Developing Methods for Predicting Affect in Algorithmic Composition

Daniel Pitman

B.Mus. (Hons) 2012

Submitted in partial fulfillment of the requirements for the degree of

Master of Philosophy

Elder Conservatorium of Music
Faculty of Arts

University of Adelaide

June 2015

Contents

Abstract

Affective Algorithmic Composition (AAC) is a field that focuses on the algorithmic generation of music specifically to affect its audience in a targeted way.

This thesis presents a novel method for developing AAC systems based on collecting both perceived and induced affect data from human participants using multiple biosensor and surevey approaches, and modelling the resulting data in a predictive function based on a neural network. This in turn is used to drive the musical algorithm to generate music that can invoke any specified affective target.

These various approaches to affect measurement can be assessed and compared by their respective predictive error when used to train a neural network, providing an assessment tool for further refinement and development.

A pilot study of this method is also presented, The Affective Algorithmic Composer (AACr). AACr's predictive functions are trained using multiple forms of affect data collected from a group of participants, and can generate original music to invoke specific emotional states, physiological states, perceived content, and themes. Several generated compositions are included to demonstrate the abilities of the AACr to invoke affective states defined manually or directly taken from the user via biosensors.

The thesis concludes by reflecting on the method's strengths, areas for further development, and methods that could be used to determine the success of future AAC systems.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Daniel Pitman

2015

Acknowledgements

My supervisors Dr. Luke Harrald and Mr. Stephen Whittington, for their immense efforts and patience

Assc.Prof. Kimi Coaldrake whose guidance was so often critical

The Electronic Music Unit and Elder Conservatorium staff who offered assistance, critique, or patience

The Pain and Anaesthesia Research Clinic, Royal Adelaide Hospital, for access to and training with their EEG facilities

Friends and Family: Barry and Therese, Candice, Robert, Iran, Dan, and Meredith, for enduring support and/or proofreading

Finally I must express an immense appreciation to the volunteers (who must remain anonymous) who patiently auditioned and auditioned musical samples with uncanny resolve and no reward, yet were more critical to the project's success than anyone. -This project is literally made of you, and I can't appreciate you all enough for trusting me with your 'is-ness' in this whole process.

List of Acronyms

AAC: Affective Algorithmic Composition. A field of study where algorithmic composition is specifically designed to invoke a target affective state in the listener.

AACr: Affective Algorithmic Composer. The software and hardware developed in this study.

BCI: Brain Control Interface: A field of study, separate to AAC, where brain sensors are used to control computer systems.

BCMI: Brain Control Musical Interface. A field of study, separate to AAC, where brain sensors are used to control computer systems for music.

BIO: Biosensor (not including EEG). A descriptive abbreviation used in the AACr GUI.

BR: Bayesian Regularisation. A learning function used in neural networking

CAT: Computed Axial Tomography scanning. A brain scanning technology used in hospitals and laboratories involving X-rays.

DIY: Do It Yourself. A common term for enthusiast projects built at home or as a hobby.

ECG: Electrocardiogram. (sometimes EKG from Latin, ‘kardia’) A sensor that reports heart rate by measuring the field generated from the electrical impulses of the heart muscle.

EEG: Electroencephalogram. A sensor (or collection of sensors) that measures electric fields created by the brain’s neurons firing. Also used in the AACr GUI as a descriptive abbreviation.

EMG: Electromyography. A technique for measuring the electrical field changes created by muscles and their associated nerves.

EOG: Electrooculogram. A device that measures electrical field changes created by movement of the eyeball.

ERP: Event Related Potential. In EEG, a specific and often expected change in signals in reaction to a stimulus.

EVM: Eulerian Video Magnification. A process where the most subtle temporal changes in a digital video file are magnified to become visible, such as the change in complexion due to heart beat, or a vibrating guitar string.

fMRI: functional Magnetic Resonance Imaging. A brain scanning technology that is very common, but requires relatively large equipment and a magnetically isolated room.

GP: Genetic Programming. Specifically in this paper in regards to GP as a method of implementing symbolic regression, where solutions to a curve are formed using random symbols as pieces, and improved using a fitness routine.

GSR: Galvanic Skin Response. Another name for skin conductance.

GUI: Graphical User Interface. That part of a program that presents controls and information to the user via the screen.

HRV: Heart Rate Variance. The amount of variation from the average period of heart beats from a given sample.

IBI: Inter-Beat Intervals. The period in between each heartbeat.

LED: Light Emitting Diode. Simple, polarised, light emitting electrical component.

LIK: Likert. A descriptive abbreviation used in the AACr GUI.

LM: Levenberg-Marquardt. A learning function used in neural networking.

MEG: Magnetoencephalogram. A brain scanning technology measuring magnetic field changes, also requiring a magnetically isolated room.

MIDI: Musical Instrument Digital Interface. A simple and very common language/protocol for controlling instruments and synthesizers via digital commands.

MSE: Mean Squared Error. The average squares of the errors between data and a function trying to fit that data.

NN: Neural Network. An umbrella term for machine learning systems that emulate neural processing as found in organic brain networks.

NNS: Neural Network Server. A module of the Affective Algorithmic Composer system, described herein, that uses neural networks to calculates potential musical candidates for a given affective targets.

OSC: Open Sound Control. A sound control interface protocol that uses TCP-IP addresses and can communicate via ethernet.

R: (or R value) The correlation coefficient of two variables. If R is close to zero, the two variables in question are not related. As R approaches 1, the more the variables are related.

SCG: Scaled Conjugate Grading. A learning function used in neural networking.

SCL: Skin Conductance Level. The mean value of conductance between two electrodes on the skin of a set period of time.

SCR: Skin Conductance Response. A time measurement of the period between the beginning of a skin conductance event and the point it reaches half way back to the original level.

SQL: SQLite. An open and simple database language, natively implemented in Cycling’74’s *Max* software development environment.

Glossary of Terms

Affect: To have influence on, or potential to cause change.

Affective Target: A defined change in affective state that is required. In the AACr this involves a shopping cart type system, offering specific changes the user might like to attempt to invoke in the listener.

Algorithmic Composition: A method of composition, typically experimental, that involves defining an audible rendering of a mathematical formula or procedure rather then defining the notes themselves.

Arduino: One of many brands of small microprocessor\circuit boards that enables an electronic circuit to interface with a computer via USB or wirelessly. Other brands mentioned include *Teensy* and *OpenEEG*.

Array: A computer programming term for a list of numbers, similar to a matrix or multi-dimensional grid.

Biomusicology: A field of study that focuses on biological musical phenomena, including evolution, physiology, and empirical studies.

Biosensor Battery: Often a collection of different bio-sensors are collectively referred to as a battery, much like artillery.

Boolean: A method of mathematical analysis commonly used in programming that returns either ‘true’ or ‘false’. For example, if a = 1, and b = 2, then “a < b” is true and “a > b” is false.

Computational Critic: A part of a computer program that is responsible for assessing potential output (in this case musical passages) against the context of a defined target, usually using machine learning or some form of regression.

Induced Affect: That affect which causes changes in the listener both physiologically and psychologically. For example, a song that triggers the emotion of sadness in a listener whether the listener thinks the song is about being sad or not.

External Factors/Confounding Factors: May refer to anything which causes measured change in the participant other than the musical stimulus itself (distraction), influences that change (such as alcohol), or which causes the participant to react in ways beyond their personality/heritage (traumatic bias towards piano, illness).

Heuristics: A term for methods by which a problem can be approached and solved, typically used in computer programming and machine learning. For example, different search engines use different heuristics and return different results to the same search line.

Hierarchical Analysis: A traditional approach for analysing music using a range of levels or tiers of structure. Each tier influences the tier below it from the highest, overall form, right down to the lowest, individual motifs and notes.

Likert: A method of survey that employs answers using scales.

Music Affect: Both a contemporary and historical field of study, involving the influence of music on the human being, as well as a term for the phenomenon itself. For practical differentiation, in this study, the phenomenon studied is referred to as musical affect.

Music Algorithm: A system (typically a computer program) specifically designed to procedurally generate music.

Neutral Period: A period of time where no musical stimulus is presented at all.

Perceived Affect: Affect which is reported by the listener in terms of being expressed by the music. For example, “A song about being sad” does not necessarily trigger sadness in the listener.

Phasic: An event that occurs in phases, involves a series of smaller events, or is temporal in nature. A phasic change might involve the change of the temporal, periodical, or dynamic nature of the phenomena in question.

Predictive Function: Any mathematical or programming action that analyses a data set and extrapolates possible solutions to new enquiries, typically employing regression and machine learning techniques.

Salient: Noticeable or important details

Seed: See Structural Array

Stochastic Data: Data using samples, which inherently contains errors or noisy variation to some degree.

Structural Array: Also called a ‘seed’. Both terms are specific to this project. A structural array is the term used to refer to a list of numbers that represent all of the hierarchical variables needed to have the music algorithm generate a piece of music.

Synaesthesia: A physiological condition where various receptive systems (sight, hearing, smell etc.) are confused or influence other receptive systems in the brain. Commonly symptoms include associating colour with certain aspects music or timbre, or having texture strongly associated with certain smells.

Tag-word: A single word associated with a data entry, effectively forming a category so that all entries with a specific tag-word can be recalled using that tag-word.

List of Figures & Tables