ADAPTING THE OVERLAP-ADD METHOD TO THE SYNTHESIS OF NOISE

Pierre HANNA

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

Myriam DESAINTE-CATHERINE

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

ABSTRACT

Spectral synthesis techniques often use the OverLap-Add method. But in the case of noise synthesis, both experiments and theory show that this method introduces intensity fluctuations which imply audible artifacts. We propose here new methods to avoid these variations. The first one consists in multiplying the resulting signal by another weighting window to compensate dynamic fluctuations. The second one defines a new OLA weighting window. The third one concerns only noises synthesized with sinusoidal components and uses *time-shifting* to cancel artifacts.

1. INTRODUCTION

Many noise synthesis methods use the OverLap-Add (OLA) technique. Noises can be produced temporally by randomly drawing samples according to a constant distribution (uniform, normal, ...). Samples are then filtered. However all real-time implementation needs to synthesize sounds frame by frame. That's why OLA technique is often used to avoid clicks.

Noises can also be generated using sinusoidal components. For example, in SMS[1], the application of inverse-Fourier transform to synthesize the stochastic part implies a frame-by-frame synthesis, and therefore the application of the OLA technique.

These existing methods can be used to synthesize noises without previously having performed an analysis process. In this case, the resulting temporal signal does not taper to 0 at the boundaries of each frame.This can be the cases at times, even with an analysis process, if sounds are transformed. This is the reason why these methods utilize the OLA technique.

2. OLA TECHNIQUE IN THE SYNTHESIS OF NOISE

2.1. Definition

OLA technique involves in synthesizing overlapping temporal frames s_l of N samples (see figure 1). Let L be the number of frames:

$$x[n] = \sum_{l=0}^{L-1} x_l[n - lH] \tag{1}$$

where H is the hop size (or the time advance) and l is the frame number.

Each temporal frame s_l is multiplied by a weighting window w:

$$\begin{cases} x_l[n] = & s_l[n]w[n] \quad n \in [0, 1, .., N-1] \\ x_l[n] = & 0 \quad \forall n < 0, \quad \forall n \geq N \end{cases} \tag{2}$$

This window must satisfy the condition that the sum of all the weights of the different overlapped windows must be equal to 1:

$$\sum_l w[n - lH] = 1 \tag{3}$$

If the overlapped and added analysis windows do not sum to unity, then the output sound is amplitude modulated by a periodic signal.

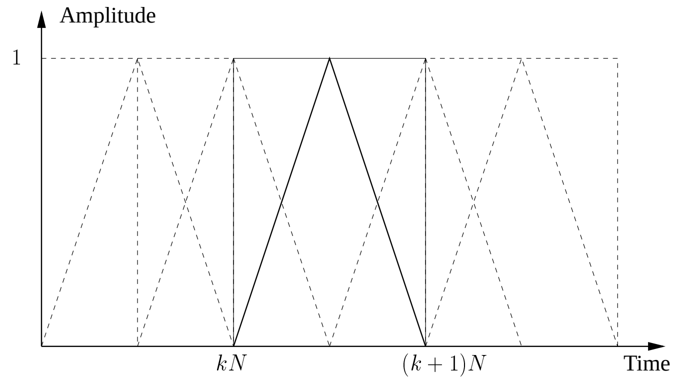

Figure 1: *Overlap-Add with Bartlett window ($H = \frac{N}{2}$).*

2.2. Limits

The use of the OLA technique introduces some auditive drawbacks if no analysis is performed.

Rectangular OLA: In the case when w is rectangular ($\forall n = [0; N-1], w[n] = 1$), some artifacts can be heard. The main artifact is due to the periodicity implied by the overlap-add process. Many clicks can also be heard in the case of sounds synthesized with very few sinusoidal components (harmonic or quasi-harmonic sounds).

Random length windows: To eliminate these artifacts, one can randomize the synthesis window length. Our experiments using this approach lead to sounds with no audible periodicity. But the clicks still remain due to the transitions between successive temporal frames.

Windowed OLA: During our experiments we also tried other type of weighting windows (Hann, Bartlett, ...). Then the clicks disappear but irrespective of the window type, new artifacts can be heard with noise-like sounds, which are described in the next section.

2.3. Experiments

2.3.1. Synthesis after analysis process

Before synthesizing a noise, some models (for example [1]) extract parameters from sounds. This analysis is often performed with overlap-add technique. For example, if a white noise is analyzed, each temporal frame s_l is not a stationary white noise: variance of this random signal periodically varies in time because of the window multiplication. This is shown with experiences in table 1. And as the variance is directly linked to the perceived intensity of the sound, it provokes audible dynamic fluctuations.

Moreover doing an OLA analysis implies that each sound segment are not *independent* (see figure 2). That's why synthesis always perfectly works in this case. But it is obvious that this operation is useless if no modification (time stretching, pitch shifting, ...) of the analyzed sound is performed before synthesis. These transformations may lead to make sound segments independent, and thus to audible artifacts (see section 2.4).

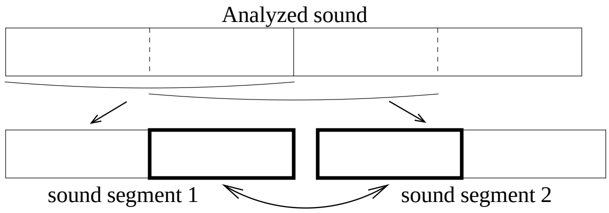

Figure 2: *Analysis: two sound segments are not independent.*

2.3.2. Synthesis without analysis process

If the OLA method is used without analysis process, some artifacts appear. For example, if two white noises are synthesized independently by random (uniform, normal, ...) draw of samples, and then are added with *classical* window (Hann, Bartlett, ...), the resulting signal is white noise with different statistical properties. The variance is not constant. Some experimental results are shown in tab 1.

(α, β)	var(s_1)	var(αs_1)	var(s_2)	var($\alpha s_1 + \beta s_2$)
$(0.5, 0.5)$	8.312	2.078	8.361	4.160
$(0.4, 0.6)$	8.297	1.328	8.349	4.342
$(0.3, 0.7)$	8.304	0.747	8.340	4.812
$(0.2, 0.8)$	8.350	0.334	8.330	5.677
$(0.1, 0.9)$	8.337	0.083	8.411	6.899
$(0.0, 1.0)$	8.323	0.000	8.341	8.341

Table 1: *Evolution of the variance of the addition of two random signals (normal distribution, Bartlett window,* 100000 *realizations).*

An example of such sounds with Bartlett window can be seen in figure 3. This example shows the periodic fluctuations of intensity. These effects can be explained perfectly as in the following by the study of statistical properties of the sounds.

2.4. Theory

We have seen in section 2 that two temporal frames are multiplied by a weighting window before being added. Any window type is, it must satisfy equation 3. By denoting x' the synthesized signal, one can write using equation 2:

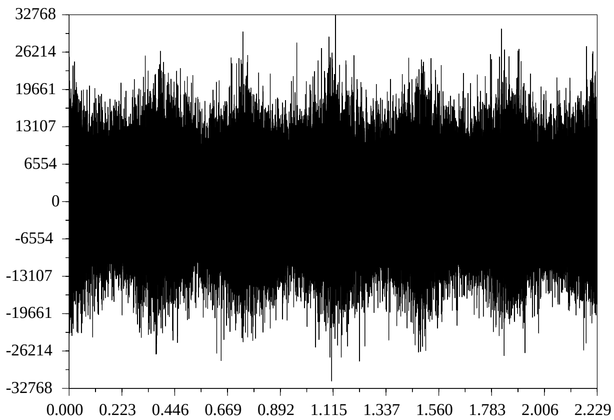

Figure 3: *Example of white noise synthesized with OLA.*

$$x'[n] = \sum_{l=0}^{L-1} s_l(n - lH)w(n - lH) \tag{4}$$

In the case of white noise, s_l ($l \in \{0, .., L-1\}$) are realizations of the same random variable X. They have the same statistical properties and we want the resulting signal $x'[n]$ to have also the same properties. Let X' be the random variable associated with $x'[n]$. Then the mean of X' stays the same but the second order moment vary. Let $V(X')$ be the variance of X':

$$V(X') = V(\sum_{l=0}^{L-1} s_l(n - lH)w(n - lH)) \tag{5}$$

If an analysis is performed but no transformation, s_l are not independent (see section 2.3.1) and we verify $V(X') = V(X)$. We consider here the case when there is no analysis process. Since s_l are then independent:

$$V(X') = \sum_{l=0}^{L-1} w^2(n - lH)V(s_l(n - lH)) \tag{6}$$

Since s_l are realizations of the same random variable X, one can write:

$$V(X') = V(X) \sum_{l=0}^{L-1} w^2(n - lH) \tag{7}$$

The definition of $w(n)$ is $\sum_l w(n - lH) = 1$, which leads to $\sum_l w^2(n - lH) = f(n)$: this sum is not constant. That's why $V(X')$ will vary periodically in time.

Assuming that $V(X')$ and $V(X)$ are equal, this calculus leads to the condition about w, $\forall n \in \mathbb{N}$:

$$\sum_{l=0}^{L-1} w^2(n - lH) = 1 \tag{8}$$

For example, in the case of $H = \frac{N}{2}$ and when weighting window w is triangular ($w(n - lH) \in [0, 1]$):

$$V(X') = V(X)(1 - \frac{4n}{N-1} + \frac{8n^2}{(N-1)^2}) \qquad (9)$$

where $n - k\frac{N}{2} \in \{0, .., \frac{N-1}{2}\}$ and $k \in \mathbb{N}$.

Figure 4 shows the variations of the variance as a function of time with $N = 16384$ and $H = \frac{N}{2}$. Variations in case (b) perfectly correspond to the intensity fluctuations of the figure 3.

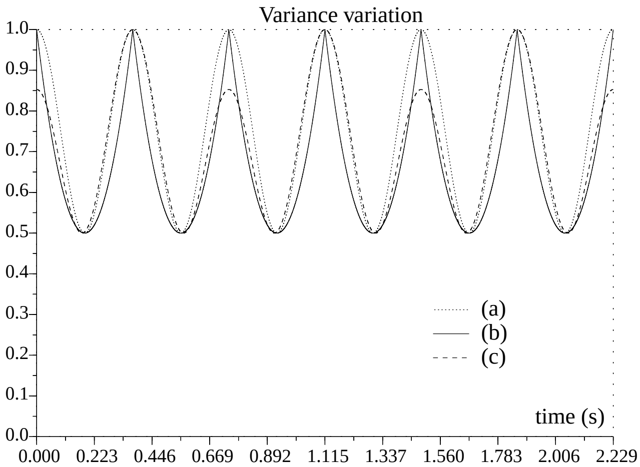

Figure 4: *Variations of the variance of the signal with (a) Hann window (b) Bartlett (triangular) window (c) Hamming window.*

3. PROPOSED METHODS

We now propose three methods to avoid these intensity fluctuations. The two first ones can be used with all type of noise synthesis whereas the third one only concerns noise synthesized with sinusoidal components.

3.1. Amplitude compensation

As we know $V(kx) = k^2 x$, it is possible to compensate the intensity fluctuations by multiplying the temporal signal by another weighting window inversely proportional to the theoretical variations (equation 9). For example, in the Bartlett window case and with $H = \frac{N}{2}$, one can apply:

$$x[n] = \sum_l \frac{1}{\sqrt{1 - 2w[n-lH] + 2w[n-lH]^2}} x'[n] \qquad (10)$$

Drawbacks of this method may appear during the synthesis of sounds composed of a little number of components (weak spectral density). The intensity fluctuations may be audible for each partial. These fluctuations are represented in figure 6.

3.2. Sinusoidal window

In section 2.4, we have demonstrated the condition (equation 8) that the weighting window must verify in order to preserve the statistical properties of the synthesized signal. But the OLA technique is based on the property described in section 2 by equation 3.

One can easily show that it is impossible to define a window w which verifies:

$$\begin{cases} w(n) + w(n + \frac{N}{2}) & = & K_1 \\ w^2(n) + w^2(n + \frac{N}{2}) & = & K_2 \end{cases} \qquad (11)$$

with K_1 and K_2 constant values.

In this section we propose a weighting window which preserves the statistical properties in the case of $H = \frac{N}{2}$. This window is simply defined with a sinusoid function and is represented in figure 5:

$$w[n] = \sin(\frac{\pi n}{N-1}) \quad \forall n \in [0, N-1] \qquad (12)$$

Indeed this window is symmetric ($w(n) = w(N - 1 - n)$) and verifies the equation 8:

$$\begin{aligned} w^2[n] + w^2[n + \tfrac{N}{2}] &= \sin^2(\tfrac{\pi n}{N-1}) + \sin^2(\tfrac{\pi n}{N-1} + \tfrac{\pi}{2}) \\ &= sin^2(\tfrac{\pi n}{N-1}) + \cos^2(\tfrac{\pi n}{N-1}) \\ &= 1 \end{aligned} \qquad (13)$$

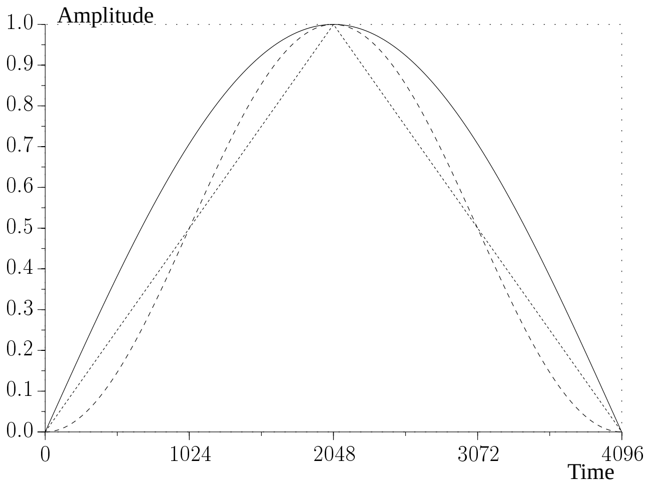

Figure 5: *Representation of the sinusoidal (plain line), Bartlett (dotted lines) and Hann (dashed lines) windows.*

But drawbacks of this method are obviously the same as the *amplitude compensation* method (see section 3.1). Using this sinusoidal window will modulate the amplitude of each sinusoidal components. But this effect can only be heard with sounds synthesized with a little number of components (weak spectral density). Let us calculate this amplitude modulation:

$$w(n) + w(n + \tfrac{N}{2}) \quad = \quad \sqrt{2}\sin(\tfrac{\pi n}{N-1} + \tfrac{\pi}{4}) \qquad (14)$$

This calculus shows that this modulation is less important than the one introduced by the *amplitude compensation* method. (see figure 6). That's why this method appears to give better quality in synthesized sounds. Moreover this method is faster. Only the multiplication by the weighting window is computed whereas this calculation *and* the compensation must be calculated in the other case.

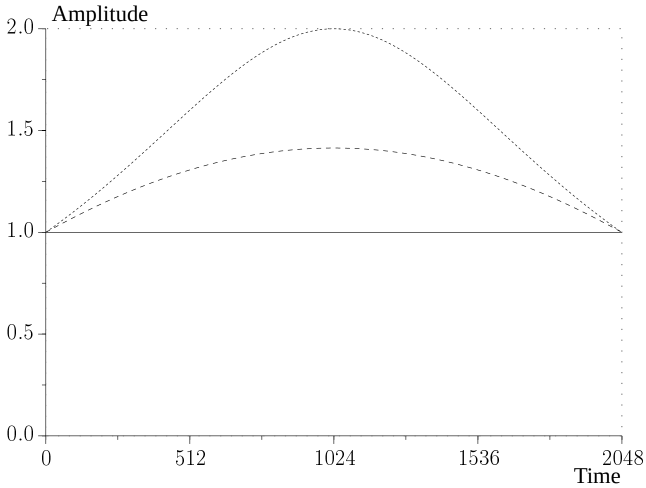

Figure 6: *amplitude modulations induced by sinusoidal window method (dashed lines) and amplitude compensation method (dotted lines).*

3.3. Sinusoidal component offsets

This method concerns the synthesis of noise by summing M sinusoidal components (denoted s^m) with random phase [2]. It consists in shifting the start of each sinusoidal component in each frame in order to distribute the intensity variations introduced by the weighting windows. Thus each component starting time (d_m with $m \in \{0, .., M-1\}$) is set to different values before being multiplied by the weighting window. A general algorithm 1 is proposed. The resulting signal x' can be written as:

$$x'(n) = \sum_{l=0}^{L-1} \sum_{m=0}^{M-1} s_l^m(n - lH - d_m) w(n - lH - d_m) \quad (15)$$

One can calculate the variance of this random signal:

$$V(X') = V(\sum_{m=0}^{M-1} t^m(n - d_m)) \quad (16)$$

where t^m are sinusoidal components. By choosing d_m equally spaced over the half window, one can show that this sum of sinusoids leads to noise with constant statistical properties.

There are many ways to determine the offset values. For example, they can be randomly drawn according to a uniform distribution. But this method may lead to artifacts because many offset may have the same value, which introduces variance fluctuations as previously seen in section 2.4. To avoid these probabilities, we prefer choosing these values by dividing the half window in M bins. So each offset d_m equals $m\frac{N}{2M}$. An example of such component offset is illustrated on figure 7.

Data : List of partials, Output buffer
begin
  for *the number of partials* **do**
    draw an offset *off*;
    synthesize the current partial;
    multiplication by weighting window;
    offset output buffer with *off*;
    add partial buffer to output buffer;
  end
end

Algorithm 1: *Algorithm for component offset*

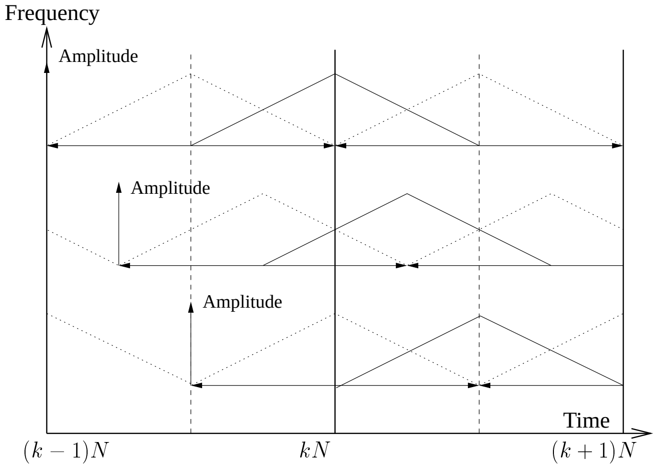

Figure 7: *Example of* 3 *sinusoidal component offsets.*

4. CONCLUSION

We have proposed methods to synthesize noisy sounds with OLA technique without dynamic artifacts. These techniques are useful in the case of dense [3] sounds and if no analysis has been performed.

We have implemented the OLA method with sinusoidal component offsets. The offset values were determined with random values chosen in bins (figure 7). We have also experimented sinusoidal window. In both cases, the synthesized sound quality are impressively improved and no more artifacts can be heard. One can also observe these results with intensity fluctuations calculus.

The offset method decreases the synthesis speed because all sinusoidal components have to be multiplied by weighting windows. One multiplication is added per partial, whereas using sinusoidal window keeps running time constant.
In the future we want to try to improve the offset method using the FFT^{-1} synthesis[4] which may increase the speed of synthesis.

5. 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] P. Hanna and M. Desainte-Catherine, “Influence of frequency distribution on intensity fluctuations of noise,” *Proceedings of the Digital Audio Effects Workshop (DAFX’01, Limerick, Ireland)*, pp. 125–129, 2001.

[4] X. Rodet and P. Depalle, “A new additive synthesis method using inverse fourier transform and spectral envelopes,” *Proceedings of International Computer Music Conference (ICMC’92, San Jose, California)*, pp. 410–411, 1992.