From Sound Modeling to Analysis-Synthesis of Sounds

R. Kronland-Martinet[1] , Ph. Guillemain[2] & S. Ystad[1]

(1) Stanford University, CCRMA, Stanford CA 94305-8180 USA
(2) CNRS – LMA 31 chemin J. Aiguier 13402 Marseille Cedex 20 France

Abstract

Sound modeling consists in designing synthesis methods to generate sounds. The process of Analysis-Synthesis consists in reconstructing a given natural sound using algorithmic techniques. This paper deals with the problem of designing methods to extract the parameters of the sound models so that the generated sound is similar to a given natural sound. We start by presenting an overview of sound models including signal, physical and hybrid models. Then we discuss the analysis problem with a and focus on time-frequency methods which are well adapted to the analysis of musical sounds. We further deal with the analysis-synthesis problem for each class of sound modeling, including additive synthesis, non-linear synthesis and digital waveguide synthesis. As a conclusion, analysis-synthesis approaches based on hybrid models are described.

1 Scope of the article

Analysis-synthesis is a set of procedures to reconstruct a given natural sound. Different methods can be used, and the success of each method depends on their adaptive possibilities and the sound effect to be produced. The «direct» analysis-synthesis process consists in reconstructing a sound signal by inversion of an analysis procedure. This is a useful process that uses an invertible analysis method to get a representation of a sound and permits various sound transformations by acting on the parameters between the analysis and the synthesis process. Nevertheless, the result of such processes is not always intuitive since the modification of the analysis parameters is generally not a valid operation from a mathematical point of view. In this article we specially pay attention to the analysis-synthesis process associated to sound modeling. Here, the representations obtained from the analysis provide parameters corresponding to given synthesis models. This brings us to the concept of algorithmic sampler which consists in simulating natural sounds through a synthesis process which is well adapted to algorithmic and real time manipulations. The resynthesis and the transformation of natural sounds are then part of the same concept, allowing the morphing of natural sounds in a way which is similar to the modification of sounds using an algorithmic synthetizer.
The paper is organized as follows. We first present the most commonly used synthesis methods. Then analysis methods such as time-frequency and wavelet transforms are described, as well as algorithms for separating and characterizing spectral components. We conclude by showing how the analysis of real sounds can be used to estimate the synthesis parameters corresponding to different classes of sound models. Most of these techniques have been developed in our group «Modeling, Synthesis and Control of Audio and Musical Signals» (acronym S2M) in Marseille, France. This paper is an updated concentrate of the previously published paper [1].

2 Sound Synthesis

Digital synthesis uses methods of signal generation that can be divided into two classes:
- signal models aimed at reconstructing a perceived effect without being concemed with the specific source that made the sound.
- physical models aimed at simulating the behavior of existing or virtual sound sources.

2.1 Signal Model Synthesis

Signal models use a purely mathematical description of sounds. They are numerically easy to implement, and they guarantee a close relation between the synthesis parameters and the resulting sound. These methods are similar to shaping and building structures from materials, and the three principal groups can be classified as follows
- additive synthesis
- subtractive synthesis
- non-linear or global synthesis

2.1.1 Additive Synthesis

A complex sound can be constructed as a superposition of elementary sounds, generally sinusoidal signals modulated in amplitude and frequency. For periodic or quasi periodic sounds, these components have average frequencies that are multiples of one fundamental frequency and are called harmonics. The periodic structure leads to electronic organ sounds if one does not consider the micro variations that can be found through the amplitude and frequency modulation laws of the components of any real sound. These dynamic laws must therefore be very precise when one reproduces a real sound. The advantages of these synthesis methods are essentially the possibilities of intimate and dynamic modifications of the sound. Granular synthesis can be considered as a special kind of additive synthesis, since it also consists in summing up elementary signals (grains) localized in both the time and the frequency domains [2].

2.1.2 Subtractive Synthesis

Like sculptor removes unwanted parts from his stone, a sound can be constructed by removing undesired components from an initial, complex sound such as a noise. This synthesis technique is closely linked to the theory of digital filtering [4] and can be related to some physical sound generation systems like for instance the speech signal [5], [6]. The advantage of this approach (if we omit the physical aspects which will be discussed when describing synthesis models by physical modeling) is the possibility of uncoupling the excitation source and the resonance system. The sound transformations related to these methods often use this property in order to make hybrid sounds or crossed synthesis of two different sounds by combining the excitation source of a sound and the resonant system of another [7][8].

2.1.3 Non-linear or Global Synthesis

Like modeling different objects from a block of clay, a simple and "inert" signal can be dynamically modeled using global synthesis models. This method is non-linear since the operations on the signals are not simple additions and amplifications. The most well-known example of global synthesis is audio Frequency Modulation (FM) updated by John Chowning [9]. The advantage of this method is that it calls for very few parameters, and that a small number of operations can generate complex spectra. This simplifies the numerical implementation and the control. However, it is difficult to control the shaping of a sound by this method, since the timbre is related to the synthesis parameters in a non-linear way and the continuous modification of these parameters may give discontinuities in the sound. Other related methods have shown to be efficient for signal synthesis, such as the waveshaping technique [10].

2.2 Physical Model Synthesis

Unlike signal models using a purely mathematical description of sounds, physical models describe the sound generation system through physical considerations. Such models can be constructed either from the equations describing the behavior of the waves propagating in the structure and their radiation in air, or from the behavior of the solution of the same equations. The first approach is costly in terms of calculations and is generally used only in connection with research work [11]. Synthesis by simulation of the solution of the propagation equation has led to the digital waveguide synthesis models [12], which have the advantage of being easy to construct with a behavior close to that of a real instrument. Thus such synthesis methods are well adapted to the modeling of acoustical instruments and can be used to simulate many different systems, such as for instance the tube representing the resonant system in wind instruments [13].

2.3 Synthesis by Hybrid Models

It is possible to combine physical and signal models into a so-called hybrid model [14], [15]. Such a model takes advantage of the positive aspects of both of the previous methods. As already mentioned the physical part of the model makes it possible to take into account physical characteristics of the sound producing system so that a physical interpretation can be linked to the model's parameters. The signal model part makes it possible to simulate systems that normally would be too time demanding or complicated to be simulated by physical models. For a large number of musical instruments this implies that a signal model should model the excitation part while a physical model generally can model the resonator part. This means that although the physics of musical instruments often is too complicated for a purely physical model to be applied, physically meaningful parameters can be included in the model, since parts of it are constructed by physical modeling. This facilitates for instance the control of the model with an interface and makes it possible to make sounds from virtual instruments with exaggerated physical characteristics like for instance gigantic strings on a violin.

The coupling between the physical and signal models is of importance for the complexity of the hybrid model. In most cases the two models will interact reciprocally like in the flute case where a non-linear signal model can model the excitation, while a digital waveguide model can model the resonator. The interaction between these two models will in this case lead to a non-linear waveguide synthesis technique [16].

3 Sound Analysis

The analysis of natural sounds calls for several methods giving a description or a representation of pertinent physical and perceptual characteristics of the sound [17]. Even though the spectral content of a sound generally is very important, the time course of its energy is at least as important. This can be shown by artificially modifying the attack of a percussive sound in order to make it "woolly", or by playing the sound backwards. In these cases the sound will be radically changed while the power spectral density remains the same. The time and frequency evolution of each partial component is also essential for the perceived characteristics of the sound. Effects like vibrato and tremolo can be analyzed from these evolutions. Such information also gives access to another perceptually important aspect of the sound, namely the different decay times of the partials of transient sounds such as sounds from plucked vibrating strings. To solve this general analysis problem of signals, a collection of methods called joint representations can be used.

The analysis methods can be divided in two principal classes: parametric methods and non-parametric methods. The parametric methods require a *a priori* knowledge of the signal, and consist in adjusting the parameters of a model. The non-parametric models do not need any knowledge of the signal to be analysed, but they often require a larger number of coefficients. We shall in this article focus on this last analysis class, since it generally corresponds to representations with physically and/or perceptually meaningful parameters.

3.1 Spectral Analysis

The best known representation is the spectral representation obtained through the Fourier transform. The signal is in this case associated with a representation giving the energy distribution as a function of frequency. As mentioned earlier, this representation is not sufficient for characterizing the timbre and the dynamic aspects of a sound. Actually, even though all the information concerning the sound is present in the Fourier transform, it is not obvious to give a sense to the phase of the transform. This is why only the modulus of the spectrum is generally considered, leading to an average energy information. Hence, the time information is dramatically missing.

3.2 Time-frequency and time-scale techniques

In what follows we describe the joint time-frequency representations considering both dynamical and frequencial aspects. The time-frequency transforms distribute the total energy of the signal in a plane similar to a musical score in which one of the axes corresponds to the time and the other to the frequency. Such representations are to the sound what the musical scores are to the melodies. They can be obtained in two different ways depending on whether the analysis acts on the energy of the signal or on the signal itself. In the first case the methods are said to be non-linear (or bilinear), giving for instance representations from the so-called “Cohen’s class”. The best known example of transformations within this class is the Wigner-Ville distribution [18]. In the second case the representations are said to be linear, leading to the Fourier transform with sliding window, the Gabor transform, or the wavelet transform. The linear methods have, at least as far as sound signals are concerned, a great advantage compared to the non-linear methods since they always make the resynthesis of signals possible and ensure that no spurious terms cause confusion during the interpretation of the analysis. We will therefore focus on the linear time-frequency methods.

By decomposing the signal into a continuous sum of elementary functions having the properties of localization both in time and in frequency, linear representations are obtained. These elementary functions correspond to the impulse response of bandpass filters. The central frequency of the analysis band is related to a frequency parameter for time-frequency transforms and to a scaling parameter for wavelet transforms. The choice of the elementary functions gives the shape of the filters.

3.2.1 Gabor transform

What we nowadays call the Gabor transform is an extension of the transform proposed by D. Gabor for signal information purposes. The elementary functions he used were exclusively Gaussian. It has now become a convention to call Gabor transform every transform of the same kind even though the elementary functions are not Gaussian. The elementary functions of the Gabor transform, also called time-frequency atoms, are all generated from a mother function (window) translated in time and in frequency. The mother function is chosen to be as well-localized as possible in time and frequency (this is limited by the Heisenberg incertainty principle) and to have finite energy (for instance a Gaussian function) (figure 1).

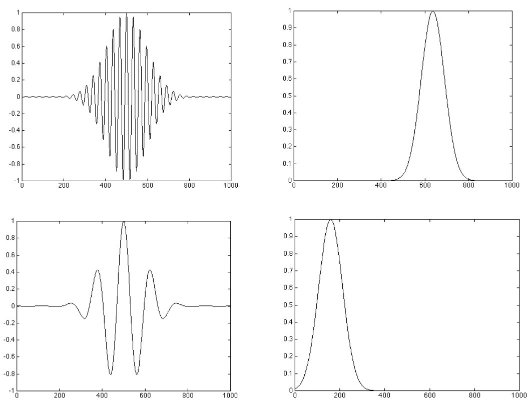

Figure 1: Two Gabor functions in the time domain (left), and their Fourier transform (right).

Each value of the transform in the time-frequency plane is obtained by comparing the signal to a time-frequency atom. This comparison is mathematically expressed by a scalar product. Each horizontal line of the Gabor transform then corresponds to a filtering of the signal by a band-pass filter centered at a given frequency with a shape that is constant as a function of frequency. The vertical lines correspond to the Fourier transform of a part of the signal, isolated by a window centered at a given time. The transform obtained this way is generally complex-valued, since the atoms themselves are complex-valued, giving two complementary images [19].

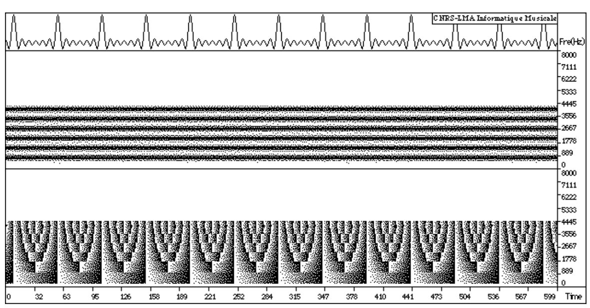

Figure 2: Gabor transform of the sum of six harmonic components. The horizontal axis is time. The vertical axis is frequency. The upper picture is the modulus, the lower is the phase, represented modulo-2π. In this case, the window is well localized in frequency, allowing the resolution of each component.

The first one is the modulus of the transform and corresponds to a classical spectrogram, the square of the modulus being interpreted as the energy distribution in the time-frequency plane. The second image corresponds to the phase of the transform and is generally less used, even though it contains a lot of information. This information mainly concerns the "oscillating part" of the signal (figures 2 and 3). Actually, the time derivative of the phase has the dimension of a frequency and leads to the frequency modulation law of the spectral components of the signal [20].

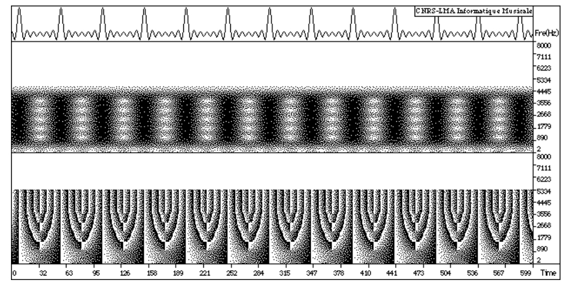

Figure 3: Gabor transform of the same signal as in figure 2, but with a window that is well localized with respect to time, leading to a bad separation of the components in the frequency domain, but showing impulses in time. In both pictures, the phase behaves similarly, showing the periodicity of each component.

3.2.2 Wavelet transform

The wavelet transform follows a principle close to that of the Gabor transform. Again the horizontal lines of the wavelet transform correspond to a filtering of the signal, but in this case the shape of the filter is independent of the scale while the bandwidth is inversely proportional to the scale. The analysis functions are all obtained from a mother wavelet by translation and change of scale (dilation) (figure 4).

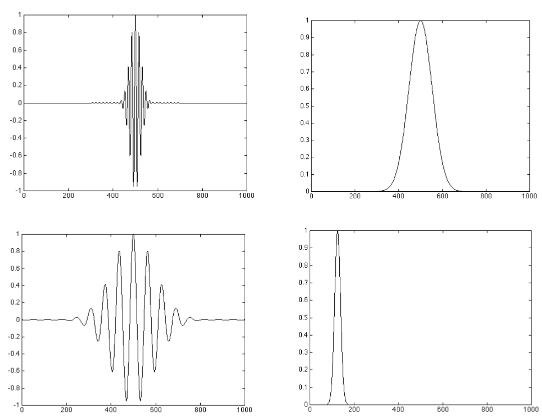

Figure 4: Two wavelets in the time domain (left), and their Fourier transform (right). All the filters of the wavelet representation are obtained through dilation of a mother function in time, yielding a constant relative bandwidth analysis.

The mother wavelet is a function with finite energy and zero mean value. These "weak" conditions offer great freedom in the choice of the mother wavelet. One can for example imagine a decomposition of a speech signal in order to detect the word "bonjour" pronounced at different pitches and with different duration. By using a mother wavelet made of two wavelets separated for example by an octave, one can detect octave chords in a musical play [3]. This corresponds to a matched filtering at different scales. One important aspect of the wavelet transform is the localization. By acting on the dilation parameter, the analyzing function is automatically adapted to the size of the observed phenomena (figure 5). A high frequency phenomenon should be analyzed with a function which is well-localized in time, whereas a low-frequency phenomenon requires a function which is well-localized in frequency. This leads to an appropriate tool for the characterization of transient signals [20]. The particular geometry of the time-scale representation, where the dilation is represented according to a logarithmic scale (in fraction of octaves) enables the transform to be interpreted like a musical score associated to the analyzed sound.

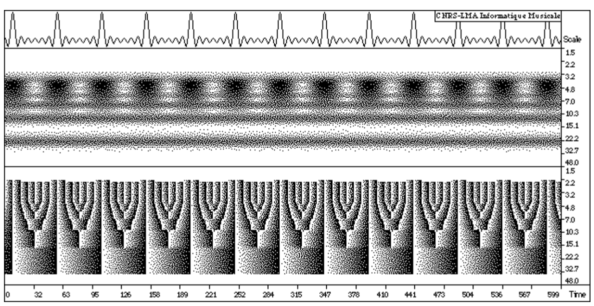

Figure 5: Wavelet transform of the sum of six harmonic components which were analyzed in figures 2 and 3. In contrast to the former representations obtained through the Gabor transform, the wavelet transform privileges the frequency accuracy at low frequencies (large scales) and the time accuracy at high frequencies (small scales).

4 Parameter extraction

The parameter extraction method makes use of the qualitative information given by the time-frequency and the time-scale transform in order to extract quantitative information from the signal. Even though the representations are not parametric, the character of the extracted information is generally determined by the supposed characteristics of the signal and by future applications. A useful representation for isolated sounds from musical instruments is the additive model. It describes the sound as a sum of elementary components modulated in amplitude and in frequency, which is relevant from a physical and a perceptual point of view (figure 6).

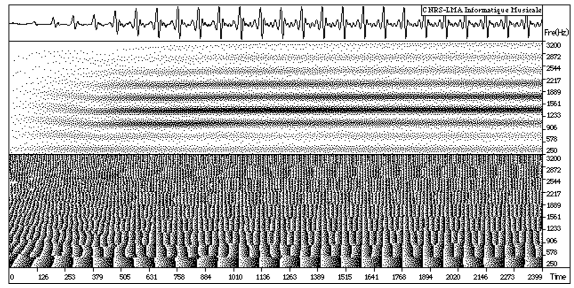

Figure 6: Gabor representation of the first 75ms of a trumpet sound. Many harmonics with different time dependencies are visible on the modulus picture. The phase picture shows different regions, around each harmonic, where the phase wraps regularly at the time period of each harmonic.

Thus, to estimate parameters for an additive resynthesis of the sound, amplitude and frequency modulation laws associated to each partial should be extracted from the transform. Of course, this process must be efficient even for extracting components that are very close to each other and which have rapidly changing amplitude modulation laws. Unfortunately all the constraints for constructing the representation make this final operation complicated. This is due to the fact that absolute accuracy both in time and in frequency is impossible because of a mathematical relation between the transform in one point of a time-frequency plane and the close vicinity of this point. Human hearing follows a rather similar "uncertainty" principle: to identify the pitch of a pure sound, it must last for a certain time. The consequences of these limitations on the additive model parameter estimation are easy to understand. A high-frequency resolution necessitates analysis functions that are well-localized in the frequency domain and therefore badly localized in the time domain. The extraction of the amplitude modulation law of a component from the modulus of the transform (on a trajectory) in the time-frequency plane smoothes the actual modulation law. This smoothing effect acts in a time interval with the same length as the analysis function. Conversely, the choice of well-localized analysis functions in the time domain generally yields oscillations in the estimated amplitude modulation laws, because of the presence of several components in the same analysis band.

To optimize the estimation of each component one can use a specific procedure based on the construction of a filter bank aiming at separating each component [20]. The procedure is done in two steps:

- precise estimation of the frequencies of the components using the phase of the transform (analytic signal)
- construction of a filter bank (by linear combinations of analysis functions) positioned at the frequencies of the components. The functions are weighted so that the filter equals one at the peak of the selected component and zero at the peaks of the other components.

Another important aspect of the musical sound is the frequency modulation of the components, in particular during the attack of the sound. Here the judicious use of the time derivative of the transform's phase offers the possibility of developing iterative algorithms tracking the modulation laws, and thus precluding the computation of the whole transform. These algorithms use frequency-modulated analysis functions, the modulations of which are automatically matched to the ones of the signal [20].

5 Fitting synthesis parameters

The extraction techniques using the time-frequency transforms directly provide a group of parameters, allowing the resynthesis of a sound with the additive model. In addition, they can be used for identification of other synthesis models. The direct parameter identification techniques for the non-linear models are complicated. Generally they do not give an exact reproduction of a given sound. The estimation criteria can be statistical (minimization of non-linear functions) [21] or psychoacoustic [22],[15]. The direct estimation of physical or subtractive model parameters requires techniques like linear prediction, used for instance in speech synthesis [23]. Another solution consists in using parameters from the additive synthesis model to estimate another set of parameters corresponding to another synthesis model. In this section we shall see how this operation can be done for some well-known models.

5.1 Parameter estimation for signal model synthesis

The parameter estimation for the additive model is the simplest one, since the parameters are determined in the analysis. The modeling of the envelopes can greatly reduce the data when one uses only perceptual criteria. The first reduction consists in associating each amplitude and frequency modulation law to a piecewise linear function [24] (figure 7). This makes it possible to automatically generate, for example, a Music V or a Csound score associated to the sound.

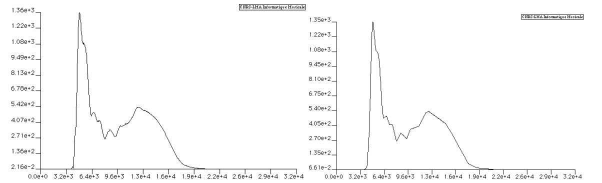

Figure 7: Original and modeled envelopes of a saxophone sound. The modeled curve is defined with 35 breakpoints and linear interpolation between them, while the original is defined on 32000 samples.

Another possible reduction consists in grouping the components from the additive synthesis (group additive synthesis) [25], [26]. This can be done by statistical methods, like principal component analysis, or by following an additive condition defined as the perceptual similarity between the amplitude modulations of the components [27]. This method offers a significant reduction in the number of synthesis parameters, since several components with a complex waveshape have the same amplitude modulation laws (figures 8 and 9).

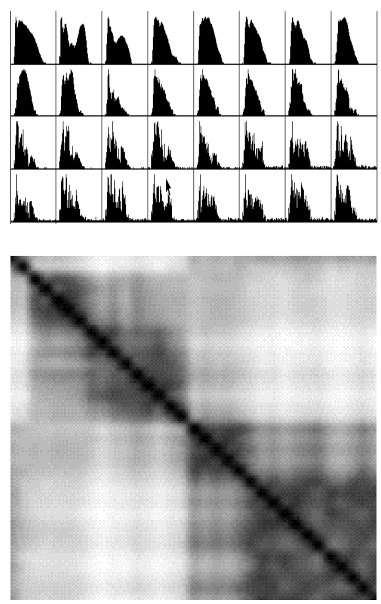

Figure 8: A whole set of envelopes of a violin sound, and the matrix showing the correlation between them. The dark regions around the diagonal correspond to curves that look similar and that correspond to components that are close in the frequency domain.

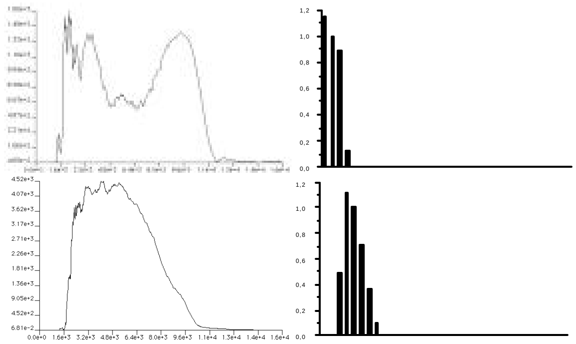

Figure 9: Two main envelopes of the group additive synthesis model, with the spectrum of their associated waveform. Psychoacoustic criteria can be used to generate a perceptively similar spectrum with non-linear techniques.

5.2 Subtractive synthesis

An evolving spectral envelope can be built by creating intermediate components obtained from the modulation laws of the additive modeling. Their amplitude modulation laws are obtained by interpolation of the envelopes of two adjacent components in the frequency domain (figure 10). These envelopes can then be used in order to "sculpt" another sound (crossed synthesis). As we already mentioned, physical modeling is sometimes close to subtractive synthesis. This aspect will be developed later.

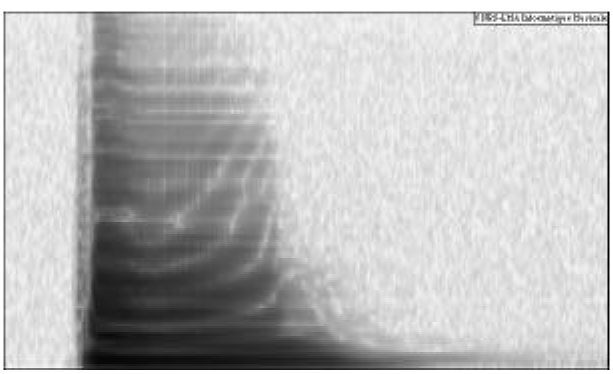

Figure 10: Spectral envelope of a saxophone sound built from the additive synthesis parameters. This envelope can be used to "sculpt" the modulus of the Gabor transform of another sound in order to perform a crossed synthesis.

5.3 Waveshaping and frequency modulation synthesis

From the parameters corresponding to the group additive synthesis (complex waves and their associated amplitude laws), one can deduce non-linear synthesis parameters [27]. The technique consists in approaching each complex wave shape by an elementary non-linear module. In the case of waveshaping, the knowledge of the complex wave allows the calculation of an exact distortion function. In the case of FM, the spectral components should be grouped, not only according to a perceptive criterion, but also according to a condition of spectral proximity. This condition is meaningful because real similarities between envelopes of neighbouring components are often observed. To generate the waveform corresponding to a group of components by an elementary FM oscillator, the perceptive approach is best suited. In that case, one can consider the energy and the spectral extent of the waveforms, which are directly related to the modulation index. Other methods based on the minimization of non-linear functions by the simulated annealing or genetic algorithms have also been explored [21]. Attempts at direct estimation of the FM parameters by extraction of frequency modulation laws from the phase of the analytic signal related to the real sound have led to interesting results [28], [29].
Recently, optimization techniques using perceptual criteria have been developed. Rather than attempting to reconstruct a signal similar to the original, the idea is in this case to look for a perceptually similar sound. Criterion such as the Tristimuli [30] can be used to define a perceptual distance. This distance is then minimized using classical methods [15] leading to the optimal value of the synthesis parameters.

5.4 Parameter estimation for physical model synthesis

The digital waveguide synthesis parameters are of a different kind. They characterize both the medium where the waves propagate and the way this medium is excited. From a physical point of view, it is difficult to separate these two aspects: the air jet of a wind instrument causes vortex sheddings interacting with the acoustic pressure in the tube [31]; the piano hammer modifies the characteristics of a string while it is in contact with it [32]. These source-resonator interactions are generally non-linear and often difficult to model physically. However, a simple, linear digital waveguide model often gives satisfactory sound results. In a general way, the study of linear wave propagation equations in a bounded medium shows that the response to a transient excitation can be written as a sum of exponentially damped sine functions. The inharmonicity is related to the dispersive characteristics of the propagation medium, the decay times are related to the dissipative characteristics of the medium, and the amplitudes are related to the spectrum of the excitation. In the same way, the impulse response of the simple digital waveguide model can be approximated by a sum of exponentially damped sinusoids whose frequencies, amplitudes, and damping rates are related in a simple way to the filter coefficients [14]. Thanks to the additive synthesis parameters one can, for transient and percussive sounds, determine the parameters of the waveguide model, and also recover the physical parameters characterizing the instrument [33] (figure 11). For sustained sounds, the estimation problem of the exciting source is crucial and necessitates the use of deconvolution techniques. This approach is entirely non parametric, but it is also possible to use parametric techniques. Indeed, the discrete time formulation of the synthesis algorithm corresponds to a modeling of the so-called ARMA type (AutoRegressive Moving Average).

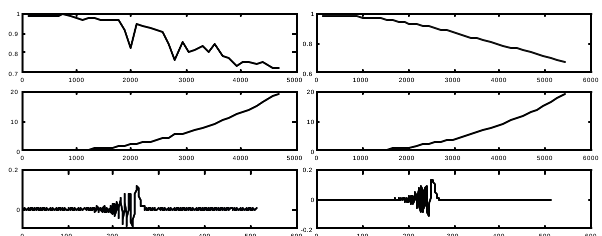

Figure 11: Parameter estimation for the digital waveguide model. Pictures on the left show the data from the estimation. Pictures on the right show the data from the movement equation of a stiff string. Respectively, from top to bottom: modulus (related to losses during the propagation); phase derivative (related to the dispersion law of the propagation medium) of the Fourier transform of the filter inside the loop; impulse response of the loop filter.

5.5 Parameter estimation for hybrid model synthesis

Hybrid models being a combination of signal and physical models, and their parameters can be estimated using most of the previous techniques. Actually, starting with a transient response of the instrument (sound generated from a rapidly closed fingerhole for example) one can estimate the parameters of the resonator's model. Then, using a deconvolution technique, one can extract the source signal from the natural sound. This source can then be modeled using signal models such as waveshaping. Two examples of such methods can be found in [14] for the flute case and [34] for the piano case.

This approach is being improved by considering a more general model which constitutes a loop system including both a linear filter and a non-linear element. This model is very general and can be considered as a "non-linear digital waveguide". Thanks to the previous method, one can estimate the parameters of such a model, even though its non-linear behavior makes the analysis-synthesis process very complicated [16].

6 Acknowledgments

This work was partly supported by the Norwegian Research Council

References

[1] Kronland-Martinet R., Guillemain, Ph & Ystad S. 1997, "Modeling of natural sounds by time-frequency and wavelet representations", Organised Sound, 2(3) : 179-191.

[2] Roads, C. Automated granular synthesis of sound. Computer Music Journal, 1978, 2, No. 2, 61-62.

[3] Kronland-Martinet, R. The use of the wavelet transform for the analysis, synthesis and processing of speech and music sounds. Computer Music Journal, 1989, 12 n° 4, 11-20 (with sound examples on disk).

[4] Rabiner L.R, Gold B., Theory and Application of Digital Signal Processing ,1975, Prentice Hall.

[5] Atal B.S, & Hanauer S.L 1971, "Speech analysis and synthesis by linear prediction of the speech wave" j. Acoust. Soc. Amer., 50, p. 637-655.

[6] Flanagan J.L, Coker C.H, Rabiner P.R, Schafer R.W., & Umeda N. , Synthetic voices for computer. I.E.E.E Spectrum, 1970 , 7, p. 22-45.

[7] Makhoul, Linear prediction, a tutorial review. Proceedings of Institute of Electrical and Electronics Engineers, 1975, 63, 561-580.

[8] Kronland-Martinet R., Digital substractive synthesis of signals based on the analysis of natural sounds. In "Etat de la Recherche Musicale (au 1er janvier 1989)" 1988, Ed. A.R.C.A.M. , Aix en Provence.

[9] Chowning, J. The synthesis of complex audio spectra by means of frequency modulation. Journal of the Audio Engineering Society, 1973, 21, 526-534.

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

[11] Chaigne, A. Trends and Challenges in Physical Modeling of Musical Instruments. Proceedings of the ICA, Volume III, p.397-400 Trondheim, Norway 26-30 june 1995.

[12] Smith, J. Physical modeling using digital waveguides. Computer Music Journal, 1992, 16 n° 4, 74-91.

[13] Cook, P.R. A meta-wind-instrument physical model controller, and a meta-controller for real-time performance control. Proceedings of the 1992 International Music Conference, San Francisco: Computer Music Association, 1992, 273-276.

[14] Ystad, S. 1998. *"Sound Modeling Using a Combination of Physical and Signal Models."* Ph. D. Thesis from the University of Aix-Marseille II.

[15] Ystad, S. 2000. "Sound Modeling Applied to Flute Sounds" *Journal of the Audio Engineering Society*, 48(9) ; 810-825.

[16] Ystad, S & Voinier, Th. Analysis-Synthesis of Flute Sounds with a Looped Non-linear Model. MOSART workshop, Barcelona, Spainm November 2001

[17] Risset J.C, & Wessel D.L, Exploration of timbre by analysis and synthesis, in D. Deutsch ed., The Psychology of Music, Academic Press 1982, 26-58.

[18] Flandrin P. Temps-fréquence. 1993, Hermès, Traité des nouvelles technologies, série traitement du signal.

[19] Kronland-Martinet, R., Morlet, J., & Grossman, A. Analysis of sound patterns through wavelet transforms. Intern. Journal of Pattern Recognition and Artificial Intelligence, 1987, 11 n° 2, 97-126.

[20] Guillemain, Ph., & Kronland-Martinet, R., Characterisation of Acoustics Signals Through Continuous Linear Time-frequency Representations Proceedings of the IEEE, Special Issue on Wavelets, 1996, Vol. 84 n°4, 561-585.

[21] Horner, A. Double-Modulator FM Matching of Instruments Tones, Computer Music Journal, 1996, V. 20 (2), pp 57-71

[22] Beauchamp, J. W. Analysis and synthesis of cornet tones using non-linear interharmonic relationships. Journal of the Audio Engineering Society, 1975, 23, 778-795.

[23] Markel, J. D, & Gray, A. H. Linear Prediction of Speech. Springer-Verlag, Communication and Cybernetics (12), 1976.

[24] Horner, A., & Beauchamp, J. Piecewise-linear approximation of additive synthesis envelopes: a comparison of various methods. Computer Music Journal, 1996, 20 n° 2, 72-95.

[25] Oates S., & Eagleston B. Analytic Methods for Group Additive Synthesis, Computer Music Journal, 1997, V. 21 (2), 21-39.

[26] Kleczkowski, P. Group additive synthesis. Computer Music Journal, 1989, 13 n° 1, 12-20.

[27] Kronland-Martinet, R., & Guillemain, P. Towards non-linear resynthesis of instrumental sounds. Proceedings of the 1993 International Music Conference, San Francisco: Computer Music Association, 1993, 86-93.

[28] Justice, J. Analytic signal processing in music computation. IEEE Transactions on Speech, Acoustics and Signal Processing, 1979, ASSP-27, 670-684.

[29] Delprat, N., Guillemain, P., & Kronland-Martinet, R. Parameter estimation for non-linear resynthesis methods with the help of a time-frequency analysis of natural sounds. Proceedings of the 1990 International Music Conference, Glasgow: Computer Music Association, 1990, 88-90.

[30] Pollard H.F. and Jansson E.V. 1982. "A Tristimulus Method for the Specification of Musical Timbre." *Acoustica*, Vol. 51 ;162-171.

[31] Verge M.P. 1995. "Aeroacoustics of Confined Jets with Applications to the Physical Modeling of Recorder-Like Instruments." PhD thesis, Eindhoven University.

[32] Weinreich, G. Coupled piano strings. Journal of the Acoustical Society of America, 1977, 62, 1474-1484.

[33] Guillemain Ph., Kronland-Martinet R., & Ystad S.. Physical Modelling Based on the Analysis of real Sounds. Proceeding of Institude of Acoustics, Edinburgh 1997, Vol 19, ISMA 97, 445-450

[34] J. Bensa, K. Jensen, R. Kronland-Martinet, S. Ystad "Perceptual and Analytical Analysis of the effect of the Hammer Impact on the Piano Tones", proceedings of the International Computer Music Conference ICMC 2000, pp.58-61, Berlin (Germany) 27 août au 1 September 2000.