Sound Modeling from the Analysis of Real Sounds

SØlvi Ystad
Philippe Guillemain
Richard Kronland-Martinet
CNRS, Laboratoire de Mécanique et d'Acoustique
31, Chemin Joseph Aiguier, 13402 Marseille cedex 20 France
{ystad, guillem, kronland }@alphalma.cnrs-mrs.fr

Abstract

This work addresses sound modeling using a combination of physical and signal models with a particular emphasis on the flute sound. For that purpose, analysis methods adapted to the non-stationary nature of sounds are developed. Further on parameters characterizing the sound from a perceptive and a physical point of view are extracted. The synthesis process is then designed to reproduce a perceptive effect and to simulate the physical behavior of the sound generating system. The correspondence between analysis and synthesis parameters is crucial and can be achieved using both mathematical and perceptive criteria. Real-time control of such models makes it possible to use specially designed interfaces mirroring already existing sound generators like traditional musical instruments.

1 Introduction

The concept of sound modeling is introduced, followed by a brief discussion of analysis techniques taking into account the dynamic and the spectral behavior of sounds.

Physical models simulating the wave propagation in unidimentional bounded media (strings and tubes) are made to simulate transient sounds. A so-called waveguide model consisting in a delay line and a loop filter is used for this purpose. The loop filter takes into account dispersive and dissipative effects due to the medium in which the waves propagate. A description of the construction of the filter taking into account these important phenomena is given.

For sustained sounds the source should be modeled separately. For this purpose, the source and the resonator have been separated by deconvolution. By using an adaptive filtering method such as the LMS algorithm, the source signal is decomposed in two contributions: a deterministic component and a stochastic component. The modeling of the deterministic part, whose behavior generally is non-linear, necessitates the use of global synthesis methods like waveshaping and, in the flute case, perceptive criteria such as the Tristimulus criterion. The stochastic component of a flute source is modeled by taking into account the probability density function and the power spectral density of the process.

An example of real-time control of a flute model is presented. A flute equipped with sensors is used as an interface to control the proposed model. Possibilities of intimate sound manipulations obtained by acting on the parameters of the model are discussed.

2 The concept of sound modeling

Sound modeling consists in constructing synthesis models to make resynthesis and transformations of natural sounds. For that, we have to design analysis methods which give representations of the real sounds. Parameters can further be extracted from these representations to feed synthesis models.

2.1 Analysis

Since sounds generally are non-stationary, and since the evolution of a sound as a function of time is important from a perceptive point of view, time-frequency representations should be used. For that we used linear representations such as the Gabor and the Wavelet transforms which consist in decomposing the signal into elementary functions. These analysis methods give time-frequency representations which can be divided into two parts - one corresponding to the modulus of the representation and the other to its phase [5]. In order to extract parameters from this representation we should use special methods like the spectral line estimation method [3] or a matched analysis method [11]. One can then associate one amplitude and one frequency modulation law to each spectral component of the sound.

2.2 Synthesis

Synthesis models can be divided into two main groups; signal models and physical models. Signal models consist in reproducing a perceptive effect using mathematical description while physical models consist in giving a physical description of the sound generating system. In this paper a combination of a signal model and a physical model has been used to model sustained sounds.

2.2.1 Modeling transient sounds

Physical models aim at simulating the behavior of existing or virtual sound sources. We have chosen to use the so-called waveguide synthesis models [8] which have the advantage of being easy to construct with a behavior close to that of a real sound generator. We here give a brief description of the method applied to the tube case by considering the solution of the wave equation [11].

In order to make a model which takes into account the theoretical phenomena, a propagative model shown in *Figure 1* was constructed. This model is a generalization of the so-called wave-guide model which was first proposed by Karplus and Strong [4].

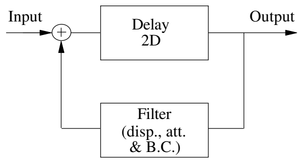

Figure 1. Waveguide model

The delay line in the model corresponds to the time the waves need to propagate " back and forth " in the resonator. The filter takes into account dispersion, attenuation and boundary conditions in the medium.

Although there already exists algorithms for constructing filters taking into account dispersion and dissipation effects [10], these algorithms are based on approximations and are not precise enough for resynthesis purposes where we want to reconstruct a sound which resembles as much as possible to the original sound. This is why we decided to design a new method for constructing a filter which consists in comparing the waveguide model with the theoretical response. The method which was used for that purpose is based on the minimization of the energy difference in a neighborhood of the resonant peaks between the response of the model and the response of the real system [11]. The corresponding impulse response is then constructed in a way similar to the inversion of a time-frequency representation. Actually it consists in summing up elementary functions adjusted and positioned in the time-frequency plane, along a curve representing the group delay. This method makes it possible to construct filters which have the exact values of dissipation and dispersion at the resonance peaks. This gives good results for resynthesis.

2.2.2 Modeling sustained sounds

The model for transient sounds is not sufficient when we want to simulate sustained sounds like wind instruments sounds. We here give an example of how to model the source of a flute. Although several attempts have been made to propose physical models of such a source, this is still an open problem. Since our aim was to simulate its perceptive effects, we decided to use a signal model for this purpose.

In order to extract the source from the the rest of the signal we assumed that the source and the resonator could be separated. Although this is not correct from a physical point of view, we shall see that it is a good assumption in this case since the aim is to reconstruct a perceptive effect. The source signal was extracted by deconvolution and then divided into a deterministic and a stochastic part which were modeled separately.

2.2.2.1 Source identification

As seen in the previous section, a physical model corresponding to the resonator of the instrument can be constructed by a filter and a delay line. The transfer function of the resonant system is of the type:

$$H(\omega) = \frac{1}{1 - F(\omega)e^{-i\omega d}}$$

and corresponds to an all-pole filter. This means that its inverse exists, and that the deconvolution between the real sound and the resonator therefore is a " legal " mathematical operation. The source signal x(t) is given by $x(t) = (y * h^{-1})(t)$
Figure 2 shows the spectrum of a flute sound obtained this way.

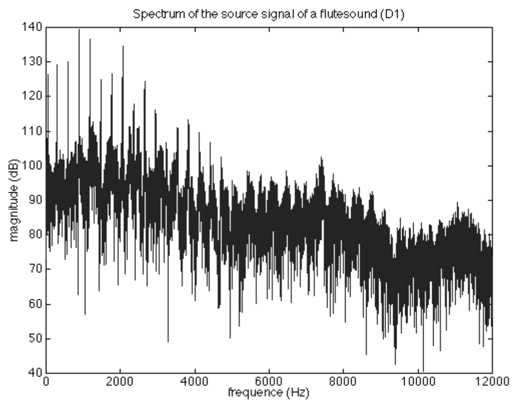

Figure 2. Spectrum of the deconvoluted signal

As we can see the source mainly consists of two contributions, a stochastic part corresponding to the noise in the signal and a deterministic part which is a sum of spectral components. To model the source signal we therefore decided to split it into a deterministic and a stochastic part and to model them independently.

In order to split the two contributions, we used the so called LMS (Least Mean Square) algorithm which consists in using an adaptive filter the aim of which is to remove from an input signal all the components which are correlated to a reference signal [9].

2.2.2.2 Modeling the deterministic part of the source

The evolution of the source spectra for different dynamic levels show that the spectra do not evolve in a linear way, since the increase in the dynamic level do not correspond to a global amplification of the spectrum. This non-linear behavior of the source is common for a lot of musical instrument, and it is of great importance from a perceptive point of view, since it means that the timbre of the flute sound changes when the dynamic level of the sound changes. To model this non-linear behavior we used a waveshaping synthesis method which has been developed by Arfib and Le Brun [1][6], and which consists in constructing a signal by a non-linear function *g* the argument of which is a monochromatic signal with amplitude *I(t)* called the index of distortion. This index has an important influence on the spectrum of the signal. The waveshaping synthesis method makes it possible to generate a wanted spectrum for a given index.

Now the great challenge is to find out how the waveshaping index should vary in order to get an evolution of the synthetic spectrum which corresponds to the evolution of the real spectrum for different dynamic levels. Since the spectral evolution cannot mathematically be modelled by the waveshaping synthesis, we used perceptive criteria to find the variation range for the index. The most well known perceptive criterion is probably the spectral centroid criterion which has been proposed by Beauchamp [2] and which is directly related to the brightness of the sound. This method is satisfying when all the spectral components change with respect to the dynamic level. In the flute case mainly the first four spectral components evolve with the dynamic level, which means that the spectral centroid changes very little although there are great changes between these components. This is the reason why the spectral centroid criterion does not work on flute sounds.

In order to find a criterion which was more adapted to the flute case, we used the tristimulus criterion which has been proposed by Pollard and Jansson [7]. This criterion consists in cutting the total loudness of the sound into three contributions; one (N1) which takes into account the fundamental component, another(N2) which takes into account the second to fourth components and a third(N3) which takes into account the rest of the components.

The tristimulus is then given by the three normalized contributions so that the sum of them is one;

This means that the tristimulus can be plotted in a two dimensional diagram where the x axis (corresponding to the normalized high frequency contribution) is the abscise and the y axis (corresponding to the mid-frequency partials) is the ordinate and where the fundamental contribution is implicit.

The tristimulus corresponding to the flute case and to a matched waveshaping sound are shown in *Figure 3*.

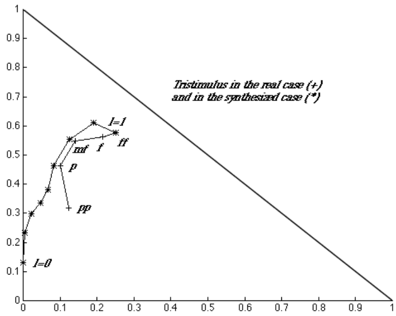

Figure 3. Tristimulus diagram with synthesized (*) and real flute sounds with different dynamic levels.

By fitting the two curves, we then find a correspondence between the index and the driving pressure. The index varies linearly with the logarithm of the driving pressure.

2.2.2.3 Modeling the stochastic part of the source

To model the stochastic part of the source we assume that the process is stationary and ergodic. This means that we can characterize the noise by its probability density function and its power spectral density. In the flute case the probability density function follows an exponential law [11]:

while the power spectral density corresponds to a low-pass filtered noise. The combination of the deterministic and the stochastic part of the source gives the complete source model.

3 A hybrid model

The signal model simulating the source combined with the physical model simulating the resonator of the instrument gives a general hybrid model which can be applied to several instruments. In this section we shall see how the hybrid model can be applied to the flute case. In order to give the possibility to play with this model, a convenient interface should be found to pilot the real-time model.

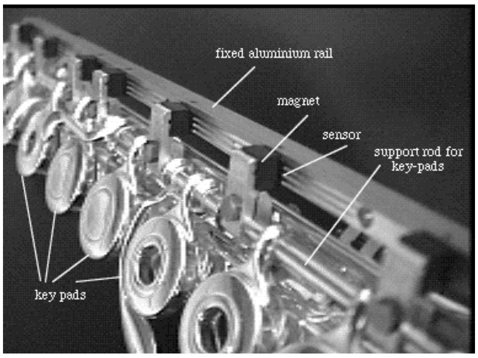

Figure 4. Close view of a flute with magnets and sensors

By choosing a traditional instrument connected to a computer by magnetic sensors detecting the finger position and a microphone at the embouchure level detecting the pressure variations, the musicians can make use of the playing techniques already aquired. The interface is illustrated in *Figure 4*.

As already mentioned, sound modeling does not only consist in resynthesizing sounds, but also in doing intimate transformations on the sound without being constrained by the mechanics of the instrument. This is the most interesting part of the digital instrument. Thus, in order to do sound transformations on this instrument, we should act on the different parameters of the model as illustrated in *Figure 5*.

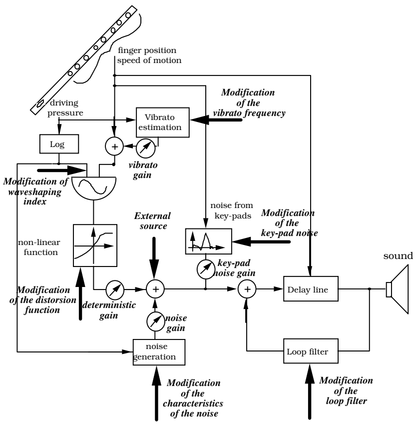

Figure 5. The hybrid flute model.

Control of the model's parameters. The bold and italic indications show the modification possibilities.

4 Conclusion

In this paper we have shown how to make a sound model using a combination of signal and physical models. The physical models take into account the most relevant physical characteristics of the sound generating system, while the signal models take into account perceptive effects. A way of designing physical models simulating the resonator of the instrument has been described. This model takes into account both dispersion and dissipation phenomena occurring during propagation in the medium. Such effects are important from a perceptive point of view. Then signal models were used to model the source of the instrument which was extracted from the sound by a deconvolution method. We further proposed to split the source signal into a deterministic and a stochastic part by the LMS algorithm, and to model these contributions independently. The deterministic part was modeled by a waveshaping synthesis method in order to take into account the non-linearities of the source signal. Perceptive criteria were then used to find the parameters to feed the synthesis model. The stochastic part was easy to implement thanks to the separation between the source and the resonator. It can be modeled by linear filtering of a white noise. In fact the stochastic part of the flute sound is colored due to the fact that a noise is propagating in the resonator.

As mentioned in the beginning, sound modeling consists in both resynthesis and transformation of natural sounds. We have therefore shown how sounds can be manipulated by the proposed model and how these manipulations can be done in real-time and piloted by an interface which we designed.

References

[1] Arfib, D. Digital synthesis of complex spectra by means of multiplication of non-linear distorted sine waves. Journal of the Audio Engineering Society, 1979, 27, pp. 757-768.

[2] Beauchamp, J. W. Synthesis by Spectral Amplitude and " Brightness " Matching of Analyzed Musical Instrument Tones. Journal of the Audio Engineering Society, 1982, Vol. 30, No. 6.

[3] Guillemain, P. Analyse et modelisation de signaux sonores par des représentations temps-fréquence linéaires. PhD thesis, Université Aix-Marseille II, juin 1994.

[4] Karplus, K. & Strong, A. Digital Synthesis of Plucked String and Drum Timbres. Computer Music Journal, 1983, 2 (7): 43-55.

[5] Kronland-Martinet, R., Morlet, J., & Grossmann, A. Analysis of sound patterns through wavelet transforms. International Journal of Pattern Recognition and Artificial Intelligence, 1987, 11 No. 2, pp. 97-126.

[6] Le Brun, M. Digital waveshaping synthesis. Journal of the Audio Engineering Society, 1979, 27, pp. 250-266.

[7] Pollard H.F. & Jansson E.V. A Tristimulus Method for the Specification of Musical Timbre. Acoustica, 1982, Vol. 51.

[8] Smith, J.O. Physical modeling using digital waveguides. Computer Music Journal, 1992, 16 No. 4, pp. 74-91.

[9] Widrow B. & Stearns S.D. Adaptive Signal Processing. Englewood Cliffs, Prentice-Hall Inc., 1985.

[10] Yegnanarayana B. Design of recursive group-delay filters by autoregressive modeling. IEEE Trans. of Acoust. Sp. and Sig. Proc. ASSP-30, 632-637.

[11] Ystad S. Sound Modeling using a combination of physical and signal models. PhD thesis, Université Aix-Marseille II, march 1998.