Affective Calibration of a Computer-Aided Composition System by Listener Evaluation

Duncan Williams*[1] Alexis Kirke[1], Eduardo Miranda[1],
Ian Daly, James Hallowell, Faustina Hwang, Asad Malik, James Weaver, Slawomir Nasuto#

*Interdisciplinary Centre for Computer Music Research, Plymouth University, UK

#Brain Embodiment Lab, School of Systems Engineering, University of Reading, UK

[1]duncan.williams.plymouth.ac.uk

ABSTRACT

Background

Affectively-driven algorithmic composition (AAC) is an emerging field combining computer music research and psychological approaches to music cognition (Mattek, 2011; Williams et al., 2014). AAC systems attempt to communicate or induce specific emotions in the listener by creating novel music. Using biophysiological readings to determine emotional induction in tandem with AAC provides a possible mechanism for reactive, feedback-driven systems that might reliably induce affective states in the listener in accordance to their own preferences and physiological responses. Here we report on a listener evaluation of one such system under development.

Aims

- To evaluate the composition of novel, affectively-driven music by an AAC system
- To investigate the inter-relationships between various musical features and their subsequent influence on listeners' affective responses to music generated by the AAC system
- To determine the widest range of affective responses that might be targeted by the AAC system

Method

Music has been shown to induce physical responses on a conscious, and unconscious level (Grewe, Nagel, Kopiez, & Altenmüller, 2005, 2007). Such measurements can be used as indicators of affective states. A prototype AAC system designed to attempt to induce specific affective states in the listener – the distinction between emotional communication and emotional induction in listeners is well documented in music psychology (Gabrielsson, 2002; Kallinen & Ravaja, 2006; Scherer, 2004) – was prototyped and evaluated in a series of listening tests. A series of affective mappings (musical features with emotional responses) were drawn from literature and implemented in an artificial intelligence driven AAC system. These mappings utilize sixteen 'modes of operation' which correspond to a time series of musical features with varying ratios, intended to evoke particular affective states on the 2-Dimensional circumplex model of affect (wherein *valence* represents positivity of the affective state, as plotted on the horizontal axis, and *arousal* represents the intensity of the state, plotted on the vertical axis) (Russell, 1980). The prototype AAC system uses these musical feature mappings to inform the generation of new music as a polyphonic piano score, aiming for a variety of affective targets in the 2-D space. This new material is generated by means of a 16-channel feed-forward artificial neural network (ANN). The generated pool of material is then further transformed according to the affective mapping, which gives control over five musical sub-features: *tempo, mode, pitch range, timbre*, and *amplitude envelope*. Specific variations in each of these musical features are used to imply different affective states in the generated material according to a 4x4 cartesian grid across the 2-D space. Co-ordinates with higher arousal generally include a larger pitch spread (range of notes), faster tempo, and harder timbres, whilst co-ordinates with higher valence generally utilize a major key. In this system, a cartesian coordinate of (*arousal [1:4]*, *valence [1:4]*) is used to refer to a given ratio of the five musical features. For example a co-ordinate of (1, 1) would refer to the lowest corner of the space (low *valence* and low *arousal*). This co-ordinate would force the transformation algorithm to create stimuli incorporating a slow tempo, a minor key, a *soft* timbre (on a piano, the timbre of the performance can be manipulated using dynamics markings, where perceptually *harder* sounds are achieved with louder performance dynamics), an amplitude envelope with considerable legato, and a spread of pitch values which are comparatively lower than those of the rest of the generated pool. The complete stimulus set is shown in Table 1.

Table 1 Stimulus set showing musical feature matrix

Stimulus #	Timbre	Key	Pitch spread	Tempo	Envelope
1-16	*Soft*	Minor	Low	Slow	Legato
17-32	*Soft*	Chromatic	Medium	Slow	Legato
33-48	*Soft*	Major	High	Slow	Legato
49-64	*Medium*	Minor	Low	Medium	None
65-80	*Medium*	Chromatic	Medium	Medium	None
81-96	*Medium*	Major	High	Medium	None
97-112	*Hard*	Minor	Low	Fast	Staccato
113-128	*Hard*	Chromatic	Medium	Fast	Staccato
129-144	*Hard*	Major	High	Fast	Staccato

A tri-stage listener evaluation was then used to inform two levels of subsequent adjustment to the feature mappings until a broad spectrum of emotional responses was achieved.

Results

The size and spread of affective responses was gradually increased by deliberate manipulation of the ratio of musical features in the affective mappings, until the mean listener responses showed a wide variety of affective states could be achieved by the mappings, as indicated in Figure 1.

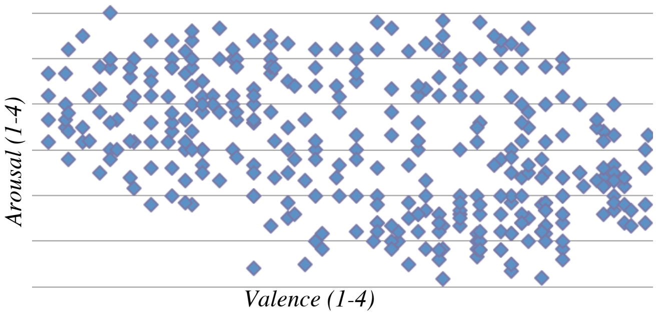

Figure 1. Mean listener responses to final stimulus set after selective manipulation of ratios of musical features in affect transformation matrix. Note uniform spread with the exception of the far left of the space (low arousal, low valence) and the top right of the space (high arousal, high valence).

Conclusions

Listener evaluation of an AAC system confirmed that affective targeting via novel musical material is possible by means of an underlying matrix of musical features. Should such a system be adaptable to real-time control by means of biophysiological estimates of affective state, a feedback driven AAC system could be created for continuous monitoring and induction of target affective states in the listener (e.g., for therapeutic means). In this work the particular ratio of these features has been explored and adjusted in response to listener evaluation but the complex nature of the inter-relationship between these musical features, and the subsequent affective responses remains the subject of considerable further work.

Keywords

Emotion, Affect, Artificial Intelligence (AI), Affective Algorithmic Composition (AAC)

Acknowledgements

This work was supported by EPSRC grants (EP/J002135/1 and EP/J003077/1).

REFERENCES

Gabrielsson, A. (2002). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, *5*(1 suppl), 123–147.

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2005). How Does Music Arouse “Chills”? *Annals of the New York Academy of Sciences*, *1060*(1), 446–449.

Grewe, O., Nagel, F., Kopiez, R., & Altenmüller, E. (2007). Emotions over time: Synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion*, *7*(4), 774–788. http://doi.org/10.1037/1528-3542.7.4.774

Kallinen, K., & Ravaja, N. (2006). Emotion perceived and emotion felt: Same and different. *Musicae Scientiae*, *10*(2), 191–213. http://doi.org/10.1177/102986490601000203

Mattek, A. (2011). Emotional Communication in Computer Generated Music: Experimenting with Affective Algorithms. In *Proceedings of the 26th Annual Conference of the Society for Electro-Acoustic Music in the United States*. Miami, Florida: University of Miami Frost School of Music.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161.

Scherer, K. R. (2004). 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.

Williams, D., Kirke, A., Miranda, E. R., Roesch, E., Daly, I., & Nasuto, S. (2014). Investigating affect in algorithmic composition systems. *Psychology of Music*. http://doi.org/10.1177/0305735614543282