Learning to Make Feelings: Expressive Performance as a Part of a Machine Learning Tool for Sound-Based Emotion Control

Alexis Kirke[1], Eduardo R. Miranda[1], and Slawomir J. Nasuto[2]

[1] Interdisciplinary Centre for Computer Music Research, Plymouth University, Plymouth, UK
{alexis.kirke,eduardo.miranda}@plymouth.ac.uk
[2] Cybernetics Research Group, University of Reading, Reading, UK
s.j.nasuto@reading.ac.uk

Abstract. We propose to significantly extend our work in EEG-based emotion detection for automated expressive performances of algorithmically composed music for affective communication and induction. This new system involves music composed and expressively performed in real-time to induce specific affective states, based on the detection of affective state in a human listener. Machine learning algorithms will learn: (1) how to use biosensors such as EEG to detect the user's current emotional state; and (2) how to use algorithmic performance and composition to induce certain trajectories through affective states. In other words the system will attempt to adapt so that it can – in real-time - turn a certain user from depressed to happy, or from stressed to relaxed, or (if they like horror movies!) from relaxed to fearful. Expressive performance is key to this process as it has been shown to increase the emotional impact of affectively-based algorithmic composition. In other words if a piece is composed by computer rules to communicate an emotion of happiness, applying expressive performance rules to humanize the piece will increase the likelihood it is perceived as happy. As well as giving a project overview, a first step of this research is presented here: a machine learning system using case-based reasoning which attempts to learn from a user how themes of different affective types combine sequentially to communicate emotions.

Keywords: Music, Emotion, Bio-signals, Affective Computing, Music Therapy, Medicine, Machine Learning, Algorithmic Composition, Computer Expressive Performance.

1 Introduction

The aim of our research is to develop technology for implementing innovative intelligent systems that can monitor a person's affective state and induce a further specific affective state through music, automatically and adaptively. [1] investigates the use of EEG to detect emotion in an individual and to then generate emotional music based on this. These ideas have been extended into a 4.5 year EPSRC research project [2] in which machine learning is used to learn, by EEG emotional feedback, what types of

M. Aramaki et al. (Eds.): CMMR 2012, LNCS 7900, pp. 490–499, 2013.

music evoke what emotions in the listener. If the positive affective state inducing capacity of music could be harnessed in a more controlled way, it would make a significant impact in various recreational and medical areas. The economic impact of a system that would enable users to enter a desired affective state using music would contribute to (a) the UK's burgeoning entertainment industry and (b) the health sector (e.g., preventive medicine). Such a system could help to enhance our quality of life and contribute towards the wellbeing of the population (e.g., help reducing levels of stress and/or anxiety). This chapter introduces the key background elements behind the project: Music and Emotion, Emotional Expressive Performance and Algorithmic Composition, and EEG Affective Analysis; then details some preparatory work being undertaken, together with the future project plans.

2 Music and Emotion

Music is commonly known to evoke various affective states (popularly referred to as "emotions"); e.g., elation, calm or cheerfulness [3]. There have been a number of questionnaire studies supporting the notion that music communicates affective states (e.g., [4, 5]) and that music can be used for affect regulation and induction (e.g., [6, 7]). However the exact nature of these phenomena is not fully understood. The literature makes a distinction between perceived and induced emotion with music being able to generate both types [4]. The differences between induced affective state and perceived affective state have been discussed by Juslin and Sloboda [3]. For example a listener may enjoy a piece of music like Barber's Adagio, which most people would describe as a "sad" piece of music. However, if they gain pleasure from listening, the induced affective state must be positive, but the perceived affective state is sadness; i.e., a negative state. Despite the differences between perceived and induced affective state, they are highly correlated [4, 8]. Zentner et al. [9] reported on research into quantifying the relationship between perceived and induced affective state in music genres. Scherer [10] discussed the underlying physical mechanisms of musically induced emotions.

3 Emotion-Based Algorithmic Composition

One area of algorithmic composition which has received more attention recently is affectively-based computer-aided composition. A common theme running through some of the affective-based systems is the representation of the valence and arousal of a participant's affective state [11]. Valence refers to the positivity or negativity of an affective state; e.g., a high valence affective state is joy or contentment, a low valence one is sadness or anger. Arousal refers to the energy level of the affective state; e.g., joy is a higher arousal affective state than happiness. Until recently the arousal-valence space was a dominant quantitative two-dimensional representation of emotions in research into musical affectivity. More recently, a new theory of emotion with the corresponding scale, referred to as GEMS (Geneva Emotional Musical Scale) has been proposed [9].

Many of the affective-based systems are actually based around re-composition rather than composition; i.e. they focus on how to transform an already composed piece of music to give a different emotional effect – e.g. make it sadder, happier, etc. This is the case with the best known and most thoroughly tested system - the Computational Music Emotion Rule System (CMERS) [11]. The rules for expressing emotions map valence and arousal onto such elements as modes and pitch class. These rules were developed based on the combining a large number of studies by psychologists into music and emotion. However it was found these needed to be supplemented by rules for expressive performance of the transformed music to express the emotion successfully. Hence CMERS is actually an integrated composition and expressive performance system. CMERS key limitation as a composition system is that it is designed for re-composition, not for generating new material.

Oliveira and Cardoso [13] also perform affective transformations on MIDI music, and utilize the valence-arousal approach to affective specification. These are to be mapped on to musical features: tempo, pitch register, musical scales, and instrumentation. A knowledge-base of musical features and emotion was developed based on musical segments with a known affective content. This knowledge-base was then used to train a generalized mapping of affective state to required music and a model was then generated based on Support Vector Machine regression. The model was tested for transforming the emotion of classical music – the current results are not as good as CMERS. One reason for this may be that Oliveira and Cardoso has the limitation that it is unable to generate expressive performances.

Although Legaspi et al. [14] utilize pre-composed music as its heart, it is more focused on composing new music. An affective model is learned based on score fragments manually labeled with their appropriate affective perception – this maps a desired affective state on to a set of musical features. The model is learned based on the machine learning approaches Inductive Logic Programming and Diverse Density Weighting Metric. This is then used as a fitness function for a Genetic Algorithm – however the GA is also constrained by some basic music theory. The GA is then used to generate the basic harmonic structure, and a set of heuristics are used to generate melodies based on the harmonic structure. The system was trained with emotion label dimensions "favourable-unfavourable", "bright-dark", "happy-sad", and "heartrending-not heartrending". Listening tests were done on a series of eight bar tunes and the results obtained were considered promising, but indicated that more development was needed. Once again, the system is lacking the ability to generate expressive performances.

4 Expressive Music Performance

The introduction of MIDI led to an explosion in the use of sequencers and computers, thanks to the new potential for connection and synchronization. These computers and sequencers performed their stored tunes in perfect metronomic time, a performance which sounded "mechanical". They sounded mechanical because human performers normally perform expressively – for example speeding up and slowing down while playing, and changing how loudly they play. The performer's changes in tempo and dynamics, and other subtle musical features, allow them to express a fixed score – hence the term expressive performance. Publications on computer expressive performance of music have lagged behind computer-aided composition by almost quarter of a century. But from the end of the 1980s onwards there was an increasing interest in automated and semi-automated Computer Systems for Expressive Music Performance (CSEMP). A CSEMP is a computer system which – given a score in some form – is able to generate expressive performances of music [15]. For example software for music typesetting will often be used to write a piece of music, but some packages play back the music in a relatively mechanical way – the addition of a CSEMP enables a more "human sounding" playback, giving a better idea of how the final performance may sound. Computer expressive music performance is used in this chapter to make performances sound less mechanical to the user, and thus increase the affective impact, as demonstrated by [11]. The particular system to be utilized is now described.

Director Musices (DM) [2] has been an ongoing project since 1982. Researchers including violinist Lars Fryden developed and tested performance rules using an analysis-by-synthesis method (later using analysis-by-measurement and studying actual performances). Currently there are around 30 rules which are written as relatively simple equations that take as input music features such as height of the current note pitch, the pitch of the current note relative to the key of the piece, or whether the current note is the first or last note of the phrase. The output of the equations defines the performance actions. For example the higher the pitch the louder the note is played, or during an upward run of notes, play the piece faster. Another DM rule is the Phrase Arch which defines a "rainbow" shape of tempo and dynamics over a phrase .The performance speeds up and gets louder towards the centre of a phrase and then tails off again in tempo and dynamics towards the end of the phrase. Each rule in DM can be weighted to give it a greater or lesser relative effect on the performance, by changing a parameter known as its k-value.

DM has also been developed to enable emotion-based expression [16]. Listening experiments were used to define the k-value settings on the DM rules for expressing emotions. The music used was a Swedish nursery rhyme and a computer-generated piece. Six rules were used from DM to generate multiple performances of each piece. Subjects were asked to identify a performance emotion from the list: fear, anger, happiness, sadness, solemnity, tenderness or no-expression. As a result parameters were found for each of the 6 rules which mould the emotion-communicating expression of a piece. For example for "tenderness": inter-onset interval is lengthened by 30%, sound level reduced by 6dB, and two other rules are used: the Final Ritardando rule (slowing down at the end of a piece) and the Duration Contrast rule (if two adjacent notes have contrasting durations, increase this contrast).

5 EEG and Emotion

EEG measurements have been found to be useful in a clinical setting for diagnosing brain damage, sleep conditions and epilepsy; e.g. [17]. It is well known in the literature that it is possible to relate different EEG spectral bandwidths (often referred to as

"EEG rhythms") to certain characteristics of mental states, such as wakefulness, drowsiness, etc. As early as the 1970s researchers have reported on the relationship between EEG asymmetry and affective state. Reviews of EEG asymmetry and affective state can be found in [18, 19] and one of the most recent sets of results can be found in [20]. Davidson [21] proposed a link between asymmetry of frontal alpha activation and the valence and arousal of a participant's affective state.

Musha and co-workers [22] developed one of the earliest computer EEG affective state detection systems and a number of detection methods have been investigated since then; e.g., [23]. More recently detection and analysis of weak synchronization patterns in EEG have been shown to be indicators of cognitive processing; growing evidence suggests that synchronization may be a carrier of information about the information processing in the brain [24]. There are different ways in which signals may co-vary. For instance, there is the hypothesis that information about many cognitive phenomena is preserved not necessarily in the intensity of the activation, but rather in the relationship between different sources of activity. There are an increasing number of studies investigating the role of synchronization in cognitive processing using various techniques, e.g. [25]. A particularly promising form of synchronization is called Phase–locking, which has been studied extensively by the third author and co-workers, e.g. [26]. Moreover, there is growing evidence supporting the role of synchronization in music perception [27] and also in response to affectively charged non-musical stimuli [28].

6 Emotional Feedback EEG Music

The above sections show that there is increasing evidence in the literature that musical traits such as rhythm, melody and tonality, can communicate specific affective states. There is also increasing evidence (e.g. [12]) that these states are detectable in the EEG of the listener. There are fewer studies into establishing which musical traits are useful for implementing a system to *induce* affective states. Amongst the techniques available, the analysis of synchronisation patterns in the EEG signal is a promising option for detecting affective states induced by music. Other techniques (such as frontal asymmetry) will also be considered in the project and the most suitable will be adopted. Thus the detection of affective state by EEG is a research area which this project will contribute to as well.

As was mentioned earlier, [1] investigates the use of EEG to detect emotion in an individual and to then generate emotion-inducing music based on this. The work done previously in [1] was not real-time and did not involve any machine learning process. The research and implementation of a real-time version of a more advanced detection method would allow us to monitor affective states induced by music on the fly. We hypothesise that once we establish – for a given context - specific musical traits associated with specific affective states, then we will be able to parameterise such traits in order to exert control in a musical composition; e.g., speed up the tempo to induce affective state X, use a "harsher" timbre to induce state Y, etc. The parameterisation of musical traits will allow for the design of algorithms capable of generating music (e.g., rule-based) embodying musical traits aimed at inducing specific EEG-observed trajectories correlated to affective states. Such a generative system can be rendered intelligent and adaptive by means of machine learning techniques (e.g., case-based reasoning and reinforcement learning) that are able to learn to recognize complex patterns and make decisions based on detected patterns in real-time.

Our initial results will be driven by more universal musical determinants of emotional response than context-specific. Thus, they will be based on results averaged across a test population. The later stages of the project will extend the former to include context-specific emotional responses. Later stages will also include the more real-time approach to learning and detection. The move towards more on-going assessment of affective state will be important because it will enable us to extend the system beyond the music composition based on manipulation of the musical traits eliciting generic affective responses, to a more adaptive individual-oriented system taking into account participants' states; thus utilising also the contextual effects of an individual and the environment.

7 Affective Structure Prototype

It has been discussed how expressive performance and various compositional musical elements will be tested for their affective impact based on context. On the music and machine learning side of the project, a Matlab prototype has been produced for investigating the effects of one musical element on emotional communication. This element is musical structure. (Communicated emotion analysis is a first step towards induced emotional analysis.)

At the heart of this prototype is a phrase generator that uses random walk with jumps [29] to generate the basic motifs. The phrases produced by the generator are then transformed in pitch height, loudness level and global tempo to investigate affective features. They can also be transformed between major and minor key modes. Once these transforms have been done Director Musices rules are applied. The following rules are utilized: Duration Contrast, Duration Contrast Articulation, Punctuation, High Loud, Phrase Arch, and Motor Errors. Although DM is capable of mimicking emotional expression, the rules are being used here to make the performances sound less mechanical to the user, and thus increase the affective impact [11].

The prototype is embedded in a test-bed which uses a pairs comparison system for ascertaining communicated valence, which is correlated to induced emotion [4, 8]. The user is presented with a piece of monophonic expressively performed music and the user is asked "Which of the following two tunes reminds you of more positive feelings?" This question is designed to ascertain the communicated valence of the tune. The user is given the options of selecting tune 1 or 2, selecting a "don't know" option, or asking to have the tunes played again. Thus at the end of the experiment a series of locally-ordered parameter set pairs (a1, a2) (b1, b2) ... etc. will have been generated. Each pair will be ordered by valence, thus leading to series of inequalities. If sufficient pairs are available for valence, then the inequalities can be used to infer a global ordering for which parameters communicate a greater valence for the user.

An algorithm for inferring the global ordering from the local ordering is incorporated into the test-bed.

The structure-based testing currently involves the following procedure. Benchmark transformations have been assigned for Happy, Sad, Stressed, and Relaxed – based on past research into the area of musical affective communication [11][30]. To initiate a structure test, two phrases are generated. Two benchmark states are randomly selected from the four above. The first phrase is transformed using the first benchmark and the second using the second. Thus a tune consisting of two affective parts is played – for example one Happy and one Sad, or one Stressed and one Happy, etc. It was found that this created a perceptible discontinuity between the two halves, so an interpolation system was developed which approximately interpolated pitches, loudness, tempo and key between the two halves – thus perceptually smoothing the transition. The whole combined theme is then also transformed using the expressive performance algorithms of DM. The original two generated phrases can then be used again and transformed to create a different affective interpolated structure. So there will now be two themes, built from the same initially generated phrases, but with different affective structures. These are then presented to the test subject one after the other, who orders them by communicated valence.

Due to the correlation between induced and communicated affect, this system will help to generate an initial core rule-set for the machine learning algorithm which we are developing. However it is also useful in learning more about the effects of musical structure on affective communication.

8 Conclusions

A new method for utilizing the emotion-inducing nature of music and sound has been introduced. The background elements have been detailed, including affective representation, computer expressive performance, affective algorithmic composition and EEG-based machine learning. Some initial steps in this research have been the development of a test-bed which utilizes computer expressive performance, and investigates the testing of musical structure effects on affective communication. The system uses a pairs-based analysis approach and structural emotion interpolation. This test-bed enables the development of a core rule-set linking musical structure and valence.

Future work in the broader project includes characterising synchrony patterns corresponding to different induced affective states from the EEG recordings while participants listen to music stimuli. Initially, the analysis and the system for learning the emotional control music generation will be developed based on the valence arousal emotional scale, due to its widespread acceptance and availability of tagged databases. We will subsequently develop a GEMS representation and will evaluate the usefulness of the two scales for developing our system.

Then, we shall progressively move towards the final goal of real-time assessment of affective states using reinforcement learning (RL). Initially, the affective state estimation will be updated at a slower time scale consistent with the computational demands of the synchronisation analysis. However, our aim is to create a system for a fast real-time assessment of affective state based on efficient analysis using feature selection and dimensionality reduction.

We plan to develop further algorithms for generating music featuring the various musical traits that have been discussed in the literature. Some musical features are more universal determinants of affective response, invariant across populations with common cultural background [9]. Other features may show more variation dependent on contextual effects of culture, personality and environment. Our initial results will be driven by more universal musical determinants of emotional response than context-specific. Thus, they will be based on results averaged across a test population. The later stages of the project will extend the former to include context-specific emotional responses.

We plan to test our initial generative music algorithms for inductive effects using an offline EEG affective state detector. The results of these tests will be used to initialize a case-based reasoning (CBR) system for affective induction by music. Then, we will extend the CBR system by investigating specific musical genres. A recent study [9] also suggested the importance of genre selection for the induction of certain affective states. The benchmark will be a classical solo piano genre, as classical music has well known computational approaches for eliciting certain affective states, but expansions on this will be investigated utilizing ideas from pop and electroacoustic music genres.

In order to have a real-time, dynamic assessment of the affective state – so as to increase accuracy and effectiveness - we will use the CBR system to initialise an automatic music generation system based on reinforcement learning (RL). RL has been successfully used in optimising the stimulation patterns in deep brain stimulation therapy of the epileptic seizures [31]. The RL system we plan to build will be used in action selection optimizing a desired affective response of this participant. The move towards more on-going assessment of affective state will be important because it will enable us to extend the system beyond the music composition based on manipulation of the musical traits eliciting generic affective responses, to a more adaptive individual-oriented system taking into account participants' states; thus utilising also the contextual effects of an individual and the environment.

Acknowledgements. This research was supported by EPSRC grant EP/J003077/1 "Brain-Computer Interface for Monitoring and Inducing Affective States".

References

1. Kirke, A., Miranda, E.R.: Combining EEG Frontal Asymmetry Studies with Affective Algorithmic Composition and Expressive Performance Models. In: Proceedings of International Computer Music Conference (ICMC 2011), Huddersfield, UK (2011)
2. Nasuto, S.J., Miranda, E.R.: Brain-Computer Interface for Monitoring and Inducing Affective States, EPSRC Grant EP/J002135/1 (2012)
3. Juslin, P., Sloboda, J.: Music and Emotion: Theory and Research. Oxford University Press (2001)

4. Juslin, P., Laukka, P.: Expression, perception, and induction of musical emotion: a review and a questionnaire study of everyday listening. Journal of New Music Research 33, 216–237 (2004)
5. Minassian, C., Gayford, C., Sloboda, J.A.: Optimal experience in musical performance: A survey of young musicians. Annual General Meeting of the Society for Education, Music and Psychology Research (2003)
6. Goethem, A.: The Functions of Music for Affect Regulation. In: International Conference on Music and Emotion, Durham, UK (2009)
7. Juslin, P.: Five Facets of Musical Expression: A Psychologist's Perspective on Music Performance. Psychology of Music 31, 273–302 (2003)
8. Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., Dacquet, A.: Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. Cognition and Emotion 19(8), 1113–1139 (2005)
9. Zentner, M., Grandjean, D., Scherer, K.R.: Emotions evoked by the sound of music: characterization, classification, and measurement. Emotion 8(4), 494–521 (2008)
10. Scherer, K.R.: Which Emotions Can be Induced by Music? What are the Underlying Mechanisms? And How Can we Measure Them? Journal of New Music Research 33(3), 239–251 (2004)
11. Livingstone, S., Muhlberger, R., Brown, A., Thompson, W.F.: Changing Musical Emotion: A Computational Rule System for Modifying Score and Performance. Computer Music Journal 34(1), 41–64 (2010)
12. Schmidt, L.A., Trainor, L.J.: Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. Cognition and Emotion 15(4), 487–500 (2001)
13. Oliveira, A.P., Cardoso, A.: Automatic Manipulation of Music to Express Desired Emotions. In: Proceedings of the 6th Sound and Music Computing Conference, Porto, Portugal, pp. 265–270 (2009)
14. Legaspi, R., Hashimoto, Y., Moriyama, K., Kurihara, S., Numao, M.: Music Compositional Intelligence with an Affective Flavor. In: Proceedings of the 2007 International Conference on Intelligent User Interfaces, Honolulu, Hawaii, USA, pp. 216–224 (2007)
15. Kirke, A., Miranda, E.R.: Guide to Computing for Expressive Music Performance. Springer, UK (2012) (in print)
16. Bresin, R., Friberg, A.: Emotional Coloring of Computer-Controlled Music Performances. Computer Music Journal 24(4), 44–63 (2000)
17. Sheerani, M., Hassan, A., Jan, A., Zaka, R.: Role of Video-EEG Monitoring in the Management of Intractable Seizures and Non-epileptic Spells. Pakistan Journal of Neurological Sciences 2(4), 207–209 (2007)
18. Silberman, E.K., Weingarter, H.: Hemispheric lateralization of functions related to emotion. Brain and Cognition 5(3), 322–353 (1986)
19. Allen, J.J.B., Kline, J.P.: Frontal EEG asymmetry, emotion, and psychopathology: the first, and the next 25 years. Biological Psychology 67(1), 1–5 (2004)
20. Wyczesany, M., Kaiser, J., Coenen, A.M.L.: Subjective mood estimation co-varies with spectral power EEG characteristics. Acta Neurobiologiae Experimentalis 68(2), 180–192 (2008)
21. Davidson, R.J.: The neuropsychology of emotion and affective style. In: Lewis, M., Haviland, J.M. (eds.) Handbook of Emotion. Guilford Press (1993)
22. Musha, T., Terasaki, Y., Haque, H.A., Ivamitsky, G.A.: Feature extraction from EEGs associated with emotions. Art. Life Robotics 1(1), 15–19 (1997)

23. Bos, D.O.: EEG-based Emotion Recognition: The Influence of Visual and Auditory Stimuli (2007), http://hmi.ewi.utwente.nl/verslagen/capita-selecta/CS-Oude_Bos-Danny.pdf
24. Fries, P.: A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends in Cognitive Sciences 9(10), 474–480 (2005)
25. Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R., Hanslmayr, S.: EEG alpha synchronization and functional coupling during top-down processing in a working memory task. Hum. Brain Map. 26(2), 148–155 (2005)
26. Sweeney-Reed, C.M., Nasuto, S.J.: A novel approach to the detection of synchronization in EEG based on empirical mode decomposition. Journal of Computational Neuroscience 23(1), 79–111 (2007)
27. Bhattacharya, J., Petsche, H., Pereda, E.: Long-range synchrony in the gamma band: role in music perception. Journal of Neuroscience 21, 6329–6337 (2001)
28. Hu, M., Li, J., Li, G., Tang, X., Freeman, W.J.: Normal and Hypoxia EEG Recognition Based on a Chaotic Olfactory Model. In: Wang, J., Yi, Z., Żurada, J.M., Lu, B.-L., Yin, H. (eds.) ISNN 2006. LNCS, vol. 3973, pp. 554–559. Springer, Heidelberg (2006)
29. Kirke, A., Miranda, E.R.: An Instance Based Model for Generating Expressive Performance During Composition. In: Proceedings of International Computer Music Conference (ICMC 2008), Belfast, UK (2008)
30. Kirke, A., Miranda, E.R.: Artificial Social Composition: A Multi-Agent System for Composing Music Performances by Emotional Communication. In: Klouche, T. (ed.) Mathematical and Computational Musicology. Springer (2010)
31. Guez, A., Vincent, R.D., Avoli, M., Pineau, J.: Adaptive Treatment of Epilepsy via Batch-mode Reinforcement Learning. In: Proceedings of the 20th Innovative Applications of Artificial Intelligence Conference, pp. 1671–1678 (2008)