A Modified Sampling Synthesis for a Realistic Simulation of Wind Instruments—The Design and Implementation

M. Pluta[1], L. J. Spalek[2], R. J. Delekta[1]

[1]Academy of Music in Krakow, ul. Sw. Tomasza 43, 31-027 Krakow, Poland
[2]Institute of Physics, Cracow University of Technology, ul.Podchorazych 1, 30-084, Krakow, Poland
Email: psi0expikx@gmail.com

Abstract

Sampling synthesis is one of the most practical and widely used approaches among the various sound synthesis methods used for creating a realistic simulation of acoustic instruments. Using numerous high quality sound samples it is possible to reproduce a sound of almost any musical instrument, including subtle variations caused by the registers of an instrument or through the use of different articulation techniques and dynamic levels. However, this method has some disadvantages. Firstly, with high fidelity reproduction systems, the repeatability of samples becomes quickly apparent for more experienced listeners. This is sometimes manually corrected by switching between several different samples of the same note. Secondly, it is standard approach to record and reproduce each note separately. It prevents samplers from reproducing natural note transitions, making fluent, connected articulations, such as legato, unnatural. Finally, samplers provide a very limited number of sound parameters to control. Therefore, it is difficult to introduce a set of purposeful fluctuations of selected parameters uniquely attributable to human performances. A synthesis system which addresses the aforementioned problems has been developed at the Academy of Music in Krakow (Poland) for a group of wind instruments as part of a symphony orchestra. The system is based on a large collection of non-standard samples. Samples contain short sequences of notes instead of single notes. In order to use them, a number of techniques have been implemented to allow the seamless connecting of recorded sequences and the control of note durations as well as tempo and dynamics envelopes. Decision-making algorithms and signal processing are applied to create melodic figures by choosing, modifying, and connecting fragments of samples, while keeping natural note transitions intact. The problem of mimicking human performances is addressed by implementing a set of performance rules. It allows to introduce context-dependent variations into the regular playback of the material contained in the musical notation to properly shape the expression of musical structures, similarly to that of live performances by musicians. This article presents the main modules of the modified sampling synthesis system designed by us as well as its general structure and principle of operation. The modules are responsible for performing musical score analyses, an automatic selection and connection of sound samples, and the application of performance rules.

Keywords

Sound Synthesis, Synthesis Methods, Sampling

1. Introduction

Sampling synthesis is a common method of reproducing instrument sounds due to its high fidelity despite a low computational cost. A sufficiently large data bank with a variety of samples with different articulations, pitches, and dynamics for a given instrument is the only requirement. However, the result is limited playback control without the use of other synthesis methods, and even so, this has to be done manually. In addition, the nature of traditional synthesizers with their single note samples makes phrases containing smooth note transitions such as those in legato seem unnatural [1] [2].

In order to remedy the problems highlighted above we propose a modified sampling synthesis system. A combination of special multi-note samples and computational processing methods allow the emulation of naturally sounding transitions. A variety of interchangeable samples solves the undesirable feeling of repeatability and the application of performance rules simulates the performances of live musicians.

A detailed account of the concept of the system was originally presented at the 2014 Forum Acusticum [3]. Implementation details with regards to score analysis and sample joining were presented at the OSA2015 conference [4] whereas the performances rules following those in [5] [6] were discussed at ISSET2015 [7]. This article presents the whole synthesis system, reflecting the feasibility of the ideas presented previously with the following self-imposed limitations:

- The system is exclusively limited to wind instruments as part of a symphony orchestra. An extension of the same principles to string instruments will be considered in the future.
- Structures only present in orchestral parts are being simulated. This approach avoids the more complex structure variability associated with solo or chamber music performances leading to larger collections of samples and more complex performance rules. Most arrangements that use samplers are created for orchestral pieces anyway. Extending the system to handle chamber music will be considered in the future.
- Performance rules for early or contemporary music are not included. Firstly, those kinds of music are less frequently arranged using samplers. Secondly, rules for contemporary music are hard to define as the example base of such performances is rather small.

Chapter 2 deals with the general structure of the modified synthesis system and the principle of its operation. The remaining chapters detail the main modules responsible for performing the musical score analyses, the automatic selection and connection of sound samples, and the application of performance rules.

2. An Overview of the Synthesis System and the Sound Samples

2.1. An Overview of the System

The system is composed of the following modules (cf. **Figure 1** for a schematic overview):

- The score analysis module responsible for carrying out a preliminary analysis. The music score is read in from a file first. The melody is then divided into fragments (musical phrases). A performance rule pattern search is performed last (e.g. phrase arches, up/down progressions, climaxes etc.).
- The figure matching module is responsible for choosing samples that will constitute components of music phrases, determining the cutting and joining positions. Specifically, which notes from the sample form the phrase and which ones notes are chosen as sample joining spots).
- The waveform generator module creates a wav file. The actual synthesis takes place at this stage through sample cutting, joining and sample processing.
- The management module which plays the role of a module coordinator. It controls the synthesis processes, *i.e.* it launches the respective modules in the appropriate order, oversees the flow of data, and coordinates and synchronizes the parameters of the voices in the case of polyphonic synthesis.

In terms of implementation, the synthesizer is as a series of Octave/Matlab scripts. This approach results in concise and readable code allowing easy modifications in the future, its reuse or integration as part of other
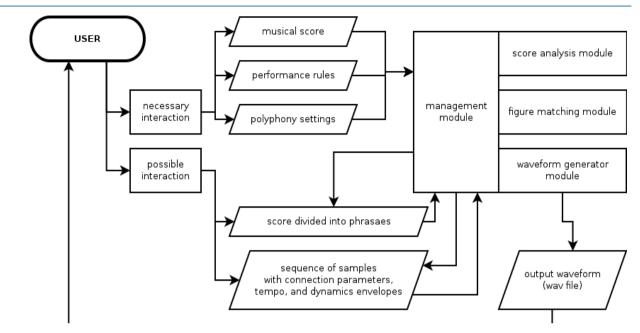
Figure 1. An overview of the synthesis system as a whole.

solutions as well as its rewriting to other programming languages. The simultaneous drawbacks include a relatively low operating speed and the possibly of reduced ease of use. However, these are of little significance during the development stages. The note analysis, motif fitting and the signal generating modules require the further implementation of certain functionalities albeit mostly work. The management module is currently being developed with a few elements already tested to be working. Looking into the future we envisage an implementation of the synthesizer in Java to support multiplatform interoperability.

2.2. Sound Samples

Two types of sounds samples are present. Single note samples contain different articulations, whereas multi-note samples contain intervals (two different pitches) from the minor second to the perfect octave as well as tetrachords (fragments of scales). Tetrachords are of importance as in the melodies of the orchestral parts of wind instruments (characterized by smooth melodics and connected sounds) parts of scales split by intervals predominate [4].

3. Modules

3.1. The Score Analysis Module

The analysis module (cf. **Figure 2** for its schematic representation) is principally responsible for the automatic division of the whole instrument part into fragments (musical phrases) which are performed as legato, *i.e.* with closely connected sounds. The module searches an appropriate spot, a phrase interruption or ending, e.g. arch endings or phrase arches like slurs or phrasing slurs, rests, repetitions or large melody jumps. Fragments which lie outside of phrases are earmarked for substitution with traditional, single note samples.

A file in the Lilypond format serves as input. The preparatory editing may be carried out in a text file or with the use of WYSIWYG programs such as Frescobaldi (cf. http://frescobaldi.org). Due to the use of formatting akin to that of the LaTeX typesetting system, the input file allows the inclusion of extra markers to be interpreted by the waveform generator. These relate to phrase arches and other performance indicators such as climaxes or changes in the tempo or dynamics etc.. The phrase search algorithm has been described in detail elsewhere [4]. The output data is in the form of a text file containing melody fragments (separate phrases and single sounds). It bears resemblance to the input file, however, with differences in the structure reflecting the divisions into phrases.
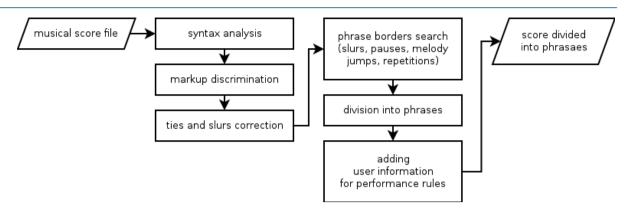
Figure 2. A schematic of the execution flow in the score analysis module.

3.2. Figure Matching Module

The figure matching module is principally responsible for (cf. **Figure 3** for its schematic representation) making appropriate choices with regards to the multi-note samples that will make up each individual phrase as well as the determination of the cutting and joining positions. With regards to the latter, notes that form phrases as well as those on which sample joining takes place need to be determined.

The input is made up of the output of the score analysis module described in Section 3.1. A text file containing a list of the sound samples used, their sequence (the indices of the respective samples, a number associated with the starting and ending note used within a sample, and whether it is connected with the next sample) as well as rhythmic, agogic (the main tempo and the tempo envelope), and dynamic (the amplitude envelope) data serves as the output.

3.3. Waveform Generator Module

The waveform generator module (cf. **Figure 4** for its schematic representation) creates a wav file through the cutting, joining and the processing of samples.

The output data from the figure matching module serve as input data here. The output data consist of a wav file with the instrument part or several files with parallel parts of instruments in the case of polyphonic synthesis.

The implementation tasks currently under way include the automatization of:
- Sample read in and amplitude normalization.
- Tempo grid determination, *i.e.* the determination of the position of rhythmical values of the musical sequence in the output wav file based on the main tempo, the tempo envelope and selected performance rules.
- The cutting out of sample fragments and the consecutive pasting in the appropriate time position of the output waveform, in line with the saved sequence and tempo grid.
- An implementation of crossfade (using the algorithm described in [4]) on the shared notes of joined pairs.
- The matching of time duration of notes registered in the samples to the target time duration in the generated waveform. Noteduration adjustment (extension or contraction) through the cutting or pasting of sustain phase fragments.
- The implementation of the amplitude envelope while taking performance rules into consideration.

3.4. Management Module

The management module (cf. **Figure 5** for its schematic representation) is responsible for coordinating all of the other modules. It controls the synthesis process by launching the appropriate modules in the appropriate sequence and manages the data flow. In the case of polyphonic synthesis it coordinates and synchronizes the parameters of the voices. The management module introduces modifications to the sequence of samples, in particular with regards to the settings of the amplitude envelope and dynamics in order to synchronize them among the instruments, if the user chose the generation of parallel audio tracks. Furthermore, in the case of polyphonic synthesis it is possible to set the phase transitioning from one instrument to another, which results in an appropriate, automatic adjustment of the parameters of the generated output signal.
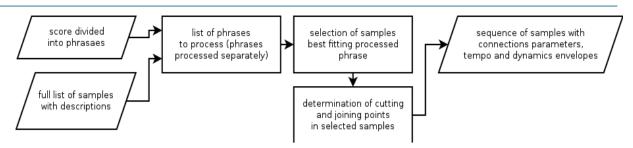
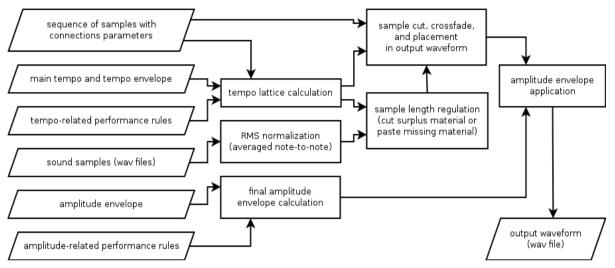**Figure 3.** A schematic of the execution flow in the figure matching module.

Figure 4. A schematic of the execution flow in the waveform generator module.
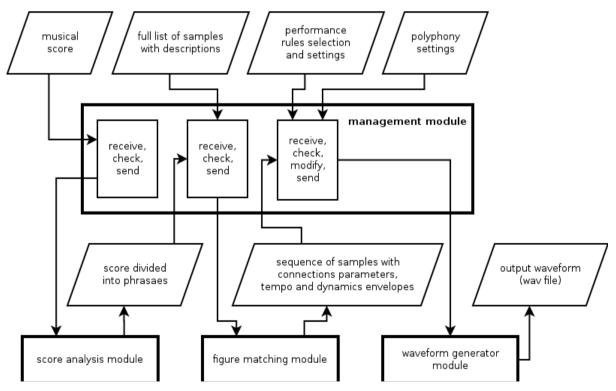

Figure 5. A schematic of the execution flow in the management module.

4. Summary

This article presented a general overview of the idea behind the synthesis system as well as a discussion of the main modules and their principles of operation. The modules are responsible for performing musical score analyses, an automatic selection and connection of sound samples, and the application of performance rules. Most of the system is implemented and ready tested except for the management module. Work on the management module is currently under way. The system in its current version serves as proof of concept. Our ultimate goal would be the implementation of a standalone utility based on the system presented here as realistic synthesis is not only used for artistic purposes but scientific purposes as well.

Acknowledgements

References

[1]   Strawn, J.M. (1985) Modeling Musical Transitions. CCRMA, Department of Music, Stanford University, Stanford, California.

[2]   Almeida, A., Chow, R., Smith, J. and Wolfe, J. (2009) The Kinetics and Acoustics of Fingering and Note Transitions on the Flute. *J. Acoust. Soc. Am.*, **126**, 3. http://dx.doi.org/10.1121/1.3179674

[3]   Delekta, R. and Pluta, M. (2014) Synthesis System for Wind Instruments Parts of the Symphony Orchestra. 7*th Forum Acusticum*, Kraków, 7-12 September 2014.

[4]   Pluta, M. and Delekta, R.J. (2015) Technique to Seamlessly Connect Sound Samples in Sampling Synthesis. *Conference proceedings OSA*2015, Wrocław-Świeradów.

[5]   Bresin, R., Friberg, A. and Sundberg, J. (2002) Director Musices: The KTH Performance Rules System. *Proceedings of the SIGMUS-46 Conference*.

[6]   Friberg, A., Bresin, R. and Sundberg, J. (2006) Overview of the KTH Rule System for Music Performance. *Advances in Experimental Psychology*, special issue on Music Performance.

[7]   Delekta, R.J. and Pluta, M. (2015) Implementacja Reguł Wykonawczych w Syntezie Dźwięku Instrumentów Dętych Zmodyfikowaną Metodą Samplingową. *Proceedings of the* 16*th International Symposium on Sound Engineering and Tonmeistering*, Warszawa, 8-10 October 2015.