:20 training:test ratio, we would have only 46 samples to train on, which comes down to around 12 samples per emotion. We decided to gather more data, and also attempt to get a higher data quality to train our emotion classifier properly.

In the second attempt of gathering training data, we used ourselves as samples again. Instead of 40 songs, we used all 120 songs of the song database (30 for each emotion happy, sad, anger and surprised). For each classified emotion, we took a four second time fragment around the indicated event by the subject (2 seconds before, and 2 seconds after). Because 1 second should be sufficient for emotion classification, we extracted 7 fragments per trial (allowing 0.5 seconds overlap between the fragments).

2.1 Data Processing

We used Matlab code as well as EEGLab to (pre)process the obtained data. We wrote a Matlab code containing functions EEGLab also uses, in order to speed up the processing for lots of data (so that we did not have to do this manually with EEGLab every time). The first preprocessing steps is channel selection (Fp1, Fp2, F7, F3, Fz, F4, F8, T7, C3, C4, T8 and Pz). Then, the data was filtered with a 1-45 Hz bandpass filter, and downsampled to 128 Hz.

For artifact removal, we used a machine learning based tool MARA[7]. This is a tool to be used in EEGLab, so we had to use this by hand for each subject. MARA automised the hand-labelling of independent components for artifact rejection. It relies on a supervised machine learning model, which takes six features from the spatial, spectral and temporal domain.

After (automatic) artifact removal, the data of each subject was split into the four emotion classes, and saved as separate EEG signal datasets. Each trial was 4 seconds long, a time window of 2 seconds before and 2 seconds after the event (when the emotion is felt by the participant). This trial is split into segments of 1 seconds, with 0.5 seconds overlap, to generate more data.

Then, two methods for feature extraction were applied to all of these datasets. First, the Power Spectral Density (PSD) approach was applied. PSD is based on Fourier transform, and indicated signal power in specific frequency ranges. The transformation converts data in time to the frequency domain and vice versa. Each EEG signal is decomposed in five frequency bands (alpha, beta, gamma, delta and theta), by the Matlab Signal Processing Toolbox. The feature is the average power over the given frequency band calculated. The average is taken because PSD represents signals in the continuous frequency domain, while we need a single feature.

Secondly, we used Fractal Dimension (FD) to extract more features. FD is the complexity of time-varying signals. Higher values mean higher brain activity, and is used more often in affective brain computing research.

These two methods gave us a total of 45 features per dataset. These 45 features of training data were used to train a classification model. We compared LDA, SVM and KNN models first, and we soon found out that LDA had a very poor accuracy (around 20%). SVM and KNN had similar accuracy rates on different try outs, but since KNN appeared to be a lot faster because its less computationally intensive, we decided to use KNN as classification model to train the final model, with 5 neighbours and a standardised seuclidean distance metric. The models are trained using 10 fold cross validation.

3 Results

3.1 Emotion classification

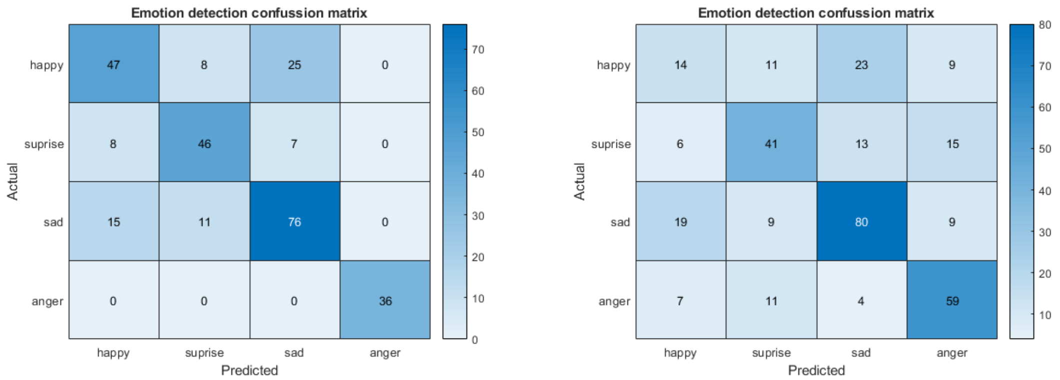
EEG signals of both experiments was combined per subject, so that more training data was available. A SVM and KNN classifier was trained on the data for both subjects, and the results are presented in table 1. We see that emotions classification for person 1 achieves a higher performance than person 2. Figure 2 shows confusion matrices of classifiers trained with SVM, for both subjects' data apart. From these confusion matrices it is already visible that classification of emotions for subject 1 achieves a higher accuracy than subject 2.

Next, a SVM was trained and tested on combined data of both subjects and both experiments. The obtained accuracy is 0.77. The confusion matrix is shown in figure 3a. We took two approached to improve classification results. We realise that this is modification of the training data, and it should not be done in real research, but our goal was to improve the classification nevertheless. First, because we observed the low accuracy of subject 2, we decided to just take the EEG signal data of subject 2 of the emotion 'anger'. Data of person 1 had a relatively low number of anger trials in the set, which could cause a low accuracy. Second, we used a KNN

[7] https://irenne.github.io/artifacts/

Subject	Accuracy with SVM
Subject 1	0.73
Subject 2	0.59

Table 1: Emotion classification results per subject with SVM



(a) Subject 1

(b) Subject 2

Figure 2: Confusion matrices of subjects 1 and 2, classifier trained with SVM

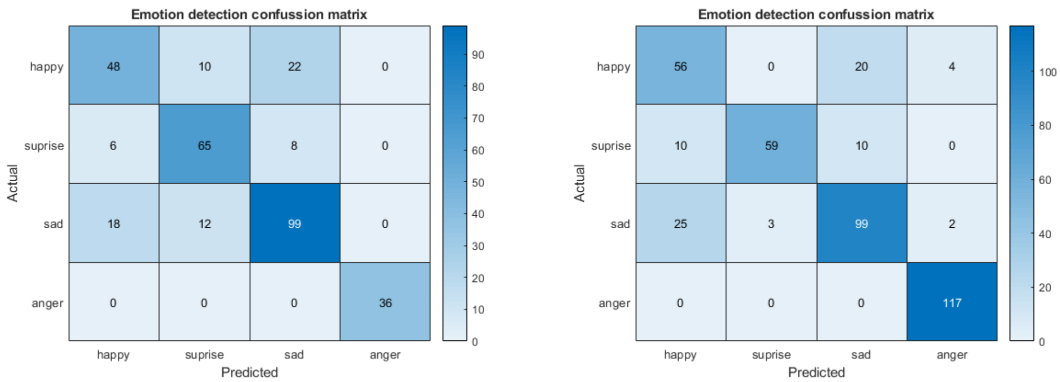
instead of SVM, because the training is faster compared to SVM, and higher accuracy was obtained. The final accuracy is 0.82, and the confusion matrix is displayed in figure 3b.

3.2 Music generation

Like mentioned before, music generation based on emotion was not the main focus of this project. Nevertheless, we adapted a machine learning model (LSTM) to learn to compose classical music based on emotion. The same four emotion classes were used. The songs from the dataset of emotions we used for the training experiments were as training data for the music composer as well. The model is accessible on this Github page. As an example, the model is trained on happy songs, and the result *'happy_ output.mid'* is included in this folder.

4 Conclusion and Discussion

Finally, an emotion classification accuracy of 82% is achieved, with the training data of subject 1 combined with 'anger' trials of subject 2. Both experiments are taken as training or test data. Only the emotion 'anger' of subject 2 is taken, because trials of other emotions for subject 2 lead to a low classification rate. This may have several causes. First, the experiments of subject two took place in a noisy environment. It was lunchtime during the experiment, which meant that the subject was distracted by moving people around her, who were talking



(a) SVM subjects combined

(b) KNN subjects combined

Figure 3: Confusion matrices SVM and KNN, subjects combined

and using the microwave in the same room. Since the experiments last for approximately half an hour, the first half of the experiments was bothered by this. In addition, the software of the music player crashed during the second of subject 2, and had to be restarted. It could be that results of subject 2 are affected by this as well. Both potential noise and/or low accuracy causes described here did not happen when the 'anger' eliciting songs were played, at the end of the experiment. This may explain why 'anger' achieves a higher classification performance rate. This is the reason to take only this emotion in the final classifier.

Another discussion point can be made about the way how emotions are classified by the subjects. Emotions should be elicited by the song snippets played, but a clear emotion was not always felt. Nevertheless, subjects had to note down an emotion for every song played.

The high EEG signal instability could be caused by body movements and electrical noise. To overcome this and generate a more stable result, other physiological measurement methods could be combined with EEG signals in a multimodel model, as Thammasan et al. [13] used.

Lastly, our method of optimising the classifier is highly debatable. As explained in the previous section, the data was optimised by ourselves because we had very little training data by hand.

5 Suggestions for future work

Research should be done to how emotion can be classified best. Suggested is to research whether separate classification of valence and arousal achieves higher performance than our method.

Second, we advise future researchers to obtain more training data. More trials of more subject most likely result in a more reliable classifier, also achieving a higher classification performance.

As mentioned in the discussion above, the experiment environment was not ideal. If subjects are less disturbed by noise from the environment, better classification results might be obtained. In addition, better selection to songs could also results in higher performance.

Due to time and knowledge limitations, we were not able to build an online emotion music composer, were emotions are classified and music is generated live. This is a goal for future research.

Finally, we researched how music could be generated from emotional state of subjects, but the opposite should also be researched. Especially when an online music generation BCI is designed, it is important to research the influence of music on emotion.

References

[1] Ian Daly et al. "Identifying music-induced emotions from EEG for use in brain-computer music interfacing". In: *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE. 2015, pp. 923–929.

[2] Paul Ekman. "An argument for basic emotions". In: *Cognition & emotion* 6.3-4 (1992), pp. 169–200.

[3] Min-Ki Kim et al. "A review on the computational methods for emotional state estimation from the human EEG". In: *Computational and mathematical methods in medicine* 2013 (2013).

[4] Stefan Koelsch. *Brain and music*. John Wiley & Sons, 2012.

[5] Stefan Koelsch. "Brain correlates of music-evoked emotions". In: *Nature Reviews Neuroscience* 15.3 (2014), p. 170.

[6] You-Yun Lee and Shulan Hsieh. "Classifying different emotional states by means of EEG-based functional connectivity patterns". In: *PloS one* 9.4 (2014), e95415.

[7] Rishi Madhok, Shivali Goel, and Shweta Garg. "SentiMozart: Music Generation based on Emotions." In: *ICAART (2)*. 2018, pp. 501–506.

[8] Raja Majid Mehmood and Hyo Jong Lee. "Emotion classification of EEG brain signal using SVM and KNN". In: *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE. 2015, pp. 1–5.

[9] *MS-EV-Exp1stimuli.pdf*. Jan. 2009. URL: https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/projects2/past-projects/coe/materials/emotion/MS-EV-Exp1stimuli.pdf/view.

[10] James A Russell. "A circumplex model of affect." In: *Journal of personality and social psychology* 39.6 (1980), p. 1161.

[11] Louis A Schmidt and Laurel J Trainor. "Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions". In: *Cognition & Emotion* 15.4 (2001), pp. 487–500.

[12] Sigur Skuli. *How to Generate Music using a LSTM Neural Network in Keras*. Dec. 2017. URL: https://towardsdatascience.com/how-to-generate-music-using-a-lstm-neural-network-in-keras-68786834d4c5.

[13] Nattapong Thammasan et al. "Continuous music-emotion recognition based on electroencephalogram". In: *IEICE TRANSACTIONS on Information and Systems* 99.4 (2016), pp. 1234–1241.

[14] Beste F Yuksel et al. "Implicit Brain-Computer Interaction Applied to a Novel Adaptive Musical Interface". In: *Dept. Computer Science, Tufts Univ.* 2015.