Mood Dependent Music Generator

Marco Scirea

IT University of Copenhagen

Abstract. Music is one of the most expressive media to show and manipulate emotions, but there have been few studies on how to generate music connected to emotions.

Such studies have always been shunned upon by musicians affirming that a machine cannot create expressive music, as it's the composer's and player's experiences and emotions that get poured into the piece. At the same time another problem is that music is highly complicated (and subjective) and finding out which elements transmit certain emotions is not an easy task.

This demo wants to show how the manipulation of a set of features can actually change the mood the music transmits, hopefully awakening an interest in this area of research.

1 Music Mood Taxonomy

The first issue we had to deal with mood generation was the mood taxonomy. The set of adjectives that describe music mood and emotional response is immense and there is no accepted standard; for example Katayose *et al.*[4] use a set of adjectives including *Gloomy, Serious, Pathetic* and *Urbane.*

Russell [9] proposed a model of affect based on two bipolar dimensions: *pleasant-unpleasant* and *arousal-sleepy*, theorizing that each affect word can be mapped in this bi-dimensional space by a combination of these two components. Thayer [10] applied Russel's model to music using as dimensions *stress* and *arousal*; although the name of the dimensions is different from Russell's their meaning is basically the same. Also, we find different names in different research while the semantic value stays the same. Since valence and arousal are the most commonly used terms in many affective computing research, we will use these terms in this paper.

Thus the music is divided in four clusters: **Anxious/Frantic** (Low Valence, High Arousal), **Depression** (Low Valence, Low Arousal), **Contentment** (High Valence, Low Arousal) and **Exuberance** (High Valence, High Arousal).

These four clusters have the advantage of being explicit and discriminable; also they are the basic music-induced moods (even if with different names) as discovered by Kreutz [5] (*Happiness*, *Sadness*, *Desire* and *Unrest*) and Lindstrom [6] (*Joy*, *Sadness*, *Anxiety* and *Calm*).

2 The Mood Modifying Features

Following Liu *et al.*[7], our current system employs three factors as the features that characterize mood: **Intensity**, **Timbre** and **Rhythm**. Liu's study was

D. Reidsma, H. Katayose, and A. Nijholt (Eds.): ACE 2013, LNCS 8253, pp. 626–629, 2013.

about mood information extraction, we applied the principles they worked with to instead generate music.

2.1 Intensity

Intensity is defined by how strong the volume of the music is; in music with low arousal we generally have a lower volume than in the ones that have high arousal.

2.2 Timbre

Timbre is what we could call the brightness of the music: how much of the audio signal is composed by bass frequencies. In previous literature MFCC (Mel-Frequency Cepstral Coefficients) and spectral shape features have been used to analyze this. For example the brightness of *Exuberance* music is generally higher than in *Depression*, this will result in greater spectral energy in the high subbands for *Exuberance*. It's generally a factor that is very dependent on the instrumentation choice; in our case we act on the synthesizers, our instruments, to generate brighter and darker sounds.

2.3 Rhythm

Finally Rhythm has been divided in rhythm strength, regularity and tempo [7]. For example in a high valence/high arousal piece of music we can observe that the rhythm is strong and steady, while in a low valence/low arousal the tempo is slow and the rhythm cannot be as easily recognized. We act on these features in different ways. To influence rhythm strength we change how much the drums are prominent in the music.

Having the notes generators create notes on the beat or the upbeat creates different feeling of regularity and irregularity, for example in Contentment music we will favor a steady rhythm with notes falling on the beats of the measure while in a Depression one we will give more space to upbeat notes. Finally to influence the tempo we just act on the BPMs (Beats Per Minute) of the music.

2.4 Remarks

We noticed that these features, originally devised to extract mood information, were enough to generate different moods. But we also realized that we could strengthen the impression by introducing dissonances in the music: for *Exuberance* and *Contentment* we use a diatonic scale while for *Anxious* and *Depression* an alterated one.

3 The System

The demo has been realized using PD (aka Pure Data) [8], a real-time graphical programming environment for audio, video, and graphical processing. In PD, programs are written as graphical graphs called *patches*, in our demo we used some patches taken from Brinkmann's website [3].

The generated music is played by 3 instruments (synthesizers) and a drum machine. The system consists of five random number generators: four of these numbers will be converted into notes for the instruments and the drums, while the last one is used as a sound effect controller. Additionally each of the 4 notes generators also generate another number that will determine the volume the note will be played.

These generators create semi-random numbers by adding (in decimal) from a start value a certain step each tick. Then each generator converts the number to another base and adds the digits of the result. This value will then become the note we'll play; at the same time through a slightly different sum we generate another number that will control the volume the note will be played.

The numbers, before being sent to the synthesizers that will generate the note, are filtered so that we can control which notes we want and which we don't. This allows us to use dissonances or maintain a diatonic feel to emphasize moods. The numbers at this point represent notes in MIDI (Musical Instrument Digital Interface) notation (they can span from 0 to 127), so to filter them we just use a modulo operation to understand what note they represent. For example, as there are 12 notes in an octave, we can see that 60 modulo 12 equals 0, this means that 60 represents a C.

At this point the synthesizers generate the note, by converting the MIDI notation to a frequency and using PD's built-in audio wave generator. The synthesizers we use are constructed in a way that we can choose the waveform of the sound, modulation and if we want to transpose octaves.

After the notes are actually generated we have a patch that controls their volumes so that we can decide if we want some instruments more prominent than the others: a mixer. Now we have all the notes at the desired intensity but before playing them we apply some effects.

At the moment we have implemented only the music generation for the four quadrants of the bi-dimensional space, we plan to expand it so that the music can be generated by choosing a point in the plane.

4 Potential Application Domains and Future Work

This project generates music in real time with seamless transition between moods, this could be used in the making of games (where, unlike in a movie, events unfold in response to the player's input) to have a dynamic soundtrack. This can be done by using libraries to interface the game engine and PD. Also, music therapy could be another possible application since music manages to transmit moods immediately and in a non ambiguous way (at least for the four quadrants). Interesting future work includes the usage of words as input, using sentiment analysis tools such as SentiWordNet [1] to extract emotional value from them. This could in the end allow us to create music generated by texts that would reflect the different moods inside it. In the future we'll have a population sample describe their emotional state using SAM (Self Assessment Manikin)

[2] before and after interacting with the demo, this will yield important data in understanding how effective our generator is.

5 The User Experience of the Demo

The demo doesn't require any specific hardware, only a computer with Pure Data installed and speakers. The users will be able to interact with it by specifying the mood they are in (or they want to listen to) and the program will generate the music.

Ideally we will have the user's mood extracted by facial recognition, but at the time of this writing this part has not been implemented yet. By interacting with the demo we hope the users to get interested in experimenting with dynamic music in various media.

The demo can be downloaded at `http://goo.gl/TAW2K9`.

References

1. Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In: Proceedings of the 7th Conference on International Language Resources and Evaluation (LREC 2010), Valletta, Malta (May 2010)
2. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry 25(1), 49–59 (1994)
3. Brinkmann, M.: Pure data patches, `http://www.martin-brinkmann.de/`
4. Katayose, H., Imai, M., Inokuchi, S.: Sentiment extraction in music. In: 9th International Conference on Pattern Recognition, 1988, pp. 1083–1087. IEEE (1988)
5. Kreutz, G., Ott, U., Teichmann, D., Osawa, P., Vaitl, D.: Using music to induce emotions: Influences of musical preference and absorption. Psychology of Music 36(1), 101–126 (2008)
6. Lindström, E., Juslin, P.N., Bresin, R., Williamon, A.: Expressivity comes from within your soul: A questionnaire study of music students' perspectives on expressivity. Research Studies in Music Education 20(1), 23–47 (2003)
7. Liu, D., Lu, L., Zhang, H.-J.: Automatic mood detection from acoustic music data. In: Proceedings of the International Symposium on Music Information Retrieval, pp. 81–87 (2003)
8. Puckette, M., et al.: Pure data: another integrated computer music environment. In: Proceedings of the Second Intercollege Computer Music Concerts, pp. 37–41 (1996)
9. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology 39(6), 1161–1178 (1980)
10. Thayer, R.E.: The biopsychology of mood and arousal. Oxford University Press on Demand (1989)