MINDMIX: Mapping brain activity to congruent audio mixing features

Duncan Williams

Acoustics Research Centre
University of Salford, Manchester UK
d.a.h.williams@salford.ac.uk

ABSTRACT

Brain-computer interfacing (BCI) offers novel methods to facilitate participation in audio engineering, providing access for individuals who might otherwise be unable to take part (either due to lack of training, or physical disability). This paper describes the development of a BCI system for conscious, or 'active', control of parameters on an audio mixer by generation of synchronous MIDI Machine Control messages. The mapping between neurophysiological cues and audio parameter must be intuitive for a neophyte audience (i.e., one without prior training or the physical skills developed by professional audio engineers when working with tactile interfaces). The prototype is dubbed MINDMIX (a portmanteau of 'mind' and 'mixer'), combining discrete and many-to-many mappings of audio mixer parameters and BCI control signals measured via Electronecephalograph (EEG). In future, specific evaluation of discrete mappings would be useful for iterative system design.

Author Keywords

BCI, sound engineering, EEG, audio mixing, accessible computing

CCS Concepts

• **Applied computing → Sound and music computing**; Performing arts; • **Information systems →** *Music retrieval*;

1. INTRODUCTION

Brain-computer interfacing (BCI) can be used to adapt various neurophysiological measurement techniques to the control of a wide range of applications, for example gaming [1]. Music can have a huge impact on our day-to-day lives, and is increasingly being correlated with mental health and wellbeing [2]. In such contexts, music has been shown to reduce stress, improve athletic performance, aid mindfullness, and increase concentration. There is a large potential user base for music amongst people who might otherwise be unable to engage in music making activities via traditional means (either due to lack of training, or physical disability), who might benefit from biophysiologically-informed computer aided interaction with music.

BCI hardware is becoming increasingly affordable and accessible, giving rise to music specific applications in the emerging field of brain-computer music interfacing, (BCMI) [3], [4]. There are many reasons why audio engineers prefer tactile control of mixing processes [5], which partially explains the significant interest, and progress being made in the field of haptic augmentation in audio and musical instrument design [6], [7]. There is a distinction to be made between active and passive BCI control. Active control means that the user must be able to take a clear agency over the resulting actions, for example, by imagining a movement and seeing a direct correlation in the resulting system behaviour. Passive control would include detection of control signals which are not directly controllable by the end user (for example, heart rate, or galvanic skin response for emotional state estimation). Beyond encouraging inclusivity and participation through facilitating access to audio engineering processes via linear mapping strategies, the potential to harness unconscious processes (passive control) suggests that augmented audio engineering, for example, individually adaptive, responsive, or context-dependent remixing, may be a possibility. Such technology could be married together with the significant advances in music information retrieval (MIR), non-linear music creation [8], and context-adaptive music selection in the future.

NIME'20, July 21-25, 2020, Royal Birmingham Conservatoire, Birmingham City University, Birmingham, United Kingdom.

2. BACKGROUND & PREVIOUS WORK

BCI is gradually becoming more established, but harnessing this technology for music making, or controlling more general interactions with music (for example, using brainwaves to select music playlists autonomously) is less common and something of an emerging field.

Significant progress towards functional Brain-Computer *Music* Interfacing, or BCMI, was made in the 1990's, for example Biomuse [9] which mapped low-level neuroelectric and myoelectric signals to the generation of MIDI data in real-time. Beyond music composition tasks, some specific work has attempted to harness the electroencephalogram, (EEG)[1] and related signals in an audio engineering context. For example, Miranda et al. attempted to create a control signal for volume automation in a basic audio mixer using a simple metric from EEG, the amplitude of alpha and beta waves [12]. This work attempted to use BCI to control the amplitude of two separate audio faders in a virtual (digital audio workstation) mixer. Beta frequencies are more often associated with active states of mind, but the process of actively mixing audio requires both attention and precise control. Therefore, this choice of parameters is somewhat incongruous for the end-use application: Becoming 'calmer' is not analogous to any mixer property, nor generally adaptable to the task of mixing a range of musical material, though specific musical examples might be more relevant - for example, in a case where a calmer state of mind would cause the level of active or energetic music to lower, and raise the level of calmer sounding music. Other work harnessing existing BCI metrics and adapting them for musical control includes use of the P300, ERP, or 'oddball' paradigm [13], measurement of specific frequency bands (activity in alpha, beta, gamma, or mu) [14], [15], steady-state visually evoked potential (SSVEP) [16], [17], and measures of asymmetry [11], [18], all via EEG.

Related research challenges include increased speed of classification, for example by machine learning techniques [19] or accuracy of the interface [20]. Generally these are challenges related

1 The Electroencephalogram (EEG) is a device for measuring electrical activity across the brain via electrodes placed across the scalp [10]. In the case of the instrument described here, the standard 10/20 arrangement was used to determine electrode placement [11].

to the type of data processing employed, either statistical data reduction or noise reduction and artefact removal techniques in other measurement paradigms, such as functional magnetic resonance imaging (fMRI) [21].

In the case of music mixing, there are many application-specific goals that need to be considered in order for the BCMI system to best serve its intended use. In, for example, a music therapy context, one advantage of a BCMI system is that it might be used by a person with no a priori experience or musical training, in order to engage in music production in context. However, in order to do this the BCMI must be capable of performing music which is well correlated with the signal being analysed as a control signal (e.g., BCI parameters mapped according to constraints of melody, harmony, rhythm, or genre) yet also allows the user enough degrees of freedom to feel that they are truly the agent of their performance. Therefore criteria for the specification of a suitable BCMI system must include consideration of both agency and conformity to production rules.

There is a marked difference between systems for controlling music directly by means of BCI, and systems for sonification or musification of brainwave data (typically EEG), [22], [23]. Sonification is a process whereby data is directly presented by auditory means (for example, an alarm, telephone ring, etc.) [24]. In terms of brain-computer music applications, sonification of EEG has become common [22], [23], [25] with many existing mappings being used. Mapping is particularly important in the design of such systems, as the range of controls available (even in hybrid systems) is still minimal in comparison to the spread of possible actions involved in music making. An overview of different types of music mapping from complex biomedical data and subsequent evaluation strategies, is given in [26]. An overview of specific mapping techniques for digital instrument design is given in [27]. Various combinations of mapping strategies exist, including one-to-one, one-to-many, and many-to-many combinations [28]. It is in the mapping stage that a system for controlling audio mixing functionality derives success through utility (or lack thereof in the case of ineffective mapping). Work to establish correlation between such parameters and plausible brain signals is beyond the scope of this paper, but would likely be well received by the applied BCI community as part of a research road map.

As mentioned above, previous attempts to use BCI to control audio mixing parameters have been designed solely to use alpha and beta activity to control the amplitude of two separate faders. Our approach is radically different in design and implementation. For the prototype under evaluation here, control metrics and mappings were selected with the intention that they would be congruent between operator and operation.

2.1 EEG Metrics

Several metrics for extracting meaningful control data from EEG are common in BCI systems. The P300 ERP (Event Related Potential, or 'oddball' paradigm) has been used to allow active control over note selection for real-time sound synthesis [29], [30]. Such methods are not dissimilar to ERP spelling systems, e.g., [31], [32], which are now increasingly common in the BCI world, though adapted to musical notes rather than text input. Stimulus-responsive input measures, for example, the SSVEP[2] [16], have been adapted to real-time score selection [34]. Active control by means of Mu frequency rhythm and motor imagery are also becoming popular as control signals for various applications, including avatar movement in virtual reality, operation of spelling devices, and neuroprostheses [15], [32]. The challenge, then, is in devising and evaluating mappings which are most suited to task-specific control – in this case, audio engineering processes, more specifically, mixing processes. MINDMIX control mappings were selected according to this philosophy. For example, once a particular channel has been selected, left or right motor imagery can be actively engaged to adjust the panorama of an audio source to move a sound image between left and right loudspeakers in a 2-channel stereo configuration. This is a many-many mapping wherein the channel is first selected by means of SSVEP, then the pan control selected by ERP, before the pan value is adjusted according to Mu L/R balance.

The range of tactile functions the MINDMIX prototype aims to augment are as follows: Transport control (play, stop, fwd, rev), fader select and level (individual channels, buss, and FX return), potentiometer select and adjust (pan, parametric EQ), and channel switching (solo, mute, insert, EQ in/out). Each of these parameters has been mapped to a sequence of actively controllable metrics, combining motor imagery (left and right), SSVEP, and ERP.

The MINDMIX prototype focusses solely on mixing (including remixing, and post-production tasks), rather than on source capture or recording. Therefore, there was no need to include functionality such the various categories of record which a transport bar might exhibit in a fully featured console or digital audio workstation.

3. SYSTEM OVERVIEW

Tables 1 and 2 show EEG metrics and mappings to parameters as implemented in the prototype system. To demonstrate the application of passive BCI measures (e.g., alpha, beta, and asymmetry), a master FX send and return was also implemented under control of the relative level between beta and alpha (greater level of alpha resulting in a "wetter balance", one with a higher ratio of effect to unaffected signal).

Table 1 Mapping between EEG metric and generic mixer control types

Mixer parameter	EEG metric
Channel select	SSVEP (specific frequency)
Switch select (e.g., select a rotary potentiometer, switch on/off)	ERP
Adjust rotary potentiometer	Motor imagery (left, right)
Adjust fader amplitude	SSVEP duration
Transport control (play/stop)	SSVEP and ERP

Combinations are accessed according to context (i.e., a channel is selected before a specific parameter is chosen and then varied according to motor imagery in the case of potentiometer, SSVEP duration in the case of faders, and ERP in the case of switches). Two common mixer parameters were not implemented in the pilot mapping: EQ width (i.e., semi-parametric EQ only), and channel FX send controls. Real-time input is analyzed and filtered, including artefact removal to produce simple control signals. Control signals are then mapped to mixer parameters; sent by Open Sound Control to generate MIDI Machine Control. The signal flow is a feedback loop comprising 16 channels EEG->data smoothing->semantic content analysis -> classify ->route to mixer parameter->EEG, shown in Figure 1.

Combinations of mappings (i.e., many-many mapping) allows for a channel to be selected using SSVEP, followed by a potentiometer (e.g., pan, or semi-parametric EQ frequency/gain) to be selected according to ERP, before the value of the potentiometer itself is set according to imagined motor imagery (i.e., left, or right). SSVEP allows users to make a selection by focusing their gaze on a visual

[2] SSVEP is a response to visual stimulation at a given frequency and integer multiples thereof, measurable in the visual cortex. For a detailed explanation of the signal characteristics under analysis, the reader is referred to [16], [17], and to [33] for a review of use in various BCMI platforms

stimulus oscillating at a given rate. As well as initial parameter selection, SSVEP also allows for second level of control by mapping the duration of the gaze with non-linear features, for example amplitude, allowing for a degree of continuous control i.e., after selecting a specific channel the duration of a user's gaze can be used to adjust the fader for the selected channel accordingly. A similar effect could be achieved using eye-tracking in a hybrid system, using duration of gaze as a secondary mapping for amplitude. The parameters which are most useful for broad user participation in terms of transport across the digital audio workstation are play, stop, select, and various level parameters. It is important to consider the most meaningful signal type for each parameter in the mapping; some of these control signals have analogous actions in a mixer, for example, motor cortex with transport controls (stop, go, fast forward, rewind), and some have analogous parameters in music (SSVEP to non-linear adjustment of amplitude via faders). However, partly due to the infancy of the use of BCI for music making, the selection of these combinations is necessarily somewhat arbitrary, and therefore methods for evaluating the success of these mappings is necessary.

Table 2 many-many mappings for mixer parameters.

Mixer parameter	**Many-many mapping**
Channel volume	SSVEP (select) and SSVEP duration (gain)
Buss volume	ERP (select) and SSVEP duration (gain)
Master FX return volume	Alpha/Beta balance
Channel pan	SSVEP (channel select), ERP (select pan), Motor imagery (adjust left-right balance)
Channel EQ freq	SSVEP (parameter select), Motor imagery (adjust frequency range low-high)
Channel EQ gain	SSVEP (parameter select), SSVEP duration (gain)
Channel insert in/out, EQ in/out, Solo	ERP

The system is implemented in OpenVibe with Max/MSP and Reaper digital audio workstations. As described in the introduction, our focus is on using relatively well-known metrics in a music control context. Any number of statistical data reduction techniques might be used in the signal analysis block - For details of previous studies using principal component analysis for music measurement when analyzing EEG, the interested reader is referred to previous work in [4], [35]–[37].

4. DISCUSSION AND FURTHER WORK

A number of paradigms for the evaluation of BCI systems exist, however they often focus on technical or methodological details. There is a tendency in BCI work to prioritise technical implementation in research reporting, for example considering increased speed or accuracy of a system, rather than the application itself. For the purposes of this work, which combines existing techniques that have already been well-documented in the BCI community, such evaluations are less relevant. Instead we suggest that readers interested in in-depth consideration of particular BCI techniques (typically regarding issues of speed, accuracy, new technological implementation etc.,) consider reviews offered in [38], [39] and most recently, [40]. Of most concern to the system presented here is the appropriateness of the mapping and the relevance and usefulness of the user interaction with the application. In the traditional audio engineering domain, this would be comparable to evaluating decisions such as whether, for example, a rotary potentiometer or a fader was most appropriate for control of a discrete audio parameter.

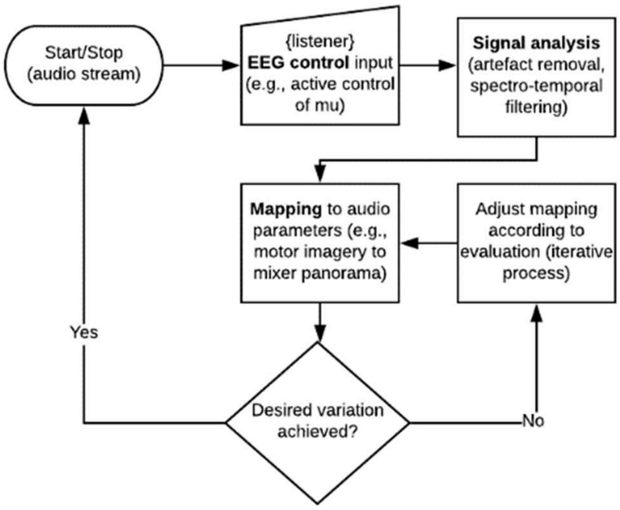

Fig. 1. Overview of signal flow and iterative evaluation process.

BCI offers (i) augmentation to listeners who might benefit from context specific or adaptive audio - such as non-linear immersive audio in the next generation of gaming and virtual reality – and (ii) participation opportunities to those who might otherwise be unable to take part in the possibilities for expression afforded by creative audio engineering, including emotional contagion, communication, and perhaps most importantly, interaction with others.

There is a serious argument to be made for the use of brain-computer devices to assist access in terms of inclusion: users who might otherwise be unable to enjoy audio engineering can potentially take part using this technology. Significant advances have already been made in the field of brain-computer music interfacing – might bio-assisted audio engineering be able to take these advances to the next level in terms of inclusive system design? Previous systems have generally not successfully been able to integrate congruent design with the sensing algorithms being used. One application adapting neurophysiological cues to the control of audio mixer parameters has been described here. As with any such application, the utility is somewhat dependent on the complexity of the mappings, and the number of meaningful, controllable features that might be extracted from the EEG. These include overall signal amplitude, frequency domain analysis derived amplitudes, and spatial distribution of both properties at specific electrode placements on the scalp (for example, denoting motor cortex activity, asymmetry and other spatial distribution metrics) [18].

Evaluation strategies for BCI-to-audio mappings, in general, are not universally agreed upon and remain a significant area for further work. An exploration of rankings across different musical genres might be a useful avenue for further work in evaluating this type of assistive technology in a real-world context.

The use of musical stimuli to mediate or entrain the listeners' brain activity (i.e., neurofeedback) also remains a fertile area for research activity [23], [36]. Neurofeedback is becoming increasingly common in the design of brain-computer music interfacing for specific purposes such as therapeutic applications. Similarly, a significant amount of further work remains in quantifying listener responses to affectively-charged music, and in measuring the impact on a given affective state that music might have on an individual already in a given state. Nevertheless, the possibility of developing affectively-responsive audio applications, using cues from BCI technology suggests that individual variability might in the future be mediated in ways that had previously been thought impossible; for example, a mapping to particular mix features which respond adaptively to the individual whilst listening to a mix.

We may then, in the future, see systems adapted to more generalizable portable mappings which that might be controlled by EEG, using adaptive mappings derived by machine learning rather than prescribed by the designers of such systems, for audio engineering applications regardless of physical ability or previous training, or even individually responsive mixes based on a listeners' biosignals. Further work establishing plausible parameters for control via brain signals, with consideration for congruence between the two, would likely provide a welcome research road map for the applied BCI community (including domains beyond sound and music computing) in future.

5. REFERENCES

[1] O. AlZoubi, I. Koprinska, and R. A. Calvo, "Classification of brain-computer interface data," in *Proceedings of the 7th Australasian Data Mining Conference-Volume 87*, 2008, pp. 123–131, Accessed: Aug. 10, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2449309.

[2] R. MacDonald, G. Kreutz, and L. Mitchell, "What is music, health, and wellbeing and why is it important," *Music, health and wellbeing*, pp. 3–11, 2012.

[3] Y. C. (Denise) Chew and E. Caspary, "MusEEGk: a brain computer musical interface," in *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*, New York, N.Y., 2011, pp. 1417–1422, doi: 10.1145/1979742.1979784.

[4] I. Daly *et al.*, "Affective brain–computer music interfacing," *Journal of Neural Engineering*, vol. 13, no. 4, pp. 46022–46035, 2016.

[5] S. Merchel, E. Altinsoy, and M. Stamm, "Tactile Music Instrument Recognition for Audio Mixers," in *Audio Engineering Society Convention 128*, May 2010, [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=15438.

[6] L. Picinali and B. F. Katz, "Spectral discrimination thresholds comparing audio and haptics," in *Proceedings of haptic and auditory interaction design workshop, Copenhagen*, 2010, pp. 1–2.

[7] S. Merchel, M. E. Altinsoy, and M. Stamm, "Touch the sound: audio-driven tactile feedback for audio mixing applications," *Journal of the Audio Engineering Society*, vol. 60, no. 1/2, pp. 47–53, 2012.

[8] A. Berndt, *Musical Nonlinearity in Interactive Narrative Environments*. Ann Arbor, MI: MPublishing, University of Michigan Library, 2009.

[9] R. B. Knapp and H. S. Lusted, "A bioelectric controller for computer music applications," *Computer music journal*, vol. 14, no. 1, pp. 42–47, 1990.

[10] E. Niedermeyer and F. H. L. da Silva, *Electroencephalography : basic principles, clinical applications, and related fields*. Philadelphia (Mass.): Lippincott Williams and Wilkins, 2005.

[11] J. J. B. Allen and J. P. Kline, "Frontal EEG asymmetry, emotion, and psychopathology: the first, and the next 25 years," *Biological Psychology*, vol. 67, no. 1–2, pp. 1–5, Oct. 2004, doi: 10.1016/j.biopsycho.2004.03.001.

[12] E. R. Miranda, "Plymouth brain-computer music interfacing project: from EEG audio mixers to composition informed by cognitive neuroscience," *International Journal of Arts and Technology*, vol. 3, no. 2, pp. 154–176, 2010.

[13] A.-M. Brouwer and J. van Erp, "A tactile P300 brain-computer interface," *Frontiers in Neuroscience*, 2010, doi: 10.3389/fnins.2010.00019.

[14] L. I. Aftanas and S. A. Golocheikine, "Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation," *Neuroscience letters*, vol. 310, no. 1, pp. 57–60, 2001.

[15] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements," *Brain topography*, vol. 12, no. 3, pp. 177–186, 2000.

[16] M. Middendorf, G. McMillan, G. Calhoun, K. S. Jones, and others, "Brain-computer interfaces based on the steady-state visual-evoked response," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 211–214, 2000.

[17] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, "Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components," *Journal of neural engineering*, vol. 2, no. 4, p. 123, 2005.

[18] J. J. B. Allen, J. A. Coan, and M. Nazarian, "Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion," *Biological Psychology*, vol. 67, no. 1–2, pp. 183–218, Oct. 2004, doi: 10.1016/j.biopsycho.2004.03.007.

[19] A. K. Liu, A. M. Dale, and J. W. Belliveau, "Monte Carlo simulation studies of EEG and MEG localization accuracy," *Human brain mapping*, vol. 16, no. 1, pp. 47–62, 2002.

[20] A. Lay-Ekuakille, P. Vergallo, D. Caratelli, F. Conversano, S. Casciaro, and A. Trabacca, "Multispectrum approach in quantitative EEG: Accuracy and physical effort," *IEEE Sensors Journal*, vol. 13, no. 9, pp. 3331–3340, 2013.

[21] R. A. Feis *et al.*, "ICA-based artifact removal diminishes scan site differences in multi-center resting-state fMRI," *Frontiers in neuroscience*, vol. 9, p. 395, 2015.

[22] G. Baier, T. Hermann, and U. Stephani, "Multi-channel sonification of human EEG," in *Proceedings of the 13th International Conference on Auditory Display*, 2007.

[23] T. Hinterberger and G. Baier, "Poser: Parametric orchestral sonification of eeg in real-time for the self-regulation of brain states," *IEEE Trans. Multimedia*, vol. 12, p. 70, 2005.

[24] P. Toharia, J. Morales, O. Juan, I. Fernaud, A. Rodríguez, and J. DeFelipe, "Musical Representation of Dendritic Spine Distribution: A New Exploratory Tool," *Neuroinform*, pp. 1–13, Jan. 2014, doi: 10.1007/s12021-013-9195-0.

[25] A. de Campo, R. Hoeldrich, G. Eckel, and A. Wallisch, "New sonification tools for EEG data screening and monitoring," in *Proc. of the ICAD, 13th International Conference on Auditory Display, Montreal, Canada*, 2007.

[26] D. Williams, "Utility Versus Creativity in Biomedical Musification," *Journal of Creative Music Systems*, vol. 1, no. 1, Sep. 2016, Accessed: Feb. 28, 2017. [Online]. Available: http://jcms.org.uk/issues/Vol1Issue1/4/4.html.

[27] C. Goudeseune, "Interpolated mappings for musical instruments," *Organised Sound*, vol. 7, no. 02, pp. 85–96, 2002.

[28] A. Hunt and R. Kirk, "Mapping strategies for musical performance," *Trends in Gestural Control of Music*, vol. 21, pp. 231–258, 2000.

[29] M. Grierson, "Composing with brainwaves: minimal trial P300b recognition as an indication of subjective preference for the control of a musical instrument," in *Proceedings of International Cryogenic Materials Conference (ICMC'08)*, 2008, Accessed: Aug. 09, 2012.

[30] M. Grierson and C. Kiefer, "Better brain interfacing for the masses," 2011, p. 1681, doi: 10.1145/1979742.1979828.

[31] D. J. Krusienski *et al.*, "A comparison of classification techniques for the P300 Speller," *Journal of Neural Engineering*, vol. 3, no. 4, pp. 299–305, Dec. 2006, doi: 10.1088/1741-2560/3/4/007.

[32] C. Neuper, G. R. Müller-Putz, R. Scherer, and G. Pfurtscheller, "Motor imagery and EEG-based control of spelling devices and neuroprostheses," *Progress in brain research*, vol. 159, 2006.

[33] D. A. H. Williams and E. R. Miranda, "BCI for Music Making: Then, Now, and Next," in *Brain-Computer Interfaces Handbook*, C. S. Nam, A. Nijholt, and F. Lotte, Eds. CRC PRESS-TAYLOR & FRANCIS GROUP, 2018, pp. 191–205.

[34] E. R. Miranda and J. Castet, Eds., *Guide to brain-computer music interfacing*. London: Springer, 2014.

[35] N. Nicolaou *et al.*, "Directed motor-auditory EEG connectivity is modulated by music tempo," *Frontiers in Human Neuroscience*, Oct. 2017, doi: 10.3389/fnhum.2017.00502.

[36] I. Daly *et al.*, "Music-induced emotions can be predicted from a combination of brain activity and acoustic features," *Brain and cognition*, vol. 101, pp. 1–11, 2015.

[37] I. Daly *et al.*, "Neural correlates of emotional responses to music: an EEG study," *Neuroscience letters*, vol. 573, pp. 52–57, 2014.

[38] G. Dornhege, *Toward brain-computer interfacing*. MIT press, 2007.

[39] R. P. Rao, *Brain-computer interfacing: an introduction*. Cambridge University Press, 2013.

[40] C. S. Nam, A. Nijholt, and F. Lotte, Eds., *Brain-computer interfaces handbook: technological and theoretical advances*. Boca Raton: Taylor & Francis, CRC Press, 2018.