Structured Additive Synthesis: Towards a Model of Sound Timbre and Electroacoustic Music Forms

Myriam Desainte-Catherine (myriam@labri.u-bordeaux.fr)
Sylvain Marchand (sm@labri.u-bordeaux.fr)

SCRIME - LaBRI - Université Bordeaux I
351 cours de la Libération, F-33405 Talence Cedex, France

Abstract

We have developed a sound model used for exploring sound timbre. This model is called Structured Additive Synthesis, or SAS for short. It has the flexibility of additive synthesis while addressing the fact that basic additive synthesis is extremely difficult to use directly for creating and editing sounds. SAS consists of a complete abstraction of sounds according to only four parameters: amplitude, frequency, color, and warping. These parameters are inspired by the vocabulary of composers of electro-acoustic music as well as by the literature and constitute a solid base for investigating scientific research on the notion of timbre. Several analyses of electro-acoustic pieces have been performed in collaboration between scientists and musicians. We have identified the need for a certain number of manipulations of sound, that we have determined to be straightforward in our model. Applications of the SAS model are numerous. A new language for musical composition has been implemented and should provide a way to validate and enrich the model.

1 Introduction

The SCRIME is an organization for scientific researchers in computer science at the University and music composers of the Conservatoire to collaborate. Projects of the SCRIME should not only be scientifically valid, but also musically relevant. Research projects of this structure are mainly situated in the field of the assistance for composition of electro-acoustic music. We observe and try to understand actual practices of electro-acoustic composers in order to provide our research in sound and music modeling with new elements. One motivation is the study of sound timbre from a perceptual and musical point of view, in collaboration with psycho-acousticians. Another motivation is to provide composers with tools well adapted to their actual needs.

In this paper, we present three research subjects that are relevant in order to reach our objectives. The second section presents the analysis of electro-acoustic music which is studied in close collaboration with composers. The third section presents the SAS sound model that has been implemented and is been validated in collaboration with psycho-acousticians and composers. The fourth section shows the applications of this model in compositional and educational contexts.

2 Music Analysis

A musical analysis of a piece consists first in segmenting the piece in order to discover a temporal organization between several sound objects. Musical discourse can be analyzed on the basis of those sound objects by pointing out relations between parameters of different parts of the piece.

When the piece is written, the analysis is based on the musical score which provides the initial segmentation. Electro-acoustic musical pieces constitute a very special case because they are not written. Their support is magnetic or numeric.

Among the analyses identified by François Delalande [Del86] we chose to perform the poietic (production-oriented) and aesthesic (reception-oriented) ones.

Poietic analysis is based on production. Such an analysis can be carried out in collaboration with the composer of the piece to analyze. Its objective is to study the musical discourse in order to find out information about the production of the piece, that is, the tools and the practices that were used in order to build the piece.

Aesthesic analysis is based on listening. Such an analysis can be performed by a composer or by a listener who is very familiar with electro-acoustic music, or by conducting experiments involving several listeners. This kind of analysis provides information on the way listeners understand electro-acoustic music.

As a matter of fact, only poietic and aesthesic analyses provide information concerning models that are in the composer's mind when he composes music or when he listens to the music. We conducted the following two analyses in collaboration with composers.

We performed an aesthesic analysis of the second movement "Balancement" of the "Variations pour une porte et un soupir" by Pierre Henry (this work has been carried out in collaboration with the composer Edgar Nicouleau [DCN98]). In that movement, inflexions of the grating door are very close to voice modulation so that they remind listeners of melodic, rhythmic and dynamic structures that are usually analyzed in that case. Sound objects of the piece have been itemized and then grouped in several families. A common formalism permits the description of the evolutions of frequencies, durations and amplitudes for all the sound objects. This analysis leads to quite classical results since it involves well-known structures like melody, dynamic and rhythm. Of course, the case of the analyzed piece is very particular and such results cannot be obtained with any electro-acoustic piece. Anyway, this research may continue with the study of the timbre structures.

We also performed a poietic analysis of the second movement of “La chute d'Icare” by Jean-Michel Rivet (this work has been carried out in collaboration with him [DCR98]). A first segmentation is proposed as well as a classification of the sounds according to the production of the piece. Then, several segmentations based on this classification are studied and a temporal structure is discovered. This analysis has pointed out structures that were pertinent for the composer. For example, a classification of sounds was obtain according to the composer's criteria for choosing one sound rather than another. Those criteria may vary from one composer to another and according to his objectives so that it is necessary to make the same kind of collaboration with several composers. The objective is on the one hand, to find out some musical elements which could be useful to several composers, and to help us in sound modeling on the other hand.

All these analyses of electro-acoustic pieces have been performed in collaboration between scientists and musicians of the SCRIME. We have identified the need for a certain number of manipulations of sounds. Among these are modulation, mixing, filtering, time stretching, cross-synthesis, morphing, as well as new ways to create hybrid sounds. The problem was yet to find a sound model allowing the composers to perform these manipulations in an intuitive and musical way.

3 The SAS Model

The Structured Additive Synthesis (SAS) model is a spectral sound model based on additive synthesis. The SAS parameters are inspired by the vocabulary of composers of electro-acoustic music as well as by the literature. We propose to focus on the perception of the sound rather than its physical cause, in order to unify sound (microscopic) and music (macroscopic). We propose as well to consider the musical intention of the instrumentalist instead of his physical action on the instrument.

3.1 Additive Synthesis

Additive synthesis is the original spectrum modeling technique. It is rooted in Fourier's theorem, which states that any periodic function can be modeled as a sum of sinusoids at various amplitudes and harmonic frequencies. For pseudo-periodic sounds, these amplitudes and frequencies evolve slowly with time, controlling a set of pseudo-sinusoidal oscillators commonly called *partials*. The audio signal a can be calculated from these additive parameters using the following equations:

$$a(t) = \sum_{p=1}^{P} a_p(t) \; cos(\phi_p(t)) \tag{1}$$

$$\phi_p(t) = \phi_p(0) + 2\pi \int_0^t f_p(u) \; du \tag{2}$$

where P is the number of partials and f_p, a_p, and ϕ_p are respectively the instantaneous frequency, amplitude and phase of the p-ieth partial. The P pairs (f_p, a_p) are the parameters of the additive model and represent points in the frequency-amplitude space, as shown in figure 1.

Any sound can be faithfully synthesized in real time from the model equations containing these parameters. The real-time synthesis has been implemented in the ReSpect software tool [MS99]. The difficulty is then to obtain these parameters from real, existing sounds. For that reason, we have developed an analysis method capable of converting sampled sounds into the SAS parameters, implemented in the InSpect program [MS99]. It is of course possible to eliminate analysis entirely, and create new sounds directly, using the parameters of our model. This is indeed possible because there is a close correspondence between these parameters and real music perception.

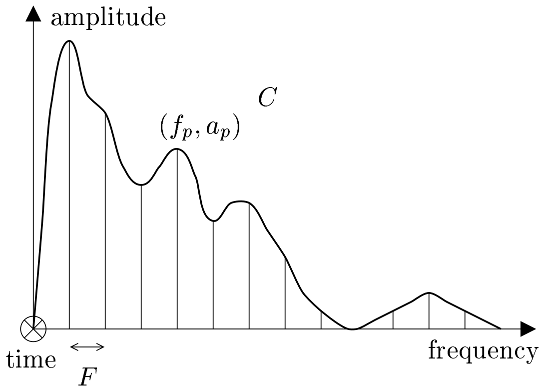

Figure 1: the spectrum of an harmonic sound.

3.2 Structured Additive Synthesis

The additive synthesis model is extremely difficult to use directly for creating and editing sounds. The reason for this difficulty is the huge number of model parameters which are only remotely related to musical parameters as perceived by a listener. The Structured Additive Synthesis (SAS) model has the flexibility of additive synthesis while addressing these problems. It imposes constraints on the additive parameters, giving birth to structured parameters as close to perception and musical terminology as possible, thus reintroducing a perceptive and musical consistency back into the model. The remaining of this section quickly presents the SAS model. An extended presentation can be found in [DCM99].

3.2.1 Structured Parameters

SAS consists of a complete abstraction of sounds according to only four physical parameters, functions closely related to perception:

Amplitude A : time $\rightarrow$ amplitude
Human beings perceive amplitude on a logarithmic scale. Amplitude can be calculated from the additive parameters like this: $A(t) = \sum_{p=1}^{P} a_p(t)$. Calculating the volume in dB from the amplitude is easy: $20\ log_{10}(\frac{A}{A_{0dB}})$.

Frequency F : time $\rightarrow$ frequency
The way of calculating the frequency from the additive parameters is trickier, and can be found in [DCM99]. Anyway, for harmonic sounds F coincides with the fundamental, possibly missing or "virtual". The frequency is also perceived on a logarithmic scale. For example, the MIDI pitch is a function of frequency: $57 + 12\ log_2(\frac{F}{440})$.

Color C : frequency $\times$ time $\rightarrow$ amplitude
Color coincides with an interpolated version of the spectral envelope [Ris86]. We call it color by analogy between audible and visible spectra. This analogy is already well-known for noises (white, blue, etc.).

Warping W : frequency $\times$ time $\rightarrow$ frequency
Generally, the partial frequencies are not exactly multiples of the fundamental frequency F. Warping gives the real frequency of a partial from the theoretical one it should have had if the sound had been harmonic. Of course for all harmonic sounds $W(t) = \mathbf{Id}$, that is, $\forall t, W(f,t) = f$.

3.2.2 Structured Equations

From the four structured parameters, we can calculate the the audio signal a:

$$a(t) = A(t)\frac{\sum_{p=1}^{P} C(W(pF(t),t),t)cos(\phi_p(t))}{\sum_{p=1}^{P} C(W(pF(t),t),t)}$$

where $P = max_t\{\lfloor\frac{F_{max}}{F(t)}\rfloor\}$
(F_{max} is the highest audible frequency) and

$$\phi_p(t) = \phi_p(0) + 2\pi\int_0^t W(pF(u),u)\ du$$

These equations are the "structured" version of equations 1 and 2. All these equations require approximately the same computation time.

3.3 Noise and Transients

SAS can faithfully reproduce a wide variety of sounds – as additive synthesis does – provided they are monophonic. However it can not produce noises or transients. Recent modeling techniques like SMS [Ser97] (*Spectral Modeling Synthesis*) and S+T+N [VM98] (*Sinusoids+Noise+Transients*) were proposed to extend the additive model. The SAS model can be extended to include noises, since every noise can be modeled as a filtered (or colored) white noise at a certain amplitude. The amplitude and color parameters exist also for noises and are sufficient to define any of them. White noise has a white color ($C = 1$), and every noise named after an analogy with a light spectrum matches this correspondence of terminology. On the other hand very short sounds like transients can not be represented in this spectral model.

4 Applications of the Model

SAS constitutes a solid base for investigating scientific and musical research on the notion of timbre. Applications of the model are numerous. A new sound synthesis language for musical composition has been implemented and should provide a way to validate and enrich the model. A pedagogical tool for early-learning electro-acoustic music is based on this model. It provides sound controls that are well-suited for young children because they are based on sound listening rather than signal synthesis.

4.1 Creation

The SAS parameters are closely related to the musical ones. Figure 2 shows the french *a* vowel sung on three notes. We can clearly see the dynamic and the melody of the song respectively in the A and F sound parameters. The arbitrary distinction between music and sound parameters simply disappears.

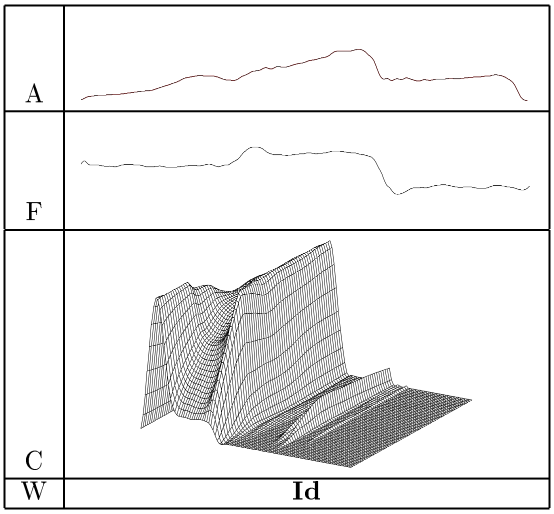

Figure 2: singing voice in the SAS model.

Most of musical transformations can be simply expressed as SAS parameter variations. Depending on the rate of these variations, composers can modify both the micro-structure and the macro-structure of musical pieces in a multi-scale composition [Vag98, DCM99]. When the variations are slow enough, they can be written on a score. This is the domain of writing. When they are too fast to be written, we enter the control (or interpretation) domain. Figure 3 gives a brief summary of the relations between musical terminology and SAS for these two domains. Multi-scale composition using SAS can be found in [DCM99].

	Writing (0-≈8Hz)	**Control** (≈8-20Hz)
Amplitude	Dynamic, Crescendo	Tremolo, Roughness
Frequency	Melody, Trill	Vibrato, Scintillating
Color	Orchestral Sonority	Spectral Envelope
Warping	Chords, Aggregates	Spectral Harmonicity

Figure 3: some relations between musical terminology and the four SAS parameters, for two ranges of variation rates.

To ease the composition with SAS, symbolic structures must be added on the top of the sub-symbolic sound model, like the hierarchic temporal organization of musical structures proposed by Balaban [BS93].

4.2 Education

Since SAS is based on perceptive and musical criteria, we think it is well-suited for computer-assisted early-learning music. *Dolabip* is a multi-field project whose objective is the creation of a meta-instrument to be used for early-learning electro-acoustic music. Practical experience in nursery school is lead by a musician, a teacher, a psychologist and a music teaching specialist.

This project is composed mainly of two parts. The first one is an hardware device consisting of potentiometers and buttons well-suited for manipulation by children. The second one is a software tool producing sounds according to the data sent by the device. The software tool allows the user to change the way the data get interpreted. The development teams of the software and hardware parts build tools that are necessary for experimenting the pedagogical program.

5 Conclusion

In this paper we have presented the Structured Additive Synthesis (SAS) model. This model represents sounds as temporal evolutions of parameters close to perception and musical terminology, thus favoring the unification of sound and music at a sub-symbolic level.

We are developing a sound synthesis language based on SAS, that has been used by Jean-Michel Rivet to produce interesting sounds in a piece of his. In order to use SAS for the whole compositional process, a hierarchic model must be designed on the top of SAS and incorporated in the language.

Again the SCRIME allows scientists and musicians to help each other. Composers lead researchers to understand their musical practices, while researchers try to provide composers with real tools for assistance for composition in harmony with their musical needs.

References

[BS93] M. Balaban and C. Samoun. Hierarchy, Time and Inheritance in Music Modeling. *Languages of design*, 1(3), 1993.

[DCM99] Myriam Desainte-Catherine and Sylvain Marchand. Vers un modèle pour unifier musique et son dans une composition multiéchelle. In *Proceedings of the Journées d'Informatique Musicale (JIM'99)*, pages 59–68, 1999.

[DCN98] M. Desainte-Catherine and Edgard Nicouleau. Segmentation et formalisation d'une oeuvre électroacoustique. Submitted to Musurgia, 1998.

[DCR98] M. Desainte-Catherine and J.M. Rivet. A la recherche de modèles poïétiques. Submitted to Musurgia, 1998.

[Del86] François Delalande. En l'absence de partition. *Analyse Musicale*, pages 54–58, 1986.

[MS99] S. Marchand and R. Strandh. InSpect and ReSpect: spectral modeling, analysis and real-time synthesis software tools for researchers and composers. In *Proceedings of the International Computer Music Conference (ICMC'99, Beijing)*, 1999.

[Ris86] J.C. Risset. Timbre et synthèse de sons. *Analyse musicale*, pages 9–19, 1986.

[Ser97] Xavier Serra. *Musical Signal Processing*, chapter Musical Sound Modeling with Sinusoids plus Noise, pages 91–122. Studies on New Music Research. Swets & Zeitlinger, Lisse, the Netherlands, 1997.

[Vag98] H. Vaggione. Transformations morphologiques (...). In *Proceedings of the Journées d'Informatique Musicale (JIM'98)*, page G1, 1998.

[VM98] Tony S. Verma and Teresa H.Y. Meng. Time Scale Modification Using a Sines+Transients+Noise Signal Model. In *Proceedings of the Digital Audio Effects Workshop (DAFX'98, Barcelona)*, pages 49–52, 1998.