David Rector
COSC 174 Project Proposal
Winter 2013

Local Musical Features and Their Physiological Affect

Problem.
For as long as man has produced musical sounds, he has attempted to divine the rules relating those sounds to the affect they produce in him. Modern computing, audio recording, and biometric data collection technology has seemingly made it possible to more precisely relate musical features to their corresponding physiological responses; however, few studies have explored anything more complex than the simplest such relationships (e.g., the emotional valence associated with major vs. minor chords). Legitimate reasons exist for this hesitancy: 1) it is difficult to know, much less to prove, how similarly music is perceived across cultures or even across a single culture; 2) a comprehensive quantitative answer to the question is undoubtedly very complex, perhaps incalculably so; and 3) collecting repeatable data on subtle, nuanced responses is difficult given a) the probability of high baseline noise in biometric data and b) the difficulty of replicating the precise listening conditions seemingly necessary to obtain similar results – putting a listener in very closely the same state as another listener at another time requires a very carefully designed study.

Proposal.
This project aims to be an exploratory study in which biometric data will be collected at many time slices during music listening, feature vectors containing music theory and audio spectral information at each time slice will be produced, and a machine learning algorithm will be used to determine the parameters relating the two. The study will attempt to venture into this "mushy" field on relatively firm ground by isolating as many variables as possible. If interesting results are obtained, however, they will serve primarily as an aid in the design of broader studies, not as proof of any generalizable phenomena in and of themselves.

Methods and Data Set.
The experimenter will himself be the sole test subject. The music to which he will listen will be the full set or some arbitrary subset of the 371 Bach chorales, each adjusted to be in the same key and each played back via identical organ samples on a computer program at the same tempo. This music set is chosen not only because of its likelihood in producing some degree of affect, but because a) the experimenter is not particularly familiar with it (and will exclude any examples he feels he recalls from the past); b) it is solo instrumental music, all of similar tempo and mood, isolating from consideration the possibility of affect-altering lyrics, tempi, instruments or instrument combinations, etc.; and c) very similar sounding, and realistic-sounding, audio files of different chorales can easily be generated from midi versions of the pieces, isolating from consideration affect produced solely by features of a certain performance or recording technique.

The experimenter will ensure each chorale is listened to twice, or possibly more, time permitting, to ensure repeatable data. Specifically, he will independently randomly arrange two (or more) sequences containing the computer-generated audio files of the chorales, and will append the sequences to create a large playlist. He will conduct listening sessions, over the course of two weeks, every morning for approximately one hour, beginning each time at the spot in playlist at which he stopped the session before. He will listen via headphones at a consistent gain level and in a room isolated from external noise. Each session will occur a set time after a consistent period of sleep and a consistent serving of

food and caffeine. (If ever this routine is varied, or he feels his baseline emotional level is substantially different from usual, he will skip the session for that day.) He will attempt attentive listening, performing no other task during each listening session. Four seconds of silence, followed by seven seconds of pink noise faded in then out, followed by four more seconds of silence, will be added between each track, so as to "cleanse the emotional palette" maximally.
Before the listening sessions, he will determine the biometric sensor available to him that seems to correlate the most with perceived affect. This will be done by listening to several affective pieces of music (not within the set to be used during data collection) with as many biometric sensors as he can obtain – galvanic skin response, heart rate, EEG, etc. - and observing the results after listening for the device that recorded the greatest "arousal" at moments (or shortly after moments) remembered as affective. This sensor (or perhaps a combination – e.g. a weighted sum or sensors) will serve as the "y" data.

The "x" data, or feature vector at each point in time, will be considered and designed only after listening, so as not to excessively pollute the listening process by priming the mind to listen for a certain feature instead of listening generally. Examples of features likely to be used are "spectral roughness" (a measure of the combined dissonances of every pair of component frequencies above a certain threshold at each time slice, weighted by their amplitudes), and relationship to the "tonic" chord (a music theory concept). However, several further issues will be considered in constructing this feature set. First, the lag time between a musical feature and its physiological effect: if there seems to be a consistent lag time between a musical feature perceived to be affective and an apparent arousal in the biometric data, the feature set corresponding to each time period will simply be "shifted" by that lag time, so its label is an earlier place in the music. If, however, a variable lag time seems to be present, it may be advisable to include feature components from previous time points (-250 ms, -500 ms, etc. up to a certain threshold) in each feature vector to accurately account for delayed effects. Additionally, some features having nothing to do with music theory or audio spectral features should be included, such as the time since the beginning of a listening session, so that one may, for example, perceive if fatigue affects physiological response; including such features will both expose and possibly minimize any skewing of results due to such effects.

The ultimate feature vector will mostly be a linear set of the possible features, but introspection, research, and experimentation may result in the addition of some features that are combinations of other features. However, a full quadratic feature vector, for example, will be likely be far too large. The design of the feature vector and algorithm may also be adjusted based on concepts learned later in the course.

Milestone
The biggest challenge will be proper data collection; this must be accomplished by the milestone and done first, so that consideration of the features does not prime and pollute the mind (and thus the biometric data) during listening/data collection. It is also advisable to have begun some consideration and calculation of the components of the feature vector by the milestone.