Exploring the Benefits of Interaction with Audio Continuators

Amar Lal
New York University
amarmalal@gmail.com

Introduction

I recently completed my Master’s thesis in Music Technology, which focused on creating an Interactive Audio Continuator (IAC) (Lal, 2014). The Continuator is an interactive music system proposed by (Pachet, 2002), which automatically learns musical patterns from MIDI information using Variable-Length Markov Chains (VLMC). A user is then allowed to interact with the learned patterns in order to automatically create new, stylistically consistent music. In the work of (Kosta, 2012) and (Marchini, 2010), VLMC are applied to timbre and pitch features extracted from audio in order to create Audio Continuators; however, these systems are not interactive. In my work, I re-created the majority of Kosta’s framework, improved some of the machine learning methods, and proposed a paradigm and basic proof-of-concept for interacting with Audio Continuators. I would like to continue researching and exploring interaction with IACs, and propose that interacting with these systems may yield educational and therapeutic benefits.

Background

I. Educational Potential

After creating the Continuator, Pachet created a number of studies, evaluations, and reflections on his work, most notably (Pachet, 2004) and (Addessi, Ferarri, Carlotti & Pachet, 2006). In the former, Pachet frames his systems as what he calls Interactive Reflexive Music Systems (IRMS). The focus of IRMS research is not the quality of the music these systems create, but the “difference between what is produced with the system and what the user would produce without it” (Pachet 2004). Pachet defines several modes of interaction with IRMS; for our purposes, we focus on the Collaborative mode, in which an IRMS plays a stream of music based on learned material, which the user can then play along with or interact with, shaping and refining the momentary material. Interacting with an IRMS in this way can lead to a Flow experience, as discussed in (Addessi et al, 2006). Children interacting with variations of the Continuator experience focused attention, concentration, clear-cut feedback, control of a (musical) situation, excitement, change in the perception of time, and involvement, which are all hallmarks of Csiksentmihalyi’s theory of Flow.

Increased Flow has been linked to heightened creativity in musical settings (MacDonald, 2006). This can be valuable from an educational standpoint, where the emphasis generally lies on music playing, and lacks development in the areas of music making and improvisation, as discussed by MacDonald. There exists a demonstrated link
between learning about improvisation and quality of musical creations (Koutsoupidou & Hargreaves, 2009), so we propose that interaction with the malleable momentary material created by an IRMS could be used to develop improvisation, and thus music-making skills in children.

II. Therapeutic Potential

In their work investigating timbre, (Zacharakis, Reiss, Pastiadis, & Papadelis, 2011) attempt to determine a group of salient timbre descriptors as agreed upon by a focus group and statistical analysis. Their findings show that timbre is a complex language, and that listeners often use very different language to interpret similar timbral phenomena. In this way, timbre perception is highly subjective, and professional musicians often exploit this precept, using timbral variation to convey emotion and affect, metaphors, and other imagery in musical performances (Holmes, 2011).

Traditionally, this kind of expressive communication via timbre has only been available to instrumental professionals who have learned to produce a range of sounds from their instruments. However, with the advent of widely available synthesis programs for computers and mobile devices, the infinite timbral possibilities afforded by digital synthesis are now accessible to an increasing population. Applications contain progressively user-friendly interfaces for the control and manipulation of these infinite timbres, and more and more people are discovering the possibilities of creating music that communicates expressivity via timbre.

The system created in (Lal, 2014) allows for the timbre of automatically generated momentary material to be manipulated expressively. We propose that, with the correct interface, it could be possible to allow people without musical training to communicate expressively through the manipulation of momentary timbre generated by an IAC. The system’s timbral manipulation leverages the concept of Feature Modulation Synthesis (FMS), as proposed by (Park, 2007), and uses a simple scaling of the timbral envelope to achieve expressivity. This could be controlled by a simple interface with a single continuous output.

In the setting of music therapy, it seems useful to strive for the most transparent interface possible, one that allows immediate and uninhibited expression using intuitive gestures and ideas. Music and interaction in music therapy need to be “scaled to [a patient’s] cognitive and emotional capacity” (Ansdell, 1995). One potential interface that meets these needs is the Squeezables; malleable, handheld balls that allow continuous control over musical parameters (Weinberg & Gan, 2001). They are interesting from a transparency standpoint, in that the balls are not dissimilar to stress balls. Being able to convey tension or stress through squeezing a stress ball seems intuitive and transparent, and the Squeezables and IAC paradigms allow this gesture to be meaningfully sonified.

The IAC system is also rhythm- and tempo-agnostic, allowing these decisions to be made by the user. However, we propose that in the context of music therapy, these be tuned to the patient. This concept is described as entrainment; by mirroring a patient’s “fundamental rhythms (pulse and breathing), [one] can come alongside … aspect[s] of action and identity that a person has” (Ansdell, 1995).

It is important to note, that in the famous “Herr G.” case that entrainment arose from, melody is noted as being the “uniquely musical and therefore human element” (Ansdell, 1995). Creating an “auditory image” of the patient’s pain through such melodies is part of
both the assessment phase and the patient’s awareness acquirement phase in music therapy (Dileo & Bradt, 1999). This melodic material could be automatically generated material that is stylistically consistent with music that the patient identifies with, is familiar with, or feels expresses their desired emotions.

So, by tying the IAC’s automatically generated melodic material (based on the preferences of the patient) with the ability to manipulate its timbre using a Squeezable-like transparent interface, and the concept of entraining the rhythms of the music to the fundamental rhythms of the patient, the IAC could present a powerful framework around which to build music therapy applications.

III. Collaborative Potential

In both instructive and therapeutic settings, group music making is known to be beneficial. (Sawyer, 2006) discusses the principle of Emergence in group settings, in which musicians may be inspired to play things they would not have been able to play alone, or would not have thought of without the inspiration of the group. This mirrors Pachet’s IRMS research intents, uncovering what can be played with the system that would not otherwise be possible.

The jam2jam system is a collaborative ‘jamming’ tool that allows for musicians to interact with each other over the internet (Dillon, 2006). The creators posit that this fosters a “contemporary musicianship, in which the computer is embraced as an instrument that can be used skillfully in live performance with both acoustic/electric instruments and other network users.” Collaboration over the internet allows for new musical collaborations that would not have otherwise been possible, making group music making accessible to the masses.

In his most recent work, Pachet has expanded the Continuator paradigm to be an interactive group experience with the VirtualBand (Moreira, Roy & Pachet, 2013). The system leverages source-separation techniques from the MIR community to separate an audio file into multiple sound sources, before modeling the stylistic patterns of each source. The interdependencies between sources are complex in this modeling process, but the VirtualBand presents an exciting collaborative direction for IACs.

By leveraging this concept of source separation before computational modeling in an IAC, and interacting with the resulting models collaboratively over a network as in the jam2jam paradigm, it may be possible to foster Dillon’s “contemporary musicianship” and enable Sawyer’s Emergence principle, further enriching music education and therapy experiences involving IACs.

Goals

The primary goal of this research is to explore the use of interaction with Interactive Audio Continuators to enhance music education and music therapy experiences. By deploying interaction prototypes in educational settings, we hope to teach children about the malleability of musical form and momentary material in order to foster music-making and improvisation skills. Our interaction prototypes for music therapy settings will leverage the concept of entrainment and the interface paradigm of the Squeezables to allow patients to communicate their pain via the manipulation of musical material that is generated in sync with their fundamental rhythms. Finally, we hope to expand our
prototypes to be both portable (leveraging the recent popularity of embedded Linux systems) and to have the potential to be linked for collaboration either locally or over the internet.

Preliminary Bibliography


