University of Huddersfield Repository

Gibson, Ian and Howard, David

Intuitive and dynamic control of synthesised sounds by voice

Original Citation

Intuitive and Dynamic Control of Synthesized Sounds by Voice

I.S.Gibson and D.M.Howard

Department of Electronics, University of York, Heslington, York YO1 5DD, United Kingdom.
e-mail: isg100@unix.york.ac.uk

ABSTRACT: This paper describes an interactive computer software environment for creative mapping of parameters estimated from the performers voice alone onto sound synthesis parameters. The composer is able to construct instruments which provide intuitive simultaneous vocal control of multiple parameters for a range of digital sound synthesis techniques. The system is written in C and runs on a Silicon Graphics Indigo (SGI) Iris 4D workstation in real-time.

Introduction

Traditional instruments offer subtle and complex methods of gestural capture [Bailey et al]. Digital sound synthesis techniques provide the composer with a great variety of timbres but often an intuitive control interface is lacking. The voice is a rich source of time-varying parameters which may be used to control synthesized sound parameters.

Voice Analysis

The system samples voice input at 16kHz. Once sufficient data has been received, the system performs several voice analysis routines. The resulting data is made available to user-defined *instruments*. Some of the voice analysis techniques available are:

- fundamental frequency estimation using the peak-picking method [Howard,D.M.],
- linear predictive coding using an autocorrelation algorithm [Witten,I.H.] and
- Fast Fourier Transform (FFT).

Instruments

Instruments are defined in a modular format using a graphical interface. Modules fall into the following categories:

- system input (e.g. voice, graphical and MIDI),
- data processing (e.g. sequencing, iteration and conditional execution) and
- output (MIDI).

A graphical *workspace* is provided to enable design, testing and playing of instruments. Icons are placed within the workspace and linked together to form a graphical representation of the instrument.

The Performance Environment

The *performance environment* allows a piece to be created using multiple instruments. The performer can sequence a number of musical events on different virtual tracks (called *parts*) resulting in possible concurrent transmission of data on all MIDI channels. Control of a performance is achieved using a combination of voice input, MIDI input and on-screen graphical interaction. At any one time the performer can control one or more instruments using these techniques.

System Output

System output is configurable by the user allowing customization for any MIDI sound module. Controller and System Exclusive messages are used to control sound synthesis parameters in real-time. The system has been developed in conjunction with a synthesizer which allows all sound program parameters (e.g. those used for wavetable selection, amplitude, low-pass filtering, stereo panning, low frequency oscillation and enveloping) to be modified in this way.

Further Work

System response times could be improved. The existing system uses the same sample input frame size for all analysis routines. However, some routines only require part of that data. Therefore, implementing different software buffer sizes for each algorithm would ensure that output variables from these routines were updated as quickly as possible. Another possible approach would be to implement the system on a multi-processor array. This would enable voice analysis routines to run simultaneously.

Higher level musical interaction could be explored within the performance environment. For example, a piece could be composed which reacts to changes in voice parameters. As a vocal sound decreased in harmonic to noise ratio some parts might increase in amplitude or extra note events might be triggered.

More complex instrument design could be achieved with the use of *compound icons*. A set of icons on the workspace would be represented as a single icon with a single set of inputs and outputs. This would also allow easy storage and retrieval of user-defined icon libraries.

Control of MIDI devices other than sound modules could be considered. This could include effects processing (e.g. equalization, reverb, delay, flanging and chorusing). Also, MIDI devices which do not output sound (such as those used to control stage lighting) might be triggered from voice parameters during a performance.

References

[Bailey et al]   BAILEY N.J., PURVIS A., BOWLER I.W., MANNING P.D. (1993). Applications of the phase vocoder in the control of real-time electronic musical instruments. *Interface*, 22, 259 - 273.

[Howard,D.M.]   HOWARD, D.M. (1989). Peak-picking fundamental period estimation for hearing prosthesis, *Journal of Acoustical Society of America*, 86, 902-910.

[Witten,I.H.]   WITTEN I.H. (1982). *Principles of Computer Speech*. London, UK: Academic Press Inc. (London) Ltd.