BELGRADE 89

* 13th INTERNATIONAL CONGRESS ON ACOUSTICS *
* YUGOSLAVIA * 1989 *

WAVEFORM TRANSFORMATION TECHNIQUES FOR GRANULAR SYNTHESIS

G. De Poli^, A. Piccialli^^

^ C.S.C.-DEI, University of Padova, Via Gradenigo 6a, 35131 Padova, Italy
^^Dept. of Physics, University of Naples, Mostra d' Oltremare Pad. 20, 80100 Napoli, Italy

Past work has indicated the powerful musical use of granular synthesis[1-9]. This method, suggested by Gabor and then developed by Roads and others, uses a number of short duration acoustic elements, called *grains*, to form musical events. However, without an interpretation from the point of view of digital signal processing theory, this approach do not allow to reconstruct a given arbitrary signal. The recent method of signal analysis by *wavelets* (functions limited both in time and frequency domain) justifies granular synthesis and permits to reconstruct signals by summation of opportune grains. Moreover there are other DSP approaches in order to extract specific information (e.g. pitch, formant structure etc.). We can adapt the grains in order to synthesize sounds using these representations.

These interpretations allows then a more musical use of granular synthesis. In fact we can identify a grain as the impulse response of a filter. In general case we have a parallel structure of sequences of grains. The synthesized signal is given by

$$s(n)= \Sigma_i \Sigma_j g_{ij}(n-d_{ij})$$

where $g_{ij}(n)$ is the j-th grain of the i-th parallel sequence and d_{ij} is its temporal location. The shape of each grain can be characterized by the duration, the waveform and the amplitude envelope. The temporal location of the grain can be choosen according various criteria. The use of fixed intervals often presents the problem of intermodulation and lack of phase continuity. Various kind of solutions has been proposed: phase shifting in the frequency domain, euristic variations, algorithmic variations (eg. maximization of similarity functions), choice during the analysis. Generally they are computationally expensive and not easy to use. Another approach is to use the grain pitch sinchronously[9]. The Short Time Fourier Transform and Wavelet Transform can be represented by this structure with suitable choice of grain waveform and of d_{ij}. Moreover, the pitch synchronous approach imply a source-filter underlying model. Thus it is more specialized, but also more powerful. In this work we propose and discuss some techniques in time domain to generate and modify the grain shape.

Basic transformations

In order to obtain an efficient and quite flexible system of granular synthesis, we propose to start from a few prototype grains (impulse responses). Then we define some general transformations of one or more waveforms. These transformations can be applied both to prototype and to transformed waveforms.

Let p(n) and q(n) be two waveforms and P(f) and Q(f) their Fourier Tranforms. We call *weighted additive transformation* the following expression

$$r(n) = A[p(n),q(n),a1,a2] = a1\cdot p(n) + a2\cdot q(n)$$

where a1,a2 are constants, that are fixed for the entire duration of the grain, but that can change in different grains. The spectrum of r(n) will be given by $R(f) = a1\cdot P(f) + a2\cdot Q(f)$. It can be noticed that if p(n) and q(n) are zero phase, r(n) will be also zero phase and the addition of the spectra is a real operation. Thus disturbing interferences and cancellations in the spectrum are avoided. We employed this kind of transformation to modify various parts of the shape of a formant. For example the added grain can control the skirt of transition band. An other application is designing a grain for each different spectrum region and then combining the resulting weveforms using this transformation. In such a way we obtain a simple control of complex grains and an easier musical use. Our experience suggests that zero phase prototype waveforms are to be preferred.

The *multiplicative transformation* is given by

$$r(n) = M[p(n),q(n)] = p(n) \cdot q(n)$$

The spectrum of r(n) is the convolution of P(f) with Q(f). We now examine some important particular cases. Let p(n) be the impulse response of a low-pass prototype and q(n) the complex exponential function, then r(n) will be a bandpass version of p(n). Considering the real signal and a generic phase ϕ of the exponential, we obtain

$$r(n) = p(n) \cdot \cos(2\pi f_c n+\phi)$$

where f_c is the normalized center frequency of the bandpass region. This operation corresponds to a frequency translation of the prototype spectrum and allows an accurate control of formant frequency. Another very important application is when q(n) is a constant (generally different for each grain). This allows, for example, the dynamic control of the amplitude of formants. Other choices of q(n) are possible. But it is advisable to use waveforms where the energy is localized in few regions of the spectrum, in order to have an intuitive idea of the resulting spectrum. In any case they must not be high pass in frequency domain, to avoid foldover.

The *time scale transformation* is defined by

$$r(n) = S[p(n),q(n)] = b \cdot p(b\cdot n)$$

The spectrum is given by $R(f)=P(f/b)$. Therefore a reduction of the temporal scale determines an inverse widening of the frequency scale. An important application is to control the bandwidth of a low pass prototype. In this case b will be equal to the ratio between the desired bandwidth and the original bandwidth of the prototype. If the prototype waveform is stored in a table, the time scale change can be obtained by reading the table scanned with a step value b. Thus it is advisable to have more detailed tables containing various intermendiate values. Note that often a more efficient implementation can be obtained, grouping the products of waveform for constants, that appear in the employed transformation. When the multiplicative (by a cosine) transformation is applied to a time scaled prototype, we have an easy dynamic control of the bandwidth and central frequency of formants. It can be noticed that if the time scaled waveform corresponds to a bandpass signal, we modify both the bandwidth and the center frequency. Their ratio remains constant. So we have a constant-Q synthesis, analogous to the one used in wavelet synthesis, that is particularly attractive, because is close to the hearing model.

Figure 1, on the left, shows the spectrum of a bandpass grain that is obtained from a maximally flat low pass prototype (bandwidth=0.0075) frequency translated by multiplicative tranformation (f_c=0.1). This bandpass signal, time scaled with b=3, is shown on the right.

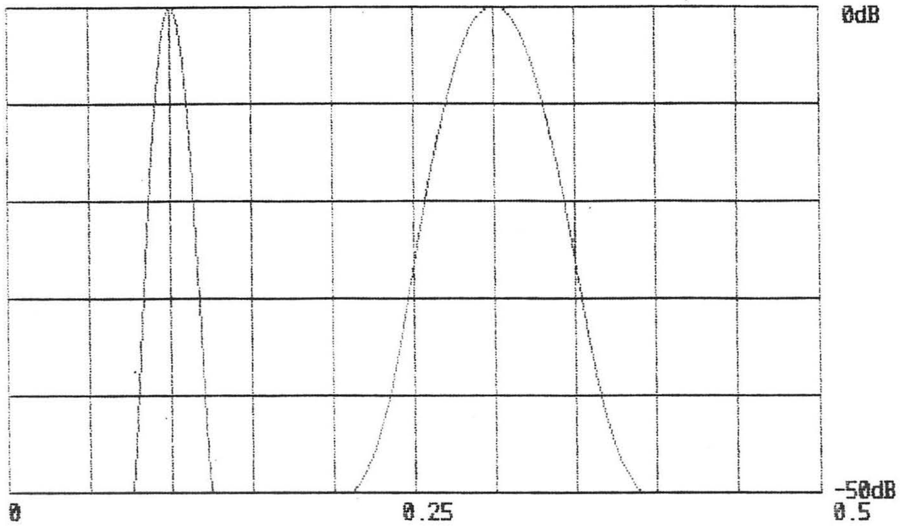

Figure 1

Non linear tranformations

Besides these tranformations that have a precise interpretation in digital signal processing theory, we propose some useful non linear tranformations. *Time scale distortion* and *non linear amplitude distortion* are given by

$$r(n) = N[p(n),q(n),c] = p[\ c\cdot q(n)\ ]$$

where in the first case q(n) is the function expressing the time scale distortion of p(n) and in the second case p(n) is the non linear amplitude distorting function applied to q(n). The factor c acts as a modulation index and allows complex variations of the grain spectrum. Small distortions of this type allow an efficient control of fine details of the sound spectrum. Note that in the implementation of this transformation, normally the functions are stored in tables and linear interpolation is used to compute values at non integer abscissae.

Among the variuos possibilities we found particularly interesting the distortion of the time scale of the shifting cosine before its use in multiplicative tranformation. It results a phase modulated sinusoidal signal, enveloped by a grain waveform

$$r(n) = p(n) \cdot \cos\,[2\pi f_c n + c\cdot\phi(n)]$$

The control of the modulation index and of the time scale of modulating function allows a continous and smooth control of the spectrum shape of the generated signal. Figure 2 shows the spectrum obtained when the same prototype, employed in figure 1, is used both as phase modulating signal and grain envelope.

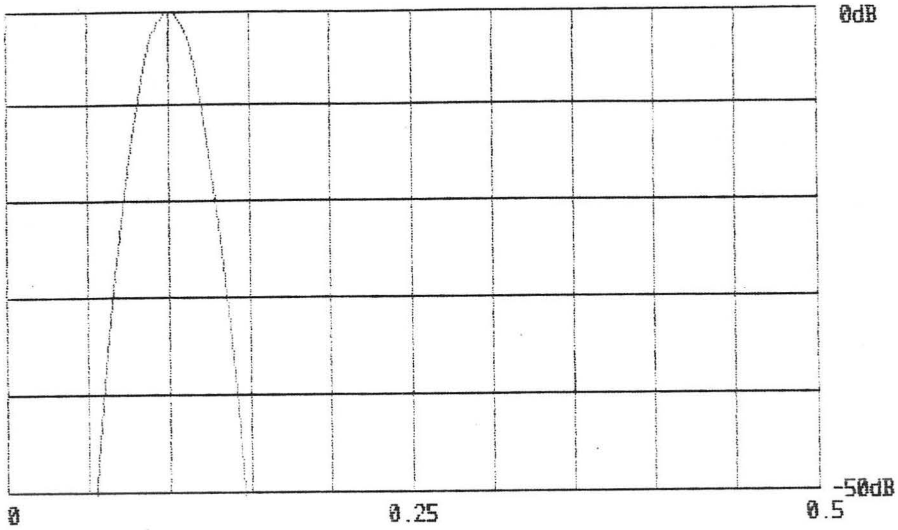

Figure 2

Conclusion

In this paper some principles and realizations of granular synthesis systems, using prototype waveforms, were investigated. When these waveforms satisfy some well known conditions, granular synthesis coincides with wavelet synthesis. Thus we have a rigorous analysis framework, that allows the reproduction of any natural sound and augments the power in musical applications.

However, more compact sound representations are required in musical praxis; the control parameters must be pertinent to the physical phenomena of sound production and to musical perception. Such representations can be realized in granular synthesis, using the wave transform techniques above described. From this point of view granular synthesis could constitute an unifying approach of variuos theoretical and practical experiences, developed in digital signal processing and musical sound synthesis.

References

[1] Gabor D., Theory of communication, J. of IEE, 93, 429-441, 1946

[2] Roads C., Automated granular synthesis of sounds, Comp. Music J., 2(2), 61-62, 1978

[3] Kaegi W., Tempelaars S., VOSIM: a new synthesis system, J.Audio Eng. Soc., 26(6), 418-424, 1978

[4] Rodet X., Time domain formant wave functions synthesis, Comp. Music J., 8(3), 9-14, 1985

[5] Cavalieri S., Piccialli A., Phase modulation with interpolated time function: synthesis by formants, Proc ICMC86, 293-297, 1986

[6] Truax B., Real time granular synthesis with the DMX-1000, Proc. ICMC86, 231-235, 1986.

[7] De Poli G., Piccialli A., Dynamic control of FIR filters for sound synthesis, Proc EUSIPCO 88, 559-562, 1988

[8] Kronland-Martinet R., Morlet J., Grossmann A., Analysis of sound patterns through wavelet transform, Int.J.Patt.Rec.& A.I., 1, 273-320, 1987

[9] De Poli G., Piccialli A., Pitch synchronous granular synthesis, Int. Work. Models of Musical Sounds, Sorrento(Italy), Oct.1988.