

Applicant

Cornelius Pöpel (cp@khm.de)
Academy of Media Arts Cologne
Peter-Welter-Platz 2
D-50676 Köln

Host Institution

Music Technology Area
Schulich School of Music, McGill University
550 Sherbrooke Street West
Montreal, Quebec, H3A 1E3
Canada
Supervisor: Prof. Dr. Gary Scavone (gary@music.mcgill.ca)

Period: 10th – 30th of August 2006

Purpose of the visit:

1. To investigate audio signal parameter extraction for the generation of control parameters for common synthesis methods. It was planned to investigate a formal definition of the parameter noisiness and to develop methods of tracking it. Further, it was planned to list playing gestures and relations to the measured signal parameters pitch, amplitude, noisiness and spectral centroid in order to get a better view on how an intuitive mapping between measured parameters of an audio signal and the input of a sound synthesis algorithm could be accomplished.
2. To develop synthesis methods (by the modification of existing methods) that can be mainly driven with the raw audio signal as a "control parameter".

I wanted to discuss these topics and evolving questions with the high expertise of the members and professors in the Music Technology Area, particular with Gary Scavone, Vincent Verfaille, Marcelo M. Wanderley and Philippe Depalle.

Work carried out:

The first part of the STMS included an introduction to the lab and its facilities, discussions of the aims and objectives of the STSM, as well as the commencement of the research indicated in item #1 above. Due to the short amount of time available, we chose to make use of the existing sound synthesis software systems MaxMSP and STK (Synthesis ToolKit). Tasks carried out in the first phase were:

1. Parameter extraction from the signal of an electric viola using existing Max/MSP objects and an analysis of the physical gestures they correspond to.
2. Check which of the parameters I can best control when playing the electric viola.
3. Try out different mappings (of those parameters) to control STK instruments.

MSP objects to extract signal parameters were collected and compared. The objects were: fiddle~, yin~, analyzer~, noisiness~, brightness~, IRCAM iana object. The parameters that could

be extracted were pitch, amplitude, noisiness (spectral flatness measure), brightness (spectral centroid), and partials. It is obviously no problem to control pitch and amplitude with a viola signal. We expected that it would be similar with brightness and noisiness. However, recording signals with different performed amounts of brightness and noisiness (the author was the only performer and sounds were played more or less bright or noisy according to a subjective interpretation of those descriptors) showed different results:

1. Spectral flatness measurements (noisiness) did not well correlate to different amounts of performed noisiness.
2. Spectral centroid (brightness) did correspond in a small band (varying between 1700 Hz and 2200 Hz when playing the note g' (fundamental 391 Hz)) with different amounts of performed brightness.

This led to the question how noisiness could be defined and what a player might understand to be noisiness. To investigate further, definitions of noisiness from different fields of science were collected and discussed in the context of the above results. One idea was that noisiness may be understood by the player in parallel to its physical definition, but that it does not necessarily have to be. It might be appropriate to use the approach of Shannon that defines noisiness as the disturbance of a signal in a transmitting medium. The player can define what the signal is. Since we do not know in advance what the player will define as the signal, a method of tracking it would have to adapt to the definition of the player. Developing such an adaptive method would have been a task beyond the scope of this STSM. Therefore this task was skipped.

Instead, available methods of noise tracking were collected and studied. For the ongoing work of this STSM, measurements available in MSP (zero crossing, spectral flatness measure) were used.

In order to map extracted parameters to STK instruments, MSP PerColate objects (which are STK instruments ported to MSP) were first studied and selected. In a second step, a modified version of STK (by Gary Scavone) was used to test STK instruments with continuous control parameters (the actual version allows to drive the instruments with MIDI NoteOn's).

Different methods how to map the extracted parameters to the different instruments were implemented and tested. It took more time than expected to simply drive the instruments with pitch and amplitude due to the inter-dependency of its input parameters (using physical modeling synthesis methods). Eventually, however, a more intuitive mapping was obtained by scaling the noisiness measurement and mapping the result to the noise gain. This increased the adequacy of the transient phase of some tones using for example the STK instrument 'Clarinet'.

Concerning the relationship between physical playing gestures and extracted signal parameters, we first had to define the physical gestures one might want to focus on. Since it was the idea of this STSM to look at the playing process from the perspective of the player, the basic physical gestures were taken from widely accepted literature of violin pedagogy. Ivan Galamian defines the basic gestures in his book 'Principles of violin playing and teaching' (Prentice-Hall, 1962). Basic gestures include things like: "Horizontal movements of left hand fingers within one position" or "Vertical motion of the right upper arm" etc. Due to the fact that similar gestures (e.g. changes of pitch can be done either by another finger on the string or by sliding with the same finger to a new position) it was not possible to put these gestures in a one-to-one relation with extracted audio parameters. However, on a higher level of "gestures" which one might call 'physical playing parameters' it may be possible to say that bow speed and bow pressure affect amplitude and bow position and bow pressure affect brightness of sound. Since these signal

parameters are affected by different physical playing parameters as well, a one-to-one relation between measurement and physical playing parameter could not be stated. The problem was discussed with Prof. Dr. Caroline Traube (Université de Montréal) who has worked a lot in gesture and timbre relation. She proposed to record sounds with different gestures, to analyse differences in visualisations of the sound files, and to create specific methods of measurement according to the analysed differences. This would be a task for future work.

The second part of the STSM was focusing the approach to drive synthesis methods mainly by the raw and non-analysed audio signal. It included discussions about the different approaches (parameter based and signal based driven synthesis), the discussion and selection of known synthesis methods in order to modify them to be driven by the signal, the implementation of these methods and the evaluation.

The following methods were taken into account during the discussions:

- Multiple Carrier-, Multiple Modulator-, Feedback FM Synthesis,
- Granular Synthesis
- Physical Modeling
- Waveshaping
- Waveterrain Synthesis

The methods selected were:

- Multiple Modulator FM Synthesis
- Granular Synthesis
- Waveshaping
- Scanned Synthesis
- Concatenative Synthesis

Physical Modeling was estimated to be of less use since it would have to be reduced to a delay with feedback to be driven by the audio signal. In this case it would be similar to a comb filter effect. The best results came from the use of Multiple Modulator FM. An already started implementation was finished that allowed one sine oscillator to be modulated by two signals. It was possible to use the raw audio signal as a modulator as well as the filtered audio signal. Granular Synthesis with a realtime input and Waveshaping Synthesis with a wavetable that is modified by extracted parameters showed interesting results as well. However, given time constraints, these approaches could only be implemented in experimental states and Concatenative Synthesis was not implemented due to missing objects in the MaxMSP environment. It will be a part of the future work to implement and modify this method and to refine the already existing experimental approaches.

A presentation of my research work was given on the 28th of August for the students and professors of the Music Technology Area at the McGill University.

Main results obtained:

Measurement of musical playing parameters may not necessarily correlate with the physical measurement of these playing parameters since a musician might have another definition of the parameter than the one used in the measurement method.

Measuring pitch, amplitude, spectral flatness, and spectral centroid may not lead to a one-to-one correlation with the basic physical gestures of a violin player.

Trying to modify known synthesis methods according to the approach of Audio Signal Driven Sound Synthesis, modified versions of Waveshaping, Multiple Modulator FM Synthesis and Granular Synthesis showed results that were estimated to be of use.

Future collaboration:

I was asked to return to the Music Technology Area in order to go on with this research. I would like to do so, however this depends on the possibilities I will find to support a future collaboration.