Pasquale Mainolfi

Appunti per il corso di Musica Elettronica

Modeling Sound (part 2)

Amplitude modulation AM - RM

Amplitude modulation AM

È un modello di sintesi a trasformazione *nonlineare*, appartenente alla famiglia delle sintesi per modulazione, estremamente efficiente e, nel contempo, a basso carico computazionale. Consiste nell'utilizzo di un segnale per il controllo nel tempo dell'ampiezza di un secondo segnale: *ogni campione del primo segnale, detto modulatore M, moltiplica un campione corrispondente del secondo segnale, detto portante C, distorcendolo e creando nuove componenti spettrali.*

Modulare l'ampiezza di un segnale, significa dunque, far variare l'ampiezza del segnale portante proporzionalmente al valore istantaneo del segnale modulante.

Se consideriamo:

$$C(t) = A_c \sin(\omega_c t) \qquad M(t) = A_m \sin(\omega_m t)$$

il segnale modulato risultate avrà la forma

$$y(t) = A_m(t) \sin(\omega_c t)$$

La frequenza è la stessa del segnale portante, mentre l'ampiezza viene fatta variare nel tempo $A_m(t)$ secondo il valore istantaneo di M, secondo la regola:

$$A_m(t) = A_c + k v_m(t)$$

dove:

$A_c =$ ampiezza della portante in assenza di modulazione
$k \leq 1 =$ costante di proporzionalità del modulatore
$v_m(t) =$ valore istantaneo del segnale modulante

Ne consegue che

$$y(t) = (A_c + kA_m \sin(\omega_m t)) \sin(\omega_c t)$$

Il rapporto tra la massima deviazione di ampiezza del segnale modulato e l'ampiezza del segnale portante in assenza di modulazione si definisce indice di modulazione o di profondità m.

$$m = \frac{kA_m}{A_c}$$

Possiamo, allora, riscrivere $y(t)$ come

$$y(t) = A_c(1 + m \sin(\omega_m t)) \sin(\omega_c t)$$

y lo si definisce *segnale a portante non soppressa* e, corrisponde alla sovrapposizione del segnale portante con due segnali sinusoidali di pulsazione $\omega_c \pm \omega_m$. Vi sono, nella pratica, due bande laterali simmetriche rispetto a ω_c di ampiezza $\frac{mA_c}{2}$: *il termine generico dello sviluppo in serie di Fourier del segnale risultante dal prodotto di due segnali contiene i termini somma e differenza delle frequenze e fasi iniziali dei segnali originali.*

Si noti che:

- per il teorema del campionamento al fine di evitare fenomeni di aliasing, la somma tra f_c ed f_m deve sempre essere uguale o inferiore alla soglia di Nyquist
- se $f_m < 20Hz$ il prodotto tra i due segnali produrrà solo una lenta e periodica variazione dell'ampiezza della portante (tremolo)
- nel caso disegnali complessi, sia i segnali di partenza, sia la risultante sono complessi e, il prodotto deve intendersi come prodotto complesso. In questo caso, la modulazione di ampiezza produce uno spostamento di frequenza: ogni parziale dei due segnali genera, in coppia con ogni parziale dell'altro, una componente che ha come frequenza la somma algebrica delle frequenze originarie. Il numero di componenti generate è pari al risultato del prodotto tra il numero di parziali dei due segnali (posto un segnale a con N parziali e, un segnale b con M parziali, il segnale risultante $c = ab$ avrà NM parziali)
- nel caso di segnali reali, sia i segnali di partenza, sia la risultante sono reali e, il prodotto deve intendersi in senso ordinario

Ring modulation RM
La modulazione ad anello (ring modulation) è una tipologia di modulazione di ampiezza che ha come risultante un segnale in cui non vi è traccia della frequenza portante, cosiddetto *a portante soppressa.*

$$y(t) = (A_c m \sin(\omega_m t)) \sin(\omega_c t)$$

Frequency modulation FM

È un procedimento di sintesi per trasformazione nonlineare, introdotto da J. Chowning nel 1973, che consente di creare e controllare segnali complessi a partire da due segnali elementari (anche in questo caso, segnale portante e segnale modulante), propriamente detto ad inviluppo costante $A_c =$ costante.

Posti i segnali, portante C e modulate M

$$C(t) = A_c \sin(\omega_c t) \qquad M(t) = A_m \sin(\omega_m t)$$

$$y(t) = A_c \sin\left(\omega_c t + \frac{kA_m}{\omega_m} \sin(\omega_m t)\right) \rightarrow \text{ modulato}$$

L'ampiezza del segnale modulato $y(t)$ è mantenuta costante ed uguale al valore della portante a riposo A_c, mentre la frequenza viene fatta variare proporzionalmente al valore istantaneo dell'ampiezza del segnale modulante, con uno scarto massimo Δf rispetto a quella del segnale portante.

$$\Delta f = \frac{kA_m}{2\pi} \rightarrow \text{ deviazione istantanea di frequenza}$$

Il rapporto tra il massimo scarto frequenziale e la frequenza del segnale modulante, si definisce indice di modulazione m.

$$m = \frac{kA_m}{\omega_m} = \frac{\Delta f}{2\pi f_m}$$

$$y(t) = A_c \sin(\omega_c t + m \sin(\omega_m t))$$

con frequenza istantanea $f(t) = 2\pi(f_c + \Delta f \sin(\omega_m t))$

Spettro di un segnale modulato in frequenza
Applicando la relazione trigonometrica $\cos(\alpha + \beta) = \cos(a)\cos(b) - \sin(a)\sin(b)$ a $A_c \sin(\omega_c t + m \sin(\omega_m t))$, quest'ultima diventa

$$A_c \cos(\omega_c t) \cos(m \sin(\omega_m t)) - A_c \sin(\omega_c t) \sin(m \sin(\omega_m t))$$

in cui $\cos(m \sin(\omega_m t))$ e $\sin(m \sin(\omega_m t))$ sono *funzioni di Bessel* di prima specie, indicate con $J_n(m)$ e con $n =$ ordine.

Ne consegue che: un segnale sinusoidale modulato in frequenza da una modulante sinusoidale $f_m > 20Hz$, è rappresentato da infinite sinusoidi di ampiezza pari al prodotto dell'ampiezza a riposo A_c per il valore delle funzioni di Bessel associate $(J_0, J_1, J_2, J_3, J_4, ...)$ e, che dipende (il valore delle funzioni di B.) dall'indice di modulazione m (la componente n-esima avrà ampiezza $A_n = A_c J_n(m)$)

$$
\begin{gathered}
y(t) = A_c J_0(m)\cos(\omega_c t) + \\
-A_c J_1(m)(\cos((\omega_c + \omega_m)t) + \cos((\omega_c - \omega_m)t)) + \\
+A_c J_2(m)(\cos((\omega_c + 2\omega_m)t) + \cos((\omega_c - 2\omega_m)t)) + \\
-A_c J_3(m)(\cos((\omega_c + 3\omega_m)t) + \cos((\omega_c - 3\omega_m)t)) + \\
+A_c J_4(m)(\cos((\omega_c + 4\omega_m)t) + \cos((\omega_c - 4\omega_m)t)) + ...
\end{gathered}
$$

In teoria, dunque, un segnale modulato in frequenza è un segnale a banda infinita. Si noti bene, però, che quando si parla di larghezza di banda di un segnale modulato in frequenza, non è questo ciò a cui ci si sta riferendo, bensì all'insieme delle componenti frequenziali di valore significativo, ossia, di ampiezza uguale o superiore all'1 della portante non modulata.
Quando l'indice cresce, aumenta in corrispondenza anche il numero di funzioni significative. Il numero di componenti laterali L di ampiezza maggiore di un centesimo è dato da

$$L = m + 2.4 \cdot m^{0.27}$$

Modulando un segnale in frequenza, si genera un insieme di componenti disposte a destra (upper sidebands) e a sinistra (lower sidebands) della frequenza portante secondo la regola

$$f_c \pm n f_m \qquad n = 1, 2, 3, 4, 5, ...$$

In altre parola, somma e differenza differenza di tutti i multipli della modulante e, con una banda utile calcolabile approssimativamente come:

$$BW_{FM} = 2(\Delta f + f_m) \text{ (formula di Carson)}$$

C:M ratio
La distanza tra le componenti dipende dal rapporto tra f_c ed f_m. Controllare detto rapporto, significa poter determinare le caratteristiche spettrali del segnale risultante.

Cominciamo analizzando le componenti $f_c + n f_m$.
Consideriamo il caso $1 : 1 \rightarrow f_m = f_c = 1$

$$f_0 = 1, f_1 = 1, f_2 = 2, f_3 = 3, ...$$

Detto rapporto $1 : 1$ è un particolare caso in cui a generarsi è tutta le serie dei multipli della fondamentale (si dice *perfettamente armonico*).

Consideriamo ora il caso $1 : 2$

$$f_0 = 1, f_1 = 3, f_2 = 5, f_3 = 7, ...$$

Sono presenti solo le componenti a frequenza dispari multipli di f_0 (si dice spettro *armonico*).

E, ancora, consideriamo il caso 2 : 5

$$f_0 = 2, f_1 = 7, f_2 = 12, f_3 = 17, f_4 = 22, ...$$

In questo caso, lo spettro generato è un particolare spettro che si definisce *armonico-inarmonico*, in cui solo alcune componenti sono multipli della frequenza fondamentale.

Il rapporto 7 : 5

$$f_0 = 7, f_1 = 12, f_2 = 17, f_3 = 22, f_4 = 27, ...$$

genera invece uno spettro cosiddetto *del tutto inarmonico*.

Passiamo, ora ad analizzare le componenti $f_c - nf_m$, ricordando che le componenti con segno meno si riflettono nella parte positiva $|-n| = n$.

Consideriamo ancora il rapporto 1 : 2

$$f_1 = |-1|, f_2 = |-3|, f_3 = |-5|, ...$$

Nulla di particolare. L'ampiezza delle formanti ribaltate si somma con quelle a destra, senza aggiungere alcun'altra componente estranea.

Se consideriamo, però, il rapporto 7 : 5, il discorso cambia. Vediamo perchè.

$$f_1 = 2, f_2 = |-3|, f_3 = |-8|, f_4 = |-13|, ...$$

È chiaro che a generarsi è una componente con ampiezza e frequenza inferiore ad f_0. In questo caso deve considerarsi componente fondamentale, la componente a frequenza più bassa contenuta nello spettro generato.

Ciò detto:

- affinché la componente portante coincida con la componente fondamentale del segnale risultante, è necessario che la frequenza della componente modulante sia almeno il doppio della portante o, in rapporto 1 : 1
- affinché le componenti a sinistra della componente portante coincidano frequenzialmente con quelle a destra, è necessario che i rapporti tra portante e modulante seguano la regola $N : 1$ o $N : 2$
- tutti i rapporti diversi da $N : 1$ ed $N : 2$ generano spettri complessi a righe asimmetriche

Rapporto in-forma-normale e in-forma-non-normale
Si parla di rapporto in forma normale, quando la frequenza della componente portante coincide con la componente fondamentale dello spettro risultante $f_c = f_0$. Si parla, invece, di rapporto in-forma-non-normale quando quando la frequenza del segnale modulante è minore del doppio della frequenza portante $f_m < 2f_c$.
Tale distinzione di forma ci permette di dedurre:

1. se lo spettro risultante sarà di tipo armonico, armonico-inarmonico o, del tutto inarmonico
2. la famiglia di appartenza del rapporto

Riguardo al primo punto: tutti i rapporti espressi in forma normale del tipo $1 : N$ producono spettri armonici, viceversa, quelli in forma normale del tipo non $1 : N$, spettri inarmonici-armonici o, del tutto inarmonici.
Nel caso, invece, di un rapporto espresso in-forma-non-normale, è possbile risalirvi traducendo il rapporto in forma normale mediante un procedimento di riduzione che consiste in una semplice operazione di sottrazione:

$$f_c - f_m = \text{ nuova } f_c$$

e, si continua a ridurre fino a portare il rapporto in forma normale.

Riguardo al secondo punto: ad ogni rapporto in forma normale corrisponde un set di rapporti, chiamato *famiglia*, in grado di generare la stessa serie di componenti laterali (in ordine non identico) del rapporto in forma normale considerato. È possibile risalire al set di rapporti appartenenti alla stessa famiglia per mezzo di:

$$(f_c + Nf_m) : f_m \text{ e } (f_c - Nf_m) : f_m \qquad n = 1, 2, 3, ...$$

Inoltre, è possibile, a partire da una certa e nota componente frequenziale portante in un certo e noto rapporto espresso in forma normale, calcolare la componente portante per i membri della stessa famiglia di rapporti, in modo che tutti gli spettri generati abbiamo la stessa componente fondamentale, secondo la regola:

$$f_{c \text{ new}} = f_{\text{ ratio}} \frac{f_c}{f_{\text{ratio in normal form}}}$$

per ogni forma del tipo $1 : N$

$$f_{c \text{ new}} = f_{\text{ ratio}} \cdot f_c$$

per rapporti in forma normale

$$f_{c\text{ new}} = f_{\text{ratio in normal form}}$$

Viceversa, nota la frequenza della portante possiamo conoscere la frequenza della componente fondamentale dello spettro risultante da rapporti in-forma-non-normale, svolgendo:

$$f_0 = f_c \frac{f_{\text{ratio in normal form}}}{f_{ratio}}$$

In ultima analisi, nota la frequenza fondamentale f_0 è possibile risalire alla frequenza del segnale portante, calcolando

$$f_c = \frac{\frac{f_0}{f_{ratio}}}{f_{\text{ratio in normal form}}}$$

Phase modulation PM

È un procedimento di sintesi nonlineare che consiste nel variare l'angolo di fase di un segnale portante, proporzionalmente al valore istantaneo si un segnale modulante.

$$y(t) = A_c \sin(\omega_c t + \phi(t)) \qquad \phi(t) = \phi_c + kA_m(t)$$

$k =$ costante di proporzionalità del modulatore
$\phi_c =$ costante di fase della portante

Definiamo, $\Delta\phi$ la massima deviazione di fase del segnale, in radianti.

$$\Delta\phi = kA_m$$

La frequenza istantanea del segnale modulato, sarà:

$$\omega(t) = \omega_c + \Delta\phi \cdot \omega_m \cos(\omega_m t)$$

e, massima deviazione di frequenza

$$\Delta f = \Delta\phi f_m$$

Dipende, quindi, sia dall'ampiezza che dalla frequenza del segnale modulante e, rappresenta nella modulazione di fase l'equivalente dell'indice di modulazione in un procedimento FM. Quanto detto per la FM vale allo stesso modo per un segnale modulato nella sua fase: ad un indice di modulazione FM $m = 12$ corrisponde uno spettro generato da un segnale modulato in fase con $\Delta\phi = 12$: *la*

modulazione della fase di una portante ha come spettro la medesima portante e coppie di righe simmetriche rispetto alla frequenza della portante, spaziate di una quantità pari alla frequenza della modulante.
Esiste, tuttavia, una sostanziale differenza tra uno spettro generato da un segnale modulato in frequenza e, uno modulato nella sua fase: in un segnale modulato in frequenza l'argomento delle funzioni di Bessel è dipendente sia dall'ampiezza che dalla frequenza del segnale modulante $\frac{\Delta f}{f_m}$, mentre, nella modulazione di fase è in funzione solo dell'ampiezza e indipendente dalla frequenza kA_m.

Modulare la fase di un segnale, qualunque sia la frequenza della componente modulante, con ampiezza costante restituisce sempre lo stesso valore di $\Delta\phi$. Per un certo valore di A_m, qualunque sia f_m, il numero di righe dello spettro resta uguale in ciascun caso. Raddoppiando f_m, il numero di componenti spettrali non varia, raddoppia solo la distanza Δf.

Modulazione di fase a portante zero
È un particolare modello di modulazione di fase in cui la frequenza della portante è, praticamente, pari a zero. Tale procedimento genera solo pettini armonici con pulsazione fondamentale ω_m (pulsazione della modulante).

$$\text{zero port.} = J_{(0,A_{PM})}\cos(\phi_c) + 2\sum_{k=1}^{\infty} J_{(k,A_{PM})}\cos\left(\phi_c + \frac{\pi}{2}k\right)\cos(K(\omega_m t + \phi_m))$$

Sintesi per distorsione nonlineare DNL

Il primo esperimento di sintesi per distorsione non lineare risale al 1969 da parte di J. C. Risset presso i laboratori Bell nel New Jersey. L'idea alla base era quella di far passare un segnale x attraverso un blocco distorcente. La funzione $h(x)$ mappa ogni valore di input x nell'intervallo $[-1,+1]$ a un valore di output $h(x)$ nello stesso intervallo.
Associabile, sotto alcuni versi FM e alla PM, ma, con una sostanziale differenza strutturale: non vi è traccia dell'oscillatore modulante.

Tipicamente, un sistema DNL si compone di un oscillatore sinusoidale (nella maggior parte dei casi, ma potrebbe trattarsi di qualunque segnale) ed una funzione di trasferimento, detta funzione distorcente (shaping function).

Associando all'ampiezza del segnale in ingresso la variabile x, l'uscita del sistema sarà $f(x)$, con f funzione di trasferimento il cui valore dipende solo ed esclusivamente da x. Ad una funzione del tipo $f(x) = 0.5x$, corrisponde, ad esempio, corrisponde la metà del valore di x. Una funzione lineare che preserva la forma del segnale riducendone solo l'ampiezza (box attenuatore).

$$f(t) = k \cdot h(t) \qquad k \leq 1$$

Altro esempio di funzione lineare è il box inversore. In questo caso si interviene solo variando la fase del segnale, invertendola:

$$f(t) = k \cdot h(t) \qquad k = -1$$

Una funzione di trasferimento è, invece, non lineare, quando possiede termini che contengono x evvlevato a potenza diversa da zero o uno oppure, quando contiene funzioni di x come logaritmo o tangente.

In un processore non lineare, la relazione tra segnale d'ingresso e segnale in uscita dipende dall'ampiezza del segnale in ingresso e, dalla natura della non linearità: un aumento dell'ampiezza del segnale in ingresso causerà una variazione della forma d'onda e, dunque, del suo contenuto spettrale.

Un waveshaper (o processore non lineare) è caratterizzato da una funzione di trasferimento, propriamente detta funzione di sagomatura, matematicamente espressa come:

$$y = f(a(t) \cdot x(t))$$

dove f è la funzione di sagomature, $a(t)$ l'indice e $x(t)$ l'input.
La relazione tra segnale in ingresso e segnale in uscita dipende dall'ampiezza del segnale: la modifica sull'ampiezza del segnale induce una modifica sul contenuto in frequenza.

Il grande problema alla base dello sviluppo e della stesura di una teoria sulla tecnica di sintesi DNL da parte del suo ideatore J. C. Risset nel 1969, fu l'impossibilità di controllare e prevedere l'effetto risultante. I segnali ottenuti, applicando un generico operatore erano in generale a banda non limitata (aliasing!). La soluzione arriverà nel 1979 con Arfib e Le Brun: è possibile ottenere risultanti a spettro controllato usando particolari funzioni distorcenti ottenibili dai polinomi di Cebysev di primo tipo nella forma $T_k(x)$:

$T_0 := 1$
$T_1 := x$
$T_2 = 2x^2 - 1$
$T_3 = 4x^3 - 3x$
$T_4 = 8x^4 - 8x^2 + 1$
$T_{k+1} = 2xT_k(x) - T_{k-1}(x)$

Per evitare l'effetto aliasing una funzione di trasferimento dovrà essere espressa come un polinomio del tipo:

$$f(x) = d_0 + d_1x + d_2x^2 + ... + d_Nx^N$$

con grado del polinomio dato da N e che equivale al valore del maggiore esponente e, che determina la N-esima formante.

Quando un processore *non lineare* è controllato da un segnale sinusoidale, l'ampiezza delle armoniche presenti nello spettro risultante può essere calcolata tramite il polinomio di trasferimento usando il *triangolo di Pascal*, con uno schema aritmetico (array triangolare di coefficienti binomiali) per la previsione dell'ampiezza di ogni singola armonica generata dal rispettivo termine nel polinomio. Quando l'ampiezza della sinusoide è pari a 1:

	DIV	h_0	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}	h_{11}
x_0	0.5	1											
x_1	1		1										
x_2	2	2		1									
x_3	4		3		1								
x_4	8	6		4		1							
x_5	16		10		5		1						
x_6	32	20		15		6		1					
x_7	64		35		21		7		1				
x_8	128	70		56		28		8		1			
x_9	256		126		84		36		9		1		
x_{10}	512	252		210		120		45		10		1	
x_{11}	1024		462		330		165		55		11		1

Ad una funzione del tipo:

$$f(x) = x^5$$

corrisponde un segnale dato dalla sovrapposizione di prima, terza e quinta componente, con ampiezze relative, pari a:

$h_1 = \frac{1}{16}(10) = 0.625$
$h_3 = \frac{1}{16}(5) = 0.3125$
$h_5 = \frac{1}{16}(1) = 0.0625$

In presenza di un polinomio composto da termini multipli, il risultato è dato dalla somma dei contributi di ciascun termine. Ad esempio:

$$f(x) = x + x^2 + x^3 + x^4 + x^5$$

$\frac{h_0}{2} = \frac{\frac{1}{2}\cdot 2 + \frac{1}{8}\cdot 6}{2} = 0.875$
$h_1 = 1 + \frac{1}{4}(3) + \frac{1}{16}(10) = 2.375$
$h_2 = \frac{1}{2}(1) + \frac{1}{8}(4) = 1$

$h_3 = \frac{1}{4}(1) + \frac{1}{16}(5) = 0.5625$
$h_4 = \frac{1}{8}(1) = 0.125$
$h_5 = \frac{1}{16}(1) = 0.0625$

Ad esponente pari, armoniche pari; ad esponente dispari, armoniche dispari.

$$f(x) = d_0 + d_1 ax + d_2 a^2 x^2 + ... + d_N a^N x^N$$

con a indice di modulazione e controllo della distorsione: a controlla la risultante. Nello specifico: *determina la relazione che lega l'ampiezza dell'onda sinusoidale in ingresso con l'ampiezza di una data armonica.*

Sintesi granulare

L'idea alla base della sintesi granulare è quella di creare suoni assai complessi a partire da una grossa quantità di elementi acustici elementari in successione e/o in sovrapposizione; un'alta densità di particelle sonore (grani sonori), con inviluppi generalmente simmetrici e di durata che oscilla tra $9 - 120ms$.

Per comprendere bene il procedimento bisogna partire da Dennis Gabor, anno 1947. Secondo Gabor qualsiasi segnale acustico può essere definito mediante l'appropriata combinazione di migliaia di *quanta acustici*. La teoria di Fourier da sola non basta a spiegare il fenomeno, semplicemente perchè non considera l'aspetto temporale (palese). È importante dare una reinterpretazione che prenda in considerazione tanto l'analisi di Fourier quanto la necessità di quantizzare il suono dal punto di vista temporale.
In altre parole, se consideriamo il periodo di una sinusoide generato da un oscillatore, questo si ripeterà uguale nel tempo. Se, a questo punto, campioniamo un intervallo di tempo pari allo stesso periodo e replichiamo il campione un numero di volte necessario a creare continuità percettiva nell'ascoltatore, allora, potremo ottenere lo stesso segnale generato dall'oscillatore ma in forma granulare.
Posto in questi termini, dal punto di vista pratico, il procedimento non sarebbe così tanto accattivante se non dal punto di vista teorico di poter controllare i singoli parametri della risintesi di ciascun grano sonoro. Per intenderci: un segnale rumoroso, non sarebbe altro che un susseguirsi di grani con un inviluppo a campana e con potenza randomica nel dominio del tempo. Di fatto, le teorie di Gabor non ebbero un grande successo e, probabilmente nessuno se ne ricorderebbe se non fosse per il compositore greco Iannis Xenakis che, invece, ne rimase affascinato.
Nel 1971, nel trattato *Formalized Music: Thought and Mathematics in Composition*, Iannis X. concretizza le ricerche di Gabor da un punto di vista musicale: fornisce un esempio di *cluster* sonoro e formalizza il *grano* nelle sue caratteristiche fondamentali: ampiezza, durata, spettro, inviluppo etc....

"*Ogni suono, così come anche le sue variazioni più continue, è assimilabile ad un insieme di un numero sufficientemente alto di particelle elementari: durante* *l'attacco, il sostegno e il decadimento di un suono complesso, migliaia di suoni puri appaiono in intervalli di tempi più o meno brevi*" (Iannis Xenakis).

Sono cinque le principali forme di sintesi granulare:

1. Fourier and Wavelet Grids
2. pitch-Synchrounus overlapping streams
3. quasi-Synchrounus streams
4. aSynchrounus clouds
5. time-Granulated or stream with overlapped quasi-Syncrhounus or aSynchrounus playback

I principali parametri di controllo sono:

- tempo d'inizio e durata della nuvola sonora (migliaia di grani distribuiti in modo irregolare all'interno di una maschera che delimita una regione di spazio tempo-frequenza-ampiezza)
- durata dei grani
- densità dei grani per secondo (basse densità producono effetti di suono puntilistico, viceversa, alte densità provocano sovrapposizione dei grani fino a produrre suoni continui)
- forma d'inviluppo del grano
- forma d'onda del grano
- posizione spaziale del grano
- larghezza di banda della nuvola (la frequenza di ogni grano viene definita specificando i limiti inferiori e superiori entro cui i grani sono dispersi. Quando i limiti coincidono, il risultato è una linea continua, un suono ad altezza determinata)
- tipo di nuvola (nuvola sincrona, in cui la distanza tra i grani è costante; nuvola asincrona, in cui la distanza tra i grani non è costante)

È possibile riassumere il procedimento di sintesi come:

$$y(t) = \sum_{k=1}^{N} a_k g_k(t - t_k) \qquad g(t) = w(t) \sin(2\pi f_0 t)$$

$w =$ finestra che definisce forma e durata del grano

$g_k(t - t_k) =$ k-esimo grano al tempo t_k

$a_k =$ coefficiente d'ampiezza del k-esimo grano

$N =$ numero totale di grani combinati

In pratica, un grano sonoro consente di fondere insieme le informazioni desumibili da una rappresentazione nel dominio del tempo con quelle nel dominio della frequenza.