INFLUENCE OF FREQUENCY DISTRIBUTION ON INTENSITY FLUCTUATIONS OF NOISE

Pierre HANNA

SCRIME - LaBRI
Université de Bordeaux 1
F-33405 Talence Cedex, France
hanna@labri.u-bordeaux.fr

Myriam DESAINTE-CATHERINE

SCRIME - LaBRI
Université de Bordeaux 1
F-33405 Talence Cedex, France
myriam@labri.u-bordeaux.fr

ABSTRACT

This article introduces mathematical parameters that generalize the frequency distribution of the statistical and spectral model. We show experiments that determine the influence of these parameters on the intensity fluctuations of the synthesized noise, accordingly to the theory developed by psycho-acoustic works. The main goal of these experiments is to be able to analyse and synthesize bands of noise with different spectral densities.

1. INTRODUCTION

Electro-acoustic composers often use real-world or environmental sounds in their works. The existing analysis/synthesis models (like for example [1]) can reproduce and transform pseudo-harmonic sounds. All these models can be used with non-deterministic sounds with the assumption that everything which can not be represented by sinusoids whose amplitudes and frequencies evolve slowly in time, is represented as filtered (or colored) white noise. Limits of these approaches can be clearly heard with applications on natural sounds. These sounds are not only very difficult to define mathematically but their perception is not completely understood. In this paper we try to give a general form for the frequency distribution of a statistical model and define mathematical parameters. The influence of these parameters on the intensity fluctuations are then compared with experiments.

2. PERCEPTION

2.1. Spectral density

Intensity fluctuations are stated as a perceptually relevant property of any sound. For example, an audible sine tone is perceived as a steady sound whereas it may have many cycles per second. This stability is described by the fluctuations of temporal envelope. This parameter is very useful to describe bands of noise [2]. One main application of the envelope fluctuations is the explanation of the ability for listeners to discriminate sounds with different spectral density [3]. Several psycho-acoustic experiments show that this ability can mainly be explained by two cues, one spectral and one temporal. The spectral cue corresponds to the pitch associated with the band of noise whereas the temporal one concerns the envelope fluctuations. These two parts are the central points of our paper.

2.2. Timbre and roughness

Moreover these fluctuations are stated as one of the dimensions of the timbre. Indeed they can be related to the complex concept of *roughness*. The roughness is defined by the perceptual effect from the fast beats produced by tones and it is perceived naturally [4]. However no theoretical relation exist (to our knowledge) between complex envelope fluctuations and *roughness*.

2.3. A musical parameter

This perceptual parameter also seem to be musically important. Schaeffer uses the concept of *grain* to describe musical differences between sounds [5]. It is defined between the rhythm and pitch domains, as many small irregularities at the *surface* of the sound.

In this paper, we focus on the dynamic envelope fluctuations of noise and, more generally, on the control of the spectral density. In the following, we use the word *envelope* to describe temporal envelope.

3. BACKGROUND

A few noise models have already been defined. The model we use in this paper is spectral and based on the thermal noise model.

3.1. Thermal noise

Thermal noises have been described in terms of a Fourier series [2]:

$$X(t) = \sum_{n=1}^{N} C_n \sin(\omega_n t + \Phi_n) \quad (1)$$

where N is the number of frequencies, n is an integer, ω_n are the pulsations which are equally spaced, C_n are random variable distributed according to a Rayleigh distribution and Φ_n are random variable uniformly distributed. This definition is the starting point of our work.

3.2. SMS model

The *Spectral Modeling Synthesis* (SMS) [1] separates the analyzed sound into deterministic and stochastic parts. The main assumption about the stochastic component is that it can be fully described by its spectral envelope. It is re-synthesized by inverse-Fourier transform. This transformation is mathematically described as a sum of a fixed number of sinusoids, whose amplitudes are spectral envelope values and whose frequencies are proportional to the ratio $\frac{R}{N}$. The phases are uniformly distributed. With this model, intensity fluctuations can only be controlled by modifying the spectral envelope from a temporal window to another: the intensity fluctuations are then related to the spectral envelope.

$$X(t) = \sum_{n=1}^{\frac{N}{2}} c_n \sin(2\pi n \frac{R}{N} t + \Phi_n) \tag{2}$$

3.3. Statistical model

A statistical approach have already been presented [6]. The sound is described as a sum of N sinusoids whose frequencies are random value (with a specific probability density function which is a parameter), phases are uniformly distributed and amplitude are fixed. With this model, the intensity fluctuations can't neither be directly controlled.

4. STATISTICAL MODEL FOR NOISES

In this paper, we consider sounds (sample rate R) in the spectral model as random processes X as in [6]. Each frequency component F_i is a random variable with fixed amplitude a_i and random phase Φ_i:

$$X_k = \sum_{i=0}^{N} a_i \sin(2\pi F_i \frac{k}{R} + \Phi_i) \tag{3}$$

where the frequencies F_i are distributed in a band whose width is W (Hz).

Previous psycho-acoustic works [2] show that phases contribute to intensity fluctuations. More precisely, a noise with minimal power fluctuations (low-noise noise [7]) can be synthesized with some fixed phase values [8]. Nevertheless in this paper, we focus on the distribution of the frequencies and their contribution to the intensity fluctuations.

4.1. Parameters

We propose to generate bands of noise in a statistical way. We consider a band of frequencies, divided in bins. No more than one frequency can be drawn in each bin. We define these parameters (see figure 1):

- N is the number of frequencies which are randomly chosen in a band.
- W is the bandwidth (Hz).
- M is the number of bins ($N \leq M$).
- L is the width of the uniform distribution ($L \leq \frac{W}{M}$), centered on each bin.

In the case of $N < M$, a bin is first drawn, then a frequency is determined within this bin.

By choosing M infinite, the distribution is uniform. By choosing $L = 0$, the signal generated follows the thermal noise model. Between these two extreme cases an infinity of distributions is proposed. In the scope of this paper, L is the same for all the bins. But we could easily extend the model and associate a distribution to each bin.

The spectral density is defined in this paper as $\frac{N}{W}$.

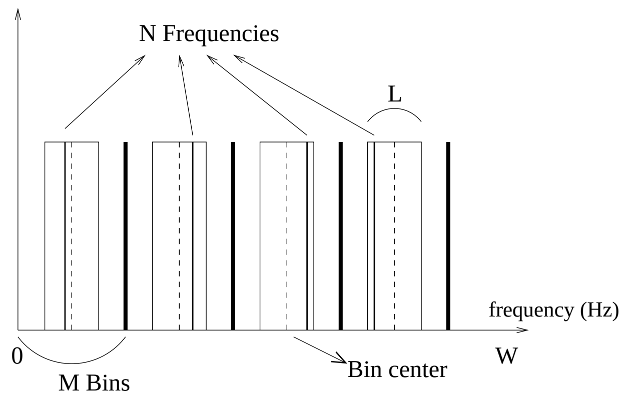

Figure 1: *Parameters of the model.*

4.2. Envelope Definition

From the above equation 3, the envelope $E(t)$ can be defined [2] as

$$X(t) = E(t)\cos(\bar{\omega}t + \phi(t)) \tag{4}$$

or using complex numbers :

$$E(t) = |\sum_{n=1}^{N} C_n e^{i(\omega_n + \phi_n)}| \tag{5}$$

Practically we can easily extract the envelope of any signal from an inverse-Fourier transform by removing the negative frequency components from the spectrum. This method was used during our experiments.

The auditory system cannot detect too fast changes of envelope. We use a model which take account into this property : the fluctuations are attenuated by a low-pass filter (equation 6).

$$h(t) = \frac{1}{\tau} e^{-\frac{t}{\tau}} \tag{6}$$

where τ is a time measure in seconds.

Psycho-acoustic experiments show that τ takes value between 2 and 3 milliseconds [2]. We have used 3 ms during our experiments. The envelope is precisely defined in [2]. The envelope power is defined by:

$$E^2(t) = \sum_{n=1}^{N} C_n^2 + 2\sum_{n=1}^{N}\sum_{n'=1}^{n-1} C_n C_{n'} \cos[(\omega_n - \omega_{n'})t + \phi_n - \phi_{n'}] \tag{7}$$

4.3. Envelope fluctuations

Hartmann [3] made psycho-acoustic experiments to understand that the ability to discriminate spectral density is related primarily to the envelope fluctuations . These envelope fluctuations are modeled in [2] so that the variance of the envelope power (normalized by the square of the mean) is related to N, M, W and τ (equation 8). In his studies, our parameter L is defined as $\frac{W}{M}$ and is fixed. In the future we could extend this model to take into account other values of the parameter L, so that another more complex equation could be defined.

$$\frac{v(E^2)}{\overline{E^4}} = \frac{(1-\frac{1}{N})}{2\pi W\tau}\frac{2}{1-\frac{1}{M}}\Bigg\{\tan^{-1}(2\pi W\tau) - \tan^{-1}(\frac{2\pi W\tau}{M}) + \frac{1}{4\pi W\tau}\Big[M\ln\big[1+(\frac{2\pi W\tau}{M})^2\big] - \ln[1+(2\pi W\tau)^2]\Big]\Bigg\} \quad (8)$$

5. EXPERIMENTS

From the model developed by Hartmann, we have decided to measure the influence of the different parameters on the envelope fluctuations by using the variance. Several experiments have been performed. Two different aspects are tested : the periodicity and the amplitude of the envelope fluctuations.

5.1. Periodicity of envelope fluctuations

From the definition of the envelope power (equation 7), one can see that it is periodic. This envelope periodicity is related to the frequencies differences. These differences can be estimated from the parameters (M, N, L, W).
Experiments show that we can hear either a pitch or beats. This pitch or beat is independent of the spectral envelope: a sound generated with $W = 18000$ Hz, centered at 11000 Hz, gives one pitch at 2000 Hz but also a low-frequency pitch or beats corresponding to the envelope fluctuations, in the case of particular spectral density. This implies that by modifying our parameters, we can synthesize bands of noise with beats or with pitch.
Our experiments show that if the number of components N increases, the periodicity decreases until low frequency beats appear. Moreover if the number of bins M or/and the width L of the probability density function increase whereas W is constant, the periodicity is no more fixed and its standard deviation increases.

The figure 2 shows the influence of the parameter L. We have used the autocorrelation function. It is known that this function is related to the periodicity perception of any sound. For example, in the case of the fixed frequencies ($L = 0$ and $M = N$), which corresponds to the thermal noise model, the autocorrelation perfectly shows the periodicity of the synthesized noise. We have measured the ratio between the second maximum and the first point of the autocorrelation function (which is the total energy of the signal and the maximum), as a function of the width L of the probability density function of the bin. This ratio is supposed to be related to the pitch perception. The figure 2 shows that increasing L decreases the ratio. That's why the periodicity (beats) is not as clearly heard as in the case of $N = M$ and $L = 0$, which corresponds to the thermal noise model and the SMS stochastic model.

5.2. Amplitude of envelope fluctuations

We have done experiments to verify the influence of the number of components N, the number of bins M, and the width L of the probability density function. In particular, it appears that :

- Increasing the number of components N increases the envelope fluctuations (this can be shown from equation 8 and can be seen figure 3).

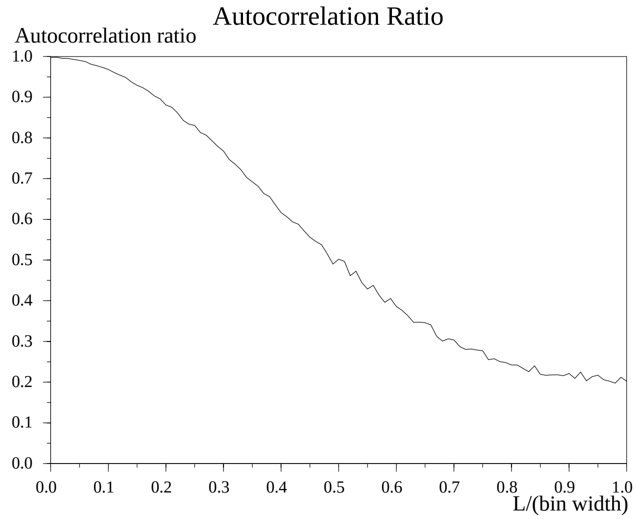

Figure 2: *Influence of the width L: autocorrelation ratio as a function of L (200 realizations of 32768 samples of a signal with $W = 2000$ Hz and $N = 10$).*

- Increasing the width L of the probability density function increases the envelope fluctuations (this can be seen figure 3.
- Increasing the number of bins M increases the envelope fluctuations (this can be seen figure 3 and shown from equation 8).

The effect of the bandwidth W of the noise is related to equation 8. It is obvious that increasing the bandwidth W (the number N of components being unchanged) results in decreasing spectral density so that envelope fluctuations may decrease or pitch may be heard due to the spectral differences (see figures 6, 4 and 5).

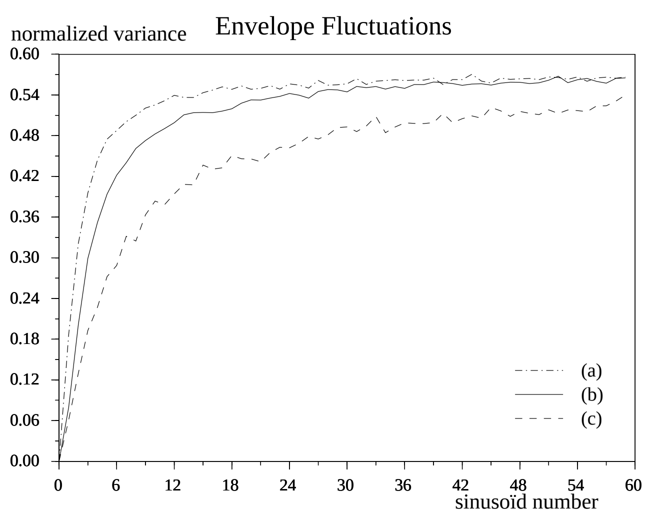

Figure 3: *Mean of the normalized variance of the power envelope as a function of the number of components N (300 realizations of 32768 samples of a signal and $W = 400$ Hz) when (a) M is infinite (b) $M = N$ and $L = \frac{W}{M}$ (c) $M = N$ and $L = 0$.*

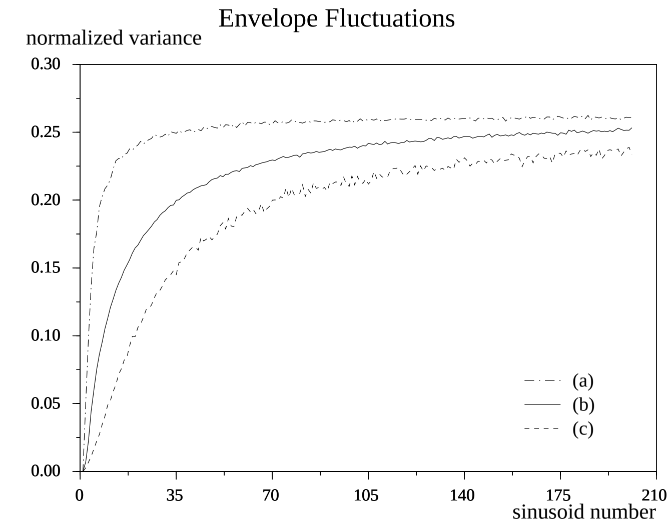

Figure 4: *Mean of the normalized variance of the power envelope as a function of the number of components N (1000 realizations of 65536 samples of a signal and $W = 2000$ Hz) when (a) M is infinite (b) $M = N$ and $L = \frac{W}{M}$ (c) $M = N$ and $L = 0$.*

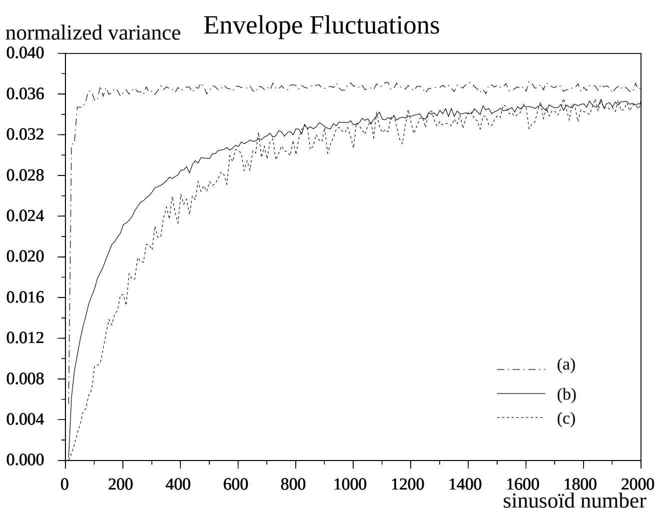

Figure 5: *Mean of the normalized variance of the power envelope as a function of the number of components N (200 realizations of 32768 samples of a signal and $W = 20000$ Hz) when (a) M is infinite (b) $M = N$ and $L = \frac{W}{M}$ (c) $M = N$ and $L = 0$.*

6. APPLICATIONS

Many applications may be based on these works about the envelope fluctuations and the influence of the statistical parameters we have defined. The first one is an analysis/synthesis model for noisy sounds.

6.1. Toward a statistical model

These studies may help to define a statistical model which could permit to control independently the spectral and the temporal envelope. Moreover the variance of the envelope fluctuations can be used as a measure of the spectral density in order to decide the number of sinusoids and the statistical parameters to draw them, and to synthesize a band of noise.

6.2. Windowing

This statistical approach may lead to a synthesis model. That's why many signals can be synthesized in successive short-time windows. The length of the windows are an interesting parameter in the case of periodic audible envelope fluctuations. A short window (less than twice the period of the envelope) *breaks* the periodicity and prevents the user from hearing beats or even pitch.

This technique is implicitly used in the inverse-Fourier transform (SMS model). For a $2k$ samples long window, this can be represented as the sum of k fixed sinusoids (at the center of the bins).This corresponds to the parameters $N = M = k$ and $L = 0$. One can demonstrate that the envelope fluctuations corresponding to the minimal frequency difference ($\frac{R}{k}$) can not be perceived because of the length of the window.

Furthermore the statistical approach seems interesting concerning the efficiency. To produce a noise with the same envelope fluctuations, or with no audible beat, one should choose L and M to reduce the number of components.

7. FUTURE WORK

7.1. Noisy sounds analysis and synthesis

Applications of the statistical parameters on the intensity fluctuations seem interesting concerning the synthesis of sounds like for example engine-, car-, traffic- or waterfall noises. Moreover many experiments may be performed on the analysis and the synthesis of consonants and whispered voices.

7.2. Phase influence

In our works, we have only considered the frequency distribution and we use an uniform distribution to draw the phases of the spectral components (as in the SMS model). Nevertheless we have begun to study the influence of the width of the probability density function of the phase. We can see on the figure 6 that increasing this width decreases the envelope fluctuations. Noises synthesized with narrow probability density function for the phase are described as impulsive noises.

Moreover some envelope fluctuations can not be synthesized with a uniform distribution of the phases. We will study the limit between a noise with large envelope fluctuations and transients and try to give a precise definition related to our future model.

As seen previously, very small values of envelope fluctuations are very useful in masking sine signals [9]. These small values can only be obtained by choosing phases. That's also why this parameter is one of the main point of our future work.

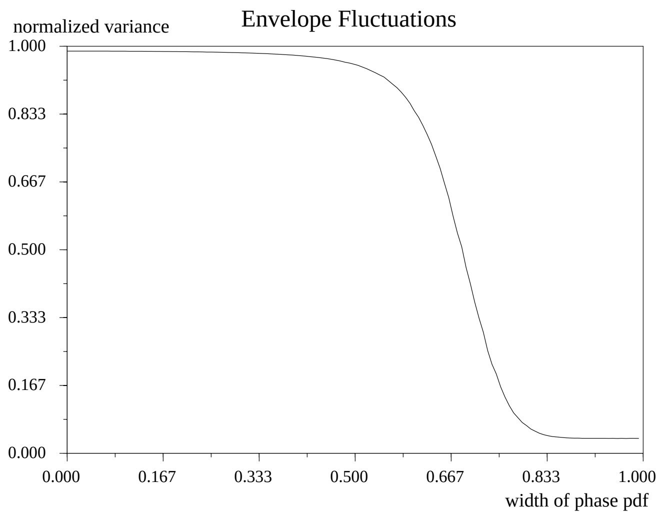

Figure 6: *Mean of the normalized variance of the power envelope as a function of the width of phase PDF (*1000 *realizations of* 32768 *samples of a signal and* M *is infinite).*

8. ACKNOWLEDGMENTS

This research was carried out in the context of the SCRIME[1] project which is funded by the DMDTS of the French Culture Ministry, the Aquitaine Regional Council, the General Council of the Gironde Department and IDDAC of the Gironde Department. SCRIME project is the result of a cooperation convention between the Conservatoire National de Rgion of Bordeaux, ENSEIRB (school of electronic and computer scientist engineers) and the University of Sciences of Bordeaux. It is composed of electroacoustic music composers and scientific researchers. It is managed by the LaBRI (laboratory of computer science of Bordeaux). Its main missions are research and creation, diffusion and pedagogy thus extending its influence.

9. REFERENCES

[1] X. Serra and J. Smith, "Spectral modeling synthesis: a sound analysis/synthesis system based on a deterministic plus stochastic decomposition," *Computer Music Journal*, vol. 14, no. 4, pp. 12–24, 1990.

[2] W.M. Hartmann, *Signals, Sound, and Sensation*, Modern Acoustics and Signal Processing AIP Press, 1997.

[3] W.M. Hartmann, S. McAdams, A. Gerzso, and P. Boulez, "Discrimination of spectral density," *Journal of Acoustical Society of America*, vol. 79, no. 6, pp. 1915–1925, 1986.

[4] D. Pressnitzer, *Perception de rugosité psychoacoustique: d'un attribut élémentaire de l'audition l'écoute musicale*, Ph.D Thesis, Université Paris VI, 1998.

[5] P. Schaeffer, *Traité des objets musicaux*, Seuil, 1966.

[6] Desainte-Catherine M. and P. Hanna, "Statistical approach for sound modeling," *Proceedings of the Digital Audio Effects Workshop (DAFX'00, Verona)*, pp. 91–96, 2000.

[7] J. Pumplin, "Low-noise noise," *Journal of Acoustical Society of America*, vol. 78, no. 1, pp. 100–104, 1985.

[8] W.M. Hartmann and J. Pumplin, "Periodic signals with minimal power fluctuations," *Journal of Acoustical Society of America*, vol. 90, no. 4, pp. 1986–1999, 1991.

[9] W.M. Hartmann and J. Pumplin, "Noise power fluctuations and the masking of sine signals," *Journal of Acoustical Society of America*, vol. 83, no. 6, pp. 2277–2289, 1988.

[1] Studio de Cration et de Recherche en Informatique et Musique lectroacoustique, www.scrime.u-bordeaux.fr