CHAOTIC PREDICTIVE MODELLING OF SOUND

J.P. Mackenzie
School of Electronic and Manufacturing Systems Engineering, University of Westminster,
115 New Cavendish St., London W1M 8JS, U.K. Tel: +44 071 911 5000 ex3642
e-mail: mackenlj1@wmin.ac.uk

Abstract

This paper presents an analysis/synthesis model for sound that is based on nonlinear dynamics, or chaos, theory. The inspiration is that since chaos and fractals can represent many complex naturally occurring forms, can the same be found for sound? Evidence is examined that shows how nonlinear dynamics plays a fundamental role in the generation of sounds, both musical and non-musical. Presented is a novel model that consists of an autonomous nonlinear feedback system and a way of analysing a sound to find parameters for the model. Encouraging results are presented showing the analysis and resynthesis of air noises, wind instrument and gong sounds

1. Introduction

It is now well known that simple nonlinear dynamical systems can produce complex and beautiful behaviour that often replicates natural phenomena. A striking example of this can be seen in the wide range of abstract and natural looking computer images that may be generated with such systems. Could the same be done for sound? For example, could a complex natural sound be represented by a simple nonlinear system defined by only a small number of parameters? If so, then the sound may be stored, regenerated and manipulated via the dynamical system offering a powerful tool for anyone involved with the creative use of sound.

There is now much evidence that nonlinear dynamics plays a central part in the mechanisms of sound generation, both musical and nonmusical, periodic and irregular. It has been shown, for example, that the sound of the wind has a 1/f form to its power spectral density i.e. the audio time series is statistically self-similar, or fractal, which indicate the presence of chaos (Mackenzie, J. 1994a). Nonlinear dynamics are fundamental to the generation of musical sustained tones arising for example in wind and string instruments (Fletcher, N. 1994). It has been found that woodwinds display period-doubling behaviour as the blowing pressure increases; the bifurcating sequence ending in noisy chaos (Gibiat, V. and Lindenman, E.). This can also be found with the artificial excitation using sinusoidal vibrations, of gongs and cymbals (Legge, K.). In all these cases, the systems that generate the sounds are known to contain significant physical nonlinearities.

It is the goal of this work to find ways in which such sounds may be represented with simple chaotic systems. This paper presents a first step towards this by demonstrating an analysis/synthesis model which is capable of capturing essential characteristics of a range of sounds with a nonlinear dynamical system. The next sections give an outline of the theory that is necessary to explain the workings of the model.

2.1 Chaos Theory

Central to nonlinear dynamics, chaos and fractals is the model of a recursive nonlinearity, or nonlinear feed-back. This simple process can be a fertile source of both simple and complex dynamic forms that may be of use for creating sound. This behaviour can be visualised by examining the way in which the state of the dynamical system moves in its state space. The technique of viewing dynamics geometrically is a very powerful one and is widely used. The tool of the phase portrait, for example,
creates suggestive images from sound signals and is very easy to generate - see Figures 2 to 6. The long term behaviour of a system is reflected in the geometry of the object to which the state is attracted in state space. Attractors having a closed loop shape correspond, for example, to periodic dynamic behaviour. Chaotic behaviour, which is complex and irregular, corresponds, typically, to an attractor having a fractal structure, i.e. having self-similar detail on a range of scales and a non-integer spatial dimension.

The model presented here works by representing a sound with a state space attractor and then recreating this with another nonlinear dynamical system. The way in which this is done relies on the assumption that the sound is produced deterministically and there are therefore fixed laws governing the evolution of a sound waveform. These laws are extracted to recreate the attractor by making a short-term prediction of the time series with a nonlinear function.

2.2. Nonlinear Dynamics, Attractors and Embedding

This section gives the mathematical detail underlying the sound model and focuses on the relationship between a nonlinear dynamical system, its state space attractor, and a time series derived from observing that system. For a good account of these topics see (Farmer, D. 1996) and (Broomhead, D.S.).

A general discrete time dynamical system may be written as

$$\hat{x}_{n+1} = F(\hat{x}_n) \quad x \in \mathbb{X}, \quad n \in \mathbb{Z}^* \quad (1)$$

where $\mathbb{X}$ is a $d$-dimensional state vector in state space $\mathbb{X}$, $n$ is discrete time and $F$ is some nonlinear mapping. Consider that the behaviour of this system is described by an attractor in state space $\mathbb{A}$, and an associated probability measure $\mu$. The attractor represents the long term behaviour of the system and is the subset of state space to which initial conditions are attracted. The measure describes the relative amount of time spent by the state in subsets of the attractor.

Chaos is a class of dynamic behaviour characterised by being complex and unpredictable in the long term, despite the system itself being simple and deterministic. This behaviour corresponds to the presence of a 'strange' attractor and associated probability measure in state space which, typically, have fractal structure.

The state vector, the mapping and attractor of a system are not, however, directly, accessible to an observer, whereas some measurement of the system usually $\hat{x}$. Let this observation process be represented by a function, $g$, of the state vector which returns a scalar value that, for each time step, produces the discrete time sequence,

$$\hat{y}_n = g(\hat{x}_n), \quad n \in \mathbb{Z}^* \quad (2)$$

Consider that the description given so far represents the acquisition of a sound signal from some chaotic system. Complete modelling of this sound would be achieved if we had access to $\hat{x}$ and $\hat{y}$; as these define the dynamics of the system that generates it. As we do not have this access, an alternative inverse problem can be stated: from a portion of the observed time series, find an equivalent chaotic system that can generate another time series that is indistinguishable from the original i.e. one that sounds the same as the original.

A means of constructing an equivalent dynamical system given only the time series is provided by the technique of 'embedding by the method of delays' (Takens, F., Broomhead, D.S. and Taylor, W.). An embedding is a mapping from one dynamical system to another which preserves essential features of the original including the topology of the system's attractor and its associated probabilities.

Given the observed time series derived from the original system,

$$y_n = g(\hat{x}_n) = g(F^{\circ n}(\hat{x}_0)), \quad n \in \mathbb{Z}^* \quad (3)$$

define a mapping of the state vector at time $n$ that comprises a sequence of $M$ observations from time $n$ backwards,
\[ H(y) = (y_{y+1}, y_{y+2}, \ldots, y_{y+n})^T \]  

(4)

Under certain conditions (Taken, F.), including the case \( M \geq 2d + 1 \), this mapping is itself an embedding. The vector formed by the mapping is known as the embedding vector and \( M \) as the embedding dimension. The embedding ensures that the trajectory of \( \vec{X} \) in state space \( X \) maps to a trajectory in the embedded space \( \mathcal{H}(X) \).

Moreover, define another mapping, \( S \), that comprises a left shift on the sequence of observations viewed through an \( M \)-length register. That is,

\[ S(y_{y+1}, y_{y+2}, \ldots, y_{y+n})^T = (y_{y+2}, y_{y+3}, \ldots, y_{y+n}, y_{y+1})^T \]  

(5)

which is equivalent to

\[ S[(g^{F^{-1}}(y_{y+1})), g^{F^{-1}}(y_{y+2}), \ldots, g^{F^{-1}}(y_{y+n})] = (g^{F^{-1}}(y_{y+2}), g^{F^{-1}}(y_{y+3}), \ldots, g^{F^{-1}}(y_{y+n}), g^{F^{-1}}(y_{y+1})) \]  

(6)

It can therefore be seen that,

\[ S[H(y)] = H^{-1}(y) \]  

(7)

In other words, the shift mapping \( S \) acting on the embedded state is equivalent to the mapping \( F \) acting on the true state. This equivalent dynamical system will possess an attractor, \( \vec{A} \) and an associated measure \( v \). Because of the qualities preserved by the embedding, the attractors \( \vec{A} \) and \( \vec{A} \) will be topologically equivalent and so quantities such as their fractal dimensions will be the same. Also, the probabilities associated with subsets of the attractors are preserved, so for some subset \( \vec{B} \) of the embedded attractor,

\[ v(\vec{B}) = \mu(\mathcal{H}^{-1}(\vec{B})) \]  

(8)

As a result, these two dynamical systems \((X, F, \vec{A}, \mu)\) and \((\mathcal{H}(X), S, \vec{A}, v)\) are statistically isomorphic (Taylor, W.). They are therefore indistinguishable to an observer who is viewing the two resulting time series: the sequence \( y_n \) from the original system, and the sequence generated by the most right-hand element of the embedding vector.

2.3. The Inverse Problem

The modeling problem therefore now becomes one of finding an approximation to the mapping \( S \) and is referred to as the inverse problem. Because of the shifting nature of this mapping, and because it must be deterministic and therefore a function of the present state, it is of the form,

\[ S[H(y_n)] = (y_{y+1}, y_{y+2}, \ldots, y_{y+n})^T \]  

(9)

and the problem reduces further to that of approximating \( I \), which is a vector to scalar function defined over the embedded state space.

The way in which this is done is to use the information provided by a sample portion of the observed time series. This is converted into a set of, say \( N \), data pairs comprising an embedded vector at time \( n \) and the value of the time series at time \( n+1 \), i.e. of the form

\[ \{H(y_n), y_{n+1}\} \]  

(10)

which must satisfy

\[ I[H(y_n)] = y_{n+1} \]  

(11)
Finding \( r \) is then a function interpolation problem. Several strategies have been suggested in work concerned with the accurate short-term prediction of time series believed to come from chaotic systems (Broomhead, D.S., Farmer, D. 1987, Casdagli, M. 1989, Taylor, W., Singer, A.C.). These include a single global nonlinear function defined over the whole embedded state space, a piece-wise linear or polynomial function and radial basis functions.

For this work, the emphasis is on having a reynthesis model that is computationally simple as many iterations of it are required to generate digital audio sequences. This concern has to be balanced with the accuracy of the approximation to \( r \) as this will determine the stability and quality of the reynthesis. The computational complexity of the analysis is also an important consideration. For these reasons a piece-wise linear function has been used to approximate \( r \).

3. The Modelling Scheme

![Diagram](image)

Figure 1. Schematic Diagram of the Modelling Scheme. The sound generated by the original physical system can be used to reconstruct the embedded system by the method of delays. The modelling problem then becomes one of finding an approximate synthetic system that generates a sound with similar properties to the original.

The complete model as shown in Figure 1 and comprises two halves: the analysis and reynthesis. The analysis comprises the following procedures:

- The original time series is embedded to form a set of embedded vectors, \( H(X) \), in embedded state space \( \mathcal{H}(X) \).
- The embedded state space is partitioned into a set of hypercuboids (\( M \)-dimensional cuboids). This is done by recursively dividing the \( M \) embedded vectors into two sets of approximately equal number until each set can be divided no further else it contains less than a pre-specified minimum number of vectors. Also formed in the process is a search tree which is used during reynthesis.

For each hypercuboid, so as to minimise the squared error, a linear function to the data pairs \( \{H(x_n), y_n\} \) which is of the form

\[
\|H(x_n) - \beta H(x_n) + b\| = \|y_n\| \quad (13)
\]

where \( \beta \) is an \( M \)-dimensional vector. The result is then a set of partitions and associated linear functions that together form an approximation to \( r \) and hence \( S \). Reynthesis consists of the following:

- Initialise a vector, \( \bar{v}_0 \), to be one of the embedded vectors.
- Use the search tree to compute which of the hypercuboids, \( \bar{v}_0 \) is in, apply the associated linear function \( S(\bar{v}_0) \) and hence calculate

\[
\bar{v}_n = S(\bar{v}_n) \quad (14)
\]

Iterate the above step as many times as is required or is possible.
4. Results

The modelling scheme has been applied to a simulated chaotic signal and a range of sounds known, or believed to come from, nonlinear dynamical systems. These sounds are: the rumble of a ventilation fan, a tuba tone, the sound of the wind, and a gong sound. The analysis was tried for a wide range of the parameters on which it depends, namely, the original time series length, embedding dimension, and the number of state space partition sets.

To first test the scheme it was used with a time series derived from numerical integration of the standard Lorenz chaotic system according to the method and parameter values given in (Hidlack, R.). A successful regeneration of a similar time series was obtained with an embedding dimension, \( m=7 \); number of data pairs, \( N=10,000 \); and a partition of the embedded state space into 128 hypercubes each containing between 50 and 80 embedded vectors. The similarity of the synthetic to the original can be seen by inspection of the time series and phase portraits shown in Figure 2. The phase portraits are constructed from three values taken from the sound time series at regular intervals of greater than one sample. This generates a three-dimensional projection of the embedded state space attractor. Notice that an exact replica of the time series is not generated, but one that shares a similar form as revealed by the phase portrait. Both time series were calculated from the same initial condition, remain the same for a short time and then diverge rapidly. This is an example of sensitive dependence on initial conditions, an important characteristic of chaotic systems.

Figures 3, 4, 5 and 6 show results for the sound time series of those experiments that gave synthetic sounds most like the original. This comparison was made aurally and by inspection of the time series, phase portraits, spectra and amplitude probability density functions. The results of these experiments show that the model can resynthesise with a good degree of perceived accuracy the sounds of the fan rumble and tuba tone. Synthetic versions of the other sounds are not so good, but various properties of their time series are preserved by the model such as the spectrum of the wind noise and the amplitude probability density function of the gong sound.

The analysis phase of the scheme is fairly computationally demanding, but the synthesis is very simple. The resynthesis involves one search of a binary tree and the calculation of a linear function per sample. On a Sun SPARCstation, the analysis took a few minutes and the resynthesis only ~3 seconds per second of generated sound. This suggests that real-time resynthesis with a DSP is a possibility.

The other important factor is the size in data terms of the model. This was found to vary widely from sound to sound. For example, the analysis of air noise used 10,000 16-bit integer samples of the original sound to create \( 40 \times 64 = 2,560 \) resynthesising parameters. For the tuba tone, a model was extracted from 8,000 samples that required only 192 parameters.

5. Conclusions

This work began with the intuition that nonlinear dynamics theory has enormous relevance to the subject of computer sound modelling. Results have been presented which confirm this idea by giving a clear indication that a nonlinear dynamical analysis/synthesis model is possible. What the author finds exciting is the range of signals that can be generated from a simple recursive nonlinearity and how both periodic and irregular sounds may be modelled with this one system.

There is still, however, much work that needs to be done to develop the model. Other forms of nonlinearity should be examined apart from piece-wise linear ones. Greater understanding is needed of when the analysis process works and when it does not. Work is also needed to reduce the size of the model as in some cases the number of parameters is unmanageable.
6. References


Figure 2. Original, left, and resynthesised, right, Lorenz time series, top, and phase portraits, bottom.
Figure 3. Original, left, and resynthesised, right, fan rumbled sound time series, top, and phase portraits, bottom.

Figure 4. Original, left, and resynthesised, right, wind noise time series, top, power spectra, middle, and phase portraits, bottom.
Figure 5. Original, left, and resynthesised, right, tuba tone time series, top, power spectra, middle, and phase portraits, bottom.

Figure 6. Original, left, and resynthesised, right, gong sound time series, top, phase portraits, middle, and amplitude probability density functions, bottom.