A Phase Vocoder Graphical Interface for Timbral Manipulation of Cellular Automata and Fractal Landscape Mappings

A.I. KATRAMI, R. KIRK, A. MYATT

Music Technology Group
University of York
York YO1 5DD, U.K.

E-mail: ak6, pkr1, am12@york.ac.uk

ABSTRACT: A graphical tool is described based on the properties of the Phase Vocoder (PV) program that allows users to manipulate PV data, prior to resynthesis. The data, displayed in a 3-D format, can be obtained from the output of Cellular Automata, Fractal Landscapes and PV analysis files. The output of any of the previously mentioned methods, can be used as a filtering function, resynthesis data, or combination of both. The proposed graphical tool offers composers the ability to select and manipulate spectral components of individual interest. Intended applications relate to graphical control of spectral morphology and manipulation of electroacoustic sound sources in the frequency or time domain.

INTRODUCTION

The representation of musical timbres through time varying spectra is a valuable technique for musical applications. The use of such data allows composers to analyse and modify the frequency, time scale and spectral content of recorded material. The PV is a technique for converting a sampled signal into a time-varying spectral format. The PV can be used to decompose an audio signal producing a representation in which temporal and spectral features can be manipulated independently. A feature of this representation is that none of the information in the original audio signal is destroyed. As a result, through resynthesis, it is always possible to reconstruct the original signal accurately, when no modification is made to it. In practice, the PV’s ability to retain phase information is what distinguishes it from its predecessor, the ”Channel Vocoder” making it suitable for musical applications.

This paper describes the way the PV is used to process audio data. The data used for resynthesis can be derived either from PV analyses files, or from synthesised images. A graphical interface is described that allows generation control and manipulation of image data. The current processes used to generate graphical data are Cellular Automata (CA) and Fractal Landscape (FL) algorithms.

CELLULAR AUTOMATA

CA is a rule based algorithm that can be used to model the evolution of complex phenomena. As stated by Wolfram [11], it evolves in discrete time steps, with the value of the variable at one site (cell) being affected by the values of variables at sites in its “neighbourhood” on the previous time step. The neighbourhood of a site is typically taken to be the site itself and all immediately adjacent sites. The variables at each site are updated simultaneously, based on the values of the variables in their neighbourhood at the preceding time step. Figure 1 shows a set of local rules for the time evolution of an elementary one-dimensional CA.

Figure 1: The figure shows rule 90 (001→01001). The first line contains all possible combinations with three bits while the second line shows the result when a specific rule was followed. The rule given is the modulo two rule: The value of a site at a particular time step is simply the sum modulo two of the values of its two neighbours at the previous time step.

This project is partially funded by SERC, and the Ontario Foundation of Genes.

ICMC 106
Some of the possible different rules can be seen below.

The current paper investigates the use of Cellular Automata output as frequency filters to the PV analysis data, or its application for direct resynthesis.

**FRACTAL APPLICATIONS**

The fractals and their resulting two dimensional colour pictures or grey scale images, can be effectively used either as a one dimensional normalisation display filter or as a two dimensional input to a PV system. In the former case amplitude normalisation can be achieved by manipulating the current grey scale values of the 2D fractal images. The latter case can be implemented with a new version of the PV program. This involves the use of both digital image processing algorithms and two dimensional Fast Fourier Transform. It was identified through the use of the PV graphical interface that the form of the displayed data could be synthesised by the application of fractals. The synthesised data could subsequently be generated to produce acoustic timbres and also be applied as filtering (Vocoding) masks to previously analysed data.

**THE PHASE VOCODER GRAPHICAL INTERFACE**

The PV is an analysis - synthesis system that has an intermediate data, the time-variant Discrete Fourier spectrum, of the input signal. The intermediate data can be transformed, with no additional loss of information, into a more conventional magnitude and frequency representation. These intermediate data can then be used to resynthesise timbres at different pitches or different rates from the original. The analysis of sound is achieved by windowing the data. Within each time-window, the PV divides the spectrum into a number of equally spaced bands known as 'channels'. It then looks in each channel to see if any sound component is present and records its amplitude and frequency. The greater the number of channels, the more detailed the analysis becomes.
In an early form, the PV graphical interface processed and displayed data from the PV analysis file. Standard PV analyses files were used as input data and the underlying program provided the graphical environment for their efficient manipulation. Figures 2 shows the typical output of this PV user interface.

Figure 3 presents a schematic representation of the current interface design. The procedures shown can be classified to describe three main categories: a) Procedures for sorting the data context of each window; b) Data displaying procedures; and c) Procedures for manipulating data. Sorting procedures were used to enable meaningful graphs of amplitude and frequency. Displaying procedures were necessary so as to visualise a maximum of 16 windows selected by the user. The current rate of 16 displayed windows is shown to be adequate in visualising spectral information. Finally, the procedures for manipulating data involve pitch stretching, time stretching and interpolation.

The new graphical interface uses additional sets of data either from CA or from FL mappings. The main processing algorithm remains the same in both cases, with the exception that in CA the images are currently stored in binary files, while in the FL mappings images are of grey scale format with pixel values ranging from 0 (black) to 255 (white). A fixed length of 512 pixels per image line is used in both cases. If the number of pixels in the image lines is not a power of two, then the line is padded out with the required number of 0's as appropriate. Finally the image is either reduced or enlarged to obtain the required 512 pixels per line using a three pixel neighbourhood interpolation.

The current graphical environment uses FL and CA images to produce time-varying filtered masks that can be applied to any audio data. In the general case of FL mappings (grey scale images) the algorithm uses the corresponding grey scale value in each pixel as a normalisation factor for each channel's amplitude value. The amplitude of the input data spectral components is scaled by the intensity of the corresponding "grey scale" filter mask components. This yields the final amplitudes for the resynthesis file.

For example, consider an amplitude value \( m \), and a grey scale intensity \( n \) in the corresponding FL image. For convenience assume that \( n = 128 \). Since the grey scale pixel values range from 0 to 255 in intensity, \( n \) has half the intensity of the maximum allowable value \( (n = 256/2) \). The resulting channel amplitude value should accordingly be 50% less than it was originally, that is \( m/2 \). A subcase of this general approach is the CA binary images where the value 0 in a image pixel cuts entirely the corresponding channel's amplitude while value 1 leaves the channel's amplitude unchanged [Figure 4]. FL and CA images have only been used as filters in the current PV implementation. Direct resynthesis from FL or CA data is achieved by using a flat spectrum as the input to this process.
FUTURE WORK

In future work, musical sound streams of certain time duration are to be viewed as 2D grey scale images that can be pre-processed using a variety of available image processing transforms (histogram modification, Median Filtering, Unsharpy Mashing, Histogram Equalisation). The resulting pictures can be displayed on the current PV interface by means of an additional 2D FFT algorithm. Various image processing techniques are proposed to be investigated as the basis of a synthesis method.

In addition the Phase Vocoder graphical interface will be installed on a NeXT computer utilising its DSP56000 processor and also on the MIDAS system [2], to aid the exploratory process by producing a real-time implementation.

BIBLIOGRAPHY
