Visual manipulation environment for sound synthesis, modification, and performance
Naotoshi Osaka  Takaumi Hikichi

NTT Communication Science Laboratories
{osaka,hikichi}@brl.ntt.co.jp
http://www.brl.ntt.co.jp/people/{osaka,hikichi}

This paper describes a visual sound manipulation system which includes sound synthesis, modification and performance functions. It can run on windows and be used basically as an off-line wave editor. Its distinct function is sinusoidal-model-based sound modification, and it also includes timbre morphing capabilities. A sound is manipulated using other sound objects and operation objects, and the procedure history for a sound object is visually displayed. The system is now being used and tested by some composers and music school students. It is becoming an extension of conventional acoustic musical instruments, and therefore will become an important tool for computer music creation and performance in the future.

1 Introduction

Sound manipulation including sound synthesis, modification, and performance is one of the most important aspects of computer music creation. One trend in sound synthesis is real-time systems, as MAX/MSP[URL1] and Kyma[URL2]. However, off-line sound synthesis is still important, and several wave editors are running in ordinary music workstations.

In some algorithms, a signal cannot be processed in real time, and therefore the system cannot run in real time. The system proposed here is basically an off-line sinusoidal model-based sound modifier which functions as a general wave editor. The system is named Otkinshi (which means system for both sound and speech in Japanese).

Created sounds are either stored on disk and used in a performance part of the system. They can also be displayed on a real track and in real time for an automatic performance. And they can be treated as macro sound objects.

The original version, which was developed in 1991[1], was written in Objective-C and runs on Next Cube. The newer system is written in C++ and runs on Windows 95, 98 and NT without any special hardware. It is intended for use by musicians and those who are interested in sound manipulation. In this paper we focus primarily on the system configuration.

2 System Feature

The system's fundamental function is sinusoidal model based sound modification, and it also includes morphing capability. Physical model based morphing is also incorporated into the system. These features ensure rich timbre synthesis.

We aim to develop a sophisticated GUI (Graphical User Interface), by taking into account the following functions:

1. Less mouse operation,
2. unified concept of "object" for both sound and operation,
3. visual procedure environment.

In function 3, a complete visual programming environment is not possible, and only sequential procedures such as successive filtering processes can be expressed.

2.1 Sound Synthesis

One of the main functions of the system is the sinusoidal model based modification of a sound. M&Q algorithm [2] is used for the analysis/synthesis. The sound can be modified by saving or deleting the partials which satisfy certain conditions, such as threshold or range of either amplitude or instantaneous frequency, length, and the partial's location in the frequency band. In the near future, vibrato extraction and addition/subtraction will also be implemented.

2.2 Timbre Morphing

Sound morphing is one of the most sophisticated technique using a sinusoidal model representation. The algorithm used here has been described in detail in our previous work[3].

Morphing is done by interpolating the model's parameters. The algorithm is based on automatic processing. The central problem that the algorithm solves is matching the members of two unequal groups in the number of members.
A physical model based morphing algorithm [4], as well as the signal model, is also being implemented. Timbre morphing is done by physical parameter interpolation. Our present study covers the sounds of struck strings, plucked strings, and elastic media.

3 System Configuration

Figure 1 depicts the total system configuration. The system is divided into two sections: sound synthesis and sound performance.

3.1 GUI (Graphical User Interface)

Only a few mouse operations are required for sound synthesis and manipulation. The system unifies operation objects and sound objects in a GUI. These objects are multi and recursive layer objects. At the top layer, simply a button with an icon is displayed (operation/sound icons). When we click an operation icon an operation is executed and when we click a sound icon a sound is made. By double clicking the icon we can reach the second layer and set the operation parameters for an operation parameters.

Figure 2 depicts a modification process of a second object. In the upper window there is a second layer display of a sound object. The lower left window is the sound modification panel. The lower right window is the pitch conversion dialog box. The upper window contains a tool box with monitoring commands and a display of procedure history. In the monitor wave and spectrum display, sound playback, recording, and editing are possible, as in an ordinary wave editor.

In the display of procedure history, operation history is shown in terms of both sound and operation icons. This provides a genealogy of the sound, and a programming notations. In this example, sequential procedures of a sound file reading, partial cutting, reversing, rate conversion, etc. are shown.

This display of procedure history is editable like a patch in Max. It allows users to redo the procedure and stop at any point as long as the stoppers are appropriately positioned.

This display of procedure history suggests that by adding some other control functions such a branch, jump, or loop, we can obtain a programming environment. The difference between MAX patch and this history description is that the former is a procedure definition, while the latter is a (sound) data definition using sound icons and operation icons.

However, currently only a sequential flow is possible and this severely restricts the programming capability.

3.2 Spectrum Display

The system is equiped with a variety of spectrum displays in order to view them from different aspects. Figure 3 depicts the three styles; original wave (upper left), 3-D display of all the frames
(lower left window), spectrogram (upper right window), and partial trajectories obtained from sinusoidal analysis (lower right window). A 2-D display of a frame is also included, although it is not shown in the figure.

3.3 Sound Synthesis and Modification Functions

The synthesis functions are: basic function generators such as sine, saw tooth, triangular functions, and white noise. Various modulations, such as frequency modulation, amplitude modulation, and ring modulation are also incorporated. Other functions include various filtering, distortions and reverberation, stereo sound manipulation, etc. However, the most powerful function is the sinusoidal model based sound modification.

3.4 Sound Performance

This is an individual unit designed for live concert use. In manual performance mode (Instrument mode), created sounds are simply made by clicking icons on the computer display. Each of these buttons is assigned to a sound generated by sound synthesis part or sound files.

A sound is made instantly when the icon is clicked. This is achieved by storing the first portion of all of the sound files into the main memory. Periodic scanings are also necessary in order to activate the sounds. This provides the convenience since there is no need to transfer sound files to other samplers or equipment.

MIDI control is also possible now, and sounds can be assigned to a musical keyboard. Other functions include the automatic sequential playback of sound files and the loop playback of each sound.

4 Monitoring of the system and Performance History

The system has been installed at Kunitachi Music College. It is used for both music and sound education. In January 1999, NTT ICC (Inter Communication Center) held a workshop named New School #6, in which the system was open and tested by twenty users.

This system has also been used at two live computer music pieces. One is for piano and computer, and the other is for violin and computer. Both were quite successful and did not experience any system trouble. Therefore, this system seems to be promising tool for a live performances.
5 Conclusion

We introduced the sound modification system Otkinh-shi (version II). Sounds can be represented by a sinusoidal model and modified. One of its sophisticated functions is morphing. It can also generally functions as a wave editor. Version II was developed for musicians and students who are interested in sounds. Currently we are distributing software and receiving feedback from monitors.

Our future work includes revising the user interface, and the physical model, as well as adding more modification functions to the system such as vibrato extraction and control, pitch conversion, or sound stretch/compress.

6 Acknowledgments

The authors would like to express their gratitude to Dr. Norihiro Hagita for his enthusiastic support.

References


[URL1] http://www.cycling74.com/